



(51) International Patent Classification:

A61K 39/12 (2006.01) C07K 14/77 (2006.01)
A61K 48/00 (2006.01) A61K 39/00 (2006.01)

(21) International Application Number:

PCT/US2016/034539

(22) International Filing Date:

27 May 2016 (27.05.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/167,709 28 May 2015 (28.05.2015) US
62/171,711 5 June 2015 (05.06.2015) US

(71) Applicant: **THE JOHNS HOPKINS UNIVERSITY**
[US/US]; 3400 North Charles Street, Baltimore, MD
21218 (US).

(72) Inventors: **RODEN, Richard**; 13 Luna Lane, Sevens
Park, MD 21146 (US). **WU, Tzyy-Chou**; 11002
Nacirema Lane, Stevenson, MD 21153 (US). **HUNG, Chi-
en-Fu**; 13 Ballyhean Court, Timonium, MD 21093 (US).

(74) Agents: **RUSSELL, Hathaway, P.** et al.; FOLEY HOAG
LLP, Seaport West, 155 Seaport Boulevard, Boston, MA
02210-2600 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report (Rule 48.2(g))

(54) Title: METHODS FOR ENHANCING ANTIGEN-SPECIFIC IMMUNE RESPONSES USING COMBINATION THERAPY
COMPRISING PAPILLOMAVIRUS CAPSID ANTIGENS

(57) Abstract: Described herein are methods comprising co-administering to a mammalian an effective amount of a combination
therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising
an antigenic peptide, to thereby enhancing an antigen specific immune response.



**METHODS FOR ENHANCING ANTIGEN-SPECIFIC
IMMUNE RESPONSES USING COMBINATION THERAPY COMPRISING
PAPILLOMAVIRUS CAPSID ANTIGENS**

CROSS REFERENCE TO RELATED APPLICATIONS

5 This application claims priority to U.S. Provisional Application No: 62/167,709, filed May 28, 2015, and U.S. Provisional Application No: 62/171,711, filed June 5, 2015, the contents of which are incorporated herein by reference in their entirety.

GOVERNMENTAL SUPPORT

10 This invention was made with some government support under grant numbers P20CA192988, RO1CA114425, 1R21AI109259, and CA118790 awarded by the National Institutes of Health. The government has certain rights in the invention. This statement is included solely to comply with 37 C.F.R. § 401.14(a)(f)(4) and should not be taken as an assertion or admission that the application discloses and/or claims only one invention.

15

BACKGROUND

 The knowledge that high risk human papillomavirus (HPV) is a necessary etiological factor for the development of cervical cancer provides the opportunity for control of cervical cancer and/or other HPV-associated malignancies through vaccination against
20 HPV. Currently, the two commercially available prophylactic vaccines, Gardasil and Cervarix, based upon virus-like particles (VLP) derived from the major capsid protein L1 effectively protect against infection by the two most common HPV types found in cervical cancer (for review see Roden R et al. *Nat Rev Cancer* 2006, 6:753-763). The minor capsid antigen L2 also shows promise for preventive vaccination in animal models, although it is
25 less immunogenic than L1 VLP. However, neither L1 VLP nor L2-based vaccines generate therapeutic effects against established HPV infection (Schiller JT et al. *Vaccine* 2008, 26 Suppl 10:K53-61). Therefore, given the significant burden of HPV-associated lesions worldwide, there is an urgent need to develop therapeutic HPV vaccines for the control of existing HPV infection and associated malignancies. Quite different from preventive
30 vaccines which target the viral capsid proteins L1 and/or L2, therapeutic HPV vaccines focus on targeting the HPV E6 and E7, since only these oncoproteins are consistently expressed in HPV-associated cancers and are responsible for the malignant transformation.

Among the various therapeutic HPV vaccines currently being tested, DNA vaccines have emerged as attractive candidates for the treatment of cervical cancer and associated malignancies. Naked DNA is relatively safe, stable, and easy to produce and transport (for review see Gurunathan S et al. *Annual review of immunology* 2000, 18:927-974).

- 5 Furthermore, DNA vaccines are capable of sustained cellular gene expression, promote MHC class I antigen presentation, and have the capacity for repeated administration since they do not lead to the generation of neutralizing antibodies. However, an important limitation of DNA vaccines is limited potency since they lack the intrinsic ability to amplify and spread *in vivo*. Therefore, it is important to consider strategies to improve
- 10 DNA vaccine potency strategies (for review, see Hung CF et al. *Methods in molecular medicine* 2006, 127:199-220; Tsen SW et al. *Expert Rev Vaccines* 2007, 6:227-239).

One strategy to enhance DNA vaccine potency is to improve antigen expression, processing, and presentation in antigen-presenting cells using intracellular targeting strategies (for review, see Hung CF et al. *Methods in molecular medicine* 2006, 127:199-

15 220; Tsen SW et al. *Expert Rev Vaccines* 2007, 6:227-239). Several of inventors' previous studies have employed DNA vaccines encoding calreticulin (CRT) linked to HPV-16 E7 antigen (CRT/E7) (Cheng WF et al. *The Journal of clinical investigation* 2001, 108:669-678; Kim JW et al. *Gene Ther* 2004; Peng S et al. *Journal of virology* 2004, 78:8468-8476; Peng S et al. *Gene Ther* 2006, 13:257-265; Peng S et al. *Human gene therapy* 2005,

20 16:584-593; Kim TW et al. *Journal of virology* 2004, 78:4638-4645) because this fusion greatly enhances the E7-specific CD8⁺ T cell immune responses in vaccinated mice. This phenomenon likely reflects the efficient recognition of calreticulin by dendritic cells and improved targeting to the MHCI pathway by targeting to the endoplasmic reticulum (Cheng WF et al. *The Journal of clinical investigation* 2001, 108:669-678).

- 25 Activation and proliferation of CD4⁺ T cells is crucial to the success of both humoral and cell-mediated responses to viral infection. Although the fusion of E7 with CRT enables such DNA vaccines to elicit significant E7-specific CD8 T cell immunity in the absence of CD4 T cell help (Peng S et al. *Cell & bioscience* 2014, 4:11), approaches to boost CD4⁺ T cell help are likely to enhance DNA vaccine potency. Indeed, CD4⁺ T cells
- 30 play a significant role in priming effector CD8⁺ T cells, thus augmenting the CD8⁺ T cell responses, as well as generation of memory T cell populations (for review see Castellino F et al. *Annual review of immunology* 2006, 24:519-540). CD4⁺ T cells help to differentiate naïve CD8⁺ T cells into effector cells by providing activation signals to dendritic cells

(DCs), most notably IL-2, thus promoting CD8⁺ T cell proliferation. Thus, strategies to induce CD4⁺ T helper cells at sites of CD8⁺ T cell priming can potentially enhance CTL immune responses.

5 **SUMMARY OF THE INVENTION**

Described herein are methods comprising co-administering to a mammalian an effective amount of a combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide, to thereby enhancing an antigen specific immune response.

10 The present invention provides for combine intracellular targeting strategies using CRT with a strategy to enhance CD4⁺ T help for the development of a therapeutic HPV DNA vaccine. Since papillomavirus L1 or L2 antigens contain CD4⁺ T cell epitopes (Steele JC et al. *British journal of cancer* 2005, 93:248-259; Mustafa W et al. *Viral immunology* 2009, 22:195-204; Pinto LA et al. *The Journal of infectious diseases* 2003, 188:327-338; 15 Tobery TW et al. *Vaccine* 2003, 21:1539-1547; Williams OM et al. *Journal of virology* 2002, 76:7418-7429), co-administration of vectors containing codon-optimized bovine papillomavirus (BPV) L1 and L2 may provide CD4⁺ T cell help and enhance the antigen-specific CD8⁺ T cell immune responses generated by DNA vaccines. An additional potential benefit would be the induction of neutralizing antibody and protective immunity 20 (Leder C et al. *Journal of virology* 2001, 75:9201-9209; Kwak K et al. *PloS one* 2013, 8:e60507).

Co-administration of vectors containing BPV L1 or L2 DNA in combination with DNA vaccines could elicit enhanced antigen-specific CD8⁺ in both CRT/E7 and ovalbumin (OVA) antigenic systems. Co-administration of vectors containing BPV1 L1 ± L2 DNA 25 with CRT/E7 DNA led to the generation of L1/L2-specific CD4⁺ T cell immune responses as well as L1-specific neutralizing antibodies. Co-administration with BPV1 L1 significantly enhances the therapeutic antitumor effects generated by CRT/E7 DNA vaccination. In addition, the observed enhancement of CD8⁺ T cell immune responses by DNA encoding L1 and L2 was also found to extend to HPV-16 L1/L2 system. Co- 30 administration of DNA encoding papillomavirus L1 or L2 can be used to enhance antigen-specific CD8⁺ T cell immune responses generated by therapeutic HPV DNA vaccination for the control of HPV infection and HPV-associated tumors. Furthermore, such methods can also generate neutralizing antibodies against papillomavirus for potential prevention against

infection. This strategy also provides the opportunity to combine preventive and therapeutic approaches. Such methods may potentially be extended to other antigenic systems for the control of infection and/or cancer.

One aspect of the invention relates to a method of inducing or enhancing an antigen-specific immune response in a mammal, comprising co-administering to the mammal an effective amount of a combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide, to thereby enhancing the antigen specific immune response.

In certain embodiments, the papillomavirus is selected from the group consisting of human papilloma virus (HPV)1, HPV2, HPV3, HPV4, HPV5, HPV6, HPV7, HPV8, HPV9, HPV10, HPV11, HPV12, HPV13, HPV14, HPV15, HPV16, HPV17, HPV18, HPV19, HPV20, HPV21, HPV22, HPV23, HPV24, HPV25, HPV26, HPV27, HPV28, HPV29, HPV30, HPV31, HPV32, HPV33, HPV34, HPV35, HPV36, HPV37, HPV38, HPV39, HPV40, HPV41, HPV42, HPV43, HPV44, HPV45, HPV46, HPV47, HPV48, HPV49, HPV50, HPV51, HPV52, HPV53, I-IPV54, HPV55, HPV56, HPV57, HPV58, HPV59, HPV60, HPV61, HPV62, HPV63, HPV64, HPV65, HPV66, HPV67, HPV68, HPV69, HPV70, HPV71, HPV72, HPV73, HPV74, HPV75, HPV76, HPV77, HPV78, HPV79, HPV80, HPV81, HPV82, HPV83, HPV84, HPV85, HPV86, HPV87, HPV88, HPV89, HPV90, HPV91, HPV92, HPV93, HPV94, HPV95, HPV96, HPV97, HPV98, HPV99, HPV100, bovine papillomavirus type 1 (BPV1), bovine papillomavirus type 2 (BPV2), bovine papillomavirus type 4 (BPV4), cottontail rabbit papillomavirus (CRPV), deer papillomavirus (DPV), European elk papillomavirus (EPPV), canine oral papillomavirus (COPV), Rhesus monkey papillomavirus (RnPV), and rabbit oral papillomavirus (ROPV).

In certain embodiments, the papillomavirus is bovine papillomavirus type 1 (BPV1).

In certain embodiments, the capsid antigen is a major and minor capsid antigen.

In certain embodiments, the major capsid antigen is L1.

In certain embodiments, the minor capsid antigen is L2.

In certain embodiments, the capsid antigen L1 is derived from BPV1, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 123.

In certain embodiments, the capsid antigen L1 is derived from BPV1, and said capsid antigen comprises a nucleotide sequence at least 60% identical to the nucleotide sequence set forth in SEQ ID NO: 122.

5 In certain embodiments, the capsid antigen L2 is derived from BPV1, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 125.

In certain embodiments, the capsid antigen L2 is derived from BPV1, and said capsid antigen comprises a nucleotide sequence at least 60% identical to the nucleotide sequence set forth in SEQ ID NO: 124.

10 In certain embodiments, the combination therapy elicits enhanced antigen-specific CD8⁺.

In certain embodiments, the DNA vaccine comprises an antigenic peptide, and said antigenic peptide is selected from the group consisting of ovalbumin (OVA), HPV16 E6, HPV16 E7, and influenza NS1.

15 In certain embodiments, the DNA vaccine comprises HPV-16 E7 (E7) antigen linked with calreticulum (CRT) (CRT/E7).

In certain embodiments, the E7 comprises an amino acid sequence set forth in RAHYNIVTF.

In certain embodiments, the DNA vaccine comprises OVA.

20 In certain embodiments, the OVA comprises an amino acid sequence set forth in ISQAVHAAHAEINEAGR.

In certain embodiments, the papillomavirus is human papillomavirus type 16 (HPV-16).

In certain embodiments, the major capsid antigen is L1.

25 In certain embodiments, the minor capsid antigen is L2.

In certain embodiments, the capsid antigen L1 is derived from HPV16, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 119.

30 In certain embodiments, the capsid antigen L1 is derived from HPV16, and said capsid antigen comprises a nucleotide sequence at least 60% identical to the nucleotide sequence set forth in SEQ ID NO: 118.

In certain embodiments, the capsid antigen L2 is derived from HPV16, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 121.

5 In certain embodiments, the capsid antigen L2 is derived from HPV16, and said capsid antigen comprises a nucleotide sequence at least 60% identical to the nucleotide sequence set forth in SEQ ID NO: 120.

In certain embodiments, the combination therapy elicits enhanced antigen-specific CD4⁺ T cell immune response.

10 In certain embodiments, the enhanced antigen-specific CD4⁺ T cell immune response is L1-, L2-, or L1/L2-specific.

In certain embodiments, the mammal is a human.

In certain embodiments, the mammal is afflicted with cancer.

In certain embodiments, the combination therapy is administered intradermally, intraperitoneally, or intravenously via injection.

15 In certain embodiments, the combination therapy is administered alone or in a pharmaceutically acceptable carrier in a biologically-effective, and/or a therapeutically-effective amount,

In certain embodiments, the antigen-specific immune response is greater in magnitude than an antigen-specific immune response induced by administration of the DNA vaccine or DNA vector alone.

20 Another aspect of the invention relates to a combination for use in inducing or enhancing an antigen-specific immune response in a mammal, comprising (a) an effective amount of at least one DNA vector comprising a papillomavirus capsid antigen to be administered; (b) an effective amount of at least one DNA vaccine comprising an antigenic peptide to be administered, and the combination induces or enhances the antigen specific immune response.

In certain embodiments, the papillomavirus is selected from the group consisting of human papilloma virus (HPV)1, HPV2, HPV3, HPV4, HPV5, HPV6, HPV7, HPV8, HPV9, HPV10, HPV11, HPV12, HPV13, HPV14, HPV15, HPV16, HPV17, HPV18, HPV19, HPV20, HPV21, HPV22, HPV23, HPV24, HPV25, HPV26, HPV27, HPV28, HPV29, HPV30, HPV31, HPV32, HPV33, HPV34, HPV35, HPV36, HPV37, HPV38, HPV39, HPV40, HPV41, HPV42, HPV43, HPV44, HPV45, HPV46, HPV47, HPV48, HPV49, HPV50, HPV51, HPV52, HPV53, I-IPV54, HPV55, HPV56, HPV57, HPV58, HPV59,

HPV60, HPV61, HPV62, HPV63, HPV64, HPV65, HPV66, HPV67, HPV68, HPV69, HPV70, HPV71, HPV72, HPV73, HPV74, HPV75, HPV76, HPV77, HPV78, HPV79, HPV80, HPV81, HPV82, HPV83, HPV84, HPV85, HPV86, HPV87, HPV88, HPV89, HPV90, HPV91, HPV92, HPV93, HPV94, HPV95, HPV96, HPV97, HPV98, HPV99, HPV100, bovine papillomavirus type 1 (BPV1), bovine papillomavirus type 2 (BPV2), bovine papillomavirus type 4 (BPV4), cottontail rabbit papillomavirus (CRPV), deer papillomavirus (DPV), European elk papillomavirus (EPPV), canine oral papillomavirus (COPV), Rhesus monkey papillomavirus (RnPV), and rabbit oral papillomavirus (ROPV).

In certain embodiments, the papillomavirus is bovine papillomavirus type 1 (BPV1).

In certain embodiments, the capsid antigen is a major and minor capsid antigen.

In certain embodiments, the major capsid antigen is L1.

In certain embodiments, the minor capsid antigen is L2.

In certain embodiments, the capsid antigen L1 is derived from BPV1, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 123.

In certain embodiments, the capsid antigen L2 is derived from BPV1, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 125.

In certain embodiments, the combination elicits enhanced antigen-specific CD8+.

In certain embodiments, the DNA vaccine comprises an antigenic peptide, and said antigenic peptide is selected from the group consisting of ovalbumin (OVA), HPV16 E6, HPV16 E7, and influenza NS1.

In certain embodiments, the DNA vaccine comprises HPV-16 E7 (E7) antigen linked with calreticulum (CRT) (CRT/E7).

In certain embodiments, the E7 comprises an amino acid sequence set forth in RAHYNIVTF.

In certain embodiments, the DNA vaccine comprises OVA.

In certain embodiments, the OVA comprises an amino acid sequence set forth in ISQAVHAAHAEINEAGR.

In certain embodiments, the papillomavirus is human papillomavirus type 16 (HPV-16).

In certain embodiments, the major capsid antigen is L1.

In certain embodiments, the minor capsid antigen is L2.

In certain embodiments, the capsid antigen L1 is derived from HPV16, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 119.

5 In certain embodiments, the capsid antigen L2 is derived from HPV16, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 121.

In certain embodiments, the combination elicits enhanced antigen-specific CD4⁺ T cell immune response.

10 In certain embodiments, the enhanced antigen-specific CD4⁺ T cell immune response is L1-, L2-, or L1/L2-specific.

In certain embodiments, the mammal is a human.

In certain embodiments, the mammal is afflicted with cancer.

15 In certain embodiments, the combination is administered intradermally, intraperitoneally, or intravenously via injection.

In certain embodiments, the combination is administered alone or in a pharmaceutically acceptable carrier in a biologically-effective, and/or a therapeutically-effective amount,

20 In certain embodiments, the antigen-specific immune response is greater in magnitude than an antigen-specific immune response induced by administration of the DNA vaccine or DNA vector alone.

25 Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

30 **Figure 1** contains two panels (A) and (B) describing characterization of antigen-specific CD8⁺ T cell immune responses generated by antigen-specific DNA vaccine mixed with vectors containing BPV1 L1 or L2 DNA. C57BL/6 mice (five per group) were vaccinated intradermally via gene gun with 2 µg/mouse of CRT/E7 or OVA DNA with or

without BPV1 L1 or L2 DNA twice at 1-week intervals. Splenocytes from vaccinated mice were collected 1 week after last immunization and the E7 or OVA-specific T cell immune responses were characterized using intracellular cytokine staining followed by flow cytometry analysis. **Panel (A)** depicts representative flow cytometry data depicting the number of E7 (upper panel) or OVA(lower panel)-specific CD8⁺ T cells. **Panel (B)** depicts a bar graph representing the number of E7 (upper panel) or OVA (lower panel) -specific CD8⁺ T cells/3x10⁵ splenocytes (mean $\pm$ SD). Data shown are representative of two experiments performed. * indicates p < 0.05.

Figure 2 contains two panels (A) and (B) describing the characterization of antigen-specific CD8⁺ T cell immune responses generated by antigen-specific DNA vaccine mixed with HPV-16 L1 or L2 DNA. C57BL/6 mice (five per group) were vaccinated intradermally via gene gun with 2 μ g/mouse of CRT/E7 or OVA DNA with or without HPV-16 L1 or L2 DNA twice at 1-week intervals. Splenocytes from vaccinated mice were collected 1 week after last immunization and the E7 or OVA-specific T cell immune responses were characterized using intracellular cytokine staining followed by flow cytometry analysis. **Panel (A)** depicts representative flow cytometry data depicting the number of E7 (upper panel) or OVA (lower panel)-specific CD8⁺ T cells. **Panel (B)** depicts a bar graph representing the number of E7 (upper panel) or OVA (lower panel)-specific CD8⁺ T cells/3x10⁵ splenocytes (mean $\pm$ SD). Data shown are representative of two experiments performed. * indicates p < 0.05.

Figure 3 contains four panels (A), (B), (C), and (D) describing characterization of BPV1 and HPV-16 L1 or L2-specific CD4⁺ T cell immune responses generated by CRT/E7 or OVA DNA mixed with DNA encoding BPV1 or HPV-16 L1 or L2. C57BL/6 mice (five per group) were vaccinated intradermally via gene gun with 2 μ g/mouse of CRT/E7 with or without BPV1 or HPV-16 L1 or L2 DNA twice at 1-week intervals. Splenocytes from vaccinated mice were collected 1 week after last immunization and pulsed with 5 μ g/mL of BPV1 or HPV-16 L1/L2 VLPs. The BPV1 L1/ L2-specific CD4⁺ T cell immune responses were characterized using intracellular cytokine staining followed by flow cytometry analysis. **Panel (A)** depicts representative flow cytometry data depicting the number of BPV1 L1/L2-specific CD4⁺ T cells. **Panel (B)** depicts a bar graph representing the number of BPV1 L1/L2-specific CD4⁺ T cells/3x10⁵ splenocytes (mean $\pm$ SD). **Panel (C)** depicts representative flow cytometry data depicting the number of HPV-16 L1/L2-specific CD4⁺ T cells. **Panel (D)** depicts a bar graph representing the number of HPV-16 L1/L2-specific

CD4⁺ T cells/ 3×10^5 splenocytes (mean[±] SD). Data shown are representative of two experiments performed. * indicates $p < 0.05$.

Figure 4 contains two panels (A) and (B) describing *in vivo* tumor treatment experiments in mice vaccinated with CRT/E7 DNA mixed with BPV1 L1 DNA. C57BL/6 mice (five per group) were first challenged with 1×10^5 /mouse of TC-1 tumor cells subcutaneously. One week after tumor challenge, mice were treated intradermally via gene gun with 2 μ g/mouse of CRT/E7 DNA alone, BPV1 L1 DNA alone or CRT/E7 DNA in combination with BPV1 L1 DNA. Vaccinated mice were boosted twice at 1-week intervals with the same dose and regimen. Mice were sacrificed on day 30 after the last vaccination. Tumor growth were monitored twice weekly by caliper measurements and palpations. **Panel (A)** depicts a line graph depicting the tumor volume over time of tumor-bearing mice treated with CRT/E7 DNA and/or BPV1 L1 DNA. **Panel (B)** depicts Kaplan-Meier survival analysis of tumor-bearing mice treated with CRT/E7 DNA and/or BPV1 L1 DNA. The data presented are from one representative experiment of the two performed. * indicates $p < 0.05$.

Figure 5 contains one panel (A) describing the characterization of BPV1-specific neutralizing antibody responses generated by mice vaccinated with CRT/E7 and/or BPV1 L1 DNA. C57BL/6 mice (three per group) were immunized on days 1, 15, and 30 intradermally using a gene gun with 2 μ g of DNA per mouse of CRT/E7 and/or BPV1 L1 DNA. In vitro neutralization assays were performed using BPV1 L1 pseudovirus on twofold dilutions of antisera collected from the mice 2 weeks after the final immunization. Endpoint titers achieving 50% neutralization are plotted and the means shown as horizontal lines. PI=Pre Immune; L1=BPV1 L1 DNA vaccine; E7=CRT/E7 DNA vaccine.

Figure 6 contains three panels (A), (B), and (C) describing the comparison of OVA-specific CD8⁺ T cell responses induced by pcDNA3-OVA vaccination with or without co-administration of BPV L1 or L2. **Panel (A)** depicts a schematic illustration of the experiment. Briefly, 5~8 weeks old female C57BL/6 mice (3 mice/group) were vaccinated with 10 μ g/mouse of pcDNA3-OVA, with either 10 μ g/mouse of pcDNA3, or pcDNA3-BPVL1, or pcDNA3-BPVL2 via intramuscular injection. The mice were boosted with the same regimen once after one week. 7 days after the last vaccination, PBMCs were collected from peripheral blood, stained with FITC-conjugated anti-mouse CD8a antibody, PE-conjugated OVA peptide (SIINFEKL) loaded H-2K^b tetramer. The data were acquired with FACSCalibur and analyzed with CellQuest. **Panel (B)** depicts representative flow

cytometry image of PBMC staining. **Panel (C)** depicts a summary of the flow cytometry data.

DETAILED DESCRIPTION

5 The inventors of the present invention have determined that that irradiated tumor microenvironments are ideal for triggering CD8-dependent antitumor effects by intratumoral administration of antigen without additional adjuvant. Using radiation therapy in combination with a therapeutic HPV recombinant vaccinia vaccine, TA-HPV, and a therapeutic HPV fusion protein vaccine, TA-CIN, to trigger E6- and E7-specific CD8+ T
10 cell immune response. Accordingly, the present invention is drawn to methods for enhancing an antigen-specific immune response in a mammal using combination of radiation therapy with local injection of therapeutic vaccines.

Partial List of Abbreviations

15 ANOVA, analysis of variance; APC, antigen presenting cell; CRT, calreticulin; CTL, cytotoxic T lymphocyte; DC, dendritic cell; DLN; draining lymph nodes; E6, HPV oncoprotein E6; E7, HPV oncoprotein E7; ELISA, enzyme-linked immunosorbent assay; HPV, human papillomavirus; IFN γ , interferon- γ ; i.m., intramuscular(ly); i.t., intratumoral(ly); i.v., intravenous(ly); luc, luciferase; mAB, monoclonal antibody; MOI,
20 multiplicity of infection; OVA, ovalbumin; p-, plasmid-; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; SD, standard deviation; TAA, tumor-associate antigen; WT, wild-type.

Capsid antigens

25 Papillomaviruses encode two capsid proteins, L1 and L2. Upon expression, the major capsid protein, L1, can spontaneously self-assemble into pentamers that further assemble into VLPs, comprised of 360 copies of L1. The minor capsid protein, L2, is not required for VLP formation, but is required for formation of infectious virions (and pseudovirions). Up to 72 copies of L2 can be incorporated in a VLP, and viral infectivity
30 correlates with L2 content (Buck, CB et al. 2008 *J Virol* 82:5190-7).

Certain aspects of the invention comprise codon-optimized capsid antigens derived from the L1 or L2 protein of different strains of HPV or papillomavirus from a variety of non-human (animal) species as otherwise set forth herein, and is highly immunogenic.

The capsid antigens may be derived or isolated from a number of different strains of papillomavirus, including strains 16, 45 and/or 58 as well as HPV strains 1, 5, 6, 11 and 18. Additional strains include, but are not limited to, papilloma virus strains HPV 1, HPV2, HPV5, HPV6, HPV8, HPV 11, HPV16, HPV18, HPV31, HPV33, HPV35, HPV39, HPV45, HPV51, HPV52, HPV56, HPV58, HPV59, HPV66, HPV68, HPV73, CRPV and BPV1 and including HPV strains 16, 45 and/or 58, as well as of HPV strains 1, 5, 6, 11 and 48, it is noted that the L1 or L2 sequence of HPV16 and BPV1 are relatively conserved across diverse papillomavirus isolates and the use of each of these sequences within the DNA vector according to the present invention may instill immunogenic protection and enhance antigenic immune responses.

In certain embodiments, L2 or L1 peptide sequences as described above are from any of HPV strains HPV1, HPV2, HPV3, HPV4, HPV5, HPV6, HPV7, HPV8, HPV9, HPV10, HPV11, HPV12, HPV13, HPV14, HPV15, HPV16, HPV17, HPV18, HPV19, HPV20, HPV21, HPV22, HPV23, HPV24, HPV25, HPV26, HPV27, HPV28, HPV29, HPV30, HPV31, HPV32, HPV33, HPV34, HPV35, HPV36, HPV37, HPV38, HPV39, HPV40, HPV41, HPV42, HPV43, HPV44, HPV45, HPV46, HPV47, HPV48, HPV49, HPV50, HPV51, HPV52, HPV53, HPV54, HPV55, HPV56, HPV57, HPV58, HPV59, HPV60, HPV61, HPV62, HPV63, HPV64, HPV65, HPV66, HPV67, HPV68, HPV69, HPV70, HPV71, HPV72, HPV73, HPV74, HPV75, HPV76, HPV77, HPV78, HPV79, HPV80, HPV81, HPV82, HPV83, HPV84, HPV85, HPV86, HPV87, HPV88, HPV89, HPV90, HPV91, HPV92, HPV93, HPV94, HPV95, HPV96, HPV97, HPV98, HPV99, HPV100; and/or animal papillomaviruses: bovine papillomavirus type 1 (BPV1), bovine papillomavirus type 2 (BPV2), bovine papillomavirus type 4 (BPV4), cottontail rabbit papillomavirus (CRPV), deer papillomavirus (DPV), European elk papillomavirus (EEPV), canine oral papillomavirus (COPV), Rhesus monkey papillomavirus (RnPV) and rabbit oral papillomavirus (ROPV) may be used in the DNA vectors according to the present invention.

Representative L2 and L2 capsid antigens for use according to the present invention include the following for each of the above-referenced HPV strains and animal PVs (note that all of the amino acid numbers listed below are referenced based upon the sequence of HPV 16 L2 or L1 and BPV1 L1 or L2, and that for each of the lengths of amino acids, each length within the length range is understood to be disclosed):

In certain embodiments, a L1 capsid antigen of BPV1 may comprise the nucleotide sequence set forth in SEQ ID NO: 122. In certain embodiments, a L1 capsid antigen of BPV1 may comprise the amino acid sequence set forth in SEQ ID NO: 123. In certain embodiments, a L1 capsid antigen of HPV16 may comprise the nucleotide sequence set forth in SEQ ID NO: 118. In certain embodiments, a L1 capsid antigen of HPV16 may comprise the amino acid sequence set forth in SEQ ID NO: 119. In certain embodiments, a L2 capsid antigen of BPV1 may comprise the nucleotide sequence set forth in SEQ ID NO: 124. In certain embodiments, a L2 capsid antigen of BPV1 may comprise the amino acid sequence set forth in SEQ ID NO: 125. In certain embodiments, a L2 capsid antigen of HPV16 may comprise the nucleotide sequence set forth in SEQ ID NO: 120. In certain embodiments, a L2 capsid antigen of HPV16 may comprise the amino acid sequence set forth in SEQ ID NO: 121.

The term "% homology" is used interchangeably herein with the term "% identity" herein and normally refers to the level of nucleic acid identity between the nucleic acid sequence of the DNA or RNA aptamers of the present inventions, when aligned using a sequence alignment program.

For example, as used herein, 80% homology means the same thing as 80% sequence identity determined by a defined algorithm, and accordingly a homologue of a given sequence has greater than 80% sequence identity over a length of the given sequence. Exemplary levels of sequence identity include, but are not limited to, 80, 85, 90, 95, 98% or more sequence identity to a given sequence as described herein.

Exemplary computer programs which can be used to determine identity between two sequences include, but are not limited to, the suite of BLAST programs, *e.g.*, BLASTN, BLASTX, and TBLASTX, BLASTP and TBLASTN, publicly accessible at www.ncbi.nlm.nih.gov/BLAST.

In certain embodiment, the method comprising co-administering to the mammal an effective amount of a combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen, said capsid antigen comprises at least 50% identical to a nucleotide sequence selected from the group consisting of : SEQ ID Nos. 118, 120, 122, and 124, or any combinations thereof. In certain embodiments, said capsid antigen comprises a nucleotide sequence that is at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98.1%, 98.2%, 98.3%, 98.4%, 98.5%, 98.6%, 98.7%, 98.8%, 98.9%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%,

99.8%, or 99.9%, identical to a nucleotide sequence selected from the group consisting of : SEQ ID Nos. 118,120, 122, and 124, or any combinations thereof.

In certain embodiment, the method comprising co-administering to the mammal an effective amount of a combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen, said capsid antigen comprises at least 50% identical to a amino acid sequence selected from the group consisting of : SEQ ID Nos. 118,121, 123, and 125, or any combinations thereof. In certain embodiments, said capsid antigen comprises a nucleotide sequence that is at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98.1%, 98.2%, 98.3%, 98.4%, 98.5%, 98.6%, 98.7%, 98.8%, 98.9%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9%, identical to a nucleotide sequence selected from the group consisting of : SEQ ID Nos. 119,121, 123, and 125, or any combinations thereof.

Other L1 and L2 sequences are readily obtained from the detailed description of the invention.

Nucleic acid (e.g., DNA) vaccines

Vaccines that may be administered to a mammal include any vaccine, e.g., a nucleic acid vaccine (e.g., a DNA vaccine). In an embodiment of the invention, a nucleic acid vaccine will encode an antigen, e.g., an antigen against which an immune response is desired. Other nucleic acids that may be used are those that increase or enhance an immune reaction, but which do not encode an antigen against which an immune reaction is desired. These vaccines are further described below.

Exemplary antigens include proteins or fragments thereof from a pathogenic organism, e.g., a bacterium or virus or other microorganism, as well as proteins or fragments thereof from a cell, e.g., a cancer cell. In one embodiment, the antigen is from a virus, such as class human papillomavirus (HPV), e.g., E7 or E6. These proteins are also oncogenic proteins, which are important in the induction and maintenance of cellular transformation and co-expressed in most HPV-containing cervical cancers and their precursor lesions. Therefore, cancer vaccines that target E7 or E6 can be used to control of HPV-associated neoplasms (Wu, T-C, *Curr Opin Immunol.* 6:746-54, 1994).

However, as noted, the present invention is not limited to the exemplified antigen(s). Rather, one of skill in the art will appreciate that the same results are expected for any antigen (and epitopes thereof) for which a T cell-mediated response is desired. The

response so generated will be effective in providing protective or therapeutic immunity, or both, directed to an organism or disease in which the epitope or antigenic determinant is involved – for example as a cell surface antigen of a pathogenic cell or an envelope or other antigen of a pathogenic virus, or a bacterial antigen, or an antigen expressed as or as part of a pathogenic molecule.

Exemplary antigens and their sequences are set forth below.

E7 Protein from HPV-16

The E7 nucleic acid sequence (SEQ ID NO:1) and amino acid sequence (SEQ ID NO:2) from HPV-16 are shown herein (see GenBank Accession No. NC_001526). The single letter code, the wild type E7 amino acid sequence (SEQ ID NO:2) is shown herein.

In another embodiment (See GenBank Accession No. AF125673, nucleotides 562-858 and the E7 amino acid sequence), the C-terminal four amino acids QDKL (and their codons) above are replaced with the three amino acids QKP (and the codons cag aaa cca), yielding a protein of 98 residues.

When an oncoprotein or an epitope thereof is the immunizing moiety, it is preferable to reduce the tumorigenic risk of the vaccine itself. Because of the potential oncogenicity of the HPV E7 protein, the E7 protein may be used in a “detoxified” form.

To reduce oncogenic potential of E7 in a construct of the present invention, one or more of the following positions of E7 is mutated:

Original residue	Mutant residue	Preferred codon mutation	nt Position (in SEQ ID NO:1)	Amino acid (in SEQ ID NO: 2)
Cys	Gly (or Ala)	TGT→GGT	70	24
Glu	Gly (or Ala)	GAG→GGG (or GCG)	77	26
Cys	Gly (or Ala)	TGC→GGC	271	91

In one embodiment, the E7 (detox) mutant sequence has the following two mutations:

a TGT→GGT mutation resulting in a Cys→Gly substitution at position 24 of SEQ ID NO: 9 and GAG→GGG mutation resulting in a Glu→Gly substitution at position 26 of the wild type E7. This mutated amino acid sequence is shown herein as SEQ ID NO:3.

These substitutions completely eliminate the capacity of the E7 to bind to Rb, and thereby nullify its transforming activity. Any nucleotide sequence that encodes the above E7 or E7(detox) polypeptide, or an antigenic fragment or epitope thereof, can be used in the

present compositions and methods, including the E7 and E7(detox) sequences which are shown herein.

E6 Protein from HPV-16

5 The wild type E6 nucleotide (SEQ ID NO:4) and amino acid sequences (SEQ ID NO:5) are shown herein (see GenBank accession Nos. K02718 and NC_001526). This polypeptide has 158 amino acids and is shown herein in single letter code as SEQ ID NO:5.

10 E6 proteins from cervical cancer-associated HPV types such as HPV-16 induce proteolysis of the p53 tumor suppressor protein through interaction with E6-AP. Human mammary epithelial cells (MECs) immortalized by E6 display low levels of p53. HPV-16 E6, as well as other cancer-related papillomavirus E6 proteins, also binds the cellular protein E6BP (ERC-55). As with E7, described below a non-oncogenic mutated form of E6 may be used, referred to as “E6(detox).” Several different E6 mutations and publications
15 describing them are discussed below.

 The amino acid residues to be mutated are underscored in the E6 amino acid sequence provided herein. Some studies of E6 mutants are based upon a shorter E6 protein of 151 nucleic acids, wherein the N-terminal residue was considered to be the Met at position 8 in the wild type E6. That shorter version of E6 is shown herein as SEQ ID NO:6.

20 To reduce oncogenic potential of E6 in a construct, one or more of the following positions of E6 is mutated:

Original residue	Mutant residue	aa position in SEQ ID NO: 5	aa position in SEQ ID NO: 6
Cys	Gly (or Ala)	70	63
Cys	Gly (or Ala)	113	106
Ile	Thr	135	128

 Nguyen *et al.*, *J Virol.* 6:13039-48, 2002, described a mutant of HPV-16 E6 deficient in binding α -helix partners which displays reduced oncogenic potential *in vivo*.

25 This mutant, which includes a replacement of Ile with Thr as position 128 (of SEQ ID NO: 6), may be used in accordance with the present invention to make an E6 DNA vaccine that has a lower risk of being oncogenic. This E6(I¹²⁸T) mutant is defective in its ability to bind at least a subset of α -helix partners, including E6AP, the ubiquitin ligase that mediates E6-dependent degradation of the p53 protein.

Cassetti MC *et al.*, *Vaccine* 22:520-52, 2004, examined the effects of mutations four or five amino acid positions in E6 and E7 to inactivate their oncogenic potential. The following mutations were examined: E6-C⁶³G and E6 C¹⁰⁶G (positions based on the wild type E6); E7-C²⁴G, E7-E²⁶G, and E7 C⁹¹G (positions based on the wild type E7).

- 5 Venezuelan equine encephalitis virus replicon particle (VRP) vaccines encoding mutant or wild type E6 and E7 proteins elicited comparable CTL responses and generated comparable antitumor responses in several HPV16 E6(+)E7(+) tumor challenge models: protection from either C3 or TC-1 tumor challenge was observed in 100% of vaccinated mice. Eradication of C3 tumors was observed in approximately 90% of the mice. The predicted
- 10 inactivation of E6 and E7 oncogenic potential was confirmed by demonstrating normal levels of both p53 and Rb proteins in human mammary epithelial cells infected with VRPs expressing mutant E6 and E7 genes.

- The HPV16 E6 protein contains two zinc fingers important for structure and function; one cysteine (C) amino acid position in each pair of C-X-X-C (where X is any
- 15 amino acid) zinc finger motifs may be mutated at E6 positions 63 and 106 (based on the wild type E6). Mutants are created, for example, using the Quick Change Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). HPV16 E6 containing a single point mutation in the codon for Cys¹⁰⁶ in the wild type E6 (= Cys 113 in the wild type E6). Cys¹⁰⁶ neither binds nor facilitates degradation of p53 and is incapable of immortalizing human mammary
- 20 epithelial cells (MEC), a phenotype dependent upon p53 degradation. A single amino acid substitution at position Cys⁶³ of the wild type E6 (=Cys⁷⁰ in the wild type E6) destroys several HPV16 E6 functions: p53 degradation, E6TP-1 degradation, activation of telomerase, and, consequently, immortalization of primary epithelial cells.

- Any nucleotide sequence that encodes these E6 polypeptides, one of the mutants
- 25 thereof, or an antigenic fragment or epitope thereof, can be used in the present invention. Other mutations can be tested and used in accordance with the methods described herein including those described in Cassetti *et al.*, *supra*. These mutations can be produced from any appropriate starting sequences by mutation of the coding DNA.

- The present invention also includes the use of a tandem E6-E7 vaccine, using one or
- 30 more of the mutations described herein to render the oncoproteins inactive with respect to their oncogenic potential *in vivo*. VRP vaccines (described in Cassetti *et al.*, *supra*) comprised fused E6 and E7 genes in one open reading frame which were mutated at four or five amino acid positions. Thus, the present constructs may include one or more epitopes

of E6 and E7, which may be arranged in their native order or shuffled in any way that permits the expressed protein to bear the E6 and E7 antigenic epitopes in an immunogenic form. DNA encoding amino acid spacers between E6 and E7 or between individual epitopes of these proteins may be introduced into the vector, provided again, that the
5 spacers permit the expression or presentation of the epitopes in an immunogenic manner after they have been expressed by transduced host cells.

Influenza hemagglutinin (HA)

A nucleic acid sequence encoding HA is shown herein as SEQ ID NO: 7. The
10 amino acid sequence of HA is shown herein as SEQ ID NO: 8, with the immunodominant epitope underscored.

Ovalbumin (OVA)

An amino acid sequence encoding a representative OVA is shown herein as
15 SEQ ID NO:9.

Other exemplary antigens

Exemplary antigens are epitopes of pathogenic microorganisms against which the host is defended by effector T cells responses, including CTL and delayed type
20 hypersensitivity. These typically include viruses, intracellular parasites such as malaria, and bacteria that grow intracellularly such as *Mycobacterium* and *Listeria* species. Thus, the types of antigens included in the vaccine compositions used in the present invention may be any of those associated with such pathogens as well as tumor-specific antigens. It is noteworthy that some viral antigens are also tumor antigens in the case where the virus is
25 a causative factor in the tumor.

In fact, the two most common cancers worldwide, hepatoma and cervical cancer, are associated with viral infection. Hepatitis B virus (HBV) (Beasley, R.P. *et al.*, *Lancet* 2:1129-1133 (1981) has been implicated as etiologic agent of hepatomas. About 80-90% of cervical cancers express the E6 and E7 antigens (discussed above and exemplified
30 herein) from one of four “high risk” human papillomavirus types: HPV-16, HPV-18, HPV-31 and HPV-45 (Gissmann, L. *et al.*, *Ciba Found Symp.* 120:190-207, 1986; Beaudenon, S., *et al. Nature* 321:246-9, 1986, incorporated by reference herein). The HPV E6 and E7 antigens are the most promising targets for virus associated cancers in immunocompetent

individuals because of their ubiquitous expression in cervical cancer. In addition to their importance as targets for therapeutic cancer vaccines, virus-associated tumor antigens are also ideal candidates for prophylactic vaccines. Indeed, introduction of prophylactic HBV vaccines in Asia have decreased the incidence of hepatoma (Chang, MH *et al.* *New Engl. J. Med.* 336, 1855-1859 (1997), representing a great impact on cancer prevention.

Among the most important viruses in chronic human viral infections are HPV, HBV, hepatitis C Virus (HCV), retroviruses such as human immunodeficiency virus (HIV-1 and HIV-2), herpes viruses such as Epstein Barr Virus (EBV), cytomegalovirus (CMV), HSV-1 and HSV-2, and influenza virus. Useful antigens include HBV surface antigen or HBV core antigen; ppUL83 or pp89 of CMV; antigens of gp120, gp41 or p24 proteins of HIV-1; ICP27, gD2, gB of HSV; or influenza hemagglutinin or nucleoprotein (Anthony, LS *et al.*, *Vaccine* 1999; 17:373-83). Other antigens associated with pathogens that can be utilized as described herein are antigens of various parasites, including malaria, *e.g.*, malaria peptide based on repeats of NANP.

In certain embodiments, the invention includes methods using foreign antigens in which individuals may have existing T cell immunity (such as influenza, tetanus toxin, herpes etc). In other embodiments, the skilled artisan would readily be able to determine whether a subject has existing T cell immunity to a specific antigen according to well known methods available in the art and use a foreign antigen to which the subject does not already have an existing T cell immunity.

In alternative embodiments, the antigen is from a pathogen that is a bacterium, such as *Bordetella pertussis*; *Ehrlichia chaffeensis*; *Staphylococcus aureus*; *Toxoplasma gondii*; *Legionella pneumophila*; *Brucella suis*; *Salmonella enterica*; *Mycobacterium avium*; *Mycobacterium tuberculosis*; *Listeria monocytogenes*; *Chlamydia trachomatis*; *Chlamydia pneumoniae*; *Rickettsia rickettsii*; or, a fungus, such as, *e.g.*, *Paracoccidioides brasiliensis*; or other pathogen, *e.g.*, *Plasmodium falciparum*.

As used herein, the term “cancer” includes, but is not limited to, solid tumors and blood borne tumors. The term cancer includes diseases of the skin, tissues, organs, bone, cartilage, blood and vessels. A term used to describe cancer that is far along in its growth, also referred to as “late stage cancer” or “advanced stage cancer,” is cancer that is metastatic, *e.g.*, cancer that has spread from its primary origin to another part of the body. In certain embodiments, advanced stage cancer includes stages 3 and 4 cancers. Cancers are ranked into stages depending on the extent of their growth and spread through the body;

stages correspond with severity. Determining the stage of a given cancer helps doctors to make treatment recommendations, to form a likely outcome scenario for what will happen to the patient (prognosis), and to communicate effectively with other doctors.

There are multiple staging scales in use. One of the most common ranks cancers
5 into five progressively more severe stages: 0, I, II, III, and IV. Stage 0 cancer is cancer that is just beginning, involving just a few cells. Stages I, II, III, and IV represent progressively more advanced cancers, characterized by larger tumor sizes, more tumors, the aggressiveness with which the cancer grows and spreads, and the extent to which the cancer has spread to infect adjacent tissues and body organs.

10 Another popular staging system is known as the TNM system, a three dimensional rating of cancer extensiveness. Using the TNM system, doctors rate the cancers they find on each of three scales, where T stands for tumor size, N stands for lymph node involvement, and M stands for metastasis (the degree to which cancer has spread beyond its original locations). Larger scores on each of the three scales indicate more advanced
15 cancer. For example, a large tumor that has not spread to other body parts might be rated T3, N0, M0, while a smaller but more aggressive cancer might be rated T2, N2, M1 suggesting a medium sized tumor that has spread to local lymph nodes and has just gotten started in a new organ location.

Cancers that may be treated by the methods of the present invention include, but are
20 not limited to, cancer cells from the bladder, blood, bone, bone marrow, brain, breast, colon, esophagus, gastrointestinal, gum, head, kidney, liver, lung, nasopharynx, neck, ovary, prostate, skin, stomach, testis, tongue, or uterus. In addition, the cancer may specifically be of the following histological type, though it is not limited to these: neoplasm, malignant; carcinoma; carcinoma, undifferentiated; giant and spindle cell carcinoma; small cell
25 carcinoma; papillary carcinoma; squamous cell carcinoma; lymphoepithelial carcinoma; basal cell carcinoma; pilomatrix carcinoma; transitional cell carcinoma; papillary transitional cell carcinoma; adenocarcinoma; gastrinoma, malignant; cholangiocarcinoma; hepatocellular carcinoma; combined hepatocellular carcinoma and cholangiocarcinoma; trabecular adenocarcinoma; adenoid cystic carcinoma; adenocarcinoma in adenomatous
30 polyp; adenocarcinoma, familial polyposis coli; solid carcinoma; carcinoid tumor, malignant; branchiolo-alveolar adenocarcinoma; papillary adenocarcinoma; chromophobe carcinoma; acidophil carcinoma; oxyphilic adenocarcinoma; basophil carcinoma; clear cell adenocarcinoma; granular cell carcinoma; follicular adenocarcinoma; papillary and

follicular adenocarcinoma; nonencapsulating sclerosing carcinoma; adrenal cortical carcinoma; endometroid carcinoma; skin appendage carcinoma; apocrine adenocarcinoma; sebaceous adenocarcinoma; ceruminous adenocarcinoma; mucoepidermoid carcinoma; cystadenocarcinoma; papillary cystadenocarcinoma; papillary serous cystadenocarcinoma;

5 mucinous cystadenocarcinoma; mucinous adenocarcinoma; signet ring cell carcinoma; infiltrating duct carcinoma; medullary carcinoma; lobular carcinoma; inflammatory carcinoma; paget's disease, mammary; acinar cell carcinoma; adenosquamous carcinoma; adenocarcinoma w/squamous metaplasia; thymoma, malignant; ovarian stromal tumor, malignant; thecoma, malignant; granulosa cell tumor, malignant; and roblastoma,

10 malignant; sertoli cell carcinoma; leydig cell tumor, malignant; lipid cell tumor, malignant; paraganglioma, malignant; extra-mammary paraganglioma, malignant; pheochromocytoma; glomangiosarcoma; malignant melanoma; amelanotic melanoma; superficial spreading melanoma; malig melanoma in giant pigmented nevus; epithelioid cell melanoma; blue nevus, malignant; sarcoma; fibrosarcoma; fibrous histiocytoma, malignant; myxosarcoma;

15 liposarcoma; leiomyosarcoma; rhabdomyosarcoma; embryonal rhabdomyosarcoma; alveolar rhabdomyosarcoma; stromal sarcoma; mixed tumor, malignant; mullerian mixed tumor; nephroblastoma; hepatoblastoma; carcinosarcoma; mesenchymoma, malignant; brenner tumor, malignant; phyllodes tumor, malignant; synovial sarcoma; mesothelioma, malignant; dysgerminoma; embryonal carcinoma; teratoma, malignant; struma ovarii,

20 malignant; choriocarcinoma; mesonephroma, malignant; hemangiosarcoma; hemangioendothelioma, malignant; kaposi's sarcoma; hemangiopericytoma, malignant; lymphangiosarcoma; osteosarcoma; juxtacortical osteosarcoma; chondrosarcoma; chondroblastoma, malignant; mesenchymal chondrosarcoma; giant cell tumor of bone; ewing's sarcoma; odontogenic tumor, malignant; ameloblastic odontosarcoma;

25 ameloblastoma, malignant; ameloblastic fibrosarcoma; pinealoma, malignant; chordoma; glioma, malignant; ependymoma; astrocytoma; protoplasmic astrocytoma; fibrillary astrocytoma; astroblastoma; glioblastoma; oligodendroglioma; oligodendroblastoma; primitive neuroectodermal; cerebellar sarcoma; ganglioneuroblastoma; neuroblastoma; retinoblastoma; olfactory neurogenic tumor; meningioma, malignant; neurofibrosarcoma;

30 neurilemmoma, malignant; granular cell tumor, malignant; malignant lymphoma; Hodgkin's disease; Hodgkin's lymphoma; paraganuloma; malignant lymphoma, small lymphocytic; malignant lymphoma, large cell, diffuse; malignant lymphoma, follicular; mycosis fungoides; other specified non-Hodgkin's lymphomas; malignant histiocytosis; multiple

myeloma; mast cell sarcoma; immunoproliferative small intestinal disease; leukemia; lymphoid leukemia; plasma cell leukemia; erythroleukemia; lymphosarcoma cell leukemia; myeloid leukemia; basophilic leukemia; eosinophilic leukemia; monocytic leukemia; mast cell leukemia; megakaryoblastic leukemia; myeloid sarcoma; and hairy cell leukemia.

5 In addition to its applicability to human cancer and infectious diseases, the present invention is also intended for use in treating animal diseases in the veterinary medicine context. Thus, the approaches described herein may be readily applied by one skilled in the art for treatment of veterinary herpes virus infections including equine herpes viruses, bovine viruses such as bovine viral diarrhea virus (for example, the E2 antigen), bovine
10 herpes viruses, Marek's disease virus in chickens and other fowl; animal retroviral and lentiviral diseases (*e.g.*, feline leukemia, feline immunodeficiency, simian immunodeficiency viruses, *etc.*); pseudorabies and rabies; and the like.

As for tumor antigens, any tumor-associated or tumor-specific antigen (or tumor cell derived epitope) (collectively, TAA) that can be recognized by T cells, including CTL,
15 can be used. These include, without limitation, mutant p53, HER2/neu or a peptide thereof, or any of a number of melanoma-associated antigens such as MAGE-1, MAGE-3, MART-1/Melan-A, tyrosinase, gp75, gp100, BAGE, GAGE-1, GAGE-2, GnT-V, and p15 (see, for example, US Pat. 6,187,306, incorporated herein by reference).

In one embodiment, it is not necessary to include a full length antigen in a nucleic acid vaccine; it suffices to include a fragment that will be presented by MHC class I and/or
20 II. A nucleic acid may include 1, 2, 3, 4, 5 or more antigens, which may be the same or different ones.

Approaches for Mutagenesis of E6, E7, and other Antigens

25 Mutants of the antigens described here may be created, for example, using the Quick Change Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). Generally, antigens that may be used herein may be proteins or peptides that differ from the naturally-occurring proteins or peptides but yet retain the necessary epitopes for functional activity. In certain embodiments, an antigen may comprise, consist essentially of, or consist of an
30 amino acid sequence that is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to that of the naturally-occurring antigen or a fragment thereof. In certain embodiments, an antigen may also comprise, consist essentially of, or consist of an amino acid sequence that is encoded by a nucleotide

sequence that is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to a nucleotide sequence encoding the naturally-occurring antigen or a fragment thereof. In certain embodiments, an antigen may also comprise, consist essentially of, or consist of an amino acid sequence that is encoded by a nucleic acid that hybridizes under high stringency conditions to a nucleic acid encoding the naturally-occurring antigen or a fragment thereof. Hybridization conditions are further described herein.

In one embodiment, an exemplary protein may comprise, consist essentially of, or consist of, an amino acid sequence that is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to that of a viral protein, including for example E6 or E7, such as an E6 or E7 sequence provided herein. Where the E6 or E7 protein is a detox E6 or E7 protein, the amino acid sequence of the protein may comprise, consist essentially of, or consist of an amino acid sequence that is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to that of an E6 or E7 protein, wherein the amino acids that render the protein a “detox” protein are present.

Exemplary nucleic acid (e.g., DNA) vaccines encoding an Immunogenicity-Potentiating Polypeptide (IPP) and an antigen

In one embodiment, a nucleic acid vaccine encodes a fusion protein comprising an antigen and a second protein, e.g., an IPP. An IPP may act in potentiating an immune response by promoting: processing of the linked antigenic polypeptide via the MHC class I pathway or targeting of a cellular compartment that increases the processing. This basic strategy may be combined with an additional strategy pioneered by the present inventors and colleagues, that involve linking DNA encoding another protein, generically termed a “targeting polypeptide,” to the antigen-encoding DNA. Again, for the sake of simplicity, the DNA encoding such a targeting polypeptide will be referred to herein as a “targeting DNA.” That strategy has been shown to be effective in enhancing the potency of the vectors carrying only antigen-encoding DNA. See for example, the following PCT publications by Wu *et al.*: WO 01/29233; WO 02/009645; WO 02/061113; WO 02/074920; and WO 02/12281, all of which are incorporated by reference in their entirety. The other strategies include the use of DNA encoding polypeptides that promote or enhance:

- (a) development, accumulation or activity of antigen presenting cells or targeting of antigen to compartments of the antigen presenting cells leading to enhanced antigen presentation;
- (b) intercellular transport and spreading of the antigen;
- 5 (c) sorting of the lysosome-associated membrane protein type 1 (Sig/LAMP-1); or
- (d) any combination of (a)-(c).

The strategy includes use of:

- (a) a viral intercellular spreading protein selected from the group of herpes simplex virus-1 VP22 protein, Marek's disease virus UL49 (see WO 02/09645 and US Patent No. 7,318,928), protein or a functional homologue or derivative thereof;
- (b) calreticulin (CRT) and other endoplasmic reticulum chaperone polypeptides selected from the group of CRT-like molecules ER60, GRP94, gp96, or a functional homologue or derivative thereof (see WO 02/12281 and US Patent No. 7,3442,002);
- 15 (c) a cytoplasmic translocation polypeptide domains of a pathogen toxin selected from the group of domain II of *Pseudomonas* exotoxin ETA or a functional homologue or derivative thereof (see published US application 20040086845);
- (d) a polypeptide that targets the centrosome compartment of a cell selected from γ -tubulin or a functional homologue or derivative thereof;
- 20 (e) a polypeptide that stimulates dendritic cell precursors or activates dendritic cell activity selected from the group of GM-CSF, Flt3-ligand extracellular domain, or a functional homologue or derivative thereof;
- (f) a costimulatory signal, such as a B7 family protein, including B7-DC (see U.S. Serial No. 09/794,210), B7.1, B7.2, soluble CD40, *etc.*); or
- 25 (g) an anti-apoptotic polypeptide selected from the group consisting of (1) BCL-xL, (2) BCL2, (3) XIAP, (4) FLICEc-s, (5) dominant-negative caspase-8, (6) dominant negative caspase-9, (7) SPI-6, and (8) a functional homologue or derivative of any of (1)-(7). (See WO 2005/047501).

The following publications, all of which are incorporated by reference in their entirety, describe IPPs: Kim TW *et al.*, *J Clin Invest* 112: 109-117, 2003; Cheng WF *et al.*, *J Clin Invest* 108: 669-678, 2001; Hung CF *et al.*, *Cancer Res* 61:3698-3703, 2001; Chen CH *et al.*, 2000, *supra*; US Pat. 6,734,173; published patent applications WO05/081716, WO05/047501, WO03/085085, WO02/12281, WO02/074920, WO02/061113,

WO02/09645, and WO01/29233. Comparative studies of these IPPs using HPV E6 as the antigen are described in Peng, S. *et al.*, *J Biomed Sci.* 12:689-700 2005.

An antigen may be linked N-terminally or C-terminally to an IPP. Exemplary IPPs and fusion constructs encoding such are described below.

5

Lysosomal Associated Membrane Protein 1 (LAMP-1)

The DNA sequence encoding the E7 protein fused to the translocation signal sequence and LAMP-1 domain (Sig-E7-LAMP-1) is shown herein as SEQ ID NO:10. The amino acid sequence of Sig-E7-LAMP-1 is shown herein as SEQ ID

10 NO:11.

The nucleotide sequence of the immunogenic vector pcDNA3-Sig/E7/LAMP-1 is shown herein as SEQ ID NO:13, with the SigE7-LAMP-1 coding sequence in lower case and underscored.

15 *HSP70 from M. tuberculosis*

The nucleotide sequence encoding HSP70 is shown herein as SEQ ID NO:13) (i.e., nucleotides 10633-12510 of the *M. tuberculosis* genome in GenBank NC_000962). The amino acid sequence of HSP70 is shown herein as SEQ ID NO:14.

The nucleic acid sequences encoding the E7-Hsp70 chimera/fusion polypeptides are shown herein as SEQ ID NO:15 and the corresponding amino acid sequence is shown herein as SEQ ID NO:16. The E7 coding sequence is shown in upper case and underscored.

25 *ETA(dII) from Pseudomonas aeruginosa*

The complete coding sequence for *Pseudomonas aeruginosa* exotoxin type A (ETA) is shown herein as SEQ ID NO:17 (GenBank Accession No. K01397). The amino acid sequence of ETA is shown herein as SEQ ID NO:18 (GenBank Accession No. K01397).

Residues 1-25 (italicized) represent the signal peptide. The first residue of the mature polypeptide, Ala, is bolded/underscored. The mature polypeptide is residues 26-638 of SEQ ID NO:18.

Domain II (ETA(II)), translocation domain (underscored above) spans residues 247-417 of the mature polypeptide (corresponding to residues 272-442 of SEQ ID NO:18) and is presented below separately herein as SEQ ID NO:19.

The nucleotide construct in which ETA(dII) is fused to HPV-16 E7 is shown herein as SEQ ID NO:20. The corresponding amino acid sequence is shown herein as SEQ ID NO:21. The ETA(dII) sequence appears in plain font, extra codons from plasmid pcDNA3 are italicized. Nucleotides between ETA(dII) and E7 are also bolded (and result in the interposition of two amino acids between ETA(dII) and E7). The E7 amino acid sequence is underscored (ends with Gln at position 269).

Pro Leu Ile Ser Leu Asp Cys Ala Phe AMB

The nucleotide sequence of the pcDNA3 vector encoding E7 and HSP70 (pcDNA3-E7-Hsp70) is shown herein as SEQ ID NO:22.

Calreticulin (CRT)

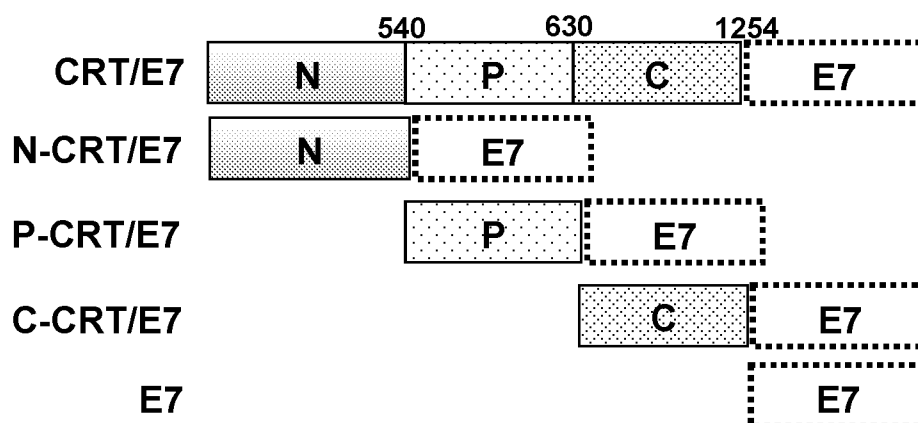
Calreticulin (CRT), a well-characterized ~46 kDa protein was described briefly above, as were a number of its biological and biochemical activities. As used herein, “calreticulin” or “CRT” refers to polypeptides and nucleic acids molecules having substantial identity to the exemplary human CRT sequences as described herein or homologues thereof, such as rabbit and rat CRT - well-known in the art. A CRT polypeptide is a polypeptide comprising a sequence identical to or substantially identical to the amino acid sequence of CRT. An exemplary nucleotide and amino acid sequence for a CRT used in the present compositions and methods are presented below. The terms “calreticulin” or “CRT” encompass native proteins as well as recombinantly produced modified proteins that, when fused with an antigen (at the DNA or protein level) promote the induction of immune responses and promote angiogenesis, including a CTL response. Thus, the terms “calreticulin” or “CRT” encompass homologues and allelic variants of human CRT, including variants of native proteins constructed by *in vitro* techniques, and proteins isolated from natural sources. The CRT polypeptides used in the present invention, and sequences encoding them, also include fusion proteins comprising non-CRT sequences, particularly MHC class I-binding peptides; and also further comprising other domains, *e.g.*, epitope tags, enzyme cleavage recognition sequences, signal sequences, secretion signals and the like.

A human CRT coding sequence is shown herein as SEQ ID NO: 23. The amino acid sequence of the human CRT protein encoded by SEQ ID NO:23 is set forth herein as SEQ ID NO:24. This amino acid sequence is highly homologous to GenBank Accession No. NM 004343.

The amino acid sequence of the rabbit and rat CRT proteins are set forth in GenBank Accession Nos. P1553 and NM 022399, respectively. An alignment of human, rabbit and rat CRT shows that these proteins are highly conserved, and most of the amino acid differences between species are conservative in nature. Most of the variation is found in the alignment of the approximately 36 C-terminal residues. Thus, for the present invention, human CRT may be used as well as, DNA encoding any homologue of CRT from any species that has the requisite biological activity (as an IPP) or any active domain or fragment thereof, may be used in place of human CRT or a domain thereof.

Cheng *et al.*, *supra*, incorporated by reference in its entirety, previously determined that nucleic acid (*e.g.*, DNA) vaccines encoding each of the N, P, and C domains of CRT chimerically linked to HPV-16 E7 elicited potent antigen-specific CD8⁺ T cell responses and antitumor immunity in mice vaccinated *i.d.*, by gene gun administration. N-CRT/E7, P-CRT/E7 or C-CRT/E7 DNA each exhibited significantly increased numbers of E7-specific CD8⁺ T cell precursors and impressive antitumor effects against E7-expressing tumors when compared with mice vaccinated with E7 DNA (antigen only). N-CRT DNA administration also resulted in anti-angiogenic antitumor effects. Thus, cancer therapy using DNA encoding N-CRT linked to a tumor antigen may be used for treating tumors through a combination of antigen-specific immunotherapy and inhibition of angiogenesis.

The constructs comprising CRT or one of its domains linked to E7 is illustrated schematically below.



The amino acid sequences of the 3 human CRT domains are shown herein as annotations of the full length protein, SEQ ID NO:24. The N domain comprises residues 1-170 (normal text); the P domain comprises residues 171-269 (underscored); and the C domain comprises residues 270-417 (bold/italic).

The sequences of the three domains are further shown as separate polypeptides herein as human N-CRT (SEQ ID NO:25), as human P-CRT (SEQ ID NO:26), and as human C-CRT (SEQ ID NO:27).

5 The present vectors may comprises DNA encoding one or more of these domain sequences, which are shown by annotation of SEQ ID NO:28 herein, wherein the N-domain sequence is upper case, the P-domain sequence is lower case/*italic*/underscored, and the C domain sequence is lower case. The stop codon is also shown but not counted.

The coding sequence for each separate domain is provided herein as human N-CRT DNA (SEQ ID NO:29), as human P-CRT DNA (SEQ ID NO:30), and as human C-CRT
10 DNA (SEQ ID NO:31). Alternatively, any nucleotide sequences that encodes these domains may be used in the present constructs. Thus, for use in humans, the sequences may be further codon-optimized.

Constructs used in the present invention may employ combinations of one or more CRT domains, in any of a number of orientations. Using the designations N^{CRT} , P^{CRT} and
15 C^{CRT} to designate the domains, the following are but a few examples of the combinations that may be used in the nucleic acid (*e.g.*, DNA) vaccine vectors used in the present invention (where it is understood that *Ag* can be any antigen, including E7(detox) or E6 (detox).

$$\begin{array}{llll} N^{CRT} - P^{CRT} - Ag; & N^{CRT} - P^{CRT} - Ag; & N^{CRT} - C^{CRT} - Ag; & N^{CRT} - N^{CRT} - Ag; \\ N^{CRT} - N^{CRT} - N^{CRT} - Ag; & P^{CRT} - P^{CRT} - Ag; & P^{CRT} - C^{CRT} - Ag; & P^{CRT} - N^{CRT} - Ag; \\ C^{CRT} - P^{CRT} - Ag; & N^{CRT} - P^{CRT} - Ag; & & etc. \end{array}$$

20 The present invention may employ shorter polypeptide fragments of CRT or CRT domains provided such fragments can enhance the immune response to an antigen with which they are paired. Shorter peptides from the CRT or domain sequences shown above that have the ability to promote protein processing via the MHC-1 class I pathway are also included, and may be defined by routine experimentation.

25 The present invention may also employ shorter nucleic acid fragments that encode CRT or CRT domains provided such fragments are functional, *e.g.*, encode polypeptides that can enhance the immune response to an antigen with which they are paired (*e.g.*, linked). Nucleic acids that encode shorter peptides from the CRT or domain sequences shown above and are functional, *e.g.*, have the ability to promote protein processing via the MHC-1 class I pathway, are also included, and may be defined by routine experimentation.

A polypeptide fragment of CRT may include at least or about 50, 100, 200, 300, or 400 amino acids. A polypeptide fragment of CRT may also include at least or about 25, 50, 75, 100, 25-50, 50-100, or 75-125 amino acids from a CRT domain selected from the group N-CRT, P-CRT, and C-CRT. A polypeptide fragment of CRT may include residues 1-50, 50-75, 75-100, 100-125, 125-150, 150-170 of the N-domain (*e.g.*, of SEQ ID NO:25). A polypeptide fragment of CRT may include residues 1-50, 50-75, 75-100, 100-109 of the P-domain (*e.g.*, of SEQ ID NO:26). A polypeptide fragment of CRT may include residues 1-50, 50-75, 75-100, 100-125, 125-138 of the C-domain (*e.g.*, of SEQ ID NO:27).

A nucleic acid fragment of CRT may encode at least or about 50, 100, 200, 300, or 400 amino acids. A nucleic acid fragment of CRT may also encode at least or about 25, 50, 75, 100, 25-50, 50-100, or 75-125 amino acids from a CRT domain selected from the group N-CRT, P-CRT, and C-CRT. A nucleic acid fragment of CRT may encode residues 1-50, 50-75, 75-100, 100-125, 125-150, 150-170 of the N-domain (*e.g.*, of SEQ ID NO:25). A nucleic acid fragment of CRT may encode residues 1-50, 50-75, 75-100, 100-109 of the P-domain (*e.g.*, of SEQ ID NO:26). A nucleic acid fragment of CRT may encode residues 1-50, 50-75, 75-100, 100-125, 125-138 of the C-domain (*e.g.*, of SEQ ID NO:27).

Polypeptide “fragments” of CRT, as provided herein, do not include full-length CRT. Likewise, nucleic acid “fragments” of CRT, as provided herein, do not include a full-length CRT nucleic acid sequence and do not encode a full-length CRT polypeptide.

In one embodiment, a vector construct of a complete chimeric nucleic acid that can be used in the present invention, is shown herein as SEQ ID NO:32. The sequence is annotated to show plasmid-derived nucleotides (lower case letters), CRT-derived nucleotides (upper case bold letters), and HPV-E7-derived nucleotides (upper case, italicized/underlined letters). Five plasmid nucleotides are found between the CRT and E7 coding sequences and that the stop codon for the E7 sequence is double underscored. This plasmid is also referred to as pNGVL4a-CRT/E7(detox). The Table below describes the structure of the above plasmid.

Plasmid Position	Genetic Construct	Source of Construct
5970-0823	E. coli ORI (ColEI)	pBR / <i>E. coli</i> -derived
0837-0881	portion of transposase (tpnA)	Common plasmid sequence Tn5/Tn903
0882-1332	β -Lactamase (Amp ^R)	pBRpUC derived plasmid
1331-2496	AphA (Kan ^R)	Tn903
2509-2691	P3 Promoter DNA binding site	Tn3/pBR322
2692-2926	pUC backbone	Common plasmid sequence pBR322-derived
2931-4009	NF1 binding and promoter	HHV-5(HCMV UL-10 IE1 gene)
4010-4014	Poly-cloning site	Common plasmid sequence
4015 - 5265	Calreticulin (CRT)	Human Calreticulin
5266-5271	GAATTC plasmid sequence	Remain after cloning
5272-5568	dE7 gene (detoxified partial)	HPV-16 (E7 gene) incl. stop codon
5569-5580	Poly-cloning site	Common plasmid sequence
551-5970	Poly-Adenylation site	Mammalian signal, pHCMV-derived

In some embodiments, an alternative to CRT is another ER chaperone polypeptide exemplified by ER60, GRP94 or gp96, well-characterized ER chaperone polypeptide that representatives of the HSP90 family of stress-induced proteins (see WO 02/012281,

5 incorporated herein by reference). The term “endoplasmic reticulum chaperone polypeptide” as used herein means any polypeptide having substantially the same ER chaperone function as the exemplary chaperone proteins CRT, tapasin, ER60 or calnexin. Thus, the term includes all functional fragments or variants or mimics thereof. A polypeptide or peptide can be routinely screened for its activity as an ER chaperone using
10 assays known in the art. While the present invention is not limited by any particular mechanism of action, *in vivo* chaperones promote the correct folding and oligomerization of many glycoproteins in the ER, including the assembly of the MHC class I heterotrimeric molecule (heavy (H) chain, β 2m, and peptide). They also retain incompletely assembled MHC class I heterotrimeric complexes in the ER (Hauri FEBS Lett. 476:32-37, 2000).

15

Intercellular spreading proteins

The potency of naked nucleic acid (*e.g.*, DNA) vaccines may be enhanced by their ability to amplify and spread *in vivo*. VP22, a herpes simplex virus type 1 (HSV-1) protein and its "homologues" in other herpes viruses, such as the avian Marek's Disease Virus
20 (MDV) have the property of intercellular transport that provide an approach for enhancing

vaccine potency. The present inventors have previously created novel fusions of VP22 with a model antigen, human papillomavirus type 16 (HPV-16) E7, in a nucleic acid (*e.g.*, DNA) vaccine which generated enhanced spreading and MHC class I presentation of antigen.

These properties led to a dramatic increase in the number of E7-specific CD8⁺ T cell

5 precursors in vaccinated mice (at least 50-fold) and converted a less effective nucleic acid (*e.g.*, DNA) vaccine into one with significant potency against E7-expressing tumors. In comparison, a non-spreading mutant, VP22(1-267), failed to enhance vaccine potency. Results presented in U.S. Patent Application publication No. 20040028693 (US Patent No. 7,318,928), hereby incorporated by reference in its entirety, show that the potency of DNA
10 vaccines is dramatically improved through enhanced intercellular spreading and MHC class I presentation of the antigen.

A similar study linking MDV-1 UL49 to E7 also led to a dramatic increase in the number of E7-specific CD8⁺ T cell precursors and potency response against E7-expressing tumors in vaccinated mice. Mice vaccinated with a MDV-1 UL49 DNA vaccine stimulated
15 E7-specific CD8⁺ T cell precursor at a level comparable to that induced by HSV-1 VP22/E7. Thus, fusion of MDV-1UL49 DNA to DNA encoding a target antigen gene significantly enhances the DNA vaccine potency.

In one embodiment, the spreading protein may be a viral spreading protein, including a herpes virus VP22 protein. Exemplified herein are fusion constructs that
20 comprise herpes simplex virus-1 (HSV-1) VP22 (abbreviated HVP22) and its homologue from Marek's disease virus (MDV) termed MDV-VP22 or MVP-22. Also included in the invention are the use of homologues of VP22 from other members of the herpesviridae or polypeptides from nonviral sources that are considered to be homologous and share the functional characteristic of promoting intercellular spreading of a polypeptide or peptide
25 that is fused or chemically conjugated thereto.

DNA encoding HVP22 has the sequence SEQ ID NO:33 of the longer sequence SEQ ID NO:34 (which is the full length nucleotide sequence of a vector that comprises HVP22). DNA encoding MDV-VP22 is shown herein as SEQ ID NO:35.

The amino acid sequence of HVP22 polypeptide is SEQ ID NO:36 as amino acid
30 residues 1-301 of SEQ ID NO:37 (*i.e.*, the full length amino acid encoded by the vector).

The amino acid sequence of the MDV-VP22 is shown herein as SEQ ID NO:38.

A DNA clone pcDNA3 VP22/E7, that includes the coding sequence for HVP22 and the HPV-16 protein, E7 (plus some additional vector sequence) is SEQ ID NO:34.

The amino acid sequence of E7 (SEQ ID NO:39) is residues 308-403 of SEQ ID NO:37. This particular clone has only 96 of the 98 residues present in E7. The C-terminal residues of wild-type E7, Lys and Pro, are absent from this construct. This is an example of a deletion variant as the term is described below. Such deletion variants (*e.g.*, terminal truncation of two or a small number of amino acids) of other antigenic polypeptides are examples of the embodiments intended within the scope of the fusion polypeptides that can be used in the present invention.

Homologues of IPPs

Homologues or variants of IPPs described herein, may also be used, provided that they have the requisite biological activity. These include various substitutions, deletions, or additions of the amino acid or nucleic acid sequences. Due to code degeneracy, for example, there may be considerable variation in nucleotide sequences encoding the same amino acid sequence.

A functional derivative of an IPP retains measurable IPP-like activity, including that of promoting immunogenicity of one or more antigenic epitopes fused thereto by promoting presentation by class I pathways. "Functional derivatives" encompass "variants" and "fragments" regardless of whether the terms are used in the conjunctive or the alternative herein.

The term "chimeric" or "fusion" polypeptide or protein refers to a composition comprising at least one polypeptide or peptide sequence or domain that is chemically bound in a linear fashion with a second polypeptide or peptide domain. One embodiment of compositions useful for the present invention is an isolated or recombinant nucleic acid molecule encoding a fusion protein comprising at least two domains, wherein the first domain comprises an IPP and the second domain comprises an antigenic epitope, *e.g.*, an MHC class I-binding peptide epitope. The "fusion" can be an association generated by a peptide bond, a chemical linking, a charge interaction (*e.g.*, electrostatic attractions, such as salt bridges, H-bonding, *etc.*) or the like. If the polypeptides are recombinant, the "fusion protein" can be translated from a common mRNA. Alternatively, the compositions of the domains can be linked by any chemical or electrostatic means. The chimeric molecules that can be used in the present invention (*e.g.*, targeting polypeptide fusion proteins) can also include additional sequences, *e.g.*, linkers, epitope tags, enzyme cleavage recognition

sequences, signal sequences, secretion signals, and the like. Alternatively, a peptide can be linked to a carrier simply to facilitate manipulation or identification/ location of the peptide.

Also included is a “functional derivative” of an IPP, which refers to an amino acid substitution variant, a “fragment” of the protein. A functional derivative of an IPP retains measurable activity that may be manifested as promoting immunogenicity of one or more antigenic epitopes fused thereto or co-administered therewith. “Functional derivatives” encompass “variants” and “fragments” regardless of whether the terms are used in the conjunctive or the alternative herein.

A functional homologue must possess the above biochemical and biological activity. In view of this functional characterization, use of homologous proteins including proteins not yet discovered, fall within the scope of the invention if these proteins have sequence similarity and the recited biochemical and biological activity.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In one embodiment, the method of alignment includes alignment of Cys residues.

In one embodiment, the length of a sequence being compared is at least 30%, at least 40%, at least 50%, at least 60%, and at least 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the length of the reference sequence (*e.g.*, an IPP). The amino acid residues (or nucleotides) at corresponding amino acid (or nucleotide) positions are then compared. When a position in the first sequence is occupied by the same amino acid residue (or nucleotide) as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid “identity” is equivalent to amino acid or nucleic acid “homology”). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In one embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* 48:444-453 (1970)) algorithm which has been incorporated into the

GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases, for example, to identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to IPP nucleic acid molecules. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to IPP protein molecules. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

Thus, a homologue of an IPP or of an IPP domain described above is characterized as having (a) functional activity of native IPP or domain thereof and (b) amino acid sequence similarity to a native IPP protein or domain thereof when determined as above, of at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

It is within the skill in the art to obtain and express such a protein using DNA probes based on the disclosed sequences of an IPP. Then, the fusion protein's biochemical and biological activity can be tested readily using art-recognized methods such as those described herein, for example, a T cell proliferation, cytokine secretion or a cytolytic assay, or an *in vivo* assay of tumor protection or tumor therapy. A biological assay of the

stimulation of antigen-specific T cell reactivity will indicate whether the homologue has the requisite activity to qualify as a “functional” homologue.

A “variant” refers to a molecule substantially identical to either the full protein or to a fragment thereof in which one or more amino acid residues have been replaced
 5 (substitution variant) or which has one or several residues deleted (deletion variant) or added (addition variant). A “fragment” of an IPP refers to any subset of the molecule, that is, a shorter polypeptide of the full-length protein.

A number of processes can be used to generate fragments, mutants and variants of the isolated DNA sequence. Small subregions or fragments of the nucleic acid encoding the
 10 spreading protein, for example 1-30 bases in length, can be prepared by standard, chemical synthesis. Antisense oligonucleotides and primers for use in the generation of larger synthetic fragment.

A one group of variants are those in which at least one amino acid residue and in certain embodiments only one, has been substituted by different residue. For a detailed
 15 description of protein chemistry and structure, see Schulz, GE *et al.*, *Principles of Protein Structure*, Springer-Verlag, New York, 1978, and Creighton, T.E., *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, 1983, which are hereby incorporated by reference. The types of substitutions that may be made in the protein molecule may be based on analysis of the frequencies of amino acid changes between a
 20 homologous protein of different species, such as those presented in Table 1-2 of Schulz *et al.* (*supra*) and Figure 3-9 of Creighton (*supra*). Based on such an analysis, conservative substitutions are defined herein as exchanges within one of the following five groups:

- | | |
|---|---------------------------|
| 1. Small aliphatic, nonpolar or slightly polar residues | Ala, Ser, Thr (Pro, Gly); |
| 2. Polar, negatively charged residues and their amides | Asp, Asn, Glu, Gln; |
| 3. Polar, positively charged residues | His, Arg, Lys; |
| 4. Large aliphatic, nonpolar residues | Met, Leu, Ile, Val (Cys) |
| 5. Large aromatic residues | Phe, Tyr, Trp. |

The three amino acid residues in parentheses above have special roles in protein architecture. Gly is the only residue lacking a side chain and thus imparts flexibility to the
 25 chain. Pro, because of its unusual geometry, tightly constrains the chain. Cys can participate in disulfide bond formation, which is important in protein folding.

More substantial changes in biochemical, functional (or immunological) properties are made by selecting substitutions that are less conservative, such as between, rather than

within, the above five groups. Such changes will differ more significantly in their effect on maintaining (a) the structure of the peptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Examples of such substitutions are (i) substitution of Gly and/or Pro by another amino acid or deletion or insertion of Gly or Pro; (ii) substitution of a hydrophilic residue, *e.g.*, Ser or Thr, for (or by) a hydrophobic residue, *e.g.*, Leu, Ile, Phe, Val or Ala; (iii) substitution of a Cys residue for (or by) any other residue; (iv) substitution of a residue having an electropositive side chain, *e.g.*, Lys, Arg or His, for (or by) a residue having an electronegative charge, *e.g.*, Glu or Asp; or (v) substitution of a residue having a bulky side chain, *e.g.*, Phe, for (or by) a residue not having such a side chain, *e.g.*, Gly.

Most acceptable deletions, insertions and substitutions according to the present invention are those that do not produce radical changes in the characteristics of the wild-type or native protein in terms of its relevant biological activity, *e.g.*, its ability to stimulate antigen specific T cell reactivity to an antigenic epitope or epitopes that are fused to the protein. However, when it is difficult to predict the exact effect of the substitution, deletion or insertion in advance of doing so, one skilled in the art will appreciate that the effect can be evaluated by routine screening assays such as those described here, without requiring undue experimentation.

Exemplary fusion proteins provided herein comprise an IPP protein or homolog thereof and an antigen. For example, a fusion protein may comprise, consist essentially of, or consist of an IPP or an IPP fragment, *e.g.*, N-CRT, P-CRT and/or C-CRT, or an amino acid sequence that is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence of the IPP or IPP fragment, wherein the IPP fragment is functionally active as further described herein, linked to an antigen. A fusion protein may also comprise an IPP or an IPP fragment and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids, or about 1-5, 1-10, 1-15, 1-20, 1-25, 1-30, 1-50 amino acids, at the N- and/or C-terminus of the IPP fragment. These additional amino acids may have an amino acid sequence that is unrelated to the amino acid sequence at the corresponding position in the IPP protein.

Homologs of an IPP or an IPP fragments may also comprise, consist essentially of, or consist of an amino acid sequence that differs from that of an IPP or IPP fragment by the addition, deletion, or substitution, *e.g.*, conservative substitution, of at least about 1, 2, 3, 4,

5, 6, 7, 8, 9 or 10 amino acids, or from about 1-5, 1-10, 1-15 or 1-20 amino acids.

Homologs of an IPP or IPP fragments may be encoded by nucleotide sequences that are at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleotide sequence encoding an IPP or IPP fragment, such as those described herein.

Yet other homologs of an IPP or IPP fragments are encoded by nucleic acids that hybridize under stringent hybridization conditions to a nucleic acid that encodes an IPP or IPP fragment. For example, homologs may be encoded by nucleic acids that hybridize under high stringency conditions of 0.2 to 1 x SSC at 65 °C followed by a wash at 0.2 x SSC at 65 °C to a nucleic acid consisting of a sequence described herein. Nucleic acids that hybridize under low stringency conditions of 6 x SSC at room temperature followed by a wash at 2 x SSC at room temperature to nucleic acid consisting of a sequence described herein or a portion thereof can be used. Other hybridization conditions include 3 x SSC at 40 or 50 °C, followed by a wash in 1 or 2 x SSC at 20, 30, 40, 50, 60, or 65 °C.

Hybridizations can be conducted in the presence of formaldehyde, *e.g.*, 10%, 20%, 30% 40% or 50%, which further increases the stringency of hybridization. Theory and practice of nucleic acid hybridization is described, *e.g.*, in S. Agrawal (ed.) *Methods in Molecular Biology*, volume 20; and Tijssen (1993) *Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes*, *e.g.*, part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays," Elsevier, New York provide a basic guide to nucleic acid hybridization.

A fragment of a nucleic acid sequence is defined as a nucleotide sequence having fewer nucleotides than the nucleotide sequence encoding the full length CRT polypeptide, antigenic polypeptide, or the fusion thereof. This invention includes the use of such nucleic acid fragments that encode polypeptides which retain the ability of the fusion polypeptide to induce increases in frequency or reactivity of T cells, including CD8⁺ T cells, that are specific for the antigen part of the fusion polypeptide.

Nucleic acid sequences that can be used in the present invention may also include linker sequences, natural or modified restriction endonuclease sites and other sequences that are useful for manipulations related to cloning, expression or purification of encoded protein or fragments. For example, a fusion protein may comprise a linker between the antigen and the IPP protein.

Other nucleic acid vaccines that may be used include single chain trimers (SCT), as further described in the Examples and in references cited therein, all of which are specifically incorporated by reference herein.

5 *Backbone of nucleic acid vaccine*

A nucleic acid, *e.g.*, DNA vaccine may comprise an “expression vector” or “expression cassette,” *i.e.*, a nucleotide sequence which is capable of affecting expression of a protein coding sequence in a host compatible with such sequences. Expression cassettes include at least a promoter operably linked with the polypeptide coding sequence; and, optionally, with other sequences, *e.g.*, transcription termination signals. Additional factors necessary or helpful in effecting expression may also be included, *e.g.*, enhancers.

“Operably linked” means that the coding sequence is linked to a regulatory sequence in a manner that allows expression of the coding sequence. Known regulatory sequences are selected to direct expression of the desired protein in an appropriate host cell. Accordingly, the term “regulatory sequence” includes promoters, enhancers and other expression control elements. Such regulatory sequences are described in, for example, Goeddel, *Gene Expression Technology. Methods in Enzymology*, vol. 185, Academic Press, San Diego, Calif. (1990)).

A promoter region of a DNA or RNA molecule binds RNA polymerase and promotes the transcription of an “operably linked” nucleic acid sequence. As used herein, a “promoter sequence” is the nucleotide sequence of the promoter which is found on that strand of the DNA or RNA which is transcribed by the RNA polymerase. Two sequences of a nucleic acid molecule, such as a promoter and a coding sequence, are “operably linked” when they are linked to each other in a manner which permits both sequences to be transcribed onto the same RNA transcript or permits an RNA transcript begun in one sequence to be extended into the second sequence. Thus, two sequences, such as a promoter sequence and a coding sequence of DNA or RNA are operably linked if transcription commencing in the promoter sequence will produce an RNA transcript of the operably linked coding sequence. In order to be “operably linked” it is not necessary that two sequences be immediately adjacent to one another in the linear sequence.

In one embodiment, certain promoter sequences useful for the present invention must be operable in mammalian cells and may be either eukaryotic or viral promoters. Certain promoters are also described in the Examples, and other useful promoters and

regulatory elements are discussed below. Suitable promoters may be inducible, repressible or constitutive. A “constitutive” promoter is one which is active under most conditions encountered in the cell’s environmental and throughout development. An “inducible” promoter is one which is under environmental or developmental regulation. A “tissue specific” promoter is active in certain tissue types of an organism. An example of a constitutive promoter is the viral promoter MSV-LTR, which is efficient and active in a variety of cell types, and, in contrast to most other promoters, has the same enhancing activity in arrested and growing cells. Other viral promoters include that present in the CMV-LTR (from cytomegalovirus) (Bashart, M. *et al.*, *Cell* 41:521, 1985) or in the RSV-LTR (from Rous sarcoma virus) (Gorman, C.M., *Proc. Natl. Acad. Sci. USA* 79:6777, 1982). Also useful are the promoter of the mouse metallothionein I gene (Hamer, D, *et al.*, *J. Mol. Appl. Gen.* 1:273-88, 1982; the TK promoter of Herpes virus (McKnight, S, *Cell* 31:355-65, 1982); the SV40 early promoter (Benoist, C., *et al.*, *Nature* 290:304-10, 1981); and the yeast *gal4* gene promoter (Johnston, SA *et al.*, *Proc. Natl. Acad. Sci. USA* 79:6971-5, 1982); Silver, PA, *et al.*, *Proc. Natl. Acad. Sci. (USA)* 81:5951-5, 1984)). Other illustrative descriptions of transcriptional factor association with promoter regions and the separate activation and DNA binding of transcription factors include: Keegan *et al.*, *Nature* 231:699, 1986; Fields *et al.*, *Nature* 340:245, 1989; Jones, *Cell* 61:9, 1990; Lewin, *Cell* 61:1161, 1990; Ptashne *et al.*, *Nature* 346:329, 1990; Adams *et al.*, *Cell* 72:306, 1993.

The promoter region may further include an octamer region which may also function as a tissue specific enhancer, by interacting with certain proteins found in the specific tissue. The enhancer domain of the DNA construct useful for the present invention is one which is specific for the target cells to be transfected, or is highly activated by cellular factors of such target cells. Examples of vectors (plasmid or retrovirus) are disclosed, *e.g.*, in Roy-Burman *et al.*, U.S. Patent No. 5,112,767, incorporated by reference. For a general discussion of enhancers and their actions in transcription, see, Lewin, BM, *Genes IV*, Oxford University Press pp. 552-576, 1990 (or later edition). Particularly useful are retroviral enhancers (*e.g.*, viral LTR) that is placed upstream from the promoter with which it interacts to stimulate gene expression. For use with retroviral vectors, the endogenous viral LTR may be rendered enhancer-less and substituted with other desired enhancer sequences which confer tissue specificity or other desirable properties such as transcriptional efficiency.

Thus, expression cassettes include plasmids, recombinant viruses, any form of a recombinant “naked DNA” vector, and the like. A “vector” comprises a nucleic acid which can infect, transfect, transiently or permanently transduce a cell. It will be recognized that a vector can be a naked nucleic acid, or a nucleic acid complexed with protein or lipid. The vector optionally comprises viral or bacterial nucleic acids and/or proteins, and/or membranes (*e.g.*, a cell membrane, a viral lipid envelope, *etc.*). Vectors include replicons (*e.g.*, RNA replicons), bacteriophages) to which fragments of DNA may be attached and become replicated. Vectors thus include, but are not limited to RNA, autonomous self-replicating circular or linear DNA or RNA, *e.g.*, plasmids, viruses, and the like (U.S. Patent No. 5,217,879, incorporated by reference), and includes both the expression and nonexpression plasmids. Where a recombinant cell or culture is described as hosting an “expression vector” this includes both extrachromosomal circular and linear DNA and DNA that has been incorporated into the host chromosome(s). Where a vector is being maintained by a host cell, the vector may either be stably replicated by the cells during mitosis as an autonomous structure, or is incorporated within the host’s genome.

Exemplary virus vectors that may be used include recombinant adenoviruses (Horowitz, MS, In: *Virology*, Fields, BN *et al.*, eds, Raven Press, NY, 1990, p. 1679; Berkner, KL, *Biotechniques* 6:616-29, 1988; Strauss, SE, In: *The Adenoviruses*, Ginsberg, HS, ed., Plenum Press, NY, 1984, chapter 11) and herpes simplex virus (HSV). Advantages of adenovirus vectors for human gene delivery include the fact that recombination is rare, no human malignancies are known to be associated with such viruses, the adenovirus genome is double stranded DNA which can be manipulated to accept foreign genes of up to 7.5 kb in size, and live adenovirus is a safe human vaccine organisms. Adeno-associated virus is also useful for human therapy (Samulski, RJ *et al.*, *EMBO J.* 10:3941, 1991) according to the present invention.

A nucleic acid (*e.g.*, DNA) vaccine may also use a replicon, *e.g.*, an RNA replicon, a self-replicating RNA vector. In one embodiment, a replicon is one based on a Sindbis virus RNA replicon, *e.g.*, SINrep5. The present inventors tested E7 in the context of such a vaccine and showed (see Wu *et al.*, U.S. Patent Application 10/343,719) that a Sindbis virus RNA vaccine encoding HSV-1 VP22 linked to E7 significantly increased activation of E7-specific CD8 T cells, resulting in potent antitumor immunity against E7-expressing tumors. The Sindbis virus RNA replicon vector used in these studies, SINrep5, has been described (Bredenbeek, P J *et al.*, 1993, *J. Virol.* 67:6439-6446).

Generally, RNA replicon vaccines may be derived from alphavirus vectors, such as Sindbis virus (Hariharan, M J *et al.*, 1998. J Virol 72:950-8.), Semliki Forest virus (Berglund, P M *et al.*, 1997. AIDS Res Hum Retroviruses 13:1487-95; Ying, H T *et al.*, 1999. Nat Med 5:823-7) or Venezuelan equine encephalitis virus (Pushko, P M *et al.*, 1997. Virology 239:389-401). These self-replicating and self-limiting vaccines may be administered as either (1) RNA or (2) DNA which is then transcribed into RNA replicons in cells transfected in vitro or in vivo (Berglund, P C *et al.*, 1998. Nat Biotechnol 16:562-5; Leitner, W W *et al.*, 2000. Cancer Res 60:51-5). An exemplary Semliki Forest virus is pSCA1 (DiCiommo, D P *et al.*, J Biol Chem 1998; 273:18060-6).

The plasmid vector pcDNA3 or a functional homolog thereof (SEQ ID NO:40) may be used in a nucleic acid (*e.g.*, DNA) vaccine. In other embodiments, pNGVL4a (SEQ ID NO:41) can be used.

pNGVL4a, one plasmid backbone for use in the present invention, was originally derived from the pNGVL3 vector, which has been approved for human vaccine trials. The pNGVL4a vector includes two immunostimulatory sequences (tandem repeats of CpG dinucleotides) in the noncoding region. Whereas any other plasmid DNA that can transform either APCs, including DC's or other cells which, via cross-priming, transfer the antigenic moiety to DCs, is useful in the present invention, pNGFVLA4a may be used because of the fact that it has already been approved for human therapeutic use.

The following references set forth principles and current information in the field of basic, medical and veterinary virology and are incorporated by reference: *Fields Virology*, Fields, BN *et al.*, eds., Lippincott Williams & Wilkins, NY, 1996; *Principles of Virology: Molecular Biology, Pathogenesis, and Control*, Flint, S.J. *et al.*, eds., Amer Soc Microbiol, Washington DC, 1999; *Principles and Practice of Clinical Virology*, 4th Edition, Zuckerman. A.J. *et al.*, eds, John Wiley & Sons, NY, 1999; *The Hepatitis C Viruses*, by Hagedorn, CH *et al.*, eds., Springer Verlag, 1999; *Hepatitis B Virus: Molecular Mechanisms in Disease and Novel Strategies for Therapy*, Koshy, R. *et al.*, eds, World Scientific Pub Co, 1998; *Veterinary Virology*, Murphy, F.A. *et al.*, eds., Academic Press, NY, 1999; *Avian Viruses: Function and Control*, Ritchie, B.W., Iowa State University Press, Ames, 2000; *Virus Taxonomy: Classification and Nomenclature of Viruses: Seventh Report of the International Committee on Taxonomy of Viruses*, by M. H. V. Van Regenmortel, MHV *et al.*, eds., Academic Press; NY, 2000.

Plasmid DNA used for transfection or microinjection may be prepared using methods well-known in the art, for example using the Qiagen procedure (Qiagen), followed by DNA purification using known methods, such as the methods exemplified herein.

Such expression vectors may be used to transfect host cells (*in vitro*, *ex vivo* or *in vivo*) for expression of the DNA and production of the encoded proteins which include fusion proteins or peptides. In one embodiment, a nucleic acid (*e.g.*, DNA) vaccine is administered to or contacted with a cell, *e.g.*, a cell obtained from a subject (*e.g.*, an antigen presenting cell), and administered to a subject, wherein the subject is treated before, after or at the same time as the cells are administered to the subject.

The term “isolated” as used herein, when referring to a molecule or composition, such as a translocation polypeptide or a nucleic acid coding therefor, means that the molecule or composition is separated from at least one other compound (protein, other nucleic acid, *etc.*) or from other contaminants with which it is natively associated or becomes associated during processing. An isolated composition can also be substantially pure. An isolated composition can be in a homogeneous state and can be dry or in aqueous solution. Purity and homogeneity can be determined, for example, using analytical chemical techniques such as polyacrylamide gel electrophoresis (PAGE) or high performance liquid chromatography (HPLC). Even where a protein has been isolated so as to appear as a homogenous or dominant band in a gel pattern, there are trace contaminants which co-purify with it.

Host cells transformed or transfected to express the fusion polypeptide or a homologue or functional derivative thereof are useful for the present invention. For example, the fusion polypeptide may be expressed in yeast, or mammalian cells such as Chinese hamster ovary cells (CHO) or human cells. In one embodiment, cells for expression according to the present invention are APCs or DCs. Other suitable host cells are known to those skilled in the art.

Other nucleic acids for potentiating immune responses

Methods of administering a chemotherapeutic drug and a vaccine may further comprise administration of one or more other constructs, *e.g.*, to prolong the life of antigen presenting cells. Exemplary constructs are described in the following two sections. Such constructs may be administered simultaneously or at the same time as a nucleic acid (*e.g.*,

DNA) vaccine. Alternatively, they may be administered before or after administration of the DNA vaccine or chemotherapeutic drug.

Potentiation of immune responses using siRNA directed at apoptotic pathways

5 Administration to a subject of a DNA vaccine and a chemotherapeutic drug may be accompanied by administration of one or more other agents, *e.g.*, constructs. In one embodiment, a method comprises further administering to a subject an siRNA directed at an apoptotic pathway, such as described in WO 2006/073970, which is incorporated herein in its entirety.

10 The present inventors have designed siRNA sequences that hybridize to, and block expression of the activation of Bak and Bax proteins that are central players in the apoptosis signaling pathway. Methods of treating tumors or hyperproliferative diseases involving the administration of siRNA molecules (sequences), vectors containing or encoding the siRNA, expression vectors with a promoter operably linked to the siRNA
15 coding sequence that drives transcription of siRNA sequences that are “specific” for sequences Bak and Bax nucleic acid are also encompassed within the present invention. siRNAs may include single stranded “hairpin” sequences because of their stability and binding to the target mRNA.

Since Bak and Bax are involved, among other death proteins, in apoptosis of APCs,
20 particularly DCs, the present siRNA sequences may be used in conjunction with a broad range of DNA vaccine constructs encoding antigens to enhance and promote the immune response induced by such DNA vaccine constructs, particularly CD8⁺ T cell mediated immune responses typified by CTL activation and action. This is believed to occur as a result of the effect of the siRNA in prolonging the life of antigen-presenting DCs which
25 may otherwise be killed in the course of a developing immune response by the very same CTLs that the DCs are responsible for inducing.

In addition to Bak and Bax, additional targets for siRNAs designed in an analogous manner include caspase 8, caspase 9 and caspase 3. The present invention includes compositions and methods in which siRNAs targeting any two or more of Bak, Bax,
30 caspase 8, caspase 9 and caspase 3 are used in combination, optionally simultaneously (along with a DNA immunogen that encodes an antigen), to administer to a subject. Such combinations of siRNAs may also be used to transfect DCs (along with antigen loading) to

improve the immunogenicity of the DCs as cellular vaccines by rendering them resistant to apoptosis.

siRNAs suppress gene expression through a highly regulated enzyme-mediated process called RNA interference (RNAi) (Sharp, P.A., *Genes Dev.* 15:485–90, 2001; Bernstein, E *et al.*, *Nature* 409:363–66, 2001; Nykanen, A *et al.*, *Cell* 107:309–21, 2001; Elbashir *et al.*, *Genes Dev.* 15:188–200, 2001). RNA interference is the sequence-specific degradation of homologues in an mRNA of a targeting sequence in an siNA. As used herein, the term siNA (small, or short, interfering nucleic acid) is meant to be equivalent to other terms used to describe nucleic acid molecules that are capable of mediating sequence specific RNAi (RNA interference), for example short (or small) interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), short hairpin RNA (shRNA), short interfering oligonucleotide, short interfering nucleic acid, short interfering modified oligonucleotide, chemically-modified siRNA, post-transcriptional gene silencing RNA (ptgsRNA), translational silencing, and others. RNAi involves multiple RNA-protein interactions characterized by four major steps: assembly of siRNA with the RNA-induced silencing complex (RISC), activation of the RISC, target recognition and target cleavage. These interactions may bias strand selection during siRNA-RISC assembly and activation, and contribute to the overall efficiency of RNAi (Khvorova, A *et al.*, *Cell* 115:209–216 (2003); Schwarz, DS *et al.* 115:199–208 (2003))

Considerations to be taken into account when designing an RNAi molecule include, among others, the sequence to be targeted, secondary structure of the RNA target and binding of RNA binding proteins. Methods of optimizing siRNA sequences will be evident to the skilled worker. Typical algorithms and methods are described in Vickers *et al.* (2003) *J Biol Chem* 278:7108-7118; Yang *et al.* (2003) *Proc Natl Acad Sci USA* 99:9942-9947; Far *et al.* (2003) *Nuc. Acids Res.* 31:4417-4424; and Reynolds *et al.* (2004) *Nature Biotechnology* 22:326-330, all of which are incorporated by reference in their entirety.

The methods described in Far *et al.*, *supra*, and Reynolds *et al.*, *supra*, may be used by those of ordinary skill in the art to select targeted sequences and design siRNA sequences that are effective at silencing the transcription of the relevant mRNA. Far *et al.* suggests options for assessing target accessibility for siRNA and supports the design of active siRNA constructs. This approach can be automated, adapted to high throughput and is open to include additional parameters relevant to the biological activity of siRNA. To identify siRNA-specific features likely to contribute to efficient processing at each of the

steps of RNAi noted above. Reynolds *et al.*, *supra*, present a systematic analysis of 180 siRNAs targeting the mRNA of two genes. Eight characteristics associated with siRNA functionality were identified: low G/C content, a bias towards low internal stability at the sense strand 3'-terminus, lack of inverted repeats, and sense strand base preferences

5 (positions 3, 10, 13 and 19). Application of an algorithm incorporating all eight criteria significantly improves potent siRNA selection. This highlights the utility of rational design for selecting potent siRNAs that facilitate functional gene knockdown.

Candidate siRNA sequences against mouse and human Bax and Bak are selected using a process that involves running a BLAST search against the sequence of Bax or Bak
10 (or any other target) and selecting sequences that "survive" to ensure that these sequences will not be cross matched with any other genes.

siRNA sequences selected according to such a process and algorithm may be cloned into an expression plasmid and tested for their activity in abrogating Bak/Bax function cells of the appropriate animal species. Those sequences that show RNAi activity may be used
15 by direct administration bound to particles, or recloned into a viral vector such as a replication-defective human adenovirus serotype 5 (Ad5).

One advantage of this viral vector is the high titer obtainable (in the range of 10^{10}) and therefore the high multiplicities-of-infection that can be attained. For example, infection with 100 infectious units/ cell ensures all cells are infected. Another advantage of
20 this virus is the high susceptibility and infectivity and the host range (with respect to cell types). Even if expression is transient, cells would survive, possibly replicate, and continue to function before Bak/Bax activity would recover and lead to cell death. In one embodiment, constructs include the following:

For **Bak**:

25 5'-P-UGCCUACGAACUCUUCACCCdTdT-3' (sense) (SEQ ID NO:42)

5'-P-GGUGAAGAGUUCGUAGGCAdTdT-3' (antisense) (SEQ ID NO:43),

The nucleotide sequence encoding the Bak protein (including the stop codon)

(GenBank accession No. NM_007523 is shown herein as SEQ ID NO:44 with the targeted sequence in upper case, underscored. The targeted sequence of Bak,

30 TGCCTACGAACCTCTTACC is shown herein as SEQ ID NO:45.

For **Bax**:

5'-P-UAUGGAGCUGCAGAGGAUGdTdT-3' (sense) (SEQ ID NO:46)

5'-P-CAUCCUCUGCAGCUCCAUAAdTdT-3' (antisense) (SEQ ID NO:47)

The nucleotide sequence encoding Bax (including the stop codon) (GenBank accession No. L22472 is shown below (SEQ ID NO:48) with the targeted sequence shown in upper case and underscored

The targeted sequence of Bax, TATGGAGCTGCAGAGGATG is shown herein as
5 SEQ ID NO:49

In a one embodiment, the inhibitory molecule is a double stranded nucleic acid (i.e., an RNA), used in a method of RNA interference. The following show the “paired” 19 nucleotide structures of the siRNA sequences shown above, where the symbol $\uparrow$:

Bak: 5'P- UGCCUACGAACUCUUCACCDtT-3' (sense) (SEQ ID NO: 42)

10 $\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$

3'P-dTdTACGGAUGCUUGAGAAGUGG -5' (antisense) (SEQ ID NO: 43)

Bax: 5'P- UAUGGAGCUGCAGAGGAUGdTdT-3' (sense) (SEQ ID NO: 46)

$\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$

3'P-dTdTAAUACCUCGACGUCUCCUAC -5' (antisense) (SEQ ID NO: 47)

15

Other Pro-Apoptotic Proteins to be Targeted

1. Caspase 8: The nucleotide sequence of human caspase-8 is shown herein as SEQ ID NO:50 (GenBank Access. # NM_001228). One target sequence for RNAi is underscored. Others may be identified using methods such as those described herein (and
20 in reference cited herein, primarily Far *et al.*, *supra* and Reynolds *et al.*, *supra*).

The sequences of sense and antisense siRNA strands for targeting this sequence including dTdT 3' overhangs, are:

5'-AACCUCGGGGAUACUGUCUGAdTdT-3' (sense) (SEQ ID NO:51)

5'-UCAGACAGUAUCCCCGAGGUUdTdT-3' (antisense) (SEQ ID NO:52)

25 2. Caspase 9: The nucleotide sequence of human caspase-9 is shown herein as SEQ ID NO:53 (see GenBank Access. # NM_001229). The sequence below is of “variant α ” which is longer than a second alternatively spliced variant β , which lacks the underscored part of the sequence shown below (and which is anti-apoptotic). Target sequences for RNAi, expected to fall in the underscored segment, are identified using
30 known methods such as those described herein and in Far *et al.*, *supra* and Reynolds *et al.*, *supra*) and siNAs, such as siRNAs, are designed accordingly.

3. Caspase 3: The nucleotide sequence of human caspase-3 is shown herein as SEQ ID NO: 54 (see GenBank Access. # NM_004346). The sequence below is of “variant

α“ which is the longer of two alternatively spliced variants, all of which encode the full protein. Target sequences for RNAi are identified using known methods such as those described herein and in Far *et al.*, *supra* and Reynolds *et al.*, *supra*) and siNAs, such as siRNAs, are designed accordingly.

5 Long double stranded interfering RNAs, such as miRNAs, appear to tolerate mismatches more readily than do short double stranded RNAs. In addition, as used herein, the term RNAi is meant to be equivalent to other terms used to describe sequence specific RNA interference, such as post transcriptional gene silencing, or an epigenetic phenomenon. For example, siNA molecules useful for the invention can be used to
10 epigenetically silence genes at both the post-transcriptional level or the pre-transcriptional level. In a non-limiting example, epigenetic regulation of gene expression by siNA molecules useful for the present invention can result from siNA mediated modification of chromatin structure and thereby alter gene expression (see, for example, Allshire *Science* 297:1818-19, 2002; Volpe *et al.*, *Science* 297:1833-37, 2002; Jenuwein, *Science* 297:2215-
15 18, 2002; and Hall *et al.*, *Science* 297, 2232-2237, 2002.)

An siNA can be designed to target any region of the coding or non-coding sequence of an mRNA. An siNA is a double-stranded polynucleotide molecule comprising self-complementary sense and antisense regions, wherein the antisense region comprises nucleotide sequence that is complementary to nucleotide sequence in a target nucleic acid
20 molecule or a portion thereof and the sense region has a nucleotide sequence corresponding to the target nucleic acid sequence or a portion thereof. The siNA can be assembled from two separate oligonucleotides, where one strand is the sense strand and the other is the antisense strand, wherein the antisense and sense strands are self-complementary. The siNA can be assembled from a single oligonucleotide, where the self-complementary sense
25 and antisense regions of the siNA are linked by means of a nucleic acid based or non-nucleic acid-based linker(s). The siNA can be a polynucleotide with a hairpin secondary structure, having self-complementary sense and antisense regions. The siNA can be a circular single-stranded polynucleotide having two or more loop structures and a stem comprising self-complementary sense and antisense regions, wherein the circular
30 polynucleotide can be processed either *in vivo* or *in vitro* to generate an active siNA molecule capable of mediating RNAi. The siNA can also comprise a single stranded polynucleotide having nucleotide sequence complementary to nucleotide sequence in a target nucleic acid molecule or a portion thereof (or can be an siNA molecule that does not

require the presence within the siNA molecule of nucleotide sequence corresponding to the target nucleic acid sequence or a portion thereof), wherein the single stranded polynucleotide can further comprise a terminal phosphate group, such as a 5'-phosphate (see for example Martinez *et al.* (2002) *Cell* 110, 563-574 and Schwarz *et al.* (2002)

5 *Molecular Cell* 10, 537-568), or 5',3'-diphosphate.

In certain embodiments, the siNA molecule useful for the present invention comprises separate sense and antisense sequences or regions, wherein the sense and antisense regions are covalently linked by nucleotide or non-nucleotide linkers molecules as is known in the art, or are alternately non-covalently linked by ionic interactions, hydrogen bonding, Van der Waal's interactions, hydrophobic interactions, and/or stacking interactions.

As used herein, siNA molecules need not be limited to those molecules containing only ribonucleotides but may also further encompass deoxyribonucleotides (as in the siRNAs which each include a dTdT dinucleotide) chemically-modified nucleotides, and non-nucleotides. In certain embodiments, the siNA molecules useful for the present invention lack 2'-hydroxy (2'-OH) containing nucleotides. In certain embodiments, siNAs do not require the presence of nucleotides having a 2'-hydroxy group for mediating RNAi and as such, siNAs useful for the present invention optionally do not include any ribonucleotides (*e.g.*, nucleotides having a 2'-OH group). Such siNA molecules that do not require the presence of ribonucleotides within the siNA molecule to support RNAi can however have an attached linker or linkers or other attached or associated groups, moieties, or chains containing one or more nucleotides with 2'-OH groups. Optionally, siNA molecules can comprise ribonucleotides at about 5, 10, 20, 30, 40, or 50% of the nucleotide positions. If modified, the siNAs useful for the present invention can also be referred to as "short interfering modified oligonucleotides" or "siMON." Other chemical modifications, *e.g.*, as described in Int'l Patent Publications WO 03/070918 and WO 03/074654, both of which are incorporated by reference, can be applied to any siNA sequence useful for the present invention.

In one embodiment a molecule mediating RNAi has a 2 nucleotide 3' overhang (dTdT in the sequences disclosed herein). If the RNAi molecule is expressed in a cell from a construct, for example from a hairpin molecule or from an inverted repeat of the desired sequence, then the endogenous cellular machinery will create the overhangs.

Methods of making siRNAs are conventional. *In vitro* methods include processing the polyribonucleotide sequence in a cell-free system (e.g., digesting long dsRNAs with RNase III or Dicer), transcribing recombinant double stranded DNA *in vitro*, and chemical synthesis of nucleotide sequences homologous to Bak or Bax sequences. See, e.g., Tuschl

5 *et al.*, *Genes & Dev.* 13:3191-3197, 1999. *In vivo* methods include

- (1) transfecting DNA vectors into a cell such that a substrate is converted into siRNA *in vivo*. See, for example, Kawasaki *et al.*, *Nucleic Acids Res* 31:700-07, 2003;; Miyagishi *et al.*, *Nature Biotechnol* 20:497-500, 2003;; Lee *et al.*, *Nature Biotechnol* 20:500-05, 2002; Brummelkamp *et al.*, *Science* 296:550-53, 2002; McManus *et al.*,
 10 *RNA* 8:842-50, 2002; Paddison *et al.*, *Genes Dev* 16:948-58, 2002; Paddison *et al.*, *Proc Natl Acad Sci USA* 99:1443-48, 2002; Paul *et al.*, *Nature Biotechnol* 20:505-08, 2002; Sui *et al.*, *Proc Natl Acad Sci USA* 99:5515-20, 2002; Yu *et al.*, *Proc Natl Acad Sci USA* 99:6047-52, 2002)
- (2) expressing short hairpin RNAs from plasmid systems using RNA polymerase III (pol
 15 III) promoters. See, for example, Kawasaki *et al.*, *supra*; Miyagishi *et al.*, *supra*; Lee *et al.*, *supra*; Brummelkamp *et al.*, *supra*; McManus *et al.*, *supra*), Paddison *et al.*, *supra* (both); Paul *et al.*, *supra*, Sui *et al.*, *supra*; and Yu *et al.*, *supra*; and/or
- (3) expressing short RNA from tandem promoters. See, for example, Miyagishi *et al.*, *supra*; Lee *et al.*, *supra*).

20 When synthesized *in vitro*, a typical micromolar scale RNA synthesis provides about 1 mg of siRNA, which is sufficient for about 1000 transfection experiments using a 24-well tissue culture plate format. In general, to inhibit Bak or Bax expression in cells in culture, one or more siRNAs can be added to cells in culture media, typically at about 1 ng/ml to about 10 µg siRNA/ml.

25 For reviews and more general description of inhibitory RNAs, see Lau *et al.*, *Sci Amer* Aug 2003: 34-41; McManus *et al.*, *Nature Rev Genetics* 3, 737-47, 2002; and Dykxhoorn *et al.*, *Nature Rev Mol Cell Bio* 4:457-467, 2003. For further guidance regarding methods of designing and preparing siRNAs, testing them for efficacy, and using them in methods of RNA interference (both *in vitro* and *in vivo*), see, e.g., Allshire, *Science*
 30 297:1818-19, 2002; Volpe *et al.*, *Science* 297:1833-37, 2002; Jenuwein, *Science* 297:2215-18, 2002; Hall *et al.*, *Science* 297 2232-37, 2002; Hutvagner *et al.*, *Science* 297:2056-60, 2002; McManus *et al.* *RNA* 8:842-850, 2002; Reinhart *et al.*, *Genes Dev.* 16:1616-26, 2002; Reinhart *et al.*, *Science* 297:1831, 2002; Fire *et al.* (1998) *Nature* 391:806-11, 2002; Moss,

Curr Biol 11:R772-5, 2002; Brummelkamp *et al.*, *supra*; Bass, *Nature* 411 428-9, 2001; Elbashir *et al.*, *Nature* 411:494-8; US Pat. 6,506,559; Published US Pat App. 20030206887; and PCT applications WO99/07409, WO99/32619, WO 00/01846, WO 00/44914, WO00/44895, WO01/29058, WO01/36646, WO01/75164, WO01/92513, WO 01/29058, 5 WO01/89304, WO01/90401, WO02/16620, and WO02/29858, all of which are incorporated by reference.

Ribozymes and siNAs can take any of the forms, including modified versions, described for antisense nucleic acid molecules; and they can be introduced into cells as oligonucleotides (single or double stranded), or in the form of an expression vector.

10 In one embodiment, an antisense nucleic acid, siNA (*e.g.*, siRNA) or ribozyme comprises a single stranded polynucleotide comprising a sequence that is at least about 90% (*e.g.*, at least about 93%, 95%, 97%, 98% or 99%) identical to a target segment (such as those indicted for Bak and Bax above) or a complement thereof. As used herein, a DNA and an RNA encoded by it are said to contain the same "sequence," taking into account that 15 the thymine bases in DNA are replaced by uracil bases in RNA.

Active variants (*e.g.*, length variants, including fragments; and sequence variants) of the nucleic acid-based inhibitors discussed herein are also within the scope of the present invention. An "active" variant is one that retains an activity of the inhibitor from which it is derived (*i.e.*, the ability to inhibit expression). It is to test a variant to determine for its 20 activity using conventional procedures.

As for length variants, an antisense nucleic acid or siRNA may be of any length that is effective for inhibition of a gene of interest. Typically, an antisense nucleic acid is between about 6 and about 50 nucleotides (*e.g.*, at least about 12, 15, 20, 25, 30, 35, 40, 45 or 50 nt), and may be as long as about 100 to about 200 nucleotides or more. Antisense 25 nucleic acids having about the same length as the gene or coding sequence to be inhibited may be used. When referring to length, the terms bases and base pairs (bp) are used interchangeably, and will be understood to correspond to single stranded (ss) and double stranded (ds) nucleic acids. The length of an effective siNA is generally between about 15 bp and about 29 bp in length, between about 19 and about 29 bp (*e.g.*, about 15, 17, 19, 21, 30 23, 25, 27 or 29 bp), with shorter and longer sequences being acceptable. Generally, siNAs are shorter than about 30 bases to prevent eliciting interferon effects. For example, an active variant of an siRNA having, for one of its strands, the 19 nucleotide sequence of any of SEQ ID NOs:42, 43, 46, and 47 herein can lack base pairs from either, or both, of ends

of the dsRNA; or can comprise additional base pairs at either, or both, ends of the ds RNA, provided that the total of length of the siRNA is between about 19 and about 29 bp, inclusive. One embodiment useful for the present invention is an siRNA that “consists essentially of” sequences represented by SEQ ID NOs:42, 43, 46, and 47 or complements
5 of these sequence. An siRNA useful for the present invention may consist essentially of between about 19 and about 29 bp in length.

As for sequence variants, in one embodiment, an inhibitory nucleic acid, whether an antisense molecule, a ribozyme (the recognition sequences), or an siNA, comprises a strand that is complementary (100% identical in sequence) to a sequence of a gene that it is
10 designed to inhibit. However, 100% sequence identity is not required to practice the present invention. Thus, the invention has the advantage of being able to tolerate naturally occurring sequence variations, for example, in human c-met, that might be expected due to genetic mutation, polymorphism, or evolutionary divergence. Alternatively, the variant sequences may be artificially generated. Nucleic acid sequences with small insertions,
15 deletions, or single point mutations relative to the target sequence can be effective inhibitors.

The degree of sequence identity may be optimized by sequence comparison and alignment algorithms well-known in the art (see Gribskov and Devereux, Sequence Analysis Primer, Stockton Press, 1991, and references cited therein) and calculating the
20 percent difference between the nucleotide sequences by, for example, the Smith-Waterman algorithm as implemented in the BESTFIT software program using default parameters (*e.g.*, University of Wisconsin Genetic Computing Group). In one embodiment, at least about 90% sequence identity may be used (*e.g.*, at least about 92%, 95%, 98% or 99%), or even 100% sequence identity, between the inhibitory nucleic acid and the targeted sequence of
25 targeted gene.

Alternatively, an active variant of an inhibitory nucleic acid useful for the present invention is one that hybridizes to the sequence it is intended to inhibit under conditions of high stringency. For example, the duplex region of an siRNA may be defined functionally as a nucleotide sequence that is capable of hybridizing with a portion of the target gene
30 transcript under high stringency conditions (*e.g.*, 400 mM NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA, 50°C or 70°C, hybridization for 12-16 hours), followed generally by washing.

DC-1 cells or BM-DCs presenting a given antigen X, when not treated with the siRNAs useful for the present invention, respond to sufficient numbers X-specific CD8+

CTL by apoptotic cell death. In contrast, the same cells transfected with the siRNA or infected with a viral vector encoding the present siRNA sequences survive better despite the delivery of killing signals.

Delivery and expression of the siRNA compositions useful for the present invention inhibit the death of DCs *in vivo* in the process of a developing T cell response, and thereby promote and stimulate the generation of an immune response induced by immunization with an antigen-encoding DNA vaccine vector. These capabilities have been exemplified by showing that:

- (1) co-administration of DNA vaccines encoding HPV-16 E7 with siRNA targeted to Bak and Bax prolongs the lives of antigen-presenting DCs in the draining lymph nodes, thereby enhancing antigen-specific CD8⁺ T cell responses, and eliciting potent antitumor effects against an E7-expressing tumor in vaccinated subjects.
- (2) DCs transfected with siRNA targeting Bak and Bax resist killing by T cells *in vivo*. E7-loaded DCs transfected with Bak/Bax siRNA so that Bak and Bax protein expression is downregulated resist apoptotic death induced by T cells *in vivo*. When administered to subjects, these DCs generate stronger antigen-specific immune responses and manifest therapeutic effects (compared to DCs transfected with control siRNA).

Thus, siRNA constructs are useful as a part of the nucleic acid vaccination and chemotherapy regimen described in this application.

Potentiation of immune responses using anti-apoptotic proteins

Administration to a subject of a DNA vaccine and a chemotherapeutic drug may also be accompanied by administration of a nucleic acid encoding an anti-apoptotic protein, as described in WO2005/047501 and in U.S. Patent Application Publication No. 20070026076, both of which are incorporated by reference.

The present inventors have designed and disclosed an immunotherapeutic strategy that combines antigen-encoding DNA vaccine compositions with additional DNA vectors comprising anti-apoptotic genes including bcl-2, bc-lxL, XIAP, dominant negative mutants of caspase-8 and caspase-9, the products of which are known to inhibit apoptosis (Wu, *et al.* U.S. Patent Application Publication No. 20070026076, incorporated herein by reference). Serine protease inhibitor 6 (SPI-6) which inhibits granzyme B, may also be employed in compositions and methods to delay apoptotic cell death of DCs. The present inventors have

shown that the harnessing of an additional biological mechanism, that of inhibiting apoptosis, significantly enhances T cell responses to DNA vaccines comprising antigen-coding sequences, as well as linked sequences encoding such IPPs.

Intradermal vaccination by gene gun efficiently delivers a DNA vaccine into DCs of the skin, resulting in the activation and priming of antigen-specific T cells in vivo. DCs, however, have a limited life span, hindering their long-term ability to prime antigen-specific T cells. According to the present invention, a strategy that combines combination therapy with methods to prolong the survival of DNA-transduced DCs enhances priming of antigen-specific T cells and thereby, increase DNA vaccine potency. Co-delivery of DNA encoding inhibitors of apoptosis (BCL-xL, BCL-2, XIAP, dominant negative caspase-9, or dominant negative caspase-8) with DNA encoding an antigen (exemplified as HPV-16 E7 protein) prolongs the survival of transduced DCs. More importantly, vaccinated subjects exhibited significant enhancement in antigen-specific CD8⁺ T cell immune responses, resulting in a potent antitumor effect against antigen-expressing tumors. Among these anti-apoptotic factors, BCL-XL demonstrated the greatest enhancement of both antigen-specific immune responses and antitumor effects. Thus, co-administration of a combination therapy including a DNA vaccine with one or more DNA constructs encoding anti-apoptotic proteins provides a way to enhance DNA vaccine potency.

Serine protease inhibitor 6 (SPI-6), also called Serpinb9, inhibits granzyme B, and may thereby delay apoptotic cell death in DCs. Intradermal co-administration of DNA encoding SPI-6 with DNA constructs encoding E7 linked to various IPPs significantly increased E7-specific CD8⁺ T cell and CD4⁺ Th1 cell responses and enhanced anti-tumor effects when compared to vaccination without SPI-6. Thus, in certain embodiments, combined methods are used that enhance MHC class I and II antigen processing with delivery of SPI-6 to potentiate immunity.

A similar approach employs DNA-based alphaviral RNA replicon vectors, also called suicidal DNA vectors. To enhance the immune response to an antigen, *e.g.*, HPV E7, a DNA-based Semliki Forest virus vector, pSCA1, the antigen DNA is fused with DNA encoding an anti-apoptotic polypeptide such BCL-xL, a member of the BCL-2 family. pSCA1 encoding a fusion protein of an antigen polypeptide and/BCL-xL delays cell death in transfected DCs and generates significantly higher antigen-specific CD8⁺ T-cell-mediated immunity. The antiapoptotic function of BCL-xL is important for the enhancement of antigen-specific CD8⁺ T-cell responses. Thus, in one embodiment,

delaying cell death induced by an otherwise desirable suicidal DNA vaccine enhances its potency.

Thus, the present invention is also directed to combination therapies including administering a chemotherapeutic drug with a nucleic acid composition useful as an immunogen, comprising a combination of: (a) first nucleic acid vector comprising a first sequence encoding an antigenic polypeptide or peptide, which first vector optionally comprises a second sequence linked to the first sequence, which second sequence encodes an immunogenicity-potentiating polypeptide (IPP); b) a second nucleic acid vector encoding an anti-apoptotic polypeptide, wherein, when the second vector is administered with the first vector to a subject, a T cell-mediated immune response to the antigenic polypeptide or peptide is induced that is greater in magnitude and/or duration than an immune response induced by administration of the first vector alone. The first vector above may comprise a promoter operatively linked to the first and/or the second sequence.

In the above compositions the anti-apoptotic polypeptide may be selected from the group consisting of (a) BCL-xL, (b) BCL2, (c) XIAP, (d) FLICEc-s, (e) dominant-negative caspase-8, (f) dominant negative caspase-9, (g) SPI-6, and (h) a functional homologue or a derivative of any of (a)-(g). The anti-apoptotic DNA may be physically linked to the antigen-encoding DNA. Examples of this are provided in U.S. Patent Application publication No. 20070026076, incorporated by reference, primarily in the form of suicidal DNA vaccine vectors. Alternatively, the anti-apoptotic DNA may be administered separately from, but in combination with the antigen-encoding DNA molecule. Even more examples of the co-administration of these two types of vectors are provided in U.S. Patent Application No. 10/546,810 (publication number US 2007-0026076).

Exemplary nucleotide and amino acid sequences of anti-apoptotic and other proteins are provided in the sequence listing. Biologically active homologs of these proteins and constructs may also be used. Biologically active homologs is to be understood as described herein in the context of other proteins, *e.g.*, IPPs.

The coding sequence for BCL-xL as present in the pcDNA3 vector useful for the present invention is SEQ ID NO:55; the amino acid sequence of BCL-xL is SEQ ID NO:56; the sequence pcDNA3-BCL-xL is SEQ ID NO:57 (the BCL-xL coding sequence corresponds to nucleotides 983 to 1732); a pcDNA3 vector combining E7 and BCL-xL, designated pcDNA3-E7/BCL-xL is SEQ ID NO:58 (the E7 and BCL-xL sequences correspond to nucleotides 960 to 2009); the amino acid sequence of the E7-BCL-xL

chimeric or fusion polypeptide is SEQ ID NO:59; a mutant BCL-xL ("mtBCL-xL") DNA sequence is SEQ ID NO:60; the amino acid sequence of mtBCL-xL is SEQ ID NO:61; the amino acid sequence of the E7-mtBCL-xL chimeric or fusion polypeptide is SEQ ID NO:62; in the pcDNA-mtBCL-xL [SEQ ID NO:63] vector, this mutant sequence is inserted

5 in the same position that BCL-xL is inserted in SEQ ID NO:57 and in the pcDNA-E7/mtBCL-XL [SEQ ID NO:64], this sequence is inserted in the same position as the BCL-xL sequence is in SEQ ID NO:58; the sequence of the suicidal DNA vector pSCA1-BCL-xL is SEQ ID NO:65 (the BCL-xL sequence corresponds to nucleotides 7483 to 8232); the sequence of the "combined" vector, pSCA1-E7/BCL-xL is SEQ ID NO:66 (the sequence of

10 E7 and BCL-xL corresponds to nucleotides 7461 to 8510); the sequence of pSCA1-mtBCL-xL [SEQ ID NO:67] is the same as that for the wild type BCL-xL except that the mtBCL-xL sequence is inserted in the same position as the wild type sequence in the pSCA1-mtBCL-xL vector; the sequence pSCA1-E7/mtBCL-xL [SEQ ID NO:68] is the same as that for the wild type pSCA1-E7/BCL-xL above, except that the mtBCL-xL sequence is inserted

15 in the same position as the wild type sequence; the sequence of the vector pSG5-BCL-xL is SEQ ID NO:69 (the BCL-xL coding sequence corresponds to nucleotides 1061 to 1810); the sequence of the vector pSG5-mtBCL-xL is SEQ ID NO:70 with the mutant BCL-xL sequence has the mtBCL-xL, shown above, inserted in the same location as for the wild type vector immediately above; the nucleotide sequence of the DNA encoding the XIAP

20 anti-apoptotic protein is SEQ ID NO:71; the amino acid of the vector comprising the XIAP anti-apoptotic protein coding sequence is SEQ ID NO:72; the nucleotide sequence of the vector comprising the XIAP anti-apoptotic protein coding sequence, designated PSG5-XIAP is shown in SEQ ID NO:73 (with the XIAP corresponding to nucleotides 1055 to 2553); the sequence of DNA encoding the anti-apoptotic protein FLICEc-s is SEQ ID

25 NO:74; the amino acid sequence of the anti-apoptotic protein FLICEc-s is SEQ ID NO:75; the PSG5 vector encoding the anti-apoptotic protein FLICEc-s, designated PSG5-FLICEc-s, has the sequence SEQ ID NO:76 (with the FLICEc-s sequence corresponding to nucleotides 1049 to 2443); the sequence of DNA encoding the anti-apoptotic protein Bcl2 is SEQ ID NO:77; the amino acid sequence of Bcl2 is SEQ ID NO:78; the PSG5 vector

30 encoding Bcl2, designated PSG5-BCL2, has the sequence SEQ ID NO:79 (with the Bcl2 sequence corresponding to nucleotides 1061 to 1678); the pSG5-dn-caspase-8 vector is SEQ ID NO:80 (encoding the dominant-negative caspase-8 corresponding to nucleotides 1055 to 2449); the amino acid sequence of dn-caspase-8 is SEQ ID NO:81; the pSG5-dn-

caspase-9 vector is SEQ ID NO:82 (encoding the dominant-negative caspase-9 as
 nucleotides 1055 to 2305); the amino acid sequence of dn-caspase-9 is SEQ ID NO:83; the
 nucleotide sequence of murine serine protease inhibitor 6 (SPI-6, deposited in
 GENE BANK as NM 009256) is SEQ ID NO:84; the amino acid sequence of the SPI-6
 5 protein is SEQ ID NO:85; the nucleic acid sequence of the mutant SPI-6 (mtSPI6) is SEQ
 ID NO:86; the amino acid sequence of the mutant SPI-6 protein (mtSPI-6) is SEQ ID
 NO:87; the sequence of the pcDNA3-Spi6 vector is SEQ ID NO:88 (the SPI-6 sequence
 corresponds to nucleotides 960 to 2081); and the sequence of the mutant vector pcDNA3-
 mtSpi6 vector [SEQ ID NO:89] is the same as that above, except that the mtSPI-6 sequence
 10 is inserted in the same location in place of the wild type SPI-6.

Biologically active homologs of these nucleic acids and proteins may be used.
 Biologically active homologs are to be understood as described in the context of other
 proteins, *e.g.*, IPPs, herein. For example, a vector may encode an anti-apoptotic protein that
 is at least about 90%, 95%, 98% or 99% identical to that of a sequence set forth herein.

15

MHC class I/II activators

“MHC class I/II activators” refers to molecules or complexes thereof that increase
 immune responses by increasing MHC class I or II (“I/II”) antigen presentation, such as by
 increasing MHC class I, class II or class I and class II activity or gene expression. In one
 20 embodiment, an MHC class I/II activator is a nucleic acid encoding a protein that enhances
 MHC class I/II antigen presentation. Exemplary MHC class I/II activators include nucleic
 acids encoding an MHC class II associated invariant chain (Ii), in which the CLIP region is
 replaced with a T cell epitope, *e.g.*, a promiscuous T cell epitope, such as the Pan HLA-DR
 reactive epitope (PADRE), or a variant thereof. Other MHC class I/II activators are nucleic
 25 acids encoding the MHC class II transactivator CIITA or a variant thereof.

In one embodiment, an MHC class I/II activator is a nucleic acid, *e.g.*, an isolated
 nucleic acid, encoding a protein comprising, consisting or consisting essentially of an
 invariant (Ii) chain, wherein the CLIP region is replaced with a promiscuous CD4+ T cell
 epitope. A “promiscuous CD4+ T cell epitope” is used interchangeably with “universal
 30 CD4+ T cell epitope” and refers to peptides that bind to numerous histocompatibility
 alleles, *e.g.*, human MHC class II molecules. In one embodiment, the promiscuous CD4+ T
 cell epitope is a Pan HLA-DR reactive epitope (PADRE), thereby forming an Ii-PADRE
 protein that is encoded by an Ii-PADRE nucleic acid. In one embodiment, a nucleic acid

encodes an Ii chain, wherein amino acids 81-102 (KPVSQMRMATPLLMRPM (SEQ ID NO:92) are replaced with the PADRE sequence AKFVAAWTLKAAA (SEQ ID NO:93). An exemplary human Ii-PADRE amino acid sequence is set forth as SEQ ID NO:91, and is encoded by nucleotide sequence SEQ ID NO:90.

5 Also provided herein are variants of a protein consisting of SEQ ID NO:91. A protein may comprise, consist essentially of, or consist of an amino acid sequence that is at least about 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:91. A protein may comprise a PADRE that is identical to the PADRE of SEQ ID NO:91, i.e., consisting of SEQ ID NO:93. A protein may comprise a PADRE sequence that is at least
10 about 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:93; and/or an Ii sequence that is at least about 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the Ii sequence of SEQ ID NO:91.

 An amino acid sequence may differ from that of SEQ ID NO:91 or the Ii or PADRE sequences thereof by the addition, deletion or substitution of at least about 1, 2, 3, 4, 5, 6, 7,
15 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30 or more amino acids. In certain embodiments, a protein lacks one or more, *e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids at the C- and/or N-terminus and/or internal relative to that of SEQ ID NO:91 or the Ii or PADRE region thereof. In certain embodiments, an amino acid sequence differs from that of SEQ ID NO:93 or from that of the Ii sequence by the addition, deletion or substitution of
20 at least about 1, 2, 3, 4, or 5 amino acids.

 Variants of SEQ ID NO:91 or the PADRE or Ii regions thereof preferably have a biological activity. Such variants are referred to as “functional homologs” or “functional variants.” Functional homologs include variants of SEQ ID NO:91 that increase an immune response, *e.g.*, an antigen specific immune response, in a subject to whom it is
25 administered, or has any of the biological activities set forth in the Examples pertaining to Ii-PADRE. Variants of the PADRE sequence or the Ii sequence may have a biological activity that is associated with that of the wild type PADRE or Ii sequences, respectively. Biological activities can be determined as known in the art or as set forth in the Examples. In addition, comparison (or alignment) of the Ii and PADRE sequences from different
30 species is expected to be helpful in determining which amino acids may be varied and which ones should preferably not be varied.

 Other proteins provided herein comprise a PADRE amino acid sequence that replaces a larger portion of Ii, *e.g.*, wherein Ii is lacking about amino acids 81-103, 81-104,

81-105, 81-106, 81-107, 81-108, 81-109, 81-110 or more; is lacking about amino acids 70-102, 71-102, 72-102, 73-102, 74-102, 75-102, 76-102, 77-102, 78-102, 79-102, 80-102 or more.

- Other promiscuous CD4⁺ T cell epitopes that may be used instead of PADRE are
- 5 listed in Table 1.

Table 1: Exemplary promiscuous CD4⁺ T cell epitopes

Promiscuous CD4 ⁺ T cell epitopes	Reference
EBV-latent membrane protein 1(LMP1 ₁₅₉₋₁₇₅) YLQQNWWTLLVDLLWLL	(1)
MAGE-A6 ₁₇₂₋₁₈₇ ; IGHVYIFATCLGLSYD Mycoplasma penetrans HF-2 ₂₁₉₋₂₂₆ ; IYIFAACL	(2)
six-transmembrane epithelial antigen of prostate (STEAP) STEAP ₁₀₂₋₁₁₆ HQQYFYKIPILVINK STEAP ₁₉₂₋₂₀₆ LLNWAYQQVQQNKED	(3)
Taxol-resistance-associated gene-3 (TRAG3) ₃₅₋₄₈ EFHACW PAFTVLGE	(4)
Survivin ₁₀₋₂₄ WQPFLKDHRISTFKN	(5)
HPV 18-E6 ₅₂₋₆₆ ; LFVVYRDSIPHAACH HPV18-E6 ₉₇₋₁₁₁ ; GLYNLLIRCLRCQKP	(6)
Carcinoembryonic antigen ₁₇₇₋₁₈₉ ; LWWVNNQSLPVSP	(7)
mycobacterial antigen MPB70 MPB70 ₁₀₆₋₁₃₀ ; FSKLPASTIDELKTNSSLLTSILTY MPB70 ₁₆₆₋₁₉₃ ; GNADVVCGGVSTANATVYMIDSVLMPPA	(8)
HER-2 ₇₇₆₋₇₈₈ GSPYVSRLLGICL	(9)
HER-2 ₈₃₃₋₈₄₉ KVPIKWMALLESILRRRF	(10)
NY-ESO-1 ₁₁₉₋₁₄₃ PGVLLKEFTVSGNILTIRLTAADHR	(11)
Tetanus toxin ₁₀₈₄₋₁₀₉₉ VSIDKFRIFCKANPK Tetanus toxin ₁₁₇₄₋₁₁₈₉ LKFIKRYTPNNEIDS Tetanus toxin ₁₀₆₄₋₁₀₇₉ IREDNNITLKLDRCN Tetanus toxin ₉₄₇₋₉₆₇ FNNFTVSFWLRVPKVSASHLE Tetanus toxin ₈₃₀₋₈₄₃ QYIKANSKFIGITE HBV nuclear capsid ₅₀₋₆₉ PHHTALRQAILCWGELMTLA Influenza haemagglutinin ₃₀₇₋₃₁₉ PKYVKQNTLKLAT HBV surface antigen ₁₉₋₃₃ -FFLLTRILTIPQSLD Influenza matrix ₁₇₋₃₁ YSGPLKAEIAQRLEDV <i>P. falciparum</i> CSP ₃₈₀₋₃₉₈ EKKIAKMEKASSVFNVVN	(12)

1. Kobayashi, H., T. Nagato, M. Takahara, K. Sato, S. Kimura, N. Aoki, M. Azumi,
10 M. Tateno, Y. Harabuchi, and E. Celis. 2008. Induction of EBV-latent membrane protein 1-
specific MHC class II-restricted T-cell responses against natural killer lymphoma cells.
Cancer Res 68:901-908.

2. Vujanovic, L., M. Mandic, W. C. Olson, J. M. Kirkwood, and W. J. Storkus. 2007. A mycoplasma peptide elicits heteroclitic CD4⁺ T cell responses against tumor antigen MAGE-A6. *Clin Cancer Res* 13:6796-6806.
3. Kobayashi, H., T. Nagato, K. Sato, N. Aoki, S. Kimura, M. Murakami, H. Iizuka,
5 M. Azumi, H. Kakizaki, M. Tateno, and E. Celis. 2007. Recognition of prostate and melanoma tumor cells by six-transmembrane epithelial antigen of prostate-specific helper T lymphocytes in a human leukocyte antigen class II-restricted manner. *Cancer Res* 67:5498-5504.
4. Janjic, B., P. Andrade, X. F. Wang, J. Fourcade, C. Almunia, P. Kudela, A. Brufsky,
10 S. Jacobs, D. Friedland, R. Stoller, D. Gillet, R. B. Herberman, J. M. Kirkwood, B. Maillere, and H. M. Zarour. 2006. Spontaneous CD4⁺ T cell responses against TRAG-3 in patients with melanoma and breast cancers. *J Immunol* 177:2717-2727.
5. Piesche, M., Y. Hildebrandt, F. Zettl, B. Chapuy, M. Schmitz, G. Wulf, L. Trumper, and R. Schroers. 2007. Identification of a promiscuous HLA DR-restricted T-cell epitope
15 derived from the inhibitor of apoptosis protein survivin. *Hum Immunol* 68:572-576.
6. Facchinetti, V., S. Seresini, R. Longhi, C. Garavaglia, G. Casorati, and M. P. Protti. 2005. CD4⁺ T cell immunity against the human papillomavirus-18 E6 transforming protein in healthy donors: identification of promiscuous naturally processed epitopes. *Eur J Immunol* 35:806-815.
- 20 7. Campi, G., M. Crosti, G. Consogno, V. Facchinetti, B. M. Conti-Fine, R. Longhi, G. Casorati, P. Dellabona, and M. P. Protti. 2003. CD4(+) T cells from healthy subjects and colon cancer patients recognize a carcinoembryonic antigen-specific immunodominant epitope. *Cancer Res* 63:8481-8486.
8. Al-Attayah, R., F. A. Shaban, H. G. Wiker, F. Oftung, and A. S. Mustafa. 2003.
25 Synthetic peptides identify promiscuous human Th1 cell epitopes of the secreted mycobacterial antigen MPB70. *Infect Immun* 71:1953-1960.
9. Sotiriadou, R., S. A. Perez, A. D. Gritzapis, P. A. Sotiropoulou, H. Echner, S. Heinzl, A. Mamalaki, G. Pawelec, W. Voelter, C. N. Baxevanis, and M. Papamichail. 2001. Peptide HER2(776-788) represents a naturally processed broad MHC class II-
30 restricted T cell epitope. *Br J Cancer* 85:1527-1534.
10. Kobayashi, H., M. Wood, Y. Song, E. Appella, and E. Celis. 2000. Defining promiscuous MHC class II helper T-cell epitopes for the HER2/neu tumor antigen. *Cancer Res* 60:5228-5236.

11. Zarour, H. M., B. Maillere, V. Brusica, K. Coval, E. Williams, S. Pouvelle-Moratille, F. Castelli, S. Land, J. Bennouna, T. Logan, and J. M. Kirkwood. 2002. NY-ESO-1 119-143 is a promiscuous major histocompatibility complex class II T-helper epitope recognized by Th1- and Th2-type tumor-reactive CD4⁺ T cells. *Cancer Res* 62:213-218.

5 12. Falugi, F., R. Petracca, M. Mariani, E. Luzzi, S. Mancianti, V. Carinci, M. L. Melli, O. Finco, A. Wack, A. Di Tommaso, M. T. De Magistris, P. Costantino, G. Del Giudice, S. Abrignani, R. Rappuoli, and G. Grandi. 2001. Rationally designed strings of promiscuous CD4(+) T cell epitopes provide help to Haemophilus influenzae type b oligosaccharide: a model for new conjugate vaccines. *Eur J Immunol* 31:3816-3824.

10 The CLIP region in an Ii molecule, e.g., having the amino acid sequence of the Ii portion set forth in SEQ ID NO:91, may be replaced with any of the peptides in Table 2 or other promiscuous epitopes set forth in the references of Table 2, or functional variants thereof. Preferred epitopes include those from tetanus toxin and influenza. Any other promiscuous CD4⁺ T cell epitopes may be used, e.g., those described in the following references:

15 1. Campi, G., M. Crosti, G. Consogno, V. Facchinetti, B. M. Conti-Fine, R. Longhi, G. Casorati, P. Dellabona, and M. P. Protti. 2003. CD4(+) T cells from healthy subjects and colon cancer patients recognize a carcinoembryonic antigen-specific immunodominant epitope. *Cancer Res* 63:8481-8486.

20 2. Castelli, F. A., M. Leleu, S. Pouvelle-Moratille, S. Farci, H. M. Zarour, M. Andrieu, C. Auriault, A. Menez, B. Georges, and B. Maillere. 2007. Differential capacity of T cell priming in naive donors of promiscuous CD4⁺ T cell epitopes of HCV NS3 and Core proteins. *Eur J Immunol* 37:1513-1523.

25 3. Consogno, G., S. Manici, V. Facchinetti, A. Bachi, J. Hammer, B. M. Conti-Fine, C. Rugarli, C. Traversari, and M. P. Protti. 2003. Identification of immunodominant regions among promiscuous HLA-DR-restricted CD4⁺ T-cell epitopes on the tumor antigen MAGE-3. *Blood* 101:1038-1044.

4. Depil, S., O. Morales, F. A. Castelli, N. Delhem, V. Francois, B. Georges, F. Dufosse, F. Morschhauser, J. Hammer, B. Maillere, C. Auriault, and V. Pancre. 2007.

30 Determination of a HLA II promiscuous peptide cocktail as potential vaccine against EBV latency II malignancies. *J Immunother (1997)* 30:215-226.

5. Facchinetti, V., S. Seresini, R. Longhi, C. Garavaglia, G. Casorati, and M. P. Protti. 2005. CD4⁺ T cell immunity against the human papillomavirus-18 E6 transforming protein

in healthy donors: identification of promiscuous naturally processed epitopes. *Eur J Immunol* 35:806-815.

6. Kobayashi, H., T. Nagato, K. Sato, N. Aoki, S. Kimura, M. Murakami, H. Iizuka, M. Azumi, H. Kakizaki, M. Tateno, and E. Celis. 2007. Recognition of prostate and melanoma tumor cells by six-transmembrane epithelial antigen of prostate-specific helper T lymphocytes in a human leukocyte antigen class II-restricted manner. *Cancer Res* 67:5498-5504.
7. Kobayashi, H., M. Wood, Y. Song, E. Appella, and E. Celis. 2000. Defining promiscuous MHC class II helper T-cell epitopes for the HER2/neu tumor antigen. *Cancer Res* 60:5228-5236.
8. Mandic, M., C. Almunia, S. Vicel, D. Gillet, B. Janjic, K. Coval, B. Maillere, J. M. Kirkwood, and H. M. Zarour. 2003. The alternative open reading frame of LAGE-1 gives rise to multiple promiscuous HLA-DR-restricted epitopes recognized by T-helper 1-type tumor-reactive CD4+ T cells. *Cancer Res* 63:6506-6515.
9. Neumann, F., C. Wagner, S. Stevanovic, B. Kubuschok, C. Schormann, A. Mischo, K. Ertan, W. Schmidt, and M. Pfreundschuh. 2004. Identification of an HLA-DR-restricted peptide epitope with a promiscuous binding pattern derived from the cancer testis antigen HOM-MEL-40/SSX2. *Int J Cancer* 112:661-668.
10. Ohkuri, T., M. Sato, H. Abe, K. Tsuji, Y. Yamagishi, H. Ikeda, N. Matsubara, H. Kitamura, and T. Nishimura. 2007. Identification of a novel NY-ESO-1 promiscuous helper epitope presented by multiple MHC class II molecules found frequently in the Japanese population. *Cancer Sci* 98:1092-1098.
11. Piesche, M., Y. Hildebrandt, F. Zettl, B. Chapuy, M. Schmitz, G. Wulf, L. Trumper, and R. Schroers. 2007. Identification of a promiscuous HLA DR-restricted T-cell epitope derived from the inhibitor of apoptosis protein survivin. *Hum Immunol* 68:572-576.
12. Sotiriadou, R., S. A. Perez, A. D. Gritzapis, P. A. Sotiropoulou, H. Echner, S. Heinzl, A. Mamalaki, G. Pawelec, W. Voelter, C. N. Baxevanis, and M. Papamichail. 2001. Peptide HER2(776-788) represents a naturally processed broad MHC class II-restricted T cell epitope. *Br J Cancer* 85:1527-1534.
13. Texier, C., S. Pouvle-Moratille, C. Buhot, F. A. Castelli, C. Pecquet, A. Menez, F. Leynadier, and B. Maillere. 2002. Emerging principles for the design of promiscuous HLA-DR-restricted peptides: an example from the major bee venom allergen. *Eur J Immunol* 32:3699-3707.

14. Vujanovic, L., M. Mandic, W. C. Olson, J. M. Kirkwood, and W. J. Storkus. 2007. A mycoplasma peptide elicits heteroclitic CD4⁺ T cell responses against tumor antigen MAGE-A6. *Clin Cancer Res* 13:6796-6806.
15. Zarour, H. M., B. Maillere, V. Brusic, K. Coval, E. Williams, S. Pouvelle-Moratille,
5 F. Castelli, S. Land, J. Bennouna, T. Logan, and J. M. Kirkwood. 2002. NY-ESO-1 119-143 is a promiscuous major histocompatibility complex class II T-helper epitope recognized by Th1- and Th2-type tumor-reactive CD4⁺ T cells. *Cancer Res* 62:213-218.
16. Gao, M., H. P. Wang, Y. N. Wang, Y. Zhou, and Q. L. Wang. 2006. HCV-NS3 Th1
10 minigene vaccine based on invariant chain CLIP genetic substitution enhances CD4(+) Th1 cell responses in vivo. *Vaccine* 24:5491-5497.
17. Nagata, T., T. Aoshi, M. Suzuki, M. Uchijima, Y. H. Kim, Z. Yang, and Y. Koide. 2002. Induction of protective immunity to *Listeria monocytogenes* by immunization with plasmid DNA expressing a helper T-cell epitope that replaces the class II-associated invariant chain peptide of the invariant chain. *Infect Immun* 70:2676-2680.
- 15 18. Nagata, T., T. Higashi, T. Aoshi, M. Suzuki, M. Uchijima, and Y. Koide. 2001. Immunization with plasmid DNA encoding MHC class II binding peptide/CLIP-replaced invariant chain (Ii) induces specific helper T cells in vivo: the assessment of Ii p31 and p41 isoforms as vehicles for immunization. *Vaccine* 20:105-114.
19. Toda, M., M. Kasai, H. Hosokawa, N. Nakano, Y. Taniguchi, S. Inouye, S.
20 Kaminogawa, T. Takemori, and M. Sakaguchi. 2002. DNA vaccine using invariant chain gene for delivery of CD4⁺ T cell epitope peptide derived from Japanese cedar pollen allergen inhibits allergen-specific IgE response. *Eur J Immunol* 32:1631-1639.
20. van Bergen, J., M. Camps, R. Offringa, C. J. Melief, F. Ossendorp, and F. Koning. 2000. Superior tumor protection induced by a cellular vaccine carrying a tumor-specific T
25 helper epitope by genetic exchange of the class II-associated invariant chain peptide. *Cancer Res* 60:6427-6433.
21. van Tienhoven, E. A., C. T. ten Brink, J. van Bergen, F. Koning, W. van Eden, and C. P. Broeren. 2001. Induction of antigen specific CD4⁺ T cell responses by invariant chain based DNA vaccines. *Vaccine* 19:1515-1519.
- 30 In certain embodiments, the CLIP region of Ii is replaced with a T cell epitope, e.g., a CD4⁺ T cell epitope, such as a promiscuous CD4⁺ T cell epitope, with the *proviso* that the resulting construct is not one that has been publicly disclosed previously, e.g., one year prior to the filing of the priority application of the instant application. For example, in

certain embodiments, the epitope that replaces the CLIP region is not a promiscuous CD4+ T cell epitope from an HCV antigen, Listeria LLO antigen, ovalbumin antigen, Japanese cedar pollen allergen, MuLV env/gp70-derived helper epitope, and Heat Shock Protein 60 (described in references 16-21 above), or epitopes replacing CLIP regions that are

5 described in publications that are referenced to in the Examples.

In certain embodiments, a nucleic acid comprises, consists essentially of, or consists of the nucleotide sequence set forth in SEQ ID NO:90, or comprises a nucleotide sequence encoding the PADRE or Ii portion thereof. A nucleic acid may also comprise a nucleotide sequence that is at least about 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99%

10 identical to SEQ ID NO:90 and/or to the PADRE and/or to the Ii portion thereof. Nucleic acids may differ by the addition, deletion or substitution of one or more, *e.g.*, 1, 3, 5, 10, 15, 20, 25, 30 or more nucleotides, which may be located at the 5' end, 3' end, and/or internally to the sequence.

In certain embodiments, a nucleic acid encodes a protein that is a functional

15 homolog of an Ii-PADRE protein, with the proviso that the Ii sequence and/or PADRE sequence is (or are) not the wild-type or a naturally-occurring sequence, *e.g.*, the wild-type or naturally-occurring human sequence.

In another embodiment, an MHC class I/II activator is a protein that enhances MHC class II expression, *e.g.*, an MHC class II transactivator (CIITA). The nucleotide and

20 amino acid sequences of human CIITA are set forth as GenBank Accession Nos. P33076, NM_000246.3 and NP_000237.2 and set forth as SEQ ID NOs:94 and 95, respectively (GeneID: 4261)).

Variants of the protein may also be used. Exemplary variants comprise, consist essentially of, or consist of an amino acid sequence that is at least about 80%, 85%, 90%,

25 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:95. An amino acid sequence may differ from that of SEQ ID NO:95 by the addition, deletion or substitution of at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30 or more amino acids. In certain embodiments, a protein lacks one or more, *e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids at the C- and/or N-terminus and/or internally relative to that of SEQ ID NO:95.

30 The locations at which amino acid changes (i.e., deletions, additions or substitutions) may be made may be determined by comparing, i.e., aligning, the amino acid sequences of CIITA homologues, *e.g.*, those from various animal species.

Exemplary amino acids that may be changed include S286, S288 and S293. Indeed, as described in Greer et al., mutation of these amino acids results in a stronger transactivation function relative to the wild-type protein. Changes are preferably not made in the guanine-nucleotide binding motifs within residues 420-561, as these appear to be
5 necessary for CIITA activity (see Chin et al. (1997) PNAS 94:2501). Amino acids 59-94 have also been shown to be necessary for CIITA activity, as further described herein. Additional structure/function data are provided, *e.g.*, in Chin et al., *supra*.

In certain embodiments, a nucleic acid comprises, consists essentially of, or consists of the nucleotide sequence set forth in SEQ ID NO:94. A nucleic acid may also comprise a
10 nucleotide sequence that is at least about 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO:94. Nucleic acids may differ by the addition, deletion or substitution of one or more, *e.g.*, 1, 3, 5, 10, 15, 20, 25, 30 or more nucleotides, which may be located at the 5' end, 3' end, and/or internally to the sequence.

In certain embodiments, a nucleic acid encodes a protein that is a functional
15 homolog of a CIITA protein, with the *proviso* that the sequence is not the wild-type or a naturally-occurring sequence, *e.g.*, the wild-type or naturally-occurring human sequence.

Other nucleic acids encoding MHC class I/II activators that may be used include those that hybridize, *e.g.*, under stringent hybridization conditions to a nucleic acid encoding an MHC class I/II activator described herein, *e.g.*, consisting of SEQ ID NO:90 or
20 94 or portions thereof. Hybridization conditions are further described herein.

Nucleic acids encoding an MHC class I/II activator may be included in plasmids or expression vectors, such as those further described herein in the context of DNA vaccines.

In one embodiment, a nucleic acid encoding an Ii-PADRE protein or functional homolog thereof is administered to a subject who is also receiving a nucleic acid encoding a
25 CIITA protein or functional homolog thereof. The nucleic acids may be administered simultaneously or consecutively. The nucleic acids may also be linked, *i.e.*, forming one nucleic acid molecule. For example, one or more nucleotide sequences encoding an Ii-PADRE protein or a functional variant thereof; one or more nucleotide sequences encoding an antigen or a fusion protein comprising an antigen; one or more nucleotide sequences
30 encoding a CIITA protein or a functional variant thereof may be linked to each other, *i.e.*, present on one nucleic acid molecule.

Chemotherapeutic drugs

Drugs may also further be administered to a mammal in accordance with the methods and compositions taught herein. Generally, any drug that reduces the growth of cells without significantly affecting the immune system may be used, or at least not suppressing the immune system to the extent of eliminating the positive effects of a DNA vaccine that is administered to the subject. In one embodiment, the drugs are chemotherapeutic drugs.

A wide variety of chemotherapeutic drugs may be used, provided that the drug stimulates the effect of a vaccine, *e.g.*, DNA vaccine. In certain embodiments, a chemotherapeutic drug may be a drug that (a) induces apoptosis of cells, in particular, cancer cells, when contacted therewith; (b) reduces tumor burden; and/or (c) enhances CD8+ T cell-mediated antitumor immunity. In certain embodiments, the drug must also be one that does not inhibit the immune system, or at least not at certain concentrations.

In one embodiment, the chemotherapeutic drug is epigallocatechin-3-gallate (EGCG) or a chemical derivative or pharmaceutically acceptable salt thereof. Epigallocatechin gallate (EGCG) is the major polyphenol component found in green tea. EGCG has demonstrated antitumor effects in various human and animal models, including cancers of the breast, prostate, stomach, esophagus, colon, pancreas, skin, lung, and other sites. EGCG has been shown to act on different pathways to regulate cancer cell growth, survival, angiogenesis and metastasis. For example, some studies suggest that EGCG protects against cancer by causing cell cycle arrest and inducing apoptosis. It is also reported that telomerase inhibition might be one of the major mechanisms underlying the anticancer effects of EGCG. In comparison with commonly-used antitumor agents, including retinoids and doxorubicin, EGCG has a relatively low toxicity and is convenient to administer due to its oral bioavailability. Thus, EGCG has been used in clinical trials and appears to be a potentially ideal antitumor agent.

Exemplary analogs or derivatives of EGCG include (-)-EGCG, (+)-EGCG, (-)-EGCG-amide, (-)-GCG, (+)-GCG, (+)-EGCG-amide, (-)-ECG, (-)-CG, genistein, GTP-1, GTP-2, GTP-3, GTP-4, GTP-5, Bn-(+)-epigallocatechin gallate (US 2004/0186167, incorporated by reference), and dideoxy-epigallocatechin gallate (Furuta, *et al.*, Bioorg. Med. Chem. Letters, 2007, 11: 3095-3098). For additional examples, see US 2004/0186167 (incorporated by reference in its entirety); Waleh, *et al.*, Anticancer Res., 2005, 25: 397-402; Wai, *et al.*, Bioorg. Med. Chem., 2004, 12: 5587-5593; Smith, *et al.*,

Proteins: Struc. Func. & Bioinform., 2003, 54: 58-70; US 7,109,236 (incorporated by reference in its entirety); Landis-Piwowar, *et al.*, Int. J. Mol. Med., 2005, 15: 735-742; Landis-Piwowar, *et al.*, J. Cell. Phys., 2007, 213: 252-260; Daniel, *et al.*, Int. J. Mol. Med., 2006, 18: 625-632; Tanaka, *et al.*, Ang. Chemie Int., 2007, 46: 5934-5937.

5 Another chemotherapeutic drug that may be used is (a) 5,6 di-methylxanthenone-4-acetic acid (DMXAA), or a chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include xanthenone-4-acetic acid, flavone-8-acetic acid, xanthen-9-one-4-acetic acid, methyl (2,2-dimethyl-6-oxo-1,2-dihydro-6H-3,11-dioxacyclopenta[α]anthracen-10-yl)acetate, methyl (2-methyl-6-oxo-1,2-dihydro-6H-3,11-dioxacyclopenta[α]anthracen-10-yl)acetate, methyl (3,3-dimethyl-7-oxo-3H,7H-4,12-dioxabenz[α]anthracen-10-yl)acetate, methyl-6-alkyloxyxanthen-9-one-4-acetates (Gobbi, *et al.*, 2002, J. Med. Chem., 45: 4931) or a . For additional examples, see WO 2007/023302 A1, WO 2007/023307 A1, US 2006/9505, WO 2004/39363 A1, WO 2003/80044 A1, AU 2003/217035 A1, and AU 2003/282215 A1, each incorporated by
15 reference in their entirety.

A chemotherapeutic drug may also be cisplatin, or a chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include dichloro[4,4'-bis(4,4,4-trifluorobutyl)-2,2'-bipyridine]platinum (Kyler *et al.*, Bioorganic & Medicinal Chemistry, 2006, 14: 8692-8700), cis-[Rh2(-
20 O2CCH3)2(CH3CN)6]2+ (Lutterman *et al.*, J. Am. Chem. Soc., 2006, 128: 738 -739), (+)-cis-(1,1-Cyclobutanedicarboxylato)((2R)-2-methyl-1,4-butanediamine-N,N')platinum (O'Brien *et al.*, Cancer Res., 1992, 52: 4130-4134), cis-bisneodecanoato-trans-R,R-1,2-diaminocyclohexane platinum(II) (Lu *et al.*, J. of Clin. Oncol., 2005, 23: 3495-3501), carboplatin (Woloschuk, Drug Intell. Clin. Pharm., 1988, 22: 843-849), sebriplatin
25 (Kanazawa *et al.*, Head & Neck, 2006, 14: 38-43), satraplatin (Amorino *et al.*, Cancer Chemother. and Pharmacol., 2000, 46: 423-426), azane (dichloroplatinum) (CID: 11961987), azanide (CID: 6712951), platinol (CID: 5702198), lopac-P-4394 (CID: 5460033), MOLI001226 (CID: 450696), trichloroplatinum (CID: 420479), platinate(1-), amminetrichloro-, ammonium (CID: 160995), triammineplatinum (CID: 119232),
30 biocisplatinum (CID: 84691), platiblastin (CID: 2767) and pharmaceutically acceptable salts thereof. For additional examples, see US 5922689, US 4996337, US 4937358, US 4808730, US 6130245, US 7232919, and US 7038071, each incorporated by reference in their entirety.

Another chemotherapeutic drug that may be used is apigenin, or a chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include acacetin, chrysin, kampherol, luteolin, myricetin, naringenin, quercetin (Wang *et al.*, Nutrition and Cancer, 2004, 48: 106-114), puerarin (US 5 2006/0276458, incorporated by reference in its entirety) and pharmaceutically acceptable salts thereof. For additional examples, see US 2006/189680 A1, incorporated by reference in its entirety).

Another chemotherapeutic drug that may be used is doxorubicin, or a chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary 10 analogs or derivatives include anthracyclines, 3'-deamino-3'-(3-cyano-4-morpholinyl)doxorubicin, WP744 (Faderl, *et al.*, Cancer Res., 2001, 21: 3777-3784), annamycin (Zou, *et al.*, Cancer Chemother. Pharmacol., 1993, 32:190-196), 5-imino-daunorubicin, 2-pyrrolinodoxorubicin, DA-125 (Lim, *et al.*, Cancer Chemother. Pharmacol., 1997, 40: 23-30), 4-demethoxy-4'-O-methyldoxorubicin, PNU 152243 and 15 pharmaceutically acceptable salts thereof (Yuan, *et al.*, Anti-Cancer Drugs, 2004, 15: 641-646). For additional examples, see EP 1242438 B1, US 6630579, AU 2001/29066 B2, US 4826964, US 4672057, US 4314054, AU 2002/358298 A1, and US 4301277, each incorporated by reference in their entirety);

Other chemotherapeutic drugs that may be used are anti-death receptor 5 antibodies 20 and binding proteins, and their derivatives, including antibody fragments, single-chain antibodies (scFvs), Avimers, chimeric antibodies, humanized antibodies, human antibodies and peptides binding death receptor 5. For examples, see US 2007/31414 and US 2006/269554, each incorporated by reference in their entirety.

Another chemotherapeutic drug that may be used is bortezomib, or a chemical 25 derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include MLN-273 and pharmaceutically acceptable salts thereof (Witola, *et al.*, Eukaryotic Cell, 2007, doi:10.1128/EC.00229-07). For additional possibilities, see Groll, *et al.*, Structure, 14:451.

Another chemotherapeutic drug that may be used is 5-aza-2-deoxycytidine, or a 30 chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include other deoxycytidine derivatives and other nucleotide derivatives, such as deoxyadenine derivatives, deoxyguanine derivatives, deoxythymidine derivatives and pharmaceutically acceptable salts thereof.

Another chemotherapeutic drug that may be used is genistein, or a chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include 7-O-modified genistein derivatives (Zhang, *et al.*, Chem. & Biodiv., 2007, 4: 248-255), 4',5,7-tri[3-(2-hydroxyethylthio)propoxy]isoflavone, genistein glycosides (Polkowski, Cancer Letters, 2004, 203: 59-69), other genistein derivatives (Li, *et al.*, Chem & Biodiv., 2006, 4: 463-472; Sarkar, *et al.*, Mini. Rev. Med. Chem., 2006, 6: 401-407) or pharmaceutically acceptable salts thereof. For additional examples, see US 6541613, US 6958156, and WO/2002/081491, each incorporated by reference in their entirety.

Another chemotherapeutic drug that may be used is celecoxib, or a chemical derivative or analog thereof or a pharmaceutically acceptable salt thereof. Exemplary analogs or derivatives include N-(2-aminoethyl)-4-[5-(4-tolyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl]benzenesulfonamide, 4-[5-(4-aminophenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl]benzenesulfonamide, OSU03012 (Johnson, *et al.*, Blood, 2005, 105: 2504-2509), OSU03013 (Tong, *et al.*, Lung Cancer, 2006, 52: 117-124), dimethyl celecoxib (Backhus, *et al.*, J. Thorac. and Cardiovasc. Surg., 2005, 130: 1406-1412), and other derivatives or pharmaceutically acceptable salts thereof (Ding, *et al.*, Int. J. Cancer, 2005, 113: 803-810; Zhu, *et al.*, Cancer Res., 2004, 64: 4309-4318; Song, *et al.*, J. Natl. Cancer Inst., 2002, 94: 585-591). For additional examples, see US 7026346, incorporated by reference in its entirety.

One of skill in the art will readily recognize that other chemotherapeutics can be used with the methods disclosed in the present invention, including proteasome inhibitors (in addition to bortezomib) and inhibitors of DNA methylation. Other drugs that may be used include Paclitaxel; selenium compounds; SN38, etoposide, 5-Fluorouracil; VP-16, cox-2 inhibitors, Vioxx, cyclooxygenase-2 inhibitors, curcumin, MPC-6827, tamoxifen or flutamide, etoposide, PG490, 2-methoxyestradiol, AEE-788, aglycon protopanaxadiol, apolidine, ARQ-501, arsenic trioxide, BMS-387032, canertinib dihydrochloride, canfosfamide hydrochloride, combretastatin A-4 prodrug, idronoxil, indisulam, INGN-201, mapatumumab, motexafin gadolinium, oblimersen sodium, OGX-011, patupilone, PXD-101, rubitecan, tipifarnib, trabectedin PXD-101, methotrexate, Zerumbone, camptothecin, MG-98, VX-680, Ceflatonin, Oblimersen sodium, motexafin gadolinium, 1D09C3, PCK-3145, ME-2 and apoptosis-inducing-ligand (TRAIL/Apo-2 ligand). Others are provided in a report entitled "competitive outlook on apoptosis in oncology, Dec. 2006, published by

Bioseeker, and available, *e.g.*, at

http://bizwiz.bioseeker.com/bw/Archives/Files/TOC_BSG0612193.pdf.

Generally, any drug that affects an apoptosis target may also be used. Apoptosis targets include the tumour-necrosis factor (TNF)-related apoptosis-inducing ligand

5 (TRAIL) receptors, the BCL2 family of anti-apoptotic proteins (such as Bcl-2), inhibitor of apoptosis (IAP) proteins, MDM2, p53, TRAIL and caspases. Exemplary targets include B-cell CLL/lymphoma 2, Caspase 3, CD4 molecule, Cytosolic ovarian carcinoma antigen 1, Eukaryotic translation elongation factor 2, Farnesyltransferase, CAAX box, alpha; Fc fragment of IgE; Histone deacetylase 1; Histone deacetylase 2; Interleukin 13 receptor,
10 alpha 1; Phosphodiesterase 2A, cGMP-stimulated Phosphodiesterase 5A, cGMP-specific; Protein kinase C, beta 1; Steroid 5-alpha-reductase, alpha polypeptide 1; 8.1.15 Topoisomerase (DNA) I; Topoisomerase (DNA) II alpha; Tubulin, beta polypeptide; and p53 protein.

In certain embodiments, the compounds described herein, *e.g.*, EGCG, are

15 naturally-occurring and may, *e.g.*, be isolated from nature. Accordingly, in certain embodiments, a compound is used in an isolated or purified form, *i.e.*, it is not in a form in which it is naturally occurring. For example, an isolated compound may contain less than about 50%, 30%, 10%, 1%, 0.1% or 0.01% of a molecule that is associated with the compound in nature. A purified preparation of a compound may comprise at least about
20 50%, 70%, 80%, 90%, 95%, 97%, 98% or 99% of the compound, by molecule number or by weight. Compositions may comprise, consist essentially of consist of one or more compounds described herein. Some compounds that are naturally occurring may also be synthesized in a laboratory and may be referred to as “synthetic.” Yet other compounds described herein are non-naturally occurring.

25 In certain embodiments, the chemotherapeutic drug is in a preparation from a natural source, *e.g.*, a preparation from green tea.

Pharmaceutical compositions comprising 1, 2, 3, 4, 5 or more chemotherapeutic drugs or pharmaceutically acceptable salts thereof are also provided herein. A

pharmaceutical composition may comprise a pharmaceutically acceptable carrier. A

30 composition, *e.g.*, a pharmaceutical composition, may also comprise a vaccine, *e.g.*, a DNA vaccine, and optionally 1, 2, 3, 4, 5 or more vectors, *e.g.*, other DNA vaccines or other constructs, *e.g.*, described herein.

Compounds may be provided with a pharmaceutically acceptable salt. The term “pharmaceutically acceptable salts” is art-recognized, and includes relatively non-toxic, inorganic and organic acid addition salts of compositions, including without limitation, therapeutic agents, excipients, other materials and the like. Examples of pharmaceutically acceptable salts include those derived from mineral acids, such as hydrochloric acid and sulfuric acid, and those derived from organic acids, such as ethanesulfonic acid, benzenesulfonic acid, p-toluenesulfonic acid, and the like. Examples of suitable inorganic bases for the formation of salts include the hydroxides, carbonates, and bicarbonates of ammonia, sodium, lithium, potassium, calcium, magnesium, aluminum, zinc and the like.

Salts may also be formed with suitable organic bases, including those that are non-toxic and strong enough to form such salts. For purposes of illustration, the class of such organic bases may include mono-, di-, and trialkylamines, such as methylamine, dimethylamine, and triethylamine; mono-, di- or trihydroxyalkylamines such as mono-, di-, and triethanolamine; amino acids, such as arginine and lysine; guanidine; N-methylglucosamine; N-methylglucamine; L-glutamine; N-methylpiperazine; morpholine; ethylenediamine; N-benzylphenethylamine; (trihydroxymethyl)aminoethane; and the like. See, for example, *J. Pharm. Sci.*, 66:1-19 (1977).

Also provided herein are compositions and kits comprising one or more DNA vaccines and one or more chemotherapeutic drugs, and optionally one or more other constructs described herein.

Therapeutic compositions and their administration

The methods of the present invention can be practiced by coadministering a combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide, described herein alone or in a pharmaceutically acceptable carrier in a biologically-effective and/or a therapeutically-effective amount. The DNA vector may comprise major and minor capsid antigens from human papillomavirus type 16 (HPV16), bovine papillomavirus type 1, or any combination thereof. The DNA protein may be used in combination with a DNA vaccine, wherein the DNA vaccine comprises an antigenic peptide. In some embodiments, said antigenic peptide is selected from HPV16 E7, HPV16E6, OVA, CRT, HPV16 E7/CRT, or HPV16 E6/CRT.

Certain conditions as described herein are disclosed in the Examples. The composition may be given alone or in combination with another protein or peptide such as an immunostimulatory molecule. Treatment may include administration of an adjuvant, used in its broadest sense to include any nonspecific immune stimulating compound such as an interferon. Adjuvants contemplated herein include resorcinols, non-ionic surfactants such as polyoxyethylene oleyl ether and n-hexadecyl polyethylene ether.

A therapeutically effective amount is a dosage that, when given for an effective period of time, achieves the desired immunological or clinical effect.

A therapeutically active amount of the combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of the fusion protein to elicit a desired response in the individual. Dosage regimes may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation. A therapeutically effective amount of the protein, in cell associated form may be stated in terms of the protein or cell equivalents.

Thus an effective amount of at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide may be between about 1 nanogram and about 1 gram per kilogram of body weight of the recipient, between about 0.1 µg/kg and about 10mg/kg, between about 1 µg/kg and about 1 mg/kg. Dosage forms suitable for internal administration may contain (for the latter dose range) from about 0.1 µg to 100 µg of active ingredient per unit. The active ingredient may vary from 0.5 to 95% by weight based on the total weight of the composition. Those skilled in the art of immunotherapy will be able to adjust these doses without undue experimentation.

The at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide may be administered via gene gun. In other embodiments, they combination thereapy may be packaged into retrovirus vectors using packaging cell lines that produce replication-defective retroviruses, as is well-known in the art (e.g., Cone, R.D. *et al.*, *Proc Natl Acad Sci USA* 81:6349-53, 1984; Mann, RF *et al.*, *Cell* 33:153-9, 1983; Miller, AD *et al.*, *Molec Cell Biol* 5:431-7, 1985; Sorge, J, *et al.*, *Molec Cell Biol* 4:1730-7, 1984; Hock, RA *et al.*, *Nature* 320:257, 1986; Miller, AD *et al.*, *Molec Cell Biol* 6:2895-2902 (1986). Newer packaging cell lines which are efficient an

safe for gene transfer have also been described (Bank *et al.*, US Pat. 5,278,056, incorporated by reference).

The above approach can be utilized in a site specific manner to deliver the retroviral vector to the tissue or organ of choice. Thus, for example, a catheter delivery system can be used (Nabel, EG *et al.*, *Science* 244:1342 (1989)). Such methods, using either a retroviral vector or a liposome vector, are particularly useful to deliver the at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide to a blood vessel wall, or into the blood circulation of a tumor.

Other pharmaceutically acceptable carriers for the at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide according to the present invention are liposomes, pharmaceutical compositions in which the active protein is contained either dispersed or variously present in corpuscles consisting of aqueous concentric layers adherent to lipidic layers. The active protein may be present in the aqueous layer and in the lipidic layer, inside or outside, or, in any event, in the non-homogeneous system generally known as a liposomic suspension. The hydrophobic layer, or lipidic layer, generally, but not exclusively, comprises phospholipids such as lecithin and sphingomyelin, steroids such as cholesterol, more or less ionic surface active substances such as dicetylphosphate, stearylamine or phosphatidic acid, and/or other materials of a hydrophobic nature. Those skilled in the art will appreciate other suitable embodiments of the present liposomal formulations.

Embodiments disclosed herein also relate to methods of administering an at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide described herein to a subject in order to contact *in vivo* cells with such compositions. The routes of administration can vary with the location and nature of the cells to be contacted, and include, *e.g.*, intravascular, intradermal, transdermal, parenteral, intravenous, intramuscular, intranasal, subcutaneous, regional, percutaneous, intratracheal, intraperitoneal, intraarterial, intravesical, intratumoral, inhalation, perfusion, lavage, direct injection, and oral administration and formulation. In other embodiments, the routes of administration of the at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide may include (a) intratumoral, peritumoral, and/or intradermal delivery using gene gun, (b) intramuscularly (i.m.) injection using a conventional syringe needle; and (c) use of a needle-free biojector

such as the Biojector 2000 (Bioject Inc., Portland, OR) which is an injection device consisting of an injector and a disposable syringe. The orifice size controls the depth of penetration.

The term "systemic administration" refers to administration of at least one DNA
5 vector comprising a papillomavirus capsid antigen and at least one DNA vaccine
comprising an antigenic peptide as described herein, in a manner that results in the
introduction of the composition into the subject's circulatory system or otherwise permits
its spread throughout the body. "Regional" administration refers to administration into a
specific, and somewhat more limited, anatomical space, such as intraperitoneal, intrathecal,
10 subdural, or to a specific organ. "Local administration" refers to administration of a
composition or drug into a limited, or circumscribed, anatomic space, such as intratumoral
injection into a tumor mass, subcutaneous injections, intradermal or intramuscular
injections. Those of skill in the art will understand that local administration or regional
administration may also result in entry of a composition into the circulatory system i.e.,
15 rendering it systemic to one degree or another. For example, the term "intravascular" is
understood to refer to delivery into the vasculature of a patient, meaning into, within, or in
a vessel or vessels of the patient, whether for systemic, regional, and/or local
administration. In certain embodiments, the administration can be into a vessel considered
to be a vein (intravenous), while in others administration can be into a vessel considered to
20 be an artery. Veins include, but are not limited to, the internal jugular vein, a peripheral
vein, a coronary vein, a hepatic vein, the portal vein, great saphenous vein, the pulmonary
vein, superior vena cava, inferior vena cava, a gastric vein, a splenic vein, inferior
mesenteric vein, superior mesenteric vein, cephalic vein, and/or femoral vein. Arteries
include, but are not limited to, coronary artery, pulmonary artery, brachial artery, internal
25 carotid artery, aortic arch, femoral artery, peripheral artery, and/or ciliary artery. It is
contemplated that delivery may be through or to an arteriole or capillary.

Injection into the tumor vasculature is specifically contemplated for discrete, solid,
accessible tumors. Local, regional or systemic administration also may be appropriate. For
tumors of greater than about 4 cm, the volume to be administered can be about 4-10 ml
30 (preferably 10 ml), while for tumors of less than about 4 cm, a volume of about 1-3 ml can
be used (preferably 3 ml). Multiple injections delivered as single dose comprise about 0.1
to about 0.5 ml volumes. The at least one DNA vector comprising a papillomavirus capsid
antigen and at least one DNA vaccine comprising an antigenic peptide may advantageously

be contacted by administering multiple injections to the tumor, spaced at approximately 1 cm intervals.

Continuous administration also may be applied where appropriate. Such continuous administration, such as intravenous injection, may take place for a period of 9 days with
5 periodic injections every 3 days. Generally, the dose of the therapeutic composition via continuous administration will be equivalent to that given by a single or multiple injections, adjusted over a period of time during which the treatment occurs. Other routes of administration include oral, intranasal or rectal or any other route known in the art.

Depending on the route of administration, the at least one DNA vector comprising a
10 papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide may be coated in a material to protect the compound from the action of enzymes, acids and other natural conditions which may inactivate the compound. Thus it may be necessary to coat the composition with, or co-administer the composition with, a material to prevent its inactivation. For example, an enzyme inhibitors of nucleases or proteases (*e.g.*,
15 pancreatic trypsin inhibitor, diisopropylfluorophosphate and trasylol) or in an appropriate carrier such as liposomes (including water-in-oil-in-water emulsions as well as conventional liposomes (Strejan *et al.*, *J. Neuroimmunol* 7:27, 1984).

A chemotherapeutic drug may be administered in doses that are similar to the doses that the chemotherapeutic drug is used to be administered for cancer therapy.
20 Alternatively, it may be possible to use lower doses, *e.g.*, doses that are lower by 10%, 30%, 50%, or 2, 5, or 10 fold lower. Generally, the dose of chemotherapeutic agent is a dose that is effective to increase the effectiveness of the combination therapy comprising at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide, but less than a dose that results in significant
25 immunosuppression or immunosuppression that essentially cancels out the effect of the combination effects of the at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide.

The route of administration of chemotherapeutic drugs may depend on the drug. For use in the methods described herein, a chemotherapeutic drug may be used as it is
30 commonly used in known methods. Generally, the drugs will be administered orally or they may be injected. The regimen of administration of the drugs may be the same as it is commonly used in known methods. For example, certain drugs are administered one time, other drugs are administered every third day for a set period of time, yet other drugs are

administered every other day or every third, fourth, fifth, sixth day or weekly. The Examples provide exemplary regimens for administering the drugs, as well as at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide. The combination thereapy may be administered via
5 intraperitoneal gene gun injection two times at a three day interval. The intraperitoneal injection of the the combination therapy may be spread out over a period of 1 week, 2 weeks, 3 weeks, 4 weeks or longer. Likewise, the combination can be repeated administered over a 1 day, 2 day, 3 day, 4 day, or more intereval.

The compositions of the present invention, may be administered simultaneously or
10 subsequently. When administered simultaneously, the different components may be administered as one composition. Accordingly, also provided herein are compositions, *e.g.*, pharmaceutical compositions comprising one or more agents.

In one embodiment, a subject first receives one or more doses of the at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine
15 comprising an antigenic peptide. One may administer 1, 2, 3, 4, 5 or more doses of the at least one DNA vector comprising a papillomavirus capsid antigen and at least one DNA vaccine comprising an antigenic peptide.

A method may further comprise subjecting a subject to another cancer treatment, *e.g.*, radiotherapy, an anti-angiogenesis agent and/or a hydrogel-based system.

20 As used herein “pharmaceutically acceptable carrier” includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the therapeutic compositions is
25 contemplated. Supplementary active compounds can also be incorporated into the compositions.

Pharmaceutically acceptable diluents include saline and aqueous buffer solutions. Pharmaceutical compositions suitable for injection include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of
30 sterile injectable solutions or dispersion. Isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride may be included in the pharmaceutical composition. In all cases, the composition should be sterile and should be fluid. It should be stable under the conditions of manufacture and storage and must include preservatives

that prevent contamination with microorganisms such as bacteria and fungi. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations may contain a preservative to prevent the growth of microorganisms.

5 The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants.

10 Prevention of the action of microorganisms in the pharmaceutical composition can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like.

 Compositions may be formulated in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form refers to physically discrete units suited as unitary
15 dosages for a mammalian subject; each unit contains a predetermined quantity of active material (*e.g.*, DNA vector comprising a papillomavirus capsid antigen or DNA vaccine comprising an antigenic peptide) calculated to produce the desired therapeutic effect, in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on (a) the unique
20 characteristics of the active material and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such an active compound for the treatment of, and sensitivity of, individual subjects.

 For lung instillation, aerosolized solutions are used. In a sprayable aerosol preparations, the active protein may be in combination with a solid or liquid inert carrier
25 material. This may also be packaged in a squeeze bottle or in admixture with a pressurized volatile, normally gaseous propellant. The aerosol preparations can contain solvents, buffers, surfactants, and antioxidants in addition to the protein of the invention.

 Diseases that may be treated as described herein include hyper proliferative diseases, *e.g.*, cancer, whether localized or having metastasized. Exemplary cancers
30 include head and neck cancers and cervical cancer. Any cancer can be treated provided that there is a tumor associated antigen that is associated with the particular cancer. Other cancers include skin cancer, lung cancer, colon cancer, kidney cancer, breast cancer,

prostate cancer, pancreatic cancer, bone cancer, ovarian cancer, brain cancer, as well as blood cancers, *e.g.*, myeloma, leukemia and lymphoma.

Treating a subject includes curing a subject or improving at least one symptom of the disease or preventing or reducing the likelihood of the disease to return. For example, 5 treating a subject having cancer could be reducing the tumor mass of a subject, *e.g.*, by about 10%, 30%, 50%, 75%, 90% or more, eliminating the tumor, preventing or reducing the likelihood of the tumor to return, or partial or complete remission.

All references cited herein are all incorporated by reference herein, in their entirety, whether specifically incorporated or not. All publications, patents, patent applications, 10 GenBank sequences and ATCC deposits, cited herein are hereby expressly incorporated by reference for all purposes. In particular, all nucleotide sequences, amino acid sequences, nucleic constructs, DNA vaccines, methods of administration, particular orders of administration of DNA vaccines and agents that are described in the patents, patent applications and other publications referred to herein or authored by one or more of the 15 inventors of this application are specifically incorporated by reference herein. In case of conflict, the definitions within the instant application govern.

Having now fully described this invention, it will be appreciated by those skilled in the art that the same can be performed within a wide range of equivalent parameters, concentrations, and conditions without departing from the spirit and scope of the invention 20 and without undue experimentation.

The present description is further illustrated by the following examples, which should not be construed as limiting in any way.

EXAMPLES

25 DNA vaccines have emerged as attractive candidates for the control of HPV-associated cancers. However, DNA vaccines suffer from limited immunogenicity and strategies to enhance DNA vaccine potency are needed. It was previously shown that DNA vaccines encoding HPV-16 E7 antigen linked to calreticulin (CRT/E7) can elicit potent E7-specific CD8⁺ T cell antitumor immunity against E7-expressing tumors. Here, the inventors 30 report a novel vaccine technology consisting of co-administration of DNA encoding CRT/E7 with DNA encoding papillomavirus major and minor capsid antigens L1 and L2. The inventors found that this vaccine technology generated antigen-specific CD4⁺ T cell help as well as enhanced antigen-specific CD8⁺ T cell responses. Furthermore, the novel

vaccine technology generated potent therapeutic antitumor effect as well as L1/L2-specific neutralizing antibodies. Such strategy may potentially be extended to other antigenic systems for the control of infection and/or cancer.

The vaccine technology consists of DNA encoding HPV-16 E7 antigen linked to calreticulin (CRT/E7) and DNA encoding papillomavirus L1 or L2. DNA vaccine encoding CRT/E7 has been shown to elicit potent E7-specific CD8⁺ T cell response and significantly enhanced the antitumor effect. It was shown that co-administration with DNA encoding L1 or L2 in addition to CRT/E7 DNA induces even more potent E7-specific CD8⁺ T cell responses compared to vaccination with CRT/E7 alone. Furthermore, co-administration with DNA encoding L1/L2 led to the generation of antigen-specific CD4⁺ T cell help and L1/L2-specific neutralizing antibodies. In addition, the current vaccine technology can be applied through different routes of administration including intradermal as well as intramuscular administration.

Example 1: Co-administration with vectors encoding papillomavirus L1 or L2 significantly enhances the antigen-specific CD8⁺ T cell immune responses generated by CRT/E7 or OVA DNA vaccination

In order to characterize the antigen-specific CD8⁺ T cell immune responses generated by vaccination with CRT/E7 or OVA DNA in combination with vectors containing codon-optimized BPV1 L1 or L2 DNA, C57BL/6 mice (five per group) were vaccinated intradermally via gene gun with CRT/E7 or OVA DNA with or without BPV1 L1 or L2 DNA twice at 1-week intervals. Splenocytes from vaccinated mice were collected 1 week after last immunization and the E7 or OVA-specific T cell immune responses were characterized using intracellular cytokine staining followed by flow cytometry analysis. As shown in **Figure 1**, mice vaccinated with CRT/E7 or OVA DNA vaccine in combination with BPV1-L1 or L2 DNA generated significantly higher E7-specific and OVA-specific CD8⁺ T cell immune responses compared to mice vaccinated with CRT/E7 or OVA DNA alone. To evaluate whether the enhancement of antigen-specific CD8⁺ T cell immune responses by co-administration of BPV1 L1 or L2 DNA is also observed in other papillomavirus systems, HPV16 L1 or L2 DNA was co-administered with CRT/E7 or OVA DNA vaccination. Of note, co-administration of HPV16 L1 or L2 DNA with CRT/E7 or OVA DNA vaccination also generated significantly higher E7-specific CD8⁺ T cell responses compared to CRT/E7 or OVA DNA vaccination alone in mice (*Figure 2*). Thus, the data indicated that co-administration with papillomavirus L1 or L2 DNA significantly

enhanced the antigen-specific CD8⁺ T cell immune responses generated by DNA vaccination.

Example 2: Co-administration of papillomavirus L1 or L2 DNA with CRT/E7 or OVA DNA led to the generation of L1/L2-specific CD4⁺ T cell immune responses

5 In order to determine whether the co-administration with BPV1 L1 or L2 DNA with CRT/E7 DNA will lead to the generation of L1 or L2-specific CD4⁺ T cell immune responses, C57BL/6 mice (five per group) were vaccinated intradermally via gene gun with CRT/E7 DNA with BPV1 L1 or L2 DNA. Mice vaccinated with CRT/E7 DNA alone were used as negative controls. Splenocytes from vaccinated mice were collected 1 week after
10 last immunization and incubated with BPV1 L1/L2 virus-like particles (VLPs). The L1 or L2-specific CD4⁺ T cell immune responses were characterized using intracellular cytokine staining followed by flow cytometry analysis. As shown in **Figure 3A and B**, mice vaccinated with BPV1 L1 in combination with CRT/E7 DNA led to the generation of L1-specific CD4⁺ T cell immune responses. Similarly, vaccination with BPV1 L2 with CRT/E7
15 DNA led to significant level of L2-specific CD4⁺ T cell immune responses compared to vaccination with CRT/E7 alone. Again, to observe whether the CD4⁺ T cell responses elicited by this vaccination strategy extend to other papillomavirus systems, HPV16 L1 or L2 DNA with CRT/E7 or OVA DNA vaccination were co-administered. As shown in **Figures 3C and D**, compared to vaccination with CRT/E7 or OVA DNA alone, mice
20 vaccinated with HPV-16 L1 or L2 DNA in combination with CRT/E7 or OVA DNA led to the generation of L1 or L2-specific CD4⁺ T cell immune responses. These results suggested that co-administration with DNA encoding papillomavirus L1 or L2 with CRT/E7 DNA was able to generate appreciable L1 or L2-specific CD4⁺ T cell immune responses respectively in vaccinated mice.

25 **Example 3: Co-administration with BPV1 L1 DNA significantly enhances the therapeutic antitumor effects generated by CRT/E7 DNA vaccination**

In order to determine if the observed enhancement of antigen-specific CD8⁺ T cell immune responses by co-administration of BPV1 L1 DNA can translate into potent therapeutic antitumor effects, in vivo tumor treatment experiments were performed using an
30 HPV-16 E7-expressing murine tumor cell line, TC-1. TC-1 also expresses HPV16 E6, but does not contain either L1 or L2. C57BL/6 mice (five per group) were first challenged with TC-1 tumor cells subcutaneously. One week after tumor challenge, mice were treated intradermally via gene gun with CRT/E7 DNA alone, BPV1 L1 DNA alone or CRT/E7

DNA in combination with BPV1 L1 DNA. Vaccinated mice were boosted twice at 1-week intervals with the same dose and regimen. Tumor growth were monitored twice weekly by caliper measurements and palpations. As shown in **Figure 4**, tumor-bearing mice treated with CRT/E7 DNA vaccine in combination with BPV1 L1 DNA generated significantly reduced tumor volume and prolonged survival compared to mice treated with CRT/E7 DNA alone or BPV1 L1 DNA alone. Thus, the data indicated that co-administration with BPV1 L1 DNA is capable of significantly enhancing the therapeutic antitumor effects generated by CRT/E7 DNA vaccination.

Example 4: The enhancement in E7-specific CD8⁺ T cell immune responses are contributed by CD4⁺ T helper cells

In order to determine the mechanism underlying the observed enhancement of antigen-specific CD8⁺ T cell immune responses generated by coadministration with L1 or L2 DNA vectors, pcDNA3 encoding the reverse sequences of L1 (BPV L1 (-)) or L2 DNA (BPV L2 (-)) were generated. C57BL/6 mice (five per group) were vaccinated intradermally via gene gun with CRT/E7 DNA with or without the reverse sequence of BPV1 L1 or L2 DNA twice at 1-week intervals. Splenocytes from vaccinated mice were collected 1 week after last immunization and the E7-specific T cell immune responses were characterized using intracellular cytokine staining followed by flow cytometry analysis. It was found that mice vaccinated with CRT/E7 DNA vaccine in combination with the reverse sequence BPV1-L1 or L2 DNA did not lead to the increased frequency of E7-specific CD8⁺ T cell immune responses observed in mice vaccinated with CRT/E7 DNA with BPV1-L1 or L2 DNA (data not shown). The insert sequences does not express into L1 and L2 that activate CD4⁺ T cell help, thereby cannot enhance the E7-specific CD8⁺ T cell responses of the CRT/E7 vaccine. The result suggested that CD4⁺ T cell help generated by co-expression of L1 or L2 protein promoted E7-specific CD8⁺ T cell immune responses to CRT/E7 DNA vaccination.

Example 5: Co-administration of BPV1 L1 DNA with CRT/E7 DNA led to the generation of L1-specific neutralizing antibodies

In order to determine if co-administration of BPV1 L1 DNA with CRT/E7 DNA will lead to the generation of BPV1 L1-specific neutralizing antibodies, C57BL/6 mice (three per group) were immunized on days 1, 15, and 30 intradermally via gene gun with CRT/E7 and/or BPV1 L1 DNA. *In vitro* neutralization assays were performed using BPV1 L1 pseudovirus on twofold dilutions of antisera collected from the mice 2 weeks after the

final immunization. Mice vaccinated with CRT/E7 in combination with BPV1 L1 DNA were found to generate similar neutralizing antibody responses compared to mice vaccinated with BPV1 L1 DNA alone (**Figure 5**).

Example 6: Co-administration of BPV1 L1 or L2 DNA with OVA DNA vaccination

5 generated OVA-specific CD8⁺ T cell response through intramuscular administration

Finally, whether co-administration with papillomavirus L1 or L2 DNA can elicit potent antigen-specific CD8⁺ T cell responses when applied through different route of administration was evaluated. C57BL/6 mice (three per group) were vaccinated intramuscularly with OVA DNA with or without BPV1 L1 or L2 DNA twice at one-week
10 intervals. One week after the last immunization, PBMCs were collected and the OVA-specific CD8⁺ T cell immune responses were characterized through flow cytometry analysis. As shown in **Figure 6**, mice vaccinated intramuscularly with OVA DNA in combination with BPV1 L1 or L2 DNA generated significantly higher percentage of OVA-specific CD8⁺ T cells compared to mice vaccinated intramuscularly with OVA DNA alone.
15 This data indicated that co-administration with papillomavirus L1 or L2 DNA can lead to enhanced antigen-specific CD8⁺ T cell immune responses through the intramuscular route of administration.

Discussion

In the current study, it was showed that co-administration of vectors containing
20 codon-optimized BPV1 or HPV-16 L1 or L2 in combination with DNA vaccines could elicit enhanced antigen-specific CD8⁺ in both CRT/E7 and ovalbumin (OVA) antigenic systems. It was also demonstrated that co-administration of BPV1 or HPV-16 L1 or L2 DNA with CRT/E7 DNA led to the generation of L1/L2-specific CD4⁺ T cell immune responses. In addition, the observed enhancement of E7-specific CD8⁺ T cell immune
25 responses by L1 DNA also conferred improved therapeutic antitumor effects against an E7-expressing tumor. Moreover, co-administration of BPV1 L1 DNA induced generation of L1-specific neutralizing antibodies, which may serve to prevent further papillomavirus infections. Of note, the enhanced antigen-specific CD8⁺ T cell responses are observed in intradermal as well as intramuscular routes of administration. Indeed, it was shown that
30 vaccination with vector expressing L1 is protective against vaginal challenge with papillomavirus pseudovirions (Kwak K et al. *PLoS one* 2013, 8:e60507). Taken together, the studies suggested the promise of this approach for future clinical translation, and it can potentially be applied to other antigenic systems.

Here, it was shown that CD4⁺ T cell help plays an important role in the enhancement of antigen-specific CD8⁺ T cell immune responses observed in the vaccination regimen as co-administration with the reverse sequence of BPV1 L1 or L2 did not enhance the antigen-specific CD8⁺ T cell responses. Studies have shown that L1 and L2 are generally much more effective in eliciting CD4⁺ T cell responses (Steele JC et al. *British journal of cancer* 2005, 93:248-259; Mustafa W et al. *Viral immunology* 2009, 22:195-204; Pinto LA et al. *The Journal of infectious diseases* 2003, 188:327-338; Tobery TW et al. *Vaccine* 2003, 21:1539-1547; Williams OM et al. *Journal of virology* 2002, 76:7418-7429). While it is possible that BPV L1 or L2 could elicit some CD8⁺ T cell responses, it is believe the CD4⁺ T cell help generated by L1 or L2 DNA is the main contributor to the enhanced antigen-specific CD8⁺ T cell response observed. Furthermore, it has been shown that CD4⁺ T cells can help generate memory T cells (Shedlock DJ et al. *Science* 2003, 300:337-339; Overstreet MG et al. *PloS one* 2011, 6:e15948). Though the current study focuses on characterizing the antigen-specific CD8⁺ T cell therapeutic antitumor effect by the inventors' vaccination strategy, it will be of interest for future studies to further characterize the complete effect of CD4⁺ T cell and its ability to generate memory T cell responses for prolonged protection.

The present vaccination strategy was able to generate a potent therapeutic antitumor effect in tumor-bearing mice. Interestingly, co-administration of BPV1 L1 DNA with CRT/E7 DNA generated L1-specific neutralizing antibodies, which confers prophylactic value. It has also been shown that vaccination with L1 DNA induced L1-specific neutralizing antibodies in Balb/c mice (Kwak K et al. *PloS one* 2013, 8:e60507). Therefore, the vaccination strategy of co-administration of L1 DNA can generate potent antibody responses in more than one genetic background. Of note, studies have shown that papillomavirus L2 is generally not as effective in generating L2-specific neutralizing antibodies (Kwak K et al. *PloS one* 2013, 8:e60507; Cason J et al. *The Journal of general virology* 1993, 74 (Pt 12):2669-2677), thus co-administration with L1 DNA should be prioritized in future translation.

In the current study, it was shows that co-administration of BPV1 L1 DNA with CRT/E7 led to potent therapeutic antitumor effects and prolonged survival due to the enhanced antigen-specific CD8⁺ T cell responses by CD4⁺ T cell help. Since co-administration with HPV16 L1 DNA with CRT/E7 also significantly enhances antigen-specific CD8⁺ T cell responses by CD4⁺ T cell help, it was believe that potent therapeutic

antitumor effects should also be observed. To further promote clinical translation, subsequent investigations focusing on directly characterizing the antitumor effects of co-administrating HPV16 L1 with CRT/E7, and whether reversing the sequence of HPV16 L1 abolishes the antigen-specific CD8⁺ T cell response enhancement should be conducted. In addition, since the vaccination regimen achieved complete tumor suppression, the frequency of vaccination may be modified and further studied to determine the optimal vaccination regimen. Importantly, the data showed that the current vaccination strategy can generate enhanced antigen-specific CD8⁺ T cell responses by both intradermal and intramuscular vaccination. Since DNA vaccines are commonly applied intramuscularly in the clinic, future translation of the current vaccination technology may focus on the intramuscular route of administration.

Several strategies have been employed to induce CD4⁺ T helper cells to enhance antigen-specific CD8⁺ T cell-mediated immune responses generated by therapeutic HPV DNA vaccines. It was previously demonstrated that a DNA vaccine encoding invariant chain (Ii) with the class II-associated invariant peptide (CLIP) region replaced with the pan HLA-DR binding epitope (PADRE) could elicit potent PADRE-specific CD4⁺ T cell responses in vaccinated mice (Hung CF et al. *Mol Ther* 2007, 15:1211-1219). In addition, a co-administration of this DNA construct (Ii-PADRE) with DNA encoding HPV-16 E7 generated significantly greater CD8⁺ T cell immune responses relative to a co-administration of DNA encoding HPV-16 E7 with DNA encoding unmodified Ii (Hung CF et al. *Mol Ther* 2007, 15:1211-1219). Thus, the current study represented another promising approach to enhance CD4⁺ T help for the improvement of DNA vaccine potency, but with the added benefit of inducing prophylactic immunity too. It will be of interest in the future to perform a head-to-head comparison between vectors containing BPV L1/L2 DNA with Ii-PADRE DNA for their ability to enhance antigen-specific CD8⁺ T cell responses by therapeutic HPV DNA vaccines. Such information will facilitate the selection of the most effective or desirable DNA construct for improving DNA vaccine potency.

Another potential mechanism for enhancing the therapeutic HPV DNA vaccine potency through the co-administration with L1/L2 DNA is the potential formation of VLPs in transfected cells. It has been shown that the expression of L1 can lead to the formation of VLPs in eukaryotic cells (Zhou J et al. *Virology* 1992, 189:592-599). Furthermore, previous studies have demonstrated that papillomavirus VLPs can directly activate dendritic cells and thereby increase expression of costimulatory markers and MHC class I and II

molecules [26-28]. Thus, the co-administration of L1 DNA may lead to the local activation of DCs, resulting in further enhancement of antigen-specific CD8⁺ T cell immune responses generated by DNA vaccination. However, since the L2 only expression construct had a similar effect to L1 DNA, this suggested that the presence of VLP does not explain the enhanced CD8 T cell response upon co-administration with CRT-E7 or OVA constructs.

Strategies to enhance CD4⁺ T cell help may potentially be combined with other strategies to further enhance DNA vaccine potency. It was previously demonstrated that a significant enhancement of DNA vaccine potency by combining a strategy to prolong dendritic cell life and intracellular targeting strategies with a strategy to boost CD4⁺ T cell help (Kim D et al. *Human gene therapy* 2007, 18:1129-1139). Since all these strategies function via different mechanisms, the combination may potentially result in significantly enhanced antigen-specific immune responses and improved antitumor effects. For clinical translation, it will be desirable to identify the best combination of the different strategies in order to achieve the best DNA vaccine potency.

Conclusions

In summary, the present study demonstrated that the employment of DNA encoding papillomavirus L1 or L2 can lead to generation of antigen-specific CD4⁺ T cells and neutralizing antibodies, resulting in the improvement of therapeutic and preventive HPV DNA vaccine potency. The strategy may potentially be extended to other antigenic systems for the control of infection and/or cancer.

Materials & Methods

Mice. C57BL/6 mice (6–8 weeks old) were purchased from the National Cancer Institute (Frederick, MD). All animals were maintained under specific pathogen-free conditions at the Johns Hopkins Hospital (Baltimore, MD). All procedures were performed according to the Johns Hopkins Institutional Care and Use Committee approved protocols and in accordance with recommendations for the proper care of laboratory animals.

Cells and DNA constructs. TC-1 cells were obtained by co-transformation of primary C57BL/6 mouse lung epithelial cells with HPV-16 E6 and E7 and an activated ras oncogene as described previously (Lin KY et al. *Cancer Res* 1996, 56:21-26). They were maintained in RPMI medium supplemented with 2 mM glutamine, 1 mM sodium pyruvate, 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 5x10⁻⁵ M β -mercaptoethanol, 100 IU ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin, 10% fetal bovine serum, and cultured at 37 °C in a humidified incubator with 5% CO₂. The pcDNA3-OVA and

pcDNA3-CRT/E7 DNA constructs were generated as described previously (Cheng WF et al. *The Journal of clinical investigation* 2001, 108:669-678; Kim TW et al. *The Journal of clinical investigation* 2003, 112:109-117).

DNA vaccination. DNA-coated gold particles were prepared as described previously (Chen CH et al. *Cancer Res* 2000, 60:1035-1042). DNA-coated gold particles were delivered to the shaved abdominal region of mice using a helium-driven gene gun (Bio-Rad Laboratories Inc., Hercules, CA, USA) with a discharge pressure of 400 p.s.i. C57BL/6 mice were immunized with 2 µg of plasmid DNA to each mouse encoding pcDNA3-CRT/E7 mixed with pcDNA3-BPV1-L1 or L2 or HPV-L1 or L2 delivered to the shaved abdomen. The mice received a homologous boost 1 week later.

C57BL/6 mice were vaccinated with 10µg of pcDNA3-OVA DNA with 10µg of pcDNA3, pcDNA3-BPVL1, or pcDNA3-BPVL2 intramuscularly in the thigh muscle. The mice received a homologous boost 1 week later.

Intracellular cytokine staining and flow cytometry analysis. Splenocytes were harvested from mice 1 week after the last vaccination. Prior to intracellular cytokine staining, 5×10^6 splenocytes from each vaccination group were incubated for 16 h with 1 µg ml⁻¹ HPV-16 E7 H-2Db epitope (RAHYNIVTF) or OVA peptide (ISQAVHAAHAEINEAGR) (Kim TW et al. *The Journal of clinical investigation* 2003, 112:109-117), or 5 µg/mL BPV1 or HPV-16 L1/L2 VLPs in the presence of GolgiPlug (BD Pharmingen) (1 µl ml⁻¹). The stimulated splenocytes were then washed once with FACScan buffer and stained with phycoerythrin- conjugated monoclonal rat anti-mouse CD8α or CD4. Cells were subjected to intracellular cytokine staining using the Cytofix/Cytoperm kit according to the manufacturer's instructions (BD Pharmingen). Intracellular IFN-γ was stained with fluorescein isothiocyanate-conjugated rat anti-mouse IFN-γ to identify the immune response and cytokine levels. PBMCs were collected from peripheral blood 1 week after last intramuscular vaccination and stained with anti-mouse CD8α and OVA peptide (SIINFEKL) loaded H-2K^b tetramer. Flow cytometry analysis was performed using FACSCalibur with CELLQuest software (BD Biosciences, Mountain View, CA, USA).

In vivo tumor treatment experiments. C57BL/6 mice (five per group) were inoculated subcutaneously with 1×10^5 TC-1 tumor cells per mouse on the left flank. After 1 week, when tumor progression is usually observed, mice were vaccinated with DNA constructs pcDNA3-BPV1 L1 or L2 or pcDNA3-HPV L1 or L2 in conjunction with pCDNA3-

CRT/E7 or pcDNA3-OVA or the control empty vector. A homologous boost was administered 1 week after the first immunization. Mice were monitored for tumor growth by measuring diameters with calipers twice a week.

Neutralization assays. The BPV1 L1, L2 and HPV L1, L2 pseudovirions with encapsulated secreted alkaline phosphatase (SEAP) were generated by co-transfection of 293TT cells with plasmids encoding BPV1 L1, L2 or HPV L1 and L2 and a SEAP reporter plasmid as described previously (Pastrana DV et al. *Virology* 2004, 321:205-216). Cells collected after transfection were treated overnight with Brij 58 (0.5%), Benzonase (0.5%) and purified by centrifugation on an Optiprep step gradient (27, 33, and 39%) at 40,000 rpm for 4.5 h. Pseudovirus neutralization assays were carried out as outlined previously (Pastrana DV et al. *Virology* 2005, 337:365-372). Briefly, the pseudovirus and the pooled mouse immune sera were incubated for 1 h and the mixture was used to infect 293TT cells. 68-72 hours post-infection, the supernatants were collected and SEAP activity in the supernatants was measured by colorimetric assay. Serum neutralization titers were defined as the highest dilution that caused at least a 50% reduction in SEAP activity, compared to control pre-immune serum samples.

Statistical analysis. All data expressed as means $\pm$ s.d. are representative of at least two different experiments. Data for intracellular cytokine staining with flow cytometry analysis and tumor treatment experiments were evaluated by analysis of variance. Comparisons between individual data points were made using Student's t-test. All p values < 0.05 were considered significant.

List of Abbreviations: HPV – Human papillomavirus; CRT – calreticulin; BPV – bovine papillomavirus; OVA – ovalbumin; DCs – dendritic cells; VLPs – virus-like particles CLIP – class II-associated invariant peptide; PADRE – pan HLA-DR binding epitope SEAP – encapsulated secreted alkaline phosphatase.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention may become apparent to those skilled in the art upon review of this specification. The full scope of the invention should be determined by reference to the claims, along with

their full scope of equivalents, and the specification, along with such variations. Such equivalents are intended to be encompassed by the following claims.

[illegible]

SEQ ID NO:3
MHGDTPTLHE YMLDLQPETT DLYGYEGLND SSEEEDEIDG PAGQAEPRDA HYNIVTFCK CDSTLRLCVQ STHVDIRTLE
DLLMGTLGIV CPICSQKP 97

SEQ ID NO:5
MHQKRTAMFQ DPQERPRKLP QLCTELQTTI HDIILECVYC KQQLLRREVY DFAFRDLCIV YRDGNPYAVC DKCLKFYSKI
SEYRHYCYSL YGTTLEQQYN KPLCDLLIRC INCQKPLCPE EKQRHLDDKKQ RFHNI¹⁵⁸RGRWT GRCMSCCRSS RTRRETQL

SEQ ID NO:7

atgaaggc aaacctactggtcctgttaagtgacactgacagctgcagatgcagacacaaatatgtataggctaccatgcgaaacaattcaaccga
cactggtgcacagctactcgagaagaatgtgacagctgacacactctgttaacctgcgtcgaagcagccacaacggaacactatgtagattaa
aaggaatgacccctactcaaatgtggggaatgtaacatcgccggtatggctcttgggaaaccagaatgcgacccactgcttcagtgagatca
tggtcctacattctgtagaacaaccaaactctgagaatggaatatgttatccgaggagatttcacgtactatgaggagctgaggggagcaattgag
ctcagtgctcatcttgcgaagatttccgaattttcccaaaaagctcatggcccaacacacaaacggaagtaacggcagcagctgtctcc
atgaggggaaaaagcagtttttaccagaatttgcctgtgacgagagaaggagggctcatccccaaagctgaaaaattcttatgtgaacaaa
aaagggaaagaagtctctgtactgtgggtattcatcaccgcctaacagtaaggaaacaacagaatatctatcagaatgaaaaatgcttatgt
ctctgtagtgtacttcaattataacaggagattatcccggaatagcagaagaacccaaagtgaagagatacaagctggggaggtagaactatt
actggaccttctgtaaaaccccgagacacaataattttggcgcaaatggaaattcaatagcaccactgtatgctcttcgcactgagtgaaggcc
tttgggtccggcatcatcacctcaaacgcatcaatgcatgagtgtaacacgaaggtgtcaaacacccctgggagctataaacagcagctctcc
ttaccagaatatcacaccagctcaaataggagagtgccaaaaatcgtcaggagtgccaaattgaggatgggttacaggactaaggaaacatc
cgtccattcaatccagaggtctattttggagccattgccggttttatgaaaggggatggactggaatgatagatgaggttgatggttatcat
catcagaatgaacagggtacaggctgtcagcggtatcaaaaaagcacacaaatgccattaacgggattacaaacaggtgaaactgttatt
cgagaataatgaacatttcaattcacagctgtgggtaaagaattcaacaaattagaaaaaaggatggaataatttaataaaaaaagttgatgatg
gatttctggacatttggacataatacgtgcagaattgttgatttctactctggaaaatgaaaggactctggatttccatgactcaaatgtgaagaat
ctgtatgagaagaagtaaaagccaatttaagaataatgcgaagaataatcggaatggatgttttgagttctaccacagctgtgacaatgaaatg
catggaagtgtaagaataatgggactttatgattatcccaatatattcagaagagtc aaagttgaacaggggaaaaggtagatggagtgaaattgg
aatcaatgggggactcatcagattcttgcgactctactcaactgtccgcagttcactggtgcttttggctctccctggggggcaatcagtttctgg
atgtgttctaatggaatttttgcagtcgacaatatgcactga

- 88 -

SEQ ID NO:9

MGSI GAASMEFCFDFVKELKVHHANENIFYCPIAIMSALAMVYLGA KDSTRTQINKVVRFDKLPFGGDSIEAQCGTSVNV
HSSLRDILNQITKPNVDVYSFSLASRLYAERYPILP EYLQCVKELYRGGLEPINFQTAADQARELINSWVESQTNIGI RN
VLQPSVSDSQ TAMVLVNAIVFKLWEKTFKDEDTPAMPFRTVEQESKPVQMMYQIGLFRVASMASEKMKILEL PFGSGTM
SMLVLLPDEVSGLEQLESII NFEKLEWTEWSSNVMEERKIKVYLPRMKMEEKYNLTSVLMAMGITDVFSSS ANLSGISSAE
SLKISQAVHAHA AEINEAGREVVGS AEAGVDAASVSEEFRA DHPFLFCIKHIATNAVLFFGRCVSP

SEQ ID NO:10

ATGGCGGCCCGCGCGCCCGCGCGCTGCTCCTGCTGCTGCTGGCAGGCCTTG CACATGGCGCCTCAGCACTCTTTGAGGATCTAATCAT
GCATGGAGATACACCTACATTGCATGAATATATGTTAGATTGCAACCAGAGACA ACTGATCTCTACTGTTATGAGCAATTAAATGACAGCT
CAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATTA CAATATTGTTACCTTTTGTGCAAGTGTGAC
TCTACGCTTCGGTTGTGCGTACAAAGCACACACGTAGACATTCTGACTTTTGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCAT
CTGTTCTCAGGATCTTAACAACATGTTGATCCCCATTGCTGTGGCGGTGCCCTGGCAGGGCTGGTCTCATCGTCTCATGCTACCTCA
TTGGCAGGAAGAGGAGTCACGCCGGCTATCAGACCATCTAG

SEQ ID NO:11

MAAPGARRPL LLLLLAGLAH GASALFEDLI MHGDTPTLHE YMLDLQPETT DLYCYEQ LND SSEEEDEIDG PAGQAE PDRA
HYNIVTFCK CDSTLRLCVQ SHVDIRPLE DLLMGLGIV CPICSDLNN MLIPIAVGGA LAGLV LIVLI AYLI GRKRSH
AGYQTI

SEQ ID NO:12

GACGGATCGGGAGATCTCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCT
TGTGTGTTGGAGGTCGCTGAGTAGTGC CGCGAGCAAAATTTAAGCTACACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAG
GGTTAGGCGTTTTCGCGTCTTCGCGATGTACGGGCCAGATACCGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGG
GGTCATTAGTTTCATTAGCTTACATGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCAACGACCCCGCCCA
TTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTACGGTAAACTGCCCA
CTTGGCAGTACATCAAGTATATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTTAGCCAGTACA
TGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGG
CGTGGATAGCGGTTTACGTCACGGGGATTCCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT
TCCAAAATGTCGTAAACAATCCGCCCATTTGACGCAAAATGGGCGGTAGGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAA
CTAGAGAACCCCATGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACGGGCCCTCTAGA
CTCGAGCGCGCCCATGCTGCTGGATATCTGCAGAATTCatggcgggccccggcgcccgggcgccgctgctctgctgctgcaaggcct
tggcacatggcgccctcagcactccttgaggatcctaatcatgcatggagatacacctacattgcatgaatatatgttagatttgcaaccagaga
caactgatctctactgttatgagcaattaaatgacagctcagaggaggagatgaaatagatggtccagctgggacaagcagaaccggacaga
gcccattacaatatgttaccctttgttgcaagtgtgactctacgcttcggttggcggtacaaagcacacagtagacattcgtactttgga
agacctgttaattgggcacacatggaattgtgccccatctgttctcaggatcttaacaacatgttgatccccattgctgtggggcggtgccc
tggcaggggtggtcctcatcgtcctcattgacctcctcattggcaggaagaggagtccacgcccggctatcagaccatctagGGATCCGAGCTC
GGTACCAAGCTTAAAGTTTAAACCGCTGATCAGCCTCGACTGTGCTTCTAGTTGGCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCTCT
TGACCTTGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGAAATTGCATCGCATTTGCTGAGTAGGTGTCACTCTATTCTGGGG
GGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGA
AAGAACCAGCTGGGGCTTAGGGGGTATCCCCACGCGCCTGTAGCGGGCGCTTAAGCGCGCGCGGTGTGGTGGTTACGCGCAGCGTGACCG
CTACACTTCCGACGCGCCTAGCGCCCGCTCCTTTCGCTTCTCCTTCTCCTTCTCGCCACGTTTCGCCGCTTTCGCCGCTCAAGCTCTAAAT
CGGGCATACCTTTAGGGTTCCGATTTAGTGCTTTACGGGACCTCGACCCCAAAACTTGATTAGGGTGTAGGTTTACGTAAGTGGGCCATC
GCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTAAATAGTGGACTCTTGTCCAAACTGGAACAACACTCAACCCTA
TCTCGGTCTATTCTTTGATTATAAGGGATTTTGGGGAATTCGGCCTATTGGTTAAAAAATGAGCTGATTAAACAAAAATTAACGCGAAT
TAATTTCTGTGAAGTGTGTCTCAGTTAGGTTGGAAAGTCCCCAGGCTCCCCAGGCAGGAGAGTATGCAAAAGCATGCATCTCAATTAGTC
AGCAACCAGGTGTGGAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCC
TAACCTCGGCCATCCCGCCCTAACTCCGCCAGTTCCGCCCATTTCTCGGCCCATGGCTGACTAATTTTTTTTATTTATGAGAGGCCGAG
GCCGCTCTGCCTCTGAGCTATTCAGAAAGTAGTGAGGAGGCTTTTTTGGAGCCCTAGGCTTTTGCAAAAGCTCCCGGGAGCTTGATATC
CATTTTCGAGTCTGATCAAGAGACAGGATGAGGATCGTTTTCGATGATTGAACAAGATGGATTGCACGAGGCTTCCGGCCGCTTGGGTGG
AGAGGCTATTTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCGGTGTCCGGCTGTACGCGCAGGGGGCGCCGGTTCTT
TTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAATCGCAGGACGAGGCGAGGCGCGGTATCGTGGCTGGCCACGACGGGCGTTCTTTCGCG
AGCTGTGCTCGAGGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTATCTCACCTTGCTC
CTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACAT
CGCATCGAGCGAGCACGTAATCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAAC
GTTCCGCTAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGCTGACCTGACCATGGCGATGCTGCTTCCGCAATCATGGTGAAATG
GTCCTTTTCGATTCTGAGTCTGACTGTGGCGGCTTGGCGGACCGGTATCAGGACATAGCGTTGGCTACCGTGTATATGCTGAAAGAG
CTTGGCGGGAATGGGCTGACCGCTTCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGAGCGCATCGCTTCTATCGCTTCTTGACGA
GTTCTTCTGAGCGGGAATCTGGGGTTTCAAAATGACCGACCAAGCGACGCCCAACCTGCCATCACAGAGATTTCGATTCCACCGCCGCTTCTA
TGAAAGGTTGGGCTTCGGAATCGTTTTCGGGACGCGCGGTGGATGATCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCAACCCCA
ACTTGTTTTATTCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTCTACTGCATTCTAGTTGT
GGTTGTGTCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCTAGCTGTTTCC
TGTGTGAAATTTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC
TCACATTAATTCGCTTGCCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGA
GGCGGTTTTCGCTATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC
AAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCGAGAAAGACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCGCGGTTGCTGGCGTTTTCCTATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
GGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTT
TCTCCCTTCGGGAAGCGTGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCTCGGTGATGGTTCGCTCCAAGCTGGGCTGTGTGC
ACGAACCCCGGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAACCCGGTAAAGACGACTTATCGCCACTGGCA
CGAGCCACTGTTAAGCAGGATAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACCTAGCTAGAAAG
GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTA
CGGTTGGTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT
CAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTTCATGAGGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATATAAAATGAAGTTT
TAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTT
GTTTCATCCATAGTTGCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA
GACCCACGCTACCGGCTCAGGATTTATCAGCAATAAACACGACGCGGAGGCGGAGGCGAGGCGAGAAGTGTCTGCAACTTATCCGCTC
CATCCAGTCTATTAATTTGTTGCGGGAAGCTAGAGTAAGTAGTTTCGCGGATTAATAGTTTTCGCGCAACGTTGTGCCATGTCTACAGGCTC
TGGTGTACGCTCGTCTGTTGGTATGGCTTCATTGAGCTCCGGTTCCTAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAA
CGGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATCTCTCT

TACTGTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT
 CTTGCCCGCGTCAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAGTGCTCATTCATTGGAAAACTTTTCGGGGCGAAAACTC
 TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTC
 TGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAATAAGGCGGACACGGAAATGTTGAATACCTCATACTCTCTTTTCAAT
 ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACA
 TTTCCCCGAAAAGTGCCACCTGACGTC

SEQ ID NO:13

atggctcg tgcggtcggg atcgacctcg ggaccaccaa ctccgtcgtc tcggttcttg aaggtggcga cccggtcgtc
 gtcgccaact ccgagggtc caggaccacc ccgtcaattg tcgctgtcgc ccgcaacggt gaggtgctgg tcggccagcc
 cgccaagaac caggcagtga ccaacgtcga tcgcaaccgt cgctcggtca agcgacacat gggcagcgac tggctccatag
 agattgacgg caagaaatac acccgccggg agatcagcgc ccgcatcttg atgaagctga agcgcgacgc cgaggcctac
 ctcggtgagg acattaccga cgcggttatc acgacgcccg cctacttcaa tgacgcccag cgtcaggcca ccaaggacgc
 cggccagatc gccggcctca acgtgctgag gatcgtaac gagccgacgc cggccgctgt ggcctacggc ctcgacaagg
 gcgagaagg gacggaatc ctggtcttcg acttgggtgg ttgcactttc gacgtttccc tgctggagat cggcgagggt
 gtggttgagg tccgtgccac ttcgggtgac aaccacctcg gcggcgacga ctgggaccag cgggtcgtcg attggctggt
 ggacaagttc aagggcacca gcggcatcga tctgaccaag gacaagatgg cgatgcagcg gctgcgggaa gccgcggaga
 aggcaaatg cagcgtgagt tcgagtcagt ccacctcgat caacctgccc tacatcacgc tcgacgcccga caagaacccg
 ttgttcttag acgagcagct gaccccgcgcg gagttccaac ggatcactca ggacctgctg gaccgcactc gcaagccgtt
 ccagtcggtg atcgctgaca ccggcatttc ggtgtcgagg atcgatcacg ttgtgctcgt ggggtggttcg acccggtatg
 ccgcggtgac cgatctggtc aaggaactca ccggcgccaa tgaacccaac aaggcgctga accccgatga ggttgcgag
 ctgggagcgc ctctgcaggc cggcgctcctc aaggcgaggg ttgaagacgt tctgctgctt gatgttacc cgtgagcct
 gggatcagc accaaggcg ggggtgatgac caggctcctc gagcgcaaca ccacgatccc caccaagcgg tcggagactt
 tcaccaccgc cgacgacaac caaccgtcgg tgcagatcca ggtctatcag ggggagcgtg agatcgccgc gcacaacaag
 ttgctcgggt ccttcgagct gacggcgcac ccggcgggat tccgagatc gaggtcactt tcgacatcga
 cgccaacggc attgtgcacg tcaccgcca ggacaaggcg accggcaagg agaacacgat ccgaatccag gaaggctcgg
 gctgttccaa ggaagacatt gacggcatga tcaaggacgc cgaagcgac gccgaggagg atcgcaagcg tcgagaggag
 gccgatgttc gtaataaagc cgagacattg gtctaccaga cggagaagtt cgtcaaagaa cagcgtgagg ccgagggtg
 ttcgaaggta cctgaagaca cgtgaacaa ggttgatgcc gcgggtggcg aagcgaaggc ggcacttggc ggatcgga
 ttctggccat caagtcggcg atggagaagc tgggccagga gtcgaggct ctggggcaag cgatctacga agcagctcag
 gctgcgtcac aggccactgg cgtgcccac ccggcgggcg agccggcgcg tgcccccccc ggctcggctg atgacgttgt
 ggacgcgagg gtggtcgacg acggccggga ggccaagtga

SEQ ID NO:14

MARAVGIDLG ITNSVSVLE GGDVPPVANS EGSRTTPSIV AFARNGEVLV GQPAKNQAVT NVDRTVRSVK RHMGSWSIE
 IDGKKYTAPF ISARILMKLK RDAEAYLGED ITDAVITTPA YFNDAGRQAT KDAGQIAGLN VLRIVNEPTA AALAYGLDKG
 EKEGRILVFD LGGGTFDVSF LEIGEGVVEV RATSGDNHLG GDDWDQVRVD WLVDKFKGTS GIDLTKDKMA MQRLEAAEK
 AKIELSSSSQS TSINLPYITV DADKNPLFLD EQLTRAEFQR ITQDLLDRTR KPFQSVIADT GISVSEIDHV VLVGGSTRMP
 AVTDLVKELT GGKEPNKGVN PDEVVAVGAA LQAGVLKGEV KDVLLEDVTP LSLGIETKGG VMTRLIERNT TIPTKRSETF
 TTADDNQPSV QIQVYQGERE IAAHNKLLGS FELTGIPPAP RGIPQIEVTF DIDANGIVHV TAKDKGTGKE NTIRIQEGSG
 LSKEDIDRMI KDAEHAHAED RKRREEADV NRQAETLVYQT EKFKVKEQREA EGGSKVPEDT LNKVDAVAE AKAAALGSDI
 SAIKSAMEKL QGESQALGQA IYEAQAASQ ATGAHPGGE PGGHPGSAD DVVDAEVVDD GREAK

SEQ ID NO:15

1/1 31/11
 ATG CAT GGA GAT ACA CCT ACA TTG CAT GAA TAT ATG TTA GAT TTG CAA CCA GAG ACA ACT
 61/21 91/31
 GAT CTC TAC TGT TAT GAG CAA TTA AAT GAC AGC TCA GAG GAG GAG GAT GAA ATA GAT GGT
 121/41 151/51
 CCA GCT GGA CAA GCA GAA CCG GAC AGA GCC CAT TAC AAT ATT GTA ACC TTT TGT TGC AAG
 181/61 211/71
 TGT GAC TCT ACG CTT CGG TTG TGC GTA CAA AGC ACA CAC GTA GAC ATT CGT ACT TTG GAA
 241/81 271/91
 GAC CTG TTA ATG GGC ACA CTA GGA ATT GTG TGC CCC ATC TGT TCT CAA GGA TCC atg gct
 301/101 331/111
 cgt gcg gtc ggg atc gac ctc ggg acc acc aac tcc gtc gtc tgc gtt ctg gaa ggt ggc
 361/121 391/131
 gac ccg gtc gtc gtc gcc aac tcc gag ggc tcc agg acc acc ccg tca att gtc gcg ttc
 421/141 451/151
 gcc cgc aac ggt gag gtg ctg gtc ggc cag ccc gcc aag aac cag gca gtg acc aac gtc
 481/161 511/171
 gat cgc acc gtg cgc tgc gtc aag cga cac atg ggc agc gac tgg tcc ata gag att gac
 541/181 571/191
 ggc aag aaa tac acc gcg ccg gag atc agc gcc cgc att ctg atg aag ctg aag cgc gac
 601/201 631/211
 gcc gag gcc tac ctc ggt gag gac att acc gag gcg gtt atc acg acg ccc gcc tac ttc
 661/221 691/231
 aat gac gcc cag cgt cag gcc acc aag gac gcc ggc cag atc gcc ggc ctc aac gtg ctg
 721/241 751/251
 cgg atc gtc aac gag ccg acc gcg gcc gcg ctg gcc tac ggc ctc gac aag ggc gag aag
 781/261 811/271
 gag cag cga atc ctg gtc ttc gac ttg ggt ggt ggc act ttc gac gtt tcc ctg ctg gag
 841/281 871/291
 atc gcc gag ggt gtg gtt gag gtc cgt gcc act tgc ggt gac aac cac ctc gcc gcc gac
 901/301 931/311
 gac tgg gac cag ccg gtc gtc gat tgg ctg gtg gac aag ttc aag ggc acc agc ggc atc
 961/321 991/331
 gat ctg acc aag gac aag atg gcg atg cag cgg ctg cgg gaa gcc gcc gag aag gca aag
 1021/341 1051/351
 atc gag ctg agt tgc agt cag tcc acc tgc atc aac ctg ccc tac atc acc gtc gac gcc

1081/361
 gac aag aac ccg ttg ttc tta gac gag cag
 1141/381
 cag gac ctg ctg gac cgc act cgc aag ccg
 1201/401
 tcg gtg tcg gag atc gat cac gtt gtg ctc
 1261/421
 acc gat ctg gtc aag gaa ctc acc ggc ggc
 1321/441
 gag gtt gtc gcg gtg gga gcc gct ctg cag
 1381/461
 gtt ctg ctg ctt gat gtt acc ccg ctg agc
 1441/481
 acc agg ctc atc gag cgc aac acc acg atc
 1501/501
 gcc gac gac aac caa ccg tcg gtg cag atc
 1561/521
 gcg cac aac aag ttg ctc ggg tcc ttc gag
 1621/541
 att ccg cag atc gag gtc act ttc gac atc
 1681/561
 aag gac aag ggc acc ggc aag gag aac acg
 1741/581
 aag gaa gac att gac cgc atg atc aag gac
 1801/601
 cgt cgc gag gag gcc gat gtt cgt aat caa
 1861/621
 ttc gtc aaa gaa cag cgt gag gcc gag ggt
 1921/641
 aag gtt gat gcc gcg gtg gcg gaa gcg aag
 1981/661
 atc aag tcg gcg atg gag aag ctg ggc cag
 2041/681
 gaa gca gct cag gct gcg tca cag gcc act
 2101/701
 AGC a

1111/371
 ctg acc cgc gcg gag ttc caa ccg atc act
 1171/391
 ttc cag tcg gtg atc gct gac acc ggc att
 1231/411
 gtg ggt ggt tcg acc cgg atg ccc gcg gtg
 1291/431
 aag gaa ccc aac aag ggc gtc aac ccc gat
 1351/451
 gcc ggc gtc ctc aag ggc gag gtg aaa gac
 1411/471
 ctg ggt atc gag acc aag ggc ggg gtg atg
 1471/491
 ccc acc aag cgg tcg gag act ttc acc acc
 1531/511
 cag gtc tat cag ggg gag cgt gag atc gcc
 1591/531
 ctg acc ggc atc ccg ccg gcg ccg cgg ggg
 1651/551
 gac gcc aac ggc att gtg cac gtc acc gcc
 1711/571
 atc cga atc cag gaa ggc tcg ggc ctg tcc
 1771/591
 gcc gaa gcg cac gcc gag gag gat cgc aag
 1831/611
 gcc gag aca ttg gtc tac cag acg gag aag
 1891/631
 ggt tcg aag gta cct gaa gac acg ctg aac
 1951/651
 gcg gca ctt ggc gga tcg gat att tcg gcc
 2011/671
 gag tcg cag gct ctg ggg caa gcg atc tac
 2071/691
 ggc gct gcc cac ccc ggc tcg gct gat gaA

SEQ ID NO:16
 1/1
 Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr
 61/21
 Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly
 121/41
 Pro Ala Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys
 181/61
 Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
 241/81
 Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln Gly Ser Met ala
 301/101
 Arg Ala Val Gly Ile Asp Leu Gly Thr Thr Asn Ser Val Val Ser Val Leu Glu Gly Gly
 361/121
 Asp Pro Val Val Val Ala Asn Ser Glu Gly Ser Arg Thr Thr Pro Ser Ile Val Ala Phe
 421/141
 Ala Arg Asn Gly Glu Val Leu Val Gly Gln Pro Ala Lys Asn Gln Ala Val Thr Asn Val
 481/161
 Asp Arg Thr Val Arg Ser Val Lys Arg His Met Gly Ser Asp Trp Ser Ile Glu Ile Asp
 541/181
 Gly Lys Lys Tyr Thr Ala Pro Glu Ile Ser Ala Arg Ile Leu Met Lys Leu Lys Arg Asp
 601/201
 Ala Glu Ala Tyr Leu Gly Glu Asp Ile Thr Asp Ala Val Ile Thr Thr Pro Ala Tyr Phe
 661/221
 Asn Asp Ala Gln Arg Gln Ala Thr Lys Asp Ala Gly Gln Ile Ala Gly Leu Asn Val Leu
 721/241
 Arg Ile Val Asn Glu Pro Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Gly Glu Lys
 781/261
 Glu Gln Arg Ile Leu Val Phe Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Leu Leu Glu
 841/281
 Ile Gly Glu Gly Val Val Glu Val Arg Ala Thr Ser Gly Asp Asn His Leu Gly Gly Asp
 901/301
 Asp Trp Asp Gln Arg Val Val Asp Trp Leu Val Asp Lys Phe Lys Gly Thr Ser Gly Ile
 961/321
 Asp Leu Thr Lys Asp Lys Met ala Met Gln Arg Leu Arg Glu Ala Ala Glu Lys Ala Lys
 1021/341
 Ile Glu Leu Ser Ser Ser Gln Ser Thr Ser Ile Asn Leu Pro Tyr Ile Thr Val Asp Ala
 1081/361
 Asp Lys Asn Pro Leu Phe Leu Asp Glu Gln Leu Thr Arg Ala Glu Phe Gln Arg Ile Thr
 1141/381
 Gln Asp Leu Leu Asp Arg Thr Arg Lys Pro Phe Gln Ser Val Ile Ala Asp Thr Gly Ile
 1201/401
 Ser Val Ser Glu Ile Asp His Val Val Leu Val Gly Gly Ser Thr Arg Met Pro Ala Val

1261/421 1291/431
 Thr Asp Leu Val Lys Glu Leu Thr Gly Gly Lys Glu Pro Asn Lys Gly Val Asn Pro Asp
 1321/441 1351/451
 Glu Val Val Ala Val Gly Ala Ala Leu Gln Ala Gly Val Leu Lys Gly Glu Val Lys Asp
 1381/461 1411/471
 Val Leu Leu Leu Asp Val Thr Pro Leu Ser Leu Gly Ile Glu Thr Lys Gly Gly Val Met
 1441/481 1471/491
 Thr Arg Leu Ile Glu Arg Asn Thr Thr Ile Pro Thr Lys Arg Ser Glu Thr Phe Thr Thr
 1501/501 1531/511
 Ala Asp Asp Asn Gln Pro Ser Val Gln Ile Gln Val Tyr Gln Gly Glu Arg Glu Ile Ala
 1561/521 1591/531
 Ala His Asn Lys Leu Leu Gly Ser Phe Glu Leu Thr Gly Ile Pro Pro Ala Pro Arg Gly
 1621/541 1651/551
 Ile Pro Gln Ile Glu Val Thr Phe Asp Ile Asp Ala Asn Gly Ile Val His Val Thr Ala
 1681/561 1711/571
 Lys Asp Lys Gly Thr Gly Lys Glu Asn Thr Ile Arg Ile Gln Glu Gly Ser Gly Leu Ser
 1741/581 1771/591
 Lys Glu Asp Ile Asp Arg Met Ile Lys Asp Ala Glu Ala His Ala Glu Glu Asp Arg Lys
 1801/601 1831/611
 Arg Arg Glu Glu Ala Asp Val Arg Asn Gln Ala Glu Thr Leu Val Tyr Gln Thr Glu Lys
 1861/621 1891/631
 Phe Val Lys Glu Gln Arg Glu Ala Glu Gly Gly Ser Lys Val Pro Glu Asp Thr Leu Asn
 1921/641 1951/651
 Lys Val Asp Ala Ala Val Ala Glu Ala Lys Ala Ala Leu Gly Gly Ser Asp Ile Ser Ala
 1981/661 2011/671
 Ile Lys Ser Ala Met Glu Lys Leu Gly Gln Glu Ser Gln Ala Leu Gly Gln Ala Ile Tyr
 2041/681 2071/691
 GLU ALA ALA GLN ALA ALA SER GLN ALA THR GLY ALA ALA HIS PRO GLY SER ALA ASP GLU
 2101/701
 Ser

SEQ ID NO:17

ctgcagctgg tcaggccgtt tccgcaacgc ttgaagtccct ggccgatata ccggcagggc cagccatcgt tcgacgaata
 aagccaccctc agccatgatg ccctttccat ccccgagcga accccgacat ggacgcaaaa gccctgctcc tcggcagcct
 ctgcctggcc gccccattcg ccgacgcggc gacgctcgac aatgctctct ccgctgctcc cgccgcccgg ctcggtgcac
 cgacacggc ggagggccag ttgcacctgc cactcaccct tgaggcccgg cgctccaccg gcgaatgcgg ctgtacctcg
 gcgctggtgc gatatcggt gctggccagg ggcgccagg ccgacagcct cgtgcttcaa gagggctgct cgatagtcgc
 caggacacgc cgcgcacgt gaccctggcg gcggagcggc gcttgccgag ccggcccgaa ctggtcgtca cctggggttg
 tcaggcgctt gactgacagg ccgggctgcc accaccaggc ccgatgggac gccctgcatg tatctctcca tcggcaagcc
 tcccgcttgc acattcacca ctctgcaatc cagttcataa atcccataaa agccctcttc cgctccccgc cagcctcccc
 gcaccccgca ccctagacgc ccgcgcgctc tccgcgcggt cgcgccgacaa gaaaaaccaa ccgctcgatc agcctcatcc
 ttacccatc acaggagcca tcgcatgcca cctgataccc cattggatcc ccttggtcgc cagcctcggc ctgctgcgcg
 gcgctcgtc cgcgtccgcc gccgaggaag ccttcgacct ctggaacgaa tgccgcaaaag cctgctgct cgacctcaag
 gacggcgtgc gttccagccg catgagcgtc gacccggcca tcgcccagac caacggccag ggctgctgc actactccat
 ggtcctggag ggccgcaag acgcgctcaa gctggccatc gacacgccc tcagcatcac cagcgacggc ctgacctacc
 gctcgaagg cgcgctcgag ccgaacaagc ccgctgcgcta cgcatacacg ccgagggcgc gcggcagttg gtcgctgaac
 tggctggtac ccatcgccca cgagaagccc tcgaacatca aggtgttcat ccacgaactg aacgcccggca accagctcag
 ccacatgtcg ccgatctaca ccatcgagat ggccgagcag ttgctggcga agctggcgcg cgatgccacc ttcttcgtca
 gggcgccaga gagcaacgag atgcagccga cgctcgccat cagccatgcc ggggtcagcg tggctcatggc ccagaccag
 ccgcgcggcg aaaagcgtg gagcgaatgg gccagcgcca aggtgttgc cctgctcgac ccgctggagc gggctacaa
 ctacctcgcc cagcaacgct gcaacctcga cgatacctgg gaaggcaaga tctaccgggt gctcgcggc aaccggcgca
 agcatgacct gacatcaaaa cccacggtca tcagtcacgt cctgcacttt cccgagggcg gcagcctggc cgcgctgacc
 gcgcaccagg cttgccacct gccgctggag actttcaccc gtcatacgcca gccgcgcggc tgggaacaac tggagcagtg
 cggctatccg gtgcagcggc tggctgcctt ctacctggcg gcgcggctgt cgtggaacca ggtcgaccag gtgatccgca
 acgcccctggc cagccccggc agcggcgcg acctggcgga agcgcacccg gagcagcccg agcagggccc tctggccctg
 accctggccg ccgcccagag cgagcgcttc gtccggcagg gcccggcga gcagcaggcc ggccggcgca acgcccagct
 ggtgagcctg acctgcccgg tcgcccggcg tgaatgcgcg ggccggcgcg acagcgcgca cgcctgctg gagcgcaact
 atcccaactg ccgagagttc ctccggcagc gcggcgacgt gacgcggct atgtgttcgt cggctaccac ggcaccttcc tcgaagcggc
 cggctgctcc aggcgcaccg ccaactggag gacgcggct gcgcgcagc gacctcgag cgatctggcg cggtttctat atcgccggc
 gcaaagcatc gtcttcggcg ggtgcgcg cgcctcgagc accaggaacc cgacgcagc ggccggatcc gcaacggtgc cctgctgcgg
 atccggcgct ggccacggc tacgcccagg accaggaacc ttctaccgca ccagcctgac cctggccggc ccgagggcg cggcgaggt
 gtctatgtgc cgcgctcgag cctgcccggc ttctaccgca ccagcctgac cctggccggc gcccagga ggaaggcggg cgctggaga
 ccaattctcg ctggccgctg gccgagcgca ccgtggtgat tccctcgcg atccccaccg acccgcgcaa cgtcgcgggc
 gacctcgacc cgtccagcat ccccgacaag gaacaggcga tcagcgcct gccggactac gccagccagc ccggcaaac
 gccgcggag gacctgaagt aactgccgg accggcggc tcccttcgca ggagccggc ttctcgggg ctgccatac
 atcaggtttt cctgatgcca gcccaatcga atatgaattc 2760

SEQ ID NO:18

MHLIPHWIPL VASLGLLAGG SSASAAEEAF DLWNECAKAC VLDLKDGVRS SRMSVDPAlA DTNGQGVLYH SMVLEGGNDA
 LKLAIDNALS ITSDGLTIRL EGGVEPNKPV RYSYTRQARG SWSLNLVPI GHEKPSNIKV FIHELNAGNQ LSHMSPIYTI
 EMGDLLAKL ARDATFFVRA HESNEMQPTL AISHAGSVV MAQTQPRREK RWSEWASGKV LCLLDPLDGV YNYLAQQRCN
 LDDTWEGKIY RVLAGNPAKH DLDIKPTVIS HRLHFPEGGS LAALTAHQAC HLPLETFTRH RQPRGWEQLE QCGYPVQRLV
 ALYLAARLSW NQVDQVIRNA LASPGSGGDL GEAIREQPEQ ARLALTAAA ESERFVRQGT GNDEAGAANA DVVSLTCPVA
 AGECAGPADS GDALLERNYP TGAEFGLDGG DVSFSTRGTQ NWTVERLLQA HRQLEERGYV FVGYHGTFFLE AAQSIVFGGV
 RARSQDLDAI WRGFYIAGDP ALAYGYAQDQ EPDARGRIRN GALLRVYVPR SSLPGFYRTS LTLAAPEAAG EVERLIGHPL
 PLRLDAITGP EEEGRLETI LGWPLAERTV VIPSAIPTDP RNVGGDLDPs SIPDKEQAlS ALPDYASQPG KPPREDLK

638

SEQ ID NO:19

RLHFPEGGSL AALTAHQACH LPLETFTRHR QPRGWEQLEQ CGYPVQRLVA LYLAARLSWN QVDQVIRNAL ASPGSGGDLG
EAIREQPEQA RLALTAAAE SERFVRQGTG NDEAGAAANAD VVSLTCPVAA GECAGPADSG DALLERNYPT GAFLGDDGD
VSFSTRGTQN W 171

SEQ ID NO:20

1/1 31/11
atg cgc ctg cac ttt ccc gag ggc ggc agc ctg gcc gcg ctg acc gcg cac cag gct tgc
61/21 91/31
cac ctg ccg ctg gag act ttc acc cgt cat cgc cag ccg cgc ggc tgg gaa caa ctg gag
121/41 151/51
cag tgc ggc tat ccg gtg cag cgg ctg gtc gcc ctc tac ctg gcg gcg cgg ctg tcg tgg
181/61 211/71
aac cag gtc gac cag gtg atc cgc aac gcc ctg gcc agc ccc ggc agc ggc ggc gac ctg
241/81 271/91
ggc gaa gcg atc cgc gag cag ccg gag cag gcc cgt ctg gcc ctg acc ctg gcc gcc gcc
301/101 331/111
gag agc gag cgc ttc gtc cgg cag ggc acc ggc aac gac gag gcc ggc gcg gcc aac gcc
361/121 391/131
gac gtg gtg agc ctg acc tgc ccg gtc gcc gcc ggt gaa tgc gcg ggc ccg gcg gac agc
421/141 451/151
ggc gac gcc ctg ctg gag cgc aac tat ccc act ggc gcg gag ttc ctc ggc gac ggc ggc
481/161 511/171
gac gtc agc ttc agc acc cgc ggc acg cag **aac gaa ttc** atg cat gga gat aca cct aca
541/181 571/191
ttg cat gaa tat atg tta gat ttg caa cca gag aca act gat ctc tac tgt tat gag caa
601/201 631/211
tta aat gac agc tca gag gag gag gat gaa ata gat ggt cca gct gga caa gca gaa ccg
661/221 691/231
gac aga gcc cat tac aat att gta acc ttt tgt tgc aag tgt gac tct acg ctt cgg ttg
721/241 751/251
tgc gta caa agc aca cac gta gac att cgt act ttg gaa gac ctg tta atg ggc aca cta
781/261 811/271
gga att gtg tgc ccc atc tgt tct caa gga tcc gag ctc ggt acc aag ctt aag ttt aaa
841/281
ccg ctg atc agc ctc gac tgt gcc ttc tag

SEQ ID NO:21

1/1 31/11
Met arg leu his phe pro glu gly gly ser leu ala ala leu thr ala his gln ala cys
61/21 91/31
His Leu Pro Leu Glu Thr Phe Thr Arg His Arg Gln Pro Arg Gly Trp Glu Gln Leu Glu
121/41 151/51
Gln Cys Gly Tyr Pro Val Gln Arg Leu Val Ala Leu Tyr Leu Ala Ala Arg Leu Ser Trp
181/61 211/71
Asn Gln Val Asp Gln Val Ile Arg Asn Ala Leu Ala Ser Pro Gly Ser Gly Gly Asp Leu
241/81 271/91
Gly Glu Ala Ile Arg Glu Gln Pro Glu Gln Ala Arg Leu Ala Leu Thr Leu Ala Ala Ala
301/101 331/111
Glu Ser Glu Arg Phe Val Arg Gln Gly Thr Gly Asn Asp Glu Ala Gly Ala Ala Asn Ala
361/121 391/131
Asp Val Val Ser Leu Thr Cys Pro Val Ala Ala Gly Glu Cys Ala Gly Pro Ala Asp Ser
421/141 451/151
Gly Asp Ala Leu Leu Glu Arg Asn Tyr Pro Thr Gly Ala Glu Phe Leu Gly Asp Gly Gly
481/161 511/171
Asp Val Ser Phe Ser Thr Arg Gly Thr Gln **Asn Glu Phe** Met His Gly Asp Thr Pro Thr
541/181 571/191
Leu His Glu Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln
601/201 631/211
Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro
661/221 691/231
Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr Leu Arg Leu
721/241 751/251
Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu
781/261 811/271
Gly Ile Val Cys Pro Ile Cys Ser Gln Gly Ser Glu Leu Gly Thr Lys Leu Lys Phe Lys
841/281
ccg ctg atc agc ctc gac tgt gcc ttc tag

SEQ ID NO:22

atg acc tct cgc cgc tcc gtg aag tcg ggt ccg cgg gag gtt ccg cgc 48
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1 5 10 15
gat gag tac gag gat ctg tac tac acc ccg tct tca ggt atg gcg agt 96
Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30
ccc gat agt ccg cct gac acc tcc cgc cgt ggc gcc cta cag aca cgc 144
Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
35 40 45
tcg cgc cag agg ggc gag gtc cgt ttc gtc cag tac gac gag tcg gat 192
Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp

50	tat gcc ctc tac ggg ggc tgc tct tcc gaa gac gac gaa cac ccg gag	240
Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu		
65	gtc ccc cgg acg cgg cgt ccc gtt tcc ggg gcg gtt ttg tcc ggc ccg	288
Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro		
85	ggg cct gcg cgg gcg cct ccg cca ccc gct ggg tcc gga ggg gcc gga	336
Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly		
100	cgc aca ccc acc acc gcc ccc cgg gcc ccc cga acc cag cgg gtg gcg	384
Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala		
115	tct aag gcc ccc gcg gcc ccg gcg gag acc acc cgc ggc agg aaa	432
Ser Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys		
130	tcg gcc cag cca gaa tcc gcc gca ctc cca gac gcc ccc gcg tcg acg	480
Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr		
145	gcg cca acc cga tcc aag aca ccc gcg cag ggg ctg gcc aga aag ctg	528
Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu		
165	cac ttt agc acc gcc ccc cca aac ccc gac gcg cca tgg acc ccc cgg	576
His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg		
180	gtg gcc ggc ttt aac aag cgc gtc ttc tgc gcc gcg gtc ggg cgc ctg	624
Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu		
195	gcg gcc atg cat gcc cgg atg gcg gct gtc cag ctc tgg gac atg tcg	672
Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser		
210	cgt ccg cgc aca gac gaa gac ctc aac gaa ctc ctt ggc atc acc acc	720
Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr		
225	atc cgc gtg acg gtc tgc gag ggc aaa aac ctg ctt cag cgc gcc aac	768
Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn		
245	gag ttg gtg aat cca gac gtg gtg cag gac gtc gac gcg gcc acg gcg	816
Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala		
260	act cga ggg cgt tct gcg gcg tcg cgc ccc acc gag cga cct cga gcc	864
Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala		
275	cca gcc cgc tcc gct tct cgc ccc aga cgg ccc gtc gag ggt acc gag	912
Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu		
290	ctc gga tcc atg cat gga gat aca cct aca ttg cat gaa tat atg tta	960
Leu Gly Ser Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu		
305	gat ttg caa cca gag aca act gat ctc tac tgt tat gag caa tta aat	1008
Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn		
325	gac agc tca gag gag gag gat gaa ata gat ggt cca gct gga caa gca	1056
Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala		
340	gaa ccg gac aga gcc cat tac aat att gta acc ttt tgt tgc aag tgt	1104
Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys		
355	gac tct acg ctt cgg ttg tgc gta caa agc aca cac gta gac att cgt	1152
Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg		
370	act ttg gaa gac ctg tta atg ggc aca cta gga att gtg tgc ccc atc	1200
Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile		
385	tgt tct cag gat aag ctt aag ttt aaa ccg ctg atc agc ctc gac tgt	1248
Cys Ser Gln Asp Lys Leu Lys Phe Lys Pro Leu Ile Ser Leu Asp Cys		
405	gcc ttc tag	1257
Ala Phe		

SEQ ID NO:23

1 **atg**ctgctat ccgtgccgct gctgctcggc ctccctcgcc tggccgtcgc cgagcccgcc
 61 gtctacttca aggagcagtt tctggacgga gacgggtgga cttcccgctg gatcgaatcc
 121 aaacacaagt cagattttgg caaattcgtt ctcagttccg gcaagttcta cgggtgacgag
 181 gagaagata aaggttttga gacaagccag gatgcacgct tttatgctct gtcggccagt
 241 ttcgagcctt tcagcaacaa aggccagacg ctggtgtgtgc agttcacggt gaaacatgag
 301 cagaacatcg actgtggggg cggtatgtg aagctgtttc ctaatagttt ggaccagaca
 361 gacatgcacg gagactcaga atacaacatc atgtttggtc ccgacatctg tggccctggc
 421 accaagaagg ttcatgtcat cttcaactac aagggaaga acgtgctgat caacaaggac
 481 atcgtttgca aggatgatga gtttacacac ctgtacacac tgattgtgag gccagacaa

541 acctatgagg tgaagattga caacagccag gtggagtccg gctccttgga agacgattgg
 601 gacttcctgc caccacaagaa gataaaggat cctgatgctt caaaaccgga agactgggat
 661 gagcgggcca agatcgatga tcccacagac tccaagcctg aggactggga caagcccagag
 721 catatccctg accctgatgc taagaagccc gaggactggg atgaagagat ggacggagag
 781 tgggaacccc cagtgtattca gaaccctgag tacaaggggtg agtgggaagcc ccggcagatc
 841 gacaacccag attacaaggg cacttggatc caccacagaaa ttgacaaccc cgagtattct
 901 cccgatccca gtatctatgc ctatgataac tttggcgtgc tgggcctgga cctctggcag
 961 gtcaagtctg gcaccatctt tgacaacttc ctcatcacca acgatgaggc atacgctgag
 1021 gagtttgga acgagacgtg gggcgtaaca aaggcagcag agaaacaaat gaaggacaaa
 1081 caggacgagg agcagaggct taaggaggag gaagaagaca agaaacgcaa agaggaggag
 1141 gaggcagagg acaaggagga tgatgaggac aaagatgagg atgaggagga tgaggaggac
 1201 aaggaggaag atgaggagga agatgtcccc ggccaggcca aggacgagct gtag 1251

SEQ ID NO:24

1 MLLSVPLLLG LLGLAVAEPA VYFKEQFLDG DGWTSRWIES KHKSDFGKFV LSSGKFYGD
 61 EKDKGLQTSQ DARFYALSAS FEPFSNKGQT LVVQFTVKHE QNIDCGGGYV KLFPSNLDQT
 121 DMHGDSEYNI MFGPDICGPG TKKVHVIFNY KGKVLINKD IRCKDDEFTH LYTLIVRPDN
 181 TYEVKIDNSQ VESGSLEDDW DFLPPKKIKD PDASKPEDWD ERAKIDDPD SKPEDWDKPE
 241 HIPDPDAKKP EDWDEEMDGE WEPPVIQNP EYKGEWKPRQI DNPDYKGTWI HPEIDNPEYS
 301 DDPSIYAYDN FGVLGLDLWQ VKSGTIFDNF LIITNDEAYAE EFGNETWGV KAAEKQMKDK
 361 QDEEQLKEE EEDKKRKEE EAEDKEDDED KDEDEEDED KEEDEEDVP GQAKDEL 417

SEQ ID NO:25

1 MLLSVPLLLG LLGLAVAEPA VYFKEQFLDG DGWTSRWIES KHKSDFGKFV LSSGKFYGD
 61 EKDKGLQTSQ DARFYALSAS FEPFSNKGQT LVVQFTVKHE QNIDCGGGYV KLFPSNLDQT
 121 DMHGDSEYNI MFGPDICGPG TKKVHVIFNY KGKVLINKD IRCKDDEFTH 170

SEQ ID NO:26

1 LYTLIVRPDN TYEVKIDNSQ VESGSLEDDW DFLPPKKIKD PDASKPEDWD ERAKIDDPD
 61 SKPEDWDKPE HIPDPDAKKP EDWDEEMDGE WEPPVIQNP EYKGEWKPRQ 109

SEQ ID NO:27

1 IDNPDKGTW IHPEIDNPEY SPDPISYAYD NFGVLGLDLW QVKSGETIFDN FLITNDEAYA
 61 EEFNETWGV TKAAEKQMKD QDEEQLKEE EEDKKRKEE EEAEDKEDDE DKDEDEEED
 121 DKEDEEEDV PGQAKDEL 138

SEQ ID NO:28

1 **ATGCTGCTAT** CCGTGCCGCT GCTGCTCGGC CTCCTCGGCC TGGCCGTCGC CGAGCCCGCC
 61 GTCTACTTCA AGGAGCAGTT TCTGGACGGA GACGGGTGGA CTTCCTCGCTG GATCGAATCC
 121 AAACACAAGT CAGATTTTGG CAAATTCGTT CTCAGTTCCG GCAAGTTCTA CGGTGACGAG
 181 GAGAAAGATA AAGGTTTGCA GACAAGCCAG GATGCACGCT TTTATGCTCT GTCGGCCAGT
 241 TTCGAGCCTT TCAGCAACAA AGGCCAGACG CTGGTGCTGC AGTTTACGGT GAAACATGAG
 301 CAGAACATCG ACTGTGGGGG CGGCTATGTG AAGCTGTTTC CTAATAGTTT GGACCAGACA
 361 GACATGCACG GAGACTCAGA ATACAACATC ATGTTTGGTC CCGACATCTG TGGCCCTGGC
 421 ACCAAGAAGG TTCATGTCTAT CTCAACTAC AAGGGCAAGA ACGTGCTGAT CAACAAGGAC
 481 ATCCGTTGCA AGGATGATGA GTTTACACAC CTGTACACAC TGATTGTGCG GCCAGACAAC
 541 acctatgagg tgaagattga caacagccag gtggagtccg gctccttgga agacgattgg
 601 gacttcctgc caccacaagaa gataaaggat cctgatgctt caaaaccgga agactgggat
 661 gagcgggcca agatcgatga tcccacagac tccaagcctg aggactggga caagcccagag
 721 catatccctg accctgatgc taagaagccc gaggactggg atgaagagat ggacggagag
 781 tgggaacccc cagtgtattca gaaccctgag tacaaggggtg agtgggaagcc ccggcagatc
 841 gacaacccag attacaaggg cacttggatc caccacagaaa ttgacaaccc cgagtattct
 901 cccgatccca gtatctatgc ctatgataac tttggcgtgc tgggcctgga cctctggcag
 961 gtcaagtctg gcaccatctt tgacaacttc ctcatcacca acgatgaggc atacgctgag
 1021 gagtttgga acgagacgtg gggcgtaaca aaggcagcag agaaacaaat gaaggacaaa
 1081 caggacgagg agcagaggct taaggaggag gaagaagaca agaaacgcaa agaggaggag
 1141 gaggcagagg acaaggagga tgatgaggac aaagatgagg atgaggagga tgaggaggac
 1201 aaggaggaag atgaggagga agatgtcccc ggccaggcca aggacgagct gtag 1251

SEQ ID NO:29

1 **ATGCTGCTAT** CCGTGCCGCT GCTGCTCGGC CTCCTCGGCC TGGCCGTCGC CGAGCCCGCC
 61 GTCTACTTCA AGGAGCAGTT TCTGGACGGA GACGGGTGGA CTTCCTCGCTG GATCGAATCC
 121 AAACACAAGT CAGATTTTGG CAAATTCGTT CTCAGTTCCG GCAAGTTCTA CGGTGACGAG
 181 GAGAAAGATA AAGGTTTGCA GACAAGCCAG GATGCACGCT TTTATGCTCT GTCGGCCAGT
 241 TTCGAGCCTT TCAGCAACAA AGGCCAGACG CTGGTGCTGC AGTTTACGGT GAAACATGAG
 301 CAGAACATCG ACTGTGGGGG CGGCTATGTG AAGCTGTTTC CTAATAGTTT GGACCAGACA
 361 GACATGCACG GAGACTCAGA ATACAACATC ATGTTTGGTC CCGACATCTG TGGCCCTGGC
 421 ACCAAGAAGG TTCATGTCTAT CTCAACTAC AAGGGCAAGA ACGTGCTGAT CAACAAGGAC
 481 ATCCGTTGCA AGGATGATGA GTTTACACAC CTGTACACAC TGATTGTGCG GCCAGACAAC

SEQ ID NO:30

1 acctatgagg tgaagattga caacagccag gtggagtccg gctccttgga agacgattgg
 61 gacttcctgc caccacaagaa gataaaggat cctgatgctt caaaaccgga agactgggat
 121 gagcgggcca agatcgatga tcccacagac tccaagcctg aggactggga caagcccagag
 181 catatccctg accctgatgc taagaagccc gaggactggg atgaagagat ggacggagag
 241 tgggaacccc cagtgtattca gaaccct 267

SEQ ID NO:31

1 gagtacaagg gtgagtggaa gccccggcag atcgacaacc cagattacaa gggcacttgg
 61 atccaccag aaattgacaa ccccgagtat tctcccgatc ccagtatcta tgcctatgat

121 aactttggcgt tgctgggcct ggacctctgg cagggtcaagt ctggcaccat ctttgacaa
 181 ttccctcatca ccaacgatga ggcatacgtc gaggagtttg gcaacgagac gtggggcgta
 241 acaaaaggcag cagagaaaca aatgaaggac aaacaggacg aggagcagag gcttaaggag
 301 gaggaagaag acaagaaacg caaagaggag gaggaggcag aggacaagga ggatgatgag
 361 gacaaagatg aggatgagga ggatgaggag gacaaggagg aagatgagga ggaagatgct
 421 cccggccagg ccaaggacga gctg **444**

SEQ ID NO:32

1 gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc
 61 gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcggtc gctctcctgt
 121 tccgacctgt ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gctggcgct
 181 ttctcatagc tcacgctgta ggtatctcag ttccggtgtg gtcgttcgct ccaagctggg
 241 ctgtgtgcac gaaccccccg ttcagcccca ccgctgcgcc ttatccggta actatcgtct
 301 tgagtccaac ccggttaagc acgacttacc gccactggca gcagccactg gtaacaggat
 361 tagcagagcg aggtatgtag cgggtgctac agagttcttg aagtgggtgc ctaactacgg
 421 ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa
 481 aagagttggg agctcttgat ccggcaaaaca aaccacgcgt ggtagcgggt gtttttttgt
 541 ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc
 601 tacgggggtc gacgctcagt ggaacgaaaa ctacagttaa gggatttttg tcatgagatt
 661 atcaaaaaag atcttccact agatcctttt aaatttaaaa tgaagtttta aatcaatcta
 721 aagtataatc gactaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat
 781 ctgagcgatc tgtctatttc gttcatccat agttgcctga ctccgggggg gggggcgctg
 841 aggtctgcct cgtgaagaag gtgttgctga ctcataccag ggcaacgttg ttgccattgc
 901 tacaggcaag ctggtgtcac gctcgtcgtt tgggtatggc tcattcagct ccggttccca
 961 acgatacaag cgagttacat gatcccccat gttgtgcaaa aaagcgggta gctccttcgg
 1021 tctccgatac gttgtcagaa gtaagttggc cgcagtgtta tcaactcatg ttatggcagc
 1081 actgcataat tctcttactg tcatgccatc cgttaagatgc ttttctgtga ctggtgagta
 1141 ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc
 1201 aatcagggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttgaaaaacg
 1261 ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgtatgaacc
 1321 cactcgtgca cctgaatcgc ccatcatcct agccagaaaag tgaggagacc acggttgatg
 1381 agagctttgt tgtagggtga ccagttggtg attttgaaat tttgctttgc cacggaacgg
 1441 tctgctgtgt cgggaagatg cgtgatctga tcttcaact cagcaaaagt tcgatttatt
 1501 caacaaagcc gccgtccgtc caagtcagcg taatgctctg ccagtggtac aaccaattaa
 1561 ccaattctga ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag
 1621 gattatcaat accatatatt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccca
 1681 ggcagttcca taggatggca agatcctggt atcgggtctgc gattccgact cgtccaacat
 1741 caatacaacc tatataattt cctcgtcaa aaataagggt atcaagttag aaatcaccat
 1801 gaitgacgac tgaatccggt gagaatggca aaagcttatg catttctttc cagacttggt
 1861 caacaggcca gccattacgc tcgtcatcaa aatcactcgc atcaacccaa ccgttattca
 1921 ttgctgattg ccctgagcgc agacgaaata cgcgacgctc gttaaaagga caattacaaa
 1981 caggaatcga atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg
 2041 aatcaggata ttcttcta atcctggaatg ctgttttccc ggggatcgca gtggtgagta
 2101 accatgcata atcaggagta cggataaaat gcttgatggt cgggaagaggc ataaattccg
 2161 tcagccagtt tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat
 2221 gtttcagaaa caactctggc gcatcgggct tcccatacaa tcgatagatt gtcgcacctg
 2281 attgcccagc attatcgcga gccattttat acccatataa atcagcatcc atgttggat
 2341 ttaatcggcg cctcgagcaa gacgtttccc gttgaatatg gctcataaca ccccttgat
 2401 tactgtttat gtaagcagac agttttattg ttcatgatga tatattttta tcttgtagca
 2461 tgtaacatca gagattttga gacacaacgt ggttttccc cccccccat tattgaagca
 2521 tttatcaggg ttattgtctc atgagcggat acatatattga atgtatttag aaaaaataac
 2581 aaatagggtt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta
 2641 ttatcatagc catataatg aaaaataggc gtatcacgag gccctttcgt ctcgcgctt
 2701 tcgggtgatg cgggtgaaaa ctctgacaca tgcagctccc ggagacgggt acagctgtc
 2761 tgtaaagcga tgcggggagc agacaagccc gtcaggcgcc gtcagcgggt gttggcgggt
 2821 gtcggggctg gcttaactat gcggcatcag agcagattgt actgagagt caccatagc
 2881 ggtgtgaaat accgcacaga tgcgttaagg gaaaataacc catcagattg gctattggcc
 2941 attgcatacg ttgtatccat atcataatat gtacatttat attggtcat gtcacacatt
 3001 accgcatagt tgacattgat tattgactag ttatataatg taatcaatta cggggtcatt
 3061 agttcatagc catatatgg agttccgctg tacataactt acggtaaatg gccgccttg
 3121 ctgaccgccc aacgaccccc gccatttgac gtcaataatg acgtatgttc ccatagtaac
 3181 gccaataggg actttccatt gacgtcaatg ggtgagatg ttacggtaaa ctgcccactt
 3241 ggcagtagat caagtgtatc atatgccaag tacgccccct attgacgtca atgacggtaa
 3301 atggcccgcc tggcattatg ccagtagcat gacctatgg gactttccta cttggcagta
 3361 catctacgta ttagtcatcg ctattaccat ggtgatgcgg ttttgccagt acatcaatgg
 3421 gcgtggatag cgggttgact cacggggatt tccaagtctc caccocattg acgtcaatgg
 3481 gagtgtttgt tggcaccaaa atcaacggga ctttccaaaa tgtcgttaaa actccgccc
 3541 attgacgcaa atggggcggt ggcgtgtacg gtgggaggtc tatataagca gagctcgtt
 3601 agtgaaccgt cagatcgctt ggagacgcca tccacgctgt tttgacctcc atagaagaca
 3661 ccgggaccca tccagctccc cgggccggga acggtgcatt ggaacgcgga ttccccgtgc
 3721 caagagtacg gtaagtaccg cctatagact ctataggcac accccttttg ctcttatgca
 3781 tgcataactg tttttggctt ggggcctata ccccccgct tctttatgct ataggtgatg
 3841 gtagatgcta gcttataggt gtgggttatt gaccattatt gaccactcca acggtggagg
 3901 gcagtgtagt cctagcagta ctcgttgctg ccgcgcgcgc caccagacat aatagctgac
 3961 agactaacag actgttcctt tccatgggtc ttttctgcag tcaccgtcgt cgacATGCTG
 4021 CTATCCGTGC CGCTGCTGCT CGGCCCTCTC GGCCTGGCCG TCGCCGAGCC TGCCGTCTAC
 4081 TTCAAGGAGC AGTTTCTGGA CGGGGACGGG TGGACTTCCC GCTGGATCGA ATCCAACAC
 4141 AAGTCAGATT TTGGCAAAAT CGTTTCTCAGT TCCGGCAAGT TCTACGGTGA CAGGAGAAA
 4201 GATAAAGGTT TGCAGACAAG CCAGGATGCA CGCTTTTATG CTCTGTCGGC CAGTTTCGAG
 4261 CCTTTCAGCA ACAAGGCCA GACGCTGGTG GTGCAGTTCA CGGTGAAACA TGAGCAGAAC

4321 ATCGACTGTG GGGGCGGCTA TGTGAAGCTG TTTCCTAATA GTTTGGACCA GACAGACATG
 4381 CACGGAGACT CAGAATACAA CATCATGTTT GGTCCCGACA TCTGTGGCCC TGGCACCAG
 4441 AAGGTTTCATG TCATCTTCAA CTACAAGGGC AAGAACGTGC TGATCAACAA GGACATCCGT
 4501 TGCAAGGATG ATGAGTTTAC ACACCTGTAC ACACCTGATTG TGCGGCCAGA CAACACCTAT
 4561 GAGGTGAAGA TTGACAACAG CCAGGTGGAG TCCGGCTCCT TGGAAGACGA TTGGGACTTC
 4621 CTGCCACCCA AGAAGATAAA GGATCCTGAT GCTTCAAAC CGGAAGACTG GGATGAGCGG
 4681 GCCAAGATCG ATGATCCAC AGACTCCAAG CCTGAGGACT GGGACAAGCC CGAGCATATC
 4741 CCTGACCCTG ATGCTAAGAA GCCCGAGGAC TGGGATGAAG AGATGGACGG AGAGTGGGAA
 4801 CCCCCAGTGA TTCAGAACCC TGAGTACAAG GGTGAGTGGA AGCCCCGCA GATCGACAAC
 4861 CCAGATTACA AGGGCACTTG GATCCACCCA GAAATTGACA ACCCCGAGTA TTCTCCCGAT
 4921 CCCAGTATCT ATGCCTATGA TAACTTTGGC GTGCTGGGCC TGGACCTCTG GCAGGTCAAG
 4981 TCTGGCACCA TCTTTGACAA CTCCTCATC ACCAACGATG AGGCATACGC TGAGGAGTTT
 5041 GGCAACGAGA CTGGGGCGT AACAAAGGCA GCAGAGAAAC AAATGAAGGA CAAACAGGAC
 5101 GAGGAGCAGA GGCTTAAGGA GGAGGAAGAA GACAAGAAAC GCAAAGAGGA GGAGGAGGCA
 5161 GAGGACAAGG AGGATGATGA GGACAAGAT GAGGATGAGG AGGATGAGGA GGACAAGGGA
 5221 GAAGATGAGG AGGAAGATGT CCCCAGGAGG GCCAAGGACG AGCTGgaatt CATGCATGGA
 5281 ATGACACTTA TATGTCATGA ATATATGTTA GATTTGCAAC CAGAGACAAC TGATCTCTAC
 5341 GGTATGGGC AATTAATGA CAGCTCAGAG GAGGAGGATG AAATAGATGG TCCAGCTGGA
 5401 CAAGCAGAAC CGACAGAGC CCATTACAAT ATTGTAACT TTTGTTGCAA GTGTGACTCT
 5461 ACGCTTCGGT TGTGCGTACA AAGCACACAC GTAGACATTC GTACTTTGGA AGACCTGTTA
 5521 ATGGGCACAC TAGGAATTGT GTGCCCCATC TGTCTCAGA AACCAATAagg atccagatct
 5581 ttttccctct gccaaaaaatt atgggggacat catgaagccc cttgagcattc tgacttcttg
 5641 ctaataaaagg aaattttattt tcattgcaat agtggtgttg aattttttgt gtctctcact
 5701 cggaaggaca tatgggaggg caaatcattt aaaacatcag aatgagtatt tggtttagag
 5761 tttggcaaca tatgcccatt ctcccgcttc ctgctcact gactcgctgc gctcggtcgt
 5821 tcggtctgag cgagcggtat cagctcactc aaagcggtta atacggttat ccacagaatc
 5881 aggggataac gcaggaaga acatgtgagc aaaagggccag caaaaggcca ggaaccgtaa
 5941 aaaggccgagc ttgctgagcgt ttttccatag 5970

SEQ ID NO:33

atg acc tct cgc cgc tcc gtg aag tgc ggt ccg cgg gag gtt ccg cgc	48
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg	
1 5 10 15	
gat gag tac gag gat ctg tac tac acc ccg tct tca ggt atg gcg agt	96
Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser	
20 25 30	
ccc gat agt ccg cct gac acc tcc cgc cgt ggc gcc cta cag aca cgc	144
Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg	
35 40 45	
tcg cgc cag agg ggc gag gtc cgt ttc gtc cag tac gac gag tgc gat	192
Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp	
50 55 60	
tat gcc ctc tac ggg ggc tgc tct tcc gaa gac gac gaa cac ccg gag	240
Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu	
65 70 75 80	
gtc ccc cgg acg cgc cgt ccc gtt tcc ggg gcg gtt ttg tcc ggc ccg	288
Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro	
85 90 95	
ggg cct gcg cgg gcg cct ccg cca ccc gct ggg tcc gga ggg gcc gga	336
Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly	
100 105 110	
cgc aca ccc acc acc gcc ccc ccg gcc ccc cga acc cag ccg gtg gcg	384
Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala	
115 120 125	
tct aag gcc ccc gcg gcc ccg gcg gcg gag acc acc ccg gcc agg aaa	432
Ser Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys	
130 135 140	
tcg gcc cag cca gaa tcc gcc gca ctc cca gac gcc ccc gcg tgc acg	480
Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr	
145 150 155 160	
gcg cca acc cga tcc aag aca ccc gcg cag ggg ctg gcc aga aag ctg	528
Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu	
165 170 175	
cac ttt agc acc gcc ccc cca aac ccc gac gcg cca tgg acc ccc ccg	576
His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg	
180 185 190	
gtg gcc ggc ttt aac aag cgc gtc ttc tgc gcc gcg gtc ggg cgc ctg	624
Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu	
195 200 205	
gcg gcc atg cat gcc ccg atg gcg gct gtc cag ctc tgg gac atg tgc	672
Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser	
210 215 220	
cgt ccg cgc aca gac gaa gac ctc aac gaa ctc ctt gcc atc acc acc	720
Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr	
225 230 235 240	
atc cgc gtg acg gtc tgc gag ggc aaa aac ctg ctt cag cgc gcc aac	768
Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn	
245 250 255	
gag ttg gtg aat cca gac gtg gtg cag gac gtc gac gcg gcc acg gcg	816

Glu	Leu	Val	Asn	Pro	Asp	Val	Val	Gln	Asp	Val	Asp	Ala	Ala	Thr	Ala	
			260					265					270			
act	cga	ggg	cgt	tct	gcg	gcg	tcg	cgc	ccc	acc	gag	cga	cct	cga	gcc	864
Thr	Arg	Gly	Arg	Ser	Ala	Ala	Ser	Arg	Pro	Thr	Glu	Arg	Pro	Arg	Ala	
		275					280					285				
cca	gcc	cgc	tcc	gct	tct	cgc	ccc	aga	cgg	ccc	gtc	gag				903
Pro	Ala	Arg	Ser	Ala	Ser	Arg	Pro	Arg	Arg	Pro	Val	Glu				
	290					295					300					
SEQ ID NO:34																
atg	acc	tct	cgc	cgc	tcc	gtg	aag	tcg	ggg	ccg	cgg	gag	gtt	ccg	cgc	48
Met	Thr	Ser	Arg	Arg	Ser	Val	Lys	Ser	Gly	Pro	Arg	Glu	Val	Pro	Arg	
1			5					10					15			
gat	gag	tac	gag	gat	ctg	tac	tac	acc	ccg	tct	tca	ggg	atg	gcg	agt	96
Asp	Glu	Tyr	Glu	Asp	Leu	Tyr	Tyr	Thr	Pro	Ser	Ser	Gly	Met	Ala	Ser	
		20					25						30			
ccc	gat	agt	ccg	cct	gac	acc	tcc	cgc	cgt	ggc	gcc	cta	cag	aca	cgc	144
Pro	Asp	Ser	Pro	Pro	Asp	Thr	Ser	Arg	Arg	Gly	Ala	Leu	Gln	Thr	Arg	
		35				40					45					
tcg	cgc	cag	agg	ggc	gag	gtc	cgt	ttc	gtc	cag	tac	gac	gag	tcg	gat	192
Ser	Arg	Gln	Arg	Gly	Glu	Val	Arg	Phe	Val	Gln	Tyr	Asp	Glu	Ser	Asp	
	50				55						60					
tat	gcc	ctc	tac	ggg	ggc	tcg	tct	tcc	gaa	gac	gac	gaa	cac	ccg	gag	240
Tyr	Ala	Leu	Tyr	Gly	Gly	Ser	Ser	Ser	Glu	Asp	Asp	Glu	His	Pro	Glu	
	65			70					75					80		
gtc	ccc	cgg	acg	cgg	cgt	ccc	gtt	tcc	ggg	gcg	gtt	ttg	tcc	ggc	ccg	288
Val	Pro	Arg	Thr	Arg	Arg	Pro	Val	Ser	Gly	Ala	Val	Leu	Ser	Gly	Pro	
			85					90					95			
ggg	cct	gcg	cgg	gcg	cct	ccg	cca	ccc	gct	ggg	tcc	gga	ggg	gcc	gga	336
Gly	Pro	Ala	Arg	Ala	Pro	Pro	Pro	Pro	Ala	Gly	Ser	Gly	Gly	Ala	Gly	
		100					105					110				
cgc	aca	ccc	acc	acc	gcc	ccc	cgg	gcc	ccc	cga	acc	cag	cgg	gtg	gcg	384
Arg	Thr	Pro	Thr	Thr	Ala	Pro	Arg	Ala	Pro	Arg	Thr	Gln	Arg	Val	Ala	
		115				120						125				
tct	aag	gcc	ccc	gcg	gcc	ccg	gcg	gag	acc	acc	cgc	ggc	agg	aaa		432
Ser	Lys	Ala	Pro	Ala	Ala	Ala	Pro	Ala	Glu	Thr	Thr	Arg	Gly	Arg	Lys	
	130				135						140					
tcg	gcc	cag	cca	gaa	tcc	gcc	gca	ctc	cca	gac	gcc	ccc	gcg	tcg	acg	480
Ser	Ala	Gln	Pro	Glu	Ser	Ala	Ala	Leu	Pro	Asp	Ala	Pro	Ala	Ser	Thr	
	145			150					155					160		
gcg	cca	acc	cga	tcc	aag	aca	ccc	gcg	cag	ggg	ctg	gcc	aga	aag	ctg	528
Ala	Pro	Thr	Arg	Ser	Lys	Thr	Pro	Ala	Gln	Gly	Leu	Ala	Arg	Lys	Leu	
			165				170					175				
cac	ttt	agc	acc	gcc	ccc	cca	aac	ccc	gac	gcg	cca	tgg	acc	ccc	cgg	576
His	Phe	Ser	Thr	Ala	Pro	Pro	Asn	Pro	Asp	Ala	Pro	Trp	Thr	Pro	Arg	
		180					185					190				
gtg	gcc	ggc	ttt	aac	aag	cgc	gtc	ttc	tgc	gcc	gcg	gtc	ggg	cgc	ctg	624
Val	Ala	Gly	Phe	Asn	Lys	Arg	Val	Phe	Cys	Ala	Ala	Val	Gly	Arg	Leu	
	195					200						205				
gcg	gcc	atg	cat	gcc	cgg	atg	gct	gtc	cag	ctc	tgg	gac	atg	tcg		672
Ala	Ala	Met	His	Ala	Arg	Met	Ala	Ala	Val	Gln	Leu	Trp	Asp	Met	Ser	
	210				215						220					
cgt	ccg	cgc	aca	gac	gaa	gac	ctc	aac	gaa	ctc	ctt	ggc	atc	acc	acc	720
Arg	Pro	Arg	Thr	Asp	Glu	Asp	Leu	Asn	Glu	Leu	Leu	Gly	Ile	Thr	Thr	
	225			230					235				240			
atc	cgc	gtg	acg	gtc	tgc	gag	ggc	aaa	aac	ctg	ctt	cag	cgc	gcc	aac	768
Ile	Arg	Val	Thr	Val	Cys	Glu	Gly	Lys	Asn	Leu	Leu	Gln	Arg	Ala	Asn	
			245					250					255			
gag	ttg	gtg	aat	cca	gac	gtg	gtg	cag	gac	gtc	gac	gcg	gcc	acg	gcg	816
Glu	Leu	Val	Asn	Pro	Asp	Val	Val	Gln	Asp	Val	Asp	Ala	Ala	Thr	Ala	
		260					265					270				
act	cga	ggg	cgt	tct	gcg	gcg	tcg	cgc	ccc	acc	gag	cga	cct	cga	gcc	864
Thr	Arg	Gly	Arg	Ser	Ala	Ala	Ser	Arg	Pro	Thr	Glu	Arg	Pro	Arg	Ala	
		275					280					285				
cca	gcc	cgc	tcc	gct	tct	cgc	ccc	aga	cgg	ccc	gtc	gag	ggg	acc	gag	912
Pro	Ala	Arg	Ser	Ala	Ser	Arg	Pro	Arg	Arg	Pro	Val	Glu	Gly	Thr	Glu	
	290					295					300					
ctc	gga	tcc	atg	cat	gga	gat	aca	cct	aca	ttg	cat	gaa	tat	atg	tta	960
Leu	Gly	Ser	Met	His	Gly	Asp	Thr	Pro	Thr	Leu	His	Glu	Tyr	Met	Leu	
	305				310					315				320		
gat	ttg	caa	cca	gag	aca	act	gat	ctc	tac	tgt	tat	gag	caa	tta	aat	1008
Asp	Leu	Gln	Pro	Glu	Thr	Thr	Asp	Leu	Tyr	Cys	Tyr	Glu	Gln	Leu	Asn	
			325					330					335			
gac	agc	tca	gag	gag	gag	gat	gaa	ata	gat	ggg	cca	gct	gga	caa	gca	1056
Asp	Ser	Ser	Glu	Glu	Glu	Asp	Glu	Ile	Asp	Gly	Pro	Ala	Gly	Gln	Ala	
		340					345					350				
gaa	ccg	gac	aga	gcc	cat	tac	aat	att	gta	acc	ttt	tgt	tgc	aag	tgt	1104
Glu	Pro	Asp	Arg	Ala	His	Tyr	Asn	Ile	Val	Thr	Phe	Cys	Cys	Lys	Cys	
		355				360						365				
gac	tct	acg	ctt	cgg	ttg	tgc	gta	caa	agc	aca	cac	gta	gac	att	cgt	1152

Asp	Ser	Thr	Leu	Arg	Leu	Cys	Val	Gln	Ser	Thr	His	Val	Asp	Ile	Arg	
	370					375					380					
act	ttg	gaa	gac	ctg	tta	atg	ggc	aca	cta	gga	att	gtg	tgc	ccc	atc	1200
Thr	Leu	Glu	Asp	Leu	Leu	Met	Gly	Thr	Leu	Gly	Ile	Val	Cys	Pro	Ile	
385					390					395				400		
tgt	tct	cag	gat	aag	ctt	aag	ttt	aaa	ccg	ctg	atc	agc	ctc	gac	tgt	1248
Cys	Ser	Gln	Asp	Lys	Leu	Lys	Phe	Lys	Pro	Leu	Ile	Ser	Leu	Asp	Cys	
				405					410					415		
gcc	ttc	tag														1257
Ala	Phe															

SEQ ID NO:35

1 atg ggg gat tct gaa agg cgg aaa tcg gaa cgg cgt cgt tcc ctt gga
 48 tat ccc tct gca tat gat gac gtc tcg att cct gct cgc aga cca tca
 96 aca cgt act cag cga aat tta aac cag gat gat ttg tca aaa cat gga
 144 cca ttt acc gac cat cca aca caa aaa cat aaa tcg gcg aaa gcc gta
 192 tcg gaa gac gtt tcg tct acc acc cgg ggt ggc ttt aca aac aaa ccc
 240 cgt acc aag ccc ggg gtc aga gct gta caa agt aat aaa ttc gct ttc
 288 agt acg gct cct tca tca gca tct agc act tgg aga tca aat aca gtg
 336 gca ttt aat cag cgt atg ttt tgc gga gcg gtt gca act gtg gct caa
 384 tat cac gca tac caa ggc gcg ctc gcc ctt tgg cgt caa gat cct ccg
 432 cga aca aat gaa gaa tta gat gca ttt ctt tcc aga gct gtc att aaa
 480 att acc att caa gag ggt cca aat ttg atg ggg gaa gcc gaa acc tgt
 528 gcc cgc aaa cta ttg gaa gag tct gga tta tcc cag ggg aac gag aac
 576 gta aag tcc aaa tot gaa cgt aca acc aaa tct gaa cgt aca aga cgc
 624 ggc ggt gaa att gaa atc aaa tcg cca gat ccg gga tct cat cgt aca
 672 cat aac cct cgc act ccc gca act tcg cgt cgc cat cat tca tcc gcc
 720 cgc gga tat cgt agc agt gat agc gaa taa 747

SEQ ID NO:36

Met	Thr	Ser	Arg	Arg	Ser	Val	Lys	Ser	Gly	Pro	Arg	Glu	Val	Pro	Arg	
1				5					10					15		
Asp	Glu	Tyr	Glu	Asp	Leu	Tyr	Tyr	Thr	Pro	Ser	Ser	Gly	Met	Ala	Ser	
			20					25					30			
Pro	Asp	Ser	Pro	Pro	Asp	Thr	Ser	Arg	Arg	Gly	Ala	Leu	Gln	Thr	Arg	
			35				40					45				
Ser	Arg	Gln	Arg	Gly	Glu	Val	Arg	Phe	Val	Gln	Tyr	Asp	Glu	Ser	Asp	
			50				55				60					
Tyr	Ala	Leu	Tyr	Gly	Gly	Ser	Ser	Ser	Glu	Asp	Asp	Glu	His	Pro	Glu	
65				70					75					80		
Val	Pro	Arg	Thr	Arg	Arg	Pro	Val	Ser	Gly	Ala	Val	Leu	Ser	Gly	Pro	
				85					90					95		
Gly	Pro	Ala	Arg	Ala	Pro	Pro	Pro	Pro	Ala	Gly	Ser	Gly	Gly	Ala	Gly	
			100					105						110		
Arg	Thr	Pro	Thr	Thr	Ala	Pro	Arg	Ala	Pro	Arg	Thr	Gln	Arg	Val	Ala	
			115				120					125				
Ser	Lys	Ala	Pro	Ala	Ala	Pro	Ala	Ala	Glu	Thr	Thr	Arg	Gly	Arg	Lys	
			130				135				140					
Ser	Ala	Gln	Pro	Glu	Ser	Ala	Ala	Leu	Pro	Asp	Ala	Pro	Ala	Ser	Thr	
145				150					155					160		
Ala	Pro	Thr	Arg	Ser	Lys	Thr	Pro	Ala	Gln	Gly	Leu	Ala	Arg	Lys	Leu	
			165					170						175		
His	Phe	Ser	Thr	Ala	Pro	Pro	Asn	Pro	Asp	Ala	Pro	Trp	Thr	Pro	Arg	
			180				185					190				
Val	Ala	Gly	Phe	Asn	Lys	Arg	Val	Phe	Cys	Ala	Ala	Val	Gly	Arg	Leu	
		195					200					205				
Ala	Ala	Met	His	Ala	Arg	Met	Ala	Ala	Val	Gln	Leu	Trp	Asp	Met	Ser	
		210				215					220					
Arg	Pro	Arg	Thr	Asp	Glu	Asp	Leu	Asn	Glu	Leu	Leu	Gly	Ile	Thr	Thr	
225				230					235					240		
Ile	Arg	Val	Thr	Val	Cys	Glu	Gly	Lys	Asn	Leu	Leu	Gln	Arg	Ala	Asn	
			245					250						255		
Glu	Leu	Val	Asn	Pro	Asp	Val	Val	Gln	Asp	Val	Asp	Ala	Ala	Thr	Ala	
			260				265					270				
Thr	Arg	Gly	Arg	Ser	Ala	Ala	Ser	Arg	Pro	Thr	Glu	Arg	Pro	Arg	Ala	
		275					280					285				
Pro	Ala	Arg	Ser	Ala	Ser	Arg	Pro	Arg	Arg	Pro	Val	Glu				
	290					295					300					

SEQ ID NO:37

Met	Thr	Ser	Arg	Arg	Ser	Val	Lys	Ser	Gly	Pro	Arg	Glu	Val	Pro	Arg	
1				5					10					15		
Asp	Glu	Tyr	Glu	Asp	Leu	Tyr	Tyr	Thr	Pro	Ser	Ser	Gly	Met	Ala	Ser	
			20					25					30			
Pro	Asp	Ser	Pro	Pro	Asp	Thr	Ser	Arg	Arg	Gly	Ala	Leu	Gln	Thr	Arg	
			35				40					45				
Ser	Arg	Gln	Arg	Gly	Glu	Val	Arg	Phe	Val	Gln	Tyr	Asp	Glu	Ser	Asp	
			50				55				60					
Tyr	Ala	Leu	Tyr	Gly	Gly	Ser	Ser	Ser	Glu	Asp	Asp	Glu	His	Pro	Glu	
65				70					75					80		

Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
 85 90 95
 Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
 100 105 110
 Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
 115 120 125
 Ser Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
 130 135 140
 Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
 145 150 155 160
 Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
 165 170 175
 His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
 180 185 190
 Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
 195 200 205
 Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
 210 215 220
 Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
 225 230 235 240
 Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
 245 250 255
 Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
 260 265 270
 Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
 275 280 285
 Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
 290 295 300
 Leu Gly Ser Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu
 305 310 315 320
 Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn
 325 330 335
 Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala
 340 345 350
 Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys
 355 360 365
 Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg
 370 375 380
 Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile
 385 390 395 400
 Cys Ser Gln Asp Lys Leu Lys Phe Lys Pro Leu Ile Ser Leu Asp Cys
 405 410 415
 Ala Phe

SEQ ID NO:38

2 Met Gly Asp Ser Glu Arg Arg Lys Ser Glu Arg Arg Arg Ser Leu Gly
 16 Tyr Pro Ser Ala Tyr Asp Asp Val Ser Ile Pro Ala Arg Arg Pro Ser
 32 Thr Arg Thr Gln Arg Asn Leu Asn Gln Asp Asp Leu Ser Lys His Gly
 48 Pro Phe Thr Asp His Pro Thr Gln Lys His Lys Ser Ala Lys Ala Val
 64 Ser Glu Asp Val Ser Ser Thr Thr Arg Gly Gly Phe Thr Asn Lys Pro
 80 Arg Thr Lys Pro Gly Val Arg Ala Val Gln Ser Asn Lys Phe Ala Phe
 96 Ser Thr Ala Pro Ser Ser Ala Ser Ser Thr Trp Arg Ser Asn Thr Val
 112 Ala Phe Asn Gln Arg Met Phe Cys Gly Ala Val Ala Thr Val Ala Gln
 128 Tyr His Ala Tyr Gln Gly Ala Leu Ala Leu Trp Arg Gln Asp Pro Pro
 144 Arg Thr Asn Glu Glu Leu Asp Ala Phe Leu Ser Arg Ala Val Ile Lys
 160 Ile Thr Ile Gln Glu Gly Pro Asn Leu Met Gly Glu Ala Glu Thr Cys
 176 Ala Arg Lys Leu Leu Glu Glu Ser Gly Leu Ser Gln Gly Asn Glu Asn
 192 Val Lys Ser Lys Ser Glu Arg Thr Thr Lys Ser Glu Arg Thr Arg Arg
 208 Gly Gly Glu Ile Glu Ile Lys Ser Pro Asp Pro Gly Ser His Arg Thr
 224 His Asn Pro Arg Thr Pro Ala Thr Ser Arg Arg His His Ser Ser Ala
 240 Arg Gly Tyr Arg Ser Ser Asp Ser Glu -- 249

SEQ ID NO:39

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

SEQ ID NO:40

gacggatcgg gagatctccc gatccctat ggtcgactct cagtacaatc tgctctgatg 60
 ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttaggggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cggtgacatt	240
gattattgac	tagttattaa	tagtaaatcaa	ttacgggggc	attagttcat	agcccatata	300
tgaggttcgc	cgtttacataa	cttacggttaa	atggcccgcg	tggtgaccg	cccaacgacc	360
ccgcccatt	gacgtcaata	atgacgtatg	ttcccattag	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tattttacggt	aaactgcccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgcccagta	catgacctta	tgggactttc	ctactttggca	gtacatctac	gtattagtca	600
tcgctattac	catgggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	attttcaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggcg	780
gtaggcggtg	acgggtggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaacgg	gccctctaga	ctcgagcggc	cgccactgtg	ctggatatct	gcagaattcc	960
accacactgg	actagtggat	ccgagctcgg	taccaagctt	aagttttaa	cgctgatcag	1020
cctcgactgt	gccttctagt	tgccagccat	ctggttgggt	cccctcccc	gtgccttcc	1080
tgaccttgga	aggtgccact	cccaactgtcc	tttcctaata	aaatgaggaa	attgcatcgc	1140
attgtctgag	taggtgtcat	tctattctgg	gggggtgggt	ggggcaggac	agcaaggggg	1200
aggattggga	agacaatagc	aggcatgctg	gggatgcggt	ggcctctatg	gcttctgagg	1260
cggaaagaac	cagctggggc	tctagggggt	atccccacgc	gccctgtagc	ggcgcattaa	1320
gcgcggcggg	tgtggtggtt	acgcgcagcg	tgaccgcgtac	acttgccagc	gccctagcgc	1380
ccgctccttt	cgttttcttc	ccttcctttc	tcgccacgtt	cgccggcttt	ccccgtcaag	1440
ctctaaatcg	gggcattcct	ttagggttcc	gatttagtgc	tttacggcac	ctcgacccca	1500
aaaaacttga	ttaggggtgat	ggttcacgta	gtgggcatc	gccctgatag	acggtttttc	1560
gccctttgac	gttggagtcc	acgtttctta	atagtgga	ctgtgtccaa	actggaacaa	1620
cactcaaccc	tatctcgttc	tattcttttg	atttataagg	gattttgggg	atttcggcct	1680
attggttaaa	aaatgagctg	atttaacaaa	aatttaacgc	gaattaattc	tgtggaatgt	1740
gtgtcagtta	gggtgtggaa	agtccccagg	ctccccaggc	aggcagaagt	atgcaaagca	1800
tgcatctcaa	ttagtacga	accaggtgtg	gaaagtcccc	aggctcccca	gcaggcagaa	1860
gtatgcaaa	gacatctc	aattagttag	caaccatag	ccgcgcccta	actccgcccc	1920
tcccgccct	aactccgccc	agttccgccc	attctccgcc	ccatggctga	ctaatttttt	1980
ttatttatgc	agagcccgag	gccgcctctg	cctctgagct	attccagaag	tagtgaggag	2040
gcttttttgg	aggcctaggc	ttttgcaaaa	agctcccggg	agcttgtata	tccattttcg	2100
gatctgatca	agagacagga	tgaggatcgt	ttcgcatgat	tgaacaagat	ggattgcacg	2160
caggtttctc	ggccgcttgg	gtggagaggg	tattcggtta	tgactgggca	caacagacaa	2220
tcggctgctc	tgatgccggc	gtgttcggcg	tgctcagcga	ggggcgcccc	gttctttttg	2280
tcaagaccga	acgttccggc	gccctgaatg	aactgcagga	cgaggcagcg	cggctatcgt	2340
ggctggccac	gacgggctgt	ccttgccgag	ctgtgctcga	cgttgcact	gaagcgggaa	2400
gggactggct	gctattgggc	gaagtgcggg	ggcaggatct	cctgtcatct	cacctgtctc	2460
ctgccgagaa	agtatccatc	atggctgatg	caatgcggcg	gctgcatacg	cttgatcccg	2520
ctacctgccc	attcgaccac	caagcgaaac	atcgcatcga	gcgagcacgt	actcggatgg	2580
aagccggtct	tgtcgatcag	gatgatctgg	acgaagagca	tcaggggctc	gcgccagccg	2640
aactgttcgc	caggtcctaa	gcgcgcctgc	ccgacggcga	ggatctcgtc	gtgaccatg	2700
gcgatgccctg	cttgcgcaat	atcatggtgg	aaaatggccg	cttttctgga	ttcatcgact	2760
gtggccggct	gggtgtggcg	gaccgctatc	aggacatagc	gttggtacc	cgtgatattg	2820
ctgaagagct	tggcggcgaa	tgggctgacc	gcttccctgt	gctttacggt	atcgccgctc	2880
ccgattcgca	gcgcacgcgc	ttctatcgcc	ttcttgacga	gttcttctga	gcgggactct	2940
ggggttcgaa	atgaccgacc	aagcgacgcc	caacctgcca	tcacgagatt	tcgattccac	3000
cgccgccttc	tatgaaaggt	tgggcttcgg	aatcgtttcc	cgggacgccc	gctggatgat	3060
cctccagcgc	ggggatctca	tgttgaggtt	cttcgcccac	cccaacttgt	ttattgcagc	3120
ttataatggt	tacaaataaa	gcaatagcat	cacaaatttc	acaaataaag	catttttttc	3180
actgcattct	agttgtggtt	tgtccaaact	catcaatgta	tcttatcatg	tctgtatacc	3240
gtcgacctct	agctagagct	tggcgtaate	atggctcatag	ctgtttcctg	tgtgaaattg	3300
ttatccgctc	acatttccac	acaacatacg	agccggaagc	ataaagtgtg	aagcctgggg	3360
tgccaatga	gtgagctaac	tcacattaat	tgcgttgcgc	tcactgccc	ctttccagtc	3420
gggaaacctg	tcgtgccagc	tgcattaatg	aatcgcccaa	cgccgcgggg	gaggcggttt	3480
cgctattggg	cgtctttccg	cttctctcgt	cactgactcg	ctgcgctcgg	tcgttcggct	3540
gcggcgagcg	gtatcaagtc	actcaaaagg	ggtaatacgg	ttatccacag	aatcagggga	3600
taacgcagga	aagaacatgt	gagcaaaaag	ccagcaaaaag	gccaggaaac	gtaaaaaggc	3660
cgcgttgctg	gcgtttttcc	atagggtccg	ccccctgac	gagcatcaca	aaaaatcgacg	3720
ctcaagtcag	aggtggcgaa	accgcagagg	actataaaga	taccaggcgt	ttccccctgg	3780
aagctccctc	gtgcgctctc	ctgttccgac	cctgcgcgtt	accggatacc	tgtccgcctt	3840
tctcccttcg	ggaagcgtgg	cgttttctca	atgctcacgc	tgtaggtatc	tcagttcggg	3900
gtaggtcgtt	cgtcccaagc	tgggctgtgt	gcacgaacct	ccggttcagc	ccgaccgctg	3960
cgccttatcc	ggttaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	4020
ggcagcagcc	actggttaaca	ggattagcag	agcgaggtat	gtaggcggtg	ctacagagtt	4080
cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	gtattttgga	tctgcgctct	4140
gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgtatccggca	aacaaaccac	4200
cgtcgttagc	ggtggttttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4260
tcaagaagat	cctttgatct	ttttctacggg	gtctgacgct	cagtggaaac	aaaactcacg	4320
taaagggatt	ttggtcatga	gattatcaaa	aaggatcttc	acctagatcc	tttttaatta	4380
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttggctcg	acagttacca	4440
atgcttaatc	agtgaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	4500
ctgactcccc	gtcgtgtaga	taactacgat	acgggagggc	ttaccatctg	gccccagtg	4560
tgcaatgata	ccgcgagacc	cacgctcacc	ggctccagat	ttatcagcaa	taaacaggcc	4620
agccggaagg	gccgagcgca	gaagtggctc	tgcaacttta	tccgcctcca	tccagtctat	4680
taattgttgc	cgggaagcta	gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	4740
tgcctattgc	acaggcatcg	tgggtgtcacg	ctcgtcgttt	ggtatggctt	cattcagctc	4800
cggttcccaa	cgatcaaggc	gagttacatg	atccccatg	ttgtgcaaaa	aagcggttag	4860
ctccttcggg	cctccgatcg	ttgtcagaag	taagttggcc	gcagtgttat	cactcatggg	4920
tatggcagca	ctgcataatt	ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	4980

tggtgagtag	tcaaccaagt	cattctgaga	atagtgtag	cggcgaccga	gttgctcttg	5040
cccgcgctca	ataggggata	ataccgcgcc	acatagcaga	actttaaaag	tgctcatcat	5100
tggaacacgt	tcttcggggc	gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	5160
gagtgtaacc	actcgtgcac	ccaactgac	ttcagcatct	tttactttca	ccagcgtttc	5220
tggtgagca	aaaacaggaa	ggcaaaatgc	cgcaaaaaag	ggaataagg	cgacacggaa	5280
atgttgaata	ctcatactct	tcctttttca	atattattga	agcatttatc	aggggtattg	5340
tctcatgagc	ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	5400
cacatttccc	cgaaaagtgc	cacctgacgt	c			5431

SEQ ID NO: 41

tgccattgc	atagcttgta	tccatatcat	aatatgtaca	tttatattgg	ctcatgtcca	60
acattaccgc	catgttgaca	ttgattattg	actagttatt	aatagtaatc	aattacgggg	120
tcatttagttc	atagcccata	tatggagttc	cgcgttacat	aacttacggg	aaatggccc	180
cctggctgac	cgcccaacga	ccccgcgcca	ttgacgtcaa	taatgacgta	tggtcccata	240
gtaacgccaa	tagggacttt	ccattgacgt	caatgggtgg	agtatttacg	gtaaactgcc	300
cacttggcag	tacatcaagt	gtatcatatg	ccaagtacgc	cccctattga	cgtcaatgac	360
ggtaaatggc	ccgcctggca	ttatgcccag	tacatgacct	tatgggactt	tcctacttgg	420
cagtacatct	acgtattagt	catcgctatt	accatgggtga	tgcggttttg	gcagtacatc	480
aatgggcgtg	gatagcgggt	tgactcacgg	ggatttccaa	gtctccacc	cattgacgtc	540
aatggggagt	tgttttggca	ccaaaatcaa	cgggactttc	caaaatgtcg	taacaactcc	600
gccccattga	cgcaaatggg	cggtaggcgt	gtacgggtgg	aggtctatat	aagcagagct	660
cgtttagtga	acgctcagat	cgccctggaga	cgccatccac	gctgttttga	cctccataga	720
agacaccggg	accgatccag	cctccgcggc	cggaacgggt	gcattggaac	gcggattccc	780
cgtgcccaaga	gtgacgtaag	taccgcctat	agagtcctata	ggcccccccc	cttggtctct	840
tatgcactgt	atactgtttt	tggtttgggg	tctatacacc	cccgcttctc	catgttatag	900
gtgatgggtat	agcttagcct	ataggtgtgg	gttattgacc	attattgacc	actccaacgg	960
tggaaggcag	tgtagtctga	gcagtactcg	ttgctgcgcg	gcgcgccacc	agacataata	1020
gctgacagac	taacagactg	ttccttttcca	tggttctttt	ctgcagtcac	cgctcgtcgac	1080
ggtatcgata	agcttgatat	cgaaattcacg	tgggcccggt	accgtatact	ctagagcggc	1140
cgcggtacca	gatctttttc	cctcgccaaa	aattatgggg	acatcatgaa	gcccttgag	1200
catctgactt	ctggctaata	aaggaaat	atctcattgc	aatagtgtgt	tggaat	1260
tggtctcttc	actcggaagg	acatatggga	gggcaaatca	ttttaaaccat	cagaatcagt	1320
atgtgtttta	gagtttgca	acatatgcc	ttcttccgct	tcctcgtc	ctgactcgct	1380
gcgctcggtc	gttcgggtgc	ggcgagcgg	atcagctcac	tcaaaaggcg	taatacgggt	1440
atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	1500
caggaaacgt	aaaaagccg	cgttgctggc	gtttttccat	aggtccgc	cccctgacga	1560
gcatcacaaa	aatcgacgct	caagtccag	gtggcgaaac	ccgacaggac	tataaagata	1620
ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	tgccgcttac	1680
cggtacacgt	tcgcgctt	tccttccggg	aagcgtggcg	ctttctcaat	gctcacgctg	1740
taggtatctc	agttcggtgt	aggtcggtc	ctccaaagctg	ggctgtgtgc	acgaaccccc	1800
cgttcagccc	gacgcgtcg	ccttatccg	taactatcgt	cttgagtcca	acccggtaa	1860
acacgactta	tcgccactgg	cagcagccac	tggttaacagg	attagcagag	cgaggtatgt	1920
aggcggtgct	acagagttct	tgaagtgggt	gcctaaactac	ggctacacta	gaaggacagt	1980
atgtgtatc	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	2040
atccggcaaa	caaacaccg	ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	2100
gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggt	ctgacgctca	2160
gtggaacgaa	aactcacgtt	aagggtat	ggtcatgaga	ttatcaaaaa	ggatcttcac	2220
ctagatcctt	ttaaattaaa	aatgaagt	taaatcaatc	taaagtatat	atgagtaaac	2280
ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	2340
tcgttcaccc	atagttgcct	gactccgggg	ggggggggcg	ctgaggtctg	cctcgtgaag	2400
aaggtgttgc	tgactcatac	caggggcaacg	ttggtgccat	tgctacaggc	atcgtgggtg	2460
cacgtctcgt	gttttggtatg	gcttcattca	gctccggttc	ccaacgatca	agggcagtta	2520
catgatcccc	catgtttgtg	aaaaaagccg	ttagctcctt	cggtcctcgc	atcgttgtca	2580
gaagtaagtt	ggccgcagtg	ttatcactca	tggttatggc	agcactgc	aattctctta	2640
ctgtcatgcc	atccgtaaga	tgcttttctg	tgactggtga	gtactcaacc	aagtcattct	2700
gagaatagtg	tatggcgga	ccgagttgct	cttgcccgcc	gtcaatacgg	gataataccg	2760
cgccacatag	cagaacttta	aaagtgtcca	tcattggaaa	acgttcttcg	ggcgcaaaac	2820
tctcaaggat	cttaccgctg	ttgagatcca	gttcgatgta	acccactcgt	gcacctgaat	2880
gcgcccatca	tccagccaga	aagttaggga	gccacggttg	atgagagctt	tggtgtaggt	2940
ggaccagttg	gtgattttga	acttttgc	tgccacggaa	cggtctgcgt	tgctcggaag	3000
atgctgtatc	tgatccttca	actcagcaaa	agttcgattt	attcaacaaa	gccgcgctcc	3060
cgtcaagtca	gcgtaatgct	ctgccagtg	tacaaccaat	taaccaattc	tgattagaaa	3120
aactcatcga	gcatacaatg	aaactgcaat	ttattcatat	caggattatc	aataccatat	3180
ttttgaaaaa	ccggtttctg	taatgaagga	gaaaaactcac	cgaggcagtt	ccataggatg	3240
gcaagatcct	ggtatcggtc	tgcgattccg	actcgtccaa	catcaataca	acctattaat	3300
ttccccctcg	caaaaataag	gttatcaagt	gagaaatcac	catgagtgc	gactgaatcc	3360
ggtgagaatg	gcaaaagctt	atgcatttct	ttccagactt	gttcaacagg	ccagccatta	3420
cgctcgtcat	caaaatcact	cgcatcaacc	aaaccgttat	tcattcgtga	ttgcgcctga	3480
gcgagacgaa	atagcgatc	gctgttaaaa	ggacaattac	aaacagggaat	cgaatgcaac	3540
cgccgcagga	acactgccag	cgcatcaaca	atattttcac	ctgaatcagg	atattcttct	3600
aatacctgga	atgctgtttt	cccggggac	gcagtggtga	gtaaccatgc	atcatcagga	3660
gtacggataa	aatgcttgat	ggtcggaaga	ggcataaatt	ccgtcagcca	gtttagtctg	3720
accatctcat	ctgtaacatc	attggcaacg	ctacctttgc	catgtttcag	aaacaactct	3780
ggcgcatcgg	gcttcccata	caatcgatag	attgtcgcac	ctgattgccc	gacattatcg	3840
cgagccatt	tatacccata	taaatcagca	tccatgttgg	aatttaacgc	cgccctcgag	3900
caagacgttt	ccggttgat	atggctcata	acaccccttg	tattactgtt	tatgtaagca	3960
gacagtttta	ctgttctatga	tgatataatt	ttatcttctg	caatgtaaca	tcagagattt	4020
tgagacacaa	cgtggctttc	ccccccccc	cattattgaa	gcatattatc	gggttattgt	4080
ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	4140
acatttcccc	gaaaagtgc	acctgacgtc	taagaaacca	ttattatcat	gacattaacc	4200

tataaaaaata	ggcgtatcac	gaggcccttt	cgctctcgcg	cgtttcggtg	atgacggtga	4260
aaacctctga	cacatgcagc	tcccgagac	ggtcacagct	tgtctgtaag	cggatgccgg	4320
gagcagacaa	gcccgtcagg	gcgcgtcagc	gggtgttggc	gggtgtcggg	gctggcttaa	4380
ctatgcggca	tcagagcaga	ttgtactgag	agtgcaccat	atgcggtgtg	aaataccgca	4440
cagatgcgta	aggagaaaat	accgcacag	attggctat			4479

SEQ ID NO:42
UGCCUACGAACUCUACCCdTdT

SEQ ID NO:43
GGUGAAGAGUUCGUAGGCAdTdT

SEQ ID NO:44
atggcatctggacaaggaccaggtccccgaaggtgggctgcgatgagtcctccctctgaacagcaggttgcccaggacacagagga
ggctctttcgaagctacgttttttacctccaccagcaggaaacaggagaccagggcgcccgctgccaaccccgagatggacaacttgcccc
tggacccaacagcatcttgggtcaggtgggtcggcagcttgcctcctcagagatgatattaaccggcgctacgacacagagttccagaat
ttactagaacagcttcagccacagccgggaaTGCCTACGAACCTCTTCACCaaagatcgctccagcctatttaagagtggcatcagctggg
ccgctgggtggctctcctggctttggctaccgtctggccctgtacgtctaccagcgtggtttgacgggcttcctgggcccaggtgacctgct
ttttggctgatcatctactgcacattacatcgccagatggatcgacacagagagcggttgggtggcagccctgaatttgcgtagagacccc
atctgacccgtaattggtgatttttgggtgtggtttctgttgggcccaattctggttacacagattcttcagatcatga 637

SEQ ID NO:45
TGCCTACGAACCTCTTCACC

SEQ ID NO:46
UAUGGAGCUGCAGAGGAUGdTdT

SEQ ID NO:47
CAUCCUCUGCAGCUCAUAdTdT

SEQ ID NO:48
atggacgggtccggggagcagcttgggagcggcgggccaccagctctgaacagatcatgaagacaggggcctttttgctacagggtttcat
ccaggatcgagcagggaggtatggctggggagacacctgagctgaccttggagcagccgccccaggatgcgtccaccaagaagctgagcaggt
gtctccggcgaattggagatgaaatggatagcaaTATGGAGCTGCAGAGGATGattgctgacgtggacacggactcccccgagaggtcttc
ttccgggtggcagctgacatgtttgctgatggcaacttcaactggggcccgctggttgcctcttctactttgctagcaaaactgggtgctcaa
ggcctgtgtcactaaagtgcacagctgatcagaacctatctggctggacacttgacttcctccgtgagcggctgcttgtctggtatccaaag
accaggggtggctgggaaggcctctctctacttccgggacccccacatggcagacagtgaccatctttgtggctggagtcctcaccgcctcg
ctcaccatctggaagaagatgggctga 589

SEQ ID NO:49
TATGGAGCTGCAGAGGATG

SEQ ID NO:50
atg gac ttc agc aga aat ctt tat gat att ggg gaa caa ctg gac agt gaa gat ctg gcc tcc ctc aag
ttc ctg agc ctg gac tac att ccg caa agg aag caa gaa ccc atc aag gat gcc ttg atg tta ttc cag
aga ctc cag gaa aag aga atg ttg gag gaa agc aat ctg tcc ttc ctg aag gag ctg ctc ttc cga att
aat aga ctg gat ttg ctg att acc tac cta aac act aga aag gag gag atg gaa agg gaa ctt cag aca
cca ggc agg gct caa att tct gcc tac agg ttc cac ttc tgc cgc atg agc tgg gct gaa gca aac agc
cag tgc cag aca cag tct gta cct ttc tgg cgg agg gtc gat cat cta tta ata agg gtc atg ctc tat
cag att tca gaa gaa gtg agc aga tca gaa ttg agg tct ttt aag ttt ctt ttg caa gag gaa atc tcc
aaa tgc aaa ctg gat gat gac atg aac ctg ctg gat att ttc ata gag atg gag aag agg gtc atc ctg
gga gaa gga aag ttg gac atc ctg aaa aga gtc tgt gcc caa atc aac aag agc ctg ctg aag ata atc
aac gac tat gaa gaa ttc agc aaa ggg gag gag ttg tgt ggg gta atg aca atc tgc gac tct cca aga
gaa cag gat agt gaa tca cag act ttg gac aaa gtt tac caa atg aaa agc aaa cct cgg gga tac tgt
ctg atc atc aac aat cac aat ttt gca aaa gca cgg gag aaa gtg ccc aaa ctt cac agc att agg gac
agg aat gga aca cag ttg gat gca ggg gct ttg acc acg acc ttt gaa gag ctt cat ttt gag atc aag
ccc cac gat gac tgc aca gta gag caa atc tat gag att ttg aaa atc tac caa ctc atg gac cac aag
aac atg gac tgc ttc atc tgc tgt atc ctc tcc cat gga gac aag ggc atc atc tat ggc act gat gga
cag gag gcc ccc atc tat gag ctg aca tct cag ttc act ggt ttg aag tgc cct tcc ctt gct gga aaa
ccc aaa gtg ttt ttt att cag gct tgt cag ggg gat aac tac cag aaa ggt ata cct gtt gag act gat
tca gag gag caa ccc tat tta gaa atg gat tta tca tca cct caa acg aga tat atc ccg gat gag gct
gac ttt ctg ctg ggg atg gcc act gtg aat aac tgt gtt tcc tac cga aac cct gca gag gga acc tgg
tac atc cag tca ctt tgc cag agc ctg aga gag cga tgt cct cga ggc gat gat att ctc acc atc ctg
act gaa gtg aac tat gaa gta agc aac aag gat gac aag aaa aac atg ggg aaa cag atg cct cag cct
act ttc aca cta aga aaa aaa ctt gtc ttc cct tct gat tga 1491

SEQ ID NO:51
AACCUCGGGGAUACUGUCUGAdTdT

SEQ ID NO:52
UCAGACAGUAUCCCCGAGGUUDTdT

SEQ ID NO:53
atg gac gaa gcg gat cgg cgg ctc ctg cgg cgg tgc cgg ctg cgg ctg gta gaa gag ctg cag gtg gac
cag ctc tgg gac gcc ctg ctg agc cgc gag ctg ttc agg ccc cat atg atc gag gac atc cag cgg gca
ggc tct gga tct cgg cgg gat cag gcc agg cag ctg atc ata gat ctg gag act cga ggg agt cag gct
ctt cct ttg ttc atc tcc tgc tta gag gac aca ggc cag gac atg ctg gct tgc ttt ctg cga act aac
agg caa gca aag ttg tgc aag cca acc cta gaa aac ctt acc cca gtg gtg ctc aga cca gag att
cgc aaa cca gag gtt ctc aga ccg gaa aca ccc aga cca gtg gac att ggt tct gga gga ttt ggt gat
gtc ggt gct ctt gag agt ttg agg gga aat gca gat ttg gct tac atc ctg agc atg gag ccc tgt ggc

cac tgc ctc att atc aac aat gtg aac ttc tgc cgt gag tcc ggg ctc cgc acc cgc act ggc tcc aac
 atc gac tgt gag aag ttg cgg cgt cgc ttc tcc tgc ctg cat ttc atg gtg gag gtg aag ggc gac ctg
 act gcc aag aaa atg gtg ctg gct ttg ctg gag ctg gcg cag cag gac cac ggt gct ctg gac tgc tgc
 gtg gtg gtc att ctc tct cac ggc tgt cag gcc agc cac ctg cag ttc cca ggg gct gtc tac ggc aca
 gat gga tgc cct gtg tgc gtc gag aag att gtg aac atc ttc aat ggg acc agc tgc ccc agc ctg gga
 ggg aag ccc aag ctc ttt ttc atc cag gcc tgt ggt ggg gag cag aaa gac cat ggg ttt gag gtg gcc
 tcc act tcc cct gaa gac gag tcc cct ggc agt aac ccc gag cca gat gcc acc ccg ttc cag gaa ggt
 ttg agg acc ttc gac cag ctg gac gcc ata tct agt ttg ccc aca ccc agt gac atc ttt gtg tcc tac
 tct act ttc cca ggt ttt gtt tcc tgg agg gac ccc aag agt ggc tcc tgg tac gtt gag acc ctg gac
 gac atc ttt gag cag tgg gct cac tct gaa gac ctg cag tcc ctc ctg ctt agg gtc gct aat gct gtt
 tgc gtg aaa ggg att tat aaa cag atg cct ggt tgc ttt aat ttc ctc cgg aaa aaa ctt ttc ttt aaa
 aca tca taa **1191**

SEQ ID NO:54

atg gag aac act gaa aac tca gtg gat tca aaa tcc att aaa aat ttg gaa cca aag atc ata cat gga
 agc gaa tca atg gac tct gga ata tcc ctg gac aac agt tat aaa atg gat tat cct gag atg ggt tta
 tgt ata ata att aat aat aag aat ttt cat aaa agc act gga atg aca tct cgg tct ggt aca gat gtc
 gat gca gca aac ctc agg gaa aca ttc aga aac ttg aaa tat gaa gtc agg aat aaa aat gat ctt aca
 cgt gaa gaa att gtg gaa ttg atg cgt gat gtt tct aaa gaa gat cac agc aaa agg agc agt ttt gtt
 tgt gtg ctt ctg agc cat ggt gaa gaa gga ata att ttt gga aca aat gga cct gtt gac ctg aaa aaa
 ata aca aac ttt ttc aga ggg gat cgt tgt aga agt cta act gga aaa ccc aaa ctt ttc att att cag
 gcc tgc cgt gac aca gaa ctg gac tgt ggc att gag aca gac agt ggt gtt gat gat gac atg gcg tgt
 cat aaa ata cca gtg gag gcc gac ttc ttg tat gca tac tcc aca gca cct ggt tat tat tct tgg cga
 aat tca aag gat ggc tcc tgg ttc atc cag tgc ctt tgt gcc atg ctg aaa cag tat gcc gac aag ctt
 gaa ttt atg cac att ctt acc cgg gtt aac cga aag gtg gca aca gaa ttt gag tcc ttt tcc ttt gac
 gct act ttt cat gca aag aaa cag att cca tgt att gtt tcc atg ctc aca aaa gaa ctc tat ttt tat
 cac taa **834**

SEQ ID NO:55

atggcggtacc	catacgatgt	tccagattac	gctagcttga	gatctacat	gtctcagagc	60
aaccgggagc	tggtggttga	ctttctctcc	tacaagcttt	cccagaaagg	atacagctgg	120
atgcagttta	gtgatgttga	agagaacagg	actgaggccc	cagaaggagc	tgaatcggag	180
atggagacc	ccagtgcct	caatggcaac	ccatcctggc	acctggcaga	cagcccgcg	240
gtgaatggag	ccactgcgca	cagcagcagt	ttggatgccc	gggaggtgat	ccccatggca	300
gcagtaaagc	aagcgctgag	ggaggcaggc	gacgagtgtg	aaactgcggt	ccggcgggca	360
ttcagtgacc	tgacatcca	gctccacatc	accccaggga	cagcatatca	gagctttgaa	420
caggtagtga	atgaactctt	ccgggatggg	gtaaactggg	gtcgatttgt	ggcctttttc	480
tccttcggcg	gggcactgtg	cgtggaaagc	gtagacaagg	agatgcaggt	attggtgagt	540
cggatcgcag	cttgatggc	cacttacctg	aatgaccacc	tagagccttg	gatccaggag	600
aacggcggt	gggatacttt	tgtggaactc	tatgggaaca	atgcagcagc	cgagagccga	660
aagggccagg	aacgcttcaa	ccgctgtgtc	ctgacgggca	tgactgtggc	cggcgtgggt	720
ctgctgggct	cactcttcag	tcggaaatga				750

SEQ ID NO:56

Met	Ala	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr	Ala	Ser	Leu	Arg	Ser	Thr
1				5					10					15	
Met	Ser	Gln	Ser	Asn	Arg	Glu	Leu	Val	Val	Asp	Phe	Leu	Ser	Tyr	Lys
			20					25					30		
Leu	Ser	Gln	Lys	Gly	Tyr	Ser	Trp	Ser	Gln	Phe	Ser	Asp	Val	Glu	Glu
			35				40					45			
Asn	Arg	Thr	Glu	Ala	Pro	Glu	Gly	Thr	Glu	Ser	Glu	Met	Glu	Thr	Pro
			50				55					60			
Ser	Ala	Ile	Asn	Gly	Asn	Pro	Ser	Trp	His	Leu	Ala	Asp	Ser	Pro	Ala
					70					75					80
Val	Asn	Gly	Ala	Thr	Ala	His	Ser	Ser	Ser	Leu	Asp	Ala	Arg	Glu	Val
					85					90				95	
Ile	Pro	Met	Ala	Ala	Val	Lys	Gln	Ala	Leu	Arg	Glu	Ala	Gly	Asp	Glu
					100				105				110		
Phe	Glu	Leu	Arg	Tyr	Arg	Arg	Ala	Phe	Ser	Asp	Leu	Thr	Ser	Gln	Leu
							120					125			
His	Ile	Thr	Pro	Gly	Thr	Ala	Tyr	Gln	Ser	Phe	Glu	Gln	Val	Val	Asn
						135					140				
Glu	Leu	Phe	Arg	Asp	Gly	Val	Asn	Trp	Gly	Arg	Ile	Val	Ala	Phe	Phe
						150				155					160
Ser	Phe	Gly	Gly	Ala	Leu	Cys	Val	Glu	Ser	Val	Asp	Lys	Glu	Met	Gln
						165				170				175	
Val	Leu	Val	Ser	Arg	Ile	Ala	Ala	Trp	Met	Ala	Thr	Tyr	Leu	Asn	Asp
						180			185				190		
His	Leu	Glu	Pro	Trp	Ile	Gln	Glu	Asn	Gly	Gly	Trp	Asp	Thr	Phe	Val
						200						205			
Glu	Leu	Tyr	Gly	Asn	Asn	Ala	Ala	Glu	Ser	Arg	Lys	Gly	Gln	Glu	
						215						220			
Arg	Phe	Asn	Arg	Trp	Phe	Leu	Thr	Gly	Met	Thr	Val	Ala	Gly	Val	Val
						230				235				240	
Leu	Leu	Gly	Ser	Leu	Phe	Ser	Arg	Lys							
						245									

SEQ ID NO:57

gacggatcgg	gagatctccc	gatcccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttggtgtgt	ggaggtcgct	gagtagtgcg	120

cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttaggggttag	gcgtttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cggtgacatt	240
gattatttgac	tggttatttaa	tagtaatacaa	ttacggggtc	attagttcat	agcccatata	300
tggaagttccg	ctgtacataaa	cttacggtaa	atggcccgc	tggtgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tattttacggt	aaactgccca	cttgccagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgcccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggctgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	ccattgacg	caaatgggcg	780
gtaggcgtgt	agcgtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaacgg	gcctctctaga	ctcgagcggc	cgccactgtg	ctggatatct	gcagaattcc	960
accacactgg	actagtggat	ctatggcgta	ccatacagat	gttccagatt	acgctagctt	1020
gagatctacc	atgtctcaga	gcaaccggga	gctggtggtt	gactttctct	cctacaagct	1080
ttcccagaaa	ggatcacagt	ggagtcagtt	tagtgatgtg	gaagagaaca	ggactgaggc	1140
ccagaaggcg	actgaatcgg	agatggagac	ccccagtgc	atcaatggca	acccatcctg	1200
gcacctggca	gacagccccc	cggtgaatgg	agccactcgg	cacagcagca	gtttggatgc	1260
ccgggagggtg	atccccatgg	cagcagtaaa	gcaagcgtg	agggaggcag	gcgacgagtt	1320
tgaactgcgg	taccggcggg	cattcagtg	cctgacatcc	cagctccaca	tcaccccagg	1380
gacagcatat	cgagactttg	aacaggtagt	gaatgaactc	ttccgggatg	gggtaaactg	1440
gggtcgcat	gtggcctttt	tctccttcgg	cggggcactg	tgcgtggaaa	gcgtagacaa	1500
ggagatgcag	gtattgttga	gtcggatcgc	agcttggatg	gccacttacc	tgaatgacca	1560
cctagagcct	tggttcagg	agaacggcgg	ctgggatact	tttgtggaa	tctatgggaa	1620
caatgcagca	gccagagacc	gaaagggcca	ggaacgcttc	aaacgctggt	tctgacggg	1680
catgactgtg	gocggcgctg	ttctgctggg	ctcactcttc	agtcggaaat	gaagatccga	1740
gctcgggtacc	aagcttaagt	ttaaaccgct	gatcagcctc	gactgtgcct	tctagtggcc	1800
agccatctgt	tgtttgcgcc	tcccccgtgc	cttcttgac	cctggaaggt	gccactccca	1860
ctgtcctttc	ctaataaaat	gaggaaaatg	catcgcatg	tctgagttag	tgctattcta	1920
ttctgggggg	tgggtggtgg	caggacagca	agggggagga	ttgggaagac	aatagcaggc	1980
atgctgggga	tcggtggg	tctatggctt	ctgagcgga	aagaaccagc	tggggctcta	2040
gggggtatcc	ccacgcgcgc	tgtagcggcg	cattaaagcg	ggcgggtgtg	gtgggtacgc	2100
gcagcgtgac	cgtacacatt	gccagcgccc	tagcgcgcgc	tcctttcgct	ttcttccctt	2160
cctttctcgc	cacgttcgcg	ggctttcccc	gtcaagctct	aaatcggggc	atccctttag	2220
ggttccgatt	tagtgcctta	cggcacctcg	acccccaaaa	acttgattag	ggtgatgggt	2280
cacgtagtgg	gccatcgccc	tgatagacgg	tttttcgccc	tttgacgttg	gagtcacagt	2340
tctttaatag	tggactcttg	ttccaaactg	gaacaacact	caaccctatc	tcggctctatt	2400
cttttgattt	ataaggattt	tgggggattt	cggcctattg	gttaaaaaat	gagctgattt	2460
aacaaaaatt	taacgcgaat	taattctgtg	gaatgtgtgt	cagttagggt	gtggaaagtc	2520
cccaggctcc	ccaggcaggc	agaagtatgc	aaagcatgca	tctcaattag	tcagcaacca	2580
ggtgtgga	gtccccagc	tcccccagc	gcagaagta	gcaaagcatg	catctcaatt	2640
agtacgcaac	catagtcgcg	cccctaactc	cgccccatcc	gcccctaact	ccgcccagtt	2700
ccgcccattc	tcgcgcccat	ggctgactaa	ttttttttat	ttatgcagag	gccgaggccg	2760
cctctgcctc	tgaagctatc	cagaagtagt	gaggaggctt	ttttggaggc	ctaggctttt	2820
gcaaaaagct	cccgggagct	tgtatatcca	ttttcggatc	tgatcaagag	acaggatgag	2880
gatcgtttcg	catgattgaa	caagatggat	tgacgcagc	ttctccggcc	gcttgggtgg	2940
agaggctatt	cggtatgac	tgggcacaa	agacaatcgg	ctgctctgat	gccgcggtgt	3000
tcggcgtg	agcgcagggg	cgcccggttc	tttttgcata	gaccgacctg	tcgggtgccc	3060
tgaatgaact	gcaggacgag	gcagcgcggc	tatcgtggct	ggccacgacg	ggcgttcctt	3120
gcgcagctgt	gctcgacgtt	gtcactgaag	cggaagggga	ctggctgcta	ttgggcgaag	3180
tgccggggca	ggatctctcg	tcactctacc	ttgctcctgc	cgagaaagta	tcacatcatg	3240
ctgatgcaat	gcggcgctg	catacgcttg	atccgctac	ctgcccattc	gaccaccaag	3300
cgaaacatcg	catcgagcga	gcacgtactc	ggatggaagc	cggtcttgct	gatcaggatg	3360
atctggacga	agagcatcag	gggtctcgcc	cagccgaact	gttcgccagg	ctcaaggcgc	3420
gcatgccgga	cgccgaggat	ctcgtctgta	cccatggcga	tgccgtcttg	ccgaatatca	3480
tggtggaaaa	tgcccgcttt	tctggattca	tcgactgtgg	ccggctgggt	gtggcggacc	3540
gctatcagga	catagcgttg	gtaccccggt	atattgctga	agagcttgcc	ggcgaatggg	3600
ctgaccgctt	cctcgtgctt	tacggtatcg	ccgctcccca	ttcgcagcgc	atcgcttctt	3660
atcgcccttt	tgacgagttc	ttctgagcgg	gactctgggg	ttcgaaatga	ccgaccaagc	3720
gacgcccac	ctgccatcac	gagatttcga	ttccacgcgc	gccttctatg	aaagggtggg	3780
cttcggaaatc	gttttccggg	acgcgcgctg	gatgatcctc	cagcgcgggg	atctcatgct	3840
ggagtctctc	gccccaccga	acttgtttat	tgcagcttat	aatggttaca	aataaagcaa	3900
tagcatcaca	aatttcacaa	ataaagcatt	tttttccactg	cattctagtt	gtggtttgct	3960
caaaactcatc	aatgtatctt	atcatgtctg	tataccgctg	acctctagct	agagcttggc	4020
gtaatcatgg	tcatagctgt	ttcctgtgtg	aaattgttat	ccgctcacia	ttccacacaa	4080
catacagagcc	ggaagcataa	agtgtaaagc	ctgggggtgc	taatgagtga	gctaactcac	4140
attaattg	ttgcgctcac	tgcccgcttt	ccagtcggga	aaactgtcgt	gccagctgca	4200
ttaatgaatc	ggccaacg	cggggagagg	cggtttcgct	attgggcgct	cttccgcttc	4260
ctcgctcact	gactcgctgc	gctcggctcg	tcggctcgcc	cgagcgggat	cagctcactc	4320
aaaaggcgta	atacggttat	ccacagaatc	aggggataac	gcaggaaaga	acatgtgagc	4380
aaaaggccag	caaaaggcca	ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttccatag	4440
gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaaaccc	4500
gcagggacta	taagataacc	aggcgtttcc	ccctgggaagc	tcctcgtg	gctctcctgt	4560
tcggaccctg	ccgcttaccc	gatacctgtc	cgctttcttc	ccttcgggaa	gcgtggcgct	4620
ttctcaatgc	tcacgctgta	ggtatctcag	ttcggtgtag	gtcgctcgct	ccaagctggg	4680
ctgtgtgcac	gaaccccccg	ttcagcccga	ccgctgcgcg	ttatccggta	actatcgtct	4740
tgagtcacac	ccggtaagac	acgacttatc	gccactggca	gcagccactg	gtaacaggat	4800
tagcagagcg	aggatgttag	gcggtgctac	agagttcttg	aagtgtgggc	ctaactacgg	4860
ctacactaga	aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	4920

aagagtttgt	agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcgggtg	gtttttttgt	4980
ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	5040
tacggggctc	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	tcatgagatt	5100
atcaaaaaag	atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	5160
aagtatatat	gagtaaacct	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	5220
ctcagcgatc	tgtctatttc	gttcattccat	agttgcttga	ctccccgtcg	tgtagataac	5280
tacgatacgg	gagggcctac	catctggccc	cagtgtctga	atgataccgc	gagacccacg	5340
ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	ggaaggggcg	agcgagaaag	5400
tggctctgca	actttatccg	cctccatcca	gtctattaat	tgttgccggg	aagctagagt	5460
aagttagttcg	ccagtttaata	gtttgcgcaa	cgttgttgcc	attgtctacag	gcacgtgggt	5520
gtcacgctcg	tgttttggtg	tggtctcatt	cagctccggt	tcccaacgat	caaggcgagt	5580
tacatgatcc	cccatgttgt	gcaaaaaagc	ggttagctcc	ttcggctcctc	cgatcgttgt	5640
cagaagtaag	ttggccgcag	tgttatcact	catggttatg	gcagcactgc	ataattctct	5700
tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactggg	gagtactcaa	ccaagtcatt	5760
ctgagaatag	tgtagtcggc	gaccgagttg	ctcttgcccg	gcgtcaatac	gggataatac	5820
cgcgccacat	agcagaactt	taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	5880
actctcaagg	atcttaccgc	tgtagagatc	cagttcgatg	taaccacatc	gtgcacccaa	5940
ctgatcttca	gcacttttta	ctttaccagg	cgtttctggg	tgagcaaaaa	caggaaggca	6000
aaatgccgca	aaaaagggaa	taagggcgac	acggaaatgt	tgaatactca	tactcttcct	6060
ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatattga	6120
atgtattttg	aaaaataaac	aaataggggt	tccgcgcaca	tttccccgaa	aagtgcaccc	6180
tgacgtc						6187

SEQ ID NO:58

gacggatcgg	gagatctccc	gataccctat	ggctgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttgtgtgtt	ggaggctcgt	gagttagtgcg	120
cgagcaaaaat	ttagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cggtgacatt	240
gattattgac	tagttattaa	tagtaataca	ttacggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccggc	tggctgaccg	cccaacgacc	360
ccccccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaaata	gggactttcc	420
attgacgtca	atgggttgac	tatttaccgt	aaactgcccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttgcca	gtacatctac	gtattagtca	600
tctgtattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaaat	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaattggcg	780
gtaggcggtg	acgggtggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaaccg	gccctctaga	ctcgagcggc	cgccactgtg	ctggatatct	gcagaattca	960
tgcatggaga	tacacctaca	ttgcatgaat	atatgttaga	tttgcaacca	gagacaactg	1020
atctctactg	ttatgagcaa	ttaaatgaca	gctcagagga	ggaggatgaa	atagatggtc	1080
cagctggaca	agcagaaccc	gacagagccc	attacaataa	tgtaaccttt	tgttgcaagt	1140
gtgactctac	gcttcggttg	tgctacaaa	gcacacacgt	agacattcgt	actttggaag	1200
acctgttaat	gggcacacta	ggaattgtgt	gccccatctg	ttctcagaaa	ccaggatcta	1260
tgcgttacc	atacgatgtt	ccagattacg	ctagcttgag	atctaccatg	tctcagagca	1320
accgggagct	ggtggttgac	tttctctcct	acaagctttc	ccagaaagga	tacagctgga	1380
gtcagtttag	tgatgtggaa	gagaacagga	ctgaggcccc	agaagggact	gaatcggaga	1440
tggagacccc	cagtggccat	aatggcaacc	catcctggca	cctggcagac	agccccgcgg	1500
tgaatggagc	gactgcgcac	agcagcagtt	tggatgcccc	ggaggtgatc	cccatggcag	1560
cagtaaagca	agcgctgagg	gaggcaggcg	acgagtttga	actgcggtac	cggcgggcat	1620
tcagtgacct	gcacatccag	ctccacatca	ccccagggac	agcatatcag	agctttgaac	1680
aggtagtgaa	tgaactcttc	cgggatgggg	taaactgggg	tgcgattgtg	gcctttttct	1740
ccttcggcgg	ggcactgtgc	gtggaaaagc	tagacaagga	gatgcaggta	ttggtgagtc	1800
ggatcgcagc	ttggatggcc	acttacctga	atgaccacct	agagccttgg	atccaggaga	1860
acggcggctg	ggatactttt	gtggaactct	atgggaacaa	tgcagcagcc	gagagccgaa	1920
agggccaggga	acgcttcaac	cgtctgttcc	tgacgggcac	gactgtggcc	ggcgtggttc	1980
tactgggctc	actcttcagt	cggaatgaa	gatccaagct	taagtttaaa	ccgctgatca	2040
gcctcgactg	tgcttctag	ttgccagcca	tctgttgttt	gccccctccc	cgtagccttc	2100
ttgacctg	aaggtgccac	tcccactgtc	ctttcctaata	aaaatgagga	aattgcatcg	2160
cattgtctga	gtagggtgca	ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	2220
gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	ggcttctgag	2280
gcggaagaa	ccagctgggg	ctctaggggg	tatccccacg	cgccctgtag	cggcgcatta	2340
agcgcgccgg	gtgtggtgtg	tacgcgcagc	gtgaccgcta	cacttgccag	cgccctagcg	2400
cccgctcctt	tgcctttctt	cccttccttt	ctcgccacgt	tgcgcggctt	tccccgtcaa	2460
gctctaaatc	ggggcatccc	tttagggttc	cgatttagtg	ctttacggca	cctcgacccc	2520
aaaaaaactg	attaggttga	tggttcacgt	agtgggccat	cgccctgata	gacggttttt	2580
cgccctttga	cgttggagtc	cacgttcttt	aatagtggac	tcttgttcca	aaactggaaca	2640
acactcaacc	ctatctcggg	ctattctttt	gatttataag	ggattttggg	gatttcggcc	2700
tattggttaa	aaaatgagct	gattttaaca	aaatttaacg	cgaatttaatt	ctgtggaatg	2760
tgtgtcagtt	aggggtgtga	aagtcgccag	gctccccagg	caggcagaag	tatgcaaaag	2820
atgcatctca	attagtcagc	aaccaggtgt	ggaaagtccc	caggctcccc	agcaggcaga	2880
agtatgcaa	gcattgcatc	caatttagta	gcaaccatag	tcccgcctct	aactccgccc	2940
atcccgcccc	taactccgcc	cagttccgcc	cattctccgc	cccatggctg	actaattttt	3000
tttatttatg	cagaggccga	ggccgcctct	gcctctgagc	tattccagaa	gtagttagga	3060
ggcttttttg	gagccctagg	cttttgcaaa	aagctcccg	gagcttgtat	atccattttc	3120
ggaatctgat	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	tggattgcac	3180
cgaggttctc	cgctcggttg	gggtgagagg	ctattcggtc	atgactgggc	acaacagaca	3240
atcggtctgt	ctgatgccgc	cgtgttccgg	ctgtcagcgc	aggggcgccc	ggttcttttt	3300
gtcaagaccg	acctgtccgg	tgcctgaat	gaactgcagg	acgaggcagc	gcggctatcg	3360

```

tggctggcca cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga 3420
agggactggc tgctattggg cgaagtgcg gggcaggatc tcctgtcatc tcaccttgct 3480
cctgccgaga aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg 3540
gctacctgcc caatcgacca ccaagcgaaa catcgcatcg agcagacacg tactcggatg 3600
gaagccgggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc 3660
gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat 3720
ggcgatgcct gcttgcgcaa tatcatgggt gaaaatggcc gcttttctgg attcatcgac 3780
tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt 3840
gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct 3900
cccgattcgc agcgcatcgc cttctatcgc cttcttgacg agttcttctg agcgggactc 3960
tggggttcga aatgaccgac caagcgacgc ccaacctgcc atcacgagat ttcgattcca 4020
ccgcccgtt ctatgaaagg ttgggcttcg gaatcgtttt ccgggacgcc ggctggatga 4080
tcctccagcg cgggatctc atgctggagt tcttcgcccc ccccaacttg tttattgcag 4140
cttataatgg ttacaaataa agcaatagca tcacaaattt cacaaataaa gcattttttt 4200
cactgcattc tagttgtggt ttgtccaaac tcatcaatgt atcttatcat gtctgtatac 4260
cgtcgacctc tagctagagc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt 4320
gttatccgct cacaattcca cacaaacata gagccggaag cataaagtgt aaagcctggg 4380
gtgcctaag agtgagctaa ctacatttaa ttgcgttgcg ctactgccc gctttccagt 4440
cgggaaacct gtcgtgccag ctgcattaat gaatcgccca acgcgcgggg agaggcgggt 4500
tgctatttgg gcgctcttcc gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc 4560
tgcggcgagc ggtatcagct cactcaaagg cggtaatagc gttatccaca gaatcagggg 4620
ataacgcagg agtgagctat tgagcaaaag gccagcaaaa ggccaggaaac cgtaaaaagg 4680
ccgcttgcct ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac 4740
gctcaagtca ggggtggcga aacccgacag gactataaag ataccaggcg tttcccccctg 4800
gaagctccct cgtgcgctct cctgttccga ccctgcgctt taccggatac ctgtccgct 4860
ttctcccttc ggggaagcgt gcgctttctc aatgctcacg ctgtaggtat ctcagttcgg 4920
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgctcag cccgacgct 4980
gcgccttacc cggtaaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac 5040
tggcagcagc cactggtaac aggattagca gacgagagta tgtaggcggt gctacagagt 5100
tctgaagtg gtggcctaac tacggctaca ctagaaggac agtatttgggt atctgcgctc 5160
tgctgaagcc agttaccttc ggaaaaaagag ttggtagctc ttgatccggc aaacaaacca 5220
ccgctggtag cgggtgtttt ttgttttga agcagcagat tacgcgcaga aaaaaaggat 5280
ctcaagaaga tcttttgatc ttttctacgg ggtctgacgc tcagtggaaac gaaaactcac 5340
gttaagggat ttgtgtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 5400
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggctc gacagttacc 5460
aatgcttaat cagttaggca cctatctcag cgatctgtct atttcgttca tccatagttg 5520
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 5580
ctgcaatgat accgcgagac ccacgctcac cggtccaga tttatcagca ataaaccagc 5640
cagccggaag gcccgagcgc agaagtggtc ctgcaacttt atccgctccc atccagtcta 5700
ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacggtg 5760
ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtagtggt tcattcagct 5820
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 5880
gctccttcgg tcctccgacg gttgtcagaa gtaagtggc cgcaagtgtta tcaactatgg 5940
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 6000
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgacgg agttgtctct 6060
gcccgcgctc aatagcggtc aataccgcg ccatagcag aactttaaaa gtgctcatca 6120
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 6180
cgaatgaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 6240
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 6300
aatgttgaat atccatactc ttcccttttc aatattattg aagcatttat cagggttatt 6360
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 6420
gcacatttcc ccgaaaagtg ccacctgacg tc 6452

```

SEQ ID NO:59

```

Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
1 5 10 15
Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
20 25 30
Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
35 40 45
Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
50 55 60
Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
65 70 75 80
Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln
85 90 95
Lys Pro Gly Ser Met Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser
100 105 110
Leu Arg Ser Thr Met Ser Gln Ser Asn Arg Glu Leu Val Val Asp Phe
115 120 125
Leu Ser Tyr Lys Leu Ser Gln Lys Gly Tyr Ser Trp Ser Gln Phe Ser
130 135 140
Asp Val Glu Glu Asn Arg Thr Glu Ala Pro Glu Gly Thr Glu Ser Glu
145 150 155 160
Met Glu Thr Pro Ser Ala Ile Asn Gly Asn Pro Ser Trp His Leu Ala
165 170 175
Asp Ser Pro Ala Val Asn Gly Ala Thr Ala His Ser Ser Ser Leu Asp
180 185 190
Ala Arg Glu Val Ile Pro Met Ala Ala Val Lys Gln Ala Leu Arg Glu
195 200 205

```

Ala Gly Asp Glu Phe Glu Leu Arg Tyr Arg Arg Ala Phe Ser Asp Leu
 210 215 220
 Thr Ser Gln Leu His Ile Thr Pro Gly Thr Ala Tyr Gln Ser Phe Glu
 225 230 235 240
 Gln Val Val Asn Glu Leu Phe Arg Asp Gly Val Asn Trp Gly Arg Ile
 245 250 255
 Val Ala Phe Phe Ser Phe Gly Gly Ala Leu Cys Val Glu Ser Val Asp
 260 265 270
 Lys Glu Met Gln Val Leu Val Ser Arg Ile Ala Ala Trp Met Ala Thr
 275 280 285
 Tyr Leu Asn Asp His Leu Glu Pro Trp Ile Gln Glu Asn Gly Gly Trp
 290 295 300
 Asp Thr Phe Val Glu Leu Tyr Gly Asn Asn Ala Ala Ala Glu Ser Arg
 305 310 315 320
 Lys Gly Gln Glu Arg Phe Asn Arg Trp Phe Leu Thr Gly Met Thr Val
 325 330 335
 Ala Gly Val Val Leu Leu Gly Ser Leu Phe Ser Arg Lys
 340 345

SEQ ID NO:60

atggcgtacc catacgatgt tccagattac gctagcttga gatctacccat gtctcagagc 60
 aaccgggagc tgggtggttga ctttctctcc tacaagcttt cccagaaagg atacagctgg 120
 agtcagttta gtgatgtgga agagaacagg actgaggccc cagaaggagc tgaatcggag 180
 atggagaccc ccagtgccat caatggcaac ccattcctggc acctggcaga cagccccgcg 240
 gtgaatggag ccactgcgca cagcagcagt ttggatgccc gggaggtgat ccccatggca 300
 gcagtaaagc aagcgctgag ggaggcaggc gacgagtttg aactgcggta ccggcgggca 360
 ttcagtgacc tgacatccca gctccacatc accccaggga cagcatatca gagctttgaa 420
 caggtagtga atgaactctt ccgggatggg gtagccattc ttgcgattgt ggcctttttc 480
 tccttcggcg gggcactgtg cgtggaaagc gtagacaagg agatgcaggt attggtgagt 540
 cggatcgag cttggatggc cacttacctg aatgaccacc tagagccttg gatccaggag 600
 aacggcggct gggatacttt tgtggaactc tatgggaaca atgcagcagc cgagagccga 660
 aagggccagg aacgcttcaa ccgctggttc ctgacgggca tgactgtggc cggcgtggtt 720
 ctgctgggct cactcttcag tcggaaatga 750

SEQ ID NO:61

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

SEQ ID NO:62

Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
 1 5 10 15
 Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
 20 25 30
 Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
 35 40 45
 Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
 50 55 60
 Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
 65 70 75 80
 Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln

SEQ ID NO:63						
gacggatcgcg	gagatctccc	gatccccatat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgcctccctg	cttgttggtt	ggaggctcgt	gagtagtcgc	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttcgccga	caattgcgatg	aagaatctgc	180
ttagggttag	gcgtttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttatttaa	tagtaaatcaa	ttacggggtc	attaaattcat	agccacataa	300
tggagttccg	ctgtacataa	cttacggtaa	atggccggcc	tggctgaccg	ccccacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgcccata	gggactttcc	420
attgacgtca	atgggttgac	tattttacggt	aaactgccca	cttggcagta	catcaagtg	480
atcatattgcc	aagtagccccc	cctattgacg	tcaattgacg	taaatggccc	gcctggcatt	540
atgccagcta	catgacctta	tgggactttc	ctactttgca	tgatcatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttagcgtcaa	tgggagtttg	ttttggcacc	720
aaaatcagca	ggactttcca	aaatgtccta	acaactccgc	cccattgacg	caaatggcgc	780
gtaggcgtgt	acgggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaacca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaacgg	gcctcttga	ctcgagcgcc	cgccactgtg	ctggatatct	gcgaattccc	960
accacactag	actagtgtat	ctatggcgta	cccatcacg	gttccagatt	acgctagctt	1020
gagatctacc	atgtctcaga	gcaaccggga	gctgggtggt	gactttctct	cctacaagct	1080
tccccagaaa	ggatacagct	ggagttagtt	tatgtatgtg	gaagagaaca	ggactgaggc	1140
ccccagaagg	actgaattcg	agaatggagc	cccagtgct	atcaatggca	accatccctg	1200
gcacctggca	gacagccccg	cggtgaaatg	agccactgcg	cacagcagca	gtttggatgc	1260
ccgggagggtg	atccccatgg	cagcagtaaa	gcaagcgctg	agggaggcag	gcgcagagtt	1320
tgaactcgcg	taccggcgcg	cattcagtha	cctgatactc	cagctccaca	tcacccccag	1380
gacagcatat	cagagactttg	aacaggtagt	gaatgaactc	ttcggggatg	gggtagccat	1440
tcttcgcatt	gtggcctttt	tctccttcgg	cggggcactg	tgcgtggaaa	gcgtagacaa	1500
ggagatcgag	gtattgtgtg	ctcggaatgc	agcttggaat	ggcaattacc	tgaatgacca	1560
ctagagacct	tggatccagg	agaaaggcg	ctgggatact	tttgtggaac	tctatgggaa	1620
caatgcagca	gcccagagcc	gaaagggcca	ggaacgcttc	aaccgctggt	tcttgacggg	1680
catgactgtg	gocggcgctg	ttctgctggg	ctactctctc	agtcggaaat	gaagatccga	1740
gctcggtacc	aagctttaag	ttaaaaacct	gacgagctc	gactgtgctt	tctagttgcc	1800
agccatctgt	tgttttgcct	tcceccgtgc	cttctctgac	cttgggaagt	gccattccca	1860
ctgtcctttc	ctaataaaat	gaggaaaatg	catcgcattg	tctgagtagg	tgtcattcta	1920
ttctgggggg	tgggggtggg	caggacagca	agggggagga	tggggaagac	aatagcaggc	1980
atgctgggga	tgcggtggcc	tctatggctt	ctgaaggcga	aagaaccagc	tggggtctta	2040
gggggtatcc	ccacgcggcc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	2100
gcagcgtgac	cgtctacatt	gcacgcgcc	tagcgcgccg	tcctttcctg	ttcttccctt	2160
cctttctcgc	cacggttcgc	ggctttcccc	gtcaagctct	aaatcggggc	atccctttag	2220
ggttcgagtt	tagtgcttta	cgcacacctg	acccccaaaa	acttgattag	gggtatggtt	2280
cacgtagtgg	gccatgcgcc	tgatagacgg	tttttcgccc	tttgacgttg	gagttccagt	2340
tcttttaatg	tggactcttg	ttccaaaact	gaaccaactc	caacctaatc	tcggtctatt	2400
cttttgattt	ataagggatt	ttggggattt	cggtactatg	gtttaaaaat	gagctgattt	2460
aacaaaaaat	taacgcgaat	taattctgtg	gaatgtgtgt	cagttagggt	gtggaaaagt	2520
cccaggtccc	ccaggcaggc	agaagtatgc	aaagcatgca	tctcaattag	tcagcaacca	2580
ggtgtggaaa	gtccccaggc	tccccagcag	gcagaagtat	gcaaacgatc	catctcaatt	2640
agtcagcaac	catagtcccg	ccctaactc	qcccatccc	qccctaact	ccqccqaatt	2700

ccgccccat	tcgccccat	ggctgactaa	ttttttttat	ttatgcagag	gccgaggccg	2760
cctctgcctc	tgagctattc	cagaagtagt	gaggaggcct	ttttggaggc	ctaggctttt	2820
gcaaaaaagc	cccgggagct	tgatatatcca	ttttcggatc	tgatcaagag	acaggatgag	2880
gatcggtttc	catgattgaa	caagatggat	tgcaacgagg	ttctccggcc	gcttgggtgg	2940
agaggctatt	cggtctatgac	tgggcacaa	agacaatcgg	ctgctctgat	gccgcggtgt	3000
tcgggctgtc	agcgagggg	cgcccggttc	tttttgtcaa	gaccgacctg	tccggtgccc	3060
tgaatgaact	gcaggacgag	gcagcgcggc	tatcgtggct	ggccacgacg	ggcggttccct	3120
gcgcagctgt	gctcgacgtt	gtcactgaag	cggaagggga	ctggctgcta	ttgggcgaag	3180
tgccggggca	ggatctctctg	tcctctcacc	ttgctcctgc	cgagaaagta	tccatcatgg	3240
ctgatgcaat	gcggcggtcg	catacgcttg	atccggctac	ctgccccattc	gaccaccaag	3300
cgaacaatcg	catcgagcga	gcacgtactc	ggatggaagc	cggtcttgct	gatcaggatg	3360
atctggacga	agagcatcag	gggtctcgcg	cagccgaact	gttcgccagg	ctcaaggcgc	3420
gcatgccccg	cgcgaggat	ctcgctcgta	cccatggcga	tgccctgctt	cgaatatca	3480
tgggtgaaaa	tggccgcttt	tctggattca	tcgactgtgg	ccggctgggt	gtggcgacc	3540
gctatcagga	catagcgttg	gtaccggtg	atattgtcta	agagcttggc	ggcgaatggg	3600
ctgaccgctt	cctcgctgtt	tacggtatcg	ccgctcccga	ttcgacgctg	atcgcttct	3660
atcgcttct	tgacgagttc	ttctgagcgg	gactctgggg	ttcgaaatga	ccgaccaagc	3720
gacgcccac	ctgcctacac	gagatttcga	ttccaccgct	gccttctatg	aaaggttggg	3780
cttcggaatc	gttttccggg	acgcgggtcg	gatgatcttc	cagcgcgggg	atctcatgct	3840
ggagttcttc	gcccacccca	acttggttat	tgacgcttat	aatggttaca	aataaagcaa	3900
tagcatcaca	aatttcacaa	ataaagcatt	tttttctact	cattctagtt	gtggtttgtc	3960
caaatctatc	atcgtatctt	atcatgtctg	tataccgtcg	acctctagct	agagcttggc	4020
gtaatcatgg	tcatagctgt	ttcctgtgtg	aaattgttat	ccgctcacia	ttccacacaa	4080
catacagagc	ggaagcataa	agtgtaaagc	ctgggggtgc	taatgagtga	gctaactcac	4140
attaatttgc	ttgcgctcac	tgcccgcttt	ccagtcggga	aacctgtcgt	gccagctgca	4200
ttaatgaatc	ggccaacgcg	cggggagagg	cggtttgcgt	attgggcgct	cttcgcgttc	4260
ctcgctcact	gactcgctgc	gctcggtcgt	tcggctcgcg	cgagcggtat	cagctcactc	4320
aaaggcggtg	atacggttat	ccacagaatc	aggggataac	gcaggaaaga	acatgtgagc	4380
aaaaggccag	caaaagcca	ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttccatag	4440
gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaacccc	4500
gcaggagcta	taagataacc	aggcgtttcc	ccctggaagc	tcctctgtgc	gctctcctgt	4560
tcgacccctg	ccgcttaccg	gatacctgtc	cgcttctctc	ccttcgggaa	gcgtggcgct	4620
ttctcaatgc	tcacgctgta	ggtatctcag	ttcggtgtag	gtcgctcgtc	ccaagctggg	4680
ctgtgtgcac	gaaccccccg	ttcagcccca	ccgctgcgcc	ttatccggta	actatcgtct	4740
tgagtccaac	ccggtaaagc	acgacttatc	gccactggca	gcagccactg	gtaacaggat	4800
tagcagagcg	aggtatgtag	gcggtgctac	agagttcttg	aagtgggtgg	ctaactacgg	4860
ctacactaga	aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	4920
aagagtttgt	agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcggtg	gtttttttgt	4980
ttgcaagcag	cagatttcgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	5040
tacggggtct	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	tcatgagatt	5100
atcaaaaaag	atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	5160
aagtatatat	gagtaaaact	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	5220
ctcagcgatc	gtcttatctt	gttcattcat	agttgacctg	ctccccgtcg	tgtagataac	5280
tacgatacgg	gagggcttac	catctggccc	cagtgtctga	atgataccgc	gagacccacg	5340
ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	ggaagggccg	agcgcagaag	5400
tggtcctgca	actttatccg	cctccatcca	gtctattaat	tggtgcccgg	aagctagagt	5460
aagtagttcg	ccagtttaata	gtttgcgcaa	cggtgttgcc	attgctacag	gcatcggtgt	5520
gtcacgctcg	tcgttttgta	tggcttcatt	cagctccggt	tcccaacgat	caaggcgagt	5580
tacatgatcc	cccatgttgt	gcaaaaaagc	ggtagctctc	ttcggtcctc	cgatcgttgt	5640
cagaagtaag	ttggccgcat	gtttatcact	catggttatg	gcagcactgc	ataattctct	5700
tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactggg	gagtactcaa	ccaagtctat	5760
ctgagaatag	tgtatgcggc	gaccgagttg	ctcttgcccg	gcgtcaatac	gggataatac	5820
cgcgccacat	acagagaact	taaaaagtgt	catcattgga	aaacgttctt	cggggcgaaa	5880
actctcaagg	atcttaccgc	tggttagatc	cagttcgatg	taaccactc	gtgcacccaa	5940
ctgatcttca	gcacttttta	ctttaccag	cgtttctggg	tgagcaaaaa	caggaaaggca	6000
aaatgccgca	aaaaagggaa	taagggcgac	acggaaatgt	tgaatactca	tactcttctt	6060
ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatttga	6120
atgtatttag	aaaaataaac	aaataggggt	tcgcgcgaca	tttccccgaa	aagtgccacc	6180
tgacgtc						6187

SEQ ID NO:64

acggatcgg	gagatctccc	gatccccat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttgtgtgtt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaaagtaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacggggtc	attagttcat	agccatata	300
tgaggttccg	cgttacataa	cttacggtaa	atggcccgcc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggt	aaactgcca	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggacttct	ctacttgcca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaaat	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccatgacg	caaatgggcg	780
gtaggcgtgt	acgggtgggag	gtctatataa	gcagagctct	ctggctaaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaacgg	gccctctaga	ctcgagcggc	cgccactgtg	ctggatatct	gcagaattca	960
tgcatggaga	tacacatcaa	ttgcatgaat	atatgttaga	tttgcaacca	gagacaactg	1020
atctctactg	ttatgagcaa	ttaaatgaca	gctcagagga	ggaggatgaa	atagatgggtc	1080
cagctggaca	agcagaacccg	gacagagccc	attacaatat	tgtaaccttt	tggtgcaagt	1140

gtgactctac	gcttcgggtg	tgcgtacaaa	gcacacacgt	agacattcgt	acttttgaag	1200
acctgttaat	gggcacacta	ggaattgtgt	gccccatctg	ttctcagaaa	ccaggatcta	1260
tggcgtagcc	atacgatgtt	ccagattacg	ctagcttgag	atctaccatg	tctcagagca	1320
accgggagct	ggtaggtgac	ttctctcctt	acaagctttc	ccagaaagga	tacagctgga	1380
gtcagtttag	tgaagtggaa	gagaacagga	ctgaggcccc	agaagggact	gaatcggaga	1440
tggagacccc	cagtgcctac	aatggcaacc	catcctggca	cctggcagac	agcccccgcg	1500
tgaatggagc	cactgcgcac	agcagcagtt	tggatgcccc	ggagggtgac	cccatggcag	1560
cagtaaaagca	agcgctgagg	gaggcaggcg	acgagtttga	actgcggtac	cggcgggcat	1620
tcagtgcact	gacatcccag	ctccacatca	ccccagggac	agcatatcag	agctttgaac	1680
aggtagtgaa	tgaactcttc	cgggatgggg	tagccattct	tgcgattgtg	gcctttttct	1740
ccttcggcgg	ggcactgtgc	gtggaaaagc	tagacaagga	gatgcaggta	ttggtgagtc	1800
ggatcgagc	ttggatggcc	acttacctga	atgaccacct	agagccttgg	atccaggaga	1860
acggcggctg	ggatactttt	gtggaactct	atgggaacaa	tgcagcagcc	gagagccgaa	1920
agggccagga	acgcttcaac	cgctgggttc	tgacgggcat	gactgtggcc	ggcgtgggtc	1980
tgtcgggtct	actcttcagt	cggaaaatgaa	gatccaaagt	taagtttaaa	ccgctgatca	2040
gcctcgactg	tgccttctag	ttgccagcca	tctgttgttt	gccccctccc	cgtgccttcc	2100
ttgacctgtg	aaggtgcac	tcccactgtc	ctttcctaata	aaaatgagga	aattgcatcg	2160
catgtcttga	gtagggtgtca	ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	2220
gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	ggcttctgag	2280
gcggaagaa	ccagctgggg	ctctaggggg	tatccccacg	cgccctgtag	cggcgcatca	2340
acgcggcg	gtgtggtggt	tacgcgcagc	gtgaccgcta	cacttgccag	cgccctagcg	2400
cgcgctctct	tgccttcttt	cccttccctt	ctcgcgcagt	tgcgcggctt	tccccgtcaa	2460
gctctaaatc	ggggcatccc	tttaggggtc	cgatttagtg	ctttacggca	cctcgacccc	2520
aaaaaacttg	attaggttga	tggttcacgt	agtgggcat	cgccctgata	gacggttttt	2580
cgcccttttg	cgttggagtc	caacttcttt	aatagtggac	tcttgttcca	aaatggaaca	2640
acactcaacc	ctatctcggg	ctatttcttt	gatttataag	ggattttggg	gatttcgggc	2700
tattgtttaa	aaaatgagct	gatttataaa	aaattttaac	cgaattaatt	ctgtggaatg	2760
tgtgtcagtt	aggggtgtga	aagtcctccg	gctccccagg	caggcagaag	tatgcaaagc	2820
atgcatctca	attagctcag	aaccagggtg	ggaaaagtccc	caggctcccc	agcaggcaga	2880
agtagcaaaa	gcactgcact	caattagtc	gcaaccatag	tccccgccct	aactccgccc	2940
atccccgccc	taactccgca	cagttccgcc	cattctccgc	cccatggctg	actaattttt	3000
tttatttatg	cagaggcgca	ggcgcctct	gcctctgagc	tattccagaa	gtagtgaaga	3060
ggcttttttg	gaggcctagg	cttttgcaaa	aagctcccg	gagcttgtat	atccattttc	3120
ggatctgac	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	tggattgcac	3180
cgaggttctc	cggccgcttg	ggtggagagg	ctattcggct	atgactgggc	acaacagaca	3240
atcggctgct	ctgatgcgcg	cgtgttccgg	ctgtcagcgc	aggggcgccc	ggttcttttt	3300
gtcaagaccg	acctgtcccg	tgcctgaat	gaactgcagg	acgaggcagc	gcggctatcg	3360
tggctggcca	cgcaggcggt	tccttgccga	gctgtgctcg	acgttgtcac	tgaagcggga	3420
agggactggc	tgtctattgg	cgaagtgcg	ggcgaggatc	tcctgtcatc	tcacctgtct	3480
cctgccgaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	gcttgatccg	3540
gctacctgcc	cattcgacca	ccaagcgaaa	catcgcatcg	agcgagcacg	tactcggatg	3600
gaagccgggt	ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	cgccgcagcc	3660
gaactgtctc	ccaggtctca	ggcgcgcatg	ccgcagcgcg	aggatctcgt	cgtgacccat	3720
ggcgatgcct	gcttgccgaa	tatcatgggt	gaaaatggcc	gcttttcttg	attcatcgac	3780
tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cgttggctac	ccgtgatatt	3840
gctgaagagc	ttggcggcga	atgggctgac	cgcttctctg	tgtcttacgg	tatcgccgct	3900
cccgattcgc	agcgcatcgc	cttctatcgc	cttcttgacg	agttcttctg	agcgggactc	3960
tggggtttca	aatgaccgac	caagcgacgc	ccaacctgcc	atcacgagat	ttcgattcca	4020
cgcgcgcctt	ctatgaaaag	ttgggcttcg	gaatcgtttt	ccgggacgcc	ggctggatga	4080
tcctccagcg	cggggatctc	atgctggagt	tcttcgccca	ccccaaactt	tttattgcag	4140
cttataatgg	ttacaaaata	agcaatagca	tcacaaaatt	cacaaaataa	gcattttttt	4200
cactgcattc	tagtttgtgt	ttgtccaaac	tcatcaatgt	atcttatcat	gtctgtatac	4260
cgtgcacctc	tagctgttat	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	4320
gttatccgct	cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	4380
gtgcctaata	agttagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	gctttccagt	4440
cgggaaacct	gtcgtgccag	ctgcattaat	gaatcgccca	acgcgcgggg	agaggcgggt	4500
tgcgtattgg	gcgctcttcc	gcttctcgcg	tcactgactc	gctgcgctcg	gtcgttcggc	4560
tgcggcgagc	ggtatcagct	cactcaaaag	cggtaatcac	gttatccaca	gaatcagggg	4620
ataacgcagg	aaagaacatg	tgagcaaaa	gccagcaaaa	ggccaggaa	cgtaaaaagg	4680
ccggtttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	4740
gctcaagtca	gaggtggcga	aaaccgacag	gactataaag	ataccaggcg	tttccccctg	4800
gaagctccct	cgtgcgctct	cctgttccga	ccctgcccgt	taccggatac	ctgtccgcct	4860
ttctcccttc	gggaagcgtg	gcgctttctc	aatgctcacg	ctgtagggtat	ctcagttcgg	4920
tgtaggctcg	tcgctccaa	ctgggctgtg	tgcacgaacc	ccccgttcag	cccagccgct	4980
gcgccttatc	cgttaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	5040
tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	5100
tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	5160
tgtgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	5220
ccgctggtag	cgttggtttt	tttgttttga	agcagcagat	tacgcgcaga	aaaaaaggat	5280
ctcaagaaga	tccttttgatc	ttttctacgg	ggtctgacgc	tcagtggaa	gaaaactcac	5340
gttaagggat	tttggctatg	agattatcaa	aaaggatctt	cactatagat	cttttaaat	5400
aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	5460
aatgcttaat	cagttaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	5520
cctgactccc	cgtcgttag	ataactacga	tacgggagg	cttaccatct	ggccccagtg	5580
ctgcaatgat	accgcgagac	ccacgctcac	cggtctcaga	tttatcagca	ataaaccagc	5640
cagccggaag	ggccgagcgc	agaagtgtgc	ctgcaacttt	atccgcctcc	atccagtcta	5700
ttaatgtttg	ccgggaagct	agagtaagta	gttcgcccag	taatagtttg	cgcaacggtt	5760
ttgccattgc	tacaggcatc	gtgggtgtcac	gctcgtcggt	tggtagggct	tcattcagct	5820
ccggttccca	acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcgggtta	5880
gctccttcgg	tcctccgata	gttgtcagaa	gtaagttggc	cgcagtgtta	tcactcatgg	5940

ttatggcgagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	6000
ctgggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgtctctt	6060
gcccgcgctc	aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	6120
ttggaaaacg	tttttcgggg	cgaaaaactct	caaggatctt	accgctgttg	agatccagtt	6180
cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	6240
ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataaag	gcgacacgga	6300
aatgttgaat	actcatactc	ttcctttttc	aatattattg	aagcatttat	cagggttatt	6360
gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	taaacaaaaa	ggggttccgc	6420
gcacatttcc	ccgaaaagtg	ccacctgacg	tc			6452

SEQ ID NO: 65

atggcgggatg	tgtgacatac	acgacgccaa	aagattttgt	tccagctcct	gccacctccg	60
ctacgcgaga	gattaaccac	ccacgatggc	cgccaaaagt	catgttgata	ttgaggctga	120
cagcccatte	atcaagctct	tgcagaaggc	atttccgtcg	ttcgagggtg	agtcattgca	180
ggtcacacca	aatgaccatg	caaatgccag	agcattttcg	cacctggcta	ccaaattgat	240
cgagcaggag	actgacaaag	acacactcat	cttggatata	ggcagtgccg	cttccaggag	300
aatgatgtct	acgcacaaat	accactgcgt	atgccctatg	cgcagcgccg	aagaccccca	360
aaggctcgat	agctacgcaa	agaaactggc	agcggcctcc	gggaagggtg	tggatagaga	420
gatcgagga	aaaatcaccc	acctgcagac	cgctatggct	acgcagacgc	ctgaatctcc	480
taccttttgc	ctgcatacag	acgtcacgtg	tctgtacggc	gccgaagtgg	ccgtatacca	540
ggacgtgtat	gctgtacatg	caccaacatc	gctgtaccat	caggcgatga	aaggtgtcag	600
aacggcgatg	ttgattgggt	ttgacaccac	cccgtttatg	tttgacgcgc	tagcaggcgc	660
gtatccaacc	tacgccacaa	actgggccga	cgagcagggt	ttacaggcca	ggaacatagg	720
actgtgtgca	gcactcctga	ctgaggggaag	actcggcaaa	ctgtccattc	tccgcaagaa	780
gcaattgaaa	cccttgccga	cagtcatgtt	ctcggtagga	tctacattgt	acactgagag	840
cagaaagcta	ctgaggagct	ggcacttacc	ctccgtatcc	cacctgaaag	gtaaacaatc	900
ctttacctgt	agggtgcgata	ccatcgtatc	atgtgaaggg	tacgtagtta	agaaaaatcac	960
tatgtgcccc	ggcctgtacg	gtaaaaacgt	agggtagccc	gtgacgtatc	acgcggaggg	1020
attccttagtg	tgaagaccca	cagacactgt	caaaggagaa	agagtctcat	tccctgtatg	1080
cacctacgtc	ccctcaacca	tctgtgatca	aatgactggc	atactagcga	ccgacgtcac	1140
accggaggac	gcacagaaat	tgttagtggt	attgaatcag	aggatagtgt	tgaacggaa	1200
aacacagcga	aacactaaca	cgatgaagaa	ctatctgctt	ccgattgttg	ccgtcgcatt	1260
tagcaagtgg	gcgagggaat	acaaggcaga	ccttgatgat	gaaaaacctc	tgggtgtccg	1320
agagagggtca	cttacttgct	gctgcttggt	ggcattttaa	acgagggaag	tgcacaccat	1380
gtacaagaaa	ccagacaccc	agacaatagt	gaagggtgct	tcagagttaa	actcgttctg	1440
catcccagc	ctatggctca	caggccctcg	aatcccagtc	agatcacgca	ttaagatgct	1500
tttggccaag	aagaccaagc	gagagttaat	acctgttctc	gacgcgtcgt	cagccaggga	1560
tgtgaacaaa	gaggagaaag	agagggttga	ggccgagctg	actagagaag	ccttaccacc	1620
cctcgtcccc	atcgcgcggg	cgagacgggg	agtcgtcgac	gtcgcagttg	aagaactaga	1680
gtatcacgca	ggtgcagggg	tctgtgaaac	acctcgcagc	gcgttgaaa	tcaccgcaca	1740
gccgaacgac	gtactactag	gaaattactg	agttctgtcc	ccgcagaccg	tgctcaagag	1800
ctccaagtgg	gcccccgctg	accctctagc	agagcagggt	aaaataataa	cacataacgg	1860
gagggccggc	ggtttaccag	tcgacggata	tgacggcagg	gtcctactac	catgtggatc	1920
ggccattccg	gtccctgagt	ttcaagcttt	gagcagagac	gccactatgg	tgtacaacga	1980
aaggggagtt	gtcaacagga	aactatacca	tattgcccgt	cacggaccgt	cgctgaacac	2040
cgacgagag	aagtcagagc	tgaaagaaat	gacgcggagt	gacgcggagt	acgtgttcga	2100
cgtagataaaa	aaatgctgct	tcaagagaga	ggaagcgtcg	ggtttggtgt	tgggtgggaga	2160
gctaaccaaac	cccccggtcc	atgaattcgc	ctacgaaggg	ctgaagatca	ggccgtcggc	2220
accatataag	actacagtag	taggagctct	tgggggtccg	ggatcaggga	agtcctgtat	2280
tatttaagagc	ctcgtgacca	aacacgatct	ggtcaccagc	ggcaagaagg	agaactgccca	2340
ggaaatagtt	aacgacgtga	agaagcaccc	cggaaggggg	acaagtaggg	aaaacagtga	2400
ctccatctctg	ctaaacgggt	gtcgtcgtgc	cgtggacatc	ctatatgtgg	acgaggcttt	2460
cgtcgtccat	tccgtgccc	tgctggccct	aattgctctt	gttaaacctc	ggagcaaaat	2520
ggtgttatgc	ggagacccca	agcaatgcgg	attcttcaat	atgatgcagc	ttaagggtga	2580
cttcaaccac	aacatctgca	ctgaagtatg	tcataaaagt	atatccagac	gttgacgcgc	2640
ttccagtacg	gccatcgtgt	ctacgttgca	ctacggaggc	aagatgcgca	cgaaccaacc	2700
gtgcaacaaa	cccataatca	tagacaccac	aggacagacc	aagcccaagg	caggagacat	2760
cgtgttaaca	tgtctccgag	gctgggcaaa	gcagctgcag	ttggactacc	gtggacacga	2820
agtcatgaca	gcagcagcat	ctcaggccct	caccgcgaaa	gggggtatag	ccgtaaaggca	2880
gaagggtgaat	gaaaatccct	tgtatgcccc	tgcgtcggag	cacgtgaatg	tactgttgac	2940
gcgcactgag	gataggctgg	tgtggaaaaa	gctggccggc	gatccctgga	ttaaggctct	3000
atcaaacatt	ccacagggtg	actttacggc	cacattggaa	gaatggcaag	aagaacacga	3060
caaaataatg	aagggtgattg	aaggaccggc	tgcgcctgtg	gacgcgttcc	agaacaaagc	3120
gaacgtgtgt	tgggcgaaaa	gcctgggtgc	tgtcctggac	actgccggaa	tcagattgac	3180
agcagaggag	tggagcacca	taattacagc	atttaaggag	gacagagctt	actctccagt	3240
ggtggccttg	aatgaaatgt	gcaccaagta	ctatggagtt	gacctggaca	gtggcctgtt	3300
ttctgccccg	aagggtgtccc	tgtattacga	gaacaaccac	tgggataaca	gacctggttg	3360
aaggatgtat	ggattcaatg	ccgcaacagc	tgccaggctg	gaagctagac	ataccttctc	3420
gaagggggcag	tggcatatcg	gcaagcaggc	agttatcgca	gaaagaaaaa	tccaaccgct	3480
ttctgtgctg	gacaatgtaa	ttcctatcaa	ccgcaggctg	ccgcacgccc	tgggtggtga	3540
gtacaagacg	gttaaaggca	gtgggttgta	gtgggtggtc	aataaaagtaa	gagggtagca	3600
cgtcctgctg	gtgagttagt	acaacctggc	tttgccctcga	cgcagggtca	cttggtgtgc	3660
accgctgaat	gtcacaggcg	ccgataggtg	ctacgacctc	agttttaggac	tgccggctga	3720
cgcggcgagg	ttcgacttgc	tctttgtgaa	cattcacacg	gaattcagaa	tccaccacta	3780
ccagcagtg	gtcgaccacg	ccatgaagct	gcagatgctt	gggggagatg	cgctacgact	3840
gctaaaaccc	ggcggtcatct	tgatgagagc	ttacggatac	gccgataaaa	tcagcgaagc	3900
cgttggtttc	tccttaagca	gaaagtcttc	gtctgcaaga	gtgttgccgc	cggattgtgt	3960
caccagcaat	acagagagtg	tcttgctgtt	ctccaaacttt	gacaaacggaa	agagaccctc	4020
tacgctacac	cagatgaata	ccaagctgag	tgccgtgtat	gccggagaag	ccatgcacac	4080
ggccgggtgt	gcaccatcct	acagagttaa	gagagcagac	atagccacgt	gcacagaagc	4140

ggctgtggtt	aacgcagcta	acgcccgtgg	aactgtaggg	gatggcgat	gcagggccgt	4200
ggcgaagaaa	tgcccgctcag	cctttaaggg	agcagcaaca	ccagtgggca	caattaaaaac	4260
agtcagtgtc	ggctcgctacc	ccgtcatcca	cgctgtagcg	cctaatttct	ctgccacgac	4320
tgaagcggaa	ggggaccgcg	aattggccgc	tgtctaccgg	gcagtggccg	ccgaagtaaa	4380
cagactgtca	gggagcagcg	tagccatccc	gctgctgtcc	acaggagtgt	tcagcggcgg	4440
aagagatagg	ctgcagcaat	ccctcaacca	tctattcaca	gcaatggacg	ccacggacgc	4500
tgacgtgacc	atctactgca	gagacaaaag	ttggggagaag	aaaatccagg	aagccattga	4560
catgaggacg	gctgtggagt	tgctcaatga	tgacgtggag	ctgaccacag	acttggtgag	4620
agtgcacccg	gacagcagcc	tggtgggtcg	taagggctac	agtaccactg	acgggtcgct	4680
gtactcgctac	tttgaaggta	cgaatttcaa	ccaggctgct	attgatattg	cagagatact	4740
gacgtttgtg	cccagactgc	aagaggcaaa	cgaacagata	tgccataacg	cgctgggcga	4800
aacaatggac	aacatcagat	ccaaatgtcc	ggtgaacgat	tccgattcat	caacacctcc	4860
caggacagtg	ccctgcctgt	gccgctacgc	aatgacagca	gaacggatcg	cccgccttag	4920
gtcacaccaa	gttaaaagca	tggtggtttg	ctcatctttt	cccctcccga	aataccatgt	4980
agatggggtg	cagaaggtaa	agtgcgagaa	ggttctctctg	ttcgaccoga	cggtaaccttc	5040
agtgttagt	ccgcgggaagt	atgccgcac	tacgacggac	cactcagatc	ggtcgttacg	5100
agggtttagc	ttggactgga	ccaccgactc	gtcttccact	gccagcgata	ccatgtcgct	5160
accagtttg	cagtcgtgtg	acatcgactc	gatctacgag	ccaatggctc	ccatagtagt	5220
gacggctgac	gtacacccctg	aaccgcgagg	catcgcgagc	ctggcggcag	atgtgcaccc	5280
tgaccccgca	gaccatgtgg	acctcgagaa	cccgatctct	ccaccgcgcc	cgaagagagc	5340
tgcatacctt	gcctcccgcg	cgccggagcg	accggtgccg	gcgccgagaa	agccgacgcc	5400
tgccccaagg	actcgcttta	ggaacaagct	gcctttgacg	ttcgggcact	ttgacgagca	5460
cgaggtcgat	gcgttggcct	ccgggattac	tttcggagac	ttcgacgacg	tcttgcgact	5520
agcccgccg	ggtgcataata	ttttctcctc	ggacactggc	agcggacatt	tacaacaaaa	5580
atccgttagg	cagcacatc	tccagtgcgc	acaactggat	gcgggtccagg	aggagaaaaa	5640
gtacccgccca	aaattggata	ctgagaggga	gaagctgttg	ctgctgaaaa	tgcatagtga	5700
cccatcggag	gctaataaga	gtcgatacca	gtctcgcaaa	gtggagaaca	tgaaagccac	5760
ggtggtggac	aggctcacat	cgggggccag	attgtacacg	ggagcggacg	taggccgcac	5820
accaacatac	cggttctcgt	accccgcgcc	cgtgtactcc	cctaccgtga	tcgaaagatt	5880
ctcaagcccc	gatgtagcaa	tcgcagcgtg	caacgaatac	ctatccagaa	attacccaac	5940
agtggcgtcg	taccagataa	cagatgaata	cgacgcatac	ttggacatgg	ttgacgggtc	6000
ggatagttgc	ttggacagtg	cgacattctg	cccgccgaag	ctccgtgtct	acccgaaaca	6060
tcattgcgtac	caccagccga	ctgtacgcag	tgccgtcccg	tcacccttcc	agaacacact	6120
acagaacgtg	ctagcggccg	ccaccaagag	aaactgcaac	gtcacgcaaa	tgcgagaact	6180
accacccatg	gactcggcag	tggtcaacgt	ggagtgtctc	aagcgtatg	cctgctccgg	6240
agaatatagg	gaagaatatg	ctaaacaacc	tatccggata	accactgaga	acatcactac	6300
ctatgtgacc	aaattgaaag	gcccgaagac	tgctgccttg	ttcgctaaga	cccacaactt	6360
ggttccgctg	caggagggtc	ccatggacag	attcacggtc	gacatgaaac	gagatgtcaa	6420
agtcactcca	gggacgaaac	acacagagga	aagacccaaa	gtccaggtaa	ttcaagcagc	6480
ggagccattg	gcgaccgctt	acctgtgcgg	catccacagg	gaattagtaa	ggagactaaa	6540
tgctgtgtta	cgccctaacc	tgacacacat	gtttgatatt	tcggccgaag	actttgacgc	6600
gatcatcgcc	tctcacttcc	acccaggaga	cccggttcta	gagacggaca	ttgcatcatt	6660
cgacaaaagc	caggacgact	ccttggctct	tacaggttta	atgatcctcg	aagatctagg	6720
ggtggatcag	tacctgctgg	acttgatcga	ggcagccttt	ggggaaatat	ccagctgtca	6780
cctaccaact	ggcacgcgtc	tcaagttcgg	agctatgatg	aaatcgggca	tgtttctgac	6840
tttgtttatt	aaactgtctt	tgaacatcac	catagcaagc	agggtactgg	agcagagact	6900
cactgactcc	gcctgtgcgg	ccttcatcgg	cgacgacaac	atcgttcacg	gagtgatctc	6960
cgacaagctg	atggcggaga	ggtgcgcgtc	gtgggtcaac	atggagggtga	agatcattga	7020
gcctgtcatg	ggcgaaaaac	ccccataatt	tttgtgggga	ttcatagtgt	ttgacagcgt	7080
cacacagacc	gcctgcctgt	tttcagaccc	acttaagcgc	ctgttcaagt	tggttaagcc	7140
gctaacagct	gaagacaagc	aggacgaaga	caggcgacga	gcactgagtg	acgaggttag	7200
caagtggttc	cggacaggtc	tgggggccga	actggagggt	gcactaacat	ctaggatatga	7260
ggtagagggc	tgcaaaagta	tcctcatagc	catggccacc	ttggcgaggg	acattaaggc	7320
gtttaagaaa	ttgagaggac	ctgtttataca	cctctacggc	ggtcctagat	tggtgcgtta	7380
atacacagaa	ttctgatttg	atcccaaacg	ggccctctag	actcgagcgg	ccgccactgt	7440
gctggatata	tgacagaattc	caccacactg	gactagtggg	tctatggcgt	acccatacga	7500
tgttccagat	tacgtatctc	tgagatctac	catgtctcag	agcaaccggg	agctggtggt	7560
tgactttctc	tcctacaagc	tttcccagaa	aggatacagc	tgagatcagt	ttagtatgtg	7620
ggaagagaa	aggactgagg	ccccagaagg	gactgaatcg	gagatggaga	ccccagtgcc	7680
catcaattgg	aaacctcctc	ggcacctggc	agacagcccc	gcgggtgaatg	gagccactgc	7740
gcacagcagc	agtttgtagt	cccgggaggt	gatccccatg	gcagcagtaa	agcaagcgct	7800
gaggggaggca	ggcgacgagt	ttgaactcgc	gtaccggcgg	gcattcagtg	acctgacatc	7860
ccagctccac	atcaccccag	ggacagcata	tcagagcttt	gaacaggtag	tgaatgaact	7920
cttccgggat	ggggtaaaact	gggttcgcgt	tgtggccttt	ttctccttcg	gcggggcact	7980
gtgcgtggaa	agcgtagaca	aggagatgca	ggtattggtg	agtcggatcg	cagcttggat	8040
ggccacttac	ctgaatgacc	acctagagcc	ttggatccag	gagaacggcg	gctgggatac	8100
ttttgtggaa	ctctatggga	acaatgcagc	agccgagagc	cgaaggggcc	aggaacgctt	8160
caaccgctgg	ttcctgacgg	gcatgactgt	ggccggcatg	gtttactctg	gctcactctt	8220
cagtccgaaa	tgaagatccg	agctcgttac	caagcttaag	tttggttaat	taattgaatt	8280
acatccctac	gcaaacgttt	tacggccgcg	ggtggcgcgc	gcgccggcgg	gcccgtcctt	8340
ggccgttgca	gtggctcccg	tggtctcccg	tcgtccccga	cttccaggcc	cagcagatgc	8400
agcaactcat	cagcgccgta	aatgcgctga	caatgagaca	gaacgcaatt	gctcctgcta	8460
ggcctcccaa	accaaaagag	aagaagacaa	ccaaaccaaa	gccgaaaacg	cagcccaaga	8520
agatcaacgg	aaaaacgcag	cagcaaaaaga	agaaaagacaa	gcaagccgac	aagaagaaga	8580
agaaacccgg	aaaaagagaa	agaatgtgca	tgaagattga	aaatgactgt	atcttcgtat	8640
gcggctagcc	acagtaacgt	agtgtttcca	gacatgtcgg	gcaccgcact	atcatgggtg	8700
cagaaaaatc	cggtgtgtct	gggggccttc	gcaatcgggc	ctatcctggt	gctggttggt	8760
gtcacttgca	ttgggtctcg	cagataagtt	aggttaggca	atggcattga	tatagcaaga	8820
aaattgaaaa	cagaaaaagt	tagggtaagc	aatggcatat	aaccataact	gtataacttg	8880
taacaaagcg	caacaagacc	tgcgcaattg	gccccgtggt	ccgcctcacg	gaaactcggg	8940

gcaactcata	ttgacacatt	aattggcaat	aattggaagc	ttacataagc	ttaattcgac	9000
gaataattgg	attttttattt	tatttttgcaa	ttggtttttta	atattttccaa	aaaaaaaaaa	9060
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	9120
agtgatcata	atcagccata	ccacatttgt	agaggtttta	cttgctttaa	aaaacctccc	9180
acacctcccc	ctgaacctga	aacataaaat	gaatgcaatt	gttggttgta	acttgtttat	9240
tgagcttat	aattggttaca	aataaagcaa	tagcatcaca	aatttcacaa	ataaagcatt	9300
tttttctactg	cattctagtt	gtggtttgtc	caaactcatc	aatgtatctt	atcatgtctg	9360
gatctagtct	gcattaatga	atcgccaac	gcgcggggag	aggcggtttg	cgtattgggc	9420
gctcttccgc	ttcctcgtc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	9480
tatcagctca	ctcaaaaggc	gtaatacgg	tatccacaga	atcaggggat	aacgcaggaa	9540
agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	9600
cgtttttcca	taggctcgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	9660
ggtggcgaaa	ccgcagacga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	9720
tgcgctctcc	tgttccgacc	ctgcgcgtta	ccggatacct	gtccgccttt	ctcccttcgg	9780
gaagcgtggc	gcttttctcaa	tgctcgcgt	gtaggtatct	cagttcgggtg	taggtcgttc	9840
gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	9900
gtactatcg	tcttgagtc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	9960
ctggtaacag	gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtgg	10020
ggcctaacta	cggctacact	agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	10080
ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	10140
gtggtttttt	tgtttgcaag	cagcagatta	cgcgagaaa	aaaaggatct	caagaagatc	10200
ctttgatctt	ttctacgggg	cattctgacg	ctcagtgga	cgaaaaactca	cgtttaaggga	10260
ttttggtcat	gagattatca	aaaaggatct	tcacctagat	cctttttaaat	taaaaatgaa	10320
gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	10380
tcagtgaagg	acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcc	10440
ccgtcgtgta	gataactacg	atacgggagg	gcttaccatc	tgccccag	gctgcaatga	10500
taccgcgaga	ccacgcgtca	ccggctccag	atttatcagc	aataaacacg	ccagccggaa	10560
gggcccagcg	cagaagtgg	cctgcaactt	tatccgcctc	catccagctc	attaattggt	10620
gcccgggaagc	tagagtaagt	agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	10680
ctacaggcat	cgtggtgtca	cgctcgtcgt	ttggtatggc	ttcattcagc	tccggttccc	10740
aacgatcaag	gcgagttaca	tgatccccc	tggttggtcaa	aaaagcggtt	agctccttcg	10800
gtcctccgat	cgttgctcaga	agtaagtgg	ccgcagtggt	atcactcatg	gttatggcag	10860
cactgcataa	ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	10920
actcaaccaa	gtcattctga	gaatagtgt	tgccggcgacc	gagttgctct	tgccccgctg	10980
caatacggga	taataccgcg	ccacatagca	gaacttttaa	agtgtctatc	attggaaaac	11040
gttcttccgg	gcgaaaaactc	tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	11100
ccactcgtgc	acccaactga	tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	11160
caaaaacagg	aaggcaaaat	gccgcaaaaa	agggaaataag	ggcgacacgg	aaatgttgaa	11220
tactcatact	cttctctttt	caatattatt	gaagcattta	tcagggttat	tgtctcatga	11280
gcgatacat	atttgaatgt	atttagaaaa	ataaacaaat	aggggttccg	cgcacatttc	11340
cccgaaaagt	gccacctgac	gtctaagaaa	ccattattat	catgacatta	acctataaaa	11400
ataggcgtat	cacgagggcc	tttctgctcg	cgcggttccg	tgatgacggt	gaaaacctct	11460
gacacatgca	gctcccgag	acgggtcacag	cttctgtcta	agcggatgcc	gggagcagac	11520
aagcccgtca	ggcgcgctca	gcgggtgttg	gcgggtgtcg	gggctggctt	aactatgcgg	11580
catcagagca	gattgtactg	agagtgcacc	atatcgacgc	tctcccttat	gcgactcctg	11640
cattaggaag	cagcccgata	ctagggttag	gcggttgagc	accgcgcgct	caagggaatg	11700
tgcatgcgta	atcaattacg	gggtcattag	ttcatagccc	atatatggag	ttccgcgtta	11760
cataacttac	ggtaaatggc	ccgcctggct	gaccgcccaa	cgaccccgcc	ccattgacgt	11820
caataatgac	glatgttccc	atagtaacgc	caatagggag	tttccattga	cgtcaatggg	11880
tgagtgattt	acggtaaaat	gcccaactgg	cagtaacatca	agtgatcat	atgccaaagta	11940
cgccccctat	tgacgtcaat	gacggtaaat	ggcccgccctg	gcattatgcc	cagtacatga	12000
ccttatggga	tttctctact	tggcagtaca	tctacgtatt	agtcacgct	attaccatgg	12060
tgatgcggtt	cttgcgctac	atcaatgggc	gtggatagcg	gtttgactca	cggggatttc	12120
caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	gcacccaaat	caacgggact	12180
ttccaaaatg	tcgtaaacac	tccgccccat	tgacgcaaat	gggcggttag	cgtgtacgg	12240
gggaggtcta	tataagcaga	gctctctggc	taactagaga	acccactgct	taactggctt	12300
atcgaaatta	atcagactca	ctatagggag	accggaagct	tgaatttc		12347

SEQ ID NO:66

atggcggatg	tgtgacatac	acgacgcca	aagattttgt	tccagctcct	gccacctccg	60
ctacgcgaga	gattaaccac	ccacgatggc	cgccaaagt	catgttgata	ttgaggctga	120
cagcccatc	atcaagtctt	tgagaaagg	atttccgctg	ttcgaggtgg	agtcattgca	180
ggtcacacca	aatgacctatg	caaatgccag	agcatttttcg	cacctggcta	ccaaattgat	240
cgagcaggag	actgacaaag	acacactcat	cttggatatac	ggcagtgccg	cttccaggag	300
aatgatgtct	acgcacaaat	accactgcgt	atgccctatg	cgacgcgcag	aagaccccg	360
aaggctcgat	agctacgcaa	agaaactggc	agcggcctcc	gggaaggtgc	tggtatagaga	420
gatcgaggga	aaaatcaccg	acctgcagac	cgtcatggct	acgccagacg	ctgaatctcc	480
taccttttgc	ctgcatacag	acgtcacgtg	tcgtacggca	gccgaagtgg	ccgtatacca	540
ggacgtgtat	gctgtacatg	caccaacatc	gctgtaccat	caggcgatga	aaggtgtcag	600
aacggcgtat	tgattgggtt	ttgacaccac	cccgtttatg	tttgacgcgc	tagcaggcgc	660
gtatcccaacc	tacgcccaaa	actggggcca	cgagcaggtg	ttacaggcca	ggaacatagg	720
actgtgtgca	gcatacctga	ctgagggga	actcggcaaa	ctgtccattc	tccgcaagaa	780
gcaattgaaa	ccttgcgaca	cagtcattgt	ctcggtagga	tctacattgt	acactgagag	840
cagaaagcta	ctgaggagct	ggcacttacc	ctccgtattc	cacctgaaag	gtaaaccaatc	900
ctttacctgt	aggtgcgata	ccatcgtatc	atgtgaagg	tacgtagtta	agaaaatcac	960
tatgtgcccc	ggcctgtacg	gtaaaacggt	agggtagccc	gtgacgtatc	acgcggagg	1020
attcctagt	tgcaagacca	cagacactgt	caaaggagaa	agagtctcat	tccctgtatg	1080
cacctacgtc	ccctcaacca	tctgtgatca	aatgactggc	atactagcca	ccgacgtcac	1140
accggaggac	gcacagaagt	tgtagtggtg	attgaatcag	aggatagtgt	tgaacggaag	1200
aacacagcga	aacactaaca	cgatgaagaa	ctatctgctt	ccgattgtgg	ccgtcgcatt	1260

tagcaagtgg	gcgaggggaat	acaaggcaga	ccttgatgat	gaaaaacctc	tgggtgtccg	1320
agagaggtca	cttacttgc	gctgcttgc	ggcattttaa	acgaggaaga	tgcacaccat	1380
gtacaagaaa	ccagacaccc	agacaatagt	gaaggtgcct	tcagagttaa	actcgttcgt	1440
catcccagagc	ctatggctca	caggccctcg	aatcccagtc	agatcacgca	ttaagatgct	1500
tttggccaag	aagaccaagc	gagagttaat	acctgttctc	gacgcgtcgt	cagccaggga	1560
tgctgaacaa	gaggagaagg	agaggttgga	ggccgagctg	actagagaag	ccttaccacc	1620
cctcgtcccc	atcgcgccgg	cggagacggg	agtcgtcgac	gtcgacgttg	aagaactaga	1680
gtatcacgca	ggtgcagggg	tcgtggaaac	acctcgagc	gcgttgaaag	tcaccgcaca	1740
gccgaacgac	gtactactag	gaaattacgt	agttctgtcc	ccgcagaccg	tgctcaagag	1800
ctccaagtgt	gccccctgtc	accctctagc	agagcaggtg	aaaataataa	cacataacgg	1860
gagggccggc	ggttaccagg	tcgacggata	tgacggcagg	gtcctactac	catgtggatc	1920
ggccattccg	gtccctgagt	ttcaagcttt	gagcgagagc	gccactatgg	tgtacaacga	1980
aaggggagttc	gtcaacaggga	aactatacca	tattgccgtt	cacggaccgt	cgctgaacac	2040
cgacgaggag	aactacgaga	aagtcagagc	tgaaagaact	gacgccgagt	acgtgttcga	2100
cgtagataaa	aaatgctgcg	tcaagagaga	ggaagcgctc	ggtttgtgtg	tgggtggaga	2160
gctaaccaac	cccccgcttc	atgaattcgc	ctacgaaggg	ctgaagatca	ggccgtcggc	2220
accatataag	actacagtag	taggagctct	tggggttccg	ggatcaggca	agtcgtctat	2280
tattaagagc	ctcgtgacca	aacacgatct	ggtcaccagc	ggcaagaagg	agaactgcca	2340
ggaaatagtt	aacgacgtga	agaagcaccg	cggggaaggg	acaagttagg	aaaacagtga	2400
ctccatcctg	ctaaacgggt	gtcgtcgtgc	cgtggacatc	ctatatgtgg	acgaggcttt	2460
cgctagccat	tccggtactc	tgctggccct	aattgctctt	gttaaaccct	ggagcaaagt	2520
ggtgttatgc	ggagacccca	agcaatgcgg	attcttcaat	atgatgcagc	ttaaggtgaa	2580
cttcaaccac	aacatctgca	ctgaagtatg	tcataaaagt	atatccagac	gttgcaacgc	2640
tcagtcacag	gccatctgtg	ctacgttgca	ctacggaggc	aagatgcgca	cgaccaaccc	2700
gtgcaacaaa	cccataatca	tagacaccac	aggacagacc	aagcccaagc	caggagacat	2760
cgtgttaaca	tgcttcggag	gctgggcaaa	gcagctgcag	ttggactacc	gtggacacga	2820
agtcatgaca	gcagcagcat	ctcagggcct	cacccgcaaa	gggggtatac	ccgttaaggca	2880
gaaggtgaat	gaaaatccct	tgtatgcccc	tgcgctcgag	cacgtgaatg	tactgtcgac	2940
gcgcactagag	gataggctgg	tgtggaaaaa	gctggccggc	gatccctgga	ttaaggtcct	3000
atcaaacatt	ccacagggtg	actttacggc	cacattggaa	gaatggcaag	aagaacacga	3060
caaaataatg	aaggtgattg	aaggaccggc	tgcgccctgt	gacgcgttcc	agaacaaagc	3120
gaacgtgtgt	tgggcgaaaa	gcctggtgcc	tgctcctggc	actgccggaa	tcagattgac	3180
agcagaggag	tgagacacca	taattacagc	atttaaggag	gacagagctt	actctccagt	3240
ggtggccctg	aatgaaatct	gcaccaagta	ctatggagtt	gacctggaca	gtggcctgtt	3300
ttctgccccg	aaggtgtccc	tgtattacga	gaacaaccac	tgggataaca	gacctgggtg	3360
aaggatgtat	ggaattcaatg	ccgcaacacg	tgccaggctg	gaagctagac	ataccttcc	3420
gaagggggcag	tggtacacgg	gcaagcaggc	agttatcgca	gaaagaaaaa	tccaaccgct	3480
ttctgtgtgt	gacaatgtaa	ttcctatcaa	ccgcaggtcg	ccgcacgccc	tggtggctga	3540
gtacaagacg	gttaaaggca	gtagggttga	gtggctggtc	aataaaagtaa	gagggtacca	3600
cgtcctgctg	gtgagttagt	acaacctggc	tttgccctga	cgcagggtca	cttggttgtc	3660
accgctgaat	gtcacaggcg	ccgataggtg	ctacgacctc	agtttaaggac	tgccggctga	3720
cgccggcgagg	ttcgacttgg	tctttgtgaa	cattcacacg	gaattcagaa	tccaccacta	3780
ccagcagtg	gtcgaccacg	ccatgaagct	gcagatgctt	gggggagatg	cgctacgact	3840
gctaaaaccc	ggcggcatct	tgatgagagc	ttacggatac	gccgataaaa	tcagcgaagc	3900
cgttgtttcc	tccttaagca	gaaagttctc	gtctgcaaga	gtgttgccgc	cggtattgtg	3960
caccagcaat	acagaagtg	tcttgctgtt	ctccaaacttt	gacaaacggaa	agagaccctc	4020
tacgctacac	cagatgaata	ccaagctgag	tgccgtgtat	gccggagaag	ccatgcacac	4080
ggccgggtgt	gcaccatcct	acagagttaa	gagagcagac	atagccacgt	gcacagaagc	4140
ggctgtggtt	aacgcagcta	acgcccgtgg	aactgtaggg	gatggcgatg	gcagggccgt	4200
ggcgaagaaa	tgcccgctag	ccttttaagg	agcagcaaca	ccagtgggcg	caattaaaaa	4260
agtcagtgtc	ggctcgtacc	ccgtcatcca	cgctgtagcg	cctaatttct	ctgccacgac	4320
tgaagcggaa	ggggaccggc	aattggccgc	tgtctacccg	gcagtgccgc	ccgaagtaaa	4380
cagactgtca	ctgacgacgc	tagccatccc	gctgctgtcc	acaggagtgt	tcagcggcgg	4440
aagagatagg	ctgcagcaat	ccctcaacca	tctattcaca	gcaatggacg	ccacggacgc	4500
tgacgtgacc	atctactgca	gagacaaaag	ttggggagaag	aaaatccagg	aagccattga	4560
catgaggacg	gctgtggagt	tgctcaatga	tgacgtggag	ctgaccacag	acttgggtgag	4620
agtgcaccgg	gacagcagct	tggtgggtcg	taagggtctac	agtaccactg	acgggtcgct	4680
gtactcgtac	tttgaaggtg	cgaaattcaa	ccaggctgct	attgatattg	cagagatact	4740
gacgttgttg	ccagactgac	aagaggcaaa	cgaacagata	tgccctatac	cgctggggcg	4800
aacaatggac	aacatcagat	ccaaatgtcc	ggtgaacgat	tccgattcat	caacacctcc	4860
caggacagtg	ccctgcctgt	gccgctacgc	aatgacagca	gaacggatcg	cccgccttag	4920
gtcacaccaa	gttaaaagca	tggttggtttg	ctcatctttt	cccctccoga	aataccatgt	4980
agatgggggtg	cagaaggtaa	agtgcgagaa	ggttctcctg	ttcgaccoga	cggtaccttc	5040
agtgggttagt	ccgcgggaagt	atgccgcctc	tacgacggac	cactcagatc	ggctggttac	5100
agggtttgac	ttggactgga	ccaccgactc	gtcttccact	gccagcgata	ccatgtcgct	5160
accagtttg	cagtcgtgtg	acatcgactc	gatctacgag	ccaatggctc	ccatagtagt	5220
gacggctgac	gtacaccctg	aaccgcagg	catcgccgac	ctggcggcag	atgtgcaccc	5280
tgaaccgcga	gaccatgtgg	acctcgagaa	cccgaattct	ccaccgcgc	cgaagagagc	5340
tgcataccct	gcctcccgcg	cggcggagcg	accggtgccg	gcgcgcgaaa	agccgacgcc	5400
tgcccccaagg	actgcgttta	ggaacaagct	gcctttgagc	ttcggcgact	ttgacgagca	5460
cgaggtcgat	cggttgccct	ccgggattac	tttcggagac	ttcgacgacg	tcttcgactc	5520
aggccgcgcg	ggtgcatata	ttttctcctc	ggacactggc	agcggacatt	tacaacaaaa	5580
atccgttagg	cagcacatc	tccagtcgcg	acaactggat	gcgggtccag	aggagaaaaa	5640
gtaccccgca	aaattggata	ctgagaggga	gaagctgttg	ctgctgaaaa	tcagatgca	5700
cccatcgagg	gctaataaga	gtcgatacca	gtctcgcaaa	gtggagaaca	tgaaagccac	5760
ggtggtggac	aggctcacat	cgggggcccag	attgtacacg	ggagcggacg	taggccgcat	5820
accaacatac	gcggttcggt	acccccgcc	cgtgtactcc	cctaccgtga	tcgaaaagatt	5880
ctcaagcccc	cagttagcaa	tcgcagcgtg	caacgaatac	ctatccagaa	attaccaaac	5940
agtgccgctg	taccagataa	cagatgaata	cgacgcatac	ttggacatgg	ttgacgggtc	6000
ggaatagttg	ttggacagag	cgacattctg	cccggcggaag	ctccggtgct	acccgaaaaa	6060

tcattgcgtac	caccagccga	ctgtacgcag	tgccgtcccg	tcaccctttc	agaacacact	6120
acagaaactg	ctagcggccg	ccaccaagag	aaactgcaac	gtcacgcaaa	tgcgagaact	6180
accacacatg	gactcggcag	tgttcaacgt	ggagtgcttc	aagcgctatg	cctgctcccg	6240
agaatattgg	gaagaatctg	ctaaacaacc	tatcccgata	accactgaga	acatcactac	6300
ctatgtgacc	aaattgaaag	gcccgaagc	tgctgccttg	ttcgctaaga	cccacaactt	6360
ggttccgctg	caggaggttc	ccatggacag	attcacggtc	gacatgaaac	gagatgtcaa	6420
agtcactcca	gggacgaaac	acacagagga	aagaccctaa	gtccaggtaa	ttcaagcagc	6480
ggagccattg	gcgaccgctt	acctgtgcgg	catccacagg	gaattagtaa	ggagactaaa	6540
tgctgtgtta	cgccctaacc	tgacacatt	gtttgatatg	tcggccgaag	actttgacgc	6600
gatcatcgcc	tctcacttcc	acccaggaga	cccggttcta	gagacggaca	ttgcatcatt	6660
cgacaaaagc	caggacgact	ccttggtctt	tacaggttta	atgatcctcg	aagatctagg	6720
ggtaggatcag	tacctgctgg	acttgatcga	ggcagccttt	ggggaaatat	ccagctgtca	6780
cctaccaact	ggcagcgctt	tcaagttcgg	agctatgatg	aaatcgggca	tgtttctgac	6840
tttgtttatt	aacactgttt	tgaacatcac	catagcaagc	agggtactgg	agcagagact	6900
cactgactcc	gcctgtgcgg	ccttcactcg	cgacgacaac	atcggtcacg	gagtgatctc	6960
cgacaagctg	atggcggaga	ggtgcgcgtc	gtgggtcaac	atggagggtga	agatcattga	7020
cgctgtcatg	ggcgaaaaac	ccccataatt	ttgtggggga	ttcatagtgt	ttgacagcgt	7080
cacacagacc	gcctgcggtg	tttcagaccc	acttaagcgc	ctgttcaagt	tgggtaagcc	7140
gctaacagct	gaagacaagc	aggacgaaga	caggcgacga	gcactgagtg	acgagggttag	7200
caagtggctc	cggacaggct	tgggggccga	actggagggtg	gcactaacat	ctagggtatga	7260
ggtagagggc	tgcaaaagta	tcctcatagc	catggccacc	ttggcgaggg	acattaaggc	7320
gtttaagaaa	ttgagaggac	ctgttatata	cctctacggc	ggctctagat	tgggtcggtta	7380
atacacagaa	ttctgattgg	atcccaaacg	ggccctctag	actcgagcgg	ccgccactgt	7440
gctggatata	tgagaaattc	atgcatggag	atcacactac	attgcatgaa	tatatgttag	7500
atttgcaacc	agagcaactc	gatctctact	gttatgagca	attaaatgac	agctcagagg	7560
aggaggatga	aatagatggt	ccagctggac	aagcagaacc	ggacagagcc	cattacaata	7620
ttgttaacct	ttgttgcaag	tgtgactcta	cgcttcgggt	gtgcgtacaa	agcacacacg	7680
tagacattcg	tactttggaa	gacctgttaa	tgggcacact	aggaattgtg	tgccccatct	7740
gttctcagaa	accaggatct	atggcgctacc	catacgatgt	tccagattac	gctagcttga	7800
gatctaccat	gtctcagagc	aaccgggagc	tggtggttga	ctttctctcc	tacaagcttt	7860
ccagaaaagg	atacagctgg	agtcagttta	gtgatgtgga	agagaacagg	actgaggccc	7920
cagaaggga	gatccagagc	atggagaccc	ccagtgcctat	caatggcaac	ccatcctggc	7980
acctggcaga	cagccccgcg	gtgaatggag	ccactgcgca	cagcagcagt	ttggatgccc	8040
gggagggtgat	cccataggca	gcagtaaaag	aagcgctgag	ggaggcaggc	gacgagtttg	8100
aactgcggtg	ccggcgggga	ttcagtgacc	tgacatccca	gctccacatc	accccaggga	8160
cagcatatca	gagctttgaa	caggtagtga	atgaactctt	ccgggatggg	gtaaactggg	8220
gtcgcatgtg	ggcctttttc	tccttcggcg	gggcactgtg	cgtagaaagc	gtagacaagg	8280
agatgcaggt	attggtgagt	cggtatcgag	cttggtatgg	cacttacctg	aatgaccacc	8340
tagagccttg	tcacagagag	aacggcggtt	gggatacttt	tgttgaactc	tatgggaaca	8400
atgcagcagc	cgagagccga	aaggggcagg	aacgcttcaa	ccgctggttc	ctgacgggca	8460
tgactgtggc	cggcggtggt	ctgctgggct	cactcttcag	tcggaaatga	agatccaagc	8520
ttaaagtttg	gtaattaatt	gaattacatc	cctacgcaaa	cgttttacgg	ccgccggtgg	8580
cgcccgccgc	ggcgcccgcc	tccttgcccg	ttgcaggcca	ctccggtggc	tcccgtcgtc	8640
cccgacttcc	aggccagca	gatgcagcaa	ctcatcagcg	ccgtaaatgc	gctgacaatg	8700
agacagaacg	caattgtctc	tgctaggcct	cccaaaccaa	agaagaagaa	gacaacccaa	8760
acaaagccga	aaacgcagcc	caagaagatc	aacgggaaaa	cgacagcaga	aaagaagaaa	8820
gacaagcaag	ccgacaagaa	gaagaagaaa	cccggaaaaa	gagaaagaat	gtgcatgaag	8880
attgaaaatg	actgtatctt	cgtagtcggc	tagccacagt	aacgtagtgt	ttccagacat	8940
gtcgggcacc	gcactatcat	gggtgcagaa	aatctcgggt	ggctctgggg	ccttcgcaat	9000
cgcgctatc	ctgggtgctg	ttgtggtcac	ttgcattggg	ctccgagat	aagttaggg	9060
aggcaatggc	attgatatag	caagaaaatt	gaaaacagaa	aaagttaggg	taagcaatgg	9120
catataacca	taactgtata	acttgtaaca	aagcgcaaca	agacctgcgc	aattggcccc	9180
gtggtccgcc	tcacggaaac	tcggggcaac	tcataattgac	acattaattg	gcaataattg	9240
gaagcttaca	taagcttaat	tcgacgaata	attggatttt	tattttattt	tgcaattgg	9300
ttttaataat	tcacaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	9360
aaaaaaaaaa	aaaaaaaaaa	aaactagtga	tcataatcag	ccataccaca	ttttagaggg	9420
ttttacttgc	tttaaaaaac	ctcccacacc	tccttcgtaa	cctgaacat	aaaatgaatg	9480
caattgttgt	tgtaactttg	tttattgcag	cttataatgg	ttacaaataa	agcaatagca	9540
tcacaaattt	cacaaataaa	gcattttttt	cactgcattc	tagttgtggt	ttgtccaaac	9600
tcactcaatg	atcttatcat	gtctgatat	agtctgcatt	aatgaatcgg	ccaacgcgcg	9660
gggagaggcg	gtttgcgtat	tgggcgctct	tccgcttctt	cgctcactga	ctcgctgcgc	9720
tcggtcgctt	ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	acggttatcc	9780
acagaaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	9840
aaccgtaaaa	aggccgcggt	gctggcggtt	ttccataggg	tcggccccc	tgacgagcat	9900
cacaaaaatc	gacgctcaag	tcagagggtg	cgaaaccoga	caggactata	aagataccag	9960
gcgtttcccc	ctggaagctc	cctcgctgcg	tctcctgttc	cgacctgccc	gcttacccga	10020
tacctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	ctcaatgctc	gcgctgtagg	10080
tatctcagtt	cgtgttaggt	cgttcgctcc	aagctgggct	gtgtgcacga	acccccggtt	10140
cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaaccc	ggtaagacac	10200
gacttatcgc	cactggcagc	agccactggt	aacaggatta	gcagagcgag	gtatgtaggc	10260
gggtctacag	agttcttgaa	gtggtggcct	aactacggct	acactagaag	gacagtattt	10320
ggatctctcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	10380
ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	gcaagcagca	gattacgcgc	10440
agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggcattc	tgacgctcag	10500
tggaacgaaa	actcacgtta	agggtatttt	gtcatgagat	tatcaaaaaa	gatcttcacc	10560
tagatccttt	taaatataaa	atgaagtttt	aaatcaatct	aaagtataata	tgagtaaaact	10620
tggtctgaca	gttaaccaatg	cttaatcagt	gaggcaccta	tctcagcgat	ctgtctatctt	10680
cgttcatcca	tagttgcctg	actcccgtgc	gtgtagataa	ctacgataacg	ggagggctta	10740
ccatctggcc	ccagtgctgc	aatgataccg	cgagaccacc	gctcaccggc	tccagattta	10800
tcagcaataa	accagccagc	cggaaggggc	gagcgcagaa	gtggctctgc	aactttatcc	10860

gcctccatcc	agtctattaa	ttgttgccgg	gaagctagag	taagtagttc	gccagttaat	10920
agtttgcgca	acgttggtgc	cattgctaca	ggcatcgtgg	tgtcacgctc	gtcgtttggg	10980
atggcttcat	tcagctccgg	ttcccaacga	tcaaggcgag	ttacatgata	ccccatggtg	11040
tgcaaaaaag	cggttagctc	cttcggtctc	ccgatcgttg	tcagaagtaa	gttgcccgca	11100
gtgttatcac	tcattggtat	ggcagcactg	cataattctc	ttactgtcat	gccatccgta	11160
agatgctttt	ctgtgactgg	tgagtactca	accaagtcac	tctgagaata	gtgtatgcgg	11220
cgaccgagtt	gctcttgccc	ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	11280
ttaaaaagtc	tcattcattgg	aaaacgttct	tcggggcgaa	aactctcaag	gatcttaccg	11340
ctgttgagat	ccagttcgat	gtaacccact	cgtgcaccca	actgatcttc	agcatctttt	11400
actttcacca	gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgcccg	aaaaaaggga	11460
ataaggggcg	acaggaaaatg	ttgaatactc	atactcttcc	tttttcaata	ttattgaagc	11520
atztatcagg	gttattgtct	catgagcgga	tacatatattg	aatgtattta	gaaaaataaa	11580
caaatagggg	ttccgcgcac	atttccccga	aaagtgccac	ctgacgtcta	agaaaccatt	11640
attatcatga	cattaacctta	taaaaatagg	cgtatcacga	ggccctttcg	tctcgcgcgt	11700
ttcgggtgatg	acggtgaaaa	ctcttgacac	atgcagctcc	cggagacggg	cacagcttct	11760
gtctaagcgg	atgccgggag	cagacaagcc	cgtcagggcg	cgtcagcggg	tgttgccggg	11820
tgtcgggggt	ggcttaacta	tgccgcatca	gagcagattg	tactgagagt	gcaccatata	11880
gacgtctctc	cttatgcgac	tcctgcatta	ggaagcagcc	cagtaactagg	ttgaggccgt	11940
tgagcaccgc	cgccgcaagg	aatggtgcat	gcgtaataca	ttacggggtc	attagtctcat	12000
agcccatata	tgaggttccg	cgttacataa	cttacggtaa	atggcccgcg	tggtctgaccg	12060
cccaacgacc	ccgcgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	12120
gggactttcc	atggagctca	atgggtggag	tattttacgg	aaactgcccc	cttgccagta	12180
catcaagtgt	atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaattggccc	12240
gcctggcatt	atgcccgata	catgacctta	tgggactttc	ctacttgcca	gtacatctac	12300
gtattagtca	tcgctattac	catgggtgatg	cgggtttggc	agtacatcaa	tgggcgtgga	12360
tagcgggtttg	actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	12420
ttttggcacc	aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	12480
caaatgggcg	gtaggcggtg	acgggtggag	gtctatataa	gcagagctct	ctggctaact	12540
agagaacca	ctgcttaact	ggcttatcga	aattaatacg	actcactata	gggagaccgg	12600
aagcttgaat	tc					12612

SEQ ID NO: 67

atggcgggatg	tgtgacatac	acgacgccaa	aagattttgt	tccagctcct	gccacctccg	60
ctacgcgaga	gattaaccac	ccacgatggc	cgccaaaagt	catgttgata	ttgaggctga	120
cagcccatc	atcaagtcct	tgcaagaagg	atttccgtcg	ttcgaggtgg	agtcattgca	180
gggtcacacca	aaatgccatg	caaatgccag	agcattttcg	cacctggcta	ccaaattgat	240
cgagcaggag	actgacaaa	acacactcat	cttgatatac	ggcagtgccg	cttccaggag	300
aatgatgtct	acgcacaaat	accactgcgt	atgccctatg	cgcagcgag	aagaccccca	360
aaggctcgat	agctacgcaa	agaaactggc	agcggcctcc	gggaagggtg	tggtatagaga	420
gatcgagga	aaaatcaccc	acctgcagac	cgtcatggct	acgcagagcg	ctgaatctcc	480
taccttttgc	ctgcatacag	acgtcacgtg	tcgtacggca	gccgaagtgg	ccgtatacca	540
ggacgtgtat	gctgtacatg	caccaacatc	gctgtacctt	caggcgatga	aaggtgtcag	600
aacggcgat	tgacacacac	ttgacaccac	cccgtttatg	tttgacgcgc	tagcaggcgc	660
gtatccaacc	tacgccacaa	actgggccga	cgagcagggt	ttacaggcca	ggaacatagg	720
actgtgtgca	gcatacctga	ctgagggaa	actcggcaaa	ctgtccattc	tccgcaagaa	780
gcaattgaaa	ccttgccgca	cagtcattgt	ctcggtagga	tctacattgt	acactgagag	840
cagaaagcta	ctgaggagct	ggcacttacc	ctccgtattc	cacctgaaag	gtaaacaatc	900
ctttacctgt	aggtgcgata	ccatcgtatc	atgtgaagg	tacgtagtta	agaaaaatcac	960
tatgtgcccc	ggcctgtacg	gtaaaacggt	agggtagcgc	gtgacgtatc	acgcggaggg	1020
attcctagt	tgcaagacca	cagacactgt	caaaggagaa	agagtctcat	tccctgtatg	1080
cacctacgtc	ccctcaacca	tctgtgatca	aatgactggc	atactagcga	ccgacgtcac	1140
accggaggac	gcacagaagt	tgttagtggg	attgaatcag	aggatagtgg	tgaacggaag	1200
aacacacgca	aactacaaca	cgatgaagaa	ctatctgctt	ccgattgtgg	ccgtccgatt	1260
tagcaagtgg	gcgagggaat	acaaggcaga	ccttgatgat	gaaaaacctc	tgggtgtccg	1320
agagaggtca	cttacttgct	gctgcttggt	ggcattttaa	acgagggaag	tgcacaccat	1380
gtacaagaaa	ccagacaccc	agacaatagt	gaagggtgct	tcagagttaa	actcgttctg	1440
catcccagc	ctatggtcta	caggccctcg	aatcccagtc	agatcacgca	ttaagatgct	1500
tttggccaag	aagaccaagc	gagagttaat	acctgttctc	gacgcgtcgt	cagccaggga	1560
tgtgaacaa	gaggagaagg	agagggttga	ggccgagctg	actagagaag	ccttaccacc	1620
cctcgtcccc	atcgcgcggg	cggagacggg	agtcgtcgac	gtcgcgtttg	aagaactaga	1680
gtatcacgca	ggtgcagggg	tcgtggaac	acctcgcagc	gcgttgaaag	tcaccgcaca	1740
gccgaacgac	gtactactag	gaaattacgt	agttctgtcc	ccgcagaccg	tgctcaagag	1800
ctccaagtgg	gcccccgctg	accctctagc	agagcagggt	aaaataataa	cacataacgg	1860
gagggccggc	ggttaccagg	tcgacggata	tgacggcagg	gtcctactac	catgtggatc	1920
ggccattccg	gtccctgagt	ttcaagcttt	gagcgagagc	gccactatgg	tgtacaacga	1980
aagggagttc	gtcaacagga	aactatacca	tattgccgtt	cacggaccgt	cgctgaacac	2040
cgacgaggag	aactacgaga	aagtcagagc	tgaaagaact	gacgcggagt	acgtgttcga	2100
cgtagataaa	aaatgctgcg	tcaagagaga	ggaagcgtcg	ggtttggtgt	tggtgggaga	2160
gctaaccaac	ccccggttcc	atgaattcgc	ctacgaagg	ctgaagatca	ggccgtcggc	2220
accatataag	actacagtat	taggagctct	tggggttccg	ggatcaggca	agtcgtctat	2280
tattaagagc	ctcgtgacca	aacacgatct	ggtcacccagc	ggcaagaagg	agaactgcca	2340
ggaaatagtt	aacgacgtga	agaagcaccg	cgggaagggg	acaagtaggg	aaaacagtga	2400
ctccatcctg	ctaaacgggt	gtcgtcgtgc	cgtggacatc	ctatatgtgg	acgaggcttt	2460
cgtagcccat	tcgggtactc	tgctggccct	aattgctctt	gttaaacctc	ggagcaaaat	2520
ggtgttatgc	ggagacccca	agcaatgcgg	attcttcaat	atgatgcagc	ttaagggtgaa	2580
cttcaaccac	aacatctgca	ctgaagtatg	tcataaaagt	atatccagac	gttgacgcgg	2640
tcagtcacg	gccatcgtgt	ctacgttgca	ctacggaggc	aagatgcgca	cgaccaaccc	2700
gtgcaacaaa	ccataatca	tagacaccac	aggacagacc	aagcccaagg	caggagacat	2760
cgtgttaaca	tgcttccgag	gctgggcaaa	gcagctgcag	ttggactacc	gtggacacga	2820
agtcagtaca	gcagcagcat	ctcagggcct	cacccgcaaa	gggttatagc	ccgtaaggca	2880

gaaggtgaat	gaaaaatccct	tgtatgcccc	tgcgtcggag	cacgtgaatg	tactgtctgac	2940
gcgcactgag	gataggctgg	tgtggaaaaac	gctggccggc	gatccctgga	ttaaggctcct	3000
atcaaacatt	ccacagggtg	acttttacggc	cacattggaa	gaatggcaag	aagaacacga	3060
caaaaataatg	aaggtgattg	aaggaccggc	tgcgcctgtg	gacgcgttcc	agaacaaagc	3120
gaacgtgtgt	tggggcaaaa	gcctgggtgcc	tgtcctggac	actgccggaa	tcagattgac	3180
agcagaggag	tggagcacca	taattacagc	atttaaggag	gacagagctt	actctccagt	3240
ggtggccttg	aatgaaatgt	gcaccaagta	ctatggagtt	gacctggaca	gtggcctgtt	3300
ttctgccccg	aaggtgtccc	tgtattacga	gaacaaccac	tgggataaca	gacctggtgg	3360
aaggatgtat	ggattcaatg	ccgcaacagc	tgcaggctg	gaagctagac	ataccttcct	3420
gaagggggcag	tggcatacgg	gcaagcaggc	agttatcgca	gaaagaaaaa	tccaaccgct	3480
ttctgtgctg	gacaatgtaa	ttcctatcaa	ccgcaggctg	ccgcacgccc	tgggtggtga	3540
gtacaagacg	gttaaaggca	gtagggttga	gtggctggtc	aataaagtaa	gaggggtacca	3600
cgtcctgtctg	gtgagtga	acaacctggc	tttgccctcga	cgccagggtca	cttggttctg	3660
accgctgaat	gtcacaggcg	ccgatagggtg	ctacgacctg	agtttaggac	tggcggctga	3720
cgccggcagg	ttcgacttgg	tctttgtgaa	cattcacacg	gaattcagaa	tccaccacta	3780
ccagcagtg	gtcgaccacg	ccatgaagct	gcagatgctt	gggggagatg	cgctacgact	3840
gtcaaaaacc	ggcggtcatc	tgatgagagc	ttacggatac	gccgataaaa	tcagcgaagc	3900
cgtgttttcc	tccttaagca	gaaagtttct	gtctgcaaga	gtgttgcgcc	cggattgtgt	3960
caccagcaat	acagaagtgt	tcttgctgtt	ctccaacttt	gacaacggaa	agagacccctc	4020
tacgctacac	cagatgaata	ccaagctgag	tgcctgttat	gccggagaag	ccatgcacac	4080
ggccgggtgt	gcaccatcct	acagagttaa	gagagcagac	atagccacgt	gcacagaagc	4140
ggctgtgggt	aacgcagcta	acgcccgtgg	aactgtaggg	gatggcgat	gcagggccgt	4200
ggcgaagaaa	tggcgtcag	cctttaaggg	agcagcaaca	ccagtgggca	caattaaaac	4260
agtcatgtgc	ggctcgtacc	ccgtcatcca	cgctgtagcg	cctaatttct	ctgccacgac	4320
tgaagcgaa	ggggcagcc	aattggccgc	tgtctaccgg	gcagtggccg	ccgaagtaaa	4380
cagactgtca	ctgagcagcg	tagccatccc	gctgctgtcc	acaggagtgt	tcagcggcgg	4440
aagagatagg	ctgcagcaat	ccctcaacca	tctattcaca	gcaatggacg	ccacggacgc	4500
tgacgtgacc	atctactgca	gagacaaaag	ttggggagaag	aaaatccagg	aagccattga	4560
catgaggacg	cgtgtggagt	tgtctcaatga	tgacgtggag	ctgaccacag	acttggtgag	4620
agtgcacccg	gacagcagcc	tgggtgggtcg	taagggctac	agtaccactg	acgggtcgct	4680
gtactcgtac	tttgaaggtg	cgaatttcaa	ccaggctgct	attgatattg	cagagatact	4740
gacgttgttg	ccagactgc	aagaggcaaa	cgaacagata	tgcctatacg	cgctggcgca	4800
aacaatggac	aacatcagat	ccaaatgtcc	ggtgaacgat	tccgattcat	caacacctcc	4860
caggacagtg	ccctgcccgt	gccgctacgc	aatgacagca	gaacggatcg	cccgccttag	4920
gtcacaccaa	gttaaaagca	tgggtggttg	ctcatctttt	cccctcccga	aataccatgt	4980
agatgggggtg	cagaaggtaa	agtcgagaa	ggtttctctg	ttcgaccgca	cggtaaccttc	5040
agtgtttagt	ccgcggaaag	atgccgcac	tacgacggac	cactcagatc	ggtcgttacg	5100
agggtttgac	ttggactgga	ccaccgactc	gtcttccact	gccagcgata	ccatgtcgct	5160
accagttttg	cagtcgtgtg	acatcgactc	gatctacgag	ccaatggctc	ccatagtagt	5220
gacggctgac	gtacacccctg	aaaccgcagg	catcgcgagc	ctggcggcag	atgtgcaccc	5280
tgaacccgca	gaccatgtgg	acctcgagaa	cccgattcct	ccaccgcgcc	cgaagagagc	5340
tgcatacctt	gcctcccggc	cgccggagcg	accggtgccc	gcgccgagaa	agccgacgcc	5400
tgcaccaagg	actgcgttta	ggaacaagct	gcctttgacg	ttcgggcact	ttgacgagca	5460
cgaggtcgat	gcgttggcct	ccgggattac	tttcggagac	ttcgacgacg	tcttgcgact	5520
agggccgcgcg	ggtgcatata	ttttctcttc	ggacactggc	agcggacatt	tacaacaaaa	5580
atccgttagg	cagcacaatc	tccagtgcgc	acaactggat	gcgggtccag	aggagaaaaat	5640
gtacccgcca	aaattggata	ctgagaggga	gaagctgttg	ctgctgaaaa	tgcagatgca	5700
cccatcggag	gctaataaga	gtcgatacca	gtctcgcaaa	gtggagaaca	tgaagccac	5760
ggtggtggac	aggctcacat	cgggggccag	attgtacacg	ggagcggacg	taggcccgat	5820
accaacatac	cggttctggg	accccgcgcc	cgtgtactcc	cctaccgtga	tcgaaagatt	5880
ctcaagcccc	gatgtagcaa	tcgcagcgtg	caacgaatac	ctatccagaa	attaccaaac	5940
agtggcgtcg	taccagataa	cagatgaata	cgacgcatac	ttggacatgg	ttgacgggtc	6000
ggatgtgtgc	ttggacagag	cgacattctg	cccgccgaag	ctccggtgct	accggaaca	6060
tcatgcttac	caccagccga	ctgtacgcag	tgcctgccg	tcaccttttc	agaacacact	6120
acagaacgtg	ctagcggcgg	ccaccaagag	aaactgcaac	gtcacgcaaa	tgcgagaact	6180
accacccatg	gactcggcag	tgttcaacgt	ggagtgtctc	aagcgtatg	cctgctccgg	6240
agaatatattg	gaagaatatg	ctaaacaacc	tatccggata	accactgaga	acatcactac	6300
ctatgtgacc	aaattgaaag	gccccaaaag	tgttgccttg	ttcgctaaga	cccacaactt	6360
ggttccgctg	caggagggtt	ccatggacag	attcacggtc	gacatgaaac	gagatgtcaa	6420
agtcactcca	gggacgaaac	acacagagga	aagacccaaa	gtccaggtaa	ttcaagcagc	6480
ggagccattg	gcgaccgctt	acctgtgcgg	catccacagg	gaattagtaa	ggagactaaa	6540
tgtctgtgtta	cgccctaacg	tgcacacatt	gtttgatattg	tggccgaag	actttgacgc	6600
gatcatcgcc	tctcacttcc	acccaggaga	cccggttcta	gagacggaca	ttgcatcatt	6660
cgacaaaagc	caggacgact	ccttggctct	tacaggttta	atgatcctcg	aagatctagg	6720
ggtggatcag	tacctgctgg	acttgatcga	ggcagccttt	ggggaaatat	ccagctgtca	6780
cttaccacact	ggcacgcgct	tcaagttcgg	agctatgatg	aaatcgggca	tgtttctgac	6840
tttgtttatt	aacactgttt	tgaacatcac	catagcaagc	agggtagctg	agcagagact	6900
actgactccc	gctgtgcgg	ccttcatcgg	cgacgacaa	atcgttcacg	gagtgatctc	6960
cgacaagctg	atggcggaga	ggtgcgcgtc	gtgggtcaac	atggaggtga	agatcattga	7020
cgcgtgtcatg	ggcgaaaaac	ccccataatt	tttgtgggga	ttcatagtgt	ttgacagcgt	7080
cacacagacc	gctgcgctg	tttcagaccc	acttaagcgc	ctgttcaagt	tgggtaaagc	7140
gctaacagct	gaagacaagc	aggacgaaga	caggcgacga	gcactgagtg	acgaggttag	7200
caagtgggtc	cggacagcgt	tggggggcca	actggaggtg	gcactaacat	ctaggtatga	7260
ggtagagggc	tgcaaaagta	tctcatagc	catggccacc	ttggcgaggg	acattaaggg	7320
gtttaagaaa	ttgagaggac	ctgtttataca	cctctacggc	ggtcctagat	tgggtgcgtta	7380
atacacagaa	ttctgattgg	atcccaaacg	ggccctctag	actcgagcgg	ccgccactgt	7440
gctggatatac	tgcagaatct	caccacactg	gactagtggg	tctatggcgt	acccatacga	7500
tgttccagat	tacgtatctc	tgagatctac	catgtctcag	agcaaccggg	agctggtggg	7560
tgacttttctc	tcttacaagc	tttcccagaa	aggatacagc	tggagttagt	ttagtgtatgt	7620
ggaagagaac	aggactgagg	ccccagaagg	gactgaatcg	gagatggaga	ccccacgtgc	7680

catcaatggc	aacccatcct	ggcacctggc	agacagcccc	gcggtgaatg	gagccactgc	7740
gcacagcagc	agtttgatg	cccgaggagt	gatccccatg	gcagcagtaa	agcaagcgct	7800
gagggaggca	ggcgacgagt	ttgaactgcg	gtaccggcgg	gcattcagtg	acctgacatc	7860
ccagctccac	atcaccccag	ggacagcata	tcagagcttt	gaacaggtag	tgaatgaact	7920
cttccgggat	ggggtagcca	ttcttcgcat	tgtggccttt	ttctccttcg	gcggggcact	7980
gtgcgtggaa	agcgtagaca	aggagatgca	ggtattgggtg	agtcggatcg	cagcttggtg	8040
ggccacttac	ctgaatgacc	acctagagcc	ttggatccag	gagaacggcg	gctgggatac	8100
gatccaacgg	acaaacgcag	acaaatgcagc	agccgagagc	cgaaaaggcc	aggaacgctt	8160
caacgcgtgg	ttcctgacgg	gcatgactgt	ggccggcgctg	gttctgctgg	gctcactctt	8220
cagtcggaaa	tgaagatccg	agctcggtag	caagcttaag	tttgggtaat	taattgaatt	8280
acatccctac	gc aaacgctt	tacggccgcc	ggtggcgccc	gcgcccggcg	gcccgtcctt	8340
ggccggttga	ggccactccg	gtggctcccc	tcgtccccga	cttcaggccc	cagcagatgc	8400
agcaactcat	cagcgccgta	aatgcgctga	caatgagaca	gaacgcaatt	gctcctgcta	8460
ggcctcccaa	accaaagaag	aagaagacaa	ccaaaccaa	gccgaaaacg	cagcccaaga	8520
agatccaacg	aaaaacgcag	cagcaaaaga	agaaagacaa	gcaagccgac	aagaagaaga	8580
agaaccccg	aaaaagagaa	agaatgtgca	tgaagattga	aaatgactgt	atcttcgtat	8640
gcggctagcc	acagtaacgt	agtgtttcca	gacatgtcgg	gcaccgcact	atcatgggtg	8700
cagaaaatct	cggttggtct	gggggccttc	gcaatcggcg	ctatcctggt	gctggttggtg	8760
gtcacttgca	ttgggtcccg	cagataagtt	agggtagcca	atggcattga	tatagcaaga	8820
aaattgaaaa	cagaaaaagt	tagggtaagc	aatggcatat	aaccataact	gtataacttg	8880
taacaaagcg	caacaagacc	tgcgcaattg	gccccgtggt	ccgcctcacg	gaaactcggg	8940
gcaactcata	ttgacacatt	aattggcaat	aattggaagc	ttacataagc	ttaatcgcac	9000
gaataattgg	atttttat	tattttgcaa	ttggttttta	atattttcaa	aaaaaaaaaa	9060
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	9120
agtgatcata	atcagccata	ccacatttgt	agaggtttta	cttgctttta	aaaacctccc	9180
acacctcccc	ctgaacctga	aacataaaat	gaatgcaatt	gttgttggtta	acttgtttat	9240
tgacgcttat	aatggttaca	aataaagcaa	tagcatcaca	aatttcacaa	ataaagcatt	9300
tttttactg	cattctagtt	gtggtttgtc	caaactcatc	aatgtatctt	atcatgtctg	9360
gatctagctt	gcattaatga	atcgcccaac	gcgcggggag	aggcggtttg	cgtattggcg	9420
gctcttcgcg	ttcctcgctc	actgactcgc	tgcgctcggt	cggttcggctg	cggcgagcgg	9480
tatcagctca	ctcaaaaggc	gtaatacgtt	tatccacaga	atcaggggat	aacgcaggaa	9540
agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaacc	taaaaaaggc	gcgttgctg	9600
cgtttttcca	taggtctcgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	9660
ggtggcgaaa	cccgacagga	ctataaaagt	accaggcggt	tccccctgga	agctccctcg	9720
tgcgctctcc	tggtccgacc	ctgcgcgtta	ccggataacct	gtccgccttt	ctcccttcgg	9780
gaagcggtgc	gctttctcaa	tgtctcgctc	gtaggtatct	cagttcggtg	taggtcgttc	9840
gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	9900
gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	9960
ctggtaaacg	gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtggg	10020
ggcctaacta	cggtacactc	agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	10080
ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	10140
gtggtttttt	tggttgcaag	cagcagatta	cgcgacagaa	aaaaggatct	caagaagatc	10200
ctttgatctt	ttctacgggg	cattctgacg	ctcagtgga	cgaaaactca	cgttaaggga	10260
ttttggtcat	gagattatca	aaaaggatct	tcacctagat	cctttttaat	taaaaaatgaa	10320
gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	10380
tcagtgaggc	acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcc	10440
ccgtcgtgta	gataactacg	atacgggag	gcttaccatc	tgccccaggt	gctgcaatga	10500
taccgcgaga	ccacgcgtca	ccggctccag	atttatcagc	aataaacacg	ccagccggaa	10560
gggcccagcg	cagaagtggg	cctgcaactt	tatccgcctc	catccagctc	attaattggt	10620
gcccgggaagc	tagagttaag	agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	10680
ctacaggcat	cggtgtgtca	cgctcgtcgt	ttggtatggc	ttcattcagc	tccggttccc	10740
aacgatcaag	gcgagttaca	tgatccccca	tggttggtcaa	aaaagcggtt	agctccttcg	10800
gtcctccgat	cggtgtcaga	agtaagtggg	ccgcagtggt	atcactcatg	gttatggcag	10860
cactgcataa	ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	10920
actcaaccaa	gtcattctga	gaatagtgtg	tgccggcgacc	gagttgctct	tgccccggct	10980
caatacggga	taataccgcg	ccacatagca	gaacttttaa	agtgtctatc	attggaaaaa	11040
gttcttcggg	gcgaaaaactc	tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	11100
ccactcgtgc	acccaactga	tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	11160
aaaaaacagg	aaggcaaaat	gccgcaaaaa	aggggaataag	ggcgacacgg	aaatgttgaa	11220
tactcatact	cttctttttt	caatattatt	gaagcattta	tcagggttat	tgtctcatga	11280
gcggatacat	atttgaatgt	atttagaaaa	ataaacaaat	aggggttccg	cgcacatttc	11340
cccgaaaagt	gccacctgac	gtctaagaaa	ccattattat	catgacatta	acctataaaa	11400
ataggcgtat	cacgaggccc	tttctgtctg	cgcggttccg	tgatgacggt	gaaaacctct	11460
gacacatgca	gctcccggag	acgggtcacag	cttctgtcta	agcggatgcc	gggagcagac	11520
aagcccgtca	ggcgcgctca	gcgggtgttg	gcgggtgtcg	gggctggctt	aactatgcgg	11580
catcagagca	gattgtactg	agagtgcacc	atatcgacgc	tctcccttat	gcgactcctg	11640
cattaggaag	cagcccagta	ctaggttgag	gcggttgagc	accgcccgcg	caaggaaatg	11700
tgcatgcgta	atcaattacg	gggtcattag	ttcatagccc	atatatggag	ttccgcgtta	11760
cataacttac	ggtaaatggc	ccgcctggct	gaccgcccaa	cgacccccgc	ccattgacgt	11820
caataatgac	gtagtttccc	atagtaacgc	caatagggac	tttccattga	cgctcaatgg	11880
tgaggtattt	acggtaaaact	gcccacttgg	cagtagacatc	agtgtatcat	atgccaaagta	11940
cgccccctat	tgcagtcatt	gacggtaaat	ggcccgcctg	gcattatgcc	cagtacatga	12000
ccttatggga	tttctctact	tggcagtaca	tctacgtatt	agtcacgctt	attaccatgg	12060
tgatgcggtt	tttgcagtag	atcaatgggc	gtggatagcg	gtttgactca	cggggatttc	12120
caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	gcacccaaat	caacgggact	12180
ttccaaaatg	tcgtaaacac	tccgccccat	tgacgcaaat	gggcggttag	cgtgtacggg	12240
gggaggtcta	tataagcaga	gctctctggc	taactagaga	acccactgct	taactggctt	12300
atcgaaatta	atcagcatca	ctatagggag	accggaagct	tgaatttc		12347

SEQ ID NO:68

atggcggatg	tgtgacatac	acgacgccaa	aagattttgt	tccagctcct	gccacctccg	60
ctacgcgaga	gattaaccac	ccacgatggc	cgccaaagt	catgttgata	ttgaggctga	120
cagcccattc	atcaagcttt	tgacgaaggc	atttccgtcg	ttcagaggtg	agtcattgca	180
ggtcacacaa	aatgaccaca	caaatgccag	agcattttcg	cacctggcta	ccaaattgat	240
cgagcaggag	actgacaaag	acacactcat	cttggatata	ggcagtgccg	cttccaggag	300
aatgatgtct	acgcacaaat	accactgcgt	atgccctatg	cgacgcgcag	aagaccccga	360
aaggctcgat	agctacgcaa	agaaactggc	agcggcctcc	gggaaggtgc	tgatagaga	420
gatcgacgga	aaaatcaccc	acctgcagac	cgctcatggc	acgccagacg	ctgaattctc	480
taccttttgc	ctgcatacag	acgtcacgtg	tcgtacggca	gccgaagtgg	ccgtatacca	540
ggacgtgtat	gctgtacatg	caccaacatc	gctgtaccat	caggcgatga	aagggtgcag	600
aacggcgctat	tgattgggtt	ttgacaccac	cccgtttatg	tttgacgcgc	tagcaggcgc	660
gtatccaacc	tacgccacaa	actgggcccga	cgagcagggt	ttacaggcca	ggaacatagg	720
actgtgtgca	gcatacctga	ctgagggaag	actcggcaaa	ctgtccattc	tccgcaagaa	780
gcaattgaaa	ccttgcgaca	cagtcattgt	ctcggtagga	tctacattgt	acactgagag	840
cagaaagcta	ctgaggagct	ggcacttacc	ctccgtattc	cacctgaaag	gtaaaccaatc	900
ctttacctgt	aggtgcgata	ccatcgtatc	atgtgaagg	tacgtagtta	agaaaaatcac	960
tatgtgcccc	ggcctgtacg	gtaaaaacgt	agggtagcgc	gtgacgtatc	acgcggagg	1020
attcctagt	tgcaaggaca	cagacactgt	caaaggagaa	agagtctcat	tccctgtatg	1080
cacctacgtc	ccctcaacca	tctgtgatca	aatgactggc	atactagcga	ccgacgtcac	1140
accggaggac	gcacagaagt	tgtagtgagg	attgaatcag	aggatagttg	tgaacggaag	1200
aacacacgca	aacactaaca	cgatgaagaa	ctatctgctt	ccgattgtgg	ccgtcgcatt	1260
tagcaagtg	ctgagggaat	acaaggcaga	ccttgatgat	gaaaaacctc	tggtgttccg	1320
agagaggtca	cttacttgct	gctgcttggt	ggcattttaa	acgaggaaga	tgcacaccat	1380
gtacaagaaa	ccagacaccc	agacaatagt	gaaggtgcct	tcagagttaa	actcgttcgt	1440
catcccagc	ctatgggtca	caggcctcgc	aatcccagtc	agatcacgca	ttaagatgct	1500
tttggccaag	aagaccaagc	gagagttaat	acctgttctc	gacgcgtcgt	cagccaggga	1560
tgctgaacaa	gaggagaagg	agagggttga	ggccgagctg	actagagaag	ccttaccacc	1620
cctcgtcccc	atcgcgccgg	cggagacggg	agtcgtcgac	gtcgacgttg	agaactaga	1680
gtatcacgca	ctgagggagt	tcgtggaaac	acctcgagc	gcgttgaaag	tcaccgcaca	1740
gccgaacgac	gtactactag	gaaattacgt	agttctgtcc	ccgcagaccg	tgctcaagag	1800
ctccaagtgt	gccccctg	accctctagc	agagcagggt	aaaataataa	cacataacgg	1860
gagggcggcg	ggctaccagg	tcgacggata	tgacggcagg	gtcctactac	catgtggatc	1920
ggccattccg	gtccctgagt	ttcaagcttt	gagcgagagc	gccactatgg	tgtacaacga	1980
aagggagttc	gtcaacagga	aactatacca	tattgccgtt	cacggaccgt	cgctgaacac	2040
cgacgaggag	aactacgaga	aagtcagagc	tgaaagaact	gacgccgagt	acgtgttcga	2100
cgtagataaa	aaatgctcgc	tcaagagaga	ggaagcgtcg	ggtttgggtg	tggtgggaga	2160
gctaaccaac	ccccgttcc	atgaattcgc	ctacgaagg	ctgaagatca	ggccgtcggc	2220
accatataag	actacagttag	taggagctct	tggtgttccg	ggatcaggca	agtcgtctat	2280
tattaagagc	ctcgtgacca	aacacgatct	ggcaccagc	ggcaagaagg	agaactgcca	2340
ggaatatgtt	aacgacgtga	agaagcaccg	cggtgaagg	acaagttagg	aaaacagtga	2400
ctccatccctg	ctaaacgggt	gtcgtcgtgc	cgtggacatc	ctatatgtgg	acgaggcttt	2460
cgctagccat	tccggtactc	tgctggccct	aattgtctct	gttaaaccctc	ggagcaaagt	2520
ggtgtttatg	ggagacccca	agcaatgcgc	attcttcaat	atgatgcagc	ttaaggtgaa	2580
cttcaaccac	aacatctgca	ctgaagtatg	tcataaaagt	atatccagac	gttgacgcgc	2640
tcagtcacg	gccatcgtgt	ctacgttgca	ctacggagc	aagatgcgca	cgaccaaccc	2700
gtgcaacaaa	cccataatca	tagacaccac	aggacagacc	aagcccaagc	caggagacat	2760
cgtgttaaca	tgcttccgag	gctgggcaaa	gcagctgcag	ttggactacc	gtggacacga	2820
agtcatgaca	gcagcagcat	ctcagggcct	caccgcgcaa	gggtgtatcg	ccgttaaggca	2880
gaaggtgaat	gaaaatccct	tgtatgcccc	tgctgcggag	cacgtgaatg	tactgtgcag	2940
gcgcactgag	gtgagctggg	ttgggaaaac	gctggccggc	gatccctgga	ttaaggtcct	3000
atcaaacatt	ccacagggtg	actttacggc	cacatttgaa	gaatggcaag	aagaacacga	3060
caaaataatg	aaggtgattg	aaggaccggc	tgccgctgtg	gacgcgttcc	agaacaaaag	3120
gaacgtgtgt	tgggcgaaaa	gcctgtgtcc	tgctcctggc	actgccggaa	tcagattgac	3180
agcagaggag	tgagacacca	taattacagc	atttaaggag	gacagagctt	actctccagt	3240
ggtggccttg	aatgaaattt	gcaccaagta	ctatggagtt	gacctggaca	gtggcctgtt	3300
ttctgccccg	aaggtgtccc	tgtattacga	gaacaaccac	tggtgataca	gacctgggtg	3360
aaggatgtat	gatttcaatg	ccgcaacagc	tgccaggctg	gaagctagac	ataccttctt	3420
gaaggggacg	tggtatcagg	gcaagcaggc	agttatcgca	gaaagaaaaa	tccaacccgt	3480
ttctgtgtgt	gacaatgtaa	ttcctatcaa	ccgcaggctg	ccgcacgccc	tggtggctga	3540
gtacaagacg	gttaaaggca	gtagggttga	gtggctgggt	aataaaagta	gagggtacca	3600
cgctcctgtg	gtgagttagt	acaacctggc	tttgccctga	cgacgggtca	cttggtgtgc	3660
accgctgaat	gtcacaggcg	ccgataggtg	ctacgacctc	agttttaggac	tgccggtctg	3720
cgccggcagg	ttcgacttgg	tctttgtgaa	catttcacag	gaattcagaa	tccaccacta	3780
ccagcagtg	gtcgacgtcg	ccatgaagct	gcagatgctt	gggggagatg	cgctacgact	3840
gctaaaaccc	ggcggtcatc	tgatgagagc	ttacggatac	gccgataaaa	tcagcgaagc	3900
cgttgtttcc	tccttaagca	gaaagtcttc	gtctgcaaga	gtgttgccgc	cggattgtgt	3960
caccagcaat	acagaagtgt	tcttgcgtgt	ctccaaactt	gacaacggaa	agagaccctc	4020
tacgctacac	cagatgaata	ccaagctgag	tgccgtgtat	gccggagaag	ccatgcacac	4080
ggccgggtgt	gcaccatcct	acagagttaa	gagagcagac	atagccacct	gcacagaagc	4140
ggctgtggtt	aacgcagcta	acgcccgtgg	aactgtagg	gatggcgtat	gcagggccgt	4200
ggcgaagaaa	tgcccgtagc	cctttaaagg	agcagcaaca	ccagtggtga	caattaaaac	4260
agtcatgtgc	ggctcgtacc	ccgtcatcca	cgctgtagcg	cctaatttct	ctgccacgac	4320
tgaagcggaa	ggggaccggc	aattggccgc	tgtctaccgg	gcagtgccgc	ccgaagtaaa	4380
cagactgtca	ctgagcagcg	tagccatccc	gctgctgtcc	acaggagtgt	tcagcggcgg	4440
aagagatagg	ctgcagcaat	ccctcaacca	tctattcaca	gcaatggacg	ccacggacgc	4500
tgacgtgacc	atctactgca	gagacaaaag	ttgggagaag	aaaatccagg	aagccattga	4560
catgaggacg	gctgtggagt	tgctcaatga	tgacgtggag	ctgaccacag	acttggtgag	4620
agtgcacccg	gcagcagcgt	tggtgggtcg	taagggtctac	agtaccactg	acgggtcgtc	4680
gtactcgtac	tttgaaggtg	cgaaattcaa	ccaggctgct	attgatattg	cagagatact	4740
gacgttgtgt	ccagactgac	aagaggcaaa	cgaacagata	tgccctatag	cgctggggca	4800

aacaatggac	aacatcagat	ccaaatgtcc	ggtgaacgat	tccgattcat	caacacctcc	4860
caggacagtg	ccttcgctgt	gccgctacgc	aatgacagca	gaacggatcg	cccgccttag	4920
gtcacaccaa	gttaaaagca	tgggtggttg	ctcatctttt	cccctcccga	aataccatgt	4980
agatgggggtg	cagaaggtaa	agtgcgagaa	ggttctcctg	ttcgaccoga	cggtaaccttc	5040
agtgtttagt	ccgcggaagt	atgccgcata	tacgacggac	cactcagatc	ggtcgtttacg	5100
agggttttgac	ttggactgga	ccaccgactc	gtcttccact	gccagcgata	ccatgtcgct	5160
accagtttg	cagtcgtgtg	acatcgactc	gatctacgag	ccaatggctc	ccatagtagt	5220
gacggctgac	gtacaccctg	aaccgcgagg	catcgcgagc	ctggcgcgag	atgtgcaccc	5280
tgaacccgca	gaccatgtgg	acctcgagaa	cccgatctct	ccaccgcgcc	cgaagagagc	5340
tgcataacct	gcttccccgcg	cggcgggagcg	accgggtgcg	gcgcgcgagaa	agccgacgcc	5400
tgccccaaagg	actgcgttta	ggaacaagct	gcctttgacg	ttcggcgact	ttgacgagca	5460
cgaggctgat	gcgttggcct	ccgggattac	tttcggagac	ttcgacgacg	tcctgcgact	5520
aggccgcgcg	ggtgcatata	ttttctcctc	ggacactggc	agcggacatt	tacaacaaaa	5580
atccggttagg	cagcacaata	tccagtgcgc	acaactggat	gcgggtccagg	aggagaaaaat	5640
gtaccccgca	aaattggata	ctgagaggga	gaagctgttg	ctgctgaaaa	tgcatatgca	5700
cccatcggag	gctaataaga	gtcgatacca	gtctcgcaaa	gtggagaaaca	tgaagccac	5760
ggtggtggac	aggctcacat	cggggggcag	attgtacacg	ggagcggagc	tagggcgcat	5820
accaacatac	gcggttcggt	acccccgcc	cgtgtacttc	cctaccgtga	tcgaaagatt	5880
ctcaagcccc	gatgtagcaa	tcgcagcgtg	caacgaatac	ctatccagaa	attacccaac	5940
agtggcgtcg	taccagataa	cagatgaata	cgacgcatac	ttggacatgg	ttgacgggtc	6000
ggatagttgc	ttggacagag	cgacattctg	cccggcgagg	ctccggtgct	acccgaaaca	6060
tcatgcgtac	caccagcgca	ctgtacgcag	tgccgtcccg	tcaccctttc	agaacacact	6120
acagaacgtg	ctagcggccg	ccaccaagag	aaactgcaac	gtcacgcaaa	tgcgagaact	6180
accacccatg	gactcggcag	tgttcaacgt	ggagtgcttc	aagcgctatg	cctgctccgg	6240
agaatatattg	gaagacatgt	ctaaacaacc	tatccggata	accactgaga	acatcactac	6300
ctatgtgacc	aaattgaaag	gccccgaaagc	tgctgccttg	ttcgtaaga	cccacaactt	6360
ggttccgctg	caggagggttc	ccatggacag	attcacggtc	gacatgaaac	gagatgtcaa	6420
agtcactcca	gggacgaaac	acacagagga	aagacccaaa	gtccaggtaa	ttcaagcagc	6480
ggagccattg	gcgaccgctt	acctgtgcgg	catccacagc	gaattagtaa	ggagactaaa	6540
tgctgtgtta	cgccctaaccg	tgacacacatt	gtttgatatg	tcggccgaag	actttgacgc	6600
gatcatcgcc	tctcacttcc	acccaggaga	cccggttcta	gagacggaca	ttgcatcatt	6660
cgacaaaagc	caggacgact	ccttggtctc	tacaggttta	atgatcctcg	aagatctagg	6720
ggtggatcag	tacctgctgg	acttgatcga	ggcagccttt	ggggaataat	ccagctgtca	6780
cctaccaact	ggcacgcgct	tcaagttcgg	agctatgatg	aaatcgggca	tgtttctgac	6840
tttgtttatt	aacactgttt	tgaacatcac	catagcaagc	agggtactgg	agcagagact	6900
cactgactcc	gcctgtgcgg	ccttcacg	cgacgacaa	atcggtcacg	gagtgatctc	6960
cgacaagctg	atggcggaga	ggtgcgcgtc	gtgggtcaac	atggagggtga	agatcattga	7020
cgctgtcatg	ggcgaaaaac	ccccataatt	ttgtggggga	ttcatagttt	ttgacagcgt	7080
caacacagacc	gcctgcgctg	tttcagaccc	acttaagcgc	ctgttcaagt	tgggttaagcc	7140
gctaacagct	gaagacaagc	aggacgaaga	caggcgacga	gcactgagtg	acgagggttag	7200
caagtggttc	cggacagcgt	tggggggccga	actggagggtg	gcactaacat	ctaggatatga	7260
ggtagagggc	tgcaaaagta	tcctcatagc	catggccacc	ttggcgaggg	acattaaggc	7320
gtttaagaaa	ttgagaggac	ctgttataca	cctctacggc	ggctctagat	tggtgcgtta	7380
atacacagaa	ttctgattgg	atcccaaacg	ggccctctag	actcgagcgg	ccgccactgt	7440
gctggatata	tgagaaatc	atgcatggag	atacacctac	attgcatgaa	tatatgttag	7500
atttgcaacc	agagacaact	gatctctact	gttatgagca	attaaatgac	agctcagagg	7560
aggaggatga	aatagatggt	ccagctggac	aagcagaacc	ggacagagcc	cattacaata	7620
ttgtaacctt	ttgttgcaag	tgtgactcta	cgcttcgggt	gtgcgtacaa	agcacacacg	7680
tagacattcg	tactttggaa	gacctgttaa	tgggcacact	aggaattgtg	tgccccatct	7740
gttctcagaa	accaggatct	atggcgctacc	catacgatgt	tccagattac	gctagcttga	7800
gatctaccat	gtctcagagc	aacggggagc	tggtggttga	ctttctctcc	tacaagcttt	7860
cccagaaaag	atacagctgg	agtcagttta	gtgatgtgga	agagaacagg	actgaggccc	7920
cagaaggagc	tgaatcgag	atggagaccc	ccagtgccat	caatggcaac	ccatcctggc	7980
acctggcaga	cagccccgcg	gtgaatggag	ccactgcgca	cagcagcagt	ttggatgcc	8040
gggagggtgat	ccccatggca	gcagtaaaagc	aagcgctgag	ggaggcaggc	gacgagtttg	8100
aactgcggtg	ccggcgggga	ttcagtgacc	tgacatccca	gctccacatc	accccaggga	8160
cagcatatca	gagctttgaa	caggtagtga	atgaactctt	ccgggatggg	gtagccattc	8220
ttcgcatgtg	ggcctttttc	tccttcggcg	gggcactgtg	cgtggaagc	gtagacaagg	8280
agatgcaggt	attggtgagt	cggatcgag	cttgatggc	cacttaacctg	aatgaccacc	8340
tagagccttg	gatccaggag	aacggcgct	gggatacttt	tgtggaactc	tatgggaaca	8400
atgcagcagc	cgagagccga	aaggggcagg	aacgcttcaa	ccgctggctc	ctgacgggca	8460
tgactgtggc	cggcgtggtt	ctgctgggct	cactcttcag	tcggaaatga	agatccaagc	8520
ttaagtttg	gtaattaatt	gaattacatc	cctacgcaaa	cgttttacgg	ccgcgggttg	8580
cgcgcgcgc	cgcgggcccg	tccttggccg	ttgcaggcca	ctccggtggc	tcctgtcgtc	8640
cccgacttcc	aggcccagca	gatgcagcaa	ctcatcagcg	ccgtaaatgc	gctgacaatg	8700
agacagaacg	caattgtctc	tgctaggcct	cccaaaccaa	agaagaagaa	gacaacccaa	8760
ccaaagccga	aaacgcagcc	caagaagatc	aacggaaaaa	cgacgacgca	aaagaagaaa	8820
gacaagcaag	ccgacaagaa	gaagaagaaa	cccgaaaaaa	gagaaagaat	gtgcatgaag	8880
attgaaaaatg	actgtatctt	cgtatgcggc	tagccacagt	aacgtagtgt	ttccagacat	8940
gtcgggcacc	gcactatcat	gggtgcagaa	aatctcgggt	ggtctggggg	ccttcgcaat	9000
cgcgctatc	ctggtgctgg	ttgtggtcac	ttgcattggg	ctccgagat	aagttaggg	9060
aggcaatggc	attgatatag	caagaaaatt	gaaaacagaa	aaagttaggg	taagcaatgg	9120
catataacca	taactgtata	acttgtaaca	aagcgcaaca	agacctgcgc	aattggcccc	9180
gtggtccgcc	tcacggaaac	tcggggcaac	tcatattgac	acattaattg	gcaataattg	9240
gaagcttaca	taagcttaat	tcgacgaata	attggatttt	tatttttatt	tgcaattgg	9300
ttttaatat	ttcaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	9360
aaaaaaaaaa	aaaaaaaaaa	aaactagtga	tcataatcag	ccataaccaca	ttttagaggg	9420
ttttactg	tttaaaaaac	ctccccacac	tcctcgtaa	cctgaaacat	aaaatgaatg	9480
caattgtgtg	tgtaacttg	tttattgcag	cttataatgg	ttacaaataa	agcaatagca	9540
tcacaaattt	cacaaataaa	gcattttttt	cactgcattc	tagttgtggt	ttgtccaaac	9600

tcataaatgt	atcttatcat	gtctggatct	agtctgcatt	aatgaatcgg	ccaacgcgcg	9660
gggagaggcg	gtttgcgtat	tgggcgctct	tccgcttcct	cgctcactga	ctcgcctgcgc	9720
tcgggtcggtc	ggctgcggcg	agcggatca	gctcactcaa	aggcggtaat	acgggttatcc	9780
acagaatcac	gggataacgc	aggaaagaa	atgtgagcaa	aaggccagca	aaaggccagg	9840
aaccgtaaaa	aggccgcgtt	gtggcggttt	ttccataggc	tccgcccccc	tgacgagcat	9900
cacaaaaatc	gacgctcaag	tcagagggtg	cgaaaccgga	caggactata	aagataccag	9960
gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	gcttaccgga	10020
tacgtgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcaatgctc	gcgctgtagg	10080
tatctcagtt	cgggtgtagt	cgttcgctcc	aagctgggct	gtgtgcacga	accccccggt	10140
cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaaccc	ggtaagacac	10200
gacttatcgc	cactggcagc	agccactggt	aacaggatta	gcagagcgag	gtatgtaggc	10260
ggtgctacag	agtctctgaa	gtggtggcct	aactacggct	acactagaag	gacagtattt	10320
ggtatctcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	10380
ggcaacaaaa	ccaccgctgg	tagcgggtgt	ttttttgttt	gcaagcagca	gattacgcgc	10440
agaaaaaaag	gactctcaaga	agatcctttg	atctttttcta	cggggcattc	tgacgctcag	10500
tggagcgaag	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	10560
tagatccttt	taaatataaa	atgaagtttt	aaatcaatct	aaagtataata	tgagtaaaact	10620
tggctctgaca	gttaccaatg	cttaaatcagt	gaggcaccta	tctcagcgat	ctgtctattt	10680
cgttcatcca	tagttgctcg	actccccgtc	gtgtagataa	ctacgatacg	ggagggctta	10740
ccatctggcc	ccagtgtctg	aatgataccg	cgagaccac	gctcaccggc	tccagattta	10800
tcagcaataa	accagccagc	cggaaggggc	gagcgcagaa	gtggtcctgc	aactttatcc	10860
gcctccatcc	gttactattaa	ttgttgccgg	gaagctagag	taagttagttc	gccagttaat	10920
agtttgcgca	acgttggtgc	cattgctaca	ggcatcgtgg	tgtaacgctc	gtcgtttggt	10980
atggctttcat	tcagctccgg	ttcccaacga	tcaaggcgag	ttacatgatc	ccccatgttg	11040
tgcaaaaaag	cgggttagctc	cttcggtctc	ccgatcgttg	tcagaagtaa	gttgcccgca	11100
gtgttatcac	tcattggttat	ggcagcactg	cataattctc	ttactgtcat	gccatccgta	11160
agatgctttt	ctgtgactgg	tgagtactca	accaagtcat	tctgagaata	gtgtatgcgg	11220
gcaccgagtt	gctcttgccc	ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	11280
ttaaaaagtc	tcactattgg	aaaacgttct	tcggggcgaa	aactctcaag	gatcttaccg	11340
ctgttgagat	ccagttcgat	gtaacccact	cgtgcaccca	actgatcttc	agcatctttt	11400
actttcacca	gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgcgcg	aaaaaaggga	11460
ataaggcgca	cacggaaaatg	ttgaatactc	atactcttcc	tttttcaata	ttattgaagc	11520
atttatcagg	gttattgtct	catgagcgga	tacatatattg	aatgtattta	gaaaaataaa	11580
caaatagggg	ttccgcgcac	atttccccga	aaagtgccac	ctgacgtcta	agaaaccatt	11640
attatcatga	cattaacctta	taaaaatagg	cgtatcacga	ggccctttcg	tctcgcgcgt	11700
ttcgggtgatg	acgggtgaaa	cctctgacac	atgcagctcc	cggagacggg	cacagcttct	11760
gtctaagcgg	atgccgggag	cagacaagcc	cgtcaggggc	cgtcagcggg	tggtggcggg	11820
tgtcggggct	ggcttaacta	tgccggatca	gagcagattg	tactgagagt	gcaccatata	11880
gacgtctctc	cttatgcgac	tcttgcat	ggaagcagcg	cagtaactagg	ttgaggccgt	11940
tgagcaccgc	cgccgcaagg	aatggtgcat	gcgtaatcaa	ttacgggggtc	attagtctcat	12000
agcccatata	tgaggttccg	cgttacataa	cttacggtaa	atggcccgcc	tggtgaccg	12060
cccaacgacc	cccgcaccat	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	12120
gggaactttcc	attgacgtca	atgggtggag	tattttacgg	aaactgcccc	cttggcagta	12180
catcaagtgt	atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaagtggccc	12240
gcctggcatt	atgcccagta	catgacctta	tgggactttc	ctacttgcca	gtacatctac	12300
gtattagtca	tcgtctattac	catgggtgatg	cgggttttggc	agtacatcaa	tgggcgtgga	12360
tagcgggtttg	actcacgggg	atttccaaat	ctccacccca	ttgacgtcaa	tgggagtttg	12420
ttttggcacc	aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	12480
caaatgggcg	gtaggcggtg	acgggtggag	gtctatataa	gcagagctct	ctggctaact	12540
agagaacca	ctgcttaact	ggcttatcga	aattaatac	actcactata	gggagaccgg	12600
aagcttgaat	tc					12612

SEQ ID NO: 69

gtcgaacttct	gaggcggaag	gaaccagctg	tggaaatgtg	gtcagttagg	gtgtggaaag	60
tccccagggt	ccccagcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	120
agggtgtgga	agtccccagg	ctccccagca	ggcagaagta	tgcaaaagcat	gcatctcaat	180
tagtcagcaa	cccatagctc	gcccctaact	ccgcctatcc	cgcccttaac	tccgcccagt	240
tccgcccatt	ctccgcccga	tggtgacta	atttttttta	tttatgcaga	ggccgaggcc	300
gcctcggcct	ctgagctatt	ccagaagtag	tgaggaggct	tttttgagg	cctaggcttt	360
tgcaaaaaag	tgatcgatc	ctgagaactt	cagggtgagt	ttggggaccc	ttgattgttc	420
tttctttttc	gctattgtaa	aattcatgtt	atatggaggg	ggcaaatgtt	tcagggtgtt	480
gtttagaatg	ggaagatgtc	ccttgatcca	ccatggaccc	tcatgataat	tttgtttctt	540
tcacttttcta	ctctgttgac	aaccattgtc	tctctttatt	ttcttttcat	tttctgtaac	600
tttttcgtta	aacttttagct	tgcatttgta	acgaattttt	aaattcactt	ttgtttattt	660
gtcagattgt	aagtactttt	tctaatactt	tttttttcaa	ggcaatcagg	gtatattata	720
ttgtacttca	gcacagtttt	agagaacaa	tggtataatt	aaatgataag	gtagaatatt	780
tctgcataata	aattctggct	ggcgtggaaa	tattcttatt	ggtagaaaca	actacatcct	840
ggtcatcatc	ctgcctttct	ctttatgggt	acaatgatat	acactgtttg	agatgaggat	900
aaaatactct	gagtccaac	cgggcccctc	tgctaaccat	gttcatgcct	tcttcttttt	960
cctacagctc	ctgggcaacg	tgctggttat	tgtgtgtgtc	catcattttg	gcaaaagaatt	1020
gtaatacgac	tcactatagg	gcgaattcgg	atccagatct	atggcgtaac	catacagatg	1080
tccagattac	gctagcttga	gatctaccat	gtctcagagc	aaccgggagc	tggtggttga	1140
ctttctctcc	tacaagcttt	cccagaaagg	atacagctgg	agtcagttta	gtgagtgtga	1200
agagaacagg	actgagggcc	cagaagggac	tgaatcggag	atggagaccc	ccagtgccat	1260
caatggcaac	ccatcctggc	acctggcaga	cagccccggc	gtgaatggag	ccactgcgca	1320
cagcagcagt	ttggatgccc	gggaggtgat	ccccatggca	gcagtaaaag	aagcgtgag	1380
ggaggcaggc	gacgagtttg	aactgcggta	cgggcgggca	ttcagtgacc	tgacatccca	1440
gctccacatc	acccaggga	cagcatatca	gagctttgaa	caggtagtga	atgaactctt	1500
ccgggatggg	gtaaaactgg	gtcgcattgt	ggcctttttc	tccttcggcg	gggcactgtg	1560
cgtggaaaag	gtagacaagg	agatgcaggt	attggtgagt	cggatcgcag	cttgatggc	1620

cacttacctg	aatgaccacc	tagagccttg	gatccaggag	aacggcggct	gggatacttt	1680
tgtggaactc	tatgggaaca	atgcagcagc	cgagagccga	aagggccagg	aacgcttcaa	1740
ccgctgggtc	ctgacgggca	tgactgtggc	cggcgtgggt	ctgctgggct	cactcttcag	1800
tcggaaatga	agatcttatt	aaagcagaaa	ttgtttattg	cagcttataa	tggttacaaa	1860
taaagcaata	gcatcacaaa	tttcacaaat	aaagcatttt	tttactgca	ttctagtgtg	1920
ggtttgtcca	aactcatcaa	tgtatcttat	catgtctggt	cgactctaga	ctcttccgct	1980
tctctgctca	ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	2040
tcgaaggcgg	taatacgggt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	2100
gcaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	ttttttccat	2160
aggctccgcc	ccctgacga	gcatcacaaa	aatcgacgct	caagtacagag	gtggcgaaac	2220
ccgacaggac	tataaagata	ccaggcgttt	ccccctggaa	gctccctcgt	gcgctctcct	2280
gttccgaccc	tgcgccttac	cggatacctg	tccgcctttc	tcccttcggg	aagcgtggcg	2340
ctttctcaat	gctcacgctg	taggtatctc	agttcgggtg	aggctcgttcg	ctccaagctg	2400
ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	2460
cttgagtcca	acccggttaag	acacgactta	tcgccactgg	cagcagccac	tggtaacagg	2520
attagcagag	cgagggtatgt	aggcgggtgt	acagagtctt	tgaagtgggtg	gcctaactac	2580
ggctacacta	gaaggacagt	atttgggtatc	tgcgctctgc	tgaagccagt	taccttcgga	2640
aaaagagttg	gtagctctctg	atccggcaaaa	caaacaccgc	ctggtagcgtg	tggttttttt	2700
gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	2760
tctacggggt	ctgacgctca	gtggaacgaa	aactcacgtt	aagggatttt	ggtcatgaga	2820
ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	2880
taaaagtata	atagtaaaac	ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	2940
atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	gactccccgt	cgtgtagata	3000
actacgatac	gggagggctt	accatctggc	cccagtgctg	caatgatacc	gcgagaccga	3060
cgtccaccgg	ctccagattt	atcagcaata	aaccagccag	ccggaaggcg	cgagcgacga	3120
agtggctctg	caactttatc	cgctccatc	cagtctatta	attgttgccg	ggaagctaga	3180
gtaagtagtt	cgccagttaa	tagtttgcgc	aacggtgttg	ccattgctac	aggcatcgtg	3240
gtgtcacgct	cgtcggttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	3300
gttacatgat	ccccctggtt	gtgcaaaaaa	gcgggttagct	ccttcggtcc	tccgatcgtt	3360
gtcagaagta	agtggccgcg	agtgttatca	ctcatggtta	tggcagcact	gcataattct	3420
cttactgtca	tgcctatcgt	aagatgcttt	tctgtgactg	gtgagtactc	aaccaagtca	3480
ttctgagaat	agtgtatgcg	gcgaccgaat	tgctcttgcc	cggcgctcaat	acgggataat	3540
accgcgccac	atagcagaac	tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	3600
aaactctcaa	ggatcttacc	gctgttgaga	tccagttcga	tgtaacccac	tctgtcaccc	3660
aactgatctt	cagcatcttt	tactttcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	3720
caaaatgccg	caaaaaaggg	aataaaggcg	acacggaat	gttgaatact	catactcttc	3780
ttttttcaat	attattgaag	catttatcag	ggttattgtc	tcatgagcgg	atacatattt	3840
gaatgtattt	agaaaaataa	acaaataggg	gttccgcgca	catttccccc	aaaagtgccca	3900
ccagcgtct	aagaaaccat	tattatcatg	acattaaact	ataaaaatag	gcgtatcacg	3960
aggccccctt	cgtctcgcgc	gtttcgggtg	tgacggtgaa	aacctctgac	acatgcagct	4020
cccggagacg	gtcacagctt	gtctgtaagc	ggatgccggg	agcagacaag	cccgtcaggg	4080
cgcgtcagcg	ggtgttgccg	ggtgtcgggg	ctggcttaac	tatgcggcat	cagagcagat	4140
tgtaactaga	gtgcaccata	tgcgggtgga	aataccgcac	agatgcgtaa	ggagaaaaata	4200
ccgcatcagg	aaattgtaaa	cgtaatatatt	ttgttaaaat	tcgcgttaaa	tttttgttaa	4260
atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	tcccttataa	atcaaaagaa	4320
tagacagaga	taggttttga	gtttggtcca	gtttggaaca	agagtccact	attaaagaac	4380
gtggactcca	acgtcaaaag	gcgaaaaaac	gtctatcagg	gcgatggccc	actacgtgaa	4440
ccatcacccct	aatcaagttt	tttggggctg	aggtgccgta	aagcactaaa	tccgaaaccct	4500
aaaggagacc	cccgatattag	agcttgacgg	ggaaagccgg	cgaacgtggc	gagaaaggaa	4560
gggaagaaag	cgaaaggagc	ggcgctagg	gcgctggcaa	gtgtagcggg	cacgctgcgc	4620
gtaaccacca	cacccgccgc	gcttaatgcg	ccgctacagg	gcgctgcg	ccattcgcca	4680
ttcaggctac	gcaactgttg	ggaagggcga	tcgggtcggg	cctcttcgct	attacgccag	4740
ctggcgaaag	ggggatgtgc	tgcaaggcga	ttaagttggg	taacgccagg	gttttcccag	4800
tcacgacgtt	gtaaaacgac	ggccagtgaa	tt			4832

SEQ ID NO:70

gtcgacttct	gaggcgga	gaaccagctg	tggaaatgtg	gtcagttagg	gtgtggaaag	60
tcccagggtc	cccagcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	120
agggtgtgga	agtccccagg	ctcccagca	ggcagaagta	tgcaaaagcat	gcatctcaat	180
tagtcagcaa	ccatagtccc	gccctaact	ccgcccatac	cgcccctaac	tccgcccagt	240
tccgcccatt	ctccgcccga	tggctgacta	atttttttta	tttatgcaga	ggccgaggcc	300
gcctcgccct	ctgagctatt	ccagaagtag	tgaggaggct	tttttggagg	cctaggcttt	360
tgcaaaaagc	tggatcgatc	ctgagaactt	cagggtgagt	ttggggaccc	ttgattgttc	420
tttctttttc	gctattgtaa	aattcatggt	atatggaggg	ggcaaaagttt	tcagggtggt	480
gtttagaatg	ggaagatgtc	ccttgatca	ccatggaccc	tcatgataat	tttgtttctt	540
tcacttttcta	ctctgttgac	aaccattgtc	tctctttatt	ttcttttcat	tttctgtaac	600
tttttcgtta	aacttttagt	tgcatttgta	acgaattttt	aaattcactt	ttgtttattt	660
gtcagattgt	aagtactttc	tctaatactt	tttttttcaa	ggcaatcagg	gtataattata	720
ttgtacttca	gcacagtttt	agagaacaat	tgttataaatt	aatgataag	gtagaatatt	780
tctgcataca	aattctggct	ggcgtggaaa	tattcttatt	ggtagaaaca	actacatcct	840
ggctcatcatc	ctgcctttct	ctttatggtt	acaatgatatt	acactgtttt	agatgaggat	900
aaaatactct	gagtcacaac	cgggcccctc	tgctaaccat	gttcatgcct	tcttcttttt	960
cctacagctc	ctgggcaacg	tgctggttat	tgtgtgtgtc	catcattttt	gcaaaagaatt	1020
gtaatacgac	tcactatagg	gcgaattcgg	atccagatct	atggcgctac	catacagatg	1080
tccagattac	gctagcttga	gatctaccat	gtctcagagc	aacggggagc	tgggtggttga	1140
ctttctctcc	tacaagcttt	cccagaaaag	atacagctgg	agtcagttta	gtgatgtgga	1200
agagaacagg	actgaggccc	cagaagggac	tgaatcggag	atggagaccc	ccagtgccat	1260
caatggcaac	ccatcctggc	acctggcaga	cagcccgcg	gtgaatggag	ccactgcgca	1320
cagcagcagt	ttggatgccc	gggaggtgat	ccccatggca	gcagtaaaag	aagcgtgag	1380
ggaggcaggc	gacgagtttg	aactgcggtg	ccggcgggca	ttcagtgacc	tgacatccca	1440

gctccacatc	accccagggg	cagcatatca	gagctttgaa	caggtagtga	atgaactctt	1500
ccgggatggg	gtagccattc	ttcgcattgt	ggcctttttc	tccttcggcg	gggcactgtg	1560
cgtggaaagc	gtagacaagg	agatgcagg	attgggtgag	cggatcgag	cttggtatgg	1620
cacttacctg	aatgaccacc	tagagccttg	gatccaggag	aacggcggt	gggatacttt	1680
tgtggaactc	tatgggaaca	atgcagcagc	cgagagccga	aagggccagg	aacgcttcaa	1740
ccgctgggtc	ctgacgggca	tgactgtggc	cggcggtggt	ctgctgggct	cactcttcag	1800
tcggaaatga	agatcttatt	aaagcagaac	ttgtttattg	cagcttataa	tggttacaaa	1860
taaaagcaata	gcatcacaaa	tttcacaaat	aaagcatttt	tttactgca	ttctagtgtg	1920
ggtttgtcca	aactcatcaa	tgtatcttat	catgtctggt	cgactctaga	ctcttcgct	1980
tcctcgctca	ctgactcgct	gcgctcggtc	gttcgggtgc	ggcgagcggt	atcagctcac	2040
tcaaaggcgg	taatacgggt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	2100
gcaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	ttttttccat	2160
aggctccgcc	cccctgacga	gcatcacaaa	aatcgacgct	caagtcagag	gtggcgaaac	2220
ccgacaggac	tataaagata	ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	2280
gttccgacct	tgccgcttac	cggataacctg	tcgccttttc	tccttcgggg	aagcgtggcg	2340
ctttctcaat	gctcacgctg	taggtatctc	agttcgggtg	aggctcgttcg	ctccaagctg	2400
ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	2460
cttgagtcca	acccggtaag	acacgactta	tcgccactgg	cagcagccac	tggtaacagg	2520
attagcagag	cgaggtagtg	aggcggtgct	acagagtctt	tgaagtgggtg	gcctaactac	2580
ggctacacta	gaaggacagt	atttggtatc	tgcgctctgc	tgaagccagt	taccttcgga	2640
aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	ctggtagcgg	tggttttttt	2700
gtttgcaagc	acgagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	2760
tctacggggt	ctgacgctca	gtggaacgaa	aactcacgtt	aagggatttt	ggtcatgaga	2820
ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	2880
taaatgtata	attagttaaa	ttgggtctgac	agttaccaat	gcttaatcag	tgaggcacct	2940
atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	gactccccgt	cgtgtagata	3000
actacgatac	gggaggggct	accatctggc	cccagtgctg	caatgatacc	gcgagaccca	3060
cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	ccggaagggc	cgagcgcaga	3120
agtgttctctg	caactttatc	cgcctccatc	cagtctatta	attgttgccg	ggaagctaga	3180
gtaagtagtt	cgccagttaa	tagtttgccg	aacgttggtg	ccattgtctac	aggcatcggtg	3240
gtgtcacgct	cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	3300
gttacatgat	ccccgatgtt	gtgcacaaaa	gcggttagct	ccttcggtcc	tcgcatcggt	3360
gtcagaagta	agttggccgc	agtgttatca	ctcatgggtt	tggcagcact	gcataattct	3420
cttactgtca	tgccatccgt	aaagtgtctt	tctgtgactg	gtgagtactc	aaccaagtca	3480
ttctgagaat	agtgtatgcg	gcgaccgagt	tgtctctgce	cggcgtcaat	acgggataat	3540
accgcgccac	atagcagaac	tttaaaaagt	ctcatcattg	gaaaacgttc	ttcggggcga	3600
aaactctcaa	ggatcttacc	gctgttgaga	tccagttcga	tgtaacccac	tcgtgcaccc	3660
aaactgatct	cagcatcttt	tactttcacc	agcgtttctg	ggtgagcaaa	aacagggaag	3720
caaaatgccg	caaaaaaggc	aataaggcgc	acacggaaat	gttgaatact	catactcttc	3780
ttttttcaat	attattgaag	catttatcag	ggttatgtgc	tcatgagcgg	atacatattt	3840
gaatgtatct	agaaaaataa	acaaataggg	gttccgcgca	catttccccc	aaaagtgcga	3900
cttgacgtct	aagaaaccat	tattatcatg	acattaaact	ataaaaaatg	gcgtatcacg	3960
aggccctctt	gctctcgcg	gtttcgggtg	tgacgggtgaa	aacctctgac	acatgcagct	4020
cccgagacg	gtcacagctt	gtctgtaagc	ggatgccggg	agcagacaag	cccgtcaggg	4080
cgcgtcagcg	gggtgtggcg	gggtgtcggg	ctggcttaac	tatgcggcat	cagagcagat	4140
tgtactgaga	gtgcaccata	tgccgtgtga	aataccgcac	agatgcgtaa	ggagaaaaata	4200
ccgcatcagg	aaattgtaaa	cgtaaatatt	ttgttaaaat	tcgcgttaaa	tttttgttaa	4260
atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	tccttataaa	atcaaaagaa	4320
tagaccgaga	taggggttag	tgttgttcca	gtttggaaca	agagtccact	attaaagaac	4380
gtggactcca	acgtcaaaag	gcgaaaaaac	gtctatcagg	gcgatggccc	actacgtgaa	4440
ccatcacccct	aatcaagttt	tttggggctg	agggtccgta	aagcactaaa	tcggaaccct	4500
aaaagggagcc	ccgatttag	agcttgacgg	ggaaagccgg	cgaacgtggc	gagaaaggaa	4560
gggaagaagc	cgaaaggtag	ggcgctagg	gcgctggcaa	gtgtagcgct	cacgctgcgc	4620
gtaaccacca	cacccgcgcg	gcttaatgcg	ccgctacagg	gcgcgtcgcg	ccattcgcca	4680
ttcaggctac	gcaactgttg	ggaaggcgga	tcgggtcggg	cctcttcgct	attacgccag	4740
ctggcggaag	ggggatgtgc	tgcaaggcga	ttaaagtggg	taacgccagg	gttttcccag	4800
tcacgacggt	gtaaaaacgac	ggccagtga	tt			4832

SEQ ID NO:71

atgactttta	acagttttga	aggatctaaa	acttgtgtac	ctgcagacat	caataaggaa	60
gaagaatttg	tagaagagtt	taatagatta	aaaacttttg	ctaatttttc	aagtggtagt	120
cctgttttcag	catcaacact	ggcacagca	gggtttcttt	atactggtga	aggagatacc	180
gtgcggtgct	ttagttgtca	tgcagctgta	gatagatggc	aatatggaga	ctcagcagtt	240
ggaagacaca	ggaagatato	cccaaatgca	agattttatca	acggctttta	tcttgaaaaat	300
agtgccacgc	agtctacaaa	ttctggtatc	cagaatggtc	agtacaaagt	tgaaaaactat	360
ctgggaagca	gagatcattt	tgcccttagac	aggccatctg	agacacatgc	agactatctt	420
ttgagaactg	ggcaggttgt	agatatatca	gacaccatat	acccgaggaa	ccctgccatg	480
tattgtgaag	aagctagatt	aaagtccttt	cagaactggc	cagactatgc	tcacctaac	540
ccaagagagt	tagcaagtgc	tggactctac	tacacaggta	ttggtgacca	agtgcagtg	600
ttttgttgtg	gtggaaaaat	gaaaaattgg	gaaccttgtg	atcgtgcctg	gtcagaacac	660
aggcgactct	ttcctaattg	cttctttgtt	ttggggccgga	atcttaatat	tcgaagtga	720
tctgatgctg	tgagttctga	taggaatttc	ccaaattcaa	caaactcttc	aagaaatcca	780
tcattggcag	attatgaagc	acggatcttt	acttttggga	catggatata	ctcagttaac	840
aaggagcagc	gtgcagaagc	tggattttat	gcttttaggtg	aagggtgata	agtaaatg	900
tttactgtg	gaggagggct	aactgattgg	aagccagtg	aagacccttg	ggaacaacat	960
gctaaatggt	atccagggtg	caaatatctg	ttagaacaga	agggacaaga	atatataaac	1020
aatattcatt	taactcattc	acttgaggag	tgtctggtaa	gaactactga	gaaaacacca	1080
tcactaacta	gaagaattga	tgataccatc	ttccaaaatc	ctatggtaca	agaagctata	1140
cgaatggggt	tcagtttcaa	ggacatttaag	aaaataatgg	aggaaaaaat	tcagatatct	1200
gggagcaact	ataaatcact	tgagggtctg	gttgacagatc	tagtgaatgc	tcagaaagac	1260

```

agtatgcaag atgagtcaag tcagacttca ttacagaaa agattagtagc tgaagagcag 1320
ctaaggcgcc tgcaagagga gaagctttgc aaaatctgta tggatagaaa tattgtctatc 1380
gtttttgttc cttgtggaca tctagtcact tgtaaacaaat gtgctgaagc agttgacaag 1440
tgtcccatgt gctacacagt cactactttc aagcaaaaaa tttttatgtc ttaatctaa 1499

```

SEQ ID NO:72

```

Met Thr Phe Asn Ser Phe Glu Gly Ser Lys Thr Cys Val Pro Ala Asp
1      5      10      15
Ile Asn Lys Glu Glu Glu Phe Val Glu Glu Phe Asn Arg Leu Lys Thr
20
Phe Ala Asn Phe Pro Ser Gly Ser Pro Val Ser Ala Ser Thr Leu Ala
35      40      45
Arg Ala Gly Phe Leu Tyr Thr Gly Glu Gly Asp Thr Val Arg Cys Phe
50      55      60
Ser Cys His Ala Ala Val Asp Arg Trp Gln Tyr Gly Asp Ser Ala Val
65      70      75      80
Gly Arg His Arg Lys Val Ser Pro Asn Cys Arg Phe Ile Asn Gly Phe
85      90      95
Tyr Leu Glu Asn Ser Ala Thr Gln Ser Thr Asn Ser Gly Ile Gln Asn
100      105      110
Gly Gln Tyr Lys Val Glu Asn Tyr Leu Gly Ser Arg Asp His Phe Ala
115      120      125
Leu Asp Arg Pro Ser Glu Thr His Ala Asp Tyr Leu Leu Arg Thr Gly
130      135      140
Gln Val Val Asp Ile Ser Asp Thr Ile Tyr Pro Arg Asn Pro Ala Met
145      150      155      160
Tyr Cys Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr
165      170      175
Ala His Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr
180      185      190
Gly Ile Gly Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys
195      200      205
Asn Trp Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe
210      215      220
Pro Asn Cys Phe Phe Val Leu Gly Arg Asn Leu Asn Ile Arg Ser Glu
225      230      235      240
Ser Asp Ala Val Ser Ser Asp Arg Asn Phe Pro Asn Ser Thr Asn Leu
245      250      255
Pro Arg Asn Pro Ser Met Ala Asp Tyr Glu Ala Arg Ile Phe Thr Phe
260      265      270
Gly Thr Trp Ile Tyr Ser Val Asn Lys Glu Gln Leu Ala Arg Ala Gly
275      280      285
Phe Tyr Ala Leu Gly Glu Gly Asp Lys Val Lys Cys Phe His Cys Gly
290      295      300
Gly Gly Leu Thr Asp Trp Lys Pro Ser Glu Asp Pro Trp Glu Gln His
305      310      315      320
Ala Lys Trp Tyr Pro Gly Cys Lys Tyr Leu Leu Glu Gln Lys Gly Gln
325      330      335
Glu Tyr Ile Asn Asn Ile His Leu Thr His Ser Leu Glu Glu Cys Leu
340      345      350
Val Arg Thr Thr Glu Lys Thr Pro Ser Leu Thr Arg Arg Ile Asp Asp
355      360      365
Thr Ile Phe Gln Asn Pro Met Val Gln Glu Ala Ile Arg Met Gly Phe
370      375      380
Ser Phe Lys Asp Ile Lys Lys Ile Met Glu Glu Lys Ile Gln Ile Ser
385      390      395      400
Gly Ser Asn Tyr Lys Ser Leu Glu Val Leu Val Ala Asp Leu Val Asn
405      410      415
Ala Gln Lys Asp Ser Met Gln Asp Glu Ser Ser Gln Thr Ser Leu Gln
420      425      430
Lys Glu Ile Ser Thr Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys
435      440      445
Leu Cys Lys Ile Cys Met Asp Arg Asn Ile Ala Ile Val Phe Val Pro
450      455      460
Cys Gly His Leu Val Thr Cys Lys Gln Cys Ala Glu Ala Val Asp Lys
465      470      475      480
Cys Pro Met Cys Tyr Thr Val Ile Thr Phe Lys Gln Lys Ile Phe Met
485      490      495
Ser

```

SEQ ID NO:73

```

gtcgacttct gagggcgaaa gaaccagctg tggaatgtgt gtcagttagg gtgtggaaaag 60
tccccaggct ccccgagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc 120
aggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaaagcat gcatctcaat 180
tagtcagcaa ccatagtccc gccccctaact ccgcccatcc cgcccctaac tccgccagc 240
tccgcccatt ctccgcccga tggtgacta atttttttta tttatgcaga ggccgaggcc 300
gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt 360
tgcaaaaagc tggatcgatc ctgagaactt cagggtgagt ttgggggacc ttgattgttc 420
tttctttttc gctattgtaa aattcatgtt atatggagg ggcaaagttt tcagggtgtt 480
gtttagaatg ggaagatgtc ccttgatatc ccatggaccc tcatgataat tttgtttctt 540

```

tcacttttcta	ctctgttgac	aaccattgtc	tctctttatt	ttcttttcat	tttctgtaac	600
tttttcgtta	aacttttagct	tgcaatttgta	acgaattttt	aaatttcactt	ttgtttatatt	660
gtcagattgt	aagtactttt	tctaatacact	tttttttcaa	ggcaatcagg	gtatattata	720
ttgtactttca	gcacagtttt	agagaacaa	tggtataatt	aatgataag	gtagaatatt	780
tctgcatata	aattcttggt	ggcgtggaaa	taatcttatt	ggtagaaca	actacatcct	840
ggatcatcatc	ctgccttttct	ctttatgggt	acaatgatat	acactgtttg	agatgaggat	900
aaaaactct	gagtcctaac	cgggcccctc	tgctaaccat	gttcatgcct	tcttcttttt	960
cctacagctc	ctgggcaacg	tgctggttat	tggtctgtct	catcattttg	gcaaagaatt	1020
gtaatacgac	tcactatagg	gcgaattcgg	atccatgact	tttaacagtt	ttgaaggatc	1080
taaaacttgt	gtacctgcag	acatcaataa	ggaagaagaa	tttgtagaag	agtttaatat	1140
attaaaaact	tttgttaatt	ttccaagtgg	tagtcctggt	tcagcatcaa	cactggcacg	1200
agcagggttt	ctttatactg	gtgaaggaga	taccgtgcgg	tgctttagtt	gtcatgcagc	1260
tgtagataga	tggaataatg	gagactcagc	agttggaaga	cacaggaaag	tatccccaaa	1320
ttgcagattt	atcaacggct	tttatcttga	aaatagtgc	acgcagtcta	caaattcttg	1380
tatccagaat	ggtcagtaaa	aaagtgaata	ctatctggga	agcagagatc	attttgcctt	1440
agacaggcca	tctgagacac	atgcagacta	tcttttgaga	actgggcagg	ttgtagatat	1500
atcagacacc	atataccga	ggaacccctg	catgtattgt	gaagaagcta	gattaaagtc	1560
ctttcagaac	tgccagact	atgctcacct	aaccccaaga	gagttagcaa	gtgctggact	1620
ctactacaca	ggtattggtg	accaagtgc	gtgcttttgt	tggtggtgaa	aactgaaaaa	1680
ttgggaacct	tgtagctgtg	cctggtcaga	acacaggcga	cactttccta	attgcttctt	1740
tgttttgggc	cgaatcttta	atattcgaa	tgaatctgat	gctgtgagtt	ctgataggaa	1800
tttcccaaat	tcacaagaat	ttccaagaaa	tccatccatg	gcagattatg	aagcacggat	1860
ctttactttt	gggacatgga	tatactcagt	taacaaggag	cagcttgcaa	gagctggatt	1920
ttatgcttta	gttgaaggtg	ataaagtaaa	gtgctttcac	tggtggaggag	ggctaactga	1980
ttgggaaccc	agtgaagacc	cttgggaaca	acatgctaaa	tggtatccag	ggtgcaataa	2040
tctgttagaa	cagaaggagc	aagaatatat	aaacaatatt	cattttaactc	attcacttga	2100
ggagtgtctg	gtaagaacta	ctgagaaaaa	accatcacta	actagaagaa	ttgatgatac	2160
catcttccaa	aatccctatg	tacaagaagc	tatacgaatg	gggttcagtt	tcaaggacat	2220
taagaaaaata	atggagaaa	aaattcagat	atctgggagc	aactataaat	cacttgaggt	2280
tctggttgca	gatctagtga	atgctcagaa	agacagtatg	caagatgagt	caagtcagac	2340
ttcattacag	aaagagatta	gtactgaaga	gcagctaagg	cgccctgcaag	aggagaagct	2400
ttgcaaaaatc	tgtaggata	gaaatatgtc	tatcgttttt	gttccctgtg	gacatctagt	2460
cacttgtaaa	caatgtgctg	aagcagttga	caagtgtccc	atgtgtctaca	cagtcattac	2520
tttcaagcaa	aaaattttta	tgtcttaatc	taaagatctt	attaaagcag	aacttgttta	2580
ttgcagctta	taatggttac	aaataaaagca	atagcatcac	aaatttcaca	aataaagcat	2640
ttttttcact	gcattctagt	tggtggttgt	ccaaactcat	caatgtatct	tatcatgtct	2700
ggtcgactct	agactcttcc	gcttcctcgc	tcactgactc	gctgcgctcg	gtcgttcggc	2760
tgccgagcgc	ggtatcagct	cactcaaaag	cggtaataacg	gttatccaca	gaatcagggg	2820
ataacgcagg	aaagaacatg	tgagcaaaa	gccagcaaaa	ggccaggaa	cgtaaaaagg	2880
ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	2940
gctcaagtca	gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	3000
gaagctccct	cgtgcctctg	cctgttccga	ccctgcctgt	taccggatac	ctgtccgcct	3060
tttcccttc	gggaagcgtg	gcgctttctc	aatgctcacg	ctgtaggtat	ctcagttcgg	3120
tgtaggtcgt	tcgctccaag	ctgggctgtg	tgacgaacc	ccccgttcag	cccgcacgct	3180
gccccttatc	cgttaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	3240
tgccagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gtcacagagt	3300
tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agttatttgt	atctgcgctc	3360
tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	3420
ccgctggtag	cgttggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	3480
ctcaagaaga	tcttttgatc	ttttctacgg	ggtctgacgc	tcagtggaa	gaaaactcac	3540
gttaagggat	tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	cttttaaat	3600
aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aaacttggtc	gacagttacc	3660
aatgcttaat	cagtgtaaca	cctatctcag	cgatctgtct	atttctgtct	tccatagttg	3720
cctgactccc	cgctgtgtag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	3780
ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaacagac	3840
cagccggaag	ggccgagcgc	agaagtggct	ctgcaacttt	atccgcctcc	atccagtcta	3900
ttaatgttg	ctgggaagct	agagtaagta	gttcgcccag	taatatgttt	cgcaacggtg	3960
ttgccaattgc	tacaggcatc	gtggtgtcac	gctcgtcggt	tggtatggct	tcattcagct	4020
cgggttccca	acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcgggtta	4080
gctccttcgg	tctccgac	gtgttcagaa	gtaagttggc	cgcatgtgta	tcactcatgg	4140
ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	4200
ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	4260
gcccggcgct	aatacgggat	aataccgcgc	cacatagcag	aaactttaaa	gtgctcatca	4320
ttggaaaacg	ttcttcgggg	cgaaaaactct	caaggatctt	accgctgttg	agatccagtt	4380
cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcggtt	4440
ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataaag	gcgacacgga	4500
aatgttgaa	actcactact	ttcttttttc	aatattattg	aagcatttat	cagggttatt	4560
gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	taaacaaata	ggggttccgc	4620
gcacatttcc	ccgaaaagtg	ccacctgacg	tctaagaaac	cattattatc	atgacattaa	4680
cctataaaaa	taggcgtatc	acgaggcccc	tttctgtctg	cgcggttcgg	tgatgacggt	4740
gaaaacctct	gacatcacag	gctcccggag	acggtcacag	cttgtctgta	agcggatgcc	4800
gggagcagac	aagcccgta	gggcgcgtca	gcgggtgttg	gcgggtgtgc	gggctggctt	4860
aactatgcgg	catcagagca	gattgtactg	agagtgacac	atatgcgggt	tgaaataaccg	4920
cacagatcg	taaggagaaa	ataccgcatc	aggaaattgt	aaacgttaat	attttgttaa	4980
aattcgcgtt	aaatttttgt	taaatcagct	cattttttta	ccaataggcc	gaaatcggca	5040
aaatccctta	taaatcaaaa	gaatagacgg	agatagggtt	gagtggttgt	ccagtttgga	5100
acaagagtc	actattaaag	aacgtggact	ccaacgtcaa	agggcgaaaa	accgtctatc	5160
agggcagtg	gaacctacgt	gaacctcac	cctaatacag	ttttttgggg	tcgaggtgcc	5220
gtaaaagcact	aaatcggaac	cctaaaggga	gcccccgatt	tagagcttga	cggggaaagc	5280
cggcgaaacgt	ggcgagaaa	gaagggaaga	aagcgaaaag	agcgggcgct	agggcgctgg	5340

caagtgtagc	ggtcacgtcg	cgcgtaacca	ccacacccgc	cgcgcttaat	gcgccgctac	5400
agggcgcgtc	gcgccattcg	ccattcaggc	tacgcaactg	ttgggaaggg	cgatcggtgc	5460
gggcctcttc	gctattacgc	cagctggcga	aggggggatg	tgctgcaagg	cgattaagtt	5520
gggtaacgcc	agggttttcc	cagtcacgac	gttgtaaaac	gacggccagt	gaatt	5575

SEQ ID NO:74

atggacttca	gcagaaatct	ttatgatatt	ggggaacaac	tggacagtga	agatctggcc	60
tccctcaagt	tccctgagcct	ggactacatt	ccgcaaagga	agcaagaacc	catcaaggat	120
gccttgatgt	tattccagag	actccaggaa	aagagaatgt	tggaggaaag	caatctgtcc	180
ttcctgaagg	agctgctctt	ccgaattaat	agactggatt	tgctgattac	ctacctaaac	240
actagaaagg	aggagatgga	aagggaaactt	cagacaccag	gcagggctca	aattttctgcc	300
tacagggtca	tgctctatca	gatttcagaa	gaagtgagca	gatcagaatt	gaggtctttt	360
aagttttctt	tgcaagagga	aatctccaaa	tgcaaaactgg	atgatgacat	gaacctgtctg	420
gatattttca	tagagatgga	gaagagggtc	atcctgggag	aaggaaagt	ggacatcctg	480
aaaagagtct	gtgcccgaat	caacaagagc	ctgctgaaga	taatcaacga	ctatgaagaa	540
ttcagcaaa	gggaggagtt	gtgtggggta	atgacaatct	cggactctcc	aagagaacag	600
gatagtgaat	cacagacttt	ggacaaagtt	taccaaata	aaagcaaac	tcggggatac	660
tgctctgatca	tcaacaatca	caatttttgca	aaagcacggg	agaaagtgcc	caaacttcac	720
agcattaggg	acaggaatgg	aacacacttg	gatgcagggg	ctttgaccac	gacctttgaa	780
gagcttcatt	ttgagatcaa	gccccacgat	gactgcacag	tagagcaaat	ctatgagatt	840
ttgaaaatct	accaactcat	ggaccacagt	aacatggact	gcttcactctg	ctgtatcctc	900
ttccatggag	acaagggcat	catctatggc	actgatggac	aggaggcccc	catctatgag	960
ctgacatctc	agttcactgg	tttgaagtgc	ccttcccttg	ctggaaaacc	caaagtgttt	1020
tttattcagg	ctgtgcaggg	ggataactac	cagaaaggta	tacctgttga	gactgattca	1080
gaggagcaac	cctattttaga	aatggattta	tcataccttc	aaacgagata	tatcccgat	1140
gaggctgact	ttctgctggg	gatggccact	gtgaataact	gtgtttccta	ccgaaaccct	1200
gcagagggaa	cctggtacat	ccagtcactt	tgccagagcc	tgagagagcg	atgtcctcga	1260
ggcgatgata	ttctcaccat	cctgactgaa	gtgaactatg	aagtaagcaa	caaggatgac	1320
aaagaaaaaca	tggggaaaaca	gatgcctcag	cctactttca	cactaagaaa	aaaacttgtc	1380
ttcccttctg	attga					1395

SEQ ID NO:75

Met	Asp	Phe	Ser	Arg	Asn	Leu	Tyr	Asp	Ile	Gly	Glu	Gln	Leu	Asp	Ser
1				5					10					15	
Glu	Asp	Leu	Ala	Ser	Leu	Lys	Phe	Leu	Ser	Leu	Asp	Tyr	Ile	Pro	Gln
			20					25					30		
Arg	Lys	Gln	Glu	Pro	Ile	Lys	Asp	Ala	Leu	Met	Leu	Phe	Gln	Arg	Leu
		35					40					45			
Gln	Glu	Lys	Arg	Met	Leu	Glu	Ser	Asn	Leu	Ser	Phe	Leu	Lys	Glu	
		50				55				60					
Leu	Leu	Phe	Arg	Ile	Asn	Arg	Leu	Asp	Leu	Leu	Ile	Thr	Tyr	Leu	Asn
65					70					75				80	
Thr	Arg	Lys	Glu	Glu	Met	Glu	Arg	Glu	Leu	Gln	Thr	Pro	Gly	Arg	Ala
			85					90						95	
Gln	Ile	Ser	Ala	Tyr	Arg	Val	Met	Leu	Tyr	Gln	Ile	Ser	Glu	Glu	Val
			100					105					110		
Ser	Arg	Ser	Glu	Leu	Arg	Ser	Phe	Lys	Phe	Leu	Leu	Gln	Glu	Glu	Ile
			115				120					125			
Ser	Lys	Cys	Lys	Leu	Asp	Asp	Met	Asn	Leu	Leu	Asp	Ile	Phe	Ile	
			130			135					140				
Glu	Met	Glu	Lys	Arg	Val	Ile	Leu	Gly	Glu	Gly	Lys	Leu	Asp	Ile	Leu
145					150				155					160	
Lys	Arg	Val	Cys	Ala	Gln	Ile	Asn	Lys	Ser	Leu	Leu	Lys	Ile	Ile	Asn
			165					170						175	
Asp	Tyr	Glu	Glu	Phe	Ser	Lys	Gly	Glu	Glu	Leu	Cys	Gly	Val	Met	Thr
			180					185					190		
Ile	Ser	Asp	Ser	Pro	Arg	Glu	Gln	Asp	Ser	Glu	Ser	Gln	Thr	Leu	Asp
			195				200					205			
Lys	Val	Tyr	Gln	Met	Lys	Ser	Lys	Pro	Arg	Gly	Tyr	Cys	Leu	Ile	Ile
			210				215					220			
Asn	Asn	His	Asn	Phe	Ala	Lys	Ala	Arg	Glu	Lys	Val	Pro	Lys	Leu	His
225					230					235				240	
Ser	Ile	Arg	Asp	Arg	Asn	Gly	Thr	His	Leu	Asp	Ala	Gly	Ala	Leu	Thr
			245					250						255	
Thr	Thr	Phe	Glu	Glu	Leu	His	Phe	Glu	Ile	Lys	Pro	His	Asp	Asp	Cys
			260				265						270		
Thr	Val	Glu	Gln	Ile	Tyr	Glu	Ile	Leu	Lys	Ile	Tyr	Gln	Leu	Met	Asp
			275				280					285			
His	Ser	Asn	Met	Asp	Cys	Phe	Ile	Cys	Cys	Ile	Leu	Ser	His	Gly	Asp
			290				295					300			
Lys	Gly	Ile	Ile	Tyr	Gly	Thr	Asp	Gly	Gln	Glu	Ala	Pro	Ile	Tyr	Glu
305					310					315				320	
Leu	Thr	Ser	Gln	Phe	Thr	Gly	Leu	Lys	Cys	Pro	Ser	Leu	Ala	Gly	Lys
			325					330						335	
Pro	Lys	Val	Phe	Phe	Ile	Gln	Ala	Cys	Gln	Gly	Asp	Asn	Tyr	Gln	Lys
			340					345					350		
Gly	Ile	Pro	Val	Glu	Thr	Asp	Ser	Glu	Glu	Gln	Pro	Tyr	Leu	Glu	Met
			355				360					365			
Asp	Leu	Ser	Ser	Pro	Gln	Thr	Arg	Tyr	Ile	Pro	Asp	Glu	Ala	Asp	Phe
370						375					380				

Leu Leu Gly Met Ala Thr Val Asn Asn Cys Val Ser Tyr Arg Asn Pro
 385 390 395 400
 Ala Glu Gly Thr Trp Tyr Ile Gln Ser Leu Cys Gln Ser Leu Arg Glu
 405 410 415
 Arg Cys Pro Arg Gly Asp Asp Ile Leu Thr Ile Leu Thr Glu Val Asn
 420 425 430
 Tyr Glu Val Ser Asn Lys Asp Asp Lys Lys Asn Met Gly Lys Gln Met
 435 440 445
 Pro Gln Pro Thr Phe Thr Leu Arg Lys Lys Leu Val Phe Pro Ser Asp
 450 455 460

SEQ ID NO:76

gtcgacttct gaggcggaag gaaccagctg tggaaatgtgt gtcagttagg gtgtggaaaag 60
 tccccaggct ccccgagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc 120
 aggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaaagcat gcatctcaat 180
 tagtcagcaa ccatagctccc gccctaaact ccgcccattc cggcccctaac tccgcccagt 240
 tccgcccatt ctccgcccga tggctgacta atttttttta tttatgcaga ggccgaggcc 300
 gccctggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt 360
 tgcaaaaagc tggatcgatc ctgagaactt cagggtgagt ttggggaccc ttgattgttc 420
 tttctttttc gctattgtaa aattcatgtt atatggagg ggcaaatgtt tcagggtgtt 480
 gtttagaatg ggaagatgct ccttgtatca ccatggaccc tcatgataat tttgtttctt 540
 tcacttttca ctctgttgac aaccattgtc tcctcttatt ttcttttcat tttctgtaac 600
 tttttcgtaa aacttttagct tgcatttgta acgaattttt aaattcactt ttgtttatatt 660
 gtcagattgt aactactttc tctaatactt tttttttcaa ggcaatcagg gtatattata 720
 ttgtacttca gcagagtttt agagaacaat tgttataatt aaatgataag gtagaatatt 780
 tctgcatata aattctggct ggctgggaaa tattcttatt ggtagaaca actacatcct 840
 ggtcatcatc ctgcctttct ctttatgtgt acaatgatat acactgtttg agatgaggat 900
 aaaatactct gagtccaaac cgggcccctc tgctaaccat gttcatgect tcttcttttt 960
 cctacagctc ctgggcaacg tgcctgttat tgtgctgtct catcattttg gcaaaagaatt 1020
 gtaatacgac tcatatagg gcgaattcat ggacttcagc agaaatcttt atgatattgg 1080
 ggaacaactg gacagtgaag atctggcctc cctcaagtgc ctgagcctgg actacattcc 1140
 gcaaggaag caagaaccca tcaaggatgc cttgatgtta ttccagagac tccaggaaaa 1200
 gagaatgttg gaggaaagca atctgtcctt cctgaaggag ctgctcttcc gaattaatag 1260
 actggatttg ctgattacct acctaaacac tagaaaagg gagatggaaa gggaaactca 1320
 gacaccaggc agggctcaaa tttctgccta cagggtcatg ctctatcaga tttcagaaga 1380
 agtgagcaga tcaagaattga ggtcttttaa gtttcttttg caagaggaaa tctccaaatg 1440
 caaactggat gatgacatga acctgctgga tattttcata gagatggaga agagggatcat 1500
 cctgggagaa ggaagtgttg acatcctgaa aagagctgtt gcccaaatca acaagagcct 1560
 gctgaagata atcaacgact atgaagaatt cagcaaaagg gaggagtgtg gtggggtaat 1620
 gacaatctcg gactctccaa gagaacagga tagtgaatca cagacttttg acaaagttaa 1680
 ccaaatgaaa agcaaacctc gggatactgt ctgatcatca acaatcacia ttttgcaaaa 1740
 gcaaggagaa aagtgcacca aacttcacag cattagggac aggaatggaa cacacttgga 1800
 tgcaagggtc ttgaccacga cttttgaaga gtttcatttt gagatcaagc cccacgatga 1860
 ctgcacagta gagcaaatct atgagatttt gaaaatctac caactcatgg accacagtaa 1920
 catggactgc ttcatctgct gtatcctctc ccatggagac aagggcatca tctatggcac 1980
 tgatggacag gaggccccca tctatgagct gacatctcag ttcaactggt tgaaagtccc 2040
 ttcccttgct ggaacccca aagtgttttt tattcaggct tgtcaggggg ataactacca 2100
 gaaaggtata cctgttgaga ctgattcaga ggagcaaccc tatttagaaa tggatttatc 2160
 atcacctcaa acgagatata tcccgatga ggctgacttt ctgctgggga tggccactgt 2220
 gaataactgt gtttctacc gaaacctgac agaggggaac tggtagatcc agtcactttg 2280
 ccagagcctg agagagcgat gtccctcgagg cgatgatatt ctccacctcc tgactgaagt 2340
 gaactatgaa gtaagcaaca aggatgacaa gaaaaacatg gggaaacaga tgcctcagcc 2400
 tacttttaca ctcaaaaaaa aacttgtctt cccttctgat tgaggatoca gatcttatta 2460
 aagcagaact tgtttattgc agcttataat gggtacaaat aaagcaaatg catcacaaat 2520
 ttacacaaata aagcattttt ttactgcat tctagtgttg gtttgtccaa actcatcaat 2580
 gactcttate atgtctgtgc gactctagac tcttccgctt cctcgctcac tgactcgctg 2640
 cgctcggctg ttccgctgcg gcgagcggtc tcagctcact caaaggcggt aatacggtta 2700
 tccacagaat caggggataa cgaggaaaag aacatgtgag caaaaggcca gcaaaaggcc 2760
 aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 2820
 catcacaaaa atcgacgctc aagtcaagg tgccgaaacc cgacaggact ataaagatac 2880
 caggcggttc ccctggaag ctccctcgct cgctctcctg ttccgacctt gccgttacc 2940
 ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctacgctgt 3000
 aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 3060
 gttcagcccg accgtcgcg cttatccggt aactatcgct ttgagtccaa cccggttaaga 3120
 cacgacttat cgcactggc agcagccact ggtaacagga ttagcagagc gaggatagta 3180
 ggcggtgcta cagagtctct gaagtgtgg accctaacag gctacactag aaggacagta 3240
 tttggatatc gcgctctgct gaagccagtt accttcggaa aaagagtgg tagctcttga 3300
 tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 3360
 cgagaaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 3420
 tggaacgaaa actcacgta agggattttg gtcattgagt tatcaaaaag gatcttcacc 3480
 tagactcttt taataataaa atgaagtttt aaatcaatct aaagtataa tgagtaaaact 3540
 tggctctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctatatt 3600
 cgttcatcca tagtgectg actccccgct ggttagataa ctacgatacg ggagggtta 3660
 ccactctgcc ccagtctgc aatgatacgc cgagaccacg gctcacggcg tccagattta 3720
 tcagcaataa accagccagc cgaaggggcc gagcgagaa gtggtcctgc aactttatcc 3780
 gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtgttc gccagttaat 3840
 agtttgccga acgttgttgc cattgctaca ggcactcgtg tgtcacgctc gtcgtttggt 3900
 atggcttcat tcaagctcgc ttcccaacga tcaaggcgag ttacatgac ccccatgttg 3960
 tgcaaaaaaag cgttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 4020
 gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 4080

```

agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg 4140
cgaccgagtt gctcttggcc gccgtcaata cgggataata ccgcgccaca tagcagaact 4200
ttaaagtgct tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 4260
ctgttgagat ccagttcgat gtaaccacact cgtgcaccca actgatcttc agcatctttt 4320
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 4380
ataagggcga caccgaaaatg ttgaatactc atactcttct tttttcaata ttattgaagc 4440
atztatcagg gttattgtct catgagcgga tacatatattg aatgtattta gaaaaataaa 4500
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 4560
attatcatga cattaaccta taaaaatagg cgtatcacga ggcccctttc gtctcgcgcg 4620
tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg 4680
tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttgccgg 4740
gtgtcggggc tggcttaact atgcggcatc agagcagatt gtactgagag tgcaccatat 4800
gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacagga aattgtaaac 4860
tgtaatatatt tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt ttttaaccaa 4920
taggccgaaa tcggcaaaat cccttataaa tcaaaaagaat agaccgagat aggggtgagt 4980
gtgtttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa cgtcaaaggg 5040
cgaaaaacgg tctatcaggg cgtatggcca ctacgtgaac catcaccta atcaagtttt 5100
ttggggtcga ggtgccgtaa agcactaaat cggaaacccta aaggagccc ccgatttaga 5160
gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc gaaaggagcg 5220
ggcgctaggg cgttggaag ttagcggtc acgctgcgg taaccaccac acccgccgcg 5280
cttaatgcgc cgtacaggg cgcgtcgcgc cattcggcat tcaggctacg caactgttgg 5340
gaagggcgat cgtgcgggc ctcttcgcta ttacgccagc tggcgagggg gggatgtgct 5400
gcaaggcgat taagtgggt aacgccaggg ttttcccagt cacgacgttg taaaacgacg 5460
gccagtgaat t 5471

```

SEQ ID NO:77

```

atggcgacag ctgggagaa agggtagcat aaccgggaga tagtgatgaa gtacatccat 60
tataagctgt cgcagagggg ctacgagtg gatgcgggag atgtggggcg cgcgcccccg 120
ggggccggcc ccgcaccggg catcttctcc tcccagccc ggacacggcc ccatccagcc 180
gcatcccggg acccggtcgc caggacctcg ccgctgcaga ccccggtgc ccccgccgccc 240
gcgcggggc ctgcgctcag ccggtgccca cctgtgttcc acctgacctt ccgccaggcc 300
ggcgacgact tctcccgcgc ctaccgcgcg gacttcgcgc agatgtccag ccagctgcac 360
ctgacgcctt tcaccgcgcg gggacgcttt gccacggtgg tggaggagct cttcaggggc 420
ggggtgaact ggggagagat tttggccttc tttgagttcg gtggggtcat gtgtgtggag 480
agcgtcaacc gggagatgtc gcccctggtg gacaacatcg ccctgtggat gactgagtac 540
ctgaaccggc acctgcacac ctggatccag gataacggag gctgggttagg tgcacttggg 600
gatgtgagtc tgggctga 618

```

SEQ ID NO:78

```

Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
1 5 10 15
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
20 25 30
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
35 40 45
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
50 55 60
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
65 70 75 80
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Thr
85 90 95
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Arg Asp Phe
100 105 110
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
115 120 125
Arg Phe Ala Thr Val Val Glu Leu Phe Arg Asp Gly Val Asn Trp
130 135 140
Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145 150 155 160
Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
165 170 175
Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
180 185 190
Gly Gly Trp Val Gly Ala Leu Gly Asp Val Ser Leu Gly
195 200 205

```

SEQ ID NO:79

```

gtcgacttct gaggggaaa gaaccagctg tggaatgtgt gtcagttagg gtgtggaaaag 60
tcccaggct cccagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc 120
agggtgtgaa agtccccagg ctccccagca ggcagaagta tgcaaagcat gcatctcaat 180
tagtcagcaa ccatagtcct gcccttaact ccgccatcc cgcccctaac tccgccagc 240
tccgccatt ctccgcccga ttgctgacta atttttttta tttatgcaga ggccgaggcc 300
gctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt 360
tgcaaaaagc ttgatcgatc ctgagaactt cagggtgagt ttggggaccc ttgattgttc 420
tttctttttc gctattgtaa aattcatggt atatggaggg ggcaaagttt tcagggtgtt 480
gtttagaatg ggaagatgtc cttgtatca ccatggaccc tcatgataat tttgtttctt 540
tcacttttcta ctctgttgac aaccattgtc tcctcttatt ttcttttcat tttctgtaac 600
tttttcgtta aactttagct tgcatttgta acgaattttt aaattcactt ttgtttattt 660
gtcagattgt aagtactttc tctaatactt tttttttcaa ggcaatcagg gtatattata 720
ttgtacttca gcacagtttt agagaacaa ttgtataatt aaatgataag gtagaatatt 780

```

tctgcatata	aattcttggt	ggcgtgga	tattcttatt	ggtagaaca	actacatcct	840
ggatcatcat	ctgcctttct	ctttatgggt	acaatgat	acactgtttg	agatgaggat	900
aaaatactct	gagtcacaa	cggggccctc	tgctaaccat	gttcatgcct	tcttcttttt	960
cctacagctc	ctgggcaacg	tgctggttat	tgctgtgtct	catcattttg	gcaaagaatt	1020
gtaatacgac	tcactatagg	gcgaattcgg	atccagatct	atggcgacag	ctgggagAAC	1080
agggtagcat	aaccgggaga	tagtgatgaa	gtacatccat	tataagctgt	cgagagggg	1140
ctacgagtgg	gatgggggag	atgtggggcg	cgcgccccc	ggggccgccc	ccgcaccggg	1200
catcttctcc	tcccagcccg	ggcacacgccc	ccatccagcc	gcaccccggg	acccggtcgc	1260
caggacctcg	ccgctgcaga	ccccggctgc	ccccggcgcc	gcgcgggggc	ctgcgctcag	1320
cccggtgcca	cctgtgtgtc	acctgacctt	ccgcccaggcc	ggcgacgact	tctcccgccg	1380
ctacgcgcgc	gacttcgccc	agatgtccag	ccagctgcac	ctgacgcctt	tcaccgcgcg	1440
gggacgcttt	gccacgggtg	tggaggagct	cttcaggggac	gggggtgaact	gggggaggat	1500
tggtggccttc	tttgagttcg	gtgggggtcat	gtgtgtggag	agcgtcaacc	gggagatgtc	1560
gcccctgggtg	gacaacatcg	ccctgtggat	gactgagtac	ctgaaccggc	acctgcacac	1620
ctggatccag	gataacggag	gctgggtagg	tgcaacttgg	gatgtgagtc	tgggctgaag	1680
atcttatttaa	agcagaactt	gtttattgca	gcttataatg	gttacaataa	aagcaatagc	1740
atcacaat	tcacaaataa	agcatttttt	tcactgcatt	ctagtgtggg	tttgtccaaa	1800
ctcatcaatg	tatcttatac	tgtctgtctg	actctagact	cttccgcttc	ctcgtcactt	1860
gactcgtctg	gctcggctcg	tcggctgcgc	cgagcggtat	cagctcactc	aaaggcggtg	1920
atacgggttat	ccacagaatc	aggggataac	gcaggaaaga	acatgtgagc	aaaaggccag	1980
caaaaggcca	ggaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccc	2040
ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagagggtg	gcgaaacccg	acaggactat	2100
aaagatacca	ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctgtt	ccgacctctg	2160
cgtttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cggtggcgctt	tctcaatgct	2220
cacgtctgtag	tgatctcagt	tcgggtgtag	tcgttcgctc	caagctgggc	tggtgtgcacg	2280
aaccccccg	tcagcccgac	cgctgcgcct	tatccggtaa	ctatcgtctt	gagtcacacc	2340
cggtaaagaca	cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	2400
ggatgttagg	cgtgtctaca	gagttcttga	agtgggtggc	taactacggc	tacactagaa	2460
ggacagtat	tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	2520
gctcttgatc	cgcaaaacaa	accaccgctg	gtagcgggtg	tttttttgg	tgcaagcagc	2580
agattacgcg	cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	acggggtctg	2640
acgctcagtg	gaacgaaaa	tcacgttaag	ggattttgg	catgagatta	tcaaaaagga	2700
tcttcaccta	gatcctttta	aattaaaaat	gaagttttta	atcaatctaa	agtatatatg	2760
agtaaaactg	gtctgacagt	taccaatgct	taatcagtga	ggcacctatc	tcagcgatct	2820
gtctatttctg	ttcatccata	gttgcttgac	tccccgctgc	gtagataact	acgatacggg	2880
agggctttacc	atctggcccc	agtgtgcgca	tgataaccgc	agacccacgc	tcaccggctc	2940
cagattttatc	agcaataaac	cagccagccg	gaagggccga	gcgcagaggt	ggctcctgca	3000
ctttatccgc	ctccatccag	tctattaatt	gttgccggga	agctagagta	agtagttcgc	3060
cagttaaatag	tttgccgaac	gtgtttgcca	ttgctacagg	catcgtgggtg	tcacgctcgt	3120
cgtttggtat	ggcttcattc	agctccgggt	cccaacgac	aaggcgaggt	acatgatccc	3180
ccatgtttgtg	caaaaaagcg	gttagctcct	tcgggtcctc	gatcgttgtc	agaagtaagt	3240
tggccgcag	gttatcactc	atgggttatg	cagcactgca	taattctctt	actgtcatgc	3300
catccgtaag	atgcttttct	gtgactgggtg	agtactcaac	caagctattc	tgagaatagt	3360
gtatgcccgc	accgagttgc	tcttgcccgc	cgtcaatcac	ggataatacc	gcgccacata	3420
gcagaacttt	aaaagtgtc	atcattggaa	aacgttcttc	ggggcgaaaa	ctctcaagga	3480
tcttaccgct	gttcgagatc	agttcgatgt	aacccactcg	tgcaaccac	tgatcttcag	3540
catctttttac	tttcaccagc	gtttctgggt	gagcaaaaa	aggaaggcaa	aatgccgcaa	3600
aaaagggaat	aaggcgacac	cggaaatgtt	gaatactcat	actcttcttt	tttcaatatt	3660
attgaagcat	ttatcagggt	tattgtctca	tgagcggata	catatttgaa	tgtatttaga	3720
aaaataaaca	aatagggtt	ccgcgcacat	ttccccgaaa	agtgccacct	gacgtctaag	3780
aaaccattat	tatcatgaca	ttaacctata	aaaataggcg	tatcacgagg	cccctttcgt	3840
ctcgcgcggt	tcggtgatga	cggtgaaaac	ctctgacaca	tgacgtcccc	ggagacgggtc	3900
acagctttgtc	ttgtaagcga	tgccgggagc	agacaagccc	gtcaggggcg	gtcagcggtg	3960
gttgggcggt	gtcggggctg	gcttaactat	gcggcatcac	agcagattgt	actgagagtg	4020
caccatattgc	ggtgtgaaat	accgcacaga	tgcgtaagga	gaaaaataccg	catcaggaaa	4080
ttgtaaacgt	taataattttg	ttaaaattcg	cgttaaattt	ttgttaaattc	agctcatttt	4140
ttaaccaata	ggccgaaatc	gttataaatc	cttataaatc	aaaagaatag	accgagatag	4200
ggttgagtg	tgttccagtt	tggaacaaga	gtccactatt	aaagaacgtg	gactccaacg	4260
tcaaagggcg	aaaaaccgtc	tatcaggggc	atgggccact	acgtgaacca	tcaccctaat	4320
caagtttttt	ggggtcgagg	tgccgtaaag	cactaaatcg	gaaccctaaa	gggagcccc	4380
gatttagagc	ttgacgggga	aagccggcga	acgtggcgag	aaagggaagg	aagaaagcga	4440
aaggagcggg	cgctaggcg	ctggcaagtg	tagcggtcac	gctgcgcgta	accaccacac	4500
ccgcgcgcgt	taatgcgcgc	ctacaggggc	cgctgcgcga	ttcgccattc	aggctacgca	4560
actgtttgga	agggcgatcg	gtgcggggcct	cttcgctatt	acgccagctg	gcgaaggggg	4620
gatgtgctgc	aaggcgatta	agttgggtaa	cgccagggtt	ttcccagtc	cgacgttgta	4680
aaacgacggc	cagtgaatt					4699

SEQ ID NO:80

gtcgaacttct	gaggcgga	gaaccagctg	tggaatgtgt	gtcagttagg	gtgtggaaa	60
tccccaggct	ccccagcag	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	120
aggtgtgga	agtccccagg	ctccccagca	ggcagaagta	tgcaaaagcat	gcacttcaat	180
tagtcagcaa	ccatagtc	gccccctaac	ccgcccaccc	cgccccctaac	tccgcccagt	240
tcgcgccatt	ctccgccc	tggtgacta	attttttttt	tttatgcaga	ggccgaggcc	300
gcctcggcct	ctgagcgtt	ccagaagtag	tgaggaggct	tttttggagg	cctaggcttt	360
tgcaaaaagc	tggatcgatc	ctgagaactt	cagggtgagt	ttggggaccc	ttgattgttc	420
tttctttttc	gctattgtaa	aattcatggt	atatggagg	ggcaaaagtt	tcagggtgtt	480
gtttagaatg	ggaagatgtc	ccttgtatca	ccatggaccc	tcagtataat	tttgtttctt	540
tcactttcta	ctctgttgac	aaccattgtc	tcctcttatt	ttcttttcat	tttctgtaac	600
tttttcgtta	aacttttagc	tgcatttgtg	acgaattttt	aaattcactt	ttgttttatt	660
gtcagattgt	aagtactttc	tctaactcact	tttttttcaa	ggcaatcagg	gtatattata	720

ttgtacttca	gcacagtttt	agagaacaat	tgttataatt	aatgataag	gtagaatatt	780
tctgcatata	aattctggct	ggcgtggaaa	tattcttatt	ggtagaaaca	actacatcct	840
ggatcatcatc	gtgcctttct	ctttatggtt	acaatgatat	acactgtttg	agatgaggat	900
aaaatactct	cgtctccaa	cgggccctc	tgctaaccat	gttcacgctt	tcttcttttt	960
cctacagctc	ctgggcaacg	tgctggttat	tgtgctgtct	catcattttg	gcaaagaatt	1020
gtaatacgac	tcactatagg	gcgaattcgg	atccatggac	ttcagcagaa	atctttatga	1080
tattggggaa	caactggaca	gtgaagatct	ggcctccctc	aagttcctga	gcctggacta	1140
cattccgcaa	aggaagcaag	aacccatcaa	ggatgccttg	atgttattcc	agagactcca	1200
ggaaaagaga	atgttggagg	aaagcaatct	gtccttcctg	aaggagctgc	tcttccgaat	1260
taatagactg	gatttgctga	ttacctacct	aaacactaga	aaggaggaga	tggaagggga	1320
acttcagaca	ccaggcagg	ctcaaatttc	tgctacagag	gtcatgctct	atcagatttc	1380
agaagaagtg	agcagatcag	aattgaggct	ttttaagttt	cttttgcaag	aggaaatctc	1440
caaatgcaaa	ctggatgatg	acatgaacct	gctggatatt	ttcatagaga	tggaagagag	1500
ggatcatcctg	ggagaaggaa	agttggacat	cctgaaaaga	gtctgtgccc	aaatcaacaa	1560
gagcctgctg	aagataatca	acgactatga	agaattcagc	aaaggggagg	agttgtgtgg	1620
ggtaatgaca	atctcggact	ctccaagaga	acaggatagt	gaatcacaga	ctttggacaa	1680
agtttaccaa	atgaaaagca	aacctcgggg	atactgtctg	atcatcaaca	atcacatttt	1740
tgcaaaagca	cgggagaaag	tgcccaaat	tcacagcatt	agggacagga	atggaacaca	1800
cttggaatgca	ggggctttga	ccacgacctt	tgaagagctt	cattttgaga	tcaagcccca	1860
cgatgactgc	acagtagagc	aaatctatga	gattttgaaa	atctaccaac	tcatggacca	1920
cagtaacatg	gactgcttca	tctgctgtat	cctctcccat	ggagacaagg	gcatcatcta	1980
tggtcactgat	ggcaggatga	cccccattca	tgagctgaca	tctcagttca	ctgggttgaa	2040
gtgccccttcc	cttgctggaa	aacccaaagt	gttttttatt	caggcttctc	agggggataa	2100
ctaccagaaa	ggtataactg	ttgagactga	ttcagaggag	caacctattt	tagaaatgga	2160
tttatcatca	ctccaaacg	gatataatccc	ggatgaggct	gactttctgc	tggggatggc	2220
cactgtgaat	aactgtgttt	cctaccgaaa	ccctgcagag	ggaacctggt	acatccagtc	2280
actttgccag	agcctgagag	agcagatgtc	tcgaggcgat	gatattctca	ccatcctgac	2340
tgaagtgaac	tatgaagtaa	gcaacaagga	tgacaagaaa	aacatgggga	aacagatgcc	2400
tcagcctact	ttcacactaa	gaaaaaaact	tgtcttccct	tctgattgaa	gatcttatta	2460
aagcagaact	tgtttattgc	agcttataat	ggttacaaat	aaagcaatag	catcacaaat	2520
ttcacaaaata	aagcattttt	ttcactgcat	tctagtgttg	gtttgtccaa	actcatcaat	2580
gtatctttatc	atgtctgtgc	gactctagac	tcttccgctt	cctcgctcac	tgactcgctg	2640
cgctcggctg	ttcggtctgc	gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	2700
ttcacagaaat	caggggataa	cgaggaaaag	aacatgtgag	caaaaaggcca	gcaaaaaggcc	2760
aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	2820
catcacaaaa	atcgacgctc	aagtcaaggg	tgcgcaaaac	cgacaggact	ataaaagatac	2880
caggcggtttc	cccctggaag	ctccctcgtg	cgctctccctg	ttccgacctt	gccgcttacc	2940
ggataccgtg	ccgcctttct	cccttcggga	agcgtggcgc	tttctcaatg	ctcacgctgt	3000
aggtatctca	gttcgggtga	ggtcgttcgc	tccaaagctgg	gctgtgtgca	cgaaccccc	3060
gttcagcccg	accgtctgcg	cttatccggt	aactatcgct	ttgagtccaa	cccggtaaga	3120
cacgacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	3180
ggcggtgcta	cagagttcct	gaagtgtggt	cctaactacg	gctacactag	aaggacagta	3240
tttggtatct	cgctctgctg	gaagccagtt	accttcggaa	aaagagtttg	tagctcttga	3300
tcgggcaaac	aaaccaccgc	tggtagcggg	ggtttttttg	tttgcaagca	gcagattacg	3360
cgagaaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	3420
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	3480
tagatccttt	taaattaaaa	atgaagtgtt	aaatcaatct	aaagtatata	tgagtaaaact	3540
tggtctgaca	gttaccaatg	cttaatcagt	gaggcaccta	tctcagcgat	ctgtctatctt	3600
cgttcatcca	tagttgctcg	actccccgtc	gtgtagataa	ctacgatacg	ggagggctta	3660
ccatctggcc	ccagtgctgc	aatgataccg	cgagaccac	gctcacgggc	tccagattta	3720
tcagcaataa	accagccagc	cggaaggggc	gagcgcagaa	gtggtcctgc	aactttatcc	3780
gcctccatcc	agtctattaa	ttgttgccgg	gaagctagag	taagtgttc	gccagttaat	3840
agtttgcgca	acgttggtgc	cattgtctaca	ggcatcggtg	tgctacgctc	gtcgtttggt	3900
atggcttcat	tcagctccgg	ttcccaacga	tcaaggcgag	ttacatgac	ccccatggtg	3960
tgcaaaaaaag	cggttagctc	cttcggctct	ccgatcggtg	tcagaagtaa	gttgccgca	4020
gtgttatcac	tcatggttat	ggcagcactg	cataattctc	ttactgtcat	gccatccgta	4080
agatgctttt	ctgtgactgg	tgagtactca	accaagtcat	tctgagaata	gtgtatgcgg	4140
cgaccgagtt	gctcttgccc	ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	4200
ttaaaagtgc	tcacatttgg	aaaacgttct	tcggggcgaa	aactctcaag	gatctttaccg	4260
ctgttgagat	ccagttcgat	gtaacccact	cgtgcaccca	actgatcttc	agcatctttt	4320
actttcacca	gcgtttctgg	gtgagcaaaa	acagggaaggc	aaaatgccgc	aaaaaaggga	4380
ataagggcga	cacggaaatg	ttgaatactc	atactcttct	tttttcaata	ttattgaagc	4440
atztatcagg	gttattgtct	catgagcgga	tacataattg	aatgtattta	gaaaaataaa	4500
caaatagggg	ttccgcgcac	atttccccga	aaagtgccac	ctgacgtcta	agaaaccatt	4560
attatcatga	cattaaccta	taaaaatagg	cgtatcacga	ggcccctttc	gtctcgcgcg	4620
tttcgggtgat	gacggtgaaa	acctctgaca	catgcagctc	ccggagacgg	tcacagcttg	4680
tctgtaagcg	gatgcgggga	gcagacaagc	ccgtcagggc	gcgtcagcgg	gtgttgcgcg	4740
gtgtcggggc	tggtttaact	atgcggcatc	agagcagatt	gtactgagag	tgaccataat	4800
gcggtgtgaa	ataccgcaca	gatgcgtaag	gagaaaatac	cgcatcagga	aattgtaaac	4860
gttaatatatt	tgttaaaaat	cgcgttaaat	ttttgttaaa	tcagctcatt	ttttaaccaa	4920
taggcgaaaa	tcggcaaaat	cccttataaa	tcaaaaagaa	agaccgagat	aggggttgagt	4980
gttggtccag	tttggaacaa	gagtcactta	ttaaagaacg	tggaactcaa	cgtcaaaggg	5040
cgaaaaaccg	tctatcaggg	cgatggccca	ctacgtgaac	catcacccca	atcaagtttt	5100
ttggggtcga	gtgcccgtga	agcactaaat	cggaaacccta	aagggaagcc	ccgatttaga	5160
gcttgacggg	gaaagccggc	gaacgtggcg	agaaaggaag	ggaagaaagc	gaaaggagcg	5220
ggcgctaggg	cgctggcaag	gttagcggtc	acgctgcg	taaccaccac	acccgcccgc	5280
cttaatgcgc	cgctacaggg	cgctgcg	cattcgccat	tcaggctacg	caactgttgg	5340
gaaggcgat	cggtgcggc	ctcttcgcta	ttacgacagc	tggcgaaggg	gggatgtgct	5400
gcaaggcgat	taagtgggt	aacgccaggg	ttttccagct	cacgacgttg	taaaacgacg	5460
gccagtgat	t					5471

SEQ ID NO:81

```

Met Asp Phe Ser Arg Asn Leu Tyr Asp Ile Gly Glu Gln Leu Asp Ser
1      5      10      15
Glu Asp Leu Ala Ser Leu Lys Phe Leu Ser Leu Asp Tyr Ile Pro Gln
20      25      30
Arg Lys Gln Glu Pro Ile Lys Asp Ala Leu Met Leu Phe Gln Arg Leu
35      40      45
Gln Glu Lys Arg Met Leu Glu Glu Ser Asn Leu Ser Phe Leu Lys Glu
50      55      60
Leu Leu Phe Arg Ile Asn Arg Leu Asp Leu Leu Ile Thr Tyr Leu Asn
65      70      75      80
Thr Arg Lys Glu Glu Met Glu Arg Glu Leu Gln Thr Pro Gly Arg Ala
85      90      95
Gln Ile Ser Ala Tyr Arg Val Met Leu Tyr Gln Ile Ser Glu Glu Val
100     105     110
Ser Arg Ser Glu Leu Arg Ser Phe Lys Phe Leu Leu Gln Glu Glu Ile
115     120     125
Ser Lys Cys Lys Leu Asp Asp Asp Met Asn Leu Leu Asp Ile Phe Ile
130     135     140
Glu Met Glu Lys Arg Val Ile Leu Gly Glu Gly Lys Leu Asp Ile Leu
145     150     155     160
Lys Arg Val Cys Ala Gln Ile Asn Lys Ser Leu Leu Lys Ile Ile Asn
165     170     175
Asp Tyr Glu Glu Phe Ser Lys Gly Glu Glu Leu Cys Gly Val Met Thr
180     185     190
Ile Ser Asp Ser Pro Arg Glu Gln Asp Ser Glu Ser Gln Thr Leu Asp
195     200     205
Lys Val Tyr Gln Met Lys Ser Lys Pro Arg Gly Tyr Cys Leu Ile Ile
210     215     220
Asn Asn His Asn Phe Ala Lys Ala Arg Glu Lys Val Pro Lys Leu His
225     230     235     240
Ser Ile Arg Asp Arg Asn Gly Thr His Leu Asp Ala Gly Ala Leu Thr
245     250     255
Thr Thr Phe Glu Glu Leu His Phe Glu Ile Lys Pro His Asp Asp Cys
260     265     270
Thr Val Glu Gln Ile Tyr Glu Ile Leu Lys Ile Tyr Gln Leu Met Asp
275     280     285
His Ser Asn Met Asp Cys Phe Ile Cys Cys Ile Leu Ser His Gly Asp
290     295     300
Lys Gly Ile Ile Tyr Gly Thr Asp Gly Gln Glu Ala Pro Ile Tyr Glu
305     310     315     320
Leu Thr Ser Gln Phe Thr Gly Leu Lys Cys Pro Ser Leu Ala Gly Lys
325     330     335
Pro Lys Val Phe Phe Ile Gln Ala Ser Gln Gly Asp Asn Tyr Gln Lys
340     345     350
Gly Ile Pro Val Glu Thr Asp Ser Glu Glu Gln Pro Tyr Leu Glu Met
355     360     365
Asp Leu Ser Ser Pro Gln Thr Arg Tyr Ile Pro Asp Glu Ala Asp Phe
370     375     380
Leu Leu Gly Met Ala Thr Val Asn Asn Cys Val Ser Tyr Arg Asn Pro
385     390     395     400
Ala Glu Gly Thr Trp Tyr Ile Gln Ser Leu Cys Gln Ser Leu Arg Glu
405     410     415
Arg Cys Pro Arg Gly Asp Asp Ile Leu Thr Ile Leu Thr Glu Val Asn
420     425     430
Tyr Glu Val Ser Asn Lys Asp Asp Lys Lys Asn Met Gly Lys Gln Met
435     440     445
Pro Gln Pro Thr Phe Thr Leu Arg Lys Lys Leu Val Phe Pro Ser Asp
450     455     460

```

SEQ ID NO:82

```

gtcgacttct gaggcggaag gaaccagctg tggaatgtgt gtcagttagg gtgtggaaag      60
tccccaggct cccagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc      120
agggtgtgaa agtccccagg ctccccagca ggcagaagta tgcaaagcat gcatctcaat      180
tagtcagcaa ccatagtccc gcccttaact ccgcccaccc cgcccctaac tccgcccagt      240
tccgcccatt ctccgcccga tggctgacta atttttttta tttatgcaga ggccgaggcc      300
gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggcttt      360
tgcaaaaagc tggatcgatc ctgagaactt cagggtgagt ttggggaccc ttgattgttc      420
tttctttttc gctattgtaa aattcatggt atattggagg ggcaaagttt tcagggtgtt      480
gtttagaatg gaagatgtc ccttgatca ccatggaccc tcatgataat tttgtttctt      540
tcaactttcta ctctgttgac aaccattgtc tcctcttatt ttcttttcat tttctgtaac      600
tttttcgtta aacttttagct tgcatttgta acgaattttt aaattcactt ttgtttattt      660
gtcagattgt aagtactttc tctaatactt tttttttcaa ggcaatcagg gtatattata      720
ttgtacttca gcacagtttt agagaacaat tggtataatt aaatgataag gtagaatatt      780
tctgcatata aattctggct gccgtggaat tattcttatt ggtagaaaca actacatcct      840
ggtcacatc ctgcctttct ctttatgggt acaatgatat acactgtttg agatgaggat      900
aaaatactct gagtccaac cgggcccctc tgctaaccat gttcatgcct tcttcttttt      960
cctacagctc ctgggcaacg tgctggttat tgtgctgtct catcattttg gcaaagaatt      1020
gtaatacgac tcactatagg gcgaattcgg atccatggac gaagcggatc ggcggtcct      1080

```

```

gcggcgggtgc cggtcgccgc tgggtggaaga gctgcaggtg gaccagctctt gggacgccct 1140
gctgagccgc gagctgttca ggccccatat gatcgaggac atccagcggg caggctcttg 1200
atctcggcgg gatlacggcca ggagctgat catagatctg gagactcgag ggagtcaggc 1260
tcttcctttg ttcatctcct gcttagagga cacaggccag gacatgctgg ctctgtttct 1320
gcgaactaac aggcgaagcag caaagtgtgc gaagccaacc ctagaaaaacc ttaccccagt 1380
ggtgctcaga ccagagattc gcaaaccaga ggttctcaga ccggaaacac ccagaccagt 1440
ggacattggt tctggaggat ttggtgatgt cgggtgctctt gagagtttga ggggaaatgc 1500
agatttggtt tacatcttga gcatggagcc ctgtggccac tgccctatta tcaacaatgt 1560
gaacttctgc cgtgagtcgg ggctccgcac ccgcactggc tccaacatcg actgtgagaa 1620
gttgccggcg cgtttctcct cgtctgcat tggcgtggag gtgaaggggc acctgactgc 1680
caagaaaaatg tgctggcgtt tgctggagct ggcgccagcag gaccacgggtg ctctggactg 1740
ctgctggttg gtcattctct ctacaggctg tcaggccagc cacctgcagt tcccaggggc 1800
tgtctacggc acagatggat gccctgtgtc ggtcgagaag attgtgaaca tcttcaatgg 1860
gaccagctgc ccagccttgg gagggaaagg caagctcttt ttcatccagg cctctggttg 1920
ggagcagaaa gaccatgggt ttgaggtggc ctccacttcc cctgaagacg agtcccctgg 1980
cagtaacccc gagccagatg ccaccccggt ccaggaaggt ttgaggacct tcgaccagct 2040
ggagcccata tctagtttgc ccacaccag tgacatcttt gtgtcctact ctactttccc 2100
aggttttggt tcttggaggg accccaagag tggctcctgg tacgttgaga ccctggacga 2160
catctttgag cagtgggtgc actctgaaga cctgcagtcc ctctgctta gggtcgctaa 2220
tgtctgttgc gtgaaaggga ttataaaaca gatgcctggt tgctttaatt tcttccggaa 2280
aaaaactttt tttaaaacat cataaagatc ttattaaagc agaacttgtt tattgcagct 2340
tataatggtt acaataaag caatagcatc acaaatatga caaataaagc atttttttca 2400
ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctggctcgact 2460
ctagactctt ccgtcttctc gtcactgac tcgctgcgtc cggtcgtttc gctgcggcga 2520
ggcgtatcag ctactctaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 2580
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 2640
ctggcgtttt tccataggtc ccgccccctt gacgagcatc acaaaaaatc acgctcaagt 2700
cagaggtggc gaaacccgac aggaactaaa agataccagg cgtttccccc tggaagctcc 2760
ctcgtgcgtc cctctgttcc gaccctgcgt cttaccggat acctgtccgc ctttctccct 2820
tcgggaagcg tggcgcttcc tcaatgtctc cgtctaggtt atctcagttc ggtgtaggtc 2880
gttcgtctca agctgggtcg tgtgcacgaa ccccccggtc agcccgaccg ctgcgcctta 2940
tccggttaact atcgtcttga gtccaaaccg gtaagacacg acttatcgcc actggcagca 3000
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 3060
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgtgaaag 3120
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgctggt 3180
agcgggtggt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 3240
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaaactc acgttaaggg 3300
attttggtca tgagattatc aaaaaggatc ttacactaga tctttttaa ttaaaaatga 3360
agtttttaaat caatctaaag tatatatgag taaaactggt ctgacagtta ccaatgctta 3420
atcagtgagg cacctatctc agcgtatctg ctatctcggt catccatagt tgctgactc 3480
cccgtcgtgt agataactac gatacgggag ggtttaccat ctggccccag tgctgcaatg 3540
ataccgcgag acccacgctc accggctcca gatttatcag caataaaacca gccagccgga 3600
agggccgagc gcagaagtggt tcttgcactc ttatccgctt ccatccagtc tattaatgtt 3660
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgccgaacgt tgttgccatt 3720
gttacaggca tctgtggtgc acgctcgtcg tttggtatgg cttcattcag ctccggttcc 3780
caacgatcaa gcgagttac atgatcccc atggtgtgca aaaaagcggg tagctcctc 3840
ggtcctccga tctgtgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 3900
gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag 3960
tactcaacca agtcattctg agaatagtgt atgcccgcag cgagttgctc ttgcccggcg 4020
tcaatacggg ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa 4080
cgttcttcgg ggcgaaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa 4140
ccactcgtg caccacaact atcttcagca tcttttactt tcaccagcgt ttctgggtga 4200
gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga 4260
atactcatac tcttcttttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 4320
agcggataca tatttgaatg tatttagaaa aataaaacaaa taggggttcc gcgcacattt 4380
ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa 4440
aataggcgta tcacgaggcc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaaact 4500
ctgacacatg cagctcccg agacggtcac agcttgtctg taagcggatg ccgggagcag 4560
acaagcccggt cagggcggt cagcgggtgt tggcgggtgt cggggcgtgc ttaactatgc 4620
ggcatcagag cagattgtac tgagagtcca ccatatgcgg tgtgaaatac cgcacagatg 4680
cgtaaggaga aaataccgca tcaggaaatt gtaaacgtta atattttgtt aaaattcgcg 4740
ttaaattttt gttaaatcag ctcatttttt aaccaatagg ccgaaatcgg caaaatccct 4800
tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg gaacaagagt 4860
ccactattaa agaacgtgga ctccaacgtc aaaggcgcaa aaacgtcta tcaggcgat 4920
ggcccaactac gtgaaccatc accctaatac agtttttttg ggtcgagggt ccgtaaaagca 4980
ctaaatcgga accctaaagg gagccccga ttttagactt gacggggaaa gccgcgaac 5040
gtggcgagaa aggaagggaa gaaagcgaaa ggagcggcg ctaggcgct ggcaagtga 5100
gcggtcacgc tgcgcgtaac caccaccccc gccgcgtta atgcccgtc acagggcgcg 5160
tcgcccatt cgccattcag gctacgcaac tgttgggaag ggcgatcggt gcgggcctct 5220
tcgctattac gccagctggc gaagggggga tgtgctgcaa ggcgattaag ttgggtaacg 5280
ccaggggtttt ccagtcacg acgttgtaaa acgacggcca gtgaatt 5327

```

SEQ ID NO:83

```

Met Asp Glu Ala Asp Arg Arg Leu Leu Arg Arg Cys Arg Leu Arg Leu
1          5          10          15
Val Glu Glu Leu Gln Val Asp Gln Leu Trp Asp Ala Leu Leu Ser Arg
20          25          30
Glu Leu Phe Arg Pro His Met Ile Leu Asp Ile Gln Arg Ala Gly Ser
35          40          45
Gly Ser Arg Arg Asp Gln Ala Arg Gln Leu Ile Ile Asp Leu Glu Thr

```

50	55	60
Arg Gly Ser Gln Ala Leu Pro Leu Phe Ile Ser Cys Leu Glu Asp Thr		
65	70	75
Gly Gln Asp Met Leu Ala Ser Phe Leu Arg Thr Asn Arg Gln Ala Ala		
	85	90
Lys Leu Ser Lys Pro Thr Leu Glu Asn Leu Thr Pro Val Val Leu Arg		
	100	105
Pro Glu Ile Arg Lys Pro Glu Val Leu Arg Pro Glu Thr Pro Arg Pro		
	115	120
Val Asp Ile Gly Ser Gly Gly Phe Gly Asp Val Gly Ala Leu Glu Ser		
	130	135
Leu Arg Gly Asn Ala Asp Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys		
	145	150
Gly His Cys Leu Ile Asn Asn Val Asn Phe Cys Arg Glu Ser Gly		
	165	170
Leu Arg Thr Arg Thr Gly Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg		
	180	185
Arg Phe Ser Ser Leu His Phe Met Val Glu Val Lys Gly Asp Leu Thr		
	195	200
Ala Lys Lys Met Val Leu Ala Leu Leu Glu Leu Ala Gln Gln Asp His		
	210	215
Gly Ala Leu Asp Cys Cys Val Val Ile Leu Ser His Gly Cys Gln		
	225	230
Ala Ser His Leu Gln Phe Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys		
	245	250
Pro Val Ser Val Glu Lys Ile Val Asn Ile Phe Asn Gly Thr Ser Cys		
	260	265
Pro Ser Leu Gly Gly Lys Pro Lys Leu Phe Phe Ile Gln Ala Ser Gly		
	275	280
Gly Glu Gln Lys Asp His Gly Phe Glu Val Ala Ser Thr Ser Pro Glu		
	290	295
Asp Glu Ser Pro Gly Ser Asn Pro Glu Pro Asp Ala Thr Pro Phe Gln		
	305	310
Glu Gly Leu Arg Thr Phe Asp Gln Leu Asp Ala Ile Ser Ser Leu Pro		
	325	330
Thr Pro Ser Asp Ile Phe Val Ser Tyr Ser Thr Phe Pro Gly Phe Val		
	340	345
Ser Trp Arg Asp Pro Lys Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp		
	355	360
Asp Ile Phe Glu Gln Trp Ala His Ser Glu Asp Leu Gln Ser Leu Leu		
	370	375
Leu Arg Val Ala Asn Ala Val Ser Val Lys Gly Ile Tyr Lys Gln Met		
	385	390
Pro Gly Cys Phe Asn Phe Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser		
	405	410
		415

SEQ ID NO:84

gaattccggg	ctggattgag	aagccgcaac	tgtgactctg	catcatgaat	actctgtctg	60
aaggaaatgg	caccttttgc	atccatcttt	tgaagatgct	atgtcaaagc	aacccttcca	120
aaaatgtatg	ttattctcct	gcgagcatct	cctctgtctc	agctatgggt	ctcttgggtg	180
caaagggaca	gacggcagtc	cagatatctc	aggcacttgg	tttgaataaa	gaggaaggca	240
tcacatcagg	tttccagttg	cttctcagga	agctgaacaa	gccagacaga	aagtactctc	300
ttagagtggc	caacaggctc	tttgcagaca	aaacttgtga	agtcctccaa	acctttaagg	360
agtcctctct	tcacttctat	gactcagaga	tggagcagct	ctcctttgct	gaagaagcag	420
aggtgtccag	gcaacacata	aacacatggg	tctccaaaca	aactgaaggt	aaaattccag	480
agttgtttgc	aggtggctcc	gtcgtattcag	aaaccaggct	ggttctcctc	aatgccttat	540
attttaaagg	aaagtggcat	caaccattta	acaaagagta	cacaatggac	atgcccttta	600
aaataaacia	ggatgagaaa	aggccagtg	agatgatgtg	tcgtgaagac	acataatacc	660
tcgcctatgt	gaaggagggtg	caggcgcaag	tgtgtgtgat	gccatatgaa	ggaatggagc	720
tgaacttggg	ggttctgtct	ccagatgagg	gtgtggacct	cagcaagggtg	gaaaacaatc	780
tcacttttga	gaagttaaca	gcctggatgg	aagcagattt	tatgaagagc	actgatgttg	840
aggttttcc	tcacaaat	aaactccaag	aggattatga	catggagtct	ctgtttcagc	900
gcttgggagt	ggtggatgtc	ttccaagagg	acaaggctga	cttatcagga	atgtctccag	960
agagaaacct	gtgtgtgtcc	aagtttgttc	accagagtgt	agtgagagtc	aatgaggaag	1020
gcacagaggc	tgacgagccc	tctgccatca	tagaattttg	ctgtgcctct	tctgtcccaa	1080
cattctgtgc	tgaccacccc	ttccttttct	tcacatcagg	caacaaagca	aacagcatcc	1140
tgttctgtgg	caggttctca	tctccataaa	gacacatata	ctacacaggg	agagttctct	1200
cttcagtatc	cctaccactc	ctacagctct	gtcaagatgg	gcaagtaggg	ggaagtcatg	1260
ttctaagatg	aagacacttt	ccttctctgt	cagcctgatc	ttataatgcc	tgcatccaac	1320
tctccctgtc	ttgaatgcac	ctatgccctt	taccagggtta	tgtctaatga	tgccaaatac	1380
ctctgtctat	gtgattgatt	gtagccctag	ccagtaattt	atagccagtt	agaactgact	1440
tgactgtgca	agaatgctat	aatggagcta	gagagaaggc	acaaacacta	ggaaagggtg	1500
ctgtttttgc	agaggacaca	gggacatttc	ccaccactca	catggctgct	tacaacctct	1560
ggaaattcca	gtttctgtcc	atgacttgat	tccttttctt	ggcttctact	ggctccagca	1620
tcctgcacat	acatgtatcg	tcatttcagtt	acacacaaac	aagtaaaatt	ttaaaaataa	1680
ataaaaaatt	aaagagagag	tctaaaaatt	tagtaatggg	tagataatag	ctgctattgt	1740
gcctttttca	ggttttaatt	tcattattct	tgtgtataaa	gtcaataaatt	tataggaaaa	1800
catcagtgcc	ccggaatttc					1819

SEQ ID NO:85

Met Asn Thr Leu Ser Glu Gly Asn Gly Thr Phe Ala Ile His Leu Leu
 1 5 10 15
 Lys Met Leu Cys Gln Ser Asn Pro Ser Lys Asn Val Cys Tyr Ser Pro
 20 25 30
 Ala Ser Ile Ser Ser Ala Leu Ala Met Val Leu Leu Gly Ala Lys Gly
 35 40 45
 Gln Thr Ala Val Gln Ile Ser Gln Ala Leu Gly Leu Asn Lys Glu Glu
 50 55 60
 Gly Ile His Gln Gly Phe Gln Leu Leu Leu Arg Lys Leu Asn Lys Pro
 65 70 75 80
 Asp Arg Lys Tyr Ser Leu Arg Val Ala Asn Arg Leu Phe Ala Asp Lys
 85 90 95
 Thr Cys Glu Val Leu Gln Thr Phe Lys Glu Ser Ser Leu His Phe Tyr
 100 105 110
 Asp Ser Glu Met Glu Gln Leu Ser Phe Ala Glu Glu Ala Glu Val Ser
 115 120 125
 Arg Gln His Ile Asn Thr Trp Val Ser Lys Gln Thr Glu Gly Lys Ile
 130 135 140
 Pro Glu Leu Leu Ser Gly Gly Ser Val Asp Ser Glu Thr Arg Leu Val
 145 150 155 160
 Leu Ile Asn Ala Leu Tyr Phe Lys Gly Lys Trp His Gln Pro Phe Met
 165 170 175
 Lys Glu Tyr Thr Met Asp Met Pro Phe Lys Ile Asn Lys Asp Glu Lys
 180 185 190
 Arg Pro Val Gln Met Met Cys Arg Glu Asp Thr Tyr Asn Leu Ala Tyr
 195 200 205
 Val Lys Glu Val Gln Ala Gln Val Leu Val Met Pro Tyr Glu Gly Met
 210 215 220
 Glu Leu Ser Leu Val Val Leu Leu Pro Asp Glu Val Asp Leu Ser
 225 230 235 240
 Lys Val Glu Asn Asn Leu Thr Phe Glu Lys Leu Thr Ala Trp Met Glu
 245 250 255
 Ala Asp Phe Met Lys Ser Thr Asp Val Glu Val Phe Leu Pro Lys Phe
 260 265 270
 Lys Leu Gln Glu Asp Tyr Asp Met Glu Ser Leu Phe Gln Arg Leu Gly
 275 280 285
 Val Val Asp Val Phe Gln Glu Asp Lys Ala Asp Leu Ser Gly Met Ser
 290 295 300
 Pro Glu Arg Asn Leu Cys Val Ser Lys Phe Val His Gln Ser Val Val
 305 310 315 320
 Glu Ile Asn Glu Glu Gly Thr Glu Ala Ala Ala Ser Ala Ile Ile
 325 330 335
 Glu Phe Cys Cys Ala Ser Ser Val Pro Thr Phe Cys Ala Asp His Pro
 340 345 350
 Phe Leu Phe Phe Ile Arg His Asn Lys Ala Asn Ser Ile Leu Phe Cys
 355 360 365
 Gly Arg Phe Ser Ser Pro
 370

SEQ ID NO:86

atgaatactc tgtctgaagg aaatggcacc ttgcccattc atcttttgaa gatgctatgt 60
 caaagcaacc cttccaaaaa tgtatgttat tctcctgcga gcattctcctc tgctctagct 120
 atggttctct tgggtgcaca gggacagacg gcagtcacaga tatctcaggc acttggtttg 180
 aataaagagg aaggcatcca tcagggtttc cagttgcttc tcaggaagct gaacaagcca 240
 gacagaaagt actctcttag agtggccaac aggcctctttg cagacaaaac ttgtgaagtc 300
 ttccaaacct ttaaggagtc ctctcttcac ttctatgact cagagatgga gcagctctcc 360
 ttgtctgaag aagcagaggt gtccaggcaa cacataaaca catgggtctc caaacaact 420
 gaaggtaaaa ttccagaggt gttgtcaggt ggctccgctc attcagaaac caggctgggt 480
 ctcatcaatg ccttatattt taaaggaaag tggcatcaac catttaacaa agagtacaca 540
 atggacatgc cctttaaaat aaacaaggat gagaaaaggc cagtgcagat gatgtgtcgt 600
 gaagacacat ataacctcgc ctatgtgaag gaggtgcagg cgcaagtgt ggtgatgcca 660
 tatgaaggaa tggagctgag cttggtggtt ctgctcccag atgagggtgt ggacctcagc 720
 aaggtggaaa acaatctcac ttttgagaag ttaacagcct ggatggaagc agattttatg 780
 aagagcactg atgttgaggt ttcccttcca aaattttaaac tccaagagga ttatgacatg 840
 gagtctctgt ttccagcgtt gggagtgggt gatgtcttcc aagaggacaa ggctgactta 900
 tcaggaatgt ctccagagag aaacctgtgt gtgtccaagt ttgttccaca gagtgtagt 960
 gagatcaatg aggaaggcag agaggctgca gcagcctctg ccatcataga attttgctgt 1020
 gcctcttctg tcccaacatt ctgtgctgac caccctcttc ttttcttcat caggcacaac 1080
 aaagcaaaac gcattcctgt ctgtggcagg ttctcatctc cataa 1125

SEQ ID NO:87

Met Asn Thr Leu Ser Glu Gly Asn Gly Thr Phe Ala Ile His Leu Leu
 1 5 10 15
 Lys Met Leu Cys Gln Ser Asn Pro Ser Lys Asn Val Cys Tyr Ser Pro
 20 25 30
 Ala Ser Ile Ser Ser Ala Leu Ala Met Val Leu Leu Gly Ala Lys Gly
 35 40 45
 Gln Thr Ala Val Gln Ile Ser Gln Ala Leu Gly Leu Asn Lys Glu Glu
 50 55 60
 Gly Ile His Gln Gly Phe Gln Leu Leu Leu Arg Lys Leu Asn Lys Pro

[illegible]

gatgcggttg	gctctatggc	ttctgagggc	gaaagaacca	gctggggctc	taggggggat	2400
ccccacgcgc	cctgtagcgg	cgcattaagc	gcggcggttg	tggtggttac	gcgcagcgtg	2460
accgctacac	ttgccagcgc	cctagcgcgc	gctcctttcg	ctttcttccc	ttcctttctc	2520
gccacggttcg	ccggcttttc	ccgtcaagct	ctaaatcggg	gcatcccttt	aggggtccga	2580
tttagtgctt	tacggcacct	cgaccccaaa	aaacttgatt	aggggtgatg	ttcacgtagt	2640
gggcccacgc	cctgatatag	ggtttttcgc	cctttgacgt	tgaggtccac	gttctttaat	2700
agtggactct	tggtccaaac	tggaacaaca	ctcaacccta	tctcggctca	ttcttttgat	2760
ttataaggga	ttttggggat	ttcggcctat	tggttaaaaa	atgagctgat	ttacacaaaa	2820
tttaacgcga	attaattctg	tggaatgtgt	gtcagttagg	gtgtggaaag	tccccaggct	2880
ccccaggcag	gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	caggtgtgga	2940
aagtcccccag	gctccccagc	aggcagaagt	atgcaaaagc	tgcatctcaa	ttagtcagca	3000
accatagctc	gccccctaac	tccgcccac	ccgcccctaa	ctccgcccag	ttccgcccac	3060
tctcgcgcgc	atggctgact	aatttttttt	atttatgcag	aggccgaggc	cgccctctgc	3120
tctgagctat	tcagagaatg	gtgaggaggc	ttttttggag	gcctaggctt	ttgcaaaaag	3180
ctcccgggag	cttgatatgc	catttttcga	tctgatcaag	agacaggatg	aggatcggtt	3240
cgcatgattg	aacaagatgg	attgcacgca	ggttctccgg	ccgcttgggt	ggagaggcta	3300
ttcggctatg	actgggcaca	acagacaatc	ggctgctctg	atgccgcctg	gttccggctg	3360
tcagcgcagg	ggcgcccggt	tctttttgtc	aagaccgacc	tgctccggtg	cctgaatgaa	3420
ctgcaggacg	aggcagcgcg	gctatcgtgg	ctggccacga	cgggcgcttc	ttgcgcagct	3480
gtgctcgacg	ttgtcacgga	agcgggaagg	gactggctgc	tattggcgga	agtgccgggg	3540
caggatctcc	tgctatctca	ccttgctcct	gccgagaaag	tatccatcat	ggctgatgca	3600
atgcggcgcc	tgcatcggct	tgatccggct	acctgcccac	tcgaccacca	agcgaacat	3660
cgcatcgagc	gagcacgtac	tcggatggaa	gccggtcttg	tcgatcagga	tgatctggac	3720
gaagagcatc	atgggctcgc	gccagccgaa	ctgttcgcca	ggctcaaggc	gcgcattgcc	3780
gacggcgagg	agctcgtcgt	gacccatggc	gatgcctgct	tgccgaatat	catggtggaa	3840
aatggccgct	tttctggatt	catcgactgt	ggccggctgg	gtgtggcgga	ccgctatcag	3900
gacatagcgt	tggtctaccg	tgatattgct	gaagagcttg	gcggcgcaatg	ggctgaccgc	3960
ttcctcgtgc	tttaaggatc	cgccgctccc	gattcgcagc	gcctcgcctt	ctatcgcctt	4020
cttgacgagt	cttcttgagc	gggactctgg	ggttcgaaat	gaccgaccaa	gcgacgcccc	4080
acctgccatc	acgagatttc	gattccaccg	ccgccttcta	tgaaagggtg	ggcttcggaa	4140
tcgtttttccg	ggacgcgcgc	tggtatgacc	tccagcgcgg	ggatctcatg	ctggagttct	4200
tcgcccaccg	caacttggtt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	4260
caaatttcac	aaataaagca	tttttttcac	tgcatcttag	ttgtgggttg	ttccaaactca	4320
tcaatgtatc	ttatcatgtc	tgtataccgt	cgacctctag	ctagagcttg	gcgtaatcat	4380
ggcatagctc	gtttctctgt	tgaaattggt	atccgctcac	aattccacac	aacatacag	4440
ccggaagcat	aaagtgtaaa	gcctgggggt	cctaagttagt	gagctaaact	acattaattg	4500
cgttgcgctc	actgcccgct	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	4560
tcggccaacg	cgcggggaga	ggcggtttgc	gtattggggc	ctcttcgcgt	tcctcgtcga	4620
ctgaactcgt	cgctcgcgtc	gttcggctgc	ggcgagcggt	atcagctcac	tcaaaaggcg	4680
taatacgggt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	4740
agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	4800
cccctgacga	gcatacacia	aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	4860
tataaagata	ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	4920
tgccgcttac	cggatacctg	tcgccttttc	tccttcgggg	aagcgtggcg	ctttctcaat	4980
gctcacgctg	taggtatctc	agttcgggtg	aggctggttc	ctccaagctg	ggctgtgtgc	5040
acgaaccccc	tggttcagcc	gacgcgtcgc	ccttatccgg	taactatcgt	cttgagtcca	5100
accggttaag	acacgactta	tcgccactgg	cagcagccac	tggtaacagg	attagcagag	5160
cgaggtatgt	aggcggtgct	acagagtctc	tgaagtgggt	gcctaactac	ggctacacta	5220
gaaggacagt	atttggtatc	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	5280
gtagctcttg	atccggcaca	caaaaccacg	ctggtagcgg	tggttttttt	gtttgcgaagc	5340
agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	5400
ctgacgctca	gtggaacgaa	aaactcagtt	aagggatttt	ggctatgaga	ttatcaaaaa	5460
ggatctctac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatt	taaagtatat	5520
atgagtaaac	ttgggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	5580
tctgtctatt	tcgttcaccc	atagttgcct	gactccccgt	cgtgtagata	actacgatac	5640
gggagggtct	accatctggc	cccagtgctg	caatgatacc	gcgagaccca	cgctcaccgg	5700
ctccagattt	atcagcaata	aaccagccag	ccggaagggc	cgagcgcaga	agtggtcctg	5760
caacttttatc	cgctccatc	cagtctatta	attggtgccc	ggaagctaga	gtaagttagt	5820
cgccagttaa	tagtttgcgc	aacgttggtg	ccattgctac	aggcatcggt	gtgtcacgct	5880
cgctggtttg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	5940
cccccatggt	gtgcaaaaaa	gcgggttagct	ccttcgggtc	tccgatcggt	gtcagaagta	6000
agttggccgc	agtgttatca	ctcatgttta	tggcagcact	gcataattct	cttactgtca	6060
tgccatccgt	aagatgcttt	tctgtgactg	gtgagtactc	aaaccaagtca	ttctgagaat	6120
agtgtatcgc	gcgaccgagt	tgctcttgcc	cggcgtaaat	acgggataat	accgcgccac	6180
atagcagaac	tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	6240
ggatctttacc	gctgttgaga	tccagttcga	tgtaaccac	tcgtgacccc	aaactgatct	6300
cagcatcttt	tacttttcacc	agcgtttctg	ggtgagcaaa	aacagggaag	caaaatgccg	6360
caaaaaaggg	aataaggggc	acacggaaat	gttgaaatac	catactcttc	ctttttcaat	6420
attattgaag	catttatcag	ggttattgtc	tcatgagcgg	atacatattt	gaatgtattt	6480
agaaaaataa	acaaataggg	gttccgcgca	catttccccg	aaaagtgcga	cctgacgtc	6539

SEQ ID NO:89

gacggatcgg	gagatctccc	gatccccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aaagccagat	ctgctccctg	cttggtgtgt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgtttttcg	ctgcttcgcg	atgtacgggc	cagatatacg	cggttagact	240
gattattgac	tagttattaa	tagtaataca	ttacgggggc	attagttcat	agcccatata	300
tgaggttccg	ctgtacataa	cttagcgtaa	atggcccgcc	tggtgacgac	cccaacgacc	360
cccgcccat	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tattttacgtt	aaactgcccc	cttggcagta	catcaagtgt	480

atcatatgcc	aagtagcccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tggaactttc	ctacttgcca	gtacatctac	gtattagtca	600
tcgctattac	catgggtgatg	cggtttttggc	agtacatcaa	tgggcggtga	tagcggtttg	660
actcacgggg	attttccagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaaatcaacg	ggacttttcca	aaatgtcgta	acaactccgc	cccatgtacg	caaatgggag	780
gtaggcggtg	acgggtgggag	gtctatatata	gcagagctct	ctgggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaacgg	gccctctaga	ctcgagcggc	cgccactgtg	ctggatatct	gcagaaattca	960
tgaatactct	gtctgaagga	aatggcacct	ttgccatcca	tcttttgaag	atgctatgtc	1020
aaagcaaccc	ttccaaaaat	gtatgttatt	ctcctgcgag	catctcctct	gctctagcta	1080
tggttctctt	gggtgcgaag	ggacagacgg	cagttccagat	atctcaggca	cttggtttga	1140
ataaagagga	aggcatccat	cagggtttcc	agttgtctct	caggaagctg	aacaagccag	1200
acagaaagta	ctctcttaga	gtggccaaca	ggctctttgc	agacaaaact	tgtgaagtcc	1260
ttccaaacctt	taaggagtcc	tctcttcact	tctatgactc	agagatggag	cagctctcct	1320
ttgctgaaga	agcagaggtg	tcaggcaaac	acataaacac	atgggtctcc	aaacaaactg	1380
aaggtaaaat	tccagagttg	ttgtcagggtg	gctccgtcga	ttcagaaacc	aggctgggtc	1440
tcataaatgc	cttatatttt	aaaggaaagt	ggcatcaacc	atttaacaaa	gagtaacaaa	1500
tggacatgcc	cttttaaaata	aacaaggatg	agaaaaggcc	agtgcagatg	atgtgtcgtg	1560
aagacacata	taacctcgcc	tatgtgaagg	agggtgcaggc	gcaagtgtctg	gtgatgccat	1620
atgaaggaat	ggagctgagc	ttggtgtgtc	tgtctccaga	tgagggtgtg	gacctcagca	1680
agggtgaaaa	caatctcact	tttgagaagt	taacagcctg	gatggaagca	gattttatga	1740
agagcaatga	tgttgaggtt	ttccttccaa	aattttaaact	ccaagaggat	tatgacatgg	1800
agtctctgtt	tcagcgcttg	ggagtgtgtg	atgtcttcca	agaggacaag	gctgacttat	1860
caggaaatgtc	tcagagagga	aacctgtgtg	tgtccaaagt	tgttcaccag	agtgtagtgg	1920
agatcaatga	ggagggcaga	gagcctgcgc	cagcctctgc	catcatagaa	ttttgtctgtg	1980
cctcttctgt	cccaacatct	tgtgtgacc	accccttctc	tttcttctc	aggcacaaca	2040
aagcaaacag	catcctgttc	tgtggcaggt	tctcatctcc	ataaggatcc	gagctcggta	2100
ccaagcttaa	gtttaaacccg	ctgatcagcc	tcgactgtgc	cttctagtgtg	ccagccatct	2160
gttggtttg	ctcccccggt	gccttctctg	acccctggaag	gtgccactcc	cactgtcctt	2220
tcctaataaaa	atgaggaaat	tgcatcgcat	tgtctgagta	gggtgtcattc	tattctgggg	2280
gggtgggtgtg	ggcaggacag	caagggggag	gattgggaag	acaatagcag	gcatgtgtgg	2340
gatcggtgtg	gctctatggc	ttctgaggcg	gaaagaacca	gctggggctc	taggggggat	2400
ccccacgcgc	cctgtagcgg	cgcatcgaag	gcggcggtgtg	tgggtgttac	gcgcagcgtg	2460
accgctacac	ttgccagcgc	cctagcgcgc	gctccttctc	cttctctccc	ttccttctct	2520
gccacgttgc	ccggctttcc	ccgtcaagct	ctaaatcggg	gcaccccttt	agggttccga	2580
tttagtgctt	taaggcaccct	cgaccccaaa	aaacttgatt	agggtgatgg	ttcacgtagt	2640
gggccatcgc	cctgatagac	ggtttttctc	cctttgacgt	tggagtccac	gttctttaat	2700
agtggaactct	tgttccaaac	tgaacaacaa	ctcaacccta	tctcgttcta	ttcttttgat	2760
ttataaaggga	ttttggggat	ttcggcctat	tggtttaaaaa	atgagctgat	tttaacaaaaa	2820
tttaacgcga	attaattctg	tggaaatgtg	gtcagttagg	gtgtggaaaag	tccccaggct	2880
ccccaggcag	gcagaagtat	gcaaagcatg	catctcaatt	agtacagcaac	cagggtgtgga	2940
aagtccccag	gctccccagc	aggcagaagt	atgcaaaagca	tgcatctcaa	ttagtcagca	3000
accatagtc	cgccccctac	tcgccccctc	ccgccccctc	ctccgcccag	ttccgcccct	3060
tctccgcccc	atggctgact	aatttttttt	atttatgcag	aggccgaggc	cgccctctgc	3120
tcctgagctat	tcagaaagta	gtgaggaggg	ttttttggag	gcctaggtct	ttgcaaaaag	3180
ctcccggtg	cttcttatatc	catttttcgga	tctgatcaag	agacaggatg	aggatcgttt	3240
cgcatgattg	aacaagatgg	attgcacgca	ggttctccgg	ccgcttgggt	ggagaggcta	3300
ttcggctatg	actgggcaca	acagacaatc	ggctgtctctg	atgccgcggt	gttccggctg	3360
tcagcgcagg	ggcgccccgt	tctttttgtc	aagaccgaac	tgtccggtgc	cctgaatgaa	3420
ctgcaggacg	aggcagcggt	gctatcgtgg	ctggccacga	cgggcggttc	ttgcgcagct	3480
gtgctcgacg	ttgtcactga	agcgggaagg	gactggctgc	tattggcgga	agtgcggggg	3540
caggatctcc	tgtcatctca	ccttgcctct	gcggagaaag	tatccatcat	ggctgatgca	3600
atgcggcggt	tgcatcggct	tgatccggct	acctgcccct	tcgaccacaa	agcgaacat	3660
cgcatcgagc	gagcacgtac	tcggatggaa	gccggtcttg	tcgatcagga	tgatctggac	3720
gaagagcatc	aggggctcgc	gccagccgaa	ctgttcgcca	ggctcaaggc	gcgcagctcc	3780
gacggcgagg	atctcgtcgt	gacccatggc	gatgcctgtg	tgccgaatat	catggtggaa	3840
aatggccgct	tttctggatt	catcgactgt	ggccggctgg	gtgtggcgga	ccgctatcag	3900
gacatagcgt	tggctaccgg	tgatattgct	gaagagcttg	gcggcggaatg	ggctgaccgc	3960
ttctctgtgc	tttaacggtat	cgccgctccc	gattcgcagc	gcacgcctct	ctatcgctct	4020
cttgacaggt	tcttctgagc	gggactctgg	ggttcgaat	gaccgaccaa	gcgacgccc	4080
acctgccatc	acgagatttc	gattccaccg	ccgccttcta	tgaagggttg	ggcttcggaa	4140
tcgtttttccg	ggacgcgggc	tggatgatcc	tccagcgcgg	ggatctcatg	ctggagttct	4200
tcgcccaccc	caacttggtt	attgcagctt	ataatgggtt	caaataaagc	aatagcatca	4260
caaattttcc	aaataaaagca	tttttttcc	tgcattctag	ttgtggtttg	ttcaaaactca	4320
tcaatgtatc	ttatcatgtc	tgtataccgt	cgacctctag	ctagagcttg	gcgtaatcat	4380
ggcatagct	gtttcctgtg	tgaattgtt	atccgctcac	aattccacac	aacatacag	4440
ccggaagcat	aaagtgtaaa	gcttgggtg	cctaattgagt	gagctaaactc	acattaattg	4500
cggtgcgctc	actgcccgtc	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	4560
tcggcccaacg	cgcggggaga	ggcggtttgc	gtattggggc	ctcttcgctc	tctcgcctca	4620
ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	tcaaaggcgg	4680
taatacggtt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	4740
agcaaaaggc	caggaaacgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccggc	4800
cccctgacga	gcatacacia	aatcgacgtc	caagtacag	gtggcgaaac	ccgacaggac	4860
tataaagata	ccaggcgctt	cccctggaa	gctccctctg	gcgctctctc	gttccgaccc	4920
tgccgcttac	cggataacctg	tccgcttttc	tcccttcggg	aagcgtggcg	ctttctcaat	4980
gctcacgctg	taggtatctc	agttcggtgt	aggtcgcttc	ctccaaagctg	ggctgtgtgc	5040
acgaaccccc	cgttcagccc	gacgcgtcgc	ccttatccgg	taactatcgt	cttgagtcca	5100
accgggtaag	acacgactta	tcgccactgg	cagcagccac	tggttaacag	attagcagag	5160
cgagggtatgt	aggcggtgct	acagagttct	tgaagtgtgtg	gcctaactac	ggctacacta	5220
gaaggacagt	atttgggtatc	tcgctctcgc	tgaagccagt	taccttcgga	aaaagagttg	5280

```

gtagctcttg atccggcaaa caaaccacgg ctggtagcgg tgggtttttt gtttgcaagc 5340
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg 5400
ctgacgctca gtggaacgaa aactcacggt aagggatttt ggtcatgaga ttatcaaaaa 5460
ggatctttcac ctgagtcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 5520
atgagtaaac ttgggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 5580
tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 5640
gggagggtct accatctggc cccagtgtct caatgatacc gcgagaccca cgctcacccg 5700
ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgagca agtggtcctg 5760
caactttatc cgctccatc cagtctatta attgttgccg ggaagctaga gtaagttagt 5820
cgccagttaa tagtttgccg aacgttgttg ccattgtctac aggcacgtcg gtgtcacgct 5880
cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 5940
cccccatgtt gtgcaaaaaa gcgggttagct ccttcgggtc tccgatcggt gtcagaagta 6000
agttggccgc agtggttatca ctcatggtta tggcagcact gcataattct cttactgtca 6060
tgccatcgtt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 6120
agtgtatgcg gcgaccgagt tgctcttgcc cggcgctcaat accgcgccac 6180
atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 6240
ggatcttacc gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt 6300
cagcatcttt tactttcacg acggtttctg ggtgagcaaa aacaggaagg caaaatgccg 6360
caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 6420
attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 6480
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc a cctgacgtc 6539

```

SEQ ID NO:90

```

atggatgacc agcgcgacct tatctccaac aatgagcaac tgcccatgct gggccggcgc 60
cctggggccc cgagagcaaa gtgcagccgc ggagccctgt acacaggctt ttccatcctg 120
gtgactctgc tctcgcgtgg ccaggccacc accgcctact tcctgtacca gcagcagggc 180
cggctggaca aactgacagt caccctccag aacctgcagc tggagaacct gcgcatgaag 240
cttgccaagt tcgtggctgc ctggaccctg aaggctgccg ctgccctgcc ccaggggccc 300
atgcagaatg ccaccaagta tggcaacatg acagaggacc atgtgatgca cctgtccag 360
aatgctgacc ccttgaaggt gtacccgcca ctgaagggga gcttcccgga gaacctgaga 420
caccttaaga acaccatgga gaccatagac ttggaaggtct ttgagagctg gatgcaccat 480
tggtcctgtt ttgaaatgag caggcactcc ttggagcaaa agccactga cgctccaccg 540
aaagtactga ccaagtgcc aagaagagtc agccacatcc ctgctgtcca cccgggttca 600
ttcaggccca agtgcgacga gaacggcaac tatctgccac tccagtgtca tgggagcatc 660
ggctactgct ggtgtgtctt ccccaacggc acggaggtcc ccaacaccag aagccgagg 720
caccataact gcagtgaagt actggaactg gaggaccctg cttctgggct ggggtgtgacc 780
aagcaggatc tgggcccagt ccccatgtga 810

```

SEQ ID NO:91

```

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

```

SEQ ID NO:92

```

Lys Pro Val Ser Gln Met Arg Met Ala Thr Pro Leu Leu Met Arg Pro
1 5 10 15
Met

```

SEQ ID NO:93

Ala Lys Phe Val Ala Ala Trp Thr Leu Lys Ala Ala Ala
1 5 10

SEQ ID NO:94

```

atgcgttgcc tggctccacg ccctgctggg tcctacctgt cagagcccca aggcagctca 60
cagtgtgcca ccattggagt ggggccccta gaaggtggct acctggagct tcttaacagc 120
gatgctgacc cctgtgcttc taccacttct atgaccagat ggacctggct ggagaagaag 180
agattgagct ctactcagaa cccgacacag acaccatcaa ctgcgaccag ttcagcaggc 240
tgttgtgtga catggaaggt gatgaagaga ccagggaggc ttatgccaat atcgcggaac 300
tggaccagta tgtcttcacg gactccacgc tggagggcct gagcaaggac attttcaagc 360
acataggacc agatgaagtg atcgggtgaga gtatggagat gccagcagaa gttgggcaga 420
aaagtcagaa aagacccttc ccagaggagc ttccggcaga cctgaagcac tggaaagccag 480
ctgagccccc cactgtggtg actggcagtc tcctagtggg accagtgagc gactgctcca 540
ccctgcccctg cctgcccactg cctgcgctgt tcaaccagga gccagcctcc ggccagatgc 600
gcttgagaaa aaccgaccag attcccctgc ctttctccag ttcctcggtt agctgctga 660
atctccctga gggacccatc cagtttgtcc ccaccatctc cactctgccc catgggctct 720
ggcaaatctc tgaggctgga acaggggtct ccagtatatt catctaccat ggtgaggtgc 780
ccaggcccaag ccaagtaccc cctcccagtg gattcactgt ccacggcctc ccaacatctc 840
cagaccggcc aggtccacc agccccttcg ctccatcagc cactgacctg cccagcatgc 900
ctgaacctgc cctgacctcc cgagcaaaac tgacagagca caagacgtcc cccaccaaat 960
gcccggcagc tggagaggtc tccaacaagc ttccaaaatg gcctgagccg gtggagcagt 1020
tctaccgctc actgcaggac acgtatggtg ccgagcccg caggcccgat ggcatcctag 1080
tggaggtgga tctggtgcag gccaggctgg agaggagcag cagcaagagc ctggagcggg 1140
aactggccac ccgggactgg gcagaacggc agctggccca aggagccctg gctgaggtgc 1200
tgttggctgc caaggagcac cggcggccgc gtgagacacg agtgattgct gtgctgggca 1260
aagctggtca gggcaagagc tattgggctg gggcagtgag ccgggcccctg gcttgtggcc 1320
ggcttcccca gtacgacttt gtcttctctg tcccctgcca ttgcttgaac cgtccggggg 1380
atgcctatgg cctgcaggat ctgctcttct ccctgggccc acagccactc gtggcggccg 1440
atgaggtttt acgccacatc ttgaagagac ctgaccgcgt tctgctcctc ctgacgggct 1500
tcgaggagct ggaagcgcaa gatggcttcc tgcacagcac gtgctggacc gcaccggcgg 1560
agccctgctc cctccggggg ctgctggccg gccttttcca gaagaagctg ctccgaggtt 1620
gcaccctctc cctccacagc cggcccgggg gccgctggtt ccagagcctg agcaaggccg 1680
acgcccattt tgagctgtcc ggcttctcca tggagcaggc ccaggcatac gtgatgcgct 1740
actttgagag ctagggtatg acagagcacc aagacagagc cctgacgtcc ctccgggacc 1800
ggccacttct tctcagtcac agccacagcc ctactttgtg ccgggcagtg tgccagctct 1860
cagaggccct ctgagagctt ggggaggagc ccaagctgcc ctccacgtcc acgggactct 1920
atgtcggcct gctgggcccgt gcagccctcg acagccccc cggggccctg gcagagctgg 1980
ccaagctggc ctgggagctg ggcgcgagac atcaaaagtac cctacaggag gaccagttcc 2040
catccgcaga cgtgagagacc tggcgatgg ccaaaggctt agtccaacac ccaccggcg 2100
ccgcagagtc cgagctggcc ttcccacgtt tcctctgca atgcttctct ggggcccctgt 2160
ggctggtctt gactggcgaa atcaaggaca aggagctccc gcagtaacct gcattgaccc 2220
caaggaagaa gaggccctat gacaactggc tggagggcgt gccacgtttt ctggctgggc 2280
tgatcttcca gcctcccgc cgtgcccgtg gaggccatct cgggcccctc gcggctgct 2340
cggtgagacg gaagcagaag gtgcttgca ggtacctgaa cgggctgcag ccggggacac 2400
tgccggcgcg gcagctgtcg gagctgtcgc actgcgccc cgaggcccgag gaggtggaa 2460
tttggcagca cgtggtacag gagctcccgc gccgctcttc ttttctgggc acccgctca 2520
cgctcctga tgacatgta ctgggcaagg ccttgaggc ggccggccaa gacttctccc 2580
tggacctccg cagcactggc atttgcccct ctggattggg gagectcgtg ggactcagct 2640
gtgtcaccgg ttccagggtt gccttgagcg acacggtggc gctgtgggag tcccctgcagc 2700
agcatgggga gaccaagcta ctccaggcag cagaggagaa gttcaccatc gaccccttca 2760
aagccaagtc cctgaaggat gtggaagacc tgggaaagct tgtgcagact cagaggacga 2820
gaagtctctc ggaagacaca gctggggagc tccctgctgt tcgggacctc aagaaactgg 2880
agtttgccgt gggcccctgc tcaggccccc aggccttccc caaactgggt cggatcctca 2940
cggccttttc ctcccctgcag catctggacc tggatgcgct gagtgagaac aagatcgggg 3000
acgagggtgt ctgcagctc tcagccacct tcccctcagc gaagtccttg gaaacctca 3060
atctgtccca gaacaacatc actgacctgg gtgcctacaa actcgccgag gccctgcctt 3120
cgctcgctgc atcccgtctc aggcctaagct tgtacaataa ctgcatctgc gacgtgggag 3180
ccgagagctt ggtcgtgtg ctcccgaca tgggtgtcct ccgggtgatg gacgtccagt 3240
acaacaagtt cacggctgcc ggggcccagc agctcgctgc cagccttcgg aggtgtcctc 3300
atgtggagac gctggcgatg tggagcccca ccatccatt cagtgtccag gaacacctgc 3360
aacaacagga ttcacggatc agcctgagat ga 3392

```

SEQ ID NO:95

```

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

```

```

130      135      140
Arg Pro Phe Pro Glu Glu Leu Pro Ala Asp Leu Lys His Trp Lys Pro
145      150      155      160
Ala Glu Pro Pro Thr Val Val Thr Gly Ser Leu Leu Val Gly Pro Val
165      170      175
Ser Asp Cys Ser Thr Leu Pro Cys Leu Pro Leu Pro Ala Leu Phe Asn
180      185      190
Gln Glu Pro Ala Ser Gly Gln Met Arg Leu Glu Lys Thr Asp Gln Ile
195      200      205
Pro Met Pro Phe Ser Ser Ser Ser Leu Ser Cys Leu Asn Leu Pro Glu
210      215      220
Gly Pro Ile Gln Phe Val Pro Thr Ile Ser Thr Leu Pro His Gly Leu
225      230      235      240
Trp Gln Ile Ser Glu Ala Gly Thr Gly Val Ser Ser Ile Phe Ile Tyr
245      250      255
His Gly Glu Val Pro Gln Ala Ser Gln Val Pro Pro Pro Ser Gly Phe
260      265      270
Thr Val His Gly Leu Pro Thr Ser Pro Asp Arg Pro Gly Ser Thr Ser
275      280      285
Pro Phe Ala Pro Ser Ala Thr Asp Leu Pro Ser Met Pro Glu Pro Ala
290      295      300
Leu Thr Ser Arg Ala Asn Met Thr Glu His Lys Thr Ser Pro Thr Gln
305      310      315      320
Cys Pro Ala Ala Gly Glu Val Ser Asn Lys Leu Pro Lys Trp Pro Glu
325      330      335
Pro Val Glu Gln Phe Tyr Arg Ser Leu Gln Asp Thr Tyr Gly Ala Glu
340      345      350
Pro Ala Gly Pro Asp Gly Ile Leu Val Glu Val Asp Leu Val Gln Ala
355      360      365
Arg Leu Glu Arg Ser Ser Ser Lys Ser Leu Glu Arg Glu Leu Ala Thr
370      375      380
Pro Asp Trp Ala Glu Arg Gln Leu Ala Gln Gly Gly Leu Ala Glu Val
385      390      395      400
Leu Leu Ala Ala Lys Glu His Arg Arg Pro Arg Glu Thr Arg Val Ile
405      410      415
Ala Val Leu Gly Lys Ala Gly Gln Gly Lys Ser Tyr Trp Ala Gly Ala
420      425      430
Val Ser Arg Ala Trp Ala Cys Gly Arg Leu Pro Gln Tyr Asp Phe Val
435      440      445
Phe Ser Val Pro Cys His Cys Leu Asn Arg Pro Gly Asp Ala Tyr Gly
450      455      460
Leu Gln Asp Leu Leu Phe Ser Leu Gly Pro Gln Pro Leu Val Ala Ala
465      470      475      480
Asp Glu Val Phe Ser His Ile Leu Lys Arg Pro Asp Arg Val Leu Leu
485      490      495
Ile Leu Asp Gly Phe Glu Glu Leu Glu Ala Gln Asp Gly Phe Leu His
500      505      510
Ser Thr Cys Gly Pro Ala Pro Ala Glu Pro Cys Ser Leu Arg Gly Leu
515      520      525
Leu Ala Gly Leu Phe Gln Lys Lys Leu Leu Arg Gly Cys Thr Leu Leu
530      535      540
Leu Thr Ala Arg Pro Arg Gly Arg Leu Val Gln Ser Leu Ser Lys Ala
545      550      555      560
Asp Ala Leu Phe Glu Leu Ser Gly Phe Ser Met Glu Gln Ala Gln Ala
565      570      575
Tyr Val Met Arg Tyr Phe Glu Ser Ser Gly Met Thr Glu His Gln Asp
580      585      590
Arg Ala Leu Thr Leu Leu Arg Asp Arg Pro Leu Leu Leu Ser His Ser
595      600      605
His Ser Pro Thr Leu Cys Arg Ala Val Cys Gln Leu Ser Glu Ala Leu
610      615      620
Leu Glu Leu Gly Glu Asp Ala Lys Leu Pro Ser Thr Leu Thr Gly Leu
625      630      635      640
Tyr Val Gly Leu Leu Gly Arg Ala Ala Leu Asp Ser Pro Pro Gly Ala
645      650      655
Leu Ala Glu Leu Ala Lys Leu Ala Trp Glu Leu Gly Arg Arg His Gln
660      665      670
Ser Thr Leu Gln Glu Asp Gln Phe Pro Ser Ala Asp Val Arg Thr Trp
675      680      685
Ala Met Ala Lys Gly Leu Val Gln His Pro Pro Arg Ala Ala Glu Ser
690      695      700
Glu Leu Ala Phe Pro Ser Phe Leu Leu Gln Cys Phe Leu Gly Ala Leu
705      710      715      720
Trp Leu Ala Leu Ser Gly Glu Ile Lys Asp Lys Glu Leu Pro Gln Tyr
725      730      735
Leu Ala Leu Thr Pro Arg Lys Lys Arg Pro Tyr Asp Asn Trp Leu Glu
740      745      750
Gly Val Pro Arg Phe Leu Ala Gly Leu Ile Phe Gln Pro Ala Arg
755      760      765
Cys Leu Gly Ala Leu Leu Gly Pro Ser Ala Ala Ala Ser Val Asp Arg

```

```

770          775          780
Lys Gln Lys Val Leu Ala Arg Tyr Leu Lys Arg Leu Gln Pro Gly Thr
785          790          795          800
Leu Arg Ala Arg Gln Leu Leu Glu Leu Leu His Cys Ala His Glu Ala
805          810          815
Glu Glu Ala Gly Ile Trp Gln His Val Val Gln Glu Leu Pro Gly Arg
820          825          830
Leu Ser Phe Leu Gly Thr Arg Leu Thr Pro Pro Asp Ala His Val Leu
835          840          845
Gly Lys Ala Leu Glu Ala Ala Gly Gln Asp Phe Ser Leu Asp Leu Arg
850          855          860
Ser Thr Gly Ile Cys Pro Ser Gly Leu Gly Ser Leu Val Gly Leu Ser
865          870          875          880
Cys Val Thr Arg Phe Arg Ala Ala Leu Ser Asp Thr Val Ala Leu Trp
885          890          895
Glu Ser Leu Gln Gln His Gly Glu Thr Lys Leu Leu Gln Ala Ala Glu
900          905          910
Glu Lys Phe Thr Ile Glu Pro Phe Lys Ala Lys Ser Leu Lys Asp Val
915          920          925
Glu Asp Leu Gly Lys Leu Val Gln Thr Gln Arg Thr Arg Ser Ser Ser
930          935          940
Glu Asp Thr Ala Gly Glu Leu Pro Ala Val Arg Asp Leu Lys Lys Leu
945          950          955          960
Glu Phe Ala Leu Gly Pro Val Ser Gly Pro Gln Ala Phe Pro Lys Leu
965          970          975
Val Arg Ile Leu Thr Ala Phe Ser Ser Leu Gln His Leu Asp Leu Asp
980          985          990
Ala Leu Ser Glu Asn Lys Ile Gly Asp Glu Gly Val Ser Gln Leu Ser
995          1000          1005
Ala Thr Phe Pro Gln Leu Lys Ser Leu Glu Thr Leu Asn Leu Ser
1010          1015          1020
Gln Asn Asn Ile Thr Asp Leu Gly Ala Tyr Lys Leu Ala Glu Ala
1025          1030          1035
Leu Pro Ser Leu Ala Ala Ser Leu Leu Arg Leu Ser Leu Tyr Asn
1040          1045          1050
Asn Cys Ile Cys Asp Val Gly Ala Glu Ser Leu Ala Arg Val Leu
1055          1060          1065
Pro Asp Met Val Ser Leu Arg Val Met Asp Val Gln Tyr Asn Lys
1070          1075          1080
Phe Thr Ala Ala Gly Ala Gln Gln Leu Ala Ala Ser Leu Arg Arg
1085          1090          1095
Cys Pro His Val Glu Thr Leu Ala Met Trp Thr Pro Thr Ile Pro
1100          1105          1110
Phe Ser Val Gln Glu His Leu Gln Gln Gln Asp Ser Arg Ile Ser
1115          1120          1125
Leu Arg
1130

```

SEQ ID NO:96

```

1/1          31/11
ATG AGC CTG TGG CTG CCC AGC GAG GCC ACC GTG TAC CTG CCC CCC GTG CCC GTG AGC AAG
61/21          91/31
GTG GTG AGC ACC GAC GAG TAC GTG GCC AGG ACC AAC ATC TAC TAC CAC GCC GGC ACC AGC
121/41          151/51
AGG CTG CTG GCC GTG GGC CAC CCC TAC TTC CCC ATC AAG AAG CCC AAC AAC AAC AAG ATC
181/61          211/71
CTG GTG CCC AAG GTG AGC GGC CTG CAG TAC AGG GTG TTC AGG ATC CAC CTG CCC GAC CCC
241/81          271/91
AAC AAG TTC GGC TTC CCC GAC ACC AGC TTC TAC AAC CCC GAC ACC CAG AGG CTG GTG TGG
301/101          331/111
GCC TGC GTG GGC GTG GAG GTG GGC AGG GGC CAG CCC CTG GGC GTG GGC ATC AGC GGC CAC
361/121          391/131
CCC CTG CTG AAC AAG CTG GAC GAC ACC GAG AAC GCC AGC GCC TAC GCC AAC GCC GGC
421/141          451/151
GTG GAC AAC AGG GAG TGC ATC AGC ATG GAC TAC AAG CAG ACC CAG CTG TGC CTG ATC GGC
481/161          511/171
TGC AAG CCC CCC ATC GGC GAG CAC TGG GGC AAG GGC AGC CCC TGC ACC AAC GTG GCC GTG
541/181          571/191
AAC CCC GGC GAC TGC CCC CCC CTG GAG CTG ATC AAC ACC GTG ATC CAG GAC GGC GAC ATG
601/201          631/211
GTG GAC ACC GGC TTC GGC GCC ATG GAC TTC ACC ACC CTG CAG GCC AAC AAG AGC GAG GTG
661/221          691/231
CCC CTG GAC ATC TGC ACC AGC ATC TGC AAG TAC CCC GAC TAC ATC AAG ATG GTG AGC GAG
721/241          751/251
CCC TAC GGC GAC AGC CTG TTC TTC TAC CTG AGG AGG GAG CAG ATG TTC GTG AGG CAC CTG
781/261          811/271
TTC AAC AGG GCC GGC GCC GTG GGC GAG AAC GTG CCC GAC GAC CTG TAC ATC AAG GGC AGC
841/281          871/291
GGC AGC ACC GCC AAC CTG GCC AGC AGC AAC TAC TTC CCC ACC CCC AGC GGC AGC ATG GTG
901/301          931/311

```

ACC AGC GAC GCC CAG ATC TTC AAC AAG CCC TAC TGG CTG CAG AGG GCC CAG GGC CAC AAC
 961/321 991/331
 AAC GGC ATC TGC TGG GGC AAC CAG CTG TTC GTG ACC GTG GTG GAC ACC ACC AGG AGC ACC
 1021/341 1051/351
 AAC ATG AGC CTG TGC GCC GCC ATC AGC ACC AGC GAG ACC ACC TAC AAG AAC ACC AAC TTC
 1081/361 1111/371
 AAG GAG TAC CTG AGG CAC GGC GAG GAG TAC GAC CTG CAG TTC ATC TTC CAG CTG TGC AAG
 1141/381 1171/391
 ATC ACC CTG ACC GCC GAC GTG ATG ACC TAC ATC CAC AGC ATG AAC AGC ACC ATC CTG GAG
 1201/401 1231/411
 GAC TGG AAC TTC GGC CTG CAG CCC CCC CCC GGC GGC ACC CTG GAG GAC ACC TAC AGG TTC
 1261/421 1291/431
 GTG ACC AGC CAG GCC ATC GCC TGC CAG AAG CAC ACC CCC CCC GCC CCC AAG GAG GAC CCC
 ro
 1321/441 1351/451
 CTG AAG AAG TAC ACC TTC TGG GAG GTG AAC CTG AAG GAG AAG TTC AGC GCC GAC CTG GAC
 1381/461 1411/471
 CAG TTC CCC CTG GGC AGG AAG TTC CTG CTG CAG GCC GGC CTG AAG GCC AAG CCC AAG TTC
 1441/481 1471/491
 ACC CTG GGC AAG AGG AAG GCC ACC CCC ACC ACC AGC AGC ACC AGC ACC ACC GCC AAG AGG
 1501/501
 AAG AAG AGG AAG CTG TGA

SEQ ID NO:97

1/1 31/11
 Met ser leu trp leu pro ser glu ala thr val tyr leu pro pro val pro val ser lys
 61/21 91/31
 val val ser thr asp glu tyr val ala arg thr asn ile tyr tyr his ala gly thr ser
 121/41 151/51
 arg leu leu ala val gly his pro tyr phe pro ile lys lys pro asn asn asn lys ile
 181/61 211/71
 leu val pro lys val ser gly leu gln tyr arg val phe arg ile his leu pro asp pro
 241/81 271/91
 asn lys phe gly phe pro asp thr ser phe tyr asn pro asp thr gln arg leu val trp
 301/101 331/111
 ala cys val gly val glu val gly arg gly gln pro leu gly val gly ile ser gly his
 361/121 391/131
 pro leu leu asn lys leu asp asp thr glu asn ala ser ala tyr ala ala asn ala gly
 421/141 451/151
 val asp asn arg glu cys ile ser met asp tyr lys gln thr gln leu cys leu ile gly
 481/161 511/171
 cys lys pro pro ile gly glu his trp gly lys gly ser pro cys thr asn val ala val
 541/181 571/191
 asn pro gly asp cys pro pro leu glu leu ile asn thr val ile gln asp gly asp met
 601/201 631/211
 val asp thr gly phe gly ala met asp phe thr thr leu gln ala asn lys ser glu val
 661/221 691/231
 pro leu asp ile cys thr ser ile cys lys tyr pro asp tyr ile lys met val ser glu
 721/241 751/251
 pro tyr gly asp ser leu phe phe tyr leu arg arg glu gln met phe val arg his leu
 781/261 811/271
 phe asn arg ala gly ala val gly glu asn val pro asp asp leu tyr ile lys gly ser
 841/281 871/291
 gly ser thr ala asn leu ala ser ser asn tyr phe pro thr pro ser gly ser met val
 901/301 931/311
 thr ser asp ala gln ile phe asn lys pro tyr trp leu gln arg ala gln gly his asn
 961/321 991/331
 asn gly ile cys trp gly asn gln leu phe val thr val val asp thr thr arg ser thr
 1021/341 1051/351
 asn met ser leu cys ala ala ile ser thr ser glu thr thr tyr lys asn thr asn phe
 1081/361 1111/371
 lys glu tyr leu arg his gly glu glu tyr asp leu gln phe ile phe gln leu cys lys
 1141/381 1171/391
 ile thr leu thr ala asp val met thr tyr ile his ser met asn ser thr ile leu glu
 1201/401 1231/411
 asp trp asn phe gly leu gln pro pro pro gly gly thr leu glu asp thr tyr arg phe
 1261/421 1291/431
 val thr ser gln ala ile ala cys gln lys his thr pro pro ala pro lys glu asp pro
 1321/441 1351/451
 leu lys lys tyr thr phe trp glu val asn leu lys glu lys phe ser ala asp leu asp
 1381/461 1411/471
 gln phe pro leu gly arg lys phe leu leu gln ala gly leu lys ala lys pro lys phe
 1441/481 1471/491
 thr leu gly lys arg lys ala thr pro thr thr ser ser thr ser thr thr ala lys arg
 1501/501
 lys lys arg lys leu OPA

SEQ ID NO:98

1 atgtgcctgt atacacgggt cctgatatta cattaccatc tactacctct gtatggccca
 61 ttgtatcacc cagggccct gcctctacac agtatattgg tatacatggt acacattatt

```

121 atttgtggcc attatattat ttatttccta agaaacgtaa acgtgttccc tttttttttg
181 cagatggcctt tgtggcggcc tagtgacaat accgtatatc ttccacctcc ttctgtggca
241 agagttgttaa ataccgatga ttatgtgact ccacacagca ttttttatca tgctggcagc
301 tctagattat taactgtttg taatccatat tttagggttc ctgcagggtg tggcaataag
361 caggatattc ctaaggtttc tgcataccaa tatagagtat ttaggggtgca gttacctgac
421 ccaataaaat ttggtttacc tgatactagt atttataatc ctgaaacaca acgttttagtg
481 tgggcctgtg ctggagtggg aattggccgt ggtcagcctt taggtgttgg ccttagtggg
541 catccatttt ataataaatt agatgacact gaaagttccc atgccgccac gtctaattgt
601 tctgaggacg ttagggacaa tgtgtctgta gattataagc agacacagtt atgtattttg
661 ggctgtgccc ctgctatttg ggaacactgg gctaaaggca ctgctttaa atcgctcct
721 ttttggaaata ggcaggttac ccctttagaa cttaaaaaaca cagtttttga agatggtgat
781 atggttagata ctggatatgg tgcctaggac tttagtacat tgcaagatac taaatgtgag
841 gtaccatttg atattttgtca gtctatttgg aaatatcctg attattttaca aatgtctgca
901 gatccttatg gggattccat gtttttttgc ttacggcgtg agcagctttt tgctaggcat
961 ttttggaaata ggcaggttac tatgggtgac actgtgcctc aatccttata tattaaggc
1021 acaggtatgc ctgcttcacc tggcagctgt gtgtattctc cctctccaag tggctctatt
1081 gttacctctg actccagtt gttaataaaa ccatattgggt tacataaggc acagggcat
1141 aacaatggtg tttgctggca taatcaatta tttgttactg tggtagatac cactccagc
1201 accaatttaa caatatgtgc ttctacacag tctcctgtac ctgggcaata tgatgctacc
1261 aaatttaagc agtatagcag acatgttgag gaatatgatt tgcagtttat ttttcagttg
1321 tgtactatta ctttaactgc agatgttatg tcctatatcc atagtatgaa tagcagtatt
1381 ttagaggatt ggaactttgg tgtccccccc ccccaacta ctagtttggg ggatacatat
1441 cgttttgtac aatctgttgc tattacctgt caaaaggatg ctgcaccggc tgaaaataag
1501 gatccctatg ataagttaaa gttttggaat gtggatttaa aggaaaagtt ttctttagac
1561 ttagatcaat atcccttgg acgtaaaatt ttggttcagg ctggattgag tcgcaagccc
1621 accataggcc ctgcgaaacg ttctgtctcca tctgccacta cgtcttctaa acctgccaag
1681 cgtgtgctg tacgtgccag gaagtaa

```

SEQ ID NO:99

```

1 mcllytrvlil hyhlplygp lyhprplplh silvymvhii icghyiilfl rnvvnfpifl
61 qmalwrpsdn tvylpppsva rvvntddyvt ptsifyhags srlltvgnpy frvpagggkn
121 qdipkvsayg yrvfrvqlpd pnkfglpdts iynpetqrlv wacagveigr gqplgvglsg
181 hpfyknlddt esshaatsnv sedvrnsvs dykqtqlcil gcapaigehw akgtacksrp
241 lsqgdcpple lknvledgd mvdtygygamd fstlqdtkec vpldicqsic kypdylqmsa
301 dpygdsfmfc lrreqlfarh fwnragtmgd tvpqsllykg tgmpaspgsc vyspspsgsi
361 vtsdsqflnk pywlhkaqgh nngvcwhnql fvtvvdttps tnlcticastg spvpggydat
421 kfkqysrhve eydlqfifql ctitltadvn syihsmnssi ledwnfgvpp ppttslvdyt
481 rfvsqvaitc qkdaapaenk dpydklkfwn vdlkekfsld ldqyplgrkf lvqaglrkrp
541 tigprkrsap sattsskpak rrvrvrark

```

SEQ ID NO:100

```

1 atgtcttctg gcctaaacga cgtaaacgtg tccactatct ctttgcagat ggctttgtgg
61 cgccctaagt aaagcaaggt atacctacct ccaacacctg tttcaaagggt gatcagtagc
121 gatgtctatg tcacgcggac taatgtgtat taccatggtg gcagttctag gcttctcact
181 gtgggtcatc catattactc tataaagaag agtaataata aggtggcgtg gcccgaaggta
241 tctgggtacc aatatcgtgt atttcacgtg aagttgccag atccaaataa gtttggcctg
301 cccgatgctg atttgtatga tccagatacc cagagacttc tgtgggcgtg cgtgggagta
361 gaggtgggccc gtgggcagcc tttgggtgtg ggtgtgtctg gtcaccata ttacaataga
421 ctggatgaca ctgaaaatgc acacacacct gatacagctg atgatggcag ggaaaacatt
481 tctatggatt ataacacagc acagctgttc attctgggct gcaaaccccc tattgtgag
541 cactgggtcta agggtagcac ctgtaatggg tcttctgctg ctggtgactg cccgcccctc
601 caatttacta acacaactat tgaggacggg gatattggtt aaacagggtt cgggtgcctg
661 gatdttgcca cttctgcagt aaataagtca gatgttcctt tggatatattg tacciaatcc
721 tgtaaatatc ctgattatct gaagatggct gcagagcctt atggtgattc tatgttcttc
781 tcgctgcgta ggaacaaat gttcactcgt cattttttca atctgggtg taagatgggt
841 gacaccatcc cggatgagtt atacattaaa agtacctcag ttccaactcc aggcagtcac
901 gtttatactt cctactctag tggctctatg gtgtcctctg aacaacagtt gtttaataag
961 ccttactggc tacggagggc ccaagggcac acaaatggtg tgtgtcggg caatagggtc
1021 tttctgactg tggtagcac cacacgtagc actaatgtat ctctgtgtgc cactgaggcg
1081 tctgatacta attataaagg taccattttt aaggaatata tcaggcatat ggaggaatat
1141 gatttgcagt tcactttcca actgtgcaag ataaccctta ctctgaaat tatggcctat
1201 atacataata tggatcccca gttgttagag gattggaact tgggtgtacc cctccgccc
1261 tctgccagtt tacaggatac ctatagatat ttgcagtcctc aggcatttac atgtcaaaaa
1321 cctacacctc ctaagacccc taccgatccc tatgcctccc tgaccttttg ggatgtggat
1381 ctacgtgaaa gtttttccat ggatctggac caatttccct tgggtcgcaa gtttttctg
1441 cagcgggggg ctatgcctac cgtgtctcgc aagcgcgcgc ctgtttcggg gaccacgccc
1501 cccactagta aacgaaaaac ggtaaggcgt tag

```

SEQ ID NO:101

```

1 mscglndvnn stislqmalw rpnescvylp ptpvskvist dvvytrtnvy yhggsrllt
61 vghpyysikk snnkvaupkv sgyqrvfhw klpdnpkfgl pdadlydpdt qrlwacvgv
121 evgrgqplgv gvsghpyynr lddtenahp dtaddgreni smdykqtqlf ilgckppige
181 hwskgttcng ssaagdcppl qftnttiedg dmvetgfgal dfatlqsnks dvpldictnt
241 ckpydylkma aeptygdsfff slrreqmfr hffnlggkmg dtipdelyik stsvptpgsh
301 vytstpsgsm vsseqqlfkn pywlrraqgh nngmcwgnrv fltvvdttts tnvsllcatea
361 sdtnykatnf keylrhmeey dlqfifqlck itltpeimay ihnmddpille dwnfgvpppp
421 saslqdytyr lqsqaitcck ptppktptdp yasltfwdvd lsesfsmld qfplgrkfil
481 qrampptvsr kraavsgtpt ptkrkrvr

```

SEQ ID NO:102

1/1 31/11
 ATG AGG CAC AAG AGG AGC GCC AAG AGG ACC AAG AGG GCC AGC GCC ACC CAG CTG TAC AAG
 61/21 91/31
 ACC TGC AAG CAG GCC GGC ACC TGC CCC CCC GAC ATC ATC CCC AAG GTG GAG GGC AAG ACC
 21/41 151/51
 ATC GCC GAC CAG ATC CTG CAG TAC GGC AGC ATG GGC GTG TTC TTC GGC GGC CTG GGC ATC
 181/61 211/71
 GGC ACC GGC AGC GGC ACC GGC GGC AGG ACC GGC TAC ATC CCC CTG GGC ACC AGG CCC CCC
 241/81 271/91
 ACC GCC ACC GAC ACC CTG GCC CCC GTG AGG CCC CCC CTG ACC GTG GAC CCC GTG GGC CCC
 301/101 331/111
 AGC GAC CCC AGC ATC GTG AGC CTG GTG GAG GAG ACC AGC TTC ATC GAC GCC GGC GCC CCC
 361/121 391/131
 ACC AGC GTG CCC AGC ATC CCC CCC GAC GTG AGC GGC TTC AGC ATC ACC ACC AGC ACC GAC
 21/141 451/151
 ACC ACC CCC GCC ATC CTG GAC ATC AAC AAC ACC GTG ACC ACC GTG ACC ACC CAC AAC AAC
 81/161 511/171
 CCC ACC TTC ACC GAC CCC AGC GTG CTG CAG CCC CCC ACC CCC GCC GAG ACC GGC GGC CAC
 541/181 571/191
 TTC ACC CTG AGC AGC AGC ACC ATC AGC ACC CAC AAC TAC GAG GAG ATC CCC ATG GAC ACC
 601/201 631/211
 TTC ATC GTG AGC ACC AAC CCC AAC ACC GTG ACC AGC AGC ACC ATC CCC GGC AGC AGG
 661/221 691/231
 CCC GTG GCC AGG CTG GGC CTG TAC AGC AGG ACC ACC CAG CAG GTG AAG GTG GTG GAC CCC
 721/241 751/251
 GCC TTC GTG ACC ACC CCC ACC AAG CTG ATC ACC TAC GAC AAC CCC GCC TAC GAG GGC ATC
 781/261 811/271
 GAC GTG GAC AAC ACC CTG TAC TTC AGC AGC AAC GAC AAC AGC ATC AAC ATC GCC CCC GAC
 841/281 871/291
 CCC GAC TTC CTG GAC ATC GTG GCC CTG CAC AGG CCC GCC CTG ACC AGC AGG AGG ACC GGC
 901/301 931/311
 ATC AGG TAC AGC AGG ATC GGC AAC AAG CAG ACC CTG AGG ACC AGG AGC GGC AAG AGC ATC
 961/321 991/331
 GGC GCC AAG GTG CAC TAC TAC TAC GAC CTG AGC ACC ATC GAC CCC GCC GAG GAG ATC GAG
 1021/341 1051/351
 CTG CAG ACC ATC ACC CCC AGC ACC TAC ACC ACC ACC AGC CAC GCC GCC AGC CCC ACC AGC
 081/361 1111/371
 ATC AAC AAC GGC CTG TAC GAC ATC TAC GCC GAC GAC TTC ATC ACC GAC ACC AGC ACC ACC
 1141/381 1171/391
 CCC GTG CCC AGC GTG CCC AGC ACC AGC CTG AGC GGC TAC ATC CCC GCC AAC ACC ACC ATC
 1201/401 1231/411
 CCC TTC GGT GGC GCC TAC AAC ATC CCC CTG GTG AGC GGC CCC GAC ATC CCC ATC AAC ATC
 1261/421 1291/431
 ACC GAC CAG GCC CCC AGC CTG ATC CCC ATC GTG CCC GGC AGC CCC CAG TAC ACC ATC ATC
 1321/441 1351/451
 GCC GAC GCC GGC GAC TTC TAC CTG CAC CCC AGC TAC TAC ATG CTG AGG AAG AGG AGG AAG
 1381/461 1411/471
 AGG CTG CCC TAC TTC TTC AGC GAC GTG AGC CTG GCC GCC TGA

SEQ ID NO:103

1/1 31/11
 Met arg his lys arg ser ala lys arg thr lys arg ala ser ala thr gln leu tyr lys
 61/21 91/31
 thr cys lys gln ala gly thr cys pro pro asp ile ile pro lys val glu gly lys thr
 121/41 151/51
 ile ala asp gln ile leu gln tyr gly ser met gly val phe phe gly gly leu gly ile
 181/61 211/71
 gly thr gly ser gly thr gly gly arg thr gly tyr ile pro leu gly thr arg pro pro
 241/81 271/91
 thr ala thr asp thr leu ala pro val arg pro pro leu thr val asp pro val gly pro
 301/101 331/111
 ser asp pro ser ile val ser leu val glu glu thr ser phe ile asp ala gly ala pro
 361/121 391/131
 thr ser val pro ser ile pro pro asp val ser gly phe ser ile thr thr ser thr asp
 421/141 451/151
 thr thr pro ala ile leu asp ile asn asn thr val thr thr val thr thr his asn asn
 481/161 511/171
 pro thr phe thr asp pro ser val leu gln pro pro thr pro ala glu thr gly gly his
 541/181 571/191
 phe thr leu ser ser ser thr ile ser thr his asn tyr glu glu ile pro met asp thr
 601/201 631/211
 phe ile val ser thr asn pro asn thr val thr ser ser thr pro ile pro gly ser arg
 661/221 691/231
 pro val ala arg leu gly leu tyr ser arg thr thr gln gln val lys val val asp pro
 721/241 751/251
 ala phe val thr thr pro thr lys leu ile thr tyr asp asn pro ala tyr glu gly ile
 781/261 811/271
 asp val asp asn thr leu tyr phe ser ser asn asp asn ser ile asn ile ala pro asp
 841/281 871/291
 pro asp phe leu asp ile val ala leu his arg pro ala leu thr ser arg arg thr gly

```

901/301      931/311
ile arg tyr ser arg ile gly asn lys gln thr leu arg thr arg ser gly lys ser ile
961/321      991/331
gly ala lys val his tyr tyr tyr asp leu ser thr ile asp pro ala glu glu ile glu
1021/341     1051/351
leu gln thr ile thr pro ser thr tyr thr thr thr ser his ala ala ser pro thr ser
1081/361     1111/371
ile asn asn gly leu tyr asp ile tyr ala asp asp phe ile thr asp thr ser thr thr
1141/381     1171/391
pro val pro ser val pro ser thr ser leu ser gly tyr ile pro ala asn thr thr ile
1201/401     1231/411
pro phe gly gly ala tyr asn ile pro leu val ser gly pro asp ile pro ile asn ile
1261/421     1291/431
thr asp gln ala pro ser leu ile pro ile val pro gly ser pro gln tyr thr ile ile
1321/441     1351/451
ala asp ala gly asp phe tyr leu his pro ser tyr tyr met leu arg lys arg arg lys
1381/461     1411/471
arg leu pro tyr phe phe ser asp val ser leu ala ala OPA

```

SEQ ID NO:104

```

1 atggtatccc accgtgccgc acgacgcaaa cgggcttcgg taactgactt atataaaaca
61 tgtaaaacaat ctggtacatg tccacctgat gttgttccta aggtggaggg caccacgtta
121 gcagataaaa tattgcaatg gtcaagcctt ggtatatattt tgggtggact tggcataggt
181 actggcagtg gtacaggggg tcgtacaggg tacattccat tgggtgggcg ttccaatata
241 ctggtggatg ttggtctac acgtcccca gtggttattg aaactgtggg cccacagac
301 ccattctattg ttacattaat agaggactcc agtgtggtta catcagggtgc acctaggcct
361 acgtttactg gcacgtcttg gtttgatata acatctgcgg gtacaactac acctgcgggt
421 ttggatatca caccttcgtc tacctctgtg tctatttcca caaccaattt taccaatcct
481 tgattttctg atccgtccat tattgaagtt ccacaaactg gggaggtggc aggtaatgta
541 tttgttggtt cccctacatc tggacacatc gggtagtgagg aaataccttt acaaacattt
601 gcttcttctg gtacggggga ggaacccatt agtagtacc cattgcctac tgtgcggcgt
661 gtacaggtgc cccgccttta cagtagggcc taccaacaag tgcagtggc taaccctgag
721 tttcttacac gtccatctc ttaattaca tatgacaacc cggccttga gcctgtggac
781 actacattaa catttgatcc tcgtagtgt gttcctgatt cagattttat ggatattatc
841 cgtctacata ggctgcttt aacatccagg cgtgggactg ttcgctttag tagattaggt
901 caacgggcaa ctatgtttac ccgacgggt acacaaatag gtgctagggt tcacttttat
961 catgatataa gtcctattgc accttcccca gaatatattg aactgcagcc ttagtatct
1021 gccacggagg acaatgactt gtttgatata tatgcagatg acatggacc tgcagtgcct
1081 gtaccatcgc gttctactac ctcttttgca ttttttaaat attcgccac tatatcttct
1141 gcctcttct atagtaatgt aacggctcct ttaacctcct ctgggatgt gcctgtatac
1201 acgggtcctg atattacatt accatctact acctctgtat ggccattgt atcaccacg
1261 gccctgcct ctacacagta tattgtata catggtacac attattattt gtggccatta
1321 tattatttta ttctaagaa acgtaaacgt gtccctatt tttttgcaga tggctttgtg
1381 gcgcctag

```

SEQ ID NO:105

```

1 mvshraarrk rasvtdlykt ckqsgtceppd vvpkveggtl adkilqwssl giflglgig
61 tsgsgtgrtg yiplggrsnt vdvvgptrpp vvienvgptd psivtlieds svvtsgarpr
121 tftgtsgfdi tsagtttppav lditpsstsv sistnftnp afsdpsieiev pqtgevagnv
181 fvgtpstgth gyeeiplqtf assgtgeepi sstplptvrr vagprlysr yqqvsvanpe
241 fltrpsslit ydnparfepvd tltfdprsd vpsdmdfmdii rlhrpaltsr rgtvrfsrlg
301 qratmfrsrq tqigarvhfy hdispiaspe eyielqplvs atedndlfdi yaddmdpavp
361 vrsrsttsfa ffkysptiss assysnvtvp ltsswdvpyv tgpditlpst tsvwpivspt
421 apastqyigi hgthyyllwpl yyfipkkrkr vpyffadgfv aa

```

SEQ ID NO:106

```

1 atgtctgttg gtgattctta tccataatcgc ctttttattg ttgatgtttt atgtccgttt
61 gttaaaccac acctaacacc cccatttttt tatattgttt tgatacattt tcattttgat
121 acatttgtgt tttttttgta tttgtgcgt ttaataaac gtgcaacct gtctatacgt
181 gccaaagcgtc gaaagcgcgc ctccccaca gacctctatc gtacctgcaa gcaggcaggt
241 acctgcccc cagacattat ccaagagtg gaacagaaca ctttagcaga taaaatcctt
301 aagtggggca gtttaggtgt gttttttggg ggtctaggtta taggcaccg cagcgcgaca
361 ggggggcgta ctgggtacat tctgttaggt tcgcgacca cactgtagt tgacattggt
421 ccaacgcccc ggccgctgt tatcattgaa cctgtgggg cctctgaacc ctctattgtc
481 actttggttg aggactctag catcattaac gcaggagcgt cacatccac ctttactggt
541 actggtggct tcgaagtgc aacctccacc gttacagacc ccgcgcctt ggatatacacc
601 ccctcaggta ccagtgtgca ggtcagcagc agtagcttct ttaaccact atacactgag
661 ccagctattg tggaggtccc ccaaacagg gaagtatctg gccatgtact tgttagtaca
721 gccacctcag ggtctcatg ctatgaggaa ataccaatgc agacgtttgc cagctcggg
781 ggcagcggta cacagcctat cagtagcaca cccctccctg cgtgcggag agttgccgga
841 cccgcctgt acagttaggc caatcagcaa gtgcaagtca gggactctgc gtttcttgca
901 aggcctgctg atctagtaac atttgacaat cctgtgtatg acccagagga aactataata
961 tttcagcatc cagacttgca tgagccaccg gatcctgatt ttttgacat agtggcgtt
1021 catcgtcccg cctccacgtc cagaagggtt actgtccgtt ttagtaggtt gggacgcagg
1081 gctacactcc gcaccctgag tggtaaaca attggggcac ggggtgcact ctatcatgat
1141 attagcccta taggtactga ggagttggag atggagccac tgttgcccc agcttctact
1201 gataacacag atatgttata tgatgtttat gctgattcgg atgtccttca gccattgctt
1261 gatgagttac ccgcgccttc tcgcggttca ctctctctg ctgacactgc tgtgtctgcc
1321 acctccgcat ctacactacg ggggtccact actgtccctt tatcaagtgg tattgatgtg
1381 cctgtgtaca ccggtcctga cattgaacca ccaatgttc ctggcatggg acctctgatt

```

1441 cctgtggctc catccttacc atcgtctgtg tacatatattg ggggagatta ttatttgatg
1501 ccaagttatg tcttgtggcc taaacgacgt aaacgtgtcc actatttctt tgcagatggc
1561 tttgtggcgg cctaa

SEQ ID NO:107

1 msvgdsypnr lfivdvclpf vkphltpplf yivlihfhd tfvfflyllr fnkratmsir
61 akrrkraspt dlyrtckqag tcppdiiprv eqntladkil kwgslgvffg glgigtgsqt
121 ggrrtyipvg srpttvvdiq ptprrpviie pvgasepsiv tlvedssiin agashptftg
181 tggfevttst vtdpavldit psgtsvqvss ssflnplyte paiveapqtg evsghvlvst
241 atsgshgyee ipmgtfatsg gsgtepisst plpgvrrvag prlysrnqq vqvrpafala
301 rpadiivtfdn pvydpeetii fqhpdllhepp dpdfldival hrpaltstrg tvrfsrlgrr
361 atlrtrsgkq igarvhfyhd ispigteele mepllpast dntdmlydv adsdvlpqll
421 delpaaprgs lsladtavsa tsastlrgst tvplssgidv pvytgpdiep pnvpgmgpli
481 pvapslpssv yifggdyilm psyvlwpkrr krvhffadg fva

SEQ ID NO:108

1 atggagctga ggccttggtt gctatgggtg gtagcagcaa caggaacctt ggtcctgcta
61 gcagctgatg ctcaggccca gaaggtcttc accaacacgt gggctgtgcg catccctgga
121 ggccagcggg tggccaacag tgtggcacgg aagcatgggt tcctcaacct gggccagatc
181 ttccgggact attaccactt ctggcatcga ggagtgcga agcggtccct gtcgcctcac
241 cgcccgcgcc acagccggct gcagagggag cctcaagtac agtggctgga acagcaggtg
301 gcaaagcgac ggactaaacg ggacgtgtac caggagccca cagaccccaa gtttcctcag
361 cagtgtgacc tgtctggtgt cactcagcgg gacctgaatg tgaaggcggc ctgggcgcag
421 ggctacacag ggcacggcat tgtggtctcc attctggacg atggcatcga gaagaaccac
481 ccggacttgg caggcaatta tgatcctggg gccagttttg atgtcaatga ccaggacctt
541 gacccccagc ctccgtacac acagatgaat gacaacaggg acggcacacg gtgtgcgggg
601 gaagtggctg cgggtggccaa caacgggtgtc tgtggtgtag gtgtggccta caacgccgc
661 attggagggg tgcgcagtgt ggtggcgag gtgacagatg cagtggaggc acgctcgctg
721 ggcctgaacc ccaaccacat ccacatctac agtgccagct ggggccccga ggatgacggc
781 aagacagtgg atgggccagc ccgctcgcc gaggaggcct tcttcctggt ggtagccag
841 ggccgagggg ggtcgggctc catctttgtc tgggcctcgg ggaacggggg ccgggaacat
901 gacagctgca actcgcagcg ctacaccaa agtatctaca cgctgtccat cagcagcgcc
961 acgcagtttg gcaacgtgcc gtggtacagc gaggcctgct cgtccacact ggccacgacc
1021 tacagcagtg gcaaccagaa tgagaagcag atcgtgacga ctgacttgcg gcagaagtgc
1081 acggagtctc acacgggcac ctcagcctct ccccccttag cagccggcat cattgctctc
1141 accctggagg ccaataagaa cctcacatgg cgggacatgc aacacctggt ggtacagacc
1201 tcgaagccag ccacacctca tgccaacgac tgggccacca atggtgtggg ccggaagtgt
1261 agccactcat atggctacgg gcttttgagc gcaggcccca tgggtggcct ggccccagaat
1321 tggaccacag tggccccca cgggaagtgc atcatcgaca tcctcacga gcccaagac
1381 atcgggaaac ggctcagagt gcggaagacc gtgaccgctg gcctggcgga gcccaaccac
1441 atcactcggc tggagcacgc tcaggcgagg ctcaccctgt cctataatcg ccgtggcgac
1501 ctggccatcc acctggtcag ccccatgggc acccgctcca cctgctggc agccaggcca
1561 catgactact ccgcagatgg gtttaatgac tgggccttca tgacaactca ttcctgggat
1621 gaggatccct ctggcgatg ggtcctagag attgaaaaca ccagcgaagc caacaactat
1681 gggacgctga ccaagttcac cctcgctact tatggcaccg cccctgaggg gctgccccta
1741 cctccagaaa caggtggctg caagaccctc acgtccagtc aggcctgtgt ggtgtgcgag
1801 gaaggcttct cctgcacca gaagagctgt gtccagcact gccctccagg gttcggcccc
1861 caagtcctcg atacgcacta tagcaccgag aatgacgtgg agaccatccg ggccagcgtc
1921 tgcgccccct gccacgctc atgtgccaca tgccaggggc cgccctgac agactgcctc
1981 agctgccccca gccacgctc cttggaccct gtggagcaga cttgctccc gcaaagccag
2041 agcagccgag agtccccgcc acagcagcag ccacctcggc tgcccccgga ggtggaggcg
2101 gggcaacggc tgccggcagg gctgctgccc tcacacctgc ctgaggtggt ggccggcctc
2161 agctgcgcct tcactgtgct ggtcttcgtc actgtcttcc tggctcctga cctgctctc
2221 ggctttagtt ttcggggggg gaaggtgtac accatggacc gtggcctcat ctctacaag
2281 gggctgcccc ctgaagcctg gcaggaggag tgccctctg actcagaaga ggacgaggcg
2341 cggggcgaga ggaccgcctt tatcaaagac cagagcgccc tctga

SEQ ID NO:109

1 melrpwllwv vaatgtlvll aadaqqkvf tntwavrip g pavansvar khgflnlqgi
61 fgdyhfwlr gvtkrslsph rprhrlqre pqvwleqqv akrtrkrdvy qeptdpkfpq
121 qwylsgvtqr dlnvkaawaq gytghgivvs ilddgieknh pdlagnydpq asfdvndqdp
181 dpqprytqmn dnrhgtrcag evaavanngv cgvgvaynar iggvrmldge vtdavearsl
241 glnpnhihiy saswgpddg ktdvgparla eaffrgrvsq grgglsifv wasnggreh
301 dscncdgytn siytlssisa tqfgnvpwys eacsstlatt yssgnqnekq ivttldlrqk
361 teshtgtsas aplaagiial tleanknlw rdmqhlvvt skpahlndand watngvgrkv
421 shsygyglll agamvalagn wttvapqrkc iidiltepkd igkrlevrkt vtaclegpnh
481 itrlehaqar ltlsynrrgd laihlvspmg trstllaarp hdsadgfnf wafmtthswd
541 edpsgewvle ientseanny gtltkftlvi ygtapeglpv ppessgcktl tssqacvvce
601 egfslhqksc vqhcppgfap qvldthyste ndvetirasv capchascac cggpaltdcl
661 scpshasldp veqtcsrqsq ssresppqqq pprlppevea qgrrlagllp shlpevvagl
721 scafivlvfv tvflvlqlrs gfsfrgvkvy tmdrglisyk glppeawqee cpsdseedeg
781 rgertafikd qsal

SEQ ID NO:110

AATGGACCAGTTCTAATGT

SEQ ID NO:111

GTCAGCCCTAAATCTTC

SEQ ID NO:112

- 148 -

CCGCCAACCAATACCGCGTGTTCAGATCCAGCTGCCCGACCCAAACAGTTTCGCCCTGCCCGATCGCACCGTGCATAACCCCTCCAAGGAA
AGACTCGTCTGGGGCGTGATCGGGCTCCAGTCTCACGGGGCCAACCCCTGGGCGGCACCGTGACCGGCCATCCAACCTTCAACGCCCTCCT
GGAGCGCGAGAAGCTCAACCGGAAAGTCAACAACAAACCCAGCAGATCGCAAGCAGACCGGGCTGGACGCCAAACAGCAGCAAAATCCTCC
TCCTGGGGTGACACCCCGTGAGGGCGAGTACTGGACCACCGCTCGGCCCTGCGTGACCGACAGGCTGGAGAACGGGGCTTGTCCTCCCTCCCTG
GAGCTGAAGAATAAGCATATCGAGGACGGCGACATGATGGAGATCGGCTTCGGCGCCGCTAACTTCAAGGAGATCAACGCCCTCCAAGAGCGA
CCTGCCCTTGGATATCCAGAACGAAATTTGTCTCTATCCCGATTATCTGAAGATGGCCGAAGATGCCGCCGGCAACTCAATGTTTTTCTTCG
CCCGCAAGGAGCAAGTCTACGTGCGGCATATTTGGACACGGGGCGGGAGCGAAAAGGAGGCTCCCAACCCGACTTCTACCTGAAAAACAAC
AAGGGCGACGCTACACTGAAGATCCCATCCGTCCACTTCGGCTCCCATCCGGGAGCCTCGTCAGCACCGACAACAGATCTTCAACAGACC
ATATTGGCTGTTTTAGGGCTCAAGGGATGAATAACGGCATCGCTTGAACAACCTGCTCTTCTGACCGTCGGCGATACACAGGGGCACCA
ACCTGACAACTCTCCGTGGCTAGCGACGGCACACCCCTGACCGAATACGACTCAAGCAAGTTTAACTGTATCACCGGCACATGGAGGAGTAC
AAACTGGCTTTTCATCTTGAAGTGTGTAGCGTCGAGATTACCGCCAGACCGTCAGCCACCTCCAGGGCTGATGCCAAGCGTCTTGGAGAA
CTGGGAGATCGGCGTCCAACCAACAAGCAGCATCTTGAAGATACATACAGATACATCGAAAGCCCCGCCACCAAGTGCGCCCTCAAACG
TGATCCCCGCCAAGGAGTCCCTACGCCGGCTTCAAATCTTGGAAATATCGACCTGAAGGAGAACTGAGGCTCGATCTGGACAGTTCCCA
CTCGGCCGGCGTTCTTGCCCAACAGGGCGCTGGCTGCAGCACCGTCCGGAAGAGGGGATCTCACAAGAGACAGTTCCAAACCCGCCAA
GAAGAAGAAGTAG

SEQ ID NO:123

BPV-L1 AA:

MALWQQGQKLYLPPTPVSKVLCSETYVQRKSIIFYHAETERLLTIGHPPYVPSIGAKTVPKVSNQYRVFKIQLPDPNQFA
LPDRTVHNPSKERLVWAVIGVQVSRGQPLGGTVTGHPTFNALLDAENVNRKVTQTDDRKQTLGLDAKQQQILLGLCTPA
EGEYWTARPCVTDRLENGACPPLELKNKHIEDGDMMEIGFGAANFKEINASKSDLPLDIQNEICLYPDYLMKMAEDAAGN
SMFFEFARKEQVYVRHIWTRGGSEKEAPTDFYLNKNGDATLKI PSVHFGSPSGSLVSTDNQIFNRPYWLFRAGQGMNNGI
AWNLLFLTVGDNRGTNLTI SVASDGTPLTEYDSSKFNVYHRHMEEYKLAFILELCSVEITAQT VSHLQGLMPSVLENW
EIGVQPPTSSILEDTRYRIESPATKCSNVI PAKEDPYAGFKFNWIDLKEKLSLDLQFPLGRRFLAQQAGCSTVRKR
ISQKTSSKPAKKKK

SEQ ID NO:124

BPV-L2 DNA:

ATGAGCGCCCGCAAGAGAGTGAAGCGCGCCAGCGCTACGACCTGTACAGGACCTGCAAGCAGGCCGCGACATGTCCACCAGATGTGATCCC
CAAGGTGGAGGGCGACACCATCGCCGACAAGATCTGAAGTTTCGGCGGCCCTGGCCATCTACCTGGGCGGCCCTGGGCATCGGAACATGGTCTA
CCGGCAGGGTGGCCGCGCGCGGCTCACCAAGGTACACCCCACTGCGCACCGCGGCTCCACCTCCTCCCTGGCCTCCATCGGATCCAGAGCC
GTGACCGCCGGGACCCGCCCTCCATCGGCGCGGGCATCCCTCTGGACACCCCTGGAACTCTTGGGGCCCTGCGCCCTGGCGTGTACGAGGA
CACCGTGCTGCCGAAGCCCTGCCATCGTGACCCCTGACGCGCTGCCTGCAGACTCCGGCCTGGACGCGCTGTCCATCGGCACAGACTCCT
CCACCGAGACCCTGATCACCTGCTGGAGCCTGAGGGCCCCGAAGACATAGCCGTGCTGGAACCTCCAGCCCTGGACCGCCCAACCTGGCAG
GTGAGCAATGCTGTGCACCACTCCTCTGCCTACCACGCCCTCTCCAGCTGCAATCCTCCATCGCCGAGACATCTGGTTTAGAAAAATATTTT
TGTAGGAGGCTCGGGTTTAGGGGATACCGCGCGCGAGAACATCGAGCTGACCTACTTCGGCTCCCCCGCACAGCACCCCCCGCTCCATCG
CCTCCAAGTCCCGCGGCATCCTGAACTGGTTTCAGCAAGCGGTACTACACCCAGGTGCCACCGAAGATCCCGAAGTGTTCCTCCAGACC
TTCGCCAACCCCTGTACGAGGCCGAGCCGCGCTGCTGAAGGGCCCTAGCGGCCGCTGGGCCTGTCCCAGGTGTACAAGCCTGATACCCT
GACCACACGTAGCGGCACAGAGGTGGGCCCCAGCTGCATGTGAGGTACTCCCTGTCCACCATCCATGAGGATGTGGAGGCTATCCCTACA
CCGTGGATGAGAACACCCAGGGCCTGGCCTTCGTGCCCCGTGCATGAGGAGCAGGCCGGCTTCGAGGAGATCGAGCTCGACGATTTTCAGCGAG
ACCCATCGCTGCTGCCCCAGAACACCTCCTCCACCCCGTGGGCGAGCGCGTGCAGAGCCTGATCCCTACCCAGGAGTTTCAGCGCCAC
CCGGCCTACCGCGTGGTGACCTACGGCTCCCCCGACACCTACTCCGCTAGCCCGCTGACCGACCTGATTCTACCTCTCCTAGCCTGGTGA
TCGACGACACCAACCAACCCCATCATCATCGACGCCACACAGTGGATCTGTACAGCAGCAACTACACCTGCATCCCTCCCTGCTG
AGGAAGCGCAAGAAGCGCAAGCATGCCTAAA

SEQ ID NO:125

BPV-L2 AA:

STSSLASIGRAVTA TRPSIGAPIPLDTLETGLALRPGVYEDTVLPEAPAI VTPDAVPADSGLDALSIGTDSSTETLIT
LLEPEGPEDIAVLELQPLDRPTWQVSNVHQSSAYHAPLQLQSSIAETSGLENI FVGGSGLGDTGGENIELTYFGSPRTS
TPRSIASKSRGILNWFSKRYYTQVPTEDPEVFSSQTFANPLYEAEPAVLKGPSGRVGLSQVYKPDLTLTRSGTEVGPQLH
VRYSLSTIHEDVEAI PYTVDENTQGLAFVPLHEEQAGFEEILD DDFSETHRLLPQNTSSTPVGSGVRRSLIPTQFESATR
PTGVVITYGSPDITYASPVTDPDSTSPSLVIDDTTTPPIIIDGH TVDLYSSNYTLHPSLLRKRKRKHA

CLAIMS

1. A combination for use in inducing or enhancing an antigen-specific immune response in a mammal, comprising
 - (a) an effective amount of at least one DNA vector comprising a papillomavirus capsid antigen to be administered;
 - (b) an effective amount of at least one DNA vaccine comprising an antigenic peptide to be administered, andthe combination induces or enhances the antigen specific immune response.
2. The combination of claim 1, wherein the papillomavirus is selected from the group consisting of human papilloma virus (HPV)1, HPV2, HPV3, HPV4, HPV5, HPV6, HPV7, HPV8, HPV9, HPV10, HPV11, HPV12, HPV13, HPV14, HPV15, HPV16, HPV17, HPV18, HPV19, HPV20, HPV21, HPV22, HPV23, HPV24, HPV25, HPV26, HPV27, HPV28, HPV29, HPV30, HPV31, HPV32, HPV33, HPV34, HPV35, HPV36, HPV37, HPV38, HPV39, HPV40, HPV41, HPV42, HPV43, HPV44, HPV45, HPV46, HPV47, HPV48, HPV49, HPV50, HPV51, HPV52, HPV53, I-IPV54, HPV55, HPV56, HPV57, HPV58, HPV59, HPV60, HPV61, HPV62, HPV63, HPV64, HPV65, HPV66, HPV67, HPV68, HPV69, HPV70, HPV71, HPV72, HPV73, HPV74, HPV75, HPV76, HPV77, HPV78, HPV79, HPV80, HPV81, HPV82, HPV83, HPV84, HPV85, HPV86, HPV87, HPV88, HPV89, HPV90, HPV91, HPV92, HPV93, HPV94, HPV95, HPV96, HPV97, HPV98, HPV99, HPV100, bovine papillomavirus type 1 (BPV1), bovine papillomavirus type 2 (BPV2), bovine papillomavirus type 4 (BPV4), cottontail rabbit papillomavirus (CRPV), deer papillomavirus (DPV), European elk papillomavirus (EPPV), canine oral papillomavirus (COPV), Rhesus monkey papillomavirus (RnPV), and rabbit oral papillomavirus (ROPV).
3. The combination of claim 2, wherein the papillomavirus is bovine papillomavirus type 1 (BPV1).
4. The combination of claim 1, wherein the capsid antigen is a major and minor capsid antigen.
5. The combination of claims 4, wherein the major capsid antigen is L1.
6. The combination of claim 4, wherein the minor capsid antigen is L2.
7. The combination of claim 5, wherein the capsid antigen L1 is derived from BPV1, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 123.

8. The combination of claim 6, wherein the capsid antigen L2 is derived from BPV1, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 125.
9. The combination of claim 1, wherein the combination elicits enhanced antigen-specific CD8⁺.
10. The combination of claim 1, wherein the DNA vaccine comprises an antigenic peptide, and said antigenic peptide is selected from the group consisting of ovalbumin (OVA), HPV16 E6, HPV16 E7, and influenza NS1.
11. The combination of claim 10, wherein the DNA vaccine comprises HPV-16 E7 (E7) antigen linked with calreticulum (CRT) (CRT/E7).
12. The combination of claim 11, wherein the E7 comprises an amino acid sequence set forth in RAHYNIVTF.
13. The combination of claim 10, wherein the DNA vaccine comprises OVA.
14. The combination of claim 13, wherein the OVA comprises an amino acid sequence set forth in ISQAVHAAHAEINEAGR.
15. The combination of claim 2, wherein the papillomavirus is human papillomavirus type 16 (HPV-16).
16. The combination of claims 15, wherein the major capsid antigen is L1.
17. The combination of claim 15, wherein the minor capsid antigen is L2.
18. The combination of claim 16, wherein the capsid antigen L1 is derived from HPV16, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 119.
19. The combination of claim 17, wherein the capsid antigen L2 is derived from HPV16, and said capsid antigen comprises an amino acid sequence at least 60% identical to the amino acid sequence set forth in SEQ ID NO: 121.
20. The combination of claim 1, wherein the combination elicits enhanced antigen-specific CD4⁺ T cell immune response.
21. The combination of claim 1, wherein the enhanced antigen-specific CD4⁺ T cell immune response is L1-, L2-, or L1/L2-specific.
22. The combination of claim 1, wherein the mammal is a human.
23. The combination of claim 1, wherein the mammal is afflicted with cancer.
24. The combination of claim 1, wherein the combination is administered intradermally, intraperitoneally, or intravenously via injection.

25. The combination of claim 1, wherein the combination is administered alone or in a pharmaceutically acceptable carrier in a biologically-effective, and/or a therapeutically-effective amount,
26. The combination of claim 1, wherein the antigen-specific immune response is greater in magnitude than an antigen-specific immune response induced by administration of the DNA vaccine or DNA vector alone.

Figure 1

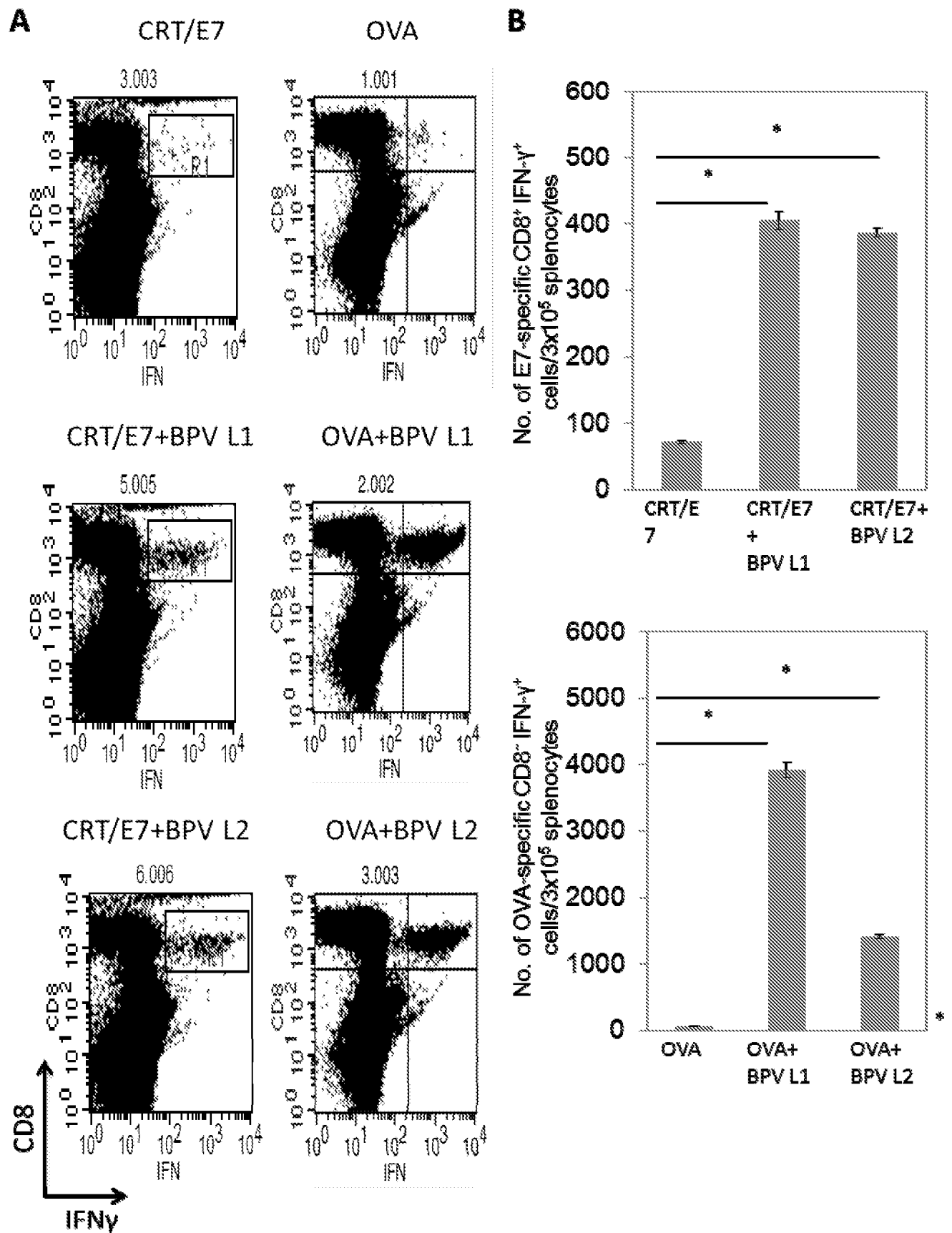


Figure 2

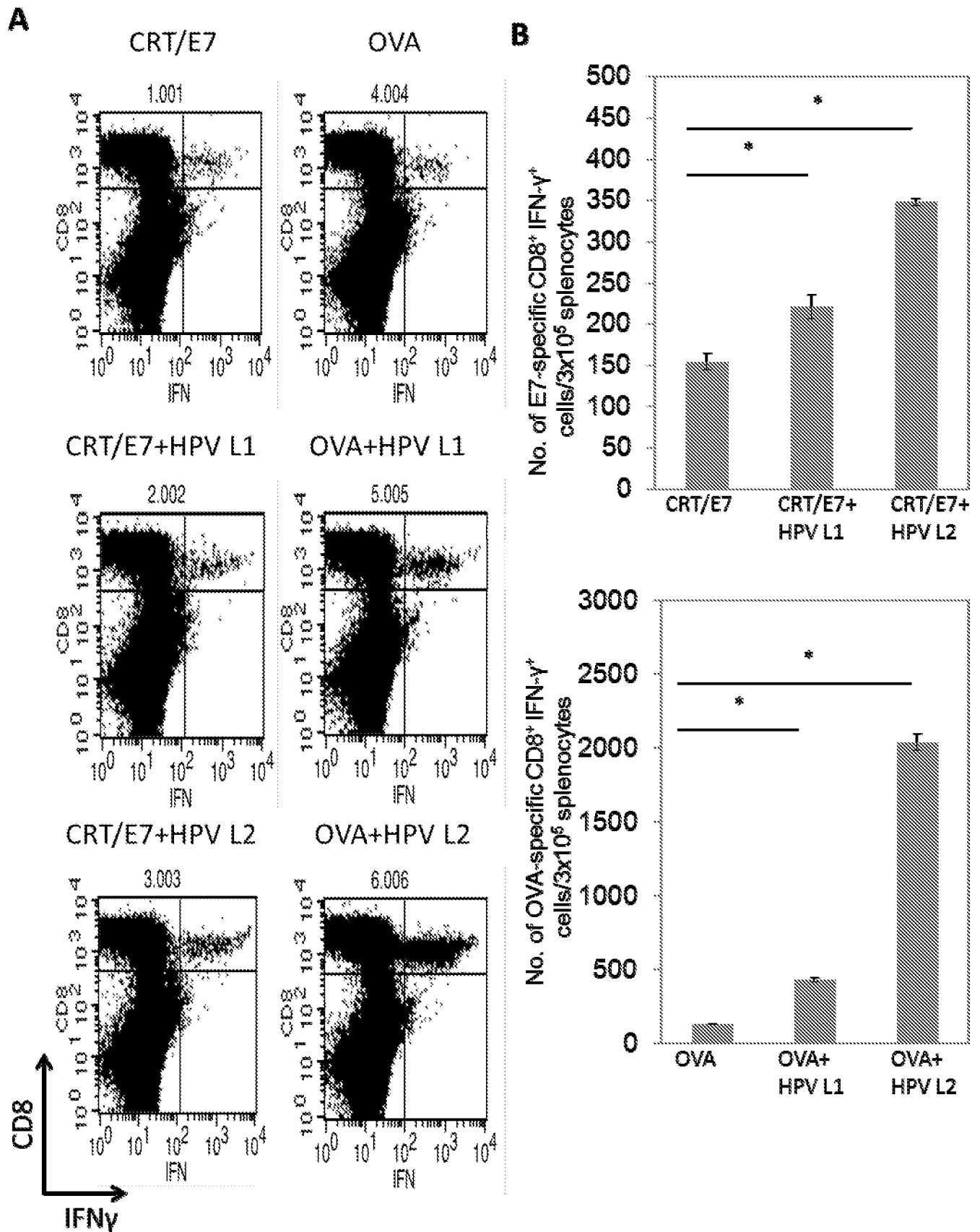


Figure 3

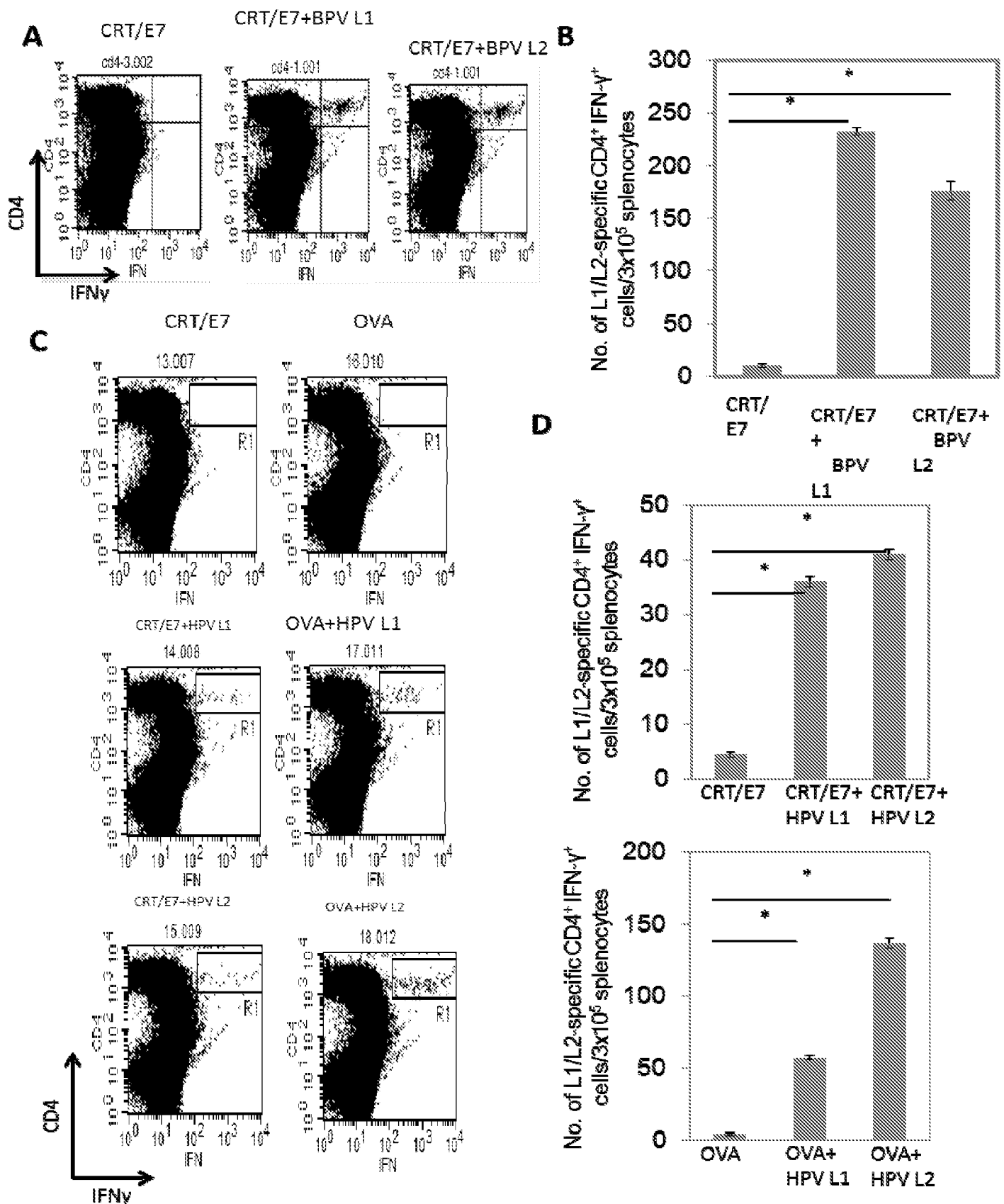


Figure 4

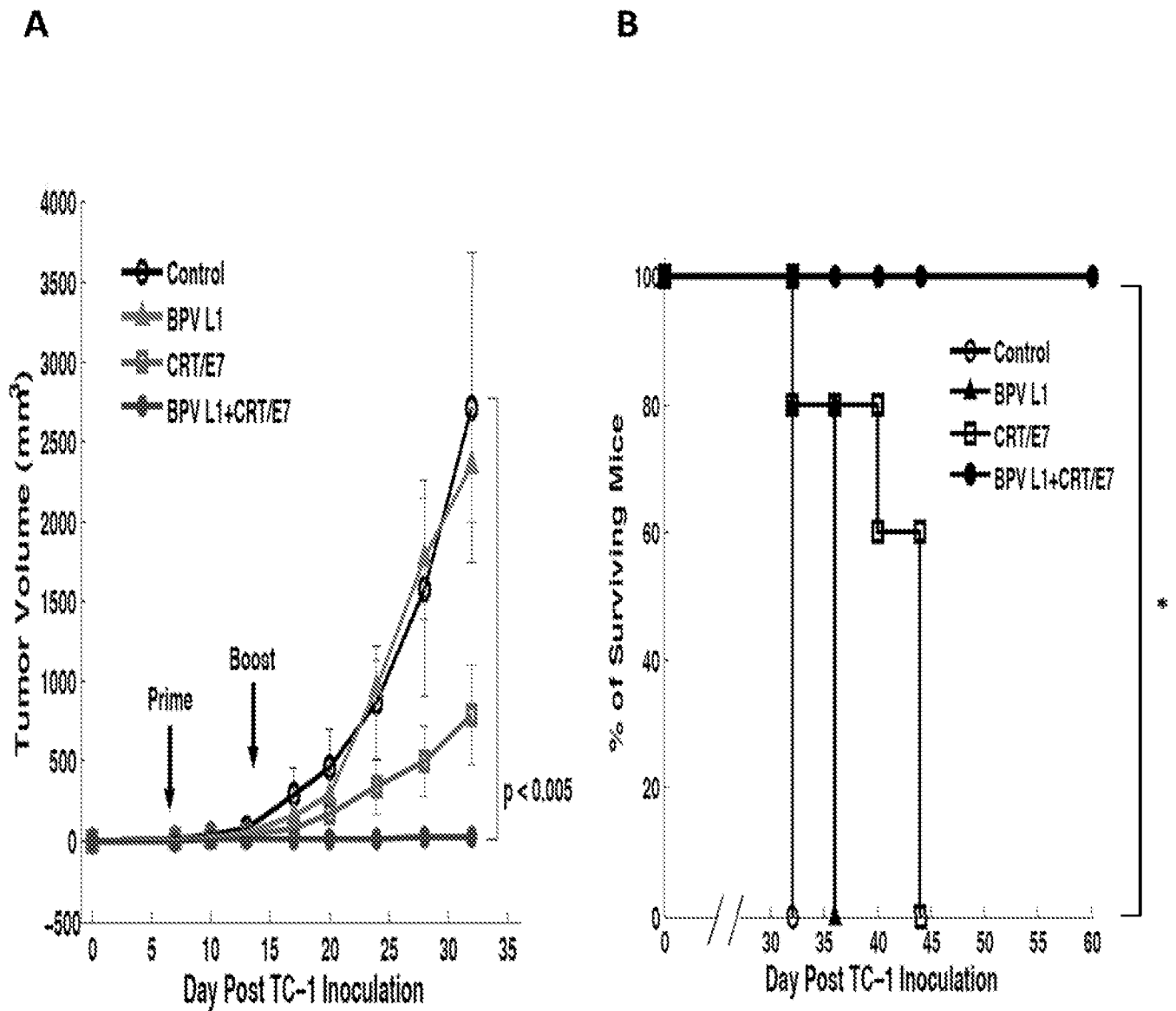


Figure 5

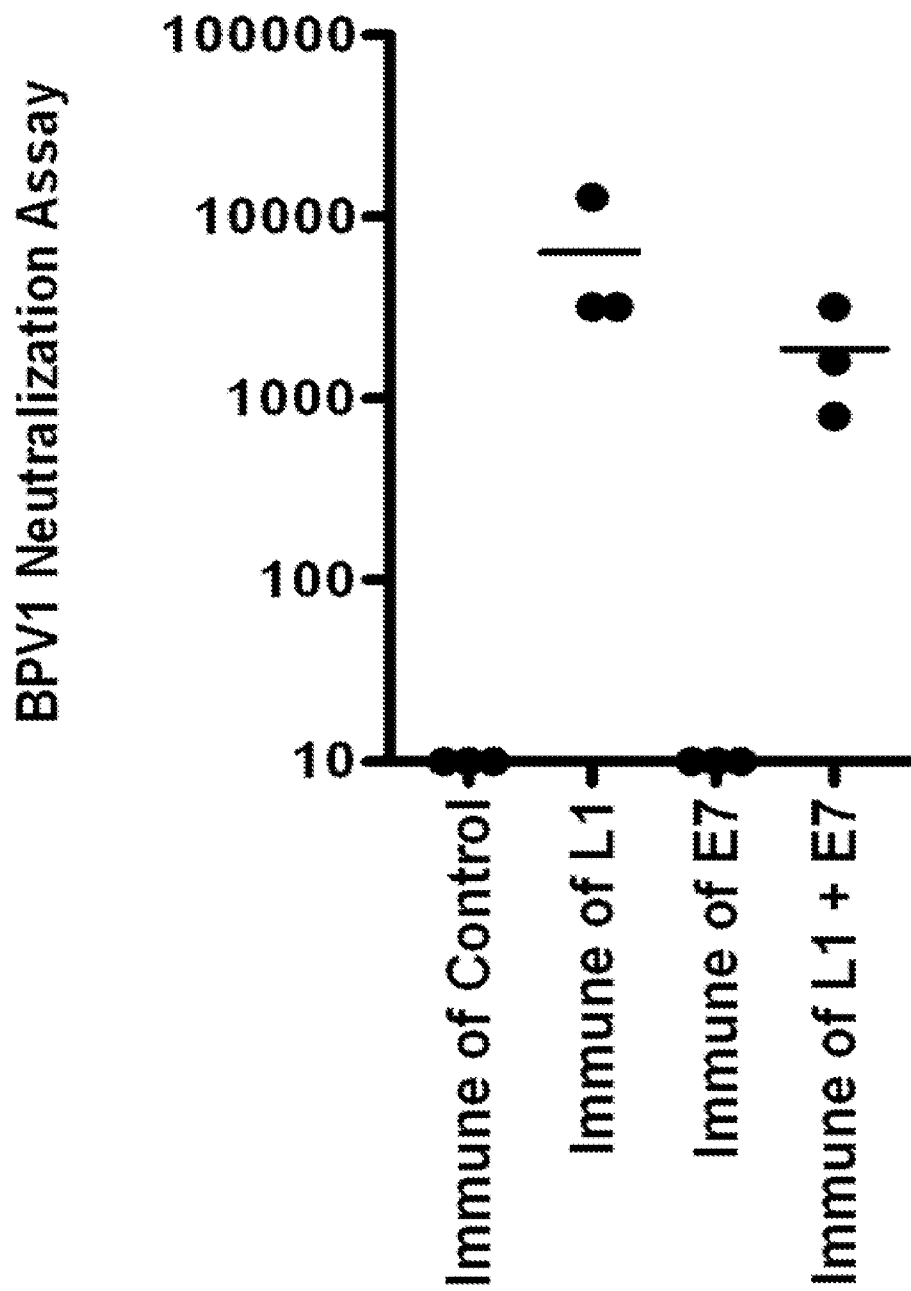


Figure 6

