



(86) Date de dépôt PCT/PCT Filing Date: 2004/09/17  
 (87) Date publication PCT/PCT Publication Date: 2005/03/31  
 (85) Entrée phase nationale/National Entry: 2006/03/16  
 (86) N° demande PCT/PCT Application No.: US 2004/030397  
 (87) N° publication PCT/PCT Publication No.: 2005/028625  
 (30) Priorités/Priorities: 2003/09/17 (US60/503,460);  
 2004/08/27 (US60/604,722)

(51) Cl.Int./Int.Cl. *A61K 39/21* (2006.01),  
*C07H 21/04* (2006.01), *C12N 15/74* (2006.01),  
*C12N 5/02* (2006.01)  
 (71) Demandeurs/Applicants:  
 DUKE UNIVERSITY, US;  
 UNIVERSITY OF ALABAMA AT BIRMINGHAM  
 RESEARCH FOUNDATION, US;  
 THE REGENTS OF THE UNIVERSITY OF CALIFORNIA,  
 US  
 (72) Inventeurs/Inventors:  
 KORBER, BETTE T., US;  
 HAHN, BEATRICE H., US;  
 SHAW, GEORGE M., US;  
 ...

(54) Titre : IMMUNOGENES CONSENSUS/ANCESTRAUX  
 (54) Title: CONSENSUS/ANCESTRAL IMMUNOGENS

(57) **Abrégé/Abstract:**

The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralizes a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.

(72) **Inventeurs(suite)/Inventors(continued)**: KOTHE, DENISE, US; LI, YING YING, US; DECKER, JULIE, US;  
HAYNES, BARTON F., US; GAO, FENG, US; LIAO, HUA-XIN, US

(74) **Agent**: GOUDREAU GAGE DUBUC

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

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
31 March 2005 (31.03.2005)

PCT

(10) International Publication Number  
**WO 2005/028625 A3**

- (51) International Patent Classification<sup>7</sup>: **A61K 39/21**, Durham, NC 27712 (US). **LIAO, Hua-Xin** [CN/US]; 200 Pebble Springs Road, Chapel Hill, NC 27514 (US).  
C07H 21/04, C12N 5/02, 15/74
- (21) International Application Number: PCT/US2004/030397
- (22) International Filing Date: 17 September 2004 (17.09.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/503,460 17 September 2003 (17.09.2003) US  
60/604,722 27 August 2004 (27.08.2004) US
- (71) Applicant (for all designated States except US): **DUKE UNIVERSITY** [US/US]; P.O. Box 90083, Durham, NC 27708-0083 (US).
- (71) Applicants and  
(72) Inventors: **KORBER, Bette, T.** [US/US]; 1290 Big Rock Loop, Los Alamos, NM 87544 (US). **HAHN, Beatrice, H.** [DE/US]; 3571 Rockhill Road, Birmingham, AL 35223 (US). **SHAW, George, M.** [US/US]; 3571 Rockhill Road, Birmingham, AL 35223 (US). **KOTHE, Denise** [US/US]; 734 39th Street, Birmingham, AL 35222 (US). **LI, Ying Ying** [CN/US]; 2529 Mountain Cove, Hoover, AL 35266 (US). **DECKER, Julie** [US/US]; 1116 Colonial Drive, Alabaster, AL 35007 (US).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): **HAYNES, Barton, F.** [US/US]; 3923 Wentworth Street, Durham, NC 27707 (US). **GAO, Feng** [US/US]; 702 Country Club Drive,
- (74) Agent: **WILSON, Mary, J.**; Nixon & Vanderhye P.C., 1100 North Glebe Road, Suite 800, Arlington, VA 22201-4714 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— with international search report
- (88) Date of publication of the international search report:  
2 June 2005
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CONSENSUS/ANCESTRAL IMMUNOGENS

(57) Abstract: The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralizes a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.



WO 2005/028625 A3

**CONSENSUS/ANCESTRAL IMMUNOGENS**

This application claims priority from Prov. Appln. No. 60/503,460, filed September 17, 2003, and Prov. Appln. No. 60/604,722, filed August 27, 2004, the entire contents of which are incorporated herein by reference.

**TECHNICAL FIELD**

The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.

**BACKGROUND**

The high level of genetic variability of HIV-1 has presented a major hurdle for AIDS vaccine development. Genetic differences among HIV-1 groups M, N, and O are extensive, ranging from 30% to 50% in *gag* and *env* genes, respectively (Gurtler et al, J. Virol. 68:1581-1585 (1994), Vanden Haesevelde et al, J. Virol. 68:1586-1596 (1994), Simon et al, Nat. Med. 4:1032-1037 (1998), Kuiken et al, Human

retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid sequences (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico)). Viruses within group M are further  
5 classified into nine genetically distinct subtypes (A-D, F-H, J and K) (Kuiken et al, Human retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid sequences  
10 (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, Robertson et al, Science 288:55-56 (2000), Robertson et al, Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino  
15 acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (2000)). With the genetic variation as high as 30% in env genes among HIV-1 subtypes, it has been  
20 difficult to consistently elicit cross-subtype T and B cell immune responses against all HIV-1 subtypes. HIV-1 also frequently recombines among different subtypes to create circulating recombinant forms (CRFs) (Robertson et al, Science 288:55-56 (2000),  
25 Robertson et al, Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505  
30 (2000), Carr et al, Human retroviruses and AIDS 1998: a compilation and analysis of nucleic acid and

amino acid sequences, eds. Korber et al (Theoretical  
Biology and Biophysics Group, Los Alamos National  
Laboratory, Los Alamos, New Mexico), pp. III-10-III-  
19 (1998)). Over 20% of HIV-1 isolates are  
5 recombinant in geographic areas where multiple  
subtypes are common (Robertson et al, Nature  
374:124-126 (1995), Cornelissen et al, J. virol.  
70:8209-8212 (1996), Dowling et al, AIDS 16:1809-  
1820 (2002)), and high prevalence rates of  
10 recombinant viruses may further complicate the  
design of experimental HIV-1 immunogens.

To overcome these challenges in AIDS vaccine  
development, three computer models (consensus,  
ancestor and center of the tree) have been used to  
15 generate centralized HIV-1 genes to (Gaschen et al,  
Science 296:2354-2360 (2002), Gao et al, Science  
299:1517-1518 (2003), Nickle et al, Science  
299:1515-1517 (2003), Novitsky et al, J. Virol.  
76:5435-5451 (2002), Ellenberger et al, Virology  
20 302:155-163 (2002), Korber et al, Science 288:1789-  
1796 (2000)). The biology of HIV gives rise to  
star-like phylogenies, and as a consequence of this,  
the three kinds of sequences differ from each other  
by 2 - 5% (Gao et al, Science 299:1517-1518 (2003)).  
25 Any of the three centralized gene strategies will  
reduce the protein distances between immunogens and  
field virus strains. Consensus sequences minimize  
the degree of sequence dissimilarity between a  
vaccine strain and contemporary circulating viruses  
30 by creating artificial sequences based on the most  
common amino acid in each position in an alignment

(Gaschen et al, Science 296:2354-2360 (2002)).  
Ancestral sequences are similar to consensus  
sequences but are generated using maximum-likelihood  
phylogenetic analysis methods (Gaschen et al,  
5 Science 296:2354-2360 (2002), Nickle et al, Science  
299:1515-1517 (2003)) . In doing so, this method  
recreates the hypothetical ancestral genes of the  
analyzed current wild-type sequences (Figure 26).  
Nickle et al proposed another method to generate  
10 centralized HIV-1 sequences, center of the tree  
(COT), that is similar to ancestral sequences but  
less influenced by outliers (Science 299:1515-1517  
(2003)).

The present invention results, at least in  
15 part, from the results of studies designed to  
determine if centralized immunogens can induce both  
T and B cell immune responses in animals. These  
studies involved the generation of an artificial  
group M consensus *env* gene (CON6), and construction  
20 of DNA plasmids and recombinant vaccinia viruses to  
express CON6 envelopes as soluble gp120 and gp140CF  
proteins. The results demonstrate that CON6 Env  
proteins are biologically functional, possess  
linear, conformational and glycan-dependent epitopes  
25 of wild-type HIV-1, and induce cytokine-producing T  
cells that recognize T cell epitopes of both HIV  
subtypes B and C. Importantly, CON6 gp120 and  
gp140CF proteins induce antibodies that neutralize  
subsets of subtype B and C HIV-1 primary isolates.

30 The iterative nature of study of the  
centralized HIV-1 gene approach is derived from the

rapidly expanding evolution of HIV-1 sequences, and the fact that sequences collected in the HIV sequence database (that is, the Los Alamos National Database) are continually being updated with new sequences each year. The CON6 gp120 envelope gene derives from Year 1999 Los Alamos National Database sequences, and Con-S derives from Year 2000 Los Alamos National Database sequences. In addition, CON6 has Chinese subtype C V1, V2, V4, and V5 Env sequences, while Con-S has all group M consensus Env constant and variable regions, that have been shortened to minimal-length variable loops. Codon-optimized genes for a series of Year 2003 group M and subtype consensus sequences have been designed, as have a corresponding series of wild-type HIV-1 Env genes for comparison, for use in inducing broadly reactive T and B cell responses to HIV-1 primary isolates.

#### SUMMARY OF THE INVENTION

The present invention relates to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response, and to nucleic acid sequences encoding same. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen.

Objects and advantages of the present invention will be clear from the description that follows.



BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1D: Generation and expression of the group M consensus env gene (CON6). The complete amino acid sequence of CON6 gp160 is shown.

5 (Fig. 1A) The five regions from the wild-type CRF08\_BC (98CN006) env gene are indicated by underlined letters. Variable regions are indicated by brackets above the sequences. Potential N-linked glycosylation sites are highlighted with bold-faced

10 letters. (Fig. 1B) Constructs of CON6 gp120 and gp140CF. CON6 gp120 and gp140CF plasmids were engineered by introducing a stop codon after the gp120 cleavage site or before the transmembrane domain, respectively. The gp120/gp41 cleavage site

15 and fusion domain of gp41 were deleted in the gp140CF protein. (Fig.1C) Expression of CON6 gp120 and gp140CF. CON6 gp120 and gp140CF were purified from the cell culture supernatants of rVV-infected 293T cells with *galanthus Nivalis argarose* lectin

20 columns. Both gp120 and gp140CF were separated on a 10% SDS-polyarylamide gel and stained with Commassie blue. (Fig. 1D.) CON6 env gene optimized based on codon usage for highly expressed human genes.

Figures 2A-2E. Binding of CON6 gp120 gp140 CF

25 to soluble CD4 (sCD4) and anti-Env mAbs. (Figs. 2A-2B) Each of the indicated mabs and sCD4 was covalently immobilized to a CM5 sensor chip (BIAcore) and CON6 gp120 (Fig. 2A) or gp140CF (Fig.

2B) (100  $\mu\text{g/ml}$  and 300  $\mu\text{g/ml}$ , respectively) were injected over each surface. Both gp120 and gp140CF proteins reacted with each anti-gp120 mabs tested except for 17b mab, which showed negligible binding to both CON6 gp120 and gp140CF. To determine induction of 17b mab binding to CON6 gp120 and gp140CF, CON6 gp120 (Fig. 2C) or gp140CF (Fig. 2D) proteins were captured (400-580 RU) on individual flow cells immobilized with sCD4 or mabs A32 or T8. Following stabilization of each of the surface, mAb 17b was injected and flowed over each of the immobilized flow cells. Overlay of curves show that the binding of mab 17b to CON6 Env proteins was markedly enhanced on both sCD4 and mab A32 surfaces but not on the T8 surface (Figs. 2C-2D). To determine binding of CON6 gp120 and gp140CF to human mabs in ELISA, stock solutions of 20 $\mu\text{g/ml}$  of mabs 447, F39F, A32, IgG1b12 and 2F5 on CON6 gp120 and gp140CF were tittered (Fig. 2E). Mabs 447 (V3), F39F (V3) A32 (gp120) and IgG1b12 (CD4 binding site) each bound to both CON6 gp120 and 140 well, while 2F5 (anti-gp41 ELDKWAS) only bound gp140CF. The concentration at endpoint titer on gp120 for mab 447 and F39F binding was <0.003  $\mu\text{g/ml}$  and 0.006  $\mu\text{g/ml}$ , respectively; for mab A32 was <0.125  $\mu\text{g/ml}$ ; for IgG1b12 was <0.002  $\mu\text{g/ml}$ ; and for 2F5 was 0.016  $\mu\text{g/ml}$ .

Figures 3A and 3B. Infectivity and coreceptor usage of CON6 envelope. (Fig. 3A) CON6 and control

env plasmids were cotransfected with HIV-1/SG3 $\Delta$ env backbone into human 293T cells to generate Env-pseudovirions. Equal amounts of each pseudovirion (5 ng p24) were used to infect JC53-BL cells. The infectivity was determined by counting the number of blue cells (infectious units, IU) per microgram of p24 of pseudovirions (IU/ $\mu$ g p24) after staining the infected cells for  $\beta$ -gal expression. (Fig. 3B) Coreceptor usage of the CON6 env gene was determined on JC53BL cells treated with AMD3100 and/or TAK-799 for 1 hr (37°C) then infected with equal amounts of p24 (5 ng) of each Env-pseudovirion. Infectivity in the control group (no blocking agent) was set as 100%. Blocking efficiency was expressed as the percentage of IU from blocking experiments compared to those from control cultures without blocking agents. Data shown are mean  $\pm$  SD.

Figure 4. Western blot analysis of multiple subtype Env proteins against multiple subtype antisera. Equal amount of Env proteins (100 ng) were separated on 10% SDS-polyacrylamide gels. Following electrophoresis, proteins were transferred to Hybond ECL nitrocellulose membranes and reacted with sera from HIV-1 infected patients (1:1,000) or guinea pigs immunized with CON6 gp120 DNA prime, rVV boost (1:1,000). Protein-bound antibody was probed with fluorescent-labeled secondary antibodies and the images scanned and recorded on an infrared imager Odyssey (Li-Cor, Lincoln, NE). Subtypes are

indicated by single-letters after Env protein and serum IDs. Four to six sera were tested for each subtype, and reaction patterns were similar among all sera from the same subtype. One representative  
5 result for each subtype serum is shown.

Figure 5. T cell immune responses induced by CON6 Env immunogens in mice. Splenocytes were isolated from individual immunized mice (5 mice/group). After splenocytes were stimulated *in*  
10 *vitro* with overlapping Env peptide pools of CON6 (black column), subtype B (hatched column), subtype C (white column), and medium (no peptide; gray column), INF- $\gamma$  producing cells were determined by the ELISPOT assay. T cell IFN- $\gamma$  responses induced  
15 by either CON6 gp120 or gp140CF were compared to those induced by subtype specific Env immunogens (JRFL and 96ZM651). Total responses for each envelope peptide pool are expressed as SFCs per million splenocytes. The values for each column are  
20 the mean  $\pm$  SEM (of IFN- $\gamma$  SFCs (n=5 mice/group)).

Figures 6A-6E. Construction of codon usage optimized subtype C ancestral and consensus envelope genes (Figs. 6A and 6B, respectively). Ancestral and consensus amino acid sequences (Figs. 6C and 6D,  
25 respectively) were transcribed to mirror the codon usage of highly expressed human genes. Paired oligonucleotides (80-mers) overlapping by 20 bp were designed to contain 5' invariant sequences including

the restriction enzyme sites EcoRI, BbsI, Bam HI and BsmBI. BbsI and BsmBI are Type II restriction enzymes that cleave outside of their recognition sequences. Paired oligomers were linked  
5 individually using PCR and primers complimentary to the 18 bp invariant sequences in a stepwise fashion, yielding 140bp PCR products. These were subcloned into pGEM-T and sequenced to confirm the absence of inadvertant mutations/deletions. Four individual  
10 pGEM-T subclones containing the proper inserts were digested and ligated together into pcDNA3.1. Multi-fragment ligations occurred repeatedly amongst groups of fragments in a stepwise manner from the 5' to the 3' end of the gene until the entire gene was  
15 reconstructed in pcDNA3.1. (See schematic in Fig. 6E.)

Figure 7. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also contain  
20 the reporter cassettes of luciferase and  $\beta$ -galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24 or 96-well plates, incubated at 37°C  
25 for 24 hours and treated with DEAE-Dextran at 37°C for 30 minutes. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-Dextran, and allowed to incubate for 3 hours at 37°C after which an additional cell media is added to each

well. Following a final 48-hour incubation at 37°C, cells are either fixed, stained using X-Gal to visualize  $\beta$ -galactosidase expressing blue foci or frozen-thawed three times to measure luciferase activity.

Figure 8. Sequence alignment of subtype C ancestral and consensus *env* genes. Alignment of the subtype C ancestral (bottom line) and consensus (top line) *env* sequences showing a 95.5% sequence homology; amino acid sequence differences are indicated. One noted difference is the addition of a glycosylation site in the C ancestral *env* gene at the base of the V1 loop. A plus sign indicates a within-class difference of amino acid at the indicated position; a bar indicates a change in the class of amino acid. Potential N-glycosylation sites are marked in blue. The position of truncation for the *gp140* gene is also shown.

Figure 9. Expression of subtype C ancestral and consensus envelopes in 293T cells. Plasmids containing codon-optimized *gp160*, *gp140*, or *gp120* subtype C ancestral and consensus genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2  $\mu$ g of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins

were transferred to a PVDF membrane and probed with HIV-1 plasma from a subtype C infected patient.

Figures 10A and 10B. Fig. 10A. *Trans* complementation of *env*-deficient HIV-1 with codon-optimized subtype C ancestral and consensus *gp160* and *gp140*. Plasmids containing codon-optimized, subtype C ancestral or consensus *gp160* or *gp140* genes were co-transfected into 293T cells with an HIV-1/SG3Δ*env* provirus. 48 hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified by centrifugation, filtered through a 0.2μM filter, and pelleted through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel for particles containing a codon-optimized envelope. 250ng of p24 was loaded per lane for particles generated by co-transfection of a rev-dependent wild-type subtype C 96ZAM651*env* gene. Differences in the amount of p24 loaded per lane were necessary to ensure visualization of the rev-dependent envelopes by Western Blot. Proteins were transferred to a PVDF membrane and probed with pooled plasma from HIV-1 subtype B and subtype C infected individuals. Fig. 10B. Infectivity of virus particles containing subtype C ancestral and consensus envelope glycoproteins. Infectivity of pseudotyped virus containing ancestral or consensus *gp160* or *gp140* envelope was determined using the

JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL  
5 cells. Following a 48-hour incubation period, cells were fixed and stained to visualize  $\beta$ -galactosidase expressing cells. Infectivity is represented as infectious units per ng of p24 to normalize for differences in the concentration of the input  
10 pseudovirions.

Figure 11. Co-receptor usage of subtype C ancestral and consensus envelopes. Pseudotyped particles containing ancestral or consensus envelope were incubated with DEAE-Dextran treated JC53-BL  
15 cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), or AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4, and YU-2, a known CCR5-using isolate, were included  
20 as controls.

Figures 12A-12C. Neutralization sensitivity of subtype C ancestral and consensus envelope glycoproteins. Equivalent amounts of pseudovirions containing the ancestral, consensus or 96ZAM651  
25 *gp160* envelopes (1,500 infectious units) were pre-incubated with a panel of plasma samples from HIV-1 subtype C infected patients and then added to the JC53-BL cell monolayer in 96-well plates. Plates



were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity is calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC<sub>50</sub>) and the actual % neutralization at each antibody dilution are then calculated for each virus. The results of all luciferase experiments are confirmed by direct counting of blue foci in parallel infections.

Figures 13A-13F. Protein expression of consensus subtype C Gag (Fig. 13A) and Nef (Fig. 13B) following transfection into 293T cells. Consensus subtype C Gag and Nef amino acid sequences are set forth in Figs. 13C and 13D, respectively, and encoding sequences are set forth in Figs. 13E and 13F, respectively.

Figures 14A-14C. Figs. 14A and 14B show the Con-S Env amino acid sequence and encoding sequence, respectively. Fig. 14C shows expression of Group M consensus Con-S Env proteins using an *in vitro* transcription and translation system.

Figures 15A and 15B. Expression of Con-S *env* gene in mammalian cells. (Fig. 15A - cell lysate, Fig. 15B - supernatant.)

Figures 16A and 16B. Infectivity (Fig. 16A) and coreceptor usage (Fig. 16B) of CON6 and Con-S *env* genes.

Figures 17A-17C. Env protein incorporation in  
5 CON6 and Con-S Env-pseudovirions. (Fig. 17A -  
lysate, Fig. 17B - supernatant, Fig. 17C pellet.)

Figures 18A-18D. Figs. 18A and 18B show  
subtype A consensus Env amino acid sequence and  
nucleic acid sequence encoding same, respectively.  
10 Figs. 18C and 18D show expression of A.con *env* gene  
in mammalian cells (Fig. 18C - cell lysate, Fig. 18D  
- supernatant).

Figures 19A-19H. M.con.gag (Fig. 19A),  
M.con.pol (Fig. 19B), M.con.nef (Fig. 19C) and  
15 C.con.pol (Fig. 19D) nucleic acid sequences and  
corresponding encoded amino acid sequences (Figs.  
19E-19H, respectively).

Figures 20A-20D. Subtype B consensus *gag* (Fig.  
20A) and *env* (Fig.20B) genes. Corresponding amino  
20 acid sequences are shown in Figs. 20C and 20D.

Figure 21. Expression of subtype B consensus  
*env* and *gag* genes in 293T cells. Plasmids  
containing codon-optimized subtype B consensus  
*gp160*, *gp140*, and *gag* genes were transfected into  
25 293T cells, and protein expression was examined by

Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2  $\mu$ g of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.

Figure 22. Co-receptor usage of subtype B consensus envelopes. Pseudotyped particles containing the subtype B consensus gp160 Env were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4 and YU-2, a known CCR5-using isolate, were included as controls.

Figures 23A and 23B. *Trans* complementation of env-deficient HIV-1 with codon-optimized subtype B consensus gp160 and gp140 genes. Plasmids containing codon-optimized, subtype B consensus gp160 or gp140 genes were co-transfected into 293T cells with an HIV-1/SG3 $\Delta$ env provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a tabletop centrifuge, filtered through a 0.2 $\mu$ M filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen

assay; 25 ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. *Trans*  
5 complementation with a rev-dependent NL4.3 env was included for control. Figure 23B. Infectivity of virus particles containing the subtype B consensus envelope. Infectivity of pseudotyped virus containing consensus B gp160 or gp140 was determined  
10 using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation  
15 period, cells were fixed and stained to visualize  $\beta$ -galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

Figures 24A-24D. Neutralization sensitivity of virions containing subtype B consensus gp160  
20 envelope. Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1  
25 subtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by

dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC<sub>50</sub>) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections. Fig. 24A. Neutralization of Pseudovirions containing Subtype B consensus Env (gp160). Fig. 24B. Neutralization of Pseudovirions containing NL4.3 Env (gp160). Fig. 24C. Neutralization of Pseudovirions containing Subtype B consensus Env (gp160). Fig. 24D. Neutralization of Pseudovirions containing NL4.3 Env (gp160).

Figures 25A and 25B. Fig. 25A. Density and p24 analysis of sucrose gradient fractions. 0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml. Fig. 25B. VLP production by co-transfection of subtype B consensus *gag* and *env* genes. 293T cells were co-transfected with subtype B consensus *gag* and

env genes. Cell supernatants were harvested 48-  
hours post-transfection, clarified through at 20%  
sucrose cushion, and further purified through a 20-  
60% sucrose gradient. Select fractions from the  
5 gradient were pooled, added to 20ml of PBS, and  
centrifuged overnight at 100,000 x g. Resuspended  
pellets were loaded onto a 4-20% SDS-PAGE gel,  
proteins were transferred to a PVDF membrane, and  
probed with plasma from an HIV-1 subtype B infected  
10 individual.

Figures 26A and 26B. Fig. 26A. 2000 Con-S  
140CFI.ENV. Fig. 26B. Codon-optimized Year 2000  
Con-S 140CFI.seq.

Figure 27. Individual C57BL/6 mouse T cell  
15 responses to HIV-1 envelope peptides. Comparative  
immunogenicity of CON6 gp140CFI and Con-S gp140CFI  
in C57BL/C mice. Mice were immunized with either  
HIV5305 (Subtype A), 2801 (Subtype B), CON6 or Con-S  
Envelope genes in DNA prime, rVV boost regimens, 5  
20 mice per group. Spleen cells were assayed for IFN- $\gamma$   
spot-forming cells 10 days after rVV boost, using  
mixtures of overlapping peptides from Envs of HIV-1  
UG37(A), MN(B), Ch19(C), 89.6(B) SF162(B) or no  
peptide negative control.

25 Figures 28A-28C. Fig. 28A. Con-B 2003 Env. pep  
(841 a.a.). Amino acid sequence underlined is the  
fusion domain that is deleted in 140CF design and

the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 28B. Con-B-140CF.pep (632 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 28C. Codon-optimized Con-B 140CF.seq (1927 nt.).

Figures 29A-29C. Fig. 29A. CON\_OF\_CONS-2003 (829 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 29B. ConS-2003 140CF.pep (620 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 29C. CODON-OPTIMIZED ConS-2003 140CF.seq (1891 nt.).

Figures 30A-30C. Fig. 30A. CONSENSUS\_A1-2003 (845 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 30B. Con-A1-2003 140CF.pep (629 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 30C. CODON-OPTIMIZED Con-A1-2003.seq.

Figures 31A-31C. Fig. 31A. CONSENSUS\_C-2003 (835 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 31B. Con-C 2003 140CF.pep (619 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 31C. CODON-OPTIMIZED Con-C-2003 (140 CF (1,888 nt.).

Figures 32A-32C. Fig. 32A. CONSENSUS\_G-2003 (842 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 32B. Con-G-2003 140CF.pep (626 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 32C. CODON-OPTIMIZED Con-G-2003.seq.

Figures 33A-33C. Fig. 33A. CONSENSUS\_01\_AE-2003 (854 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 33B. Con-AE01-2003 140CF.pep (638 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage



site. Fig. 33C. CODON-OPTIMIZED Con-AE01-2003.seq.  
(1945 nt.).

Figures 34A-34C. Fig. 34A. Wild-type subtype  
A Env. 00KE\_MSA4076-A (Subtype A, 891 a.a.). Amino  
5 acid sequence underlined is the fusion domain that  
is deleted in 140CF design and the "W" underlined  
is the last amino acid at the C-terminus, all amino  
acids after the "W" are deleted in the 140CF design.  
Fig. 34B. 00KE\_MSA4076-A 140CF.pep (647 a.a.).  
10 Amino acids in bold identify the junction of the  
deleted fusion cleavage site. Fig. 34C. CODON-  
OPTIMIZED 00KE\_MSA4076-A 140CF.seq. (1972 nt.).

Figures 35A-35C. Fig. 35A. Wild-type subtype  
B. QH0515.1g gp160 (861 a.a.). Amino acid sequence  
15 underlined is the fusion domain that is deleted in  
140CF design and the "W" underlined is the last  
amino acid at the C-terminus, all amino acids after  
the "W" are deleted in the 140CF design. Fig. 35B.  
QH0515.1g 140CF (651 a.a.). Amino acids in bold  
20 identify the junction of the deleted fusion cleavage  
site. Fig. 35C. CODON-OPTIMIZED QH0515.1g  
140CF.seq (1984 nt.).

Figures 36A-36C. Fig. 36A. Wild-type subtype  
C. DU123.6 gp160 (854 a.a.). Amino acid sequence  
25 underlined is the fusion domain that is deleted in  
140CF design and the "W" underlined is the last  
amino acid at the C-terminus, all amino acids after

the "W" are deleted in the 140CF design. Fig. 36B.  
DU123.6 140CF (638 a.a.). Amino acids in bold  
identify the junction of the deleted fusion cleavage  
site. Fig. 36C. CODON-OPTIMIZED DU123.6 140CF.seq  
5 (1945 nt.).

Figures 37A-37C. Fig. 37A. Wild-type subtype  
CRF01\_AE. 97CNGX2F-AE (854 a.a.). Amino acid  
sequence underlined is the fusion domain that is  
deleted in 140CF design and the "W" underlined is  
10 the last amino acid at the C-terminus, all amino  
acids after the "W" are deleted in the 140CF design.  
Fig. 37B. 97CNGX2F-AE 140CF.pep (629 a.a.). Amino  
acids in bold identify the junction of the deleted  
fusion cleavage site. Fig. 37C. CODON-OPTIMIZED  
15 97CNGX2F-AE 140CF.seq (1921 nt.).

Figures 38A-38C. Fig. 38A. Wild-type DRCBL-G  
(854 a.a.). Amino acid sequence underlined is the  
fusion domain that is deleted in 140CF design and  
the "W" underlined is the last amino acid at the  
20 C-terminus, all amino acids after the "W" are  
deleted in the 140CF design. Fig. 38B. DRCBL-G  
140CF.pep (630 a.a.). Amino acids in bold identify  
the junction of the deleted fusion cleavage site.  
Fig. 38C. CODON-OPTIMIZED DRCBL-G 140CF.seq (1921  
25 nt.).

Figures 39A and 39B. Fig. 39A. 2003 Con-S  
Env. Fig. 39B. 2003 Con-S Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 40A and 40B. Fig. 40A. 2003 M.  
5 Group.Anc Env. Fig. 40B. 2003 M. Group.anc  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

Figures 41A and 41B. Fig. 41A. 2003 CON\_A1  
Env. Fig. 41B. 2003 CON\_A1 Env.seq.opt.  
10 (Seq.opt. = codon optimized encoding sequence.)

Figures 42A and 42B. Fig. 42A. 2003 A1.Anc  
Env. Figs. 42B. 2003 A1.anc Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 43A and 43B. Fig. 43A. 2003 CON\_A2  
15 Env. Fig. 43B. 2003 CON\_A2 Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 44A and 44B. Fig. 44A. 2003 CON\_B  
Env. Fig. 44B. 2003 CON\_B Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 45A and 45B. Fig. 45A. 2003 B.anc  
20 Env. Figs. 45B. 2003 B.anc Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 46A and 46B. Fig. 46A. 2003 CON\_C  
Env. Fig. 46B. 2003 CON\_C Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 47A and 47B. Fig. 47A. 2003 C.anc  
5 Env. Fig. 47B. 2003 C.anc Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 48A and 48B. Fig. 48A. 2003 CON\_D  
Env. Fig. 48B. 2003 CON\_D Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

10 Figures 49A and 49B. Fig. 49A. 2003 CON\_F1  
Env. Fig. 49B. 2003 CON\_F1 Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 50A and 50B. Fig. 50A. 2003 CON\_F2  
Env. Fig. 50B. 2003 CON\_F2 Env.seq.opt.  
15 (Seq.opt. = codon optimized encoding sequence.)

Figures 51A and 51B. Fig. 51A. 2003 CON\_G  
Env. Fig. 51B. 2003 CON\_G Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 52A and 52B. Fig. 52A. 2003 CON\_H  
20 Env. Fig. 52B. 2003 CON\_H Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 53A and 53B. Fig. 53A. 2003 CON\_01\_AE  
Env. Fig. 53B. 2003 CON\_01\_AE Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 54A and 54B. Fig. 54A. 2003 CON\_02\_AG  
5 Env. Fig. 54B. 2003 CON\_02\_AG Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 55A and 55B. Fig. 55A. 2003 CON\_03\_AB  
Env. Fig. 55B. 2003 CON\_03\_AB Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

10 Figures 56A and 56B. Fig. 56A. 2003  
CON\_04\_CPX Env. Fig. 56B. 2003 CON\_04\_CPX  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

Figures 57A and 57B. Fig. 57A. 2003  
15 CON\_06\_CPX Env. Fig. 57B. 2003 CON\_06\_CPX  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

Figures 58A and 58B. Fig. 58A. 2003 CON\_08\_BC  
Env. Fig. 58B. 2003 CON\_08\_BC Env.seq.opt.  
20 (Seq.opt. = codon optimized encoding sequence.)

Figures 59A and 59B. Fig. 59A. 2003 CON\_10\_CD  
Env. Fig. 59B. 2003 CON\_10\_CD Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 60A and 60B. Fig. 60A. 2003  
CON\_11\_CPX Env. Fig. 60B. 2003 CON\_11\_CPX  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

5 Figures 61A and 61B. Fig. 61A. 2003 CON\_12\_BF  
Env. Fig. 61B. 2003 CON\_12\_BF Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 62A and 62B. Fig. 62A. 2003 CON\_14\_BG  
Env. Fig. 62B. 2003 CON\_14\_BG Env.seq.opt.  
10 (Seq.opt. = codon optimized encoding sequence.)

Figures 63A and 63B. Fig. 63A. 2003\_CON\_S  
gag.PEP. Fig. 63B. 2003\_CON\_S gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 64A and 64B. Fig. 64A.  
15 2003\_M.GROUP.anc gag.PEP. Fig. 64B.  
2003\_M.GROUP.anc gag.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 65A-65D. Fig. 65A. 2003\_CON\_A1  
gag.PEP. Fig. 65B. 2003\_CON\_A1 gag.OPT. Fig. 65C.  
20 2003\_A1.anc gag.PEP. Fig. 65D. 2003\_A1.anc  
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 66A and 66B. Fig. 66A. 2003\_CON\_A2  
gag.PEP. Fig. 66B. 2003\_CON\_A2 gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 67A-67D. Fig. 67A. 2003\_CON\_B  
5 gag.PEP. Fig. 67B. 2003\_CON\_B gag.OPT. Fig. 67C.  
2003\_B.anc gag.PEP. Fig. 67D. 2003\_B.anc gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 68A-68D. Fig. 68A. 2003\_CON\_C  
gag.PEP. Fig. 68B. 2003\_CON\_C gag.OPT. Fig. 68C.  
10 2003\_C.anc.gag.PEP. Fig. 68D. 2003\_C.anc.gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 69A and 69B. Fig. 69A. 2003\_CON\_D  
gag.PEP. Fig. 69B. 2003\_CON\_D gag.OPT.  
(OPT = codon optimized encoding sequence.)

15 Figures 70A and 70B. Fig. 70A. 2003\_CON\_F  
gag.PEP. Fig. 70B. 2003\_CON\_F gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 71A and 71B. Fig. 71A. 2003\_CON\_G  
gag.PEP. Fig. 71B. 2003\_CON\_G gag.OPT.  
20 (OPT = codon optimized encoding sequence.)

Figures 72A and 72B. Fig. 72A. 2003\_CON\_H  
gag.PEP. Fig. 72B. 2003\_CON\_H gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 73A and 73B. Fig. 73A. 2003\_CON\_K  
gag.PEP. Fig. 73B. 2003\_CON\_K gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 74A and 74B. Fig. 74A. 2003\_CON\_01\_AE  
5 gag.PEP. Fig. 7B. 2003\_CON\_01\_AE gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 75A and 75B. Fig. 75A. 2003\_CON\_02\_AG  
gag.PEP. Fig. 75B. 2003\_CON\_02\_AG gag.OPT.  
(OPT = codon optimized encoding sequence.)

10 Figures 76A and 76B. Fig. 76A.  
2003\_CON\_03\_ABG gag.PEP. Fig. 76B. 2003\_CON\_03\_ABG  
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 77A and 77B. Fig. 77A.  
2003\_CON\_04\_CFX gag.PEP. Fig. 77B. 2003 CON\_04\_CFX  
15 gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 78A and 78B. Fig. 78A.  
2003\_CON\_06\_CPX gag.PEP. Fig. 78B. 2003\_CON\_06\_CPX  
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 79A and 79B. Fig. 79A. 2003\_CON\_07\_BC  
20 gag.PEP. Fig. 79B. 2003\_CON\_07\_BC gag.OPT.  
(OPT = codon optimized encoding sequence.)



Figures 80A and 80B. Fig. 80A. 2003\_CON\_08\_BC  
gag.PEP. Fig. 80B. 2003\_CON\_08\_BC gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 81A and 81B. Fig. 81A. 2003\_CON\_10\_CD  
5 gag.PEP. Fig. 81B. 2003\_CON\_10\_CD gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 82A and 82B. Fig. 82A.  
2003\_CON\_11\_CPX gag.PEP. Fig. 82B. 2003\_CON\_11\_CPX  
gag.OPT. (OPT = codon optimized encoding sequence.)

10 Figures 83A and 83B. Fig. 83A.  
2003\_CON\_12\_BF.gag.PEP. Fig. 83B.  
2003\_CON\_12\_BF.gag.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 84A and 84B. Fig. 84A. 2003\_CON\_14\_BG  
15 gag.PEP. Fig. 84B. 2003\_CON\_14\_BG gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 85A and 85B. Fig. 85A. 2003\_CONS  
nef.PEP. Fig. 85B. 2003\_CONS nef.OPT.  
(OPT = codon optimized encoding sequence.)

20 Figures 86A and 86B. Fig. 86A. 2003\_M  
GROUP.anc nef.PEP. Fig. 86B. 2003\_M  
GROUP.anc.nef.OPT. (OPT = codon optimized encoding  
sequence.)

Figures 87A and 87B. Fig. 87A. 2003\_CON\_A  
nef.PEP. Fig. 87B. 2003\_CON\_A nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 88A-88D. Fig. 88A. 2003\_CON\_A1  
5 nef.PEP. Fig. 88B. 2003\_CON\_A1 nef.OPT. Fig. 88C.  
2003\_A1.anc nef.PEP. Fig. 88D. 2003\_A1.anc  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 89A and 89B. Fig. 89A. 2003\_CON\_A2  
nef.PEP. Fig. 89B. 2003\_CON\_A2 nef.OPT.  
10 (OPT = codon optimized encoding sequence.)

Figures 90A-90D. Fig. 90A. 2003\_CON\_B  
nef.PEP. Fig. 90B. 2003\_CON-B nef.OPT. Fig. 90C.  
2003\_B.anc nef.PEP. Fig. 90D. 2003\_B.anc nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 91A and 91B. Fig. 91A. 2003\_CON\_02\_AG  
15 nef.PEP. Fig. 91B. 2003\_CON\_02\_AG nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 92A-92D. Fig. 92A. 2003\_CON\_C  
nef.PEP. Fig. 92B. 2003\_CON\_C nef.OPT. Fig. 92C.  
20 2003\_C.anc nef.PEP. Fig. 92D. 2003\_C.anc nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 93A and 93B. Fig. 93A. 2003\_CON\_D  
nef.PEP. Fig. 93B. 2003\_CON\_D nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 94A and 94B. Fig. 94A. 2003\_CON\_F1  
5 nef.PEP. Fig. 94B. 2003\_CON\_F1 nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 95A and 95B. Fig. 95A. 2003\_CON\_F2  
nef.PEP. Fig. 95B. 2003\_CON\_F2 nef.OPT.  
(OPT = codon optimized encoding sequence.)

10 Figures 96A and 96B. Fig. 96A. 2003\_CON\_G  
nef.PEP. Fig. 96B. 2003\_CON\_G nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 97A and 97B. Fig. 97A. 2003\_CON\_H  
nef.PEP. Fig. 97B. 2003\_CON\_H nef.OPT.  
15 (OPT = codon optimized encoding sequence.)

Figures 98A and 98B. Fig. 98A. 2003\_CON\_01\_AE  
nef.PEP. Fig. 98B. 2003\_CON\_01\_AE nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 99A and 99B. Fig. 99A. 2003\_CON\_03\_AE  
20 nef.PEP. Fig. 99B. 2003\_CON\_03\_AE nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 100A and 100B. Fig. 100A.

2003\_CON\_04\_CFX nef.PEP. Fig. 100B.

2003\_CON\_04\_CFX nef.OPT. (OPT = codon optimized encoding sequence.)

5 Figures 101A and 101B. Fig. 101A.

2003\_CON\_06\_CFX nef.PEP. Fig. 101B.

2003\_CON\_06\_CFX nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 102A and 102B. Fig. 102A.

10 2003\_CON\_08\_BC nef.PEP. Fig. 102B. 2003\_CON\_08\_BC  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 103A and 103B. Fig. 103A.

2003\_CON\_10\_CD nef.PEP. Fig. 103B. 2003\_CON\_10\_CD  
nef.OPT. (OPT = codon optimized encoding sequence.)

15 Figures 104A and 104B. Fig. 104A.

2003\_CON\_11\_CFX nef.PEP. Fig. 104B.

2003\_CON\_11\_CFX nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 105A and 105B. Fig. 105A.

20 2003\_CON\_12\_BF nef.PEP. Fig. 105B. 2003\_CON\_12\_BF  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 106A and 106B. Fig. 106A.  
2003\_CON\_14\_BG nef.PEP. Fig. 106B. 2003\_CON\_14\_BG  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 107A and 107B. Fig. 107A. 2003\_CON\_S  
5 pol.PEP. Fig. 107B. 2003\_CON\_S pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 108A and 108B. Fig. 108A. 2003\_M  
GROUP anc pol.PEP. Fig. 108B. 2003\_M.GROUP anc  
pol.OPT. (OPT = codon optimized encoding sequence.)

10 Figures 109A-109D. Fig. 109A. 2003\_CON\_A1  
pol.PEP. Fig. 109B. 2003\_CON\_A1 pol.OPT.  
Fig. 109C. 2003\_A1.anc pol.PEP. Fig. 109D.  
2003\_A1.anc pol.OPT. (OPT = codon optimized  
encoding sequence.)

15 Figures 110A and 110B. Fig. 110A. 2003\_CON\_A2  
pol.PEP. Fig. 110B. 2003\_CON\_A2 pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 111A-111D. Fig. 111A. 2003\_CON\_B  
pol.PEP. Fig. 111B. 2003\_CON\_B pol.OPT. Fig.  
20 111C. 2003\_B.anc pol.PEP. Fig. 111D. 2003\_B.anc  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 112A-112D. Fig. 112A. 2003\_CON\_C  
pol.PEP. Fig. 112B. 2003\_CON\_C pol.OPT.

Fig. 112C. 2003\_C.anc pol.PEP. Fig. 112D.  
2003\_C.anc pol.OPT. (OPT = codon optimized encoding  
sequence.)

Figures 113A and 113B. Fig. 113A. 2003\_CON\_D  
5 pol.PEP. Fig. 113B. 2003\_CON\_D pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 114A and 114B. Fig. 114A. 2003\_CON\_F1  
pol.PEP. Fig. 114B. 2003\_CON\_F1 pol.OPT.  
(OPT = codon optimized encoding sequence.)

10 Figures 115A and 115B. Fig. 115A. 2003\_CON\_F2  
pol.PEP. Fig. 115B. 2003\_CON\_F2 pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 116A and 116B. Fig. 116A. 2003\_CON\_G  
pol.PEP. Fig. 116B. 2003\_CON\_G pol.OPT.  
15 (OPT = codon optimized encoding sequence.)

Figures 117A and 117B. Fig. 117A. 2003\_CON\_H  
pol.PEP. Fig. 117B. 2003\_CON\_H pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 118A and 118B. Fig. 118A.  
20 2003\_CON\_01\_AE pol.PEP. Fig. 118B. 2003\_CON\_01\_AE  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 119A and 119B. Fig. 119A.

2003\_CON\_02\_AG pol.PEP. Fig. 119B. 2003\_CON\_02\_AG  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 120A and 120B. Fig. 120A.

5 2003\_CON\_03\_AB pol.PEP. Fig. 120B. 2003\_CON\_03\_AB  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 121A and 121B. Fig. 121A.

2003\_CON\_04\_CPX pol.PEP. Fig. 121B.  
2003\_CON\_04\_CPX pol.OPT. (OPT = codon optimized  
10 encoding sequence.)

Figures 122A and 122B. Fig. 122A.

2003\_CON\_06\_CPX pol.PEP. Fig. 122B.  
2003\_CON\_06\_CPX pol.OPT. (OPT = codon optimized  
encoding sequence.)

15 Figures 123A and 123B. Fig. 123A.

2003\_CON\_08\_BC pol.PEP. Fig. 123B. 2003\_CON\_08\_BC  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 124A and 124B. Fig. 124A.

2003\_CON\_10\_CD pol.PEP. Fig. 124B. 2003\_CON\_10\_CD  
20 pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 125A and 125B. Fig. 125A.

2003\_CON\_11\_CPX pol.PEP. Fig. 125B.

2003\_CON\_11\_CPX pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 126A and 126B. Fig. 126A.

2003\_CON\_12\_BF pol.PEP. Fig. 126B. 2003\_CON\_12\_BF  
5 pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 127A and 127B. Fig. 127A.

2003\_CON\_14\_BG pol.PEP. Fig. 127B. 2003\_CON\_14\_BG  
pol.OPT. (OPT = codon optimized encoding sequence.)

#### DETAILED DESCRIPTION OF THE INVENTION

10 The present invention relates to an immunogen that induces antibodies that neutralize a wide spectrum of human immunodeficiency virus (HIV) primary isolates and/or that induces a T cell response. The immunogen comprises at least one  
15 consensus or ancestral immunogen (e.g., Env, Gag, Nef or Pol), or portion or variant thereof. The invention also relates to nucleic acid sequences encoding the consensus or ancestral immunogen, or portion or variant thereof. The invention further  
20 relates to methods of using both the immunogen and the encoding sequences. While the invention is described in detail with reference to specific consensus and ancestral immunogens (for example, to a group M consensus Env), it will be appreciated  
25 that the approach described herein can be used to generate a variety of consensus or ancestral



immunogens (for example, envelopes for other HIV-1 groups (e.g., N and O)).

In accordance with one embodiment of the invention, a consensus *env* gene can be constructed by generating consensus sequences of *env* genes for each subtype of a particular HIV-1 group (group M being classified into subtypes A-D, F-H, J and K), for example, from sequences in the Los Alamos HIV Sequence Database (using, for example, MASE (Multiple Aligned Sequence Editor)). A consensus sequence of all subtype consensus sequences can then be generated to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). In the case of the group M consensus *env* gene described in Example 1 (designated CON6), five highly variable regions from a CRF08\_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a region in cytoplasmic domain of gp41) are used to fill in the missing regions in the sequence (see, however, corresponding regions for Con-S). For high levels of expression, the codons of consensus or ancestral genes can be optimized based on codon usage for highly expressed human genes (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)).

With the Year 1999 consensus group M *env* gene, CON6, it has been possible to demonstrate induction of superior T cell responses by CON6 versus wild-type B and C *env* by the number of ELISPOT  $\gamma$ -interferon spleen spot forming cells and the

number of epitopes recognized in two strains of mice (Tables 1 and 2 show the data in BALB/c mice). The ability of CON6 Env protein to induce neutralizing antibodies to HIV-1 primary isolates has been compared to that of several subtype B Env. The target of neutralizing antibodies induced by CON6 includes several non-B HIV-1 strains.

Table 1. T cell epitope mapping of CON6, JRFL and 96ZM651 Env immunogen in BALB/c mice

| Peptide                    | Immunogen |          |             | T cell response |
|----------------------------|-----------|----------|-------------|-----------------|
|                            | CON6      | JRFL (B) | 96ZM651 (C) |                 |
| CON 6 (group M consensus)  |           |          |             |                 |
| 16 DTEVHNWVWATHACVP        | +         |          | +           | CD4             |
| 48 KNSSEYRLINCNTS          | +         |          | +           | CD4             |
| 49 EYYRLINCNTSAITQ         |           |          |             |                 |
| 53 CPKVSFEPIPIHYCA         | +         |          |             | CD4             |
| 54 SFEPIPIHYCAPAGF         |           |          |             |                 |
| 62 NVSTVQCTHGKIPVV         | +         |          |             | CD4             |
| 104 ETITLPCRKQIINM         | +         |          |             | CD8             |
| 105 LPCRIKQIINMWQGV        |           |          |             |                 |
| 130 GIVQQSNLLRAIEA         | +         |          |             | CD4             |
| 131 VQQSNLLRAIEAQQHL       |           |          |             |                 |
| 134 AQQHLLQLTVWGIKQLQ      | +         |          |             | CD4             |
| 135 LQLTVWGIKQLQARVL       |           |          |             |                 |
| Subtype B (MN)             |           |          |             |                 |
| 6223 AKAYDTEVHNWVWATQ      | +         |          |             | CD4             |
| 6224 DTEVHNWVWATQACVP      |           |          |             |                 |
| 6261 ACPKISFEPIPIHYC       | +         |          |             | CD4             |
| 6262 ISFEPIPIHYCAPAG       |           |          |             |                 |
| 6286 RKRIHIGPGRAFYTIT      |           | +        |             | CD8             |
| 6287 HIGPGRAFYTITKNII      |           |          |             |                 |
| 6346 IVQQSNLLRAIEAQ        | +         |          |             | CD4             |
| 6347 QNNLLRAIEAQQHML       |           |          |             |                 |
| Subtype C (Chn19)          |           |          |             |                 |
| 4834 VPVWKEAKTTLFCASDAKSY  |           |          | +           | CD4             |
| 4836 GKEVHNWVWATHACVPTDNP  | +         |          | +           | CD4             |
| 4846 SSENSEYRLINCNTSAIT    | +         |          | +           | CD4             |
| 4854 STVQCTHGKIPVVSTQLLN   | +         |          |             | CD4             |
| 4884 QQSNNLLRAIEAQQHLLQLTV | +         |          |             | CD4             |
| 4885 AQQHLLQLTVWGIKQLQTRV  | +         |          |             | CD4             |

Table 2. T cell epitope mapping of CON6.gp120 immunogen in C57BL/6 mice

| Peptide            | Peptide sequence      | T cell response |
|--------------------|-----------------------|-----------------|
| CON 6 (consensus)  |                       |                 |
| 2                  | GIQRNCQHLWRWGT        | CD8             |
| 3                  | NCQHLWRWGTMI LGM      |                 |
| 16                 | DTEVHNVWATHACVP       | CD4             |
| 53                 | CPKVSFEPIPIHYCA       | CD4             |
| 97                 | FYCNTSGLFNSTW MF      | CD8             |
| 99                 | FNSTW MFNGTYMFNG      | CD8             |
| Subtype B (MN)     |                       |                 |
| 6210               | GIRRN YQHWWGWGT       | CD8             |
| 6211               | NYQHWWGWGTMLLGL       |                 |
| 6232               | NMWKNNMVEQM HEDI      | CD4             |
| 6262               | ISFEPIPIHYCAPAG       | CD4             |
| 6290               | NIIGTIRQAHCNISR       | CD4             |
| 6291               | TIRQAHCNISRAKWN       |                 |
| Subtype C (Chn 19) |                       |                 |
| 4830               | MRVTGIRKN YQHLWRWGTML | CD8             |
| 5446               | RWGTMLLGLMLICSA AEN   | CD8             |
| 4836               | GKEVHNVWATHACVPTDPNP  | CD4             |
| 4862               | GDIRQAHCNISKDKWNETLQ  | CD4             |
| 4888               | LLGIWGC SGKLICTTTPWN  | CD8             |

For the Year 2000 consensus group M env gene, 5' Con-S, the Con-S envelope has been shown to be as immunogenic as the CON6 envelope gene in T cell  $\gamma$  interferon ELISPOT assays in two strains of mice

(the data for C57BL/6 are shown in Fig. 27).  
Furthermore, in comparing CON6 and Con-S gp140 Envs  
as protein immunogens for antibody in guinea pigs  
(Table 3), both gp140 Envs were found to induce  
5 antibodies that neutralized subtype B primary  
isolates. However, Con-S gp140 also induced robust  
neutralization of the subtype C isolates TV-1 and DU  
123 as well as one subtype A HIV-1 primary isolate,  
while CON6 did not.

TABLE 3 Ability of Group M Consensus CON6 and Con-S Envs to Induce Neutralization of HIV-1 Primary Isolates

| HIV-1 Isolate<br>(Subtype) | CON6 gp140CF |     |     |     |     | CON6 gp140 CFI |      |     |      |      | CONS gp140 CFI |      |      |      |      |      |      |      |      |      |      |      |
|----------------------------|--------------|-----|-----|-----|-----|----------------|------|-----|------|------|----------------|------|------|------|------|------|------|------|------|------|------|------|
|                            | 770          | 771 | 772 | 775 | 781 | 783            | 784  | 786 | 776  | 777  | 778            | 780  | 781  | 782  | 783  | 784  | 786  | 776  | 777  | 778  | 780  |      |
| BX08(B)                    | 520          | 257 | 428 | 189 | 218 | 164            | >540 | 199 | >540 | >540 | >540           | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 |
| QH0692 (B)                 | 46           | 55  | 58  | 77  | <20 | 91             | 100  | 76  | 109  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| SS1196(B)                  | 398          | 306 | 284 | 222 | 431 | 242            | >540 | 351 | >540 | >540 | >540           | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 | >540 |
| JRLFL(B)                   | <20          | <20 | <20 | <20 | <20 | 169            | <20  | <20 | <20  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| BG1168(B)                  | <20          | <20 | <20 | <20 | <20 | <20            | <20  | <20 | <20  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| 3988(B)                    | <20          | <20 | <20 | <20 | <20 | <20            | <20  | <20 | <20  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| 6101(B)                    | <20          | <20 | <20 | <20 | <20 | <20            | <20  | <20 | <20  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| TV-1(C)                    | <20          | <20 | <20 | <20 | <20 | <20            | <20  | <20 | <20  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| DU123(C)                   | <20          | <20 | 71  | 74  | <20 | 72             | <20  | <20 | 176  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| DU172(C)                   | <20          | <20 | 96  | 64  | <20 | <20            | <20  | <20 | <20  | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |
| ZM18108.6(C)               | ND           | ND  | ND  | ND  | <20 | <20            | <20  | <20 | 84   | <20  | <20            | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  | <20  |



As the next iteration of consensus immunogens, and in recognition of the fact that a practical HIV-1 immunogen can be a polyvalent mixture of either  
5 several subtype consensus genes, a mixture of subtype and consensus genes, or a mixture of centralized genes and wild type genes, a series of 11 subtype consensus, and wild type genes have been designed from subtypes A, B, C, CRF AE01, and G as  
10 well as a group M consensus gene from Year 2003 Los Alamos National Database sequences. The wild type sequences were chosen either because they were known to come from early transmitted HIV-1 strains (those strains most likely to be necessary to be protected  
15 against by a vaccine) or because they were the most recently submitted strains in the database of that subtype. These nucleotide and amino acid sequences are shown in Figures 28-38 (for all 140CF designs shown, 140CF gene can be flanked with the 5'  
20 sequence "TTCAGTCGACGGCCACC" that contains a Kozak sequence (GCCACCATGG/A) and *Sa*II site and 3' sequence of TAAAGATCTTACAA containing stop codon and *Bgl*II site). Shown in Figures 39-62 are 2003 centralized (consensus and ancestral) HIV-1 envelope  
25 proteins and the codon optimized gene sequences.

Major differences between CON6 gp140 (which does not neutralize non-clade B HIV strains) and Con-S gp140 (which does induce antibodies that neutralize non-clade B HIV strains) are in Con-S V1,  
30 V2, V4 and V5 regions. For clade B strains, peptides of the V3 region can induce neutralizing

antibodies (Haynes et al, J. Immunol. 151:1646-1653 (1993)). Thus, construction of Th-V1, Th-V2, Th-V4, Th-V5 peptides can be expected to give rise to the desired broadly reactive anti-non-clade B neutralizing antibodies. Therefore, the Th-V peptides set forth in Table 4 are contemplated for use as a peptide immunogen(s) derived from Con-S gp140. The gag Th determinant (GTH, Table 4) or any homologous GTH sequence in other HIV strains, can be used to promote immunogenicity and the C4 region of HIV gp120 can be used as well (KQIINMWQVVGKAMYA) or any homologous C4 sequence from other HIV strains (Haynes et al, J. Immunol. 151:1646-1653 (1993)). Con-S V1, V2, V4, V5 peptides with an N-terminal helper determinant can be used singly or together, when formulated in a suitable adjuvant such as Corixa's RC529 (Baldrige et al, J. Endotoxin Res. 8:453-458 (2002)), to induce broadly cross reactive neutralizing antibodies to non-clade B isolates.

20



| Table 4 |                      |  |
|---------|----------------------|--|
|         |                      |  |
| 1)      | GTH Con-S V1 132-150 | YKRWII LGLNKIVRMYTNNVNTNTTNNTEEKGEIKN                    |
| 2)      | GTH Con-S V2 157-189 | YKRWII LGLNKIVRMYTEIRDKKQKVYALFYRLDVPIDNNNNSSNYR         |
| 3)      | GTH Con-S V3 294-315 | YKRWII LGLNKIVRMYTRPNNNTRKSIRIGPGQAFYAT                  |
| 4)      | GTH Con-S V4 381-408 | YKRWII LGLNKIVRMYNTSGLFNSTWIGNGTKNNNNTNDTITLP            |
| 5)      | GTH Con-S V5 447-466 | YKRWII LGLNKIVRMYRDGGNNNTNETEIFRPGGGD                    |
| 6)      | GTH Con-6 V1 132-150 | YKRWII LGLNKIVRMYNVRNVSSNGTETDNEEIKN                     |
| 7)      | GTH Con-6 V2 157-196 | YKRWII LGLNKIVRMYTELDRKKQKVYALFYRLDVPIDDKNSSEISGKNSSEYYR |
| 8)      | GTH-Con6 V3 301-322  | YKRWII LGLNKIVRMYTRPNNNTRKSIHIGPGQAFYAT                  |
| 9)      | GTH Con-6 V4 388-418 | YKRWII LGLNKIVRMYNTSGLFNSTWMFNGTYMFNGTKDNSETITLP         |
| 10)     | GTH Con 6 V5 457-477 | YKRWII LGLNKIVRMYRDGGNNSNKNKTETFRPGGGD                   |

It will be appreciated that the invention includes portions and variants of the sequences specifically disclosed herein. For example, forms of codon optimized consensus encoding sequences can be constructed as gp140CF, gp140 CFI, gp120 or gp160 forms with either gp120/41 cleaved or uncleaved. For example, and as regards the consensus and ancestral envelope sequences, the invention encompasses envelope sequences devoid of V3. Alternatively, V3 sequences can be selected from preferred sequences, for example, those described in U.S. Application No. 10/431,596 and U.S. Provisional Application No. 60/471,327. In addition, an optimal immunogen for breadth of response can include mixtures of group M consensus *gag*, *pol*, *nef* and *env* encoding sequences, and as well as consist of

mixtures of subtype consensus or ancestral encoding sequences for *gag*, *pol*, *nef* and *env* HIV genes. For dealing with regional differences in virus strains, an efficacious mixture can include mixtures of  
5 consensus/ancestral and wild type encoding sequences.

A consensus or ancestral envelope of the invention can be been "activated" to expose intermediate conformations of neutralization  
10 epitopes that normally are only transiently or less well exposed on the surface of the HIV virion. The immunogen can be a "frozen" triggered form of a consensus or ancestral envelope that makes available specific epitopes for presentation to B lymphocytes.  
15 The result of this epitope presentation is the production of antibodies that broadly neutralize HIV. (Attention is directed to WO 02/024149 and to the activated/triggered envelopes described therein.)

20 The concept of a fusion intermediate immunogen is consistent with observations that the gp41 HR-2 region peptide, DP178, can capture an uncoiled conformation of gp41 (Furata et al, Nature Struct. Biol. 5:276 (1998)), and that formalin-fixed HIV-  
25 infected cells can generate broadly neutralizing antibodies (LaCasse et al, Science 283:357 (1997)). Recently a monoclonal antibody against the coiled-coil region bound to a conformational determinant of gp41 in HR1 and HR2 regions of the coiled-coil gp41  
30 structure, but did not neutralize HIV (Jiang et al, J. Virol. 10213 (1998)). However, this latter study

proved that the coiled-coil region is available for antibody to bind if the correct antibody is generated.

The immunogen of one aspect of the invention  
5 comprises a consensus or ancestral envelope either  
in soluble form or anchored, for example, in cell  
vesicles or in liposomes containing translipid  
bilayer envelope. To make a more native envelope,  
gp140 or gp160 consensus or ancestral sequences can  
10 be configured in lipid bilayers for native trimeric  
envelope formation. Alternatively, triggered gp160  
in aldrithio 1-2 inactivated HIV-1 virions can be  
used as an immunogen. The gp160 can also exist as a  
recombinant protein either as gp160 or gp140 (gp140  
15 is gp160 with the transmembrane region and possibly  
other gp41 regions deleted). Bound to gp160 or  
gp140 can be recombinant CCR5 or CXCR4 co-receptor  
proteins (or their extracellular domain peptide or  
protein fragments) or antibodies or other ligands  
20 that bind to the CXCR4 or CCR5 binding site on  
gp120, and/or soluble CD4, or antibodies or other  
ligands that mimic the binding actions of CD4.  
Alternatively, vesicles or liposomes containing CD4,  
CCR5 (or CXCR4), or soluble CD4 and peptides  
25 reflective of CCR5 or CXCR4 gp120 binding sites.  
Alternatively, an optimal CCR5 peptide ligand can be  
a peptide from the N-terminus of CCR5 wherein  
specific tyrosines are sulfated (Bormier et al,  
Proc. Natl. Acad. Sci. USA 97:5762 (2001)). The  
30 triggered immunogen may not need to be bound to a  
membrane but may exist and be triggered in solution.

Alternatively, soluble CD4 (sCD4) can be replaced by an envelope (gp140 or gp160) triggered by CD4 peptide mimetopes (Vitra et al, Proc. Natl. Acad. Sci. USA 96:1301 (1999)). Other HIV co-receptor molecules that "trigger" the gp160 or gp140 to undergo changes associated with a structure of gp160 that induces cell fusion can also be used. Ligation of soluble HIV gp140 primary isolate HIV 89.6 envelope with soluble CD4 (sCD4) induced conformational changes in gp41.

In one embodiment, the invention relates to an immunogen that has the characteristics of a receptor (CD4)-ligated consensus or ancestral envelope with CCR5 binding region exposed but unlike CD4-ligated proteins that have the CD4 binding site blocked, this immunogen has the CD4 binding site exposed (open). Moreover, this immunogen can be devoid of host CD4, which avoids the production of potentially harmful anti-CD4 antibodies upon administration to a host.

The immunogen can comprise consensus or ancestral envelope ligated with a ligand that binds to a site on gp120 recognized by an A32 monoclonal antibodies (mab) (Wyatt et al, J. Virol. 69:5723 (1995), Boots et al, AIDS Res. Hum. Retro. 13:1549 (1997), Moore et al, J. Virol. 68:8350 (1994), Sullivan et al, J. Virol. 72:4694 (1998), Fouts et al, J. Virol. 71:2779 (1997), Ye et al, J. Virol. 74:11955 (2000)). One A32 mab has been shown to mimic CD4 and when bound to gp120, upregulates (exposes) the CCR5 binding site (Wyatt et al, J.

Virol. 69:5723 (1995)). Ligation of gp120 with such a ligand also upregulates the CD4 binding site and does not block CD4 binding to gp120.

Advantageously, such ligands also upregulate the HR-  
5 2 binding site of gp41 bound to cleaved gp120, uncleaved gp140 and cleaved gp41, thereby further exposing HR-2 binding sites on these proteins - each of which are potential targets for anti-HIV neutralizing antibodies.

10 In a specific aspect of this embodiment, the immunogen comprises soluble HIV consensus or ancestral gp120 envelope ligated with either an intact A32 mab, a Fab2 fragment of an A32 mab, or a Fab fragment of an A32 mab, with the result that the  
15 CD4 binding site, the CCR5 binding site and the HR-2 binding site on the consensus or ancestral envelope are exposed/upregulated. The immunogen can comprise consensus or ancestral envelope with an A32 mab (or fragment thereof) bound or can comprise consensus or  
20 ancestral envelope with an A32 mab (or fragment thereof) bound and cross-linked with a cross-linker such as .3% formaldehyde or a heterobifunctional cross-linker such as DTSSP (Pierce Chemical Company). The immunogen can also comprise uncleaved  
25 consensus or ancestral gp140 or a mixture of uncleaved gp140, cleaved gp41 and cleaved gp120. An A32 mab (or fragment thereof) bound to consensus or ancestral gp140 and/or gp120 or to gp120 non-covalently bound to gp41, results in upregulation  
30 (exposure) of HR-2 binding sites in gp41, gp120 and uncleaved gp140. Binding of an A32 mab (or fragment

thereof) to gp120 or gp140 also results in upregulation of the CD4 binding site and the CCR5 binding site. As with gp120 containing complexes, complexes comprising uncleaved gp140 and an A32 mab (or fragment thereof) can be used as an immunogen uncross-linked or cross-linked with cross-linker such as .3% formaldehyde or DTSSP. In one embodiment, the invention relates to an immunogen comprising soluble uncleaved consensus or ancestral gp140 bound and cross linked to a Fab fragment or whole A32 mab, optionally bound and cross-linked to an HR-2 binding protein.

The consensus or ancestral envelope protein triggered with a ligand that binds to the A32 mab binding site on gp120 can be administered in combination with at least a second immunogen comprising a second envelope, triggered by a ligand that binds to a site distinct from the A32 mab binding site, such as the CCR5 binding site recognized by mab 17b. The 17b mab (Kwong et al, Nature 393:648 (1998) available from the AIDS Reference Repository, NIAID, NIH) augments sCD4 binding to gp120. This second immunogen (which can also be used alone or in combination with triggered immunogens other than that described above) can, for example, comprise soluble HIV consensus or ancestral envelope ligated with either the whole 17b mab, a Fab2 fragment of the 17b mab, or a Fab fragment of the 17b mab. It will be appreciated that other CCR5 ligands, including other antibodies (or fragments thereof), that result in the CD4 binding site being

exposed can be used in lieu of the 17b mab. This further immunogen can comprise gp120 with the 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) bound or can comprise gp120 with  
5 the 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) bound and cross-linked with an agent such as .3% formaldehyde or a heterobifunctional cross-linker, such as DTSSP (Pierce Chemical Company). Alternatively, this  
10 further immunogen can comprise uncleaved gp140 present alone or in a mixture of cleaved gp41 and cleaved gp120. Mab 17b, or fragment thereof (or other CCR5 ligand as indicated above) bound to gp140 and/or gp120 in such a mixture results in exposure  
15 of the CD4 binding region. The 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) gp140 complexes can be present uncross-linked or cross-linked with an agent such as .3% formaldehyde or DTSSP.

20 Soluble HR-2 peptides, such as T649Q26L and DP178, can be added to the above-described complexes to stabilize epitopes on consensus gp120 and gp41 as well as uncleaved consensus gp140 molecules, and can be administered either cross-linked or uncross-  
25 linked with the complex.

A series of monoclonal antibodies (mabs) have been made that neutralize many HIV primary isolates, including, in addition to the 17b mab described above, mab IgG1b12 that binds to the CD4 binding  
30 site on gp120 (Roben et al, J. Virol. 68:482 (1994), Mo et al, J. Virol. 71:6869 (1997)), mab 2G12 that

binds to a conformational determinant on gp120 (Trkola et al, J. Virol. 70:1100 (1996)), and mab 2F5 that binds to a membrane proximal region of gp41 (Muster et al, J. Virol. 68:4031 (1994)).

5 As indicated above, various approaches can be used to "freeze" fusogenic epitopes in accordance with the invention. For example, "freezing" can be effected by addition of the DP-178 or T-649Q26L peptides that represent portions of the coiled coil  
10 region, and that when added to CD4-triggered consensus or ancestral envelope, result in prevention of fusion (Rimsky et al, J. Virol. 72:986-993 (1998)). HR-2 peptide bound consensus or ancestral gp120, gp140, gp41 or gp160 can be used as  
15 an immunogen or crosslinked by a reagent such as DTSSP or DSP (Pierce Co.), formaldehyde or other crosslinking agent that has a similar effect.

"Freezing" can also be effected by the addition of 0.1% to 3% formaldehyde or paraformaldehyde, both  
20 protein cross-linking agents, to the complex, to stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160 complex, or to stabilize the "triggered" gp41 molecule, or both (LaCasse et al, Science 283:357-362 (1999)).

25 Further, "freezing" of consensus or ancestral gp41 or gp120 fusion intermediates can be effected by addition of heterobifunctional agents such as DSP (dithiobis[succinimidylpropionate]) (Pierce Co. Rockford, ILL., No. 22585ZZ) or the water soluble  
30 DTSSP (Pierce Co.) that use two NHS esters that are reactive with amino groups to cross link and



stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160 complex, or to stabilize the "triggered" gp41 molecule, or both.

Analysis of T cell immune responses in  
5 immunized or vaccinated animals and humans shows that the envelope protein is normally not a main target for T cell immune response although it is the only gene that induces neutralizing antibodies. HIV-1 Gag, Pol and Nef proteins induce a potent T  
10 cell immune response. Accordingly, the invention includes a repertoire of consensus or ancestral immunogens that can induce both humoral and cellular immune responses. Subunits of consensus or  
ancestral sequences can be used as T or B cell  
15 immunogens. (See Examples 6 and 7, and Figures referenced therein, and Figures 63-127.

The immunogen of the invention can be formulated with a pharmaceutically acceptable carrier and/or adjuvant (such as alum) using  
20 techniques well known in the art. Suitable routes of administration of the present immunogen include systemic (e.g. intramuscular or subcutaneous). Alternative routes can be used when an immune response is sought in a mucosal immune system (e.g.,  
25 intranasal).

The immunogens of the invention can be chemically synthesized and purified using methods which are well known to the ordinarily skilled artisan. The immunogens can also be synthesized by  
30 well-known recombinant DNA techniques. Nucleic acids encoding the immunogens of the invention can

be used as components of, for example, a DNA vaccine wherein the encoding sequence is administered as naked DNA or, for example, a minigene encoding the immunogen can be present in a viral vector. The  
5 encoding sequence can be present, for example, in a replicating or non-replicating adenoviral vector, an adeno-associated virus vector, an attenuated mycobacterium tuberculosis vector, a Bacillus Calmette Guerin (BCG) vector, a vaccinia or Modified  
10 Vaccinia Ankara (MVA) vector, another pox virus vector, recombinant polio and other enteric virus vector, Salmonella species bacterial vector, Shigella species bacterial vector, Venezuelan Equine Encephalitis Virus (VEE) vector, a Semliki  
15 Forest Virus vector, or a Tobacco Mosaic Virus vector. The encoding sequence, can also be expressed as a DNA plasmid with, for example, an active promoter such as a CMV promoter. Other live vectors can also be used to express the sequences of  
20 the invention. Expression of the immunogen of the invention can be induced in a patient's own cells, by introduction into those cells of nucleic acids that encode the immunogen, preferably using codons and promoters that optimize expression in human  
25 cells. Examples of methods of making and using DNA vaccines are disclosed in U.S. Pat. Nos. 5,580,859, 5,589,466, and 5,703,055.

The composition of the invention comprises an immunologically effective amount of the immunogen of  
30 this invention, or nucleic acid sequence encoding same, in a pharmaceutically acceptable delivery

system. The compositions can be used for prevention and/or treatment of immunodeficiency virus infection. The compositions of the invention can be formulated using adjuvants, emulsifiers,  
5 pharmaceutically-acceptable carriers or other ingredients routinely provided in vaccine compositions. Optimum formulations can be readily designed by one of ordinary skill in the art and can include formulations for immediate release and/or  
10 for sustained release, and for induction of systemic immunity and/or induction of localized mucosal immunity (e.g, the formulation can be designed for intranasal administration). The present compositions can be administered by any convenient  
15 route including subcutaneous, intranasal, oral, intramuscular, or other parenteral or enteral route. The immunogens can be administered as a single dose or multiple doses. Optimum immunization schedules can be readily determined by the ordinarily skilled  
20 artisan and can vary with the patient, the composition and the effect sought.

The invention contemplates the direct use of both the immunogen of the invention and/or nucleic acids encoding same and/or the immunogen expressed  
25 as minigenes in the vectors indicated above. For example, a minigene encoding the immunogen can be used as a prime and/or boost.

The invention includes any and all amino acid sequences disclosed herein and, where applicable, CF  
30 and CFI forms thereof, as well as nucleic acid

sequences encoding same (and nucleic acids complementary to such encoding sequences).

Certain aspects of the invention can be described in greater detail in the non-limiting  
5 Examples that follows.

#### EXAMPLE 1

#### Artificial HIV-1 Group M Consensus Envelope

#### EXPERIMENTAL DETAILS

10        *Expression of CON6 gp120 and gp140 proteins in recombinant vaccinia viruses (VV).* To express and purify the secreted form of HIV-1 CON6 envelope proteins, CON6 gp120 and gp140CF plasmids were constructed by introducing stop codons after the  
15 gp120 cleavage site (REKR) and before the transmembrane domain (YIKIFIMIVGGLIGLRIVFAVLSIVN), respectively. The gp120/gp41 cleavage site and fusion domain of gp41 were deleted in the gp140CF protein. Both CON6 gp120 and gp140CF DNA constructs  
20 were cloned into the pSC65 vector (from Bernard Moss, NIH, Bethesda, MD) at SalI and KpnI restriction enzyme sites. This vector contains the lacZ gene that is controlled by the p7.5 promoter. A back-to-back P E/L promoter was used to express  
25 CON6 env genes. BSC-1 cells were seeded at  $2 \times 10^5$  in each well in a 6-well plate, infected with wild-type vaccinia virus (WR) at a MOI of 0.1 pfu/cell, and 2 hr after infection, pSC65-derived plasmids

containing CON6 *env* genes were transfected into the VV-infected cells and recombinant (r) VV selected as described (Moss and Earl, Current Protocols in Molecular Biology, eds, Ausubel et al (John Wiley & Sons, Inc. Indianapolis, IN) pp. 16.15.1-16.19.9 (1998)). Recombinant VV that contained the CON6 *env* genes were confirmed by PCR and sequencing analysis. Expression of the CON6 envelope proteins was confirmed by SDS-PAGE and Western blot assay.

10 Recombinant CON6 gp120 and gp140CF were purified with agarose *galanthus Nivalis* lectin beads (Vector Labs, Burlingame, CA), and stored at -70°C until use. Recombinant VV expressing JRFL (vCB-28) or 96ZM651 (vT241R) gp160 were obtained from the NIH AIDS

15 Research and Reference Reagent Program (Bethesda, MD).

*Monoclonal Antibodies and gp120 Wild-type Envelopes.* Human mabs against a conformational

20 determinant on gp120 (A32), the gp120 V3 loop (F39F) and the CCR5 binding site (17b) were the gifts of James Robinson (Tulane Medical School, New Orleans, LA) (Wyatt et al, Nature 393;705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). Mabs 2F5,

25 447, b12, 2G12 and soluble CD4 were obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD) (Gorny et al, J. Immunol. 159:5114-5122 (1997), Nyambi et al, J. Virol. 70:6235-6243 (1996), Purtscher et al, AIDS Res. Hum. Retroviruses

30 10:1651-1658 (1994), Trkola et al, J. Virol 70:1100-1108 (1996)). T8 is a murine mab that maps to the

gp120 C1 region (a gift from P. Earl, NIH, Bethesda, MD). BaL (subtype B), 96ZM651 (subtype C), and 93TH975 (subtype E) gp120s were provided by QBI, Inc. and the Division of AIDS, NIH. CHO cell lines  
5 that express 92U037 (subtype A) and 93BR029 (subtype F) gp140 (secreted and uncleaved) were obtained from NICBS, England.

*Surface Plasmon Resonance Biosensor (SPR)*

10 *Measurements and ELISA.* SPR biosensor measurements were determined on a BIAcore 3000 instrument (BIAcore Inc., Uppsala, Sweden) instrument and data analysis was performed using BIAevaluation 3.0 software (BIAcore Inc, Upsaala, Sweden). Anti-gp120  
15 mabs (T8, A32, 17b, 2G12) or sCD4 in 10mM Na-acetate buffer, pH 4.5 were directly immobilized to a CM5 sensor chip using a standard amine coupling protocol for protein immobilization. FPLC purified CON6 gp120 monomer or gp140CF oligomer recombinant  
20 proteins were flowed over CM5 sensor chips at concentrations of 100 and 300  $\mu\text{g/ml}$ , respectively. A blank in-line reference surface (activated and deactivated for amine coupling) or non-bonding mab controls were used to subtract non-specific or bulk  
25 responses. Soluble 89.6 gp120 and irrelevant IgG was used as a positive and negative control respectively and to ensure activity of each mab surface prior to injecting the CON6 Env proteins. Binding of CON6 envelope proteins was monitored in  
30 real-time at 25°C with a continuous flow of PBS (150 mM NaCl, 0.005% surfactant P20), pH 7.4 at 10-30

$\mu$ l/min. Bound proteins were removed and the sensor surfaces were regenerated following each cycle of binding by single or duplicate 5-10  $\mu$ l pulses of regeneration solution (10 mM glycine-HCl, pH 2.9).

5 ELISA was performed to determine the reactivity of various mabs to CON6 gp120 and gp140CF proteins as described (Haynes et al, AIDS Res. Hum. Retroviruses 11:211-221 (1995)). For assay of human mab binding to rgp120 or gp140 proteins, end-point titers were

10 defined as the highest titer of mab (beginning at 20  $\mu$ g/ml) at which the mab bound CON6 gp120 and gp140CF Env proteins  $\geq$  3 fold over background control (non-binding human mab).

15 *Infectivity and coreceptor usage assays.* HIV-1/SG3 $\Delta$ env and CON6 or control env plasmids were cotransfected into human 293T cells. Pseudotyped viruses were harvested, filtered and p24 concentration was quantitated (DuPont/NEN Life

20 Sciences, Boston, MA). Equal amounts of p24 (5 ng) for each pseudovirion were used to infect JC53-BL cells to determine the infectivity (Derdeyn e al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). JC53-BL

25 cells express CD4, CCR5 and CXCR4 receptors and contain a  $\beta$ -galactosidase ( $\beta$ -gal) gene stably integrated under the transcriptional control of an HIV-1 long terminal repeat (LTR). These cells can be used to quantify the infectious titers of

30 pseudovirion stocks by staining for  $\beta$ -gal expression

and counting the number of blue cells (infectious units) per microgram of p24 of pseudovirions (IU/ $\mu$ g p24) (Derdeyn et al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). To determine the coreceptor usage of the CON6 env gene, JC53BL cells were treated with 1.2  $\mu$ M AMD3100 and 4  $\mu$ M TAK-799 for 1 hr at 37°C then infected with equal amounts of p24 (5 ng) of each Env pseudotyped virus. The blockage efficiency was expressed as the percentage of the infectious units from blockage experiments compared to that from control culture without blocking agents. The infectivity from control group (no blocking agent) was arbitrarily set as 100%.

15

*Immunizations.* All animals were housed in the Duke University Animal Facility under AALAC guidelines with animal use protocols approved by the Duke University Animal Use and Care Committee. Recombinant CON6 gp120 and gp140CF glycoproteins were formulated in a stable emulsion with RIBI-CWS adjuvant based on the protocol provided by the manufacturer (Sigma Chemical Co., St. Louis, MO). For induction of anti-envelope antibodies, each of four out-bred guinea pigs (Harlan Sprague, Inc., Chicago, IL) was given 100  $\mu$ g either purified CON6 gp120 or gp140CF subcutaneously every 3 weeks (total of 5 immunizations). Serum samples were heat-inactivated (56°C, 1 hr), and stored at -20°C until use.

30



For induction of anti-envelope T cell responses, 6-8 wk old female BALB/c mice (Frederick Cancer Research and Developmental Center, NCI, Frederick, MD) were immunized i.m. in the quadriceps  
5 with 50  $\mu$ g plasmid DNA three times at a 3-week interval. Three weeks after the last DNA immunization, mice were boosted with  $10^7$  PFU of rVV expressing Env proteins. Two weeks after the boost, all mice were euthanized and spleens were removed  
10 for isolation of splenocytes.

*Neutralization assays.* Neutralization assays were performed using either a MT-2 assay as described in Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000), a luciferase-based  
15 multiple replication cycle HIV-1 infectivity assay in 5.25.GFP.Luc.M7 cells using a panel of HIV-1 primary isolates (Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000), Bures et al, J.  
20 Virol. 76:2233-2244 (2002)), or a syncytium (fusion from without) inhibition assay using inactivated HIV-1 virions (Rossio et al, J. Virol. 72:7992-8001 (1998)). In the luciferase-based assay, neutralizing antibodies were measured as a function  
25 of a reduction in luciferase activity in 5.25.EGFP.Luc.M7 cells provided by Nathaniel R. Landau, Salk Institute, La Jolla, CA (Brandt et al, J. Biol. Chem. 277:17291-17299 (2002)). Five hundred tissue culture infectious dose 50 (TCID<sub>50</sub>) of  
30 cell-free virus was incubated with indicated serum

dilutions in 150  $\mu$ l (1 hr, at 37°C) in triplicate in 96-well flat-bottom culture plates. The 5.25.EGFP.Luc.M7 cells were suspended at a density of  $5 \times 10^5$ /ml in media containing DEAE dextran (10  $\mu$ g/ml). Cells (100  $\mu$ l) were added and until 10% of cells in control wells (no test serum sample) were positive for GFP expression by fluorescence microscopy. At this time the cells were concentrated 2-fold by removing one-half volume of media. A 50  $\mu$ l suspension of cells was transferred to 96-well white solid plates (Costar, Cambridge, MA) for measurement of luciferase activity using Bright-Glo™ substrate (Promega, Madison, WI) on a Wallac 1420 Multilabel Counter (PerkinElmer Life Sciences, Boston, MA). Neutralization titers in the MT-2 and luciferase assays were those where  $\geq 50\%$  virus infection was inhibited. Only values that titered beyond 1:20 (i.e.  $>1:30$ ) were considered significantly positive. The syncytium inhibition "fusion from without" assay utilized HIV-1 aldrithiol-2 (AT-2) inactivated virions from HIV-1 subtype B strains ADA and AD8 (the gift of Larry Arthur and Jeffrey Lifson, Frederick Research Cancer Facility, Frederick, MD) added to SupT1 cells, with syncytium inhibition titers determined as those titers where  $\geq 90\%$  of syncytia were inhibited compared to prebleed sera.

*Enzyme linked immune spot (ELISPOT) assay.*

Single-cell suspensions of splenocytes from

individual immunized mice were prepared by mincing and forcing through a 70  $\mu$ m Nylon cell strainer (BD Labware, Franklin Lakes, NJ). Overlapping Env peptides of CON6 gp140 (159 peptides, 15mers  
5 overlapping by 11) were purchased from Boston Bioscience, Inc (Royal Oak, MI). Overlapping Env peptides of MN gp140 (subtype B; 170 peptides, 15mers overlapping by 11) and Chn19 gp140 (subtype C; 69 peptides, 20mers overlapping by 10) were  
10 obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD). Splenocytes (5 mice/group) from each mouse were stimulated *in vitro* with overlapping Env peptides pools from CON6, subtype B and subtype C Env proteins. 96-well PVDF  
15 plates (MultiScreen-IP, Millipore, Billerica, MA) were coated with anti-IFN- $\gamma$  mab (5  $\mu$ g/ml, AN18; Mabtech, Stockholm, Sweden). After the plates were blocked at 37°C for 2 hr using complete Hepes buffered RPMI medium, 50 $\mu$ l of the pooled overlapping  
20 envelope peptides (13 CON6 and MN pools, 13-14 peptides in each pool; 9 Chn19 pool, 7-8 peptide in each pool) at a final concentration of 5  $\mu$ g/ml of each were added to the plate. Then 50  $\mu$ l of splenocytes at a concentration of  $1.0 \times 10^7$ /ml were  
25 added to the wells in duplicate and incubated for 16 hr at 37°C with 5% CO<sub>2</sub>. The plates were incubated with 100  $\mu$ l of a 1:1000 dilution of streptavidin alkaline phosphatase (Mabtech, Stockholm, Sweden), and purple spots developed using 100  $\mu$ l of BCIP/NBT  
30 (Plus) Alkaline Phosphatase Substrate (Moss,

Pasadena, MD). Spot forming cells (SFC) were measured using an Immunospot counting system (CTL Analyzers, Cleveland, OH). Total responses for each envelope peptide pool are expressed as SFCs per 10<sup>6</sup> splenocytes.

## RESULTS

*CON6 Envelope Gene Design, Construction and Expression.* An artificial group M consensus env gene (CON6) was constructed by generating consensus sequences of env genes for each HIV-1 subtype from sequences in the Los Alamos HIV Sequence Database, and then generating a consensus sequence of all subtype consensuses to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). Five highly variable regions from a CRF08\_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a region in cytoplasmic domain of gp41) were then used to fill in the missing regions in CON6 sequence. The CON6 V3 region is group M consensus (Figure 1A). For high levels of expression, the codons of CON6 env gene were optimized based on codon usage for highly expressed human genes (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)). (See Fig. 1D.) The codon optimized CON6 env gene was constructed and subcloned into pcDNA3.1 DNA at EcoR I and BamH I sites (Gao et al, AIDS Res. Hum. Retroviruses, 19:817-823 (2003)). High levels of protein

expression were confirmed with Western-blot assays after transfection into 293T cells. To obtain recombinant CON6 Env proteins for characterization and use as immunogens, rVV was generated to express  
5 secreted gp120 and uncleaved gp140CF (Figure 1B). Purity for each protein was  $\geq 90\%$  as determined by Coomassie blue gels under reducing conditions (Figure 1C).

10 *CD4 Binding Domain and Other Wild-type HIV-1 Epitopes are Preserved on CON6 Proteins.* To determine if CON6 proteins can bind to CD4 and express other wild-type HIV-1 epitopes, the ability of CON6 gp120 and gp140CF to bind soluble(s) CD4, to  
15 bind several well-characterized anti-gp120 mabs, and to undergo CD4-induced conformational changes was assayed. First, BIAcore CM5 sensor chips were coated with either sCD4 or mabs to monitor their binding activity to CON6 Env proteins. It was found  
20 that both monomeric CON6 gp120 and oligomeric gp140CF efficiently bound sCD4 and anti-gp120 mabs T8, 2G12 and A32, but did not constitutively bind mab 17b, that recognizes a CD4 inducible epitope in the CCR5 binding site of gp120 (Figures 2A and 2B).  
25 Both sCD4 and A32 can expose the 17b binding epitope after binding to wild-type gp120 (Wyatt et al, Nature 393;705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). To determine if the 17b epitope could be induced on CON6 Envs by either sCD4  
30 or A32, sCD4, A32 and T8 were coated on sensor chips, then CON6 gp120 or gp140CF captured, and mab

17b binding activity monitored. After binding sCD4 or mab A32, both CON6 gp120 and gp140CF were triggered to undergo conformational changes and bound mab 17b (Figures 2C and 2D). In contrast, after binding mab T8, the 17b epitope was not exposed (Figures 2C and 2D). ELISA was next used to determine the reactivity of a panel of human mabs against the gp120 V3 loop (447, F39F), the CD4 binding site (b12), and the gp41 neutralizing determinant (2F5) to CON6 gp120 and gp140CF (Figure 2E). Both CON6 rgp120 and rgp140CF proteins bound well to neutralizing V3 mabs 447 and F39F and to the potent neutralizing CD4 binding site mab b12. Mab 2F5, that neutralizes HIV-1 primary isolates by binding to a C-terminal gp41 epitope, also bound well to CON6 gp140CF (Figure 2E).

*CON6 env Gene is Biologically Functional and Uses CCR5 as its Coreceptor.* To determine whether CON6 envelope gene is biologically functional, it was co-transfected with the env-defective SG3 proviral clone into 293T cells. The pseudotyped viruses were harvested and JC53BL cells infected. Blue cells were detected in JC53-BL cells infected with the CON6 Env pseudovirions, suggesting that CON6 Env protein is biologically functional (Figure 3A). However, the infectious titers were 1-2 logs lower than that of pseudovirions with either YU2 or NL4-3 wild-type HIV-1 envelopes.

The co-receptor usage for the CON6 env gene was next determined. When treated with CXCR4 blocking

agent AMD3100, the infectivity of NL4-3 Env-pseudovirions was blocked while the infectivity of YU2 or CON6 Env-pseudovirions was not inhibited (Figure 3B). In contrast, when treated with CCR5 blocking agent TAK-779, the infectivity of NL4-3 Env-pseudovirions was not affected, while the infectivity of YU2 or CON6 Env-pseudovirions was inhibited. When treated with both blocking agents, the infectivity of all pseudovirions was inhibited. Taken together, these data show that the CON6 envelope uses the CCR5 co-receptor for its entry into target cells.

*Reaction of CON6 gp120 With Different Subtype Sera.* To determine if multiple subtype linear epitopes are preserved on CON6 gp120, a recombinant Env protein panel (gp120 and gp140) was generated. Equal amounts of each Env protein (100 ng) were loaded on SDS-polyacrylamide gels, transferred to nitrocellulose, and reacted with subtype A through G patient sera as well as anti-CON6 gp120 guinea pig sera (1:1,000 dilution) in Western blot assays. For each HIV-1 subtype, four to six patient sera were tested. One serum representative for each subtype is shown in Figure 4.

It was found that whereas all subtype sera tested showed variable reactivities among Envs in the panel, all group M subtype patient sera reacted equally well with CON6 gp120 Env protein, demonstrating that wild-type HIV-1 Env epitopes recognized by patient sera were well preserved on

the CON6 Env protein. A test was next made as to whether CON6 gp120 antiserum raised in guinea pigs could react to different subtype Env proteins. It was found that the CON6 serum reacted to its own and  
5 other subtype Env proteins equally well, with the exception of subtype A Env protein (Figure 4).

*Induction of T Cell Responses to CON6, Subtype B and Subtype C Envelope Overlapping Peptides.* To  
10 compare T cell immune responses induced by CON6 Env immunogens with those induced by subtype specific immunogens, two additional groups of mice were immunized with subtype B or subtype C DNAs and with corresponding rVV expressing subtype B or C envelope  
15 proteins. Mice immunized with subtype B (JRFL) or subtype C (96ZM651) Env immunogen had primarily subtype-specific T cell immune responses (Figure 5). IFN- $\gamma$  SFCs from mice immunized with JRFL (subtype B) immunogen were detected after stimulation with  
20 subtype B (MN) peptide pools, but not with either subtype C (Chn19) or CON6 peptide pools. IFN- $\gamma$  SFCs from mice immunized with 96ZM651 (subtype C) immunogen were detected after the stimulation with both subtype C (Chn19) and CON6 peptide pools, but  
25 not with subtype B (MN) peptide pools. In contrast, IFN- $\gamma$  SFCs were identified from mice immunized with CON6 Env immunogens when stimulated with either CON6 peptide pools as well as by subtype B or C peptide pools (Figure 5). The T cell immune responses  
30 induced by CON6 gp140 appeared more robust than



those induced by CON6 gp120. Taken together, these data demonstrated that CON6 gp120 and gp140CF immunogens were capable of inducing T cell responses that recognized T cell epitopes of wild-type subtype B and C envelopes.

*Induction of Antibodies by Recombinant CON6 gp120 and gp140CF Envelopes that Neutralize HIV-1 Subtype B and C Primary Isolates.* To determine if the CON6 envelope immunogens can induce antibodies that neutralize HIV-1 primary isolates, guinea pigs were immunized with either CON6 gp120 or gp140CF protein. Sera collected after 4 or 5 immunizations were used for neutralization assays and compared to the corresponding prebleed sera. Two AT-2 inactivated HIV-1 isolates (ADA and AD8) were tested in syncytium inhibition assays (Table 5A). Two subtype B SHIV isolates, eight subtype B primary isolates, four subtype C, and one each subtype A, D, and E primary isolates were tested in either the MT-2 or the luciferase-based assay (Table 5B). In the syncytium inhibition assay, it was found that antibodies induced by both CON 6 gp120 and gp140CF proteins strongly inhibited AT-2 inactivated ADA and AD8-induced syncytia (Table 5A). In the MT-2 assay, weak neutralization of 1 of 2 SHIV isolates (SHIV SF162P3) by two gp120 and one gp140CF sera was found (Table 5B). In the luciferase-based assay, strong neutralization of 4 of 8 subtype B primary isolates (BX08, SF162, SS1196, and BAL) by all gp120 and gp140CF sera was found, and weak neutralization of 2

of 8 subtype B isolates (6101, 0692) by most gp120 and gp140CF sera was found. No neutralization was detected against HIV-1 PAVO (Table 5B). Next, the CON6 anti-gp120 and gp140CF sera were tested against 5 four subtype C HIV-1 isolates, and weak neutralization of 3 of 4 isolates (DU179, DU368, and S080) was found, primarily by anti-CON6 gp120 sera. One gp140CF serum, no. 653, strongly neutralized DU179 and weakly neutralized S080 (Table 5B). 10 Finally, anti-CON6 Env sera strongly neutralized a subtype D isolate (93ZR001), weakly neutralized a subtype E (CM244) isolate, and did not neutralize a subtype A (92RW020) isolate.

Table 5A

**Ability of HIV-1 Group M Consensus Envelope CON6 Proteins to Induce Fusion Inhibiting Antibodies**

| Guinea Pig No.       | Immunogen | Syncytium Inhibition antibody titer <sup>1</sup> |     |
|----------------------|-----------|--|-----|
|                      |           | AD8  | ADA |
| 646                  | gp120     | 270  | 270 |
| 647                  | gp120     | 90   | 90  |
| 648                  | gp120     | 90   | 270 |
| 649                  | gp120     | 90   | 90  |
| Geometric Mean Titer |           | 119  | 156 |
| 650                  | gp140     | 270  | 270 |
| 651                  | gp140     | 90   | 90  |
| 652                  | gp140     | ≥10  | 810 |
| 653                  | gp140     | 270  | 90  |
| Geometric Mean Titer |           | 270  | 207 |

15

<sup>1</sup>Reciprocal serum dilution at which HIV-induced syncytia of Sup T1 cells was inhibited by >90% compared to pre-immune serum. All prebleed sera were negative (titer <10).

**Table 5B**  
**Ability of Group M Consensus HIV-1 Envelope CON6 gp120 and gp140CF Proteins**  
**to Induce Antibodies that Neutralize HIV Primary Isolates**

| HIV Isolate<br>(Subtype) | CON6 gp120 Protein<br>Guinea Pig No. |     |     |     |     | CON6 gp140CF Protein<br>Guinea Pig No. |        |     |       |                       | Controls |            |  |
|--------------------------|--------------------------------------|-----|-----|-----|-----|--|--------|-----|-------|-----------------------|----------|------------|--|
|                          | 646                                  | 647 | 648 | 649 | 650 | 651                                    | 652    | 653 | GMT   | TriMab <sub>2</sub> # | CD4-IgG2 | HIV+ Serum |  |
| SHIV 89.6P*(B)           | <20                                  | <20 | <20 | <20 | <20 | <20                                    | <20    | <20 | <20   | NT                    | NT       | NT         |  |
| SHIV SF162P3*(B)         | <20                                  | 30  | 48  | <20 | 27  | <20                                    | <20    | <20 | <20   | NT                    | 0.2µg/ml | NT         |  |
| BX08(B)                  | 270                                  | 183 | 254 | 55  | 199 | 64                                     | 229    | 150 | 187   | 0.7µg/ml              | NT       | 236#       |  |
| 6101(B)                  | <20                                  | 38  | 35  | <20 | <20 | 90                                     | 72     | 73  | 39    | 1.1µg/ml              | NT       | NT         |  |
| BG1168(B)                | <20                                  | <20 | <20 | <20 | 40  | <20                                    | <20    | 25  | <20   | 2.7µg/ml              | NT       | NT         |  |
| 0692(B)                  | 31                                   | 32  | 34  | <20 | 28  | 33                                     | 30     | 45  | 33    | 0.8µg/ml              | NT       | 769        |  |
| PAVO(B)                  | <20                                  | <20 | <20 | <20 | <20 | <20                                    | <20    | <20 | <20   | 2.9µg/ml              | NT       | NT         |  |
| SF162(B)                 | 2,146                                | 308 | 110 | 282 | 206 | 5,502                                  | 15,098 | 174 | 1,313 | NT                    | NT       | >540       |  |
| SS1196(B)                | 206                                  | 26  | 148 | 59  | 381 | 401                                    | 333    | 81  | 253   | NT                    | NT       | 301#       |  |
| BAL(B)                   | 123                                  | 90  | 107 | 138 | 107 | 146                                    | 136    | 85  | 116   | NT                    | NT       | 3307       |  |
| 92RW020(A)               | <20                                  | <20 | <20 | <20 | <20 | <20                                    | <20    | <20 | <20   | NT                    | NT       | 693        |  |
| DUI79(C)                 | <20                                  | 43  | <20 | 24  | <20 | <20                                    | 24     | 515 | 33    | NT                    | 0.8µg/ml | NT         |  |
| DU368(C)                 | 25                                   | 35  | 62  | <20 | <20 | <20                                    | <20    | 23  | <20   | NT                    | 2.3µg/ml | NT         |  |
| S021(C)                  | <20                                  | <20 | 33  | <20 | <20 | <20                                    | <20    | <20 | <20   | NT                    | 8.3µg/ml | NT         |  |
| S080(C)                  | 24                                   | 37  | 70  | 41  | <20 | <20                                    | <20    | 52  | <20   | NT                    | 3.4µg/ml | NT         |  |
| 93ZR001(D)               | 275                                  | 144 | 126 | 114 | 306 | 195                                    | 129    | 173 | 191   | NT                    | NT       | 693        |  |
| CM244(E)                 | 35                                   | 43  | 64  | ND  | 31  | 25                                     | 27     | 25  | 26    | NT                    | NT       | 693        |  |

\*MT-2 Assay; All other HIV isolates were tested in the M7-luciferase assay.

HIV-1 isolates QH0692, SS1196, SF162, 6101, BX08, BG1168, BAL were assayed with post-injection 5 serum; other HIV-1 isolates were assayed with post-injection 4 serum. ND = not done.

HIV+ sera was either HIV-1+ human serum (LEH3) or an anti-gp120 guinea pig serum (#) with known neutralizing activity for HIV-1 isolate SS1196. GMT = geometric mean titer of four animals per group. Neutralizing titers reported are after subtraction of any background neutralization in prebleed sera.

#TriMab<sub>2</sub> = a mixture of human mabs 2F5, b12, 2G12.

CONCLUSIONS

The production of an artificial HIV-1 Group M consensus *env* genes (encoding sequences) (CON6 and  
5 Con-S) have been described that encodes a functional Env protein that is capable of utilizing the CCR5 co-receptor for mediating viral entry. Importantly, these Group M consensus envelope genes could induce T and B cell responses that recognized epitopes of  
10 subtype B and C HIV-1 primary isolates. In addition, Con-S induces antibodies that strongly neutralize Subtype-C and A HIV-1 strains (see Table 3).

The correlates of protection to HIV-1 are not  
15 conclusively known. Considerable data from animal models and studies in HIV-1-infected patients suggest the goal of HIV-1 vaccine development should be the induction of broadly-reactive CD4+ and CD8+ anti-HIV-1 T cell responses (Letvin et al, Annu.  
20 Rev. Immunol. 20:73-99 (2002)) and high levels of antibodies that neutralize HIV-1 primary isolates of multiple subtypes (Mascola et al, J. Virol. 73:4009-4018 (1999), Mascola et al, Nat. Med. 6:270-210 (2000)).

25 The high level of genetic variability of HIV-1 has made it difficult to design immunogens capable of inducing immune responses of sufficient breadth to be clinically useful. Epitope based vaccines for T and B cell responses (McMichael et al, Vaccine  
30 20:1918-1921 (2002), Sbai et al, Curr. Drug Targets Infect, Disord. 1:303-313 (2001), Haynes, Lancet

348:933-937 (1996)), constrained envelopes reflective of fusion intermediates (Fouts et al, Proc. Natl. Acad. Sci. USA 99:11842-22847 (2002)), as well as exposure of conserved high-order  
5 structures for induction of anti-HIV-1 neutralizing antibodies have been proposed to overcome HIV-1 variability (Roben et al, J. Virol. 68:4821-4828 (1994), Sapphire et al, Science 293:1155-1159 (2001)). However, with the ever-increasing  
10 diversity and rapid evolution of HIV-1, the virus is a rapidly moving complex target, and the extent of complexity of HIV-1 variation makes all of these approaches problematic. The current most common approach to HIV-1 immunogen design is to choose a  
15 wild-type field HIV-1 isolate that may or may not be from the region in which the vaccine is to be tested. Polyvalent envelope immunogens have been designed incorporating multiple envelope immunogens (Bartlett et al, AIDS 12:1291-1300 (1998), Cho et  
20 al, J. Virol. 75:2224-2234 (2001)).

The above-described study tests a new strategy for HIV-1 immunogen design by generating a group M consensus *env* gene (CON6) with decreased genetic distance between this candidate immunogen and wild-  
25 type field virus strains. The CON6 *env* gene was generated for all subtypes by choosing the most common amino acids at most positions (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). Since only the most common  
30 amino acids were used, the majority of antibody and T cell epitopes were well preserved. Importantly,

the genetic distances between the group M consensus env sequence and any subtype env sequences was about 15%, which is only half of that between wild-type subtypes (30%) (Gaschen et al, Science 296:2354-2360 (2002)). This distance is approximately the same as that among viruses within the same subtype. Further, the group M consensus env gene was also about 15% divergent from any recombinant viral env gene, as well, since CRFs do not increase the overall genetic divergence among subtypes.

Infectivity of CON6-Env pseudovirions was confirmed using a single-round infection system, although the infectivity was compromised, indicating the artificial envelope was not in an "optimal" functional conformation, but yet was able to mediate virus entry. That the CON6 envelope used CCR5 (R5) as its coreceptor is important, since majority of HIV-1 infected patients are initially infected with R5 viruses.

BIAcore analysis showed that both CON6 gp120 and gp140CF bound sCD4 and a number of mabs that bind to wild-type HIV-1 Env proteins. The expression of the CON6 gp120 and 140CF proteins that are similar antigenically to wild-type HIV-1 envelopes is an important step in HIV-1 immunogen development. However, many wild-type envelope proteins express the epitopes to which potent neutralizing human mabs bind, yet when used as immunogens themselves, do not induce broadly neutralizing anti-HIV-1 antibodies of the specificity of the neutralizing human mabs.

The neutralizing antibody studies were encouraging in that both CON6 gp120, CON6 gp140CF and Con-S gp140CFI induced antibodies that neutralized select subtype B, C and D HIV-1 primary isolates, with Con-S gp140CFI inducing the most robust neutralization of non-subtype B primary HIV isolates. However, it is clear that the most difficult-to-neutralize primary isolates (PAVO, 6101, BG1168, 92RW020, CM244) were either only weakly or not neutralized by anti-CON6 gp120 or gp140 sera (Table 4b). Nonetheless, the Con-S envelope immunogenicity for induction of neutralizing antibodies is promising, given the breadth of responses generated with the Con-S subunit gp140CFI envelope protein for non-subtype B HIV isolates. Previous studies with poxvirus constructs expressing gp120 and gp160 have not generated high levels of neutralizing antibodies (Evans et al, J. Infect. Dis. 180:290-298 (1999), Polacino et al, J. Virol. 73:618-630 (1999), Ourmanov et al, J. Virol. 74:2960-2965 (2000), Pal et al, J. Virol 76:292-302 (2002), Excler and Plotkin, AIDS 11(Suppl A):S127-137 (1997). rVV expressing secreted CON6 gp120 and gp140 have been constructed and antibodies that neutralize HIV-1 primary isolates induced. An HIV neutralizing antibody immunogen can be a combination of Con-S gp140CFI, or subunit thereof, with immunogens that neutralize most subtype B isolates.

The structure of an oligomeric gp140 protein is critical when evaluating protein immunogenicity. In this regard, study of purified CON6 gp140CF proteins by fast performance liquid chromatography (FPLC) and analytical ultracentrifugation has demonstrated that the purified gp140 peak consists predominantly of trimers with a small component of dimers.

Thus, centralized envelopes such as CON6, Con-S or 2003 group M or subtype consensus or ancestral encoding sequences described herein, are attractive candidates for preparation of various potentially "enhanced" envelope immunogens including CD4-Env complexes, constrained envelope structures, and trimeric oligomeric forms. The ability of CON6-induced T and B cell responses to protect against HIV-1 infection and/or disease in SHIV challenge models will be studied in non-human primates.

The above study has demonstrated that artificial centralized HIV-1 genes such as group M consensus *env* gene (CON6) and Con-S can also induce T cell responses to T cell epitopes in wild-type subtype B and C Env proteins as well as to those on group M consensus Env proteins (Figure 5). While the DNA prime and rVV boost regimen with CON6 gp140CF immunogen clearly induced IFN- $\gamma$  producing T cells that recognized subtype B and C epitopes, further studies are needed to determine if centralized sequences such as are found in the CON6 envelope are significantly better at inducing cross-



clade T cell responses than wild-type HIV-1 genes  
(Ferrari et al, Proc. Natl. Acad. Sci. USA 94:1396-  
1401 (1997), Ferrari et al, AIDS Res. Hum.  
Retroviruses 16:1433-1443 (2000)). However, the  
5 fact that CON6 (and Con-S; env encoding sequence)  
prime and boosted splenocyte T cells recognized HIV-  
1 subtype B and C T cell epitopes is an important  
step in demonstration that CON6 (and Con-S) can  
induce T cell responses that might be clinically  
10 useful.

Three computer models (consensus, ancestor and  
center of the tree (COT)) have been proposed to  
generate centralized HIV-1 genes (Gaschen et al,  
Science 296:2354-2360 (2002), Gao et al, Science  
15 299:1517-1518 (2003), Nickle et al, Science  
299:1515-1517 (2003), Korber et al, Science  
288:1789-1796 (2000). They all tend to locate at  
the roots of the star-like phylogenetic trees for  
most HIV-1 sequences within or between subtypes. As  
20 experimental vaccines, they all can reduce the  
genetic distances between immunogens and field virus  
strains. However, consensus, ancestral and COT  
sequences each have advantages and disadvantages  
(Gaschen et al, Science 296:2354-2360 (2002), Gao et  
25 al, Science 299:1517-1518 (2003), Nickle et al,  
Science 299:1515-1517 (2003). Consensus and COT  
represent the sequences or epitopes in sampled  
current wild-type viruses and are less affected by  
outliers HIV-1 sequences, while ancestor represents  
30 ancestral sequences that can be significantly  
affected by outlier sequences. However, at present,

it is not known which centralized sequence can serve as the best immunogen to elicit broad immune responses against diverse HIV-1 strains, and studies are in progress to test these different strategies.

5 Taken together, the data have shown that the HIV-1 artificial CON6 and Con-S envelope can induce T cell responses to wild-type HIV-1 epitopes, and can induce antibodies that neutralize HIV-1 primary isolates, thus demonstrating the feasibility and  
10 promise of using artificial centralized HIV-1 sequences in HIV-1 vaccine design.

#### EXAMPLE 2

HIV-1 Subtype C Ancestral and Consensus Envelope  
15 Glycoproteins

#### EXPERIMENTAL DETAILS

HIV-1 subtype C ancestral and consensus *env* genes were obtained from the Los Alamos HIV Molecular Immunology Database (<http://hiv-web.lanl.gov/immunology>), codon-usage optimized for  
20 mammalian cell expression, and synthesized (Fig. 6). To ensure optimal expression, a Kozak sequence (GCCGCCGCC) was inserted immediately upstream of the initiation codon. In addition to the full-length  
25 genes, two truncated *env* genes were generated by introducing stop codons immediately after the gp41 membrane-spanning domain (IVNR) and the gp120/gp41 cleavage site (REKR), generating gp140 and gp120 form of the glycoproteins, respectively (Fig. 8).

Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. To determine if the ancestral and consensus subtype C envelopes were capable of mediating fusion and entry, *gp160* and *gp140* genes were co-transfected with an HIV-1/SG3 $\Delta$ *env* provirus and the resulting pseudovirions tested for infectivity using the JC53-BL cell assay (Fig. 7). Co-receptor usage and envelope neutralization sensitivity were also determined with slight modifications of the JC53-BL assay. Codon-usage optimized and rev-dependent 96ZAM651 *env* genes were used as contemporary subtype C controls.

## RESULTS

15

Codon-optimized subtype C ancestral and consensus envelope genes (*gp160*, *gp140*, *gp120*) express high levels of *env* glycoprotein in mammalian cells (Fig. 9).

20

Codon-optimized subtype C *gp160* and *gp140* glycoproteins are efficiently incorporated into virus particles. Western Blot analysis of sucrose-purified pseudovirions reveals ten-fold higher levels of virion incorporation of the codon-optimized envelopes compared to that of a rev-dependent contemporary envelope controls (Fig. 10A).

25

Virions pseudotyped with either the subtype C consensus *gp160* or *gp140* envelope were more infectious than pseudovirions containing the corresponding *gp160* and *gp140* ancestral envelopes.

30

Additionally, *gp160* envelopes were consistently more infectious than their respective *gp140* counterparts (Fig. 10B).

Both subtype C ancestral and consensus  
5 envelopes utilize CCR5 as a co-receptor to mediate virus entry (Fig. 11).

The infectivity of subtype C ancestral and consensus *gp160* containing pseudovirions was neutralized by plasma from subtype C infected  
10 patients. This suggests that these artificial envelopes possess a structure that is similar to that of native HIV-1 *env* glycoproteins and that common neutralization epitopes are conserved. No significant differences in neutralization potential  
15 were noted between subtype C ancestral and consensus *env* glycoproteins (*gp160*) (Fig. 12).

#### CONCLUSIONS

HIV-1 subtype C viruses are among the most prevalent circulating isolates, representing  
20 approximately fifty percent of new infections worldwide. Genetic diversity among globally circulating HIV-1 strains poses a challenge for vaccine design. Although HIV-1 Env protein is highly variable, it can induce both humoral and cellular  
25 immune responses in the infected host. By analyzing 70 HIV-1 complete subtype C *env* sequences, consensus and ancestral subtype C *env* genes have been generated. Both sequences are roughly equidistant from contemporary subtype C strains and thus

expected to induce better cross-protective immunity.  
A reconstructed ancestral or consensus sequence  
derived-immunogen minimizes the extent of genetic  
differences between the vaccine candidate and  
5 contemporary isolates. However, consensus and  
ancestral subtype C *env* genes differ by 5% amino  
acid sequences. Both consensus and ancestral  
sequences have been synthesized for analyses.  
Codon-optimized subtype C ancestral and consensus  
10 envelope genes have been constructed and the *in*  
*vitro* biological properties of the expressed  
glycoproteins determined. Synthetic subtype C  
consensus and ancestral *env* genes express  
glycoproteins that are similar in their structure,  
15 function and antigenicity to contemporary subtype C  
wild-type envelope glycoproteins.

### EXAMPLE 3

Codon-Usage Optimization of Consensus of Subtype C  
20 *gag* and *nef* Genes (C.con.gag and C.con.nef)

Subtype C viruses have become the most  
prevalent viruses among all subtypes of Group M  
viruses in the world. More than 50% of HIV-1  
25 infected people are currently carrying HIV-1 subtype  
C viruses. In addition, there is considerable  
intra-subtype C variability: different subtype C  
viruses can differ by as much as 10%, 6%, 17% and

16% of their Gag, Pol, Env and Nef proteins, respectively. Most importantly, the subtype C viruses from one country can vary as much as the viruses isolated from other parts of the world. The only exceptions are HIV-1 strains from India/China, Brazil and Ethiopia/Djibouti where subtype C appears to have been introduced more recently. Due to the high genetic variability of subtype C viruses even within a single country, an immunogen based on a single virus isolate may not elicit protective immunity against other isolates circulating in the same area.

Thus *gag* and *nef* gene sequences of subtype C viruses were gathered to generate consensus sequences for both genes by using a 50% consensus threshold. To avoid a potential bias toward founder viruses, only one sequence was used from India/China, Brazil and Ethiopia/Djibouti, respectively, to generate the subtype C consensus sequences (C.con.gag and C.con.nef). The codons of both C.con.gag and C.con.nef genes were optimized based on the codon usage of highly expressed human genes. The protein expression following transfection into 293T cells is shown in Figure 13. As can be seen, both consensus subtype C Gag and Nef proteins were expressed efficiently and recognized by Gag- and Nef-specific antibodies. The protein expression levels of both C.con.gag and C.con.nef genes are comparable to that of native subtype *env* gene (96ZM651).

EXAMPLE 4

Synthesis of a Full Length "Consensus of the  
Consensus env Gene with Consensus Variable Regions"  
5 (CON-S)

In the synthesized "consensus of the consensus"  
env gene (CON6), the variable regions were replaced  
with the corresponding regions from a contemporary  
10 subtype C virus (98CN006). A further con/con gene  
has been designed that also has consensus variable  
regions (CON-s). The codons of the Con-S env gene  
were optimized based on the codon usage of highly  
expressed human genes. (See Figs. 14A and 14B for  
15 amino acid sequences and nucleic acid sequences,  
respectfully.)

Paired oligonucleotides (80-mers) which overlap  
by 20 bp at their 3' ends and contain invariant  
sequences at their 5' and 3' ends, including the  
20 restriction enzyme sites EcoRI and BbsI as well as  
BsmBI and BamHI, respectively, were designed. BbsI  
and BamHI are Type II restriction enzymes that  
cleave outside of their recognition sequences. They  
have been positioned in the oligomers in such a way  
25 that they cleave the first four residues adjacent to  
the 18 bp invariant region, leaving 4 base 5'  
overhangs at the end of each fragment for the  
following ligation step. 26 paired oligomers were  
linked individually using PCR and primers  
30 complimentary to the 18 bp invariant sequences.

Each pair was cloned into pGEM-T (Promega) using the T/A cloning method and sequenced to confirm the absence of inadvertent mutations/deletions. pGEM-T subclones containing the proper inserts were then  
5 digested, run on a 1% agarose gel, and gel purified (Qiagen). Four individual 108-mers were ligated into pcDNA3.1 (Invitrogen) in a multi-fragment ligation reaction. The four-way ligations occurred among groups of fragments in a stepwise manner from  
10 the 5' to the 3' end of the gene. This process was repeated until the entire gene was reconstructed in the pcDNA3.1 vector.

A complete Con-S gene was constructed by ligating the codon usage optimized oligo pairs  
15 together. To confirm its open reading frame, an *in vitro* transcription and translation assay was performed. Protein products were labeled by  $s^{35}$ -methionine during the translation step, separated on a 10% SDS-PAGE, and detected by radioautography.  
20 Expected size of the expressed Con-S gp160 was identified in 4 out of 7 clones (Fig. 14C).

CONS Env protein expression in the mammalian cells after transfected into 293T cells using a Western blot assay (Figure 15). The expression level  
25 of Con-S Env protein is very similar to what was observed from the previous CON6 *env* clone that contains the consensus conservative regions and variable loops from 98CN006 virus isolate.

The Env-pseudovirions was produced by  
30 cotransfecting Con-S *env* clone and *env*-deficient SG3



proviral clone into 293T cells. Two days after transfection, the pseudovirions were harvested and infected into JC53BL-13 cells. The infectious units (IU) were determined by counting the blue cells after staining with X-gal in three independent experiments. When compared with CON6 env clone, Con-S env clones produce similar number of IU in JC53BL-13 cells (Figure 16). The IU titers for both are about 3 log higher than the SG3 backbone clone control (No Env). However, the titers are also about 2 log lower than the positive control (the native HIV-1 env gene, NL4-3 or YU2). These data suggest that both *consensus* group M env clones are biologically functional. Their functionality, however, has been compromised. The functional consensus env genes indicate that these Env proteins fold correctly, preserve the basic conformation of the native Env proteins, and are able to be developed as universal Env immunogens.

It was next determined what coreceptor Con-S Env uses for its entry into JC53-BL cells. When treated with CXCR4 blocking agent AMD3100, the infectivity of NL4-3 Env-pseudovirions was blocked while the infectivity of YU2, Con-S or CON6 Env-pseudovirions was not inhibited. In contrast, when treated with CCR5 blocking agent TAK779, the infectivity of NL4-3 Env-pseudovirions was not affected, while the infectivity of YU2, Con-S or CON6 Env-pseudovirions was inhibited. When treated with both blocking agents, the infectivity of all pseudovirions was inhibited. Taken together, these

data show that the Con-S as well as CON6 envelope uses the CCR5 but not CXCR4 co-receptor for its entry into target cells.

It was next determined whether CON6 or Con-S Env proteins could be equally efficiently incorporated in to the pseudovirions. To be able precisely compare how much Env proteins were incorporated into the pseudovirions, each pseudovirions is loaded on SDS-PAGE at the same concentraion: 5µg total protein for cell lysate, 25ng p24 for cell culture supernatant, or 150ng p24 for purified virus stock (concentrated pseudovirions after super-speed centrifugation). There was no difference in amounts of Env proteins incorporated in CON6 or Con-S Env-pseudovirions in any preparations (cell lysate, cell culture supernatant or purified virus stock) (Figure 17).

#### EXAMPLE 5

Synthesis of a *Consensus* Subtype A Full Length *env*  
(A.con.env) Gene

Subtype A viruses are the second most prevalent HIV-1 in the African continent where over 70% of HIV-1 infections have been documented. Consensus *gag*, *env* and *nef* genes for subtype C viruses that are the most prevalent viruses in Africa and in the world were previously generated. Since genetic distances between subtype A and C viruses are as high as 30% in the *env* gene, the cross reactivity or protection between both subtypes will not be

optimal. Two group M consensus *env* genes for all subtypes were also generated. However, to target any particular subtype viruses, the subtype specific consensus genes will be more effective since the genetic distances between subtype consensus genes and field viruses from the same subtype will be smaller than that between group M consensus genes and these same viruses. Therefore, consensus genes need to be generated for development of subtype A specific immunogens. The codons of the A.con.*env* gene were optimized based on the codon usage of highly expressed human genes. (See Figs. 18A and 18B for amino acid and nucleic acid sequences, respectively.)

Each pair of the oligos has been amplified, cloned, ligated and sequenced. After the open reading frame of the A.con *env* gene was confirmed by an *in vitro* transcription and translation system, the A.con *env* gene was transfected into the 293T cells and the protein expression and specificity confirmed with the Western blot assay (Figure 18). It was then determined whether A.con envelope is biologically functional. It was co-transfected with the *env*-defective SG3 proviral clone into 293T cells. The pseudotyped viruses were harvested and used to infect JC53BL cells. Blue cells were detected in JC53-BL cells infected with the A.con Env-pseudovirions, suggesting that A.con Env protein is biologically functional (Table 6). However, the infectious titer of A.con Env-psuedovirions was about 7-fold lower than that of pseudovirions with

wild-type subtype C envelope (Table 6). Taken together, the biological function A.con Env proteins suggests that it folds correctly and may induce linear and conformational T and B cell epitopes if  
 5 used as an Env immunogen.

**JC53BL13 (IU/ul)**

|              |      | 3/31/03            | 4/7/03          | 4/25/03         |
|--------------|------|--------------------|-----------------|-----------------|
|              |      | non filtered supt. | 0.22µm filtered | 0.22µm filtered |
| A.con        | +SG3 | 4                  | 8.5             | 15.3            |
| 96ZM651      | +SG3 | 87                 | 133             | 104             |
| SG3 backbone |      | 0                  | 0.07            | 0.03            |
| Neg control  |      | 0                  | 0.007           | 0               |

Table 6. Infectivity of pseudovirons with A.con env genes

EXAMPLE 6

10 Design of Full Length "*Consensus of the Consensus gag, pol and nef Genes*" (M.con.gag, M.con.pol and M.con.nef) and a Subtype C *Consensus pol Gene* (C.con.pol)

15 For the group M consensus genes, two different env genes were constructed, one with virus specific variable regions (CON6) and one with consensus variable regions (Con-S). However, analysis of T cell immune responses in immunized or vaccinated  
 20 animals and humans shows that the env gene normally is not a main target for T cell immune response

although it is the only gene that will induce neutralizing antibody. Instead, HIV-1 Gag, Pol and Nef proteins are found to be important for inducing potent T cell immune responses. To generate a repertoire of immunogens that can induce both broader humoral and cellular immune responses for all subtypes, it may be necessary to construct other group M consensus genes other than *env* gene alone. "Consensus of the consensus" *gag*, *pol* and *nef* genes (M.con.gag., M.con.pol and M.con.nef) have been designed. To generate a subtype consensus *pol* gene, the subtype C consensus *pol* gene (C.con.pol) was also designed. The codons of the M.con.gag., M.con.pol, M.con.nef and C.con.pol. genes were optimized based on the codon usage of highly expressed human genes. (See Fig. 19 for nucleic acid and amino acid sequences.)

#### EXAMPLE 7

##### Synthetic Subtype B Consensus *gag* and *env* Genes

#### 20 EXPERIMENTAL DETAILS

Subtype B consensus *gag* and *env* sequences were derived from 37 and 137 contemporary HIV-1 strains, respectively, codon-usage optimized for mammalian cell expression, and synthesized (Figs. 20A and 20B). To ensure optimal expression, a Kozak sequence (GCCGCCGCC) was inserted immediately upstream of the initiation codon. In addition to the full-length *env* gene, a truncated *env* gene was generated by introducing a stop codon immediately

after the gp41 membrane-spanning domain (IVNR) to create a *gp145* gene. Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. (Subtype B  
5 consensus Gag and Env sequences are set forth in Figs. 20C and 20D, respectively.)

To determine if the subtype B consensus envelopes were capable of mediating fusion and entry, *gp160* and *gp145* genes were co-transfected  
10 with an HIV-1/SG3 $\Delta$ env provirus and the resulting pseudovirions were tested for infectivity using the JC53-BL cell assay. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also  
15 contain the reporter cassettes of luciferase and  $\beta$ -galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24-well plates, incubated at 37°C for 24  
20 hours and treated with DEAE-Dextran at 37°C for 30min. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-dextran, and allowed to incubate for 3 hours at 37°C after which an additional 500 $\mu$ L of cell media is added to each  
25 well. Following a final 48-hour incubation at 37°C, cells are fixed, stained using X-Gal, and overlaid with PBS for microscopic counting of blue foci. Counts for mock-infected wells, used to determine background, are subtracted from counts for the  
30 sample wells. Co-receptor usage and envelope

neutralization sensitivity were also determined with slight modifications of the JC53-BL assay.

To determine whether the subtype B consensus Gag protein was capable of producing virus-like particles (VLPs) that incorporated Env glycoproteins, 293T cells were co-transfected with subtype B consensus *gag* and *env* genes. 48-hours post-transfection, cell supernatants containing VLPs were collected, clarified in a tabletop centrifuge, filtered through a 0.2µm filter, and pellet through a 20% sucrose cushion. The VLP pellet was resuspended in PBS and transferred onto a 20-60% continuous sucrose gradient. Following overnight centrifugation at 100,000 x g, 0.5 ml fractions were collected and assayed for p24 content. The refractive index of each fraction was also measured. Fractions with the correct density for VLPs and containing the highest levels of p24 were pooled and pellet a final time. VLP-containing pellets were re-suspended in PBS and loaded on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from a subtype B HIV-1 infected individual.

## RESULTS

25

Codon-usage optimized, subtype B consensus envelope (*gp160*, *gp145*) and *gag* genes express high levels of glycoprotein in mammalian cells (Fig. 21).

Subtype B *gp160* and *gp145* glycoproteins are efficiently incorporated into virus particles.

30

Western Blot analysis of sucrose-purified pseudovirions suggests at least five-fold higher levels of consensus B envelope incorporation compared to incorporation of a rev-dependent contemporary envelope (Fig.23A). Virions pseudotyped with either the subtype B consensus gp160 or gp145 envelope are more infectious than pseudovirions containing a rev-dependent contemporary envelope (Fig. 23 B).

Subtype B consensus envelopes utilize CCR5 as the co-receptor to gain entry into CD4 bearing target cells (Fig. 22).

The infectivity of pseudovirions containing the subtype B consensus gp160 envelope was neutralized by plasma from HIV-1 subtype B infected patients (Fig. 24C) and neutralizing monoclonal antibodies (Fig. 24A). This suggests that the subtype B synthetic consensus B envelopes is similar to native HIV-1 Env glycoproteins in its overall structure and that common neutralization epitopes remain intact. Figs. 24B and 24D show neutralization profiles of a subtype B control envelope (NL4.3 Env).

Subtype B consensus Gag proteins are able to bud from the cell membrane and form virus-like particles (Fig. 25A). Co-transfection of the codon-optimized subtype B consensus *gag* and *gp160* genes produces VLPs with incorporated envelope (Fig. 25B).



CONCLUSIONS

The synthetic subtype B consensus *env* and *gag* genes express viral proteins that are similar in their structure, function and antigenicity to contemporary subtype B Env and Gag proteins. It is contemplated that immunogens based on subtype B consensus genes will elicit CTL and neutralizing immune responses that are protective against a broad set of HIV-1 isolates.

10

\* \* \*

All documents and other information sources cited above are hereby incorporated in their entirety by reference. Also incorporated by reference is Liao et al, J. Virol. 78:5270 (2004)).

WHAT IS CLAIMED IS:

1. An isolated protein comprising the sequence of amino acids set forth in Fig. 1A.
2. A nucleic acid comprising a nucleotide sequence encoding CON6 HIV gp160 protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.
3. The nucleic acid according to claim 2 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 1D.
4. A nucleic acid comprising a nucleotide sequence encoding subtype C ancestral HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.
5. The nucleic acid according to claim 4 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 6A.
6. A nucleic acid comprising a nucleotide sequence encoding subtype C consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.
7. The nucleic acid according to claim 6 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 6B.
8. An isolated protein comprising the sequence of amino acids set forth in Fig. 6C or 6D.

9. A nucleic acid comprising a nucleotide sequence encoding a subtype C consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

10. The nucleic acid according to claim 9 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 13E.

11. A nucleic acid comprising a nucleotide sequence encoding a subtype C consensus HIV nef protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

12. The nucleic acid according to claim 11 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 13F.

13. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

14. The nucleic acid according to claim 13 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 14B.

15. A nucleic acid comprising a nucleotide sequence encoding subtype A consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

16. The nucleic acid according to claim 15 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 18B.

17. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

18. The nucleic acid according to claim 17 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19A .

19. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV pol protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

20. The nucleic acid according to claim 19 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19B.

21. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV nef protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

22. The nucleic acid according to claim 21 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19C.

23. A nucleic acid comprising a nucleotide sequence encoding subtype C consensus HIV pol

protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

24. The nucleic acid according to claim 23 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19D.

25. A nucleic acid comprising a nucleotide sequence encoding subtype B consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

26. The nucleic acid according to claim 25 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 20A.

27. A nucleic acid comprising a nucleotide sequence encoding subtype B consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

28. The nucleic acid according to claim 27 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 20B.

29. An isolated protein comprising the sequence of amino acids set forth in Fig. 20C or 20D.

30. An isolated protein comprising the sequence of amino acids set forth in Fig. 26A .

31. A nucleic acid comprising a nucleotide sequence that encodes the protein according to claim 30.

32. The nucleic acid according to claim 31 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 26B.

33. An isolated protein comprising the sequence of amino acids set forth in Fig. 28B.

34. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 33.

35. The nucleic acid sequence according to claim 34 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 28C.

36. An isolated protein comprising the sequence of amino acids set forth in Fig. 29B.

37. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 36.

38. The nucleic acid sequence according to claim 37 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 29C.

39. An isolated protein comprising the sequence of amino acids set forth in Fig. 30B.

40. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 39.

41. The nucleic acid sequence according to claim 40 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 30C.

42. An isolated protein comprising the sequence of amino acids set forth in Fig. 31B.

43. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 42.

44. The nucleic acid sequence according to claim 43 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 31C.

45. An isolated protein comprising the sequence of amino acids set forth in Fig. 32B.

46. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 45.

47. The nucleic acid sequence according to claim 46 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 32C.

48. An isolated protein comprising the sequence of amino acids set forth in Fig. 33B.

49. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 48.

50. The nucleic acid sequence according to claim 49 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 33C.

51. An isolated protein comprising the sequence of amino acids set forth in Fig. 34B.

52. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 51.

53. The nucleic acid sequence according to claim 52 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 34C.

54. An isolated protein comprising the sequence of amino acids set forth in Fig. 35B.

55. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 54.

56. The nucleic acid sequence according to claim 55 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 35C.

57. An isolated protein comprising the sequence of amino acids set forth in Fig. 36B.

58. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 57.

59. The nucleic acid sequence according to claim 58 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 36C.

60. An isolated protein comprising the sequence of amino acids set forth in Fig. 37B.

61. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 60.



62. The nucleic acid sequence according to claim 61 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 37C.

63. An isolated protein comprising the sequence of amino acids set forth in Fig. 38B.

64. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 63.

65. The nucleic acid sequence according to claim 64 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 38C.

66. An isolated protein comprising a CF or CFI form of the amino acid sequence set forth in any one of Figs. 39A-127A.

67. A nucleic acid comprising the nucleotide sequence set forth in Fig. 39B.

68. A nucleic acid comprising the nucleotide sequence set forth in Fig. 40B.

69. A nucleic acid comprising the nucleotide sequence set forth in Fig. 41B.

70. A nucleic acid comprising the nucleotide sequence set forth in Fig. 42B.

71. A nucleic acid comprising the nucleotide sequence set forth in Fig. 43B.

72. A nucleic acid comprising the nucleotide sequence set forth in Fig. 44B.

73. A nucleic acid comprising the nucleotide sequence set forth in Fig. 45B.

74. A nucleic acid comprising the nucleotide sequence set forth in Fig. 46B.

75. A nucleic acid comprising the nucleotide sequence set forth in Fig. 47B.

76. A nucleic acid comprising the nucleotide sequence set forth in Fig. 48B.

77. A nucleic acid comprising the nucleotide sequence set forth in Fig. 49B.

78. A nucleic acid comprising the nucleotide sequence set forth in Fig. 50B.

79. A nucleic acid comprising the nucleotide sequence set forth in Fig. 51B.

80. A nucleic acid comprising the nucleotide sequence set forth in Fig. 52B.

81. A nucleic acid comprising the nucleotide sequence set forth in Fig. 53B.

82. A nucleic acid comprising the nucleotide sequence set forth in Fig. 54B.

83. A nucleic acid comprising the nucleotide sequence set forth in Fig. 55B.

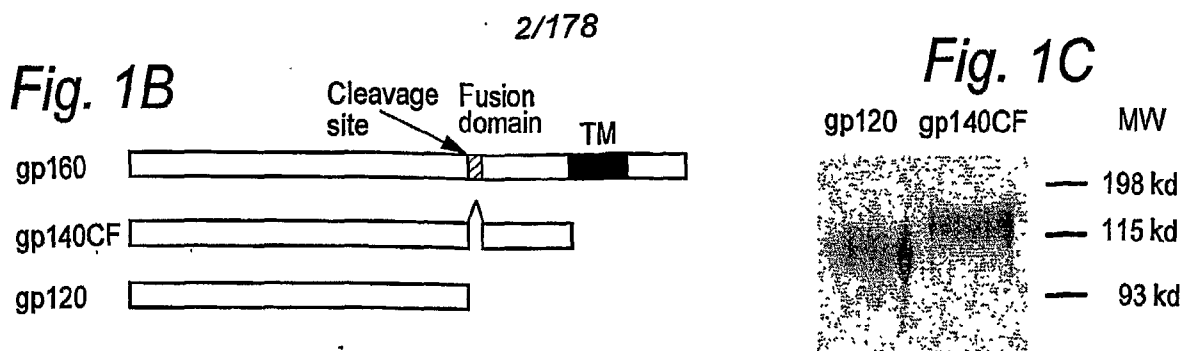
84. A nucleic acid comprising the nucleotide sequence set forth in Fig. 56B.

85. A nucleic acid comprising the nucleotide sequence set forth in Fig. 57B.
86. A nucleic acid comprising the nucleotide sequence set forth in Fig. 58B.
87. A nucleic acid comprising the nucleotide sequence set forth in Fig. 59B.
88. A nucleic acid comprising the nucleotide sequence set forth in Fig. 60B.
89. A nucleic acid comprising the nucleotide sequence set forth in Fig. 61B.
90. A nucleic acid comprising the nucleotide sequence set forth in Fig. 62B.
91. A nucleic acid comprising the nucleotide sequence set forth in any one of Figs. 63B-84B, 65D, 67D and 68D.
92. A nucleic acid comprising the nucleotide sequence set forth in any one of Figs. 85B-106B, 88D, 90D and 92D.
93. A nucleic acid comprising the nucleotide sequence set forth in any one of Figs. 107B-127B, 109D, 111D and 112D.
94. A vector comprising the nucleic acid according to any one of claims 2-7, 9-28, 31, 32, 34, 35, 37, 38, 40, 41, 43, 44, 46, 47, 49, 50, 52, 53, 55, 56, 58, 59, 61, 62, 64, 65 and 67-93.

95. A composition comprising at least one protein or nucleic acid according to any one of claims 1-93 and a carrier.

96. A method of inducing an immune response in a mammal comprising administering to said mammal an amount of at least one protein and/or nucleic acid according to any one of claims 1-93 sufficient to effect said induction.



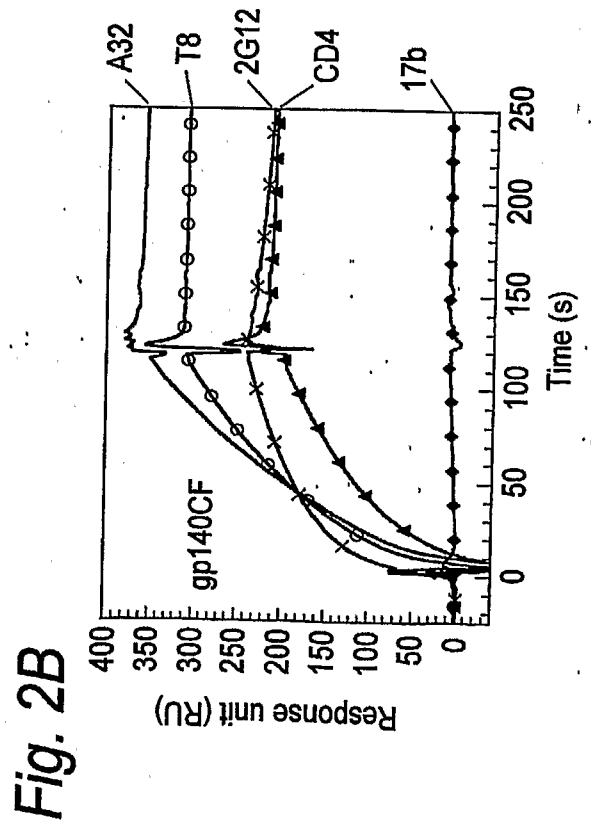
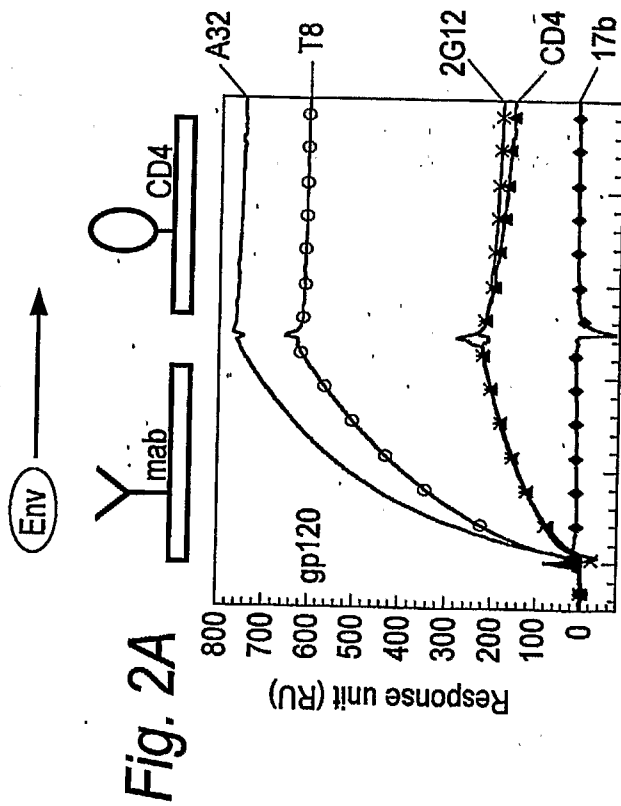
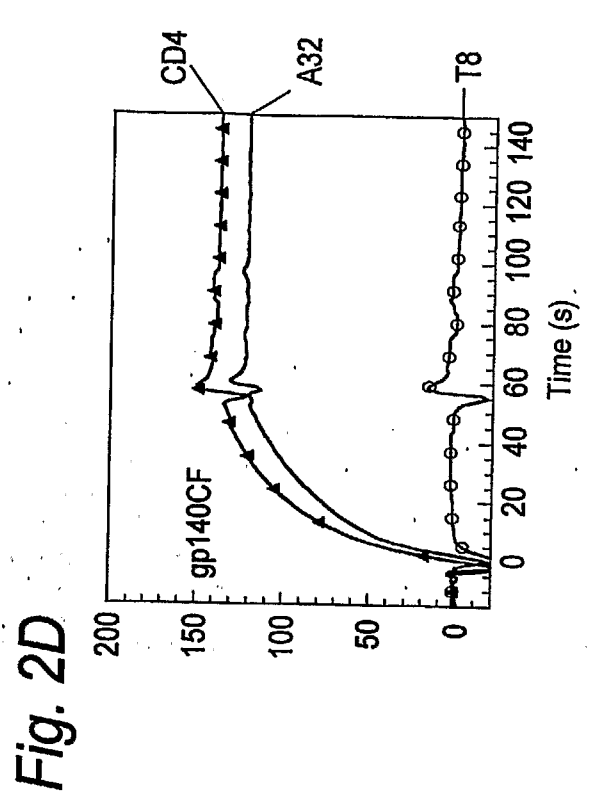
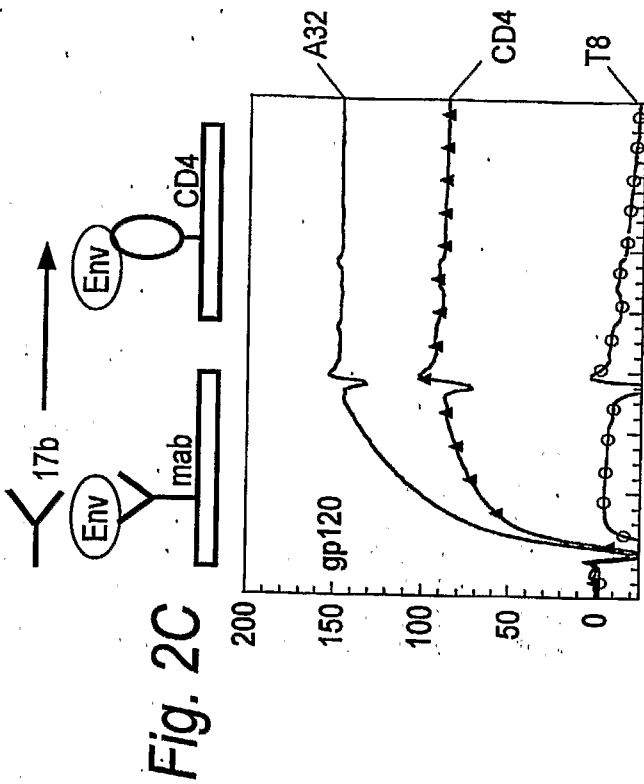
**Fig. 1D**

CON6.env (group M env consensus. This one contain five variable regions in env gene from 98CN006 virus, not in the public domain yet)

```

GCCACCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCATGATC
CTGGGCATGCTGATGATCTGCTCCGCCGCGAGAACCTGTGGGTGACCGTGTACTACGGC
GTGCCCCTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTAC
GACACCGAGGTGCACAACGTGTGGGCCACCCACGCCCTGCGTGCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAATTCAACATGTGGAAGAACAACATGGTG
GACGAGATGCACGAGGACATCATCTCCCTGTGGGACCACTCCCTGAAGCCCTGCGTGAAG
CTGACCCCTGTGCGTGACCCTGAACCTGCACCAACGTGCGCAACGTGTCTCCAACGGC
ACCGAGACCGACAACGAGGAGATCAAGAAGTCTCCTTCAACATCACCACCGAGCTGCGC
GACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGAC
AAGAACTCCTCCGAGATCTCCGGCAAGAAGTCTCCTCCGAGTACTACCGCCTGATCAACTGC
AACACCTCCGCCATCACCAGGCCTGCCCAAGGTGTCCTTCGAGCCCATCCCATCCAC
TACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTCAACGGCACC
GGCCCTGCAAGAAGCTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCC
ACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAAC
ATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGC
ACCCGCCCAACAACAACACCCGCAAGTCCATCCACATCGGCCCCGCGCCAGGCCTTCTAC
GCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCCGCACCAG
TGGAAACAAGACCTGACGAGGTTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACC
ATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGC
GGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATGTTCAAC
GGCACCTACATGTTCAACGGCACCAAGGACAACCTCCGAGACCATCACCTGCCCTGCCGC
ATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCATC
GAGGCAAGATCACCTGCAAGTCCAACATCACCGCCTGCTGCTGACCCGCGACGGCGGC
AACAACTCCAACAAGAACAAGACCGAGACCTTCCGCCCGCGCGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCTGGGCTGGCC
CCCACCAAGGCCAAGCGCCGCGTGGTGGAGCGGAGAAGCGCGCCGTGGGCTCCGGCC
GTGTTCTGGGCTTCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCCCTG
ACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGC
GCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAG
GCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC
TGCTCCGCAAGCTGATCTGCACCACCAACGTGCCCTGGAACCTCCTCCTGGTCCAACAAG
TCCCAGGACGAGATC TGGGACAAATGACCTGGATGGAGTGGGAGCGGAGATCTCCAAC
TACACCGACATCATCTACCGCTGATCGAGGAGTCCAGAACCAGCAGGAGAAGAACGAG
CAGGAGCTGCTGGCCCTGGACAAGTGGCCCTCCCTGTGGAACCTGGTTTCGACATCACCAAC
TGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCCTGATCGGCCCTGCGCATC
GTGTTCCCGTGTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTCTC
CAGACCTGATCCCCAACCCCGCGGCCCGGACCGCCCGAGGGCATCGAGGAGGAGGGC
GGCGAGCAGGGCCGCGACCGCTCCATCCGCCCTGGTGAACGGCTTCTGGCCC'TGGCCTGG
GACGACCTGCGCTCCCTGTGCC'TGTTCTCCTACCACCGCCTGCGCGACTTCACTCTGATC
GCCGCCCGCACCGTGGAGCTGCTGGGCCGCGCTCCCTGCGCGGCCCTGCAGAAGGGCTGG
GAGGCCCTGAAGTACCTGGGCAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCC
GCCATCTCCCTGCTGGACACCACCGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATC
GAGATCGTGCAGCGCGCTGCCCGCCATCCTGAACATCCCCCGCGCATCCGCCAGGGC
CTGGAGCGCGCCCTGCTGTAA

```



4/178

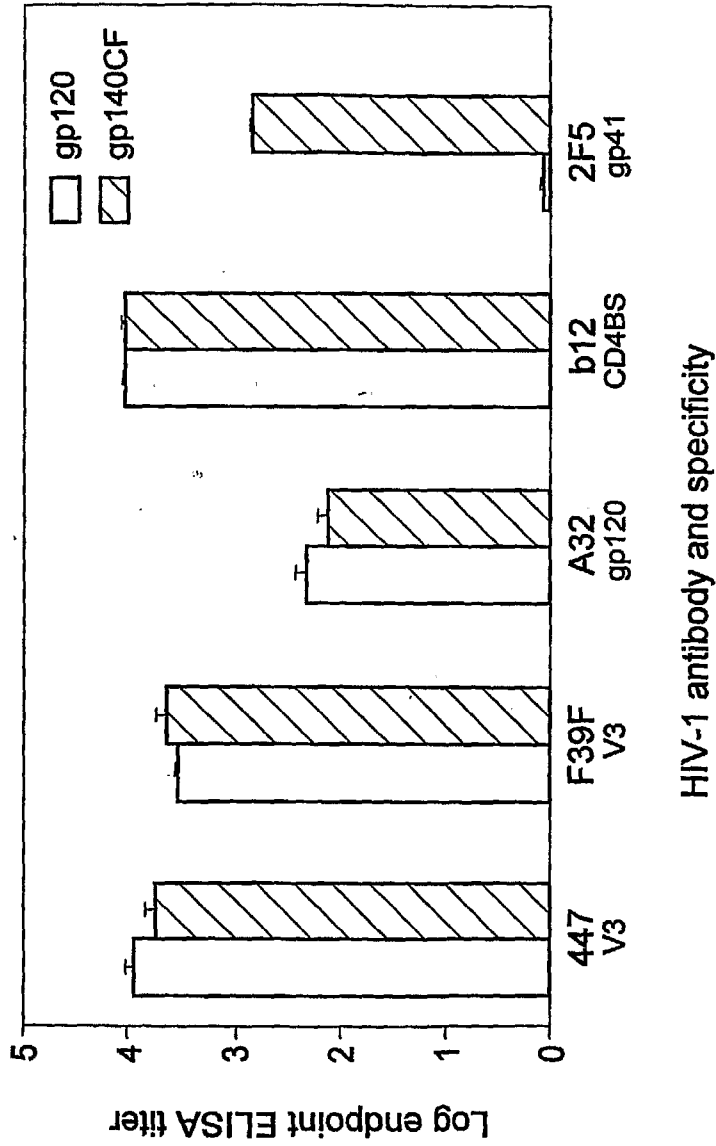


Fig. 2E



5/178

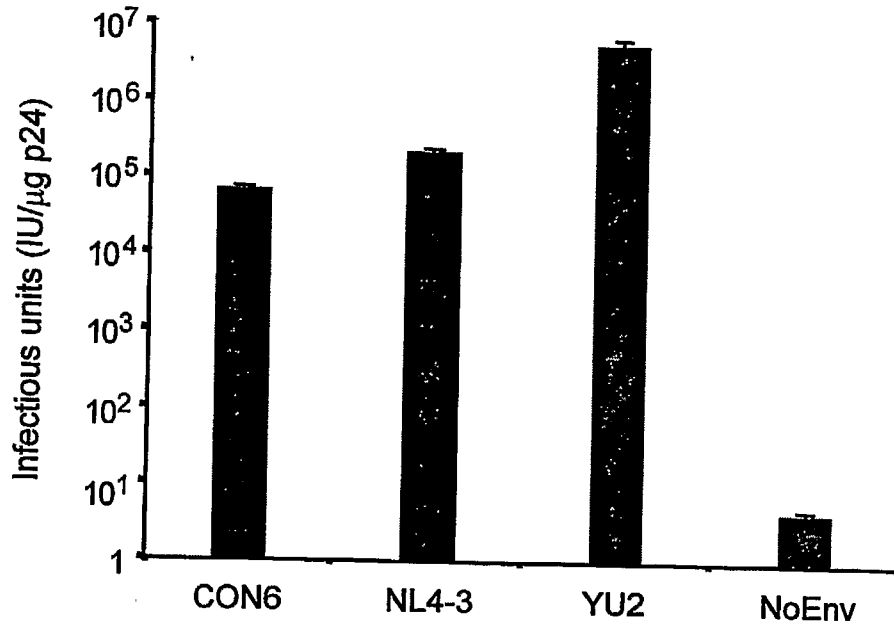


Fig. 3A

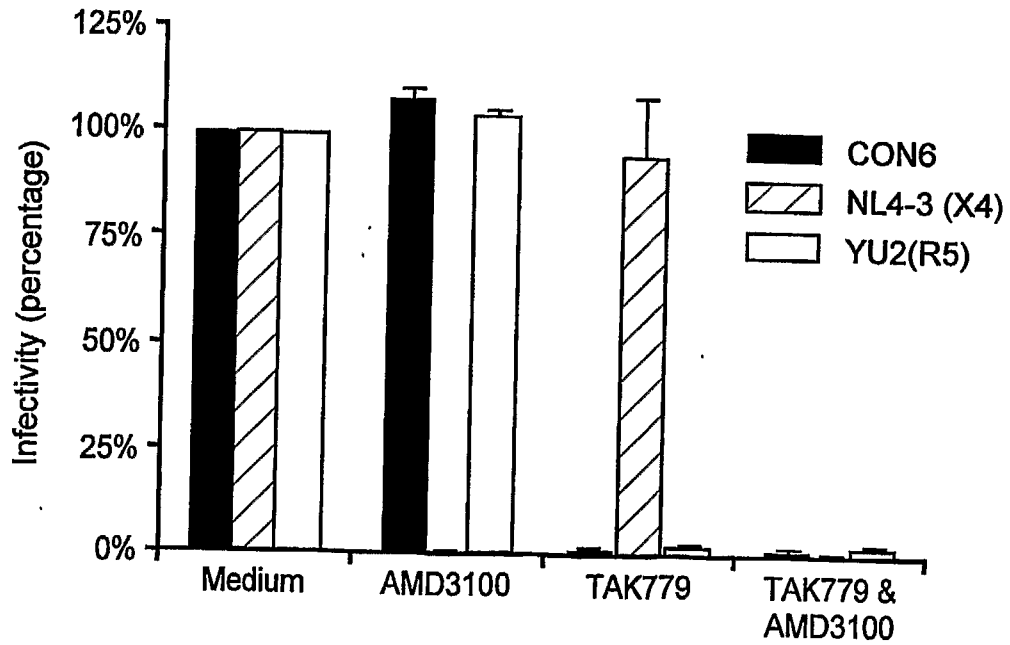


Fig. 3B

6/178

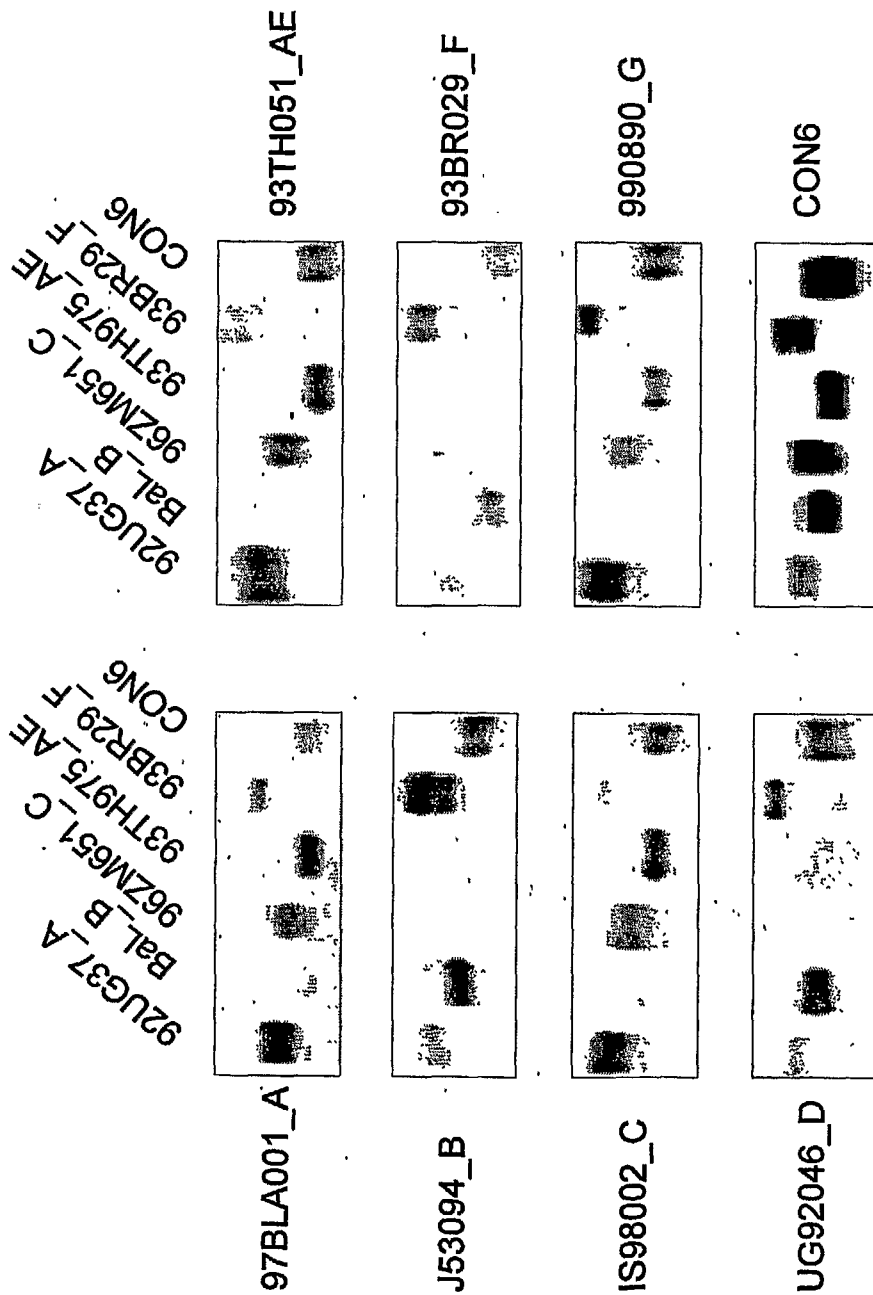
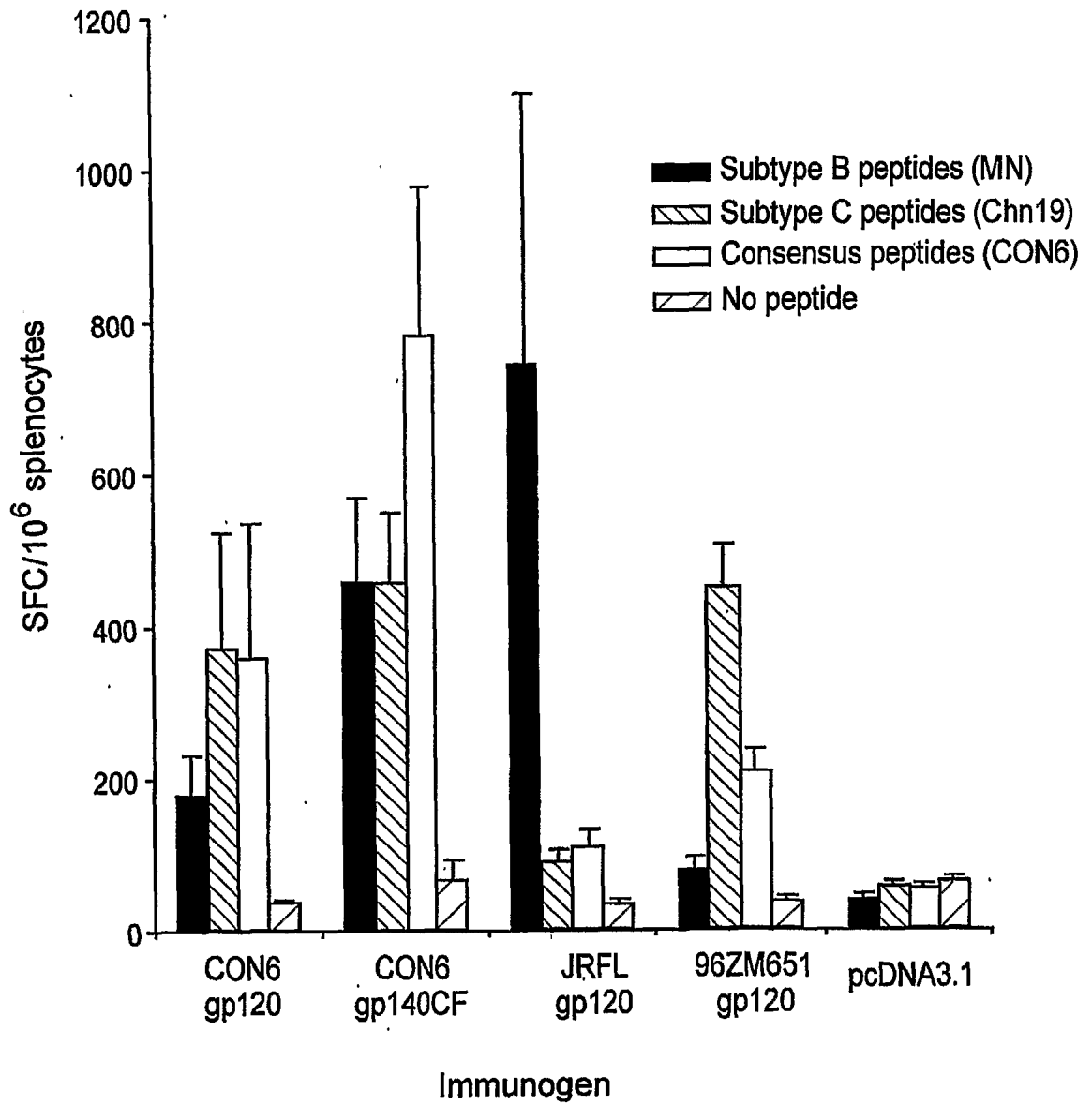


Fig. 4

7/178

*Fig. 5*

8/178

**Fig. 6A**

**C. anc. env (subtype C ancestral env. The amino acid sequence is different from Los Alamos Database August 2002)**

GCCGCCATGCGCGTGA TGGGCAT CCTGCGCAACTGCCAGCAGTGGTGGA T  
 CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCTCCGTGGTGGGCA  
 ACCTGTGGGTGACCGTGTACTACGGCGTGC CCGTGTGGAAAGGAGCCAAG  
 ACCACCCTGTTCTGCGCCTCCGACG CCAAGGCC TA CGAGCGC GAGGTGCA  
 CAACGTGTGGGCCACCACGCCTGCGTGC CACCGACCCCAA CCCCCAGG  
 AGATGGTGC TGGAGAA CGTGA CCGAGAA CT TCAACATGTGGAAGAACGAC  
 ATGGTGGACGATGCACGAGGACATCA TCTCCCTGTGGGAC CAGTCCCT  
 GAAGCCCTGCGTGAAGCTGACCCCTG TGTGCGTGA C C C T G A A C T G C A C C A  
 ACGTGA CCAACGCCAC CAACAACAC CTA CAACGCGC GAGATGAAGAATGC  
 TCCTTCAACAT CACCA CCGAG CTGC GCGACAAGAA GAAGAAGGAGTA CGC  
 CCTGTTCTA CCGCCTG GACATCGTGC C C C T G A A C G A G A A C T C C T C G A G T  
 ACCGCTTGATCAACTGCAACA C C T C G C C A T C A C C C A G G C C T G C C C A A G  
 GTGTCTTCGA C C C C A T C C C C A T C C A C T A C T G C G C C C C C G C C G G C T A C G C  
 C A T C C T G A A G T G C A A C A A C A A G A C C T T C A A C G G C A C C G G C C C T G C A A C A  
 ACGTGTCCA CCGTGCAGTGCA C C C A C G G C A T C A A G C C C G T G G T G T C C A C C  
 CAGCTGCTGCTGAA C G G C T C C T G G C C G A G G A G A T C A T C A T C G C T C  
 CGAGAACCTGACCGACAACGC CAAGACCATCATCGTGCAGCTGAA C G A G T  
 CCGTGGAGATCGTGTG C A C C C G C C C A A C A C A A C A C C C G C A A G T C C A T G  
 CGCATCGGC C C C G G C C A G A C C T T C T A C G C C A C C G G C G A C A T C A T C G G C G A  
 C A T C C G C C A G G C C C A C T G C A A C A T C T C C G A G G A C A A G T G G A A C A A G A C C C  
 TGCAGCAGGTGGCCGAGAAGCTGGGCAAGCACTTCCCAA CAAGA C C A T C  
 ACCTTCGAGC C C T C C T C C G G C G G C G A C C T G G A G A T C A C C A C C C A C T C C T T  
 CAAC TG C C G C G G C G A G T T C T T C T A C T G C A A C A C C T C C A A G C T G T T C A A C T  
 CCACCTACAACAAACA C A C C A A C T C A A C T C C A C C A T C A C C T G C C T G C  
 CGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTA  
 CGCCCCCCCATCGCCGGCAA CATCACTGCAAGTCCAACATCACCGGCC  
 TGTGCTGACC CGCGACGGCGCAAGGA GAACA C C A C C G A G A C C T T C G C  
 C C C G G C G G C G G C G A C A T G C G C G A C A A C T G G C G C T C G A G C T G T A C A A G T A  
 CAAGGTGGTGGAGATCAAGCCCTGGGCGTGGCC C C A C C G A G G C C A A G C  
 GCCGCGTGGTGAGCGCGAGAAGCGCGCCTGGGCCTGGGCGCCGTTTCT  
 CTGGGCTTCCTGGGCGCCGCCGCTCCA C C A T G G G C G C C G C C T C C A T C A C  
 CCTGACCGTGCAGGCCCGCCAGCTGCTGTCGGC A T C G T G C A G C A G C A G T  
 CCAA C C T G C T G C G C G C A T C G A G G C C C A G C A G C A C A T G C T G C A G C T G A C C  
 GTGTGGGGCATCAAGCAGCTG C A G G C C C G C G T G C T G G C C A T G G A G C G C T A  
 CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGA  
 TCTGCACCA CCGCCGTGCCCTGGAACTCCTCTGGTCCAA CAAGTCCCTG  
 GACGACATCTGGACAACATGACCTGGA TGGAGTGGGACC GC GAGATCTC  
 CAACTACACCGACACCATCTA CCGCCTGCTGGAGGAGTCC CAGAAC CAGC  
 AGGAGAAGAACGAGCAGGACCCTGGCCCTGGACTCCTGGAGAACCTG  
 TGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT  
 CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCGCCTGCTGT  
 CCATCGTGAA CCGCGTGCGCCAGGGCTA C T C C C C C C T G T C C T T C C A G A C C  
 CTGACCCCAA C C C C G C G G C C C G A C C G C C T G G A G C G C A T C G A G G A G G A  
 GGGCGGCGAGCAGGACCGCGA CCGCTCCATCCGCTGGTGTCCGGCTTC  
 TGGCCCTGG C C T G G G A C G A C C T G C G C T C C T G T G C C T G T T C T C T A C C A C  
 CGCC T G C G C G A C T T C A T C C T G A T C G C G C C G C A C C G T G G A G C T G G G  
 CCGCTCCTCCCTGCTGCGGCCTGCAGCGCGCTGGGAGGCCCTGAAGTACC  
 TGGGCTCCCTGCTGCACTACTGGG C C A G G A G C T G A A G A A G T C C G C C A T C  
 TCCCTGCTGGACACCA TCGCCATCGCCGTGGCCGAGGGCA CCGACCGCAT  
 CATCGAGGTGGTGCAGCGGCCTGC C G C G C C A T C C T G A A C A T C C C C C G C C  
 GCATCCG

9/178

*Fig. 6B*

C.con.env (subtype C consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT  
 CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTGGTGGGCA  
 ACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAG  
 ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGG AGGTGCA  
 CAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGG  
 AGATGGTGTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC  
 ATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCT  
 GAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCCGCA  
 ACGTGACCAACGCCACCAACAACACCTACAACGAGGAGATCAAG AACTGC  
 TCCTTCAACATCACCACCGAGCTGCGCGACAAGAAGAAGAAGGTGTACGC  
 CCTGTTCTACCGCCTGGACATCGTGCCCTGAACGAGAACTCCTCCGAGT  
 ACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCAAG  
 GTGTCCTTCGACCCCATCCCATCCACTACTGCGCCCCCGCCGGCTACGC  
 CATCCTGAAGTGAACAACAAGACCTTCAACGGCACCCGGCCCCCTG CAACA  
 ACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACC  
 CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTC  
 CGAGAACCTGACCAACAACGCCAAGACCATCATCGTGACCTGAACGAGT  
 CCGTGGAGATCGTGTGCACCCGCCCAACAACAACACCCGCAAGTCCATC  
 CGCATCGGCCCGCCAGACCTTCTACGCCACCCGGCGACATATCG GCGA  
 CATCCGCCAGGCCCACTGCAACATCTCCGAGGACAAGTGAACAAGACCC  
 TGCAGCGCGTGTCCAAGAAGCTGAAGGAGCACTTCCCAACAAGACCATC  
 AAGTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCAACCCACTCCTT  
 CAACTGCCGCGGCGAGTTCCTTCTACTGCAACACCTCCAAGCTGTTCAACT  
 CCACCTACAACAACAACACCAACTCCAACTCCACCATCACCCCTGCC TGC  
 CGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGTA  
 CGCCCCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCGGCC  
 TGCTGCTGACCCGCGACGGCGGCAAGAAGAACACCACCGAGATCTTCCGC  
 CCGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTA  
 CAAGGTGGTGGAGATCAAGCCCTGGGCGTGGCCCCCACCAAGGCCAA GC  
 GCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTT  
 CTGGGCTTCCCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCAC  
 CCTGACCCTGTCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGT  
 CCAACCTGCTGCGCGCCATCGAGGCCAGCAGCACATGCTGCAGCTGACC  
 GTGTGGGGCATCAAGCAGCTGCAGACCCGCGTGTGGCCATCGAGCGCTA  
 CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGA  
 TCTGCACCACCGCCGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCAG  
 GAGGACATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCTC  
 CAACTACACCGACACCATCTACCGCCTGCTGGAGGACTCCCAGAACCAGC  
 AGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACTCCTGGAAGAACCCTG  
 TGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT  
 CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCGCCGTGCTGT  
 CCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACC  
 CTGACCCCAACCCCGCGGCCCGGACCGCCTGGGCCGCGCATCGAGGAGGA  
 GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCC  
 TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCTACCAC  
 CGCCTGCGCGACTTCATCCTGGTGGCCGCCCCGCGCCGTGGAGCTGCTGGG  
 CCGCTCCTCCCTGCGCGGCTGCAGCGCGGCTGGGAGGCCCTGAAGTACC  
 TGGGCTCCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGTCCGCCATC  
 TCCTGCTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCAT  
 CATCGAGCTGATCCAGCGCATCTGCCGCGCCATCCGCAACATCCCCCGCC  
 GCATCCGCGACCGCTTCGAGCCGGCTTCAGTAA

C.anc.env (subtype C ancestral env)

MRVMGILRNCQQWIIWGIIGFWMLMI CSVVGNLWVTVYGVVPVWKEAKTTLFCASDAKAYEREVHNWVWAT  
HACVPTDPNPQEMVLENTENFNMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCRNVTNATNNT  
YNGEMKNCNFNIITELRDKKKKYYALFYRLDIVPLN ENSSEYRLINCNTSAITQACP KVSFDPIPIHYCA  
PAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEEIIIRSENLTDNAKTIIVQLN  
ESVEIVCTRPNNTRKSMRIGPGQTFYATGDIIGDIRQAHNCNISEDKWNKTLQOVAEKLGHKFFPNKTIITF  
EPSSGGDLEITTHSFNCRGEEFFYCNTSKLFNSTYNNNTNSNTITLPCR IKQIINMWQGVGQAMVAPP  
GNI TCKSNI TGLLLTRDGGKENTETFRPGGDMRDNRSELYKYKVEIKPLGVAPTEAKRRVVEREKR  
AVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLRAIEAQHMLQLTVWGIKQLQARVL  
AMERYLKDQQLLGIWGCSSGKLICTTAVPWNSSWSNKSLLDDIWDNMTWMEWDREISNYTDTIYRLLEESQN  
QQEKNEQDLLALDSWENLWVFDITNWLWYIKIFIMIVGGLIGLRIIFAVL SIVNRVROGYSPLSFQTLT  
PNRPGPDRLERIEEGEGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLGRSSLR  
GLQRGWEALKYLGSLVQYWQELKKS AISLLDTIAIAVAEGTDRIIEVVQACRAILNIPRRIRQGFEEA  
LL

Fig. 6C

C.con.env (subtype C consensus env)

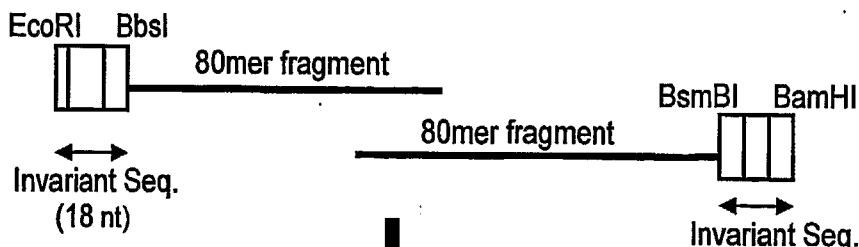
MRVMGILRNCQQWIIWGIIGFWMLMI CNVVGNLWVTVYGVVPVWKEA KTTLFCASDAKAYEREVHNWVWAT  
HACVPTDPNPQEMVLENTENFNMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCRNVTNATNNT  
YNEEIKNCNFNIITELRDKKKKYYALFYRLDIVPLNENSSEYRLINCNTSAITQACP KVSFDPIPIHYCA  
PAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEEIIIRSENLTDNAKTIIVHLN  
ESVEIVCTRPNNTRKSIRIGPGQTFYATGDIIGDIRQAHNCNISEDKWNKTLQOVSKLKEHFPNKTIKF  
EPSSGGDLEITTHSFNCRGEEFFYCNTSKLFNSTYNNNTNSNTITLPCR IKQIINMWQEVGRAMYAPP  
GNI TCKSNI TGLLLTRDGGKNTTEIFRPGGDMRDNRSELYKYKVEIKPLGVAPTEAKRRVVEREKR  
AVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLRAIEAQHMLQLTVWGI KOLQTRVL  
AIERYLKDQQLLGIWGCSSGKLICTTAVPWNSSWSNKSQED IWDNMTWQWDREISNYTDTIYRLLEDSQN  
QQEKNEKDLLALDSWKNLWVFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVROGYSPLSFQTLT  
PNRPGPDRLGRIEEGEGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILVAARAVELLGRSSLR  
GLQRGWEALKYLGSLVQYWGLELKKS AISLLDTIAIAVAEGTDRIIELIQIRICRAIRNIPRRIRQGFEEA  
LQ

Fig. 6D

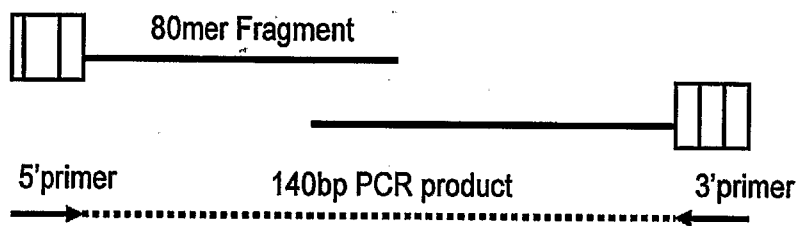
11/178

**Fig. 6E**

Synthesize entire gene in 80-mer fragments overlapping by 20 residues at the 3' end with invariant sequences at the 5' end.

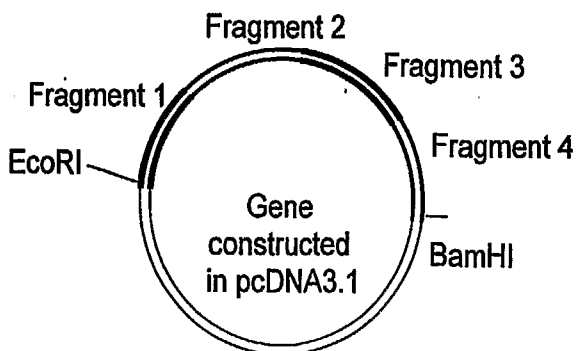


Paired 80mer oligos are connected via PCR in a stepwise manner from 5' to 3' using primers complimentary to the invariant seq.



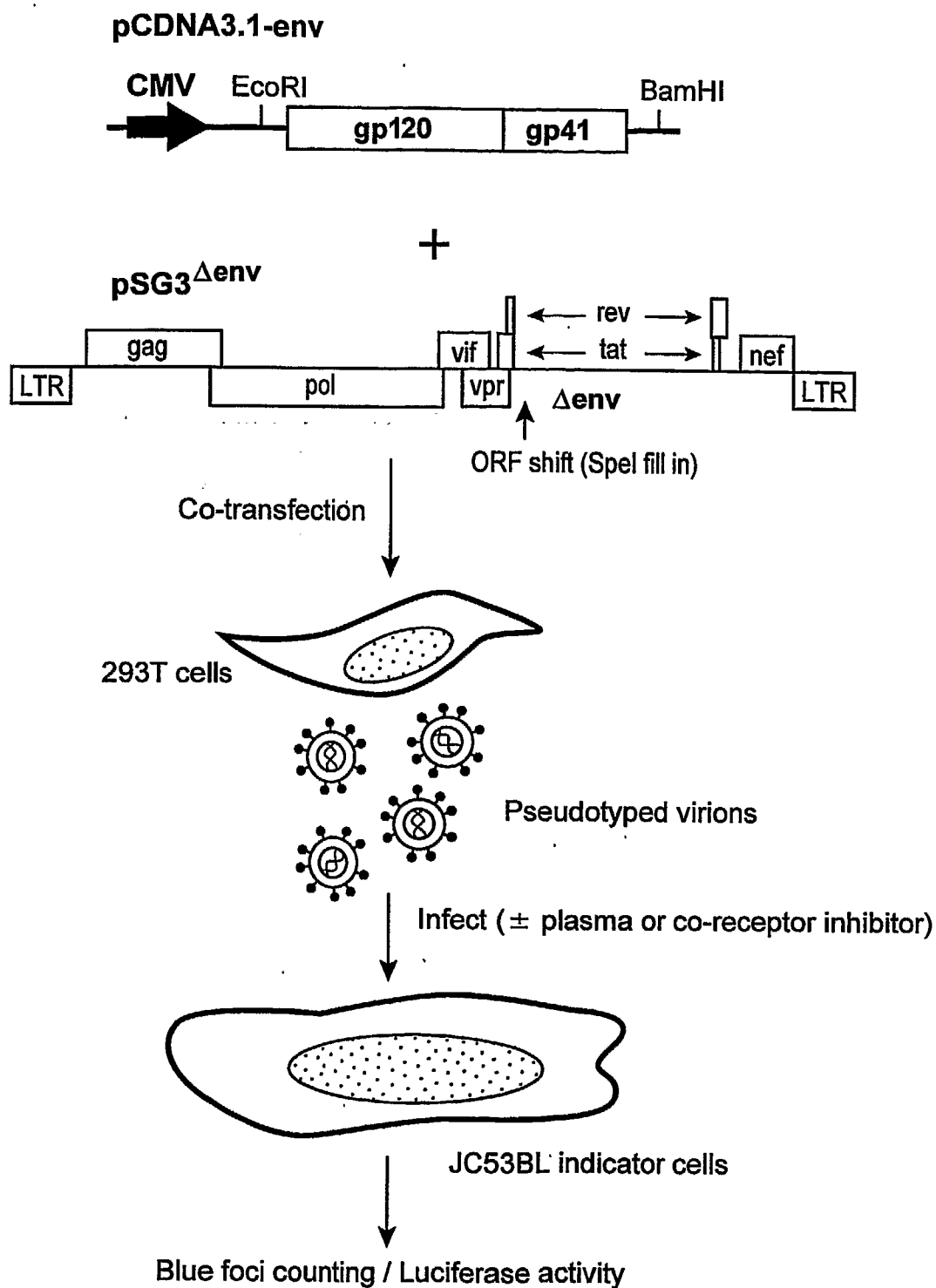
108bp PCR fragments cloned into pGEM-T and sequenced. Clones with the proper sequence will be cut with 2 restriction enzymes. 4 fragments will be ligated together with pcDNA3.1 in a stepwise manner from the 5' to 3' end of gene

| Fragments to be ligated with pcDNA3.1 (1-4 are in order from 5' to 3') | Restriction Enzymes Used to Cleave Fragment |
|--|---|
| Fragment 1   | EcoRI/BsmBI                                 |
| Fragment 2   | BbsI/BsmBI                                  |
| Fragment 3   | BbsI/BsmBI                                  |
| Fragment 4   | BbsI/BamHI                                  |
| pcDNA3.1   | EcoRI/BamHI                                 |



Ligations will be repeated stepwise 5' to 3' until the entire gene has been cloned into pcDNA3.1

12/178

*Fig. 7*





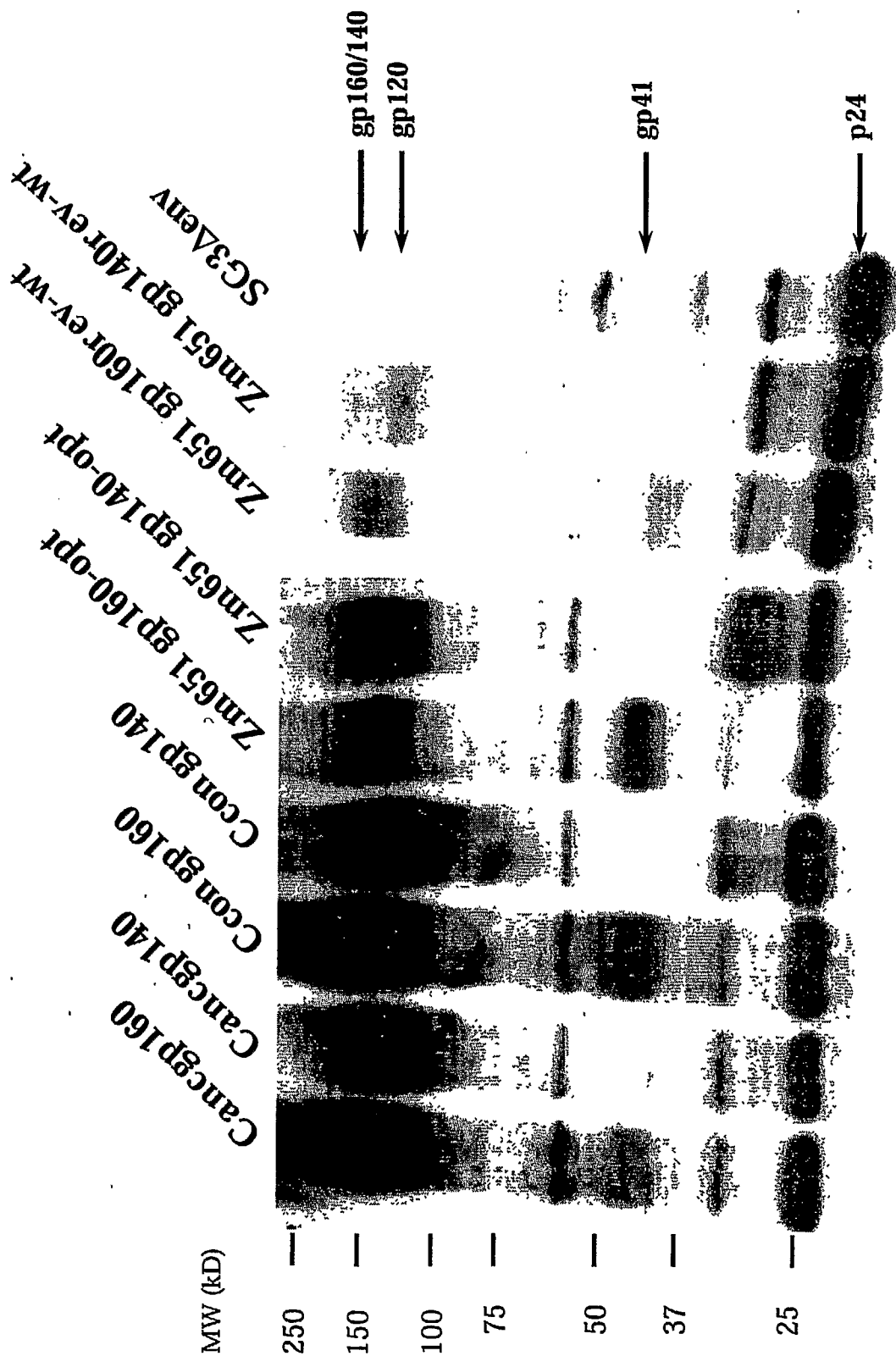
14/178



Fig. 9

15/178

Fig. 10A



16/178

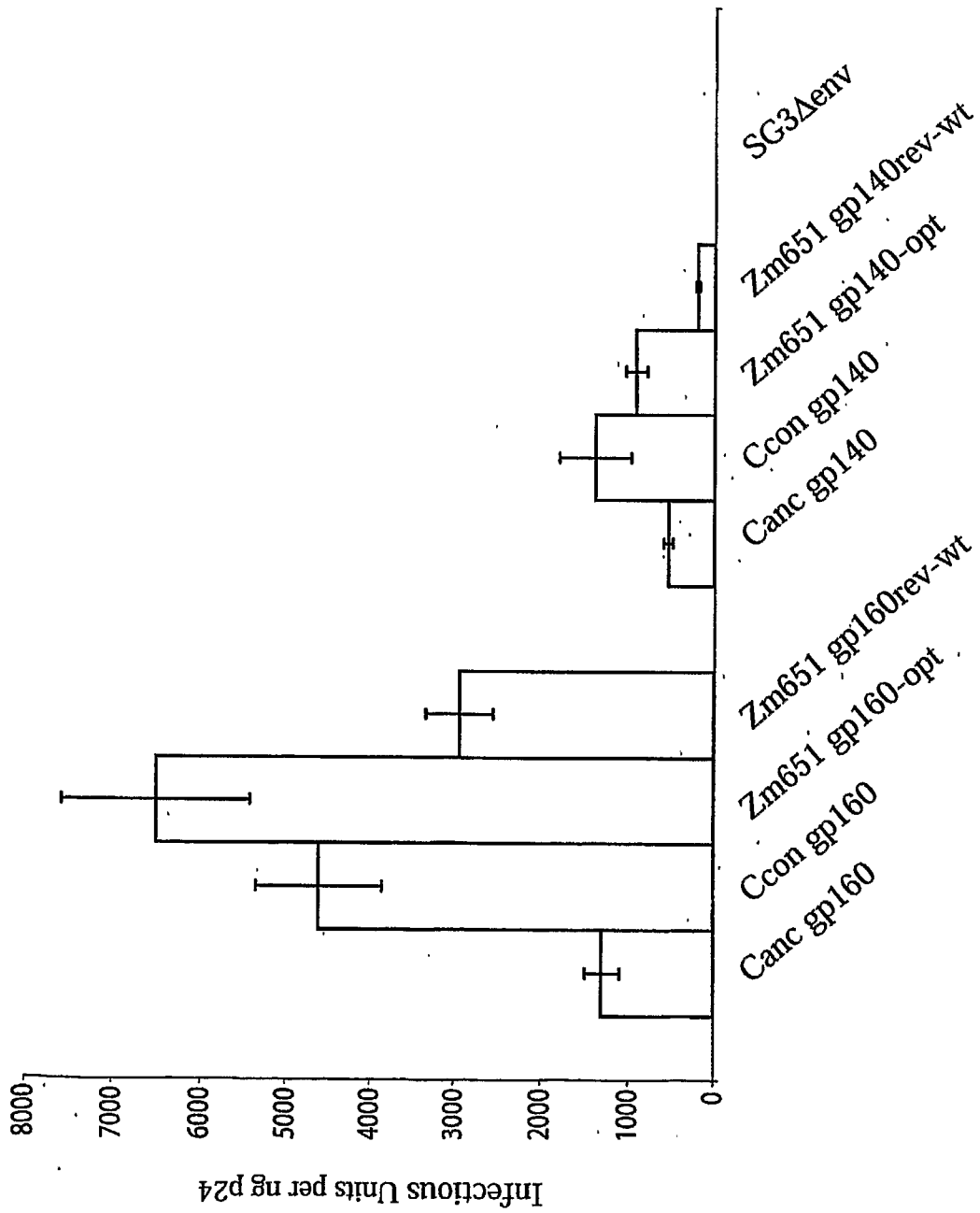


Fig. 10B

17/178

Fig. 11

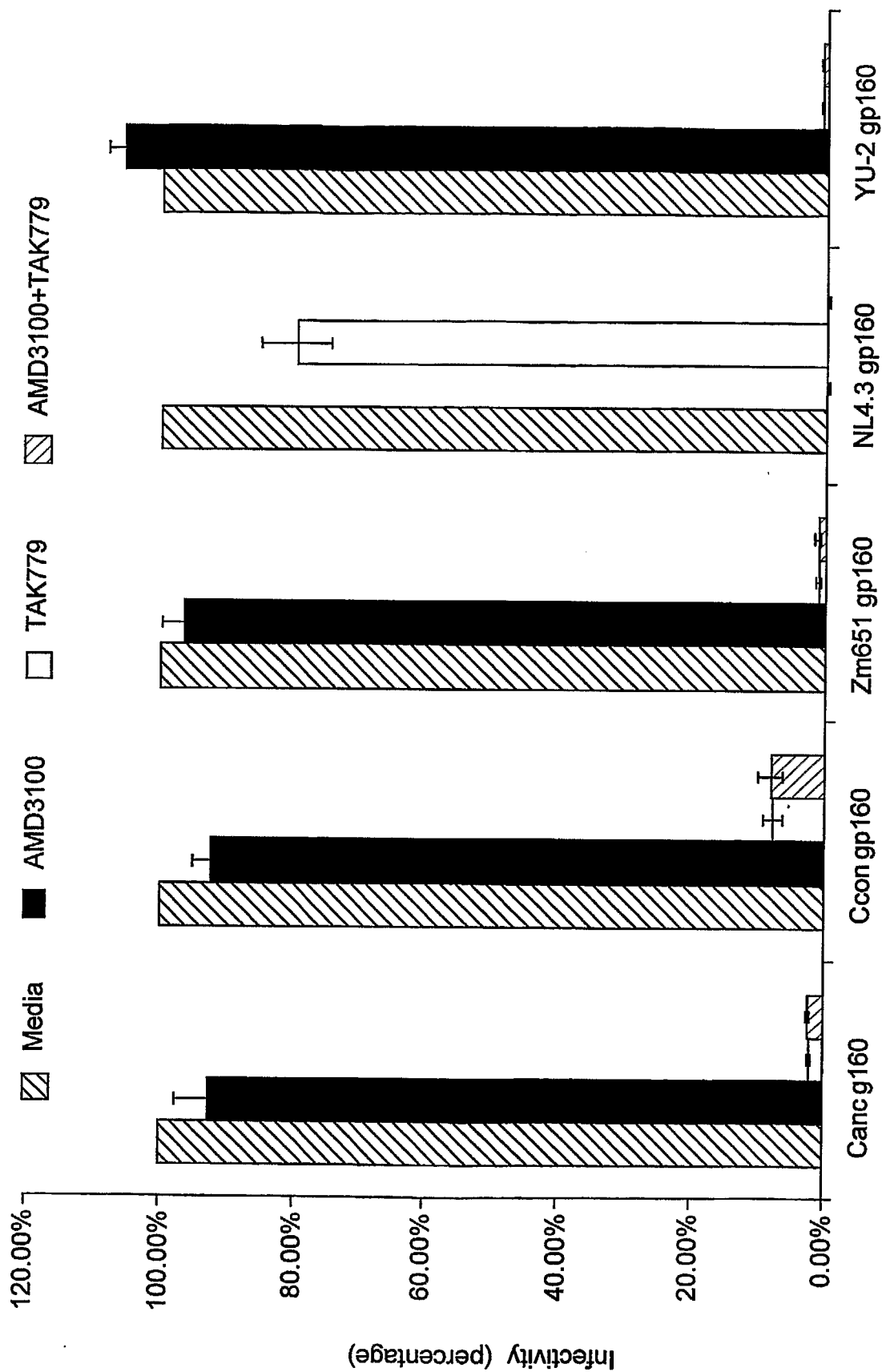


Fig. 12A

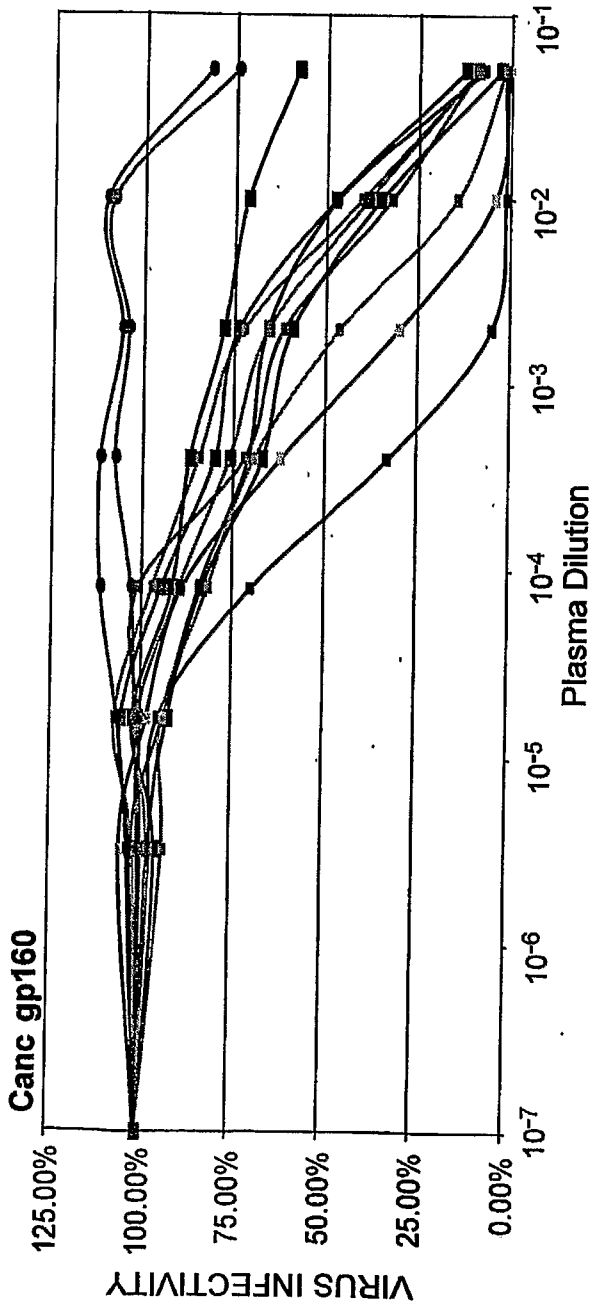
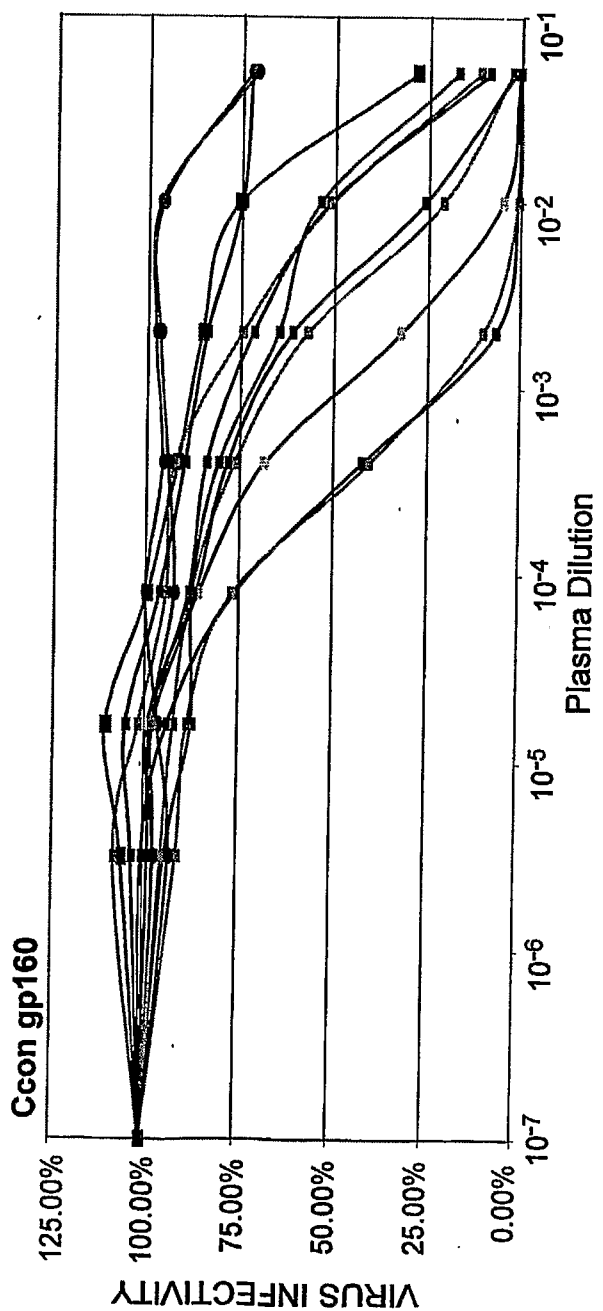


Fig. 12B



Plasma from HIV-1 subtype C infected patients

Plasma from uninfected donors

19/178

Fig. 12C

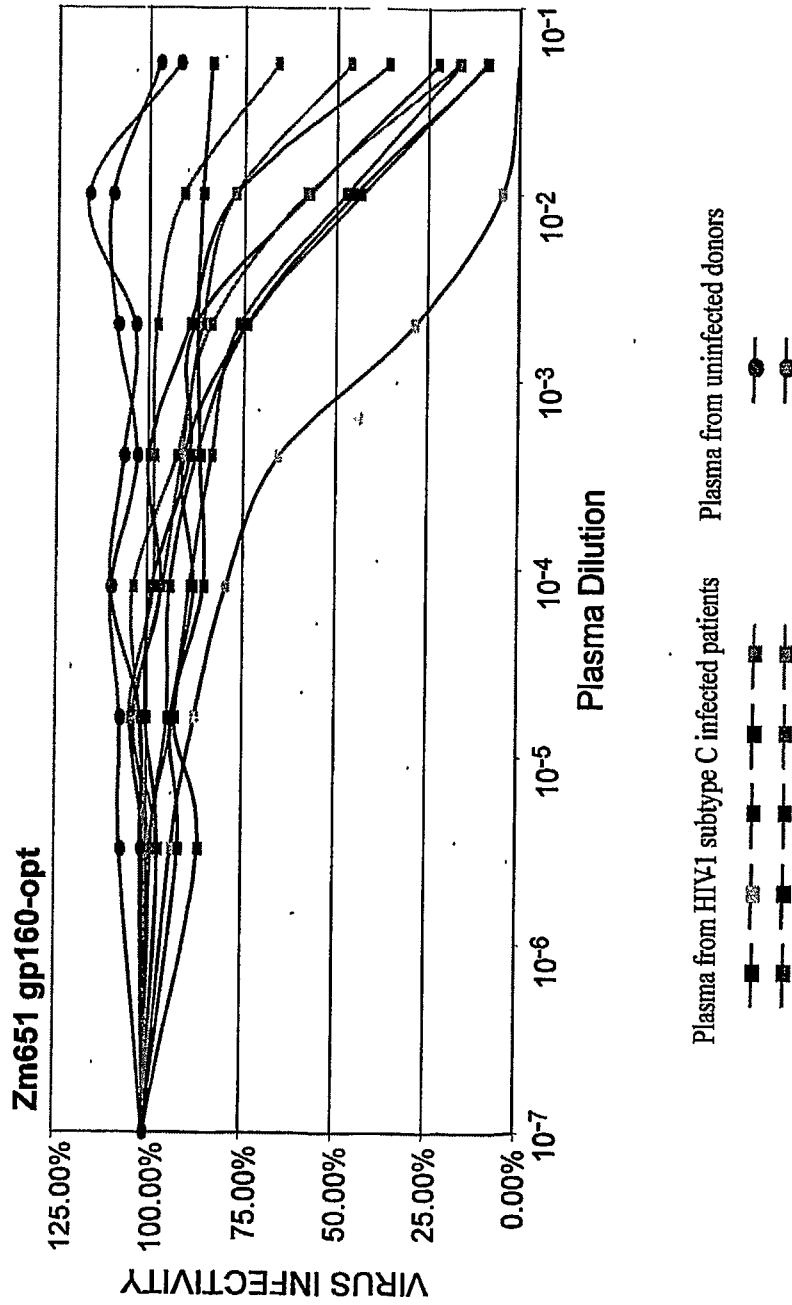
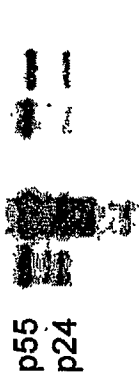


Fig. 13A

Gag

96ZM651  
 Consensus C  
 Mock  
 96ZM651  
 Consensus C  
 Mock

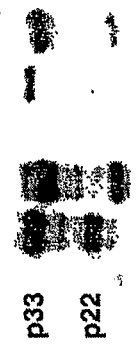


Cell lysate Supernatant

Fig. 13B

Nef

96ZM651  
 Consensus C  
 Mock  
 96ZM651  
 Consensus C  
 Mock



Cell lysate Supernatant

C.con.gag (subtype C con sensus gag)

MGARASILLRGGKLDTWKIRLRPGGKKRYMIKHLVWASRELERFALNPGLLLEETSEGCKQIMKQLQPA  
 LQTGTEELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQOKTQQAEEAADGKVSQNYPI  
 VQNLQGMVHOAISPRTLNAWKVIIEKAFSPEVIMFTALSEGATPQDLNMLNTVGGHQAAMQMLKDT  
 INEEAAEWDRLHPVHAGPIAPGQMPREPRGSDIAGTSTLQEQIAWMTSNPPVPVGDIIYKRWIIILGLNKIV  
 RMYSPVSIIDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDILLVQANPDCKTILLRALGPGASLE  
 EMMTACQGVGGPSSHARVLAEMSQANNINIMQRSNFKGPKRIVKCFNCGKEGHIARNCRAPRKKGCWK  
 CGKEGHQMKDCTERQANFLGKIWP SHKGRPGNFLQSRPEPTAPPAESFRFEETTPA  
 PKQEPKDREPLTSLKSLFGSDPLSQ

Fig. 13C

C.con.nef (subtype C consensus nef)

MGGKWSKSSIVGWPVAVRERIRRTPEAAEGVGAASQDLKYGALTSNTATNNADCAWLEAQEEEEEEV  
 GFPVRRPQVPLRPMTYKAAFDSLFFLKEKGGLEGLIYSKKRQELDLWVYHTQGFPPDWQNYTTPGPGVRY  
 LITFGWCFKLVVDPREVEEANEGENNCLLHPMSQHGMEDEDEVLKWKFDShLARRHMARELHPEYYKDC

Fig. 13D



**C.con.gag (subtype C consensus gag. Not in the public domain)**

GCCGCCCATGGGGCCCGCCAGCATCTGGCGGGCGGCAAGCTGGACACCTGGGAGAAGATCCGCC  
 TGGCCCCGGCAAGAAGCGTACATGATCAAGCACCTGGTGTGGCCAGCCGGAGCTGGAGCGCTT  
 CGCCCTGAACCCCGGCTGTGGAGACCAGCGAGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCCGCC  
 CTGCAGACCGGCACCGAGGAGCTGCGCAGCTGTACAACACCGTGGCCACCTGTACTGCGTGCACGAGA  
 AGATCGAGGTGCGGACACCAAGGAGCCCTGGACAAGATCGAGGAGGAGCAAGAACAGAGCCAGCAGAA  
 GACCCAGAGCCGAGCCCGCCGACGGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACCTGCAG  
 GGCCAGATGGTGCACAGGCCATCAGCCCGCCGACCTGAACGCTGGGTGAAGGTGATCGAGGAGAAGG  
 CCTTCAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCACCCCGCAGGACCTGAACAC  
 CATGCTGAACACCCGTGGGGCCACAGGCCGCAATGAGATGCTGAAGGACACCATCAACGAGGAGGCC  
 GCCGAGTGGACCGCTGCACCCCGTGCACCGCCGCCCCATCGCCCCGCGCAGATGCGCGAGCCCGCG  
 GCAGGCATCGCCGGCACCAACAGCACCTGCAAGGAGCAGATCGCCTGGATGACCAAGCCCGCCCGGT  
 GCCCGTGGCGACATCTACAAGCGTGGATCATCTGGGCTGAAACAAGATCGTGCAGTGTACAGCCCC  
 GTGAGCATCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGA  
 CCTGCGCGCCGAGCGCACCCAGGACGTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAGAACGC  
 CAACCCGACTGCAAGACCATCTGCGGCCCTGGCCCCGCGCCAGCTGGAGGAGATGATGACCGCC  
 TGCCAGGCGTGGCGGGCCCCAGCCACAAGGCCCGCTGCTGGCCGAGCCATGAGCCAGGCCAACACA  
 CCAACATCATGATGAGCGCAGCAACTTCAAGGGCCCCAAGCGCATCGTGAAGTCTTCAACTGCGGCAA  
 GGAGGGCCACATCGCCCGCAACTGCGGCCCCCGCAAGAAGGCTGTGGAAGTGGCGCAAGGAGGGC  
 CACCAGATGAAGACTGCACCGAGCGCCAGGCCAATCTCTGGCAAGATCTGGCCAGCCACCAAGGGCC  
 GCCCGGCAACTTCTGCAAGCGCCCGGCCCCAGGCCCAAGGCCCGCTGCTGGCCGAGCCATGAGCCAGGCCAACACA  
 GACCAACCCCGCCCCAAGCAGGAGCCCCAAGGACCGCGAGCCCTGACCAGCCCTGAAGAGCCCTGTTCGGC  
 AGCGACCCCTGAGCCAGTAA

**Fig. 13E**

**C.con.nef (subtype C consensus nef. Not in the public domain)**

GCCGCCCATGGGGCGCAAGTGGAGCAAGAGCAGCATCGTGGGCTGGCCCGCGTGGCGGAGCGCATCC  
 GCCGCACCGAGCCCGCCGAGGGCGTGGCGCCGCGCAGCCAGGACCTGGACAAGTACGGCGCCCTGAC  
 CAGCAGCAACACCGCCACCAACACCGCCGACTGCGCTGGCTGGAGGCCAGGAGGAGGAGGAGGTG  
 GGCTTCCCCGTGGCCCCCAGGTGCCCCGCCCCATGACCTACAAGGCCGCTTCGACCTGAGCTTCT  
 TCCTGAAGGAGAGGGCGGCTGGAGGGCCCTGATCTACAGCAAGAAGCCAGGAGATCTTGGACCTGTG  
 GGTGTACACACCCAGGGCTTCTTCCCCGACTGGCAGAACTACACCCCGGCCCGGCTGCGCTACCCCC  
 CTGACCTTCGGCTGGTCTTCAAGTGGTCCCGTGGACCCCGCGGAGGTGGAGGAGGCCAACGAGGGCC  
 AGAACAACTGCTGTGCACCCCATGAGCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTG  
 GAAATTGCACAGCCACCTGGCCCCCGCCACATGGCCCCGGAGCTGCACCCCGAGTACTACAAGGACTGC

TGA

**Fig. 13F**

CONs.env (gorup M consensus env gene. This one contain the consensus sequence for variable regions in env gene)

MRVRFIQRNCQHLLWRWGTLILGMLMICSAAENLWVTVYGVPPVWKEANTTLFCASDAKAYDTEVHNV  
WATHACVPTDPNPQEI VLENVTENFNWKNMVEQMHEDI ISLWDQSLKPCVKLTPLCVTLNCTNVNVTN  
TTNNTTEEKGEIKNCSFNITTEIRDKKQVYALFYRLDVVPIDNNNNSSNYRLINCNTSAITQACPVSF  
EPIPIHYCAPAGFALLKCNDDKFNKGTGPKNVSTVQCTHGIKPVVSTQLLNGSLAEFEIIIRSENIITNN  
AKTIIVQLNESVEINCRPNNTKRSIRIGPGQAFYATGDIIGDIRQAHACNISG'KWNKTLQQVAKKLR  
HFNNKTIIFKPSGGDLEITTHSFNCRGEFFYCN'TSGLFNSTWINGTKNNNTNDTI TLP'CRIKQIINM  
WQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNE'ETEIFRPGGDMRDNRSELYKYKVVKIEPLG  
VAPTKAKRRVVEREKRAVIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQOHL  
LQLTVWGIKQLQARVLAVERYLKDQQLLGIWCCSGLICTTTVPWNSSWSNKSQDEIWDNMTWMEWEREI  
NNYTDIIYSLIEESONQOKEQELALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIV  
NRVRQGYSPLSFQTLIPNPRGPPDRPEGIEEFGGQDRDRSIRLVNGFLALAWDDLLRSLCLFSYHRLRDFI  
LIIAARTVELLGRKGLRRGWEALKYLWNLLQYWGQELKNSAISLLDTTAAIAVAEGTDRVIEVVQRACRAIL  
NIPRRIRQGLERALL

Fig. 14A

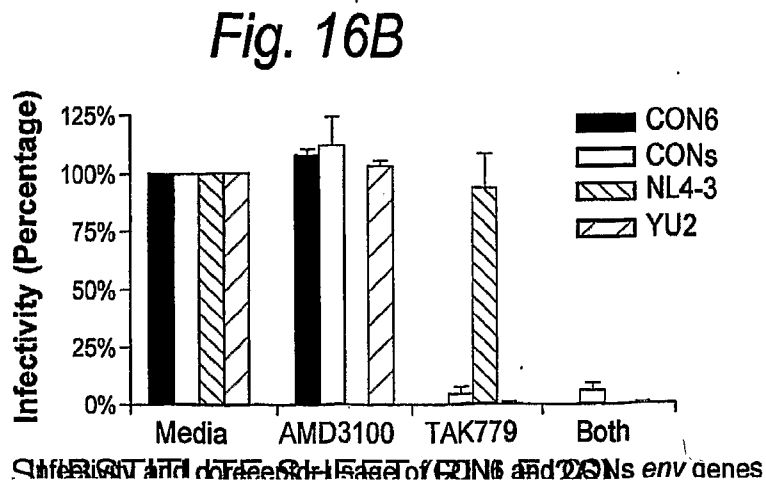
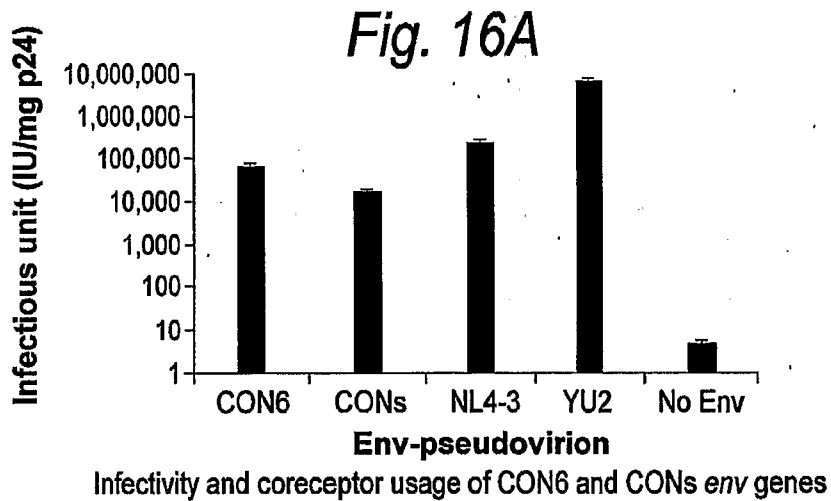
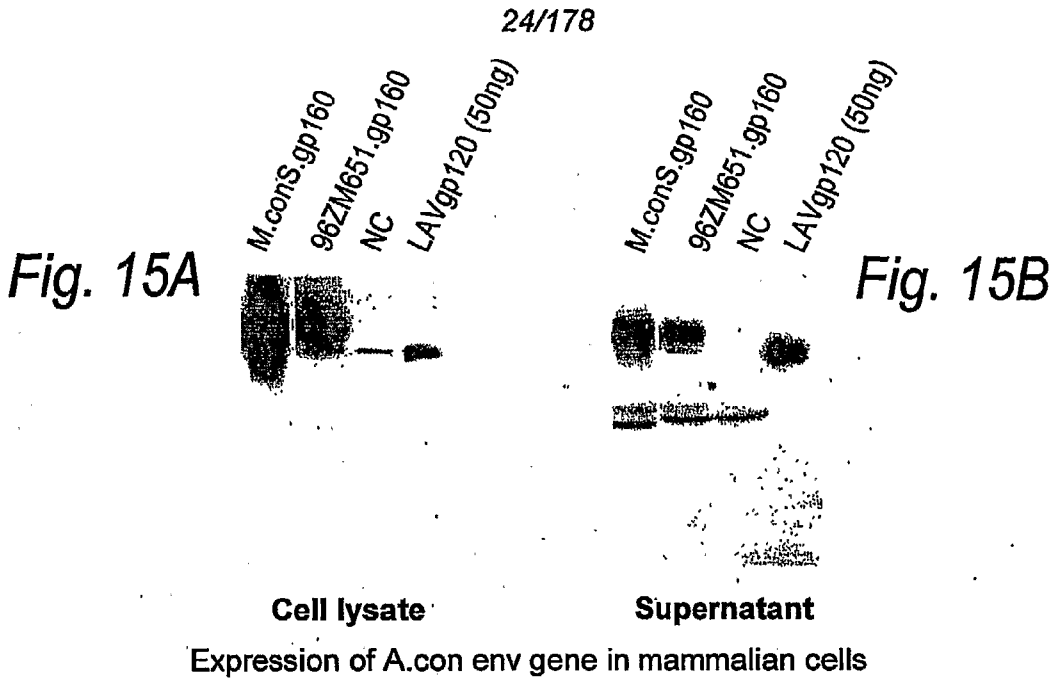


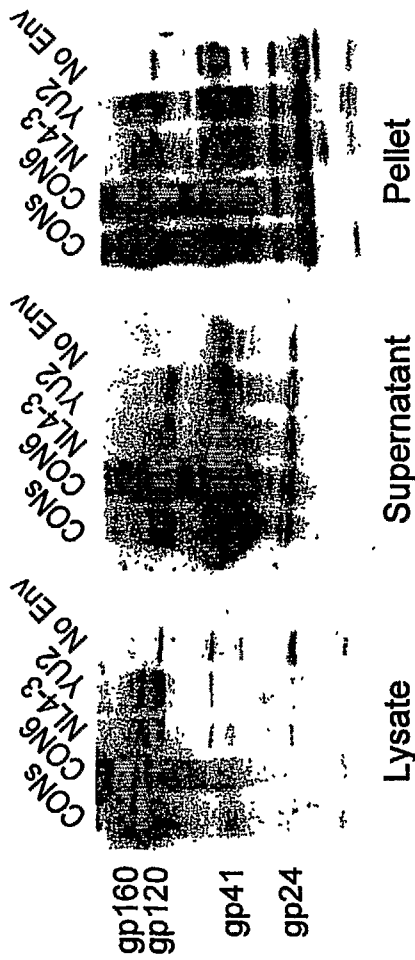
Fig. 14C

**Fig. 14B**

CONs.env (gorup M consensus env gene. This one contain the consensus sequence for variable regions in env gene. The identical amino acid sequences as in the public domain)

GCCGCCGCCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTG  
 GCGCTGGGGCACCTGATCCTGGGCATGCTGATGATCTGCTCCGCCGCCG  
 AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC  
 AACACCACCCTGTTCTGCGCTCCGACGCCAAGGCCTACGACACCGAGGT  
 GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCC  
 AGGAGATCGTGCTGGAGAACGTGACCGGAACTTCAACATGTGGAAGAAC  
 AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCATG  
 CCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAACTGCA  
 CCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGCGAG  
 ATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCA  
 GAAGGTGTACGCCCTGTTCTACCGCTGGACGTGGTGCCCATCGACGACA  
 ACAACAACAACCTCCTCCAACCTACCGCCTGATCAACTGCAACACCTCCGCC  
 ATCACCACAGGCTGCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTA  
 CTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCA  
 ACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGC  
 ATCAAGCCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA  
 GGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCA  
 TCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAAC  
 AACAAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAGGCCTTCTACGC  
 CACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCG  
 GCACCAAGTGGAAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAG  
 CACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCT  
 GGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCA  
 ACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAAC  
 AACAAACAACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT  
 CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCATCG  
 AGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCCTGCTGCTGACCCGC  
 GACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCGGCGG  
 CGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTGG  
 TGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTG  
 GTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCCTGGGCTT  
 CCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCCTGACCG  
 TGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTG  
 CTGCGGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG  
 CATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGG  
 ACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCACC  
 ACCACCGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCAGGACGAGAT  
 CTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACCTACA  
 CCGACATCATCTACTCCCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAG  
 AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACCTG  
 GTTCGACATCAACAACTGGCTGTGGTACATCAAGATCTTCATCATGATCG  
 TGGGCGCCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCCATCGTG  
 AACCGCGTGCGCCAGGGCTACTCCCCCTGTCCTTCCAGACCCCTGATCCC  
 CAACCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGGC  
 AGCAGGACCGCGACCGCTCCATCCGCTGGTGAACGGCTTCCCTGGCCCTG  
 GCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCG  
 CGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCCGCGAG  
 GCCTGCGCCGCGGCTGGGA GGCCCTGAAGTACCTGTGGAACCTGCTGCAG  
 TACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAC  
 CGCCATCGCCGTGGCCGAGGGCACCGACCGGTGATCGAGGTGGTGCAGC  
 GCGCCTGCCGCGCCATCCTGAACATCCCCCGCCGCATCCGCCAGGGCCTG  
 GAGCGCGCTGCTGTA





Env protein incorporation in CON6 and CONs Env-pseudovirions

Fig. 17A Fig. 17B Fig. 17C

A. con. env (subtype A consensus env)

MRVMGIQRNCOHLWRWGTMLGMIIISAAENLWTVVYGVVWKKDAETTLFCASDAKAYDTEVHNV  
 WATHACVPTDPNPQEIINLENVTEEFNMWKNMVEQMHTDIIISLWDQSLKPCVKLTPLCVTLNCSNVNVT  
 NITNITDNMKGKIKNCSFNMTTELDRDKKQKVSFLFYKLDVVQINKSNSSSQYRLINCNTSAITQACP  
 KVSFEPIPIHYCAPAGEAIIKCKDKEFNVTGPKNVSTVQCTHGIKPVVSTQLLLNGSLAEEVEMIRSE  
 NITNNAKNIIVQLTKPVKINCTRPNNNTRKSIIRIGPQAFYATGDIIGDIRQAHCVSRTEWNETLQK  
 VAKQLRKYFNKTIIFTNSSGGDLEITTHSFNCGGEFFYCNSTGLFNSTWNGTKKKNSTESNDTITLPC  
 RIKQIINMWQRVGQAMYAPPIQGVIRCESNITGLLLTRDGGDNNSKNETFRPGGGMDRDNWRSELYK  
 YKVVKIEPLGVAPTAKARRVVEREKRAVIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQ  
 SNLLRAIEAQQHLLKLTVWGIKQLQARVLAVERYLKDQQLLGIWCSGKLICTTNVPMNSSWSNK  
 SQSEIWDNMTWLQWDEIISNYTDIIYNLIEESQKNEQDLALDKWANLW NWFDISNWLWYKIFIMIVGGLI  
 GLRIVFAVLSVINRVRQGYSPLSFQTHTPNPGGLDRPGRIEEEGGEQGRDRSIRLVSGFLA  
 LAWDDLRSCLFSEYHRLRDFILIAARTVELLGHSSIKGLRLGWEGLYLWNLNLLYWGRELKISAINL  
 LDTIAI AVAGWTRDRIEIGQRI CRAILNIPRRIRQGLERALL

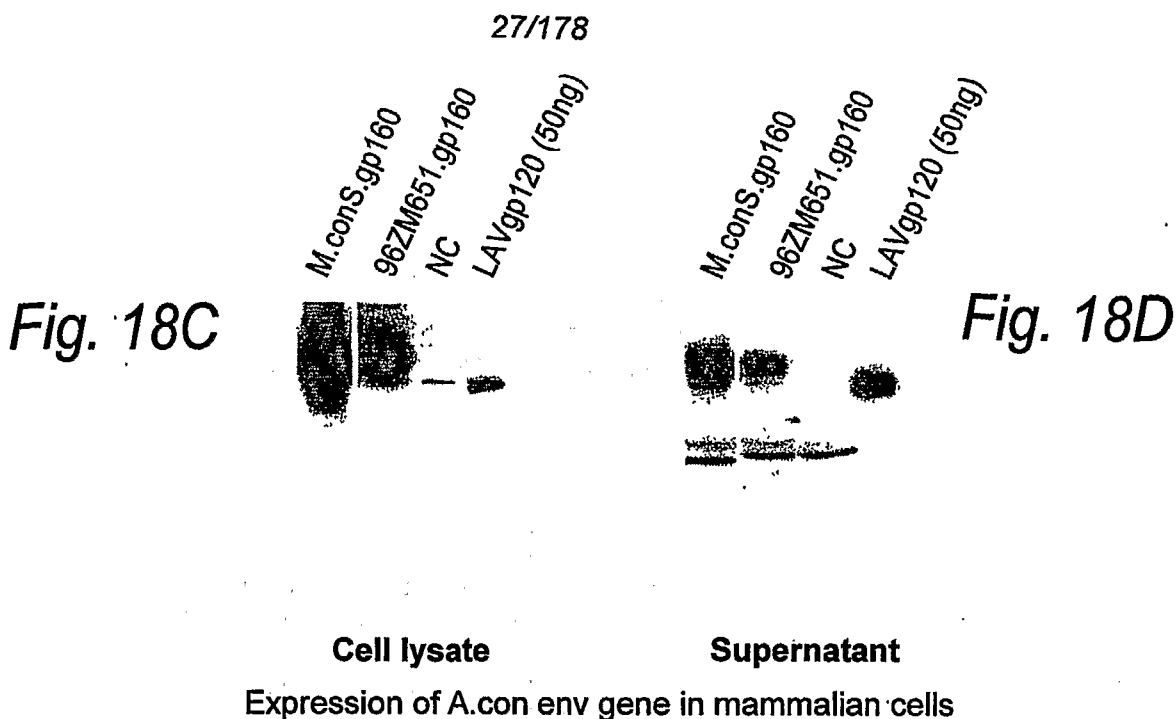
Fig. 18A

26/178

**Fig. 18B**

A.con.env (subtype A consensus env. Identical amino acid sequence to that in the public domain)

GCCGCCGCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTG  
GCGCTGGGGCACCATGATCCTGGGCATGATCATCATCTGCTCCGCCGCCG  
AGAACCCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGACGCC  
GAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCCGACCCCAACCCCC  
AGGAGATCAACCTGGAGAACGTGACCGAGGAGTTC AACATGTGGAAGAAC  
AACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGGACCAGTC  
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCCTGAACTGCT  
CCAACGTGAACGTGACCACCAACATCACCAACATCACCGACAACATGAAG  
GGCGAGATCAAGAACTGCTCCTTCAACATGACCACCGAGCTGCGCGACAA  
GAAGCAGAAGGTGTACTCCCTGTTCTACAAGCTGGACGTGGTGCAGATCA  
ACAAGTCCAACCTCCTCCAGTACCGCCTGATCAACTGCAACACCTCC  
GCCATCACCCAGGCC TGCCCAAGGTGTCCTTCGAGCCCATCCCCATCCA  
CTACTGCGCCCCCGCCGGCTTCGCATCCTGAAGTGCAAGGACAAGGAGT  
TCAACGGCACCCGGCCCTGCAAGAACGTGTCCACCGTGCACTGCACCCAC  
GGCATCAAGCCCCTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGC  
CGAGGAGGAGGTGATGATCCGCTCCGAGAACATCACCAACAACGCCAAGA  
ACATCATCGTGCAGCTGACCAAGCCCCTGAAGATCAACTGCACCCGCCCC  
AACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGGCCAGGCCTTCTA  
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGT  
CCCGCACCGAGTGGAACGAGACCCCTGCAGAAGGTGGCCAAGCAGCTGCGC  
AAGTACTTCAACAACAAGACCATCATCTTCAACCAACTCCTCCGGCGCGCA  
CCTGGAGATCACCAACCCTCCTTCAACTGCGGCGGGCGAGTTCTTCTACT  
GCAACACCTCCGGCCCTGTTCAACTCCACCTGGAACGGCAACGGCACCAAG  
AAGAAGAACTCCACCGAGTCCAACGACACCCATCACCCCTGCCCTGCCGCAT  
CAAGCAGATCATCAACATGTGGCAGCGCGTGGGCCAGGCCATGTACGCC  
CCCCATCCAGGGCGTGATCCGCTGCGAGTCCAACATCACCGGCCCTGCTG  
CTGACCCGCGACGGCGGC GACAACA ACTCCAAGAACGAGACCTTCCGCC  
CGGCGGCGGCGACATGCGCGACA ACTGGCGCTCCGAGCTGTACAAGTACA  
AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACC AAGGCCAAGCGC  
CGCGTGGTGGAGCGGAGAAGCGCGCGTGGGCATCGGCGCCGTGTTCTT  
GGGCTTCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCC  
TGACCGTGCAGGCCCGCCAGCTGCTGTCGGGCATCGTGCAGCAGCAGTCC  
AACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGAAGCTGACCGT  
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGGTGGCCGTGGAGCGCTACC  
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC  
TGCAACCACCAACGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCAGTC  
CGAGATCTGGGACAACATGA CCTGGCTGCAGTGGGACAAGGAGATCTCCA  
ACTACACCGACATCATCTACAACCTGATCGAGGAGTCCAGAACCAGCAG  
GAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGCCAACCTGTG  
GAACCTGGTTCGACATCTCCA ACTGGCTGTGGTACATCAAGATCTTCATCA  
TGATCGTGGGCGGCTGATCGCCCTGCGCATCGTGTTCGCGCTGTCTGTC  
GTGATCAACCCGCGTGCGCCAG GGCTACTCCCCCTGTCTTCCAGACCCA  
CACCCCAACCCCGCGGCCCTGACCCGCCCGGCCGCATCGAGGAGGAGG  
GCGGCGAGCAGGGCCGCGACCCGCTCCATCCGCCTGGTGTCCGGCTTCTG  
GCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCTACCACCG  
CCTGCGCGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCC  
ACTCCTCCCTGAAGGGCCTGCG CCTGGGCTGGGAGGGCCTGAAGTACCTG  
TGGAACCTGCTGCTGTACTGGGGCCGCGAGCTGAAGATCTCCGCCATCAA  
CCTGCTGGACACCATCGCCATCGCCGTGGCCGGCTGGACCGACCGCGTGA  
TCGAGATCGGCCAGCGCATCTGCCGCGCCATCCTGAACATCCCCGCGCG  
ATCCGCCGAGCCCTGAGCCCGCTGAGTCTGAA



**Fig. 19A**

M.con.gag (group M consensus gag. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGA  
 CGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCC  
 TGAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAAC  
 CCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCGGCCAGCT  
 GCAGCCCGCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA  
 CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGCACCC  
 AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAGCAGAA  
 GACCCAGCAGGCCGCCGCGACAAGGGCAACTCCTCCAAGGTGTCCAGA  
 ACTACCCCATCGTGCAGAACCCTGCAGGGCCAGATGGTGCACCAGGCCATC  
 TCCCCCGCACCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCCTT  
 CTCCCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC  
 CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC  
 ATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCG  
 CCTGCACCCCGTGCACGCCGCCCATCCCCCGGCCAGATGCGCGAGC  
 CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC  
 GCCTGGATGACCTCCAACCCCCCATCCCCGTGGGCGAGATCTACAAGCG  
 CTGGATCATCCTGGGCCTGAACAAGATCGTGCATGTACTCCCCGTGT  
 CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG  
 GACCGCTTCTCAAGACCTGCGCGCCGAGCAGGCCACCAGGACGTGAA  
 GAACTGGATGACCACACCCTGCTGGTGCAGAACGCCAACCCGACTGCA  
 AGACCATCCTGAAGGCCCTGGGCCCGGCCACCCTGGAGGAGATGATG  
 ACCGCCTGCCAGGGCGTGGGCGGCCCGGCCACAAGGCCCGGTGCTGGC  
 CGAGGCCATGTCCAGGTGACCAACGCCCATCATGATGCAGCGCGGCA  
 ACTTCAAGGGCCAGCGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAG  
 GGCCACATCGCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAA  
 GTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA  
 ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCGGCAACTTC  
 CTGCAGTCCCGCCCCGAGCCACCGCCCGCCGAGTCTTTCGGCTT  
 CGGCGAGGAGATCACCCCTCCCCAAGCAGGAGCCCAAGGACAAGGAGC  
 CCCCCCTGACCTCCTGAAGTCCCTGTTCCGGCAACGACCCCTGTCCAG  
 TCA

M.con.pol.nuc

28/178

Fig. 19B

GCCGCCGCATGCCCCAGATCACCCCTGTGGCAGCGCCCCCTGGTGACCAT  
 CAAGATCGGCGGCCAGCTGAAGGAGGCCCTGTGGCCACCGGCCCGACG  
 ACACCGTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATG  
 ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT  
 GATCGAGATCTGCGCAAGAAGGCCATCGGCACCGTGCTGGTGGGCCCCA  
 CCCCCGTGAACATCATCGGCCGAACATGCTGACCCAGATCGGCTGCACC  
 CTGAACTTCCCCATCTCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCC  
 CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGA  
 TCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC  
 TCCAAGATCGGCCCGAGAACCCTACAACACCCCATCTTCGCCATCAA  
 GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  
 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCC  
 GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC  
 CTACTTCTCCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCA  
 CCATCCCCCTCCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAAC  
 GTGCTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCTCCAT  
 GACCAAGATCCTGGAGCCCTTCCGCACCCAGAACCCCGAGATCGTGATCT  
 ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG  
 CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGCGCTGGGGCTT  
 CACCACCCCGACAAGAAGCACCAAGAAGGAGCCCCCTTCTGTGGATGG  
 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCATCCAGCTGCCC  
 GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT  
 GAACTGGGCCCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCA  
 AGCTGCTGCGCGGCCAAGGCCCTGACCGACATCGTGCCCTGACCCGAG  
 GAGGCCGAGCTGGAGCTGGCCGAGAACCAGCGAGATCCTGAAGGAGCCCGT  
 GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA  
 AGCAGGGCCAGGACCAAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG  
 AACCTCAAGACCGGCAAGTACGCCAAGATGCGCTCCGCCACACCAACGA  
 CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGTCCATCG  
 TGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCATCCAGAAGGAGACC

TGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATTCCCGAGTG  
 GGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA  
 AGGAGCCCATCGCCGCGCGAGACCTTCTACGTGGACGGCGCCGCAAC  
 CGCGAGACCAAGCTGGGCAAGGCCGGTACGTGACCGACCGCGGCCCA  
 GAAGGTGGTGTCCCTGACCGAGACCACCAACCAGAAAACCGAGCTGCAGG  
 CCATCCACCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC  
 GACTCCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCCGACAAGTCCGA  
 GTCCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG  
 TGTACCTGTCTGGGTGCCCGCCACAAGGGCATCGGCGGCAACGAGCAG  
 GTGGACAAGCTGGTGTCCACCGGCATCCGCAAGGTGCTGTTCTGGACGG  
 CATCGACAAGGCCAGGAGGAGCACGAGAAGTACCACTCCAACCTGGCGCG  
 CCATGGCCTCCGACTTCAACCTGCCCCCATCGTGGCCAAGGAGATCGTG  
 GCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGT  
 GGACTGCTCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA  
 AGATCATCCTGGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAG  
 GTGATCCCCGCGGAGACCGGCCAGGAGACCGCCTACTTCATCCTGAAGCT  
 GGCCGGCCGCTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAAC  
 TCACCTCCGCGCCGCGTGAAGGCCGCTGCTGGTGGGCGGCATCCAGCAG  
 GAGTTCGGCATCCCTACAACCCCCAGTCCAGGGCGTGGTGGAGTCCAT  
 GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGCGCGACCCAGGCCGAGC  
 ACCTCAAGACCGCCGTGCAGATGGCCGTGTTTCATCCACAACCTCAAGCGC  
 AAGGGCCGATCGGCGGCTACTCCGCGGCGAGCGCATCATCGACATCAT  
 CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCC  
 AGAACTTCCGCGTGTACTACCGGACTCCGCGGACCCCATCTGGAAGGGC  
 CCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGTATCCAGGACAA  
 CTCCGACATCAAGGTGGTGGCCCGCCGCAAGGCCAAGATCATCCGCGACT  
 ACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGGCCCGCCAGGACGAG

CA 02539325



29/178

**Fig. 19C**

**M.con.nef (group M consensus nef. Identical amino acid sequence to that in the public domain)**

GCCGCCGCCATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCC  
CGCCGTGCGCGAGCGCATCCGCCACCACCCCGCCGCCGAGGGCGTGG  
GCGCCGTGTC CCAAGACCTGGACAAGCA CGGCGCCATCACCTCTCCAAC  
ACCGCCGCCAACCAACC CCGACTGCGCCTGGCTGGAGGCCAGGAGGAGGA  
GGAGGAGGTGGGCTTC CCGTGC GCCCAGGTGC CCTGCGCCCA TGA  
CCTACAAGGCCGCCCTGGACTGTC CCACTTCCTGAAGGAGAAGGCGGC  
CTGGAGGGCCTGATCTACTCCAA GAAGCGC CAGGAGATCCTGGACTCTGT  
GGTGTAACCA CACCCAGGGCTACTTC CCGACTGGCAGAACTACACCCCG  
GCCC CGGCATCGCTA CCCC TGACTT CCGCTGGTGCTT CAAGCTGGTG  
CCCGTGGACCCCGAGGAGGTGGAGGAGCC AACGAGGGCGAGAACAACCTC  
CCTGCTGCA CCCCATGTG CCAACACGGCATGGAGGACGAGGAGCGCGAGG  
TGCTGATGTGGAAGTTCGACTCCCGCCTGGCCCTGCGCCA CACG CCGC  
GAGCTGCACCCCGAGTACTACAAGGACTGCTAA

**Fig. 19D**

**C.con.pol.nuc**

GCCGCCGCCATGCCCCAGATCACCCCTGTGGCAGCGCCCCCTGGTGTCCAT  
CAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGCCACCGGCGCCGACG  
ACACCGTGTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATG  
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT  
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCCGTGCTGGTGGGCCCCA  
CCCCGTGAACATCATCGGCCGAACATGCTGACCCAGCTGGGCTGCACC  
CTGAACCTCCCCATCTCCCCATCGAGACCGTGC CCGTGAAGCTGAAGCC  
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGA  
TCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATC  
ACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAA  
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCC  
GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGTGCTGGACGTGGGCGACGC  
CTACTTCTCCGTGCCCTGACGAGGGCTTCCGCAAGTACACCGCCTTCA  
CCATCCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAGTACAAC  
GTGCTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCCTCCAT  
GACCAAGATCCTGGAGCCCTTCCGCGCCAGAACCCCGAGATCGTGATCT  
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG  
CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGAAGTGGGGCTT  
CACCACCCCGACAAGAAGCACCAAGGAGGCCCCCTTCTGTGGATGG  
GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCATCCAGCTGCC  
GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT  
GAACTGGGCCTCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA  
AGCTGTGCGCGGCGCCAAGGCCCTGACCGACATCGTGC CCGTACCCGAG  
GAGGCGAGCTGGAGCTGGCCGAGAACC GCGAGATCCTGAAGGAGCCCGT  
GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA  
AGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG  
AACCTCAAGACCGGCAAGTACGCCAAGATGCGCACCCGCCACCAACGA  
CGTGAAGCAGCTGACCGAGGCCCGTGCAGAAGATCGCCATGGAGTCCATCG  
TGATCTGGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAGGAGACC  
TGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATTCCCGAGTG  
GGAGTTCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA  
AGGAGCCCTGCGCTGCTGAGACTTCTGCTGAGGAGCCCGCCCAAC

CGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGCGGGCCGCCA  
 GAAGATCGTGTCCCTGACCCGAGACCACCAACCCAGAAAACCGAGCTGCAGG  
 CCATCCAGCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC  
 GACTCCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGCAAGTCCGA  
 GTCCGAGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAAGGAGCCGG  
 TGTACCTGTCCCTGGTGGCCGCCCAAGGGCATCGGGGCAACGAGCAG  
 GTGGACAAGCTGGTGTCTCCGGCATCCGCAAGTGTCTTCTGGACCG  
 CATCGACAAGGCCAGGAGGACGAGAAAGTACCACTCAACTGGCCGG  
 CCAATGGCCCTCCGAGTTCAACTGCCCCCATCGTGGCCAAAGGAGATCGTG  
 GCCTCCTGGACAAGTGCAGCTGAAGGGGAGGCCATGCACTGGCCAGGT  
 GGACTGTCCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA  
 AGATCATCTCTGGTGGCCGTGCACGTTGGCTCCGGCTACATCGAGGCCGAG  
 GTGATCCCCCGGAGACCCGCCAGGAGACCCGCCCTACTTCATCCTGAAGCT  
 GGCCGGCCGCTGGCCCGTGAAGCCCGCTGTGTTGGCCCGGCATCCAGCAG  
 TCACCTCCGCCCGCTGAAGCCCGCTGTGTTGGCCCGGCATCCAGCAG  
 GAGTTCGGCATCCCTACAACCCCGCTGTGTTGGCCCGGTGGTGGAGTCCAT  
 GAACAAGGAGCTGAAGAAGATCATCGGCCAGTGTGCGGACCCAGGCCGAGC  
 ACCTCAAGACCCCGTGCAGATGGCCGTGTTCATCCACAACCTCAAGCCG  
 AAGGGCGCATCGGGCTACTCCGCCGGCAGGCCATCATCGACATCAT  
 CGCCACCGACATCCAGACCAAGGAGCTGCAGAGCAGATCATCAAGATCC  
 AGAACTTCGGGTGTACTACCGGACTCCCGGACCCCATCTGGAAGGGC  
 CCCGCCAAGCTGTGTGAAGGGCGAGGGCCCGTGTGTATCCAGGACAA  
 CTCCGACATCAAGGTGGTGCCTCCCGCCGCAAGGCCAAGATCATCAAGGACT  
 ACGGCAAGCAGATGGCCCGGCCGACTGCGTGGCCCGGCCAGGACGAG  
 GACTAA

Fig. 19D (continued)

M.con.gag (group M consensus gag)

MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPLLETSEG CKQIIGLOLPA  
 LQTGSEELRSLXNTVATLYCVHQRIEVKDTKEALEKIEEQNKSOQKTOQAAADKGNSSKVSQNYPIVQN  
 LQGQMVHQAI SPRTLNAWVKVIEEKA FSPEV I PMFSA LSEGATPQDLNLTMLNTVGGHQAAMQMLKDTINE  
 EAAEWDRLHPVHAGPI PPGQMPREPRGSDIAGTTSTLQEQI AWMTSNPPI PVGEI YKRWI IILGLNKI VRMY  
 SPVSI LDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNPANPDCKTILKALGPGATLEEMM  
 TACQGVGGPGHKARVLAEAMSQVTNAAIMMQRGNFKGQRRI IKCFNCGKEGHIARNCRAPRKKGCWKCGK  
 EGHQMKDCTERQANFLGKI WPSNKG RPNFLQSRPEPTAPAESFSGFGEETTPSPKQEPKDKPEPPLTSLK  
 SLFGNDPLSQ

Fig. 19E

M.con.pol (group M consensus pol)  
 MPQITLWQRPLVTKIGGQLKEALLATGADDTVLEEINLPKWKPKMIGGIGGFIKVRQYDQILIEICGK  
 KAIGTVLVGPTPVNIIGNMLTQIGCTLNFPIPIETVPVKLPKMGDPKVKQWPLTEEKIKALTEICTE  
 MEKEGKIKIGPENPYNTPFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHAGLKKKKSVTVLD  
 VGDAYFSPVLDDEFRKYAFTIP SINNETPIRQYQNVLPQGWKSPAFQSSMTKILEPFRFTQNP EIM  
 YQYMDL YVGS DLEIGQHRAKIEELREHLRWGFTTPDKKHQKKEPFLVMGYELHFDKWTVPQIQLP EKD  
 SWTVNDIQKLVGKLNWASQIYPGKVKQKCLRGAKALTDIVPLTEEALELAENREILKEP VHGVYYD  
 PSKDLAEIQKQGDQWYTIYQEPFKNLTKGYAKMRS AHTNDVKQLTEAVQKIATESIWIWGKTKFR  
 LPIQKETWETWTEYWAQATWIP EWFVNTPLVYLWYQLEKEPIAGAEITYV DGAANRETKLGKAGYVTD  
 RGROKVSLETNTQKTELQAIHLALQDSGSEVNTDSQYALGIQAQPKDSESELVNIIEQLIKKEK  
 VYLSWVPAHKIGGNEQVDKLVSTGIRKVLFLDGIDKAQEEHEKYHSNWRAMASDFNLPPIVAKEIVASC  
 DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGGKILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPV  
 KVIHTDNGSNFTSAAVKAACVWAGIQDEFIPYNPQSQGVVESHMKNELKIGQVRDQAEHLKTAVQMAV  
 FIHNFRRKGGIGGYSAGERIIDIAITDIQTKELQKQITKIQNFRVYRDSRDPWIKGPAKLLWKGEAW  
 IQDNSDIKAVPRRKAIRDYGKQ MAGDDCVAGRQDED

Fig. 19F

M.con.nef (group M consensus nef)  
 MGKWSKSSI VGWPAVRERIRRTHPAAE VGVAVSQDLDLKHGALTSNTAANNPDCAWLEAQEEEEVEVGFPP  
 VRPQVPLRPMTYKAALDLSHFLKEKGGLEGLIYSKKRQEIIDLWVYHTQGYFPDWNQYTPGPIRYPLTF  
 GWCFKLVVPDPEEVEEANEGENNSL LHPMCQHGMEDEREVIMWKFDLSRLA LRHIA RELHPEYYKDC

Fig. 19G

C.con.pol (subtype C consensus pol)  
 MPQITLWQRPLVSIKVGQIKEALLaTGADDTVLEEINLPKWKPKMIGGIGGFIKVRQYDQILIEICGK  
 KAIGTVLVGPTPVNIIGNMLTQLGCTLNFPIPIETVPVKLPKMGDPKVKQWPLTEEKIKALTAICEE  
 MEKEGKIKIGPENPYNTPFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHAGLKKKKSVTVLD  
 VGDAYFSPVLDDEFRKYAFTIP SINNETPIRQYQNVLPQGWKSPAFQSSMTKILEPFRFTQNP EIM  
 YQYMDL YVGS DLEIGQHRAKIEELREHLRWGFTTPDKKHQKKEPFLVMGYELHFDKWTVPQIQLP EKD  
 SWTVNDIQKLVGKLNWASQIYPGKVKRQKCLRGAKALTDIVPLTEEALELAENREILKEP VHGVYYD  
 PSKDLAEIQKQGDQWYTIYQEPFKNLTKGYAKMRTAHTNDVKQLTEAVQKIAMESIWIWGKTKFR  
 LPIQKETWETWTEYWAQATWIP EWFVNTPLVYLWYQLEKEPIAGAEITYV DGAANRETKIGKAGYVTD  
 RGROKVSLETNTQKTELQAIHLALQDSGSEVNTDSQYALGIQAQPKDSESELVNIIEQLIKKER  
 VYLSWVPAHKIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAMASEFNLPPIVAKEIVASC  
 DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGGKILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPV  
 KVIHTDNGSNFTSAAVKAACVWAGIQDEFIPYNPQSQGVVESHMKNELKIGQVRDQAEHLKTAVQMAV  
 FIHNFRRKGGIGGYSAGERIIDIAITDIQTKELQKQIIONFRVYRDSRDPWIKGPAKLLWKGEAW  
 IQDNSDIKAVPRRKAIRDYGKQ MAGADCVAGRQDED

Fig. 19H

32/178

*Fig. 20A*

**B.con.gag (subtype B consensus gag. The amino acid sequence is different from Los Alamos Database August 2002)**

GCCGCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCGAGCTGGA  
 CCGCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACAAGC  
 TGAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAAC  
 CCCGGCCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCCTGGGCCAGCT  
 GCAGCCCTCCCTGCA GACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA  
 CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACC  
 AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAA  
 GGGCCAGCAGGCCCGCCCGACACCGGCAACTCCTCCCAGGTGTCCCAGA  
 ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATC  
 TCCCCCGCACCCCTGAACGCCCTGGGTGAAGGTGGTGGAGGAGAAGGCCTT  
 CTCCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGGCGCCACC  
 CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC  
 ATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCCGCCGAGTGGGACCG  
 CCTGCACCCCCGTGCACGCCGGCCCCATCGCCCCGGCCAGATGCGCGAGC  
 CCCGCGGCTCCGACATCGCCGGCACCACTCCACCCTGCAGGAGCAGATC  
 GGCTGGATGACCAACAACCCCCCATCCCCGTGGGCGAGATCTACAAGCG  
 CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCACCT  
 CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG  
 GACCGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCCTCCAGGAGGTGAA  
 GAACTGGATGACCGAGAC CCTGCTGGTGCAGAACGCCAACCCCGACTGCA  
 AGACCATCCTGAAGGCCCTGGGCCCCCGCCACCCTGGAGGAGATGATG  
 ACCGCCCTGCCAGGGCGTGGGCGGCCCGCCACAAGGCCCGCGTGTGGC  
 CGAGGCCATGTCCAGGTGACCAACTCCGCCACCATCATGATGCAGCGCG  
 GCAACTTCCGCAACCAGCGCAAGACCGTGAAGTGCTTCAACTGCGGCAAG  
 GAGGGCCACATCGCCAAGA ACTGCCGCCCCCCCGCAAGAAGGGCTGCTG  
 GAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG  
 CCAACTTCTGGGCAAGATCTGGCCCTCCCAACAAGGGCCGCCCGGCAAC  
 TTCTGCAGTCCCGCCCCGAGCCCAACCGCCCCCCCCGAGGAGTCTTCCG  
 CTTCGGCGAGGAGACCACACCCCTCCAGAAGCAGGAGCCCATCGACA  
 AGGAGCTGTACCCCTGGCCTCCCTGCGCTCCCTGTTTCGGCAACGACCCC  
 TCCTCCAGTAA

33/178

**Fig. 20B**

**B.con.env (subtype B consensus env. The amino acid sequence is different from Los Alamos Database August 2002)**

GCCG CCGCCAT GCGCG TGAAGGG CATCC GCAAGAA CTACCAGCAC CTGTG  
 GCGCTGGGG CACCATG CTGCTGGGCATG CTGATGATCTGCTCCGC CGCCG  
 AGAAGCTGTGGGTGACCGTGTA CTACGG CGTGCCCGTGTGGAAGGAGGCC  
 ACCA CCACC CTGTTCTGCGCTC CGACG CCAAGGC CTACGACACCGAGGT  
 GCACAA CGTGTGGGCCACCCA CGCCTGCGTGCCCA CCGAC CC CAA CCCCC  
 AGGAGGTGGTGTCTGGA GAACGTGAC CGAGA ACTTCAA CATGTGGAAGAAC  
 AACATGGTGGAGCAGA TGACGAGGACATCATCTC CCTGTGGGAC CAGT C  
 CCTGAAGC CCTGCGTGAAGCTGA CCCCCTGTGCGTGACC CTGAA CTGCA  
 CCGA CCTGAAGAACAA CCTGCTGAA CAC CAACTCTCTC CGGCGAGAAG  
 ATGGAGAAGGG CGAGA TCAAGAA CTGCT CCTTCAA CATCA CCACC TC CAT  
 CCGCGA CAA GGTGCAGAAGGATACGCC CTGTTCTACAAG CTGGA CGTGG  
 TGCC CATCGACAACAA CAACAACAC CTC CTACCGC CTGAT CTCTGCAAC  
 ACCT CCGTGATCAC CCAGG CCTGC CC CAAGGTGTC CTTCGAG CCCATCCC  
 CATC CACTA CTGCGCC CCGC CGGCTTC GC CATCC TGAAG TG CAA CGACA  
 AGAAGT TCAACGGCAC CGGCC CCTGC CAC CAACGTG TCCAC CGTGCAGTGC  
 ACCACGGCAT CCGCC CCGTGGTGT CCA CC CAGCTGCTGC TGAACGGCTC  
 CCTG GC CGAGGAGGAGGTGGT GATC CGCTC CGAGA ACTTAC CGA CAACG  
 CCAAGACCAT CATCGTG CAG CTGAA CGAGTCCGTGGAGAT CAACTGCACC  
 CGCC CCAACAA CAA CA CCGCAAGT CCA TC CACAT CGGCC CCGGC CGCGC  
 CTTC TACAC CA CCGGC GAGAT CATCGGC GA CATCCGCCAGGC CCA CTGCA  
 ACATCTCCCGCCAA GTGGAACAA CAC CCTGAAG CAGAT CGTGAAGAAG  
 CTGCGCGAG CAGTT CGGCAACAAGA CCA TC GTGTT CAACCAGTCC TC CGG  
 CGGCGA CCC CGAGATCGTGATGCACT CCTTCAA CTGCGGC GG CGAGTCT  
 TCTA CTGCAACACCAC CCAGCTGTT CAA CT CCACTGGAA CGACAACGGC  
 ACCTGGAACAA CACCAAGGACAAGAACA CCATCAC CCTGC CCTGC CGCAT  
 CAAG CAGAT CATCAACATGTGGCAG GAGGTGGGCAAGGCCAT GTA CGCCC  
 CCCC CATCCGCGGC CAGATCCGCTG CTC CT CCAACATCAC CGGCC TGCTG  
 CTGAC CCGCGAC GCGCG CAACAACAACAA CGACAC CGAGA TC TTC CGCCC  
 CGGC GG CGG ACATG CGCGA CAAC TGG CG CTC CGAG CTG TA CAA GTACA  
 AGTGGTGAAGATCGA GCCCTGGG CGTGG CCCCCAC CAAGG CCAAG CGC  
 CGCG TGGTG CAGCG CGAGAAG CG CG CCGTG GGCAT CCGCG CCATG TT CCT  
 GGGCTT CCTGGGCGCCGCCGG CTCCACC ATGGGCG CCGCC TC CATGA CCC  
 TGAC CTG CAGGC CCGCCAGCTG CTGTC CGGCATCGTG CAGCAGCAGAAC  
 AACCTGCTG CG CGCCA TCGAGGC CCAGCAG CACCTGCTGCAG CTGAC CGT  
 GTGGGCAT CAAGCAG CTGCAGG CC CGC GTGCTGG CCGTGGAGCG CTACC  
 TGAAGGACCAG CAGCTGCTGGGCAT CTGGGCTGCTCCGG CAAGCTGATC  
 TGCA CCACCAC CGTGC CCTGGAA CG CCT CCTGGTC CAACAAGTCC CTGGA  
 CGAGATCTG GGA CAACATGACCTG GATGGAGTGGAGCGCGAGAT CGACA  
 ACTA CACCTCC CTGAT CTACA CCTGATCGAGGAGTCC CAGA ACCAG CAG  
 GAGAAGAAC GAGCAGGAGCTG CTGGAGCTGGACAA GTGGG CCTCC CTGTG  
 GAAC TGGTT CGACATCACCAA CTGG CTG TG GTACA TCAAGAT CTT CATCA  
 TGATCGTGGGCGGCCCTGATCGGC CTGCG CATCGTGTTCGC CGTGC TGTC  
 ATCGTGAA CCGCGTG CGCCAGGGCTACTCC CCCCCTGTCTCTC CAGAC CCG  
 CCTGCCCGCC CCGCGGGCCC CGAC CGC CC CGAGGGCATCGAGGAGGAGG  
 GCGGCGAGCGGACCG CGACCGCTC CGGCCCTGGTGGACGGCTTC CTG  
 GCCCTGATCTGGACGACCTG CGCT CCC TGTCCTGTTCTCTTAC CACCG  
 CCTGCGCGA CCTGCTG CTGATCGTGACC CG CATCGTGGAG CTGCTGGGC  
 GCCCGG CTGGGAGGTGCTGAAGTACTG GTGGAAC CTGCTGCAGTACTGG  
 TCCCAGGAGCTGAAGA ACTCCGCGTGTCC CTGCTGAACG CCACC GC CAT  
 CGCCGTGGC CGAGGGCACCGA CCGCGTGAT CGAGGTGGTG CAGCG CGCCT  
 GCCCGCCA TCCTGCA CATCC CCGCG CATCCGC CAGGG CCTGGAG CGC  
 GCCCTGCTGTAA

33/178

**Fig. 20B**

B.con.env (subtype B consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCG CCGCCATGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTG  
 GCGCTGGGGCACCATGCTGCTGGGCATGCTGATGATCTGCTCCGC CGCCG  
 AGAAGCTGTGGGTGACCGTGTACTACGGCGTGC CCGTGTGGAAGGAGGCC  
 ACCA CCACCCTGTTCTGCGCCTCGACG CCAAGGCCTACGACACCGAGGT  
 GCACAA CGTGTGGGCCACCCACGCTTGC GTGCCCA CCGACCCCAA CCCC  
 AGGAGGTGGTGCTGGA GAACGTGACCGA GAACTTCAACATGTGGAAGAAC  
 AACATGGTGGAGCAGATGCACGAGGACATCATCTCCTGTGGGAC CAGTC  
 CCTGAA GCCTGCGTGAAGCTGACCC CCGTGTGCGTGACCCTGAACTGCA  
 CCGA CTTGAAGAACAA CCTGCTGAA CACCAACTCCTCCTCGGCGAGAAG  
 ATGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAA CATCA CCACTCAT  
 CCGCGACAAGGTGCAGAAGGAGTACGCCCTGTTCTACAAGCTGGA CGTGG  
 TGCCCATCGACAACAA CAACAACACCTCCTACCGCCTGATCTCCTGCAAC  
 ACCTCCGTGATCACCCAGGCC TGC CCAAGGTGTCCTTCGAGCCCATCCC  
 CATCACTA CTGCGCC CCGCGCGGCTTCGC CATCTGAAGTGCAA CGACA  
 AGAAGTTC AACGGCACCGGCCCTG CACCAACGTGTCCACCGTGCAGTGC  
 ACCACGGCATCCGCC CCGTGGTGTCCA CCAAGCTGCTGCTGAACGGCTC  
 CCTGGCCGAGGAGGAGGTGGTGA TC CGCTCGAGA ACTTACCGA CAACG  
 CCAAGACCATCATCGTGCAGCTGAA CGAGTCCGTGGAGATCAACTGCACC  
 CGCC CCAACAA CAACA CCGCAAGTCCA TC CACATCGGCC CCGGC CGCGC  
 CTCTACAC CACCGGC GAGATCATCGGC GACATCCGCCAGGCCA CTGCA  
 ACATCTCCCGCGCCAA GTGGAACAA CACCTGAAGCAGATCTGAAGAAG  
 CTGCGCGAGCAGTTCGGCAACAAGACCA TC GTGTT CAACCAGTCTCCGG  
 CGGC GACCCGAGATCTGTATGCACTCCTTCAA CTGCGGC GGCGAGTCT  
 TCTA CTGCAACACCACCCAGCTGTTCAA CTCCA CTTGGAA CGACAACGGC  
 ACCTGGAACAA CACCAAGGACAAGAACA CCATCAC CCGTGCCTGCGCAT  
 CAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTA CGCCC  
 CCCC CATCCGCGGC CAGATCCGCTGCTCCTCCAACATCACCGGCCTGCTG  
 CTGACCCGCGACGGCGG CAACAACAACAA CGACACCGAGATGTTTCGCC  
 CGGC GGCGGCGACATGCGCGA CACTGGCGCTCCCGACCAAGGCCAAGCGC  
 CGCGTGGTGCAGCGCGAGAAGCGCG CCGTGGGCATCGGCGCCATGTTCT  
 GGGCTTCTGGGCGCCGCCGGCTCCACCATGGGCG CCGCCTCATGACC  
 TGACCGTGCAGGC CCGCCAGCTGCTGTC CGGCATCGTGCA GCAGCAGAAC  
 AACCTGCTGCGCGC ATCGAGGC CAGCAGCACCTGCTGCAGCTGACCGT  
 GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGG CCGTGGAGCGCTACC  
 TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC  
 TGCA CCAACACCGTGCCTGGAA CGCCTCTGGTCAACAAGTCCCTGGA  
 CGAGATCTGGGA CAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACA  
 ACTACACTCTGATCTACACCCTGATCGAGGAGTCCAGAACAGCAG  
 GAGAAGAACGAGCAGGAGCTGCTGGAGCTGACAAAGTGGGCTCCCTGTG  
 GAACTGGTTCGACATCACCAACTGGCTGTGTA CATCAAGATCTTCA TCA  
 TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGC CGTGTGTC  
 ATCGTGAA CCGCGTGC CGCAGGGCTACTCC CCGTGTCTTC CAGACCCG  
 CCTGCCCGCC CCGCGCGGCCCGACCGCC CCGAGGGCATCGAGGAGGAGG  
 GCGGCGAGCGCGACCGCGACCCTCGGCCGCTGGTGGACGGCTTCCTG  
 GCCCTGATCTGGGACGACCTGCGCTCCCTGTGCTGTCTCTTAC CACCG  
 CCTGCGCGA CTTGCTGCTGATCGTGACC CGCATCGTGGAGCTGCTGGGC  
 GCCGCGGCTGGGAGGTGCTGAAGTACTGGTGGAACCTGCTGCAGTACTGG  
 TCCAGGAGCTGAAGA ACTCCGC CGTGTCCCTGCTGAACG CCAACGCAT  
 CGCCGTGGC CGAGGGCACCGA CCGCGTGTGATCGAGGTGGTGCAGCGCGCT  
 GCCGCGCATCTGCA CATCCCGCGCATCCGC CAGGGCCTGGAGCGC  
 GCCCTGCTGTAA

34/178

Fig. 20C

**B. con.gag (subtype B consensus gag)**

MGARASVLSGGELDRWEKIRL RPPGKKKYLKHI VWASRELERFAVNPGLLE TSEGCRQILGQLQPSLQ T  
 GSEELRSLYNTVATLYCVHORIEVKDTKEALEKIEEEQNKSKKKAQAAADTGNSSQVSQNYPIVQN LQG  
 QMVHQAI SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLN TMLN TVGGHQAAQM LKETINEEAA  
 EWDRLHPVHAGPIAPGQMRPRGSDIAGTTSTLQEQIGWMTNPPPIPVGEIYKRWIILGLNKIV RMYSP T  
 SILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTAC  
 QGVGGPHKARVLAEMSQV TNSATIMMQRGNFRNQRKTVKCFNCGKEGHI AKNCRAPRKKGCWKCKGKEG  
 HQMKDCTERQANFLGKIWP SHKGRPGNF LQSRPEPTAPPEESFRFGEETTTPSQQEPIDKELYPLASLR  
 SLFGNDPSSQ

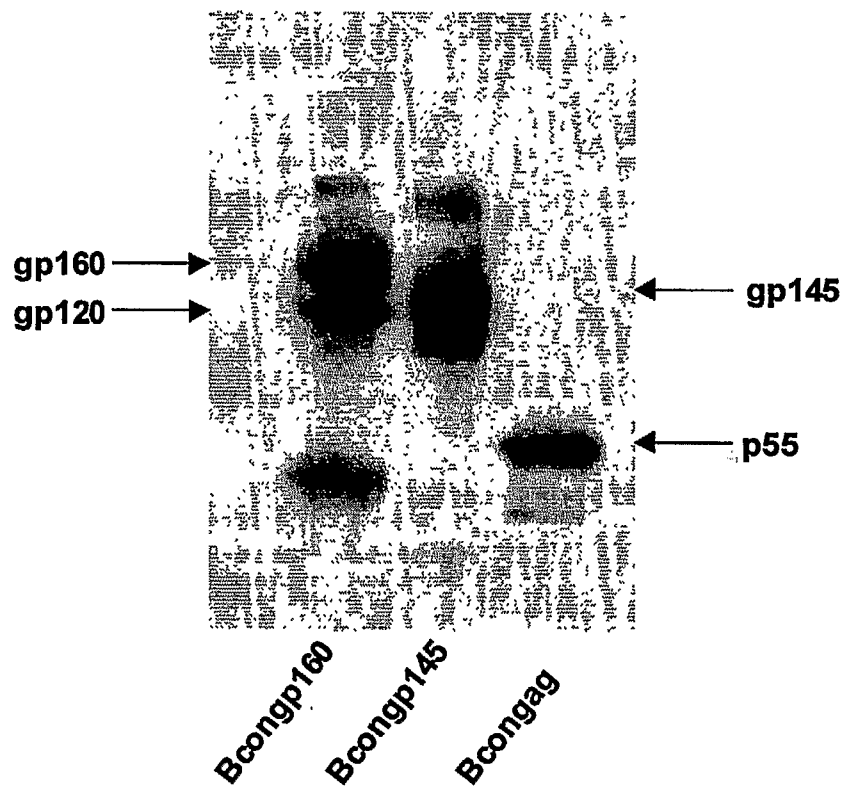
Fig. 20D

**B. con.env (subtype B consensus env)**

MRVKGIRKNYQHLWRWGTMLLGLMLMI CSAAEKLVTV YGVVPVWKEATTLFCASDAKAYDTEVHNVWAT  
 HACVPTDPNPQEVVLENV TENFNWKNMVEQM HEDIISLWDQSLKPCVKLTPLCVTLNCTDLKNNLLNT  
 NSSGKMEKGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNTSYRLIS CNTSVITQACPKVSF  
 EPIPIHYCAPAGFALIKCNDKKFNGTGPC TNVSTVQCTHGIRFPVSTQLLLNGSLAE EEVVIRSENF TDN  
 AKTIIVQLNESVEINCTRPNNTRKSIHIGPGRAFYTTGEIIGDIRQAH CNISRAKWNNTLKQIVK KLR E  
 QFGNKTIVFNQSSGGDPEI VMHSFNCGGEFFYCN TQLFNSTWNDNGTWNNTKDKNTITLPCRIKQI INM  
 WQEVGKAMYAPP IIRGOIRCSNITGLLLTRDGGNNNDTEIFRPGGDMRDNRSEL YKYKVVKIEPLGV  
 APTKAKRRVVQREKRAVGI GAMFLGFLGAAAGSTMGAASMTLTVQARQLLSGIVQQNNLLRAIEAQQHLL  
 QLTVWGIKQLQARVLAVERYLKQQQLLGIWGC SGKLICTTFVPWNASWSNKS LDEIWDNMTWMEWEREID  
 NYTSLIYTLIEESQNOQEKNEQELLELDKWASLWNWFDI TNWLWYIKIFIMI VGG LIGLRIVFAVLSIVN  
 RVRQGYSPLSFQTRLPAPRGPDRPEGIEEEGGERDRDRSGRLVDGFLALIWDDLRS LCLFSYHRLRDL LLL  
 IVTRIVELLGRRGWEVLKYWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQRACRAILHIPRR  
 IRQGLERALL

35/178

Fig. 21

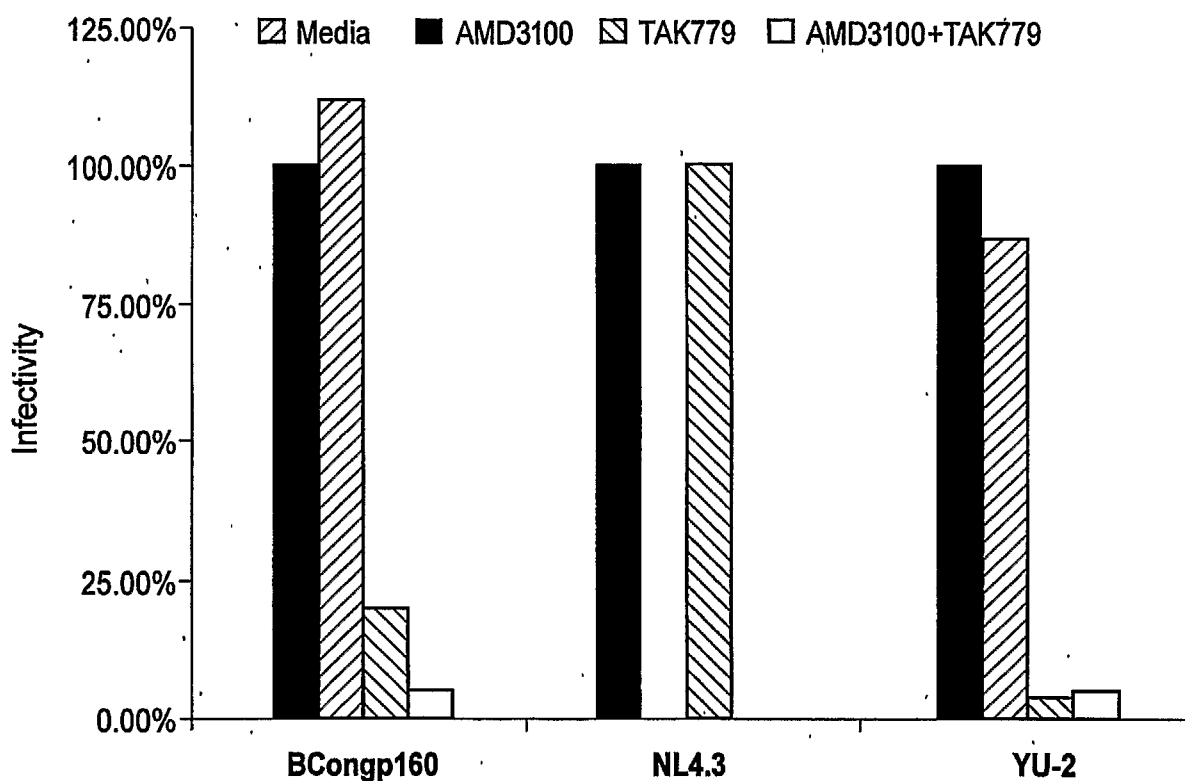


Expression of subtype B consensus *env* and *gag* genes in 293T cells. Plasmids containing codon-optimized subtype B consensus *gp160*, *gp140*, and *gag* genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2  $\mu$ g of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.



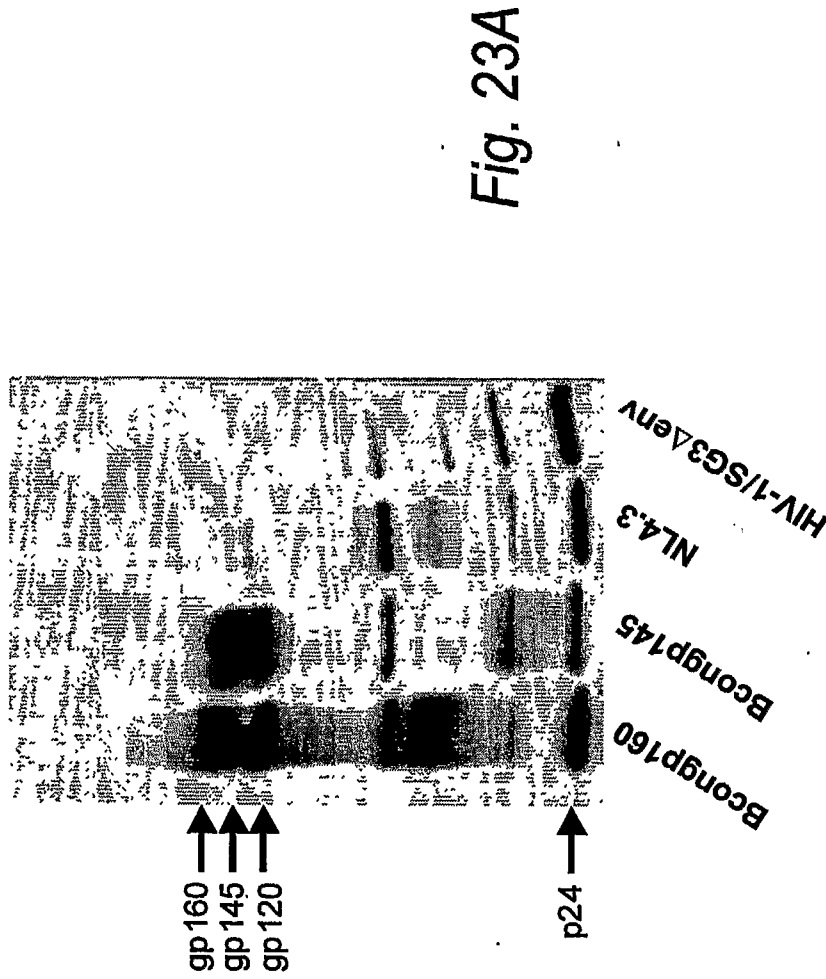
36/178

Fig. 22

**Co-receptor usage of subtype B consensus envelopes.**

Pseudotyped particles containing the subtype B consensus gp160 Env were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4 and YU-2, a known CCR5-using isolate; were included as controls.

37/178

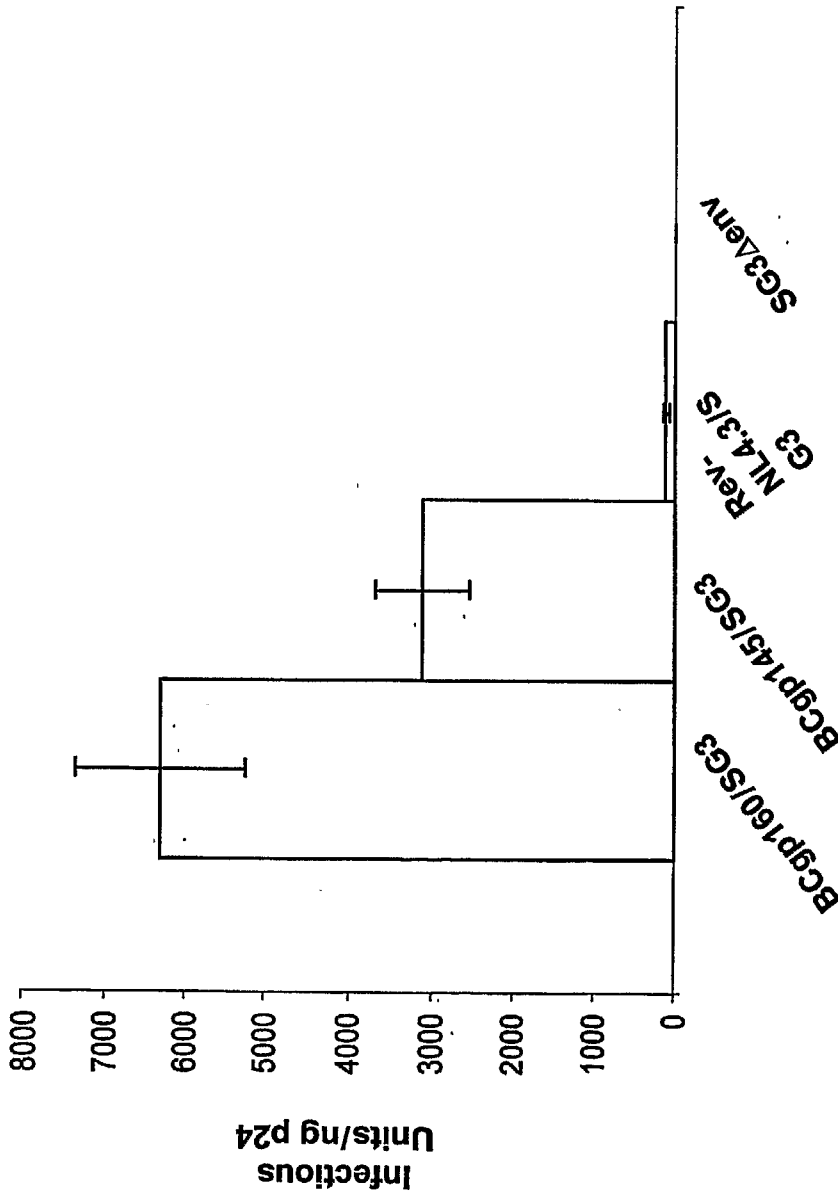


**Trans complementation of env-deficient HIV-1 with codon-optimized subtype B consensus gp160 and gp140 genes.**

Plasmids containing codon-optimized, subtype B consensus gp160 or gp140 genes were co-transfected into 293T cells with an HIV-1/SG3Δenv provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a tabletop centrifuge, filtered through a 0.2μM filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. Trans complementation with a rev-dependent NL4.3env was included for control.

38/178

Fig. 23B



**Infectivity of virus particles containing the subtype B consensus envelope.**

Infectivity of pseudotyped virus containing consensus B gp160 or gp140 was determined using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize  $\beta$ -galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

39/178

Fig. 24B

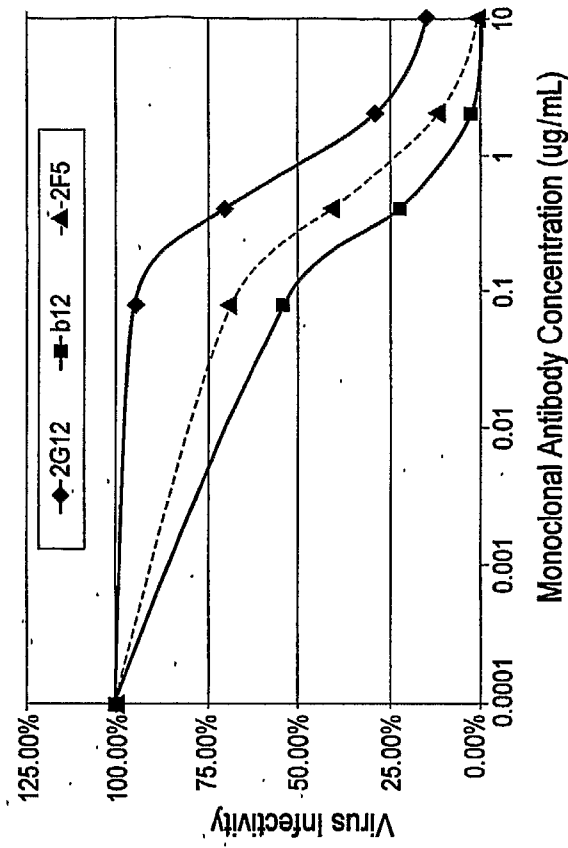


Fig. 24A

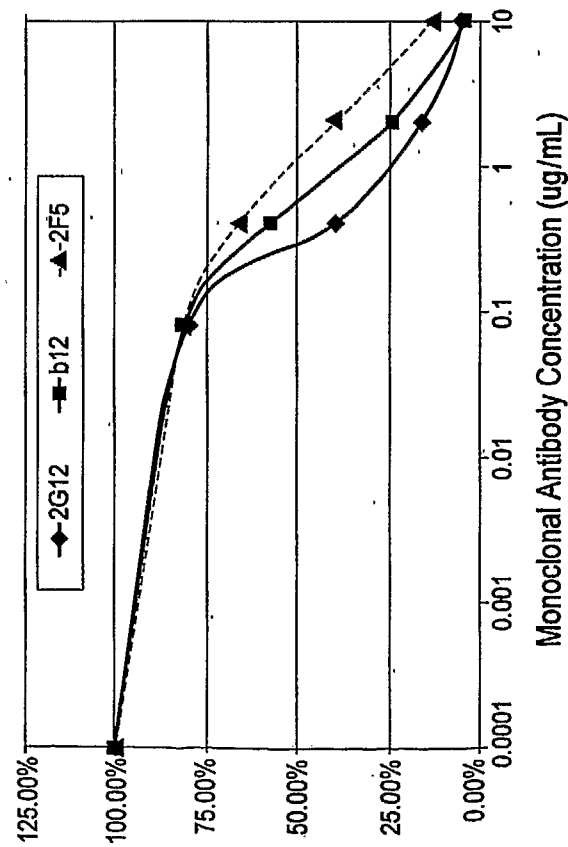
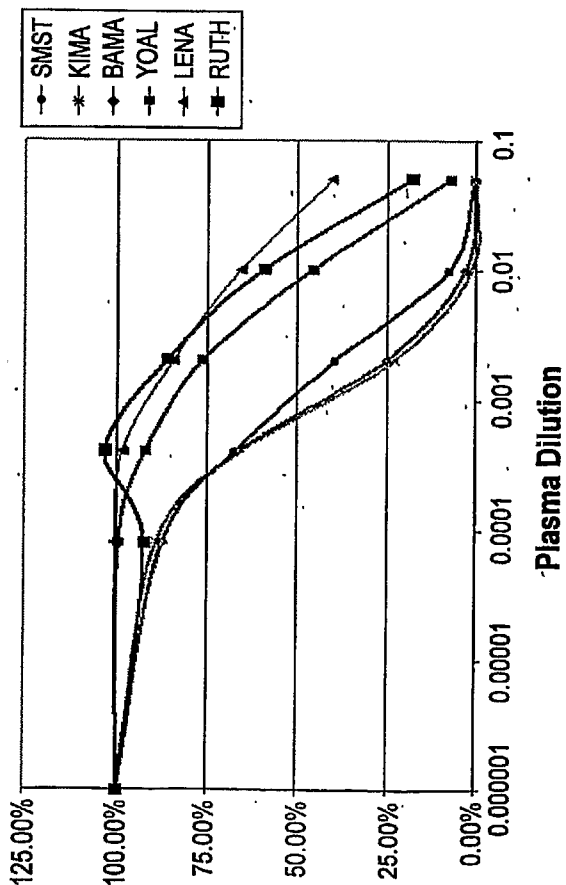
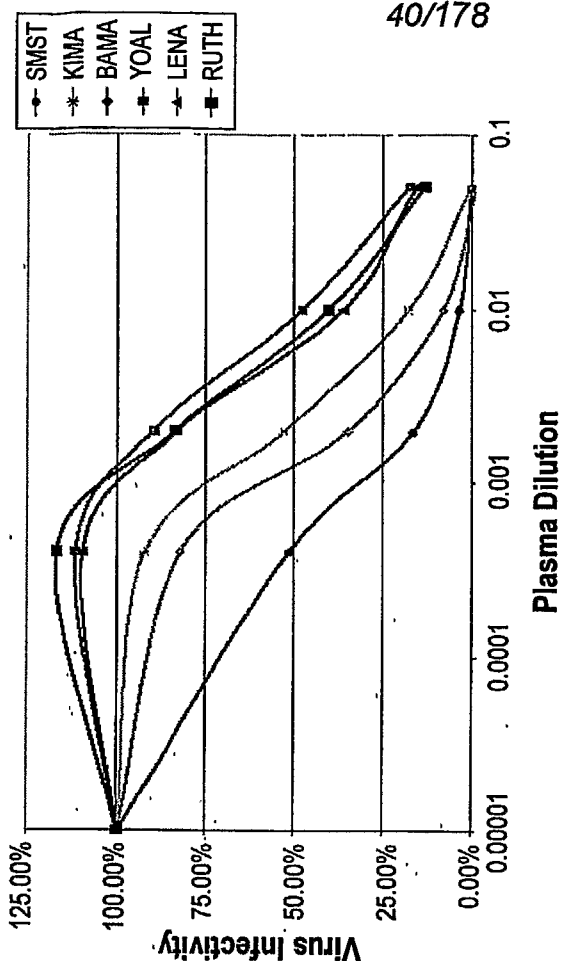


Fig. 24C



Neutralization of Pseudovirions containing Subtype B consensus Env (gp160)

Fig. 24D



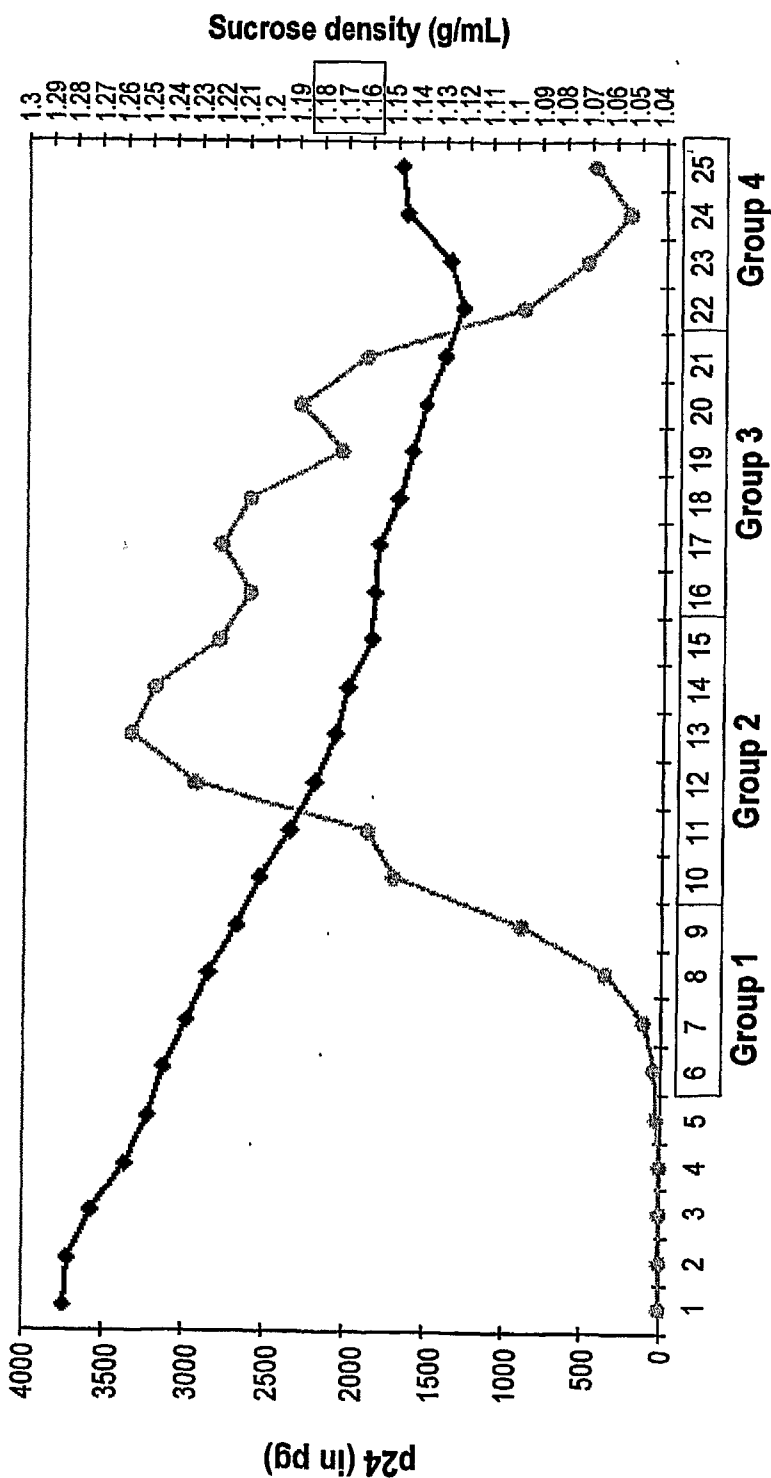
Neutralization of Pseudovirions containing NL4.3 Env (gp160)

**Neutralization sensitivity of virions containing subtype B consensus gp 160 envelope.**

Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1 subtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC<sub>50</sub>) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections.

41/178

Fig. 25A



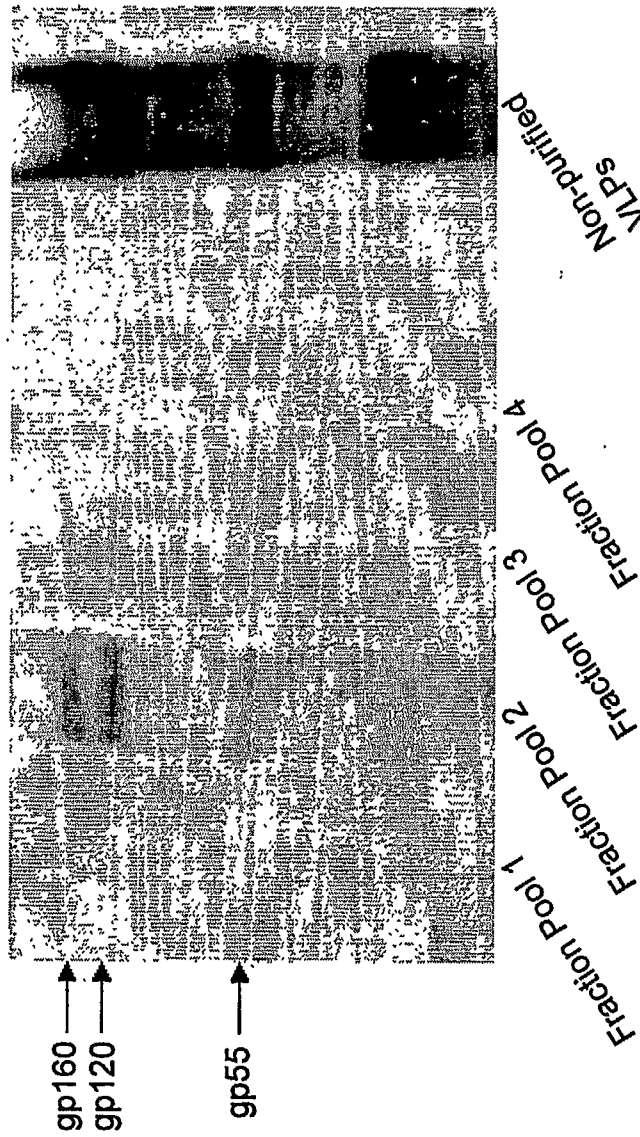
Fractions (0.5 mL increments)

**Density and p24 analysis of sucrose gradient fractions.**

0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml.

42/178

Fig. 25B



**VLP production by co-transfection of subtype B consensus gag and env genes.**

293T cells were co-transfected with subtype B consensus gag and env genes. Cell supernatants were harvested 48-hours post-transfection, clarified through a 20% sucrose cushion, and further purified through a 20-60% sucrose gradient. Select fractions from the gradient were pooled, added to 20ml of PBS, and centrifuged overnight at 100,000 x g. Resuspended pellets were loaded onto a 4-20% SDS-PAGE gel, proteins were transferred to a PVDF membrane, and probed with plasma from an HIV-1 subtype B infected individual.

43/178

**Fig. 26A****Year 2000 Con-S 140CFI.Env**

MRVRGIQRNCQHLWRWGTLILGMLMICSAEAENLWVTVYVYGVFPVWKEANTTLFCASDAKAYDTEVH  
 NVWATHACVPTDPNPQEIIVLENTENFNMWKNNMVEQMEDIISLWDQSLKPCVKLTPLCVTLNC  
 TNVNVNTNTNNTTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLINCNT  
 SAITQACPKVSFEPPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGKIPVVSTQLLLNG  
 SLAEEELIIRSENI.TNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQA  
 HCNI SGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFNCRGFEFFYCNTSGLFNSTW  
 IGNGTKNNNNNTNDTITLPCRKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTN  
 ETEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTKAKLTVQARQLLSGIVQQQSNLLRAIEAQ  
 QHLLQLTVWGIKQLQARVLAVERYLKDQOLEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEK  
 NEQELLALDKWASLWNWFDITNWLW

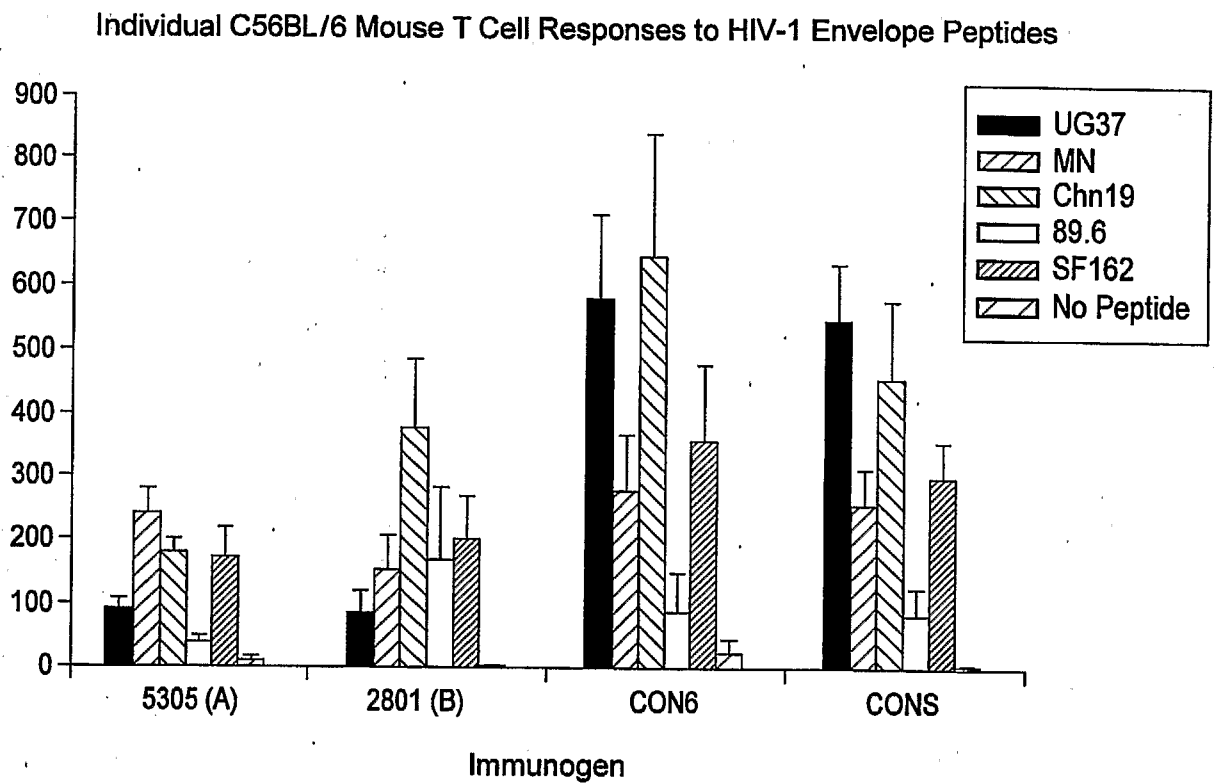
A gp140 CFI is referred to HIV-1 envelope design with the cleavage-site-deleted (C), fusion-site-deleted (F) and gp41 immunodominant region-deleted (I) in addition to the deletion of transmembrane and cytoplasmic domains.

**Fig. 26B****Codon-optimized Year 2000 Con-S 140CFI. seq**

ATGCGCGTGC GCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCCTGATCCTGGG  
 CATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGT  
 GGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCCTACGACACCGAGGTGCAC  
 AACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAA  
 CGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCT  
 CCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACCTGC  
 ACCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTC  
 CTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGG  
 ACGTGGTGCCATCGACGACAACAACAACACTCCTCCAACCTACCGCCTGATCAACTGCAACACC  
 TCCGCCATCACCAGGCCCTGCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCC  
 CGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCCTGCAAGAAGC  
 TGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGC  
 TCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCACAACGCCAAGACCATCAT  
 CGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCA  
 TCCGCATCGGCCCCGGCCAGGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCC  
 CACTGCAACATCTCCGGCACCAAGTGAACAAGACCCCTGCAGCAGGTGGCCAAGAAGCTGCGCGA  
 GCACCTCAACAACAAGACCATCATCTTCAAGCCCTCCCTCCGGCGGGCAGCCTGGAGATCACCACC  
 ACTCCTTCAACTGCCGCGGCGAGTTCCTTCTACTGCAACACCTCCGGCCCTGTTCAACTCCACCTGG  
 ATCGGCAACGGCACCAAGAACAACAACACCAACGACACCCATCACCCTGCCCTGCCGCATCAA  
 GCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGA  
 TCACCTGCAAGTCCAACATCACCAGCCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAAC  
 GAGACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAA  
 GTACAAGGTGGTGAAGATCGAGCCCCCTGGGCGTGGCCCCACCAAGGCCAAGCTTACCGTGCAGG  
 CCCGCCAGCTGCTGCTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGCGCCATCGAGGCCAG  
 CAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCCGTGG  
 GCGCTACCTGAAGACCAGCAGCTCGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCG  
 AGATCAACAACCTACACCGACATCATCTACTCCCTGATCGAGGAGTCCAGAACCAGCAGGAGAAG  
 AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCCTCCCTGTGGAACCTGGTTCGACATCACCA  
 CTGGCTGTGGTGAGGATCC



44/178

*Fig. 27*

**Fig. 28A**

**Design of expression-optimized HIV-1 envelope gp140CF**

**Con-B-2003 Env.pep (841 a.a.) \***

MRVKGIRKNYQHLLWRWGTMLLGMLMICSAAEKLWVTVYGYGVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVL  
 ENVTENFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTIIYRWRGEIKNCSEFNITTSIRDKVKQKEY  
 ALFYKLDVVPIDNDNTSYRLISCSNTSVITQACPVSFEPIPIHYCAPAGFAILKNDKFKFNGTGPCNTVSTVQCTHGIRPVVSTQ  
 LLLNGSLAEVEEVVIRSENFTDNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGRAFYTTGEIIGDIRQAHNCISRAKWNNTLKO  
 IVKKLREQFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNGTWNTEGNTILPCRIKQIINMWQEVGKAMYAPP  
 IRGQIRCSSNITGLLLTRDGGNNETEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKARRVVOREKRAVIGIGAMFLGFLGA  
 AGSTMGAASMTLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSSGKLICTTAVPW  
 NASWSNKSLEIWDNMTWMEWEREIDNYTSLIYTLIEESQOQEKNEQELLELDKWAASLWNWFDITNWLWYIKIFIMIVGGLVGL  
 RIVFAVLSIVNRVROGYSPLSFQTRLPAAPRGPDRPEGIEEGGERDRSRLVDGELALIWDDLRLSLCLFSYHRLRDLRLLLIVTR  
 IVELLGRRGWEVLKYWNLLQYWSQELKNSAVSLNATAIAVAEGTDRVIEVVQACRAILHIPRRIRQGLERALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

**Fig. 28B**

**Con-B-140CF.pep (632 a.a.)**

**Nick name: 002**

MRVKGIRKNYQHLLWRWGTMLLGMLMICSAAEKLWVTVYGYGVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVL  
 ENVTENFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTIIYRWRGEIKNCSEFNITTSIRDKVKQKEY  
 ALFYKLDVVPIDNDNTSYRLISCSNTSVITQACPVSFEPIPIHYCAPAGFAILKNDKFKFNGTGPCNTVSTVQCTHGIRPVVSTQ  
 LLLNGSLAEVEEVVIRSENFTDNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGRAFYTTGEIIGDIRQAHNCISRAKWNNTLKO  
 IVKKLREQFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNGTWNTEGNTILPCRIKQIINMWQEVGKAMYAPP  
 IRGQIRCSSNITGLLLTRDGGNNETEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKAKTLTVQARQLLSGIVQOQNNLLRA  
IEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSSGKLICTTAVPWNASWSNKSLEIWDNMTWMEWEREIDNYTSLIY  
TLIEESQOQEKNEQELLELDKWAASLWNWFDITNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

# Fig. 28C

Codon-optimized Con-B 140CF.seq (1927 nt.)

Nick name: 002

TTCAGTCGACGGCCACCATGAGGGTGAAGGGTATTCGGAAAAAATTACCAACACCTGTGGCGTGGGGAACCATGCTCCTTGGTAT  
 GCTGATGATTTGCAGTCCCGCCGAGAAACTTTGGGTAACGTGTACTACGGCGTTCTGTGTGGAAGGAAGCTACAACCACACTCTT  
 TTTTGTGCATCCGACCGTAAGCTTACGACACAGAAAGTGCATAAATGTTTGGCCACCCATGCTTGGTCCCTACAGATCCCAACC  
 CCCAGGAAGTCTGCTTGGAAATGTACAGAGAAATTTAAACATGTGGAAGAAATAATGTTAGAACAAATGCACGAAGACATAT  
 TAGCCTGTGGGACCCAGTCTTGAAGCCCTGCGTGAACACTCACTCCACTTCCGTCACTTAAGTACTGATTTGATGAACGCCA  
 ACCAACACAAATACTACTATATATAATCGCTGGAGGGGGAATCAAGAACTGCTCTTCAACATCACCACTTCCATAGGGATA  
 AGGTCAGAAAGAATAATGCCCTGTTTATAAAGTGTGATGTGGTCCCGATAGACAATGACAACACTAGCTATCGACTGATCTCTTG  
 TAACACATCCGTGATTAACCAAGTTCGCAAAAGTCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAAT  
 ATCCTCAAGTGAACGACAAAATAATCAATGGGACCGGACCTTGACAAAAGTGTCCACCGTGCATCTCCGCTCCGCTGGTTTGCC  
 CTGTTGTCAGTACCCAACTCCTTGAACGGTCTCTCGCGGAAGAGGAGTCTGATGATTAGAACGGAATACTTACCGATAACGC  
 TAAACAATCATTTGCAACTTAATGAAGCGTCAAAATTAAGTCAACAGCAACCAAAATAATCAAGAAATCTATTCACATA  
 GGGCCCGCCGGCATTTATACAACCTGGGAAATCATTTGGTGCATCAGACAAGCTCATTTGCAATATCTCCGCGGAAATGGA  
 ACAACACCTGAACAGATCGTGAAGAACTTCGAGAACAAATTCGGTATAAACAATCGTATTCACCAAAAGCTCCGGAGGCGGA  
 CCTGAGATAGTTATGCACTCATTTCAACTGTGGCGGAGTCTTCTATTGTAACACAACTCAACTTTTAAATAGCACTTGGAAAT  
 GGAACATGGAACAACAACAAGGAAACATCACTCTGCTTGTCCGATTAAGCAGATCATTAATATGTGGCAAGAAAGTGGGAAAAG  
 CTATGTACGCCCCCTATTTCGGGACAAATAAGATGCTCTAGTAATAATACCGGATTTGTTGCTGACACGCGGAGGAAATAA  
 TGAACAGAGATAATTTAGACCTGGCGGAGCGGACATGAGAGATAACTGGAGAAAGTGCAGCTTTACAAATAATAAAGTCGTAAGATA  
 GAACCATTTGGGGTAGCAACCAAGCAAAACCTTGACAGTACAGGTAGGCGTGTGAGCGGAAATCGTGCAACAAACAAA  
 ATAACTTCTCCGAGCCATAGAACACAAACATCTGTTGCAGCTGACAGTATGGGAAATCAACAGCTTTCAGGCAAGAGTGT  
 GGCCGTCGAGAGATACCTCAAAGATCAAACACTGCTGGGCATATGGGGATGTTCCGGTAAACTCATATGCATACCGCCGTGCC  
 TGGAACGGGAGCTGGTCTAATAAATCCCTGGATGAAATTTGGGACAAACATGACTTGGATGGAATGGAAACGGGAAATGACAAC  
 ATACTAGTTTGTATTTACTCTGATCGAAGAATCTCAGAACCAACAGGAGAAAACGAAACAGGAACTGCTGGAACCTGGACAAGTG  
 GGCATCATTTGTGGAAGTGGTGTGACATTAATAACTGGCTGTGGTAAAGATCTTACAA

(For all 140CF design shown here and below, 140CF gene will be flanked with the 5'  
 sequence of "TTCAGTCCGACGGCCACC" that contains a Kozak" sequence (GCCACCATGG/A) and  
 SalI site and 3' sequence of TAAAGATCTTACAA containing stop codon and EglII site.)

47/178

Fig. 29A

CON OF CON-S-2003 (829 a.a.)

MRV̄M̄GĪQR̄NC̄Q̄HL̄WR̄WḠIL̄IF̄ḠML̄IĪCS̄AĀEN̄L̄W̄TV̄YȲḠVP̄V̄W̄KĒAN̄T̄L̄FC̄AS̄DĀK̄AȲD̄TĒV̄HN̄V̄W̄ATH̄AC̄V̄PT̄DP̄NP̄QĒIV̄L̄  
 ENV̄TEN̄FN̄MW̄K̄NN̄M̄VĒQ̄MH̄EDĪIS̄L̄WD̄Q̄SL̄K̄PC̄VK̄L̄T̄PL̄CV̄TL̄NC̄TD̄VN̄AT̄NN̄T̄N̄NĒĒIK̄NC̄SF̄NĪT̄TĒIR̄DK̄KK̄V̄ȲAL̄FȲKL̄  
 DV̄VP̄ID̄DN̄NS̄ȲRL̄IN̄C̄NT̄SĀIT̄Q̄AC̄PK̄V̄SF̄ĒPĪPĪHȲCĀPAGĒFĀIL̄K̄C̄ND̄K̄K̄F̄NḠT̄ḠPC̄KN̄V̄ST̄V̄Q̄C̄TH̄GĪK̄P̄V̄V̄ST̄Q̄LL̄L̄NḠSL̄  
 ĀEĒĒIĪR̄SĒNĪTN̄NĀK̄T̄IĪV̄QL̄NĒS̄VĒIN̄CT̄RP̄NN̄TR̄KS̄ĪR̄IḠP̄Q̄AF̄ȲAT̄ḠDĪIḠDĪR̄Q̄ĀHC̄NĪS̄RT̄K̄WN̄KT̄L̄Q̄Q̄V̄AK̄KL̄RĒ  
 HF̄NK̄T̄IĪFN̄P̄SS̄GḠD̄LĒIT̄TH̄SF̄NC̄ḠGĒFF̄ȲC̄NT̄SĒL̄FN̄ST̄W̄NḠTN̄NT̄IT̄L̄PC̄RĪK̄Q̄ĪIN̄MW̄Q̄GV̄ḠQ̄AM̄ȲAP̄PĪĒḠK̄IR̄CT̄SN̄IT̄  
 ḠLL̄L̄TR̄D̄GḠNN̄T̄ĒT̄FR̄P̄GḠDM̄RD̄N̄WR̄SĒLȲKȲV̄V̄K̄IĒPL̄ḠV̄AP̄TK̄AK̄RR̄V̄VER̄EK̄RĀV̄GĪḠAV̄FL̄ḠFL̄ḠAḠST̄M̄GĀAS̄IT̄L̄  
 TV̄Q̄AR̄Q̄LL̄SḠIV̄QQ̄SN̄LL̄RĀIĒAQ̄Q̄H̄LL̄Q̄LT̄V̄WḠIK̄Q̄L̄ŌAR̄V̄L̄ĀVĒRȲL̄KD̄Q̄LL̄GĪWḠSC̄ḠK̄L̄ĪCT̄TN̄V̄P̄WN̄SS̄WS̄N̄KS̄Q̄DĒĪ  
 W̄DN̄MT̄W̄MĒWD̄KĒIN̄NȲTD̄IĪȲSL̄IĒES̄QN̄Q̄Q̄ĒK̄NĒQ̄ELL̄AL̄DK̄W̄AS̄L̄WN̄WF̄D̄IT̄N̄WL̄MȲIK̄ĪF̄IM̄IV̄ḠGL̄IḠL̄R̄IV̄FĀV̄LS̄IV̄NR̄  
 VR̄Q̄YS̄PL̄SF̄Q̄TL̄IP̄N̄PR̄ḠPD̄R̄PĒḠIĒĒḠĒQ̄DR̄DR̄S̄IR̄LV̄NḠEL̄AL̄AW̄DD̄L̄RS̄L̄C̄L̄FS̄ȲH̄RL̄DL̄L̄IĀART̄V̄ELL̄ḠRR̄ḠWĒĀ  
 L̄K̄YL̄WN̄LL̄Q̄ȲWḠQ̄EL̄K̄NS̄AIS̄L̄L̄DT̄TĀĪVĀĒGT̄DR̄V̄IĒV̄V̄Q̄RV̄C̄RĀIL̄N̄IP̄RR̄IR̄Q̄GF̄ER̄ALL̄

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 29B

CON-S-2003 140CF.pep (620 a.a.)

Nick name: 006

MRV̄M̄GĪQR̄NC̄Q̄HL̄WR̄WḠIL̄IF̄ḠML̄IĪCS̄AĀEN̄L̄W̄TV̄YȲḠVP̄V̄W̄KĒAN̄T̄L̄FC̄AS̄DĀK̄AȲD̄TĒV̄HN̄V̄W̄ATH̄AC̄V̄PT̄DP̄NP̄QĒIV̄L̄  
 ENV̄TEN̄FN̄MW̄K̄NN̄M̄VĒQ̄MH̄EDĪIS̄L̄WD̄Q̄SL̄K̄PC̄VK̄L̄T̄PL̄CV̄TL̄NC̄TD̄VN̄AT̄NN̄T̄N̄NĒĒIK̄NC̄SF̄NĪT̄TĒIR̄DK̄KK̄V̄ȲAL̄FȲKL̄  
 DV̄VP̄ID̄DN̄NS̄ȲRL̄IN̄C̄NT̄SĀIT̄Q̄AC̄PK̄V̄SF̄ĒPĪPĪHȲCĀPAGĒFĀIL̄K̄C̄ND̄K̄K̄F̄NḠT̄ḠPC̄KN̄V̄ST̄V̄Q̄C̄TH̄GĪK̄P̄V̄V̄ST̄Q̄LL̄L̄NḠSL̄  
 ĀEĒĒIĪR̄SĒNĪTN̄NĀK̄T̄IĪV̄QL̄NĒS̄VĒIN̄CT̄RP̄NN̄TR̄KS̄ĪR̄IḠP̄Q̄AF̄ȲAT̄ḠDĪIḠDĪR̄Q̄ĀHC̄NĪS̄RT̄K̄WN̄KT̄L̄Q̄Q̄V̄AK̄KL̄RĒ  
 HF̄NK̄T̄IĪFN̄P̄SS̄GḠD̄LĒIT̄TH̄SF̄NC̄ḠGĒFF̄ȲC̄NT̄SĒL̄FN̄ST̄W̄NḠTN̄NT̄IT̄L̄PC̄RĪK̄Q̄ĪIN̄MW̄Q̄GV̄ḠQ̄AM̄ȲAP̄PĪĒḠK̄IR̄CT̄SN̄IT̄  
 ḠLL̄L̄TR̄D̄GḠNN̄T̄ĒT̄FR̄P̄GḠDM̄RD̄N̄WR̄SĒLȲKȲV̄V̄K̄IĒPL̄ḠV̄AP̄TK̄AK̄TL̄TV̄Q̄AR̄Q̄LL̄SḠIV̄QQ̄SN̄LL̄RĀIĒAQ̄Q̄H̄LL̄Q̄LT̄V̄  
 W̄ḠIK̄Q̄L̄Q̄AR̄V̄L̄ĀVĒRȲL̄KD̄Q̄LL̄GĪWḠSC̄ḠK̄L̄ĪCT̄TN̄V̄P̄WN̄SS̄WS̄N̄KS̄Q̄DĒĪW̄DN̄MT̄W̄MĒWD̄KĒIN̄NȲTD̄IĪȲSL̄IĒES̄QN̄Q̄Q̄ĒK̄  
 NĒQ̄ELL̄AL̄DK̄W̄AS̄L̄WN̄WF̄D̄IT̄N̄WL̄W̄\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

48/178

## Fig. 29C

CODON-OPTIMIZED CON-S-2003 140CF.seq (1891 nt

Nick name :006

TTCAGTCGACAGCCACCATGCCGGTCAATGGGGATACAGAGGAATGCCAGCACTTGTGGAGGTGGGAAATTTGATATTCGGGAT  
 GCTCATAATCTGCTCTGCCGCTGAGAACCTGTGGTCACTGTGTATACGGCGTTCCCGTGTGAAAGAAGCTAATACTACCCCTG  
 TTTTGTGCAAGCGACCCAAAGCATAACGACACCGAAGTCCACAATGTCTGGCTACCCACGCTGTGFACTACTGATCCAAATC  
 CCAGGAAATTTTCTTGAAAACGTAAACGGAAAATTTAAACATGTGGAAGAAATAATAGTGGAGCAAAATGCACGAGGATATAAT  
 CAGCCTGTGGACCAGTCCCTCAAAACCATGCCGTTAAACTCACTCCACTGCTGACTCTGAACTGTACCGACGTGAACGCAACC  
 AATAATACAACAACAATGAGGAGATAAAGAAATTTTCATTTAATAAACCACCTGAGATACGGGATAAGAAAATAAGGTTTATG  
 CACTCTTTTACAAGCTCGACGTGGTGGCCATAGACGACAATAATAGCTACCGACTCAATTAATGCAATACTAGCGCTATAACCCA  
 GGATGCCCCAAAAGTTTCCCTCGAGCCCATACCGATTCACTACTGCCACCCCGGGATTCCGCAATTTAAATGCAATGACAAG  
 AAGTTCACGGCACCCGGACCCCTGTAAGAACGTAAGCACTGTTCAATGTACACATGGAATTAAGCCGGTAGTGTCAACCGCAGTCC  
 TCCTCAAACGGAAAGCCTTGCAGAAAGAGATCATTATACCCGCCCTAATAAACAAGAAATCACTAACCAACCGGAAACAATCATTTGAGCT  
 GAATGATCTGTAGAAATCAATTTGTACCCGCCCTAATAAACAAGAAAGTCAATTAGGATCGGATCGGACCCCGCCAGGCTTTCTAC  
 GCAACCGGAGATATCATCGGGATATACGACAGGCCCATCATCTTCAATCCAAAGTAGCGGGACCTGGAATAAACTTTGCAGCAGGTAG  
 CCAAGAACTGCGGAAACATTTTAAATAAGACAATCATCTTCAACTCAACATGGAATGGCACTAACAACTACTACACATTCCTT  
 TAACTGTGGGGCGAGTTTCTACTGTAAATACCTTGAACGTGGCAGGGTGTGGGCAAGCAATGTATGCACCACTTGAAGGCAAAATAAGAT  
 CCTTGCAGAAATAAACAGATTAACAACATGTGGCAGGGTGTGGGCAAGCAATGTATGCACCACTTGAAGGCAAAATAAGAT  
 GCACCTCCAATATACCGACTCCTCCTGACACGGGATGGGAAACAATAACCGGAGACCTTAGGCCAGGCCAAAGCTAAAACC  
 GAGAGATAACTGGCGTCCGAGCTCTATAAATAACAAGTCGTTAAGATCGAGCCCTTGGAGTTGCCCAACCCAAAGCTAAAACC  
 TTGACCGTGCAAGCCAGGCAGTTGTTGTCCAGGTATCGTACAGCAGCAATCTAATCTTTTGAAGCCATTGAGGCTCAGCAGCACC  
 TCTTGCAGCTTACCCTGTTGGGCAATCAACAACCTTCAGGCACCGCTCCCTGGCCGTAGAGCCGCTATTGAAAGACCCAACTTCT  
 CGGGATCTGGGGTGTCTGGAAAATGATCTGCACGACAAATGTGCCCTTGGAAACAGCAGCTGGTCAAATAAAGCCAAAGACGAA  
 ATATGGGATAACATGACATGGATGGATAAAGAAATTAATAATTACACTGACATTAATTTACTCACTTATCGAGGAATCAC  
 AAAATCAACAGGAAAATAAATGAACAGGAACCTTGGCTCTGGACAAAATGGGCTTCACTGTGGAACACTGGTTTCGACATCACAAAATTG  
 GCTCTGGTAAAGATCTTACAA

49/178

Fig. 30A

**CONSENSUS A1-2003 (845 a.a.)**  
 MRVNGIQRN<sup>C</sup>QHLLLRWGTMI<sup>L</sup>GMIIICSAENLWVTVYGV<sup>P</sup>VW<sup>K</sup>DAE<sup>T</sup>TLFCASDAKAYETEMHN<sup>V</sup>WATHACVPTD<sup>P</sup>NPQ<sup>E</sup>IHL  
 ENVTEEFN<sup>M</sup>WKNNMVE<sup>Q</sup>MHTDII<sup>S</sup>LWDQ<sup>S</sup>LKPCV<sup>K</sup>LTP<sup>L</sup>CVTLNCSN<sup>V</sup>NVTNNTT<sup>N</sup>THEEEI<sup>K</sup>NC<sup>S</sup>FNMTTEL<sup>R</sup>DKK<sup>Q</sup>KV<sup>S</sup>SLFY  
 RLDV<sup>V</sup>QINENNS<sup>S</sup>YRLINCNTSAIT<sup>Q</sup>ACP<sup>K</sup>VSFEPIPIHYCAPAG<sup>F</sup>AILKCKD<sup>K</sup>EFNGT<sup>G</sup>PC<sup>K</sup>NVSTV<sup>Q</sup>CTHG<sup>I</sup>K<sup>P</sup>VVST<sup>Q</sup>LL  
 L<sup>N</sup>GLAAEE<sup>E</sup>VIIRSENIT<sup>N</sup>NAKTIIV<sup>Q</sup>LTK<sup>P</sup>VKINCT<sup>R</sup>PNNT<sup>R</sup>K<sup>S</sup>IRIG<sup>P</sup>GQAF<sup>A</sup>YATGDIIG<sup>D</sup>IRQA<sup>H</sup>CN<sup>V</sup>SRSE<sup>W</sup>NK<sup>T</sup>LQ<sup>K</sup>VA  
 K<sup>Q</sup>L<sup>R</sup>K<sup>Y</sup>F<sup>K</sup>N<sup>K</sup>TII<sup>F</sup>TN<sup>S</sup>SG<sup>G</sup>DL<sup>E</sup>IT<sup>T</sup>H<sup>S</sup>FN<sup>C</sup>GG<sup>E</sup>FF<sup>Y</sup>CV<sup>N</sup>T<sup>S</sup>GL<sup>F</sup>NS<sup>T</sup>WN<sup>N</sup>GT<sup>M</sup>KN<sup>T</sup>IT<sup>L</sup>PC<sup>R</sup>IK<sup>Q</sup>IIN<sup>M</sup>W<sup>Q</sup>RAG<sup>Q</sup>AM<sup>Y</sup>AP<sup>P</sup>IQ<sup>V</sup>  
 IR<sup>C</sup>ES<sup>N</sup>IT<sup>G</sup>LL<sup>L</sup>TRD<sup>G</sup>GN<sup>N</sup>TNET<sup>F</sup>RP<sup>G</sup>GG<sup>M</sup>RD<sup>N</sup>WR<sup>S</sup>E<sup>L</sup>Y<sup>K</sup>Y<sup>K</sup>V<sup>V</sup>K<sup>I</sup>E<sup>P</sup>LG<sup>V</sup>AP<sup>T</sup>RA<sup>K</sup>RR<sup>V</sup>VER<sup>E</sup>K<sup>R</sup>AV<sup>G</sup>IG<sup>A</sup>VF<sup>L</sup>GF<sup>L</sup>GA<sup>A</sup>GS  
TMGAASITLTVQARQLLQARVLAVERILKQQLLGIWCGSGKLICTTNVPWNSS  
WSNKSQNEIWDNMTWLQWDKEISNYTHIIYNLIEESQOQEKNEQDLLALDKWANLWNWFEDISNWLWYIKIFIMIVGGLIGLRIV  
FAVLSVINRV<sup>R</sup>QGYSP<sup>L</sup>SFQ<sup>T</sup>HTPN<sup>R</sup>GLDRP<sup>G</sup>RIE<sup>E</sup>EGE<sup>Q</sup>GRDRSIR<sup>L</sup>VS<sup>G</sup>FLALAW<sup>D</sup>DL<sup>R</sup>SL<sup>C</sup>FS<sup>H</sup>RL<sup>R</sup>DR<sup>F</sup>EILIA<sup>A</sup>RT<sup>V</sup>E  
LLGHSS<sup>L</sup>KGL<sup>R</sup>L<sup>G</sup>WE<sup>G</sup>LKYL<sup>W</sup>NLLY<sup>W</sup>RELKISAIN<sup>L</sup>VD<sup>T</sup>IAI<sup>V</sup>AG<sup>W</sup>TD<sup>R</sup>VI<sup>E</sup>IG<sup>R</sup>IG<sup>R</sup>AIL<sup>H</sup>IP<sup>R</sup>RR<sup>I</sup>RQ<sup>G</sup>LE<sup>R</sup>ALL  
 \*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF  
 design and the "w" underlined with red color is the last amino acid at the C  
 terminus, and all the remaining amino acids after the "w" will be deleted in 140CF  
 design.

Fig. 30B

**Con-A1-2003 140CF.pep (629 a.a.)**  
**Nick name: 001**  
 MRVNGIQRN<sup>C</sup>QHLLLRWGTMI<sup>L</sup>GMIIICSAENLWVTVYGV<sup>P</sup>VW<sup>K</sup>DAE<sup>T</sup>TLFCASDAKAYETEMHN<sup>V</sup>WATHACVPTD<sup>P</sup>NPQ<sup>E</sup>IHL  
 ENVTEEFN<sup>M</sup>WKNNMVE<sup>Q</sup>MHTDII<sup>S</sup>LWDQ<sup>S</sup>LKPCV<sup>K</sup>LTP<sup>L</sup>CVTLNCSN<sup>V</sup>NVTNNTT<sup>N</sup>THEEEI<sup>K</sup>NC<sup>S</sup>FNMTTEL<sup>R</sup>DKK<sup>Q</sup>KV<sup>S</sup>SLFY  
 RLDV<sup>V</sup>QINENNS<sup>S</sup>YRLINCNTSAIT<sup>Q</sup>ACP<sup>K</sup>VSFEPIPIHYCAPAG<sup>F</sup>AILKCKD<sup>K</sup>EFNGT<sup>G</sup>PC<sup>K</sup>NVSTV<sup>Q</sup>CTHG<sup>I</sup>K<sup>P</sup>VVST<sup>Q</sup>LL  
 L<sup>N</sup>GLAAEE<sup>E</sup>VIIRSENIT<sup>N</sup>NAKTIIV<sup>Q</sup>LTK<sup>P</sup>VKINCT<sup>R</sup>PNNT<sup>R</sup>K<sup>S</sup>IRIG<sup>P</sup>GQAF<sup>A</sup>YATGDIIG<sup>D</sup>IRQA<sup>H</sup>CN<sup>V</sup>SRSE<sup>W</sup>NK<sup>T</sup>LQ<sup>K</sup>VA  
 K<sup>Q</sup>L<sup>R</sup>K<sup>Y</sup>F<sup>K</sup>N<sup>K</sup>TII<sup>F</sup>TN<sup>S</sup>SG<sup>G</sup>DL<sup>E</sup>IT<sup>T</sup>H<sup>S</sup>FN<sup>C</sup>GG<sup>E</sup>FF<sup>Y</sup>CV<sup>N</sup>T<sup>S</sup>GL<sup>F</sup>NS<sup>T</sup>WN<sup>N</sup>GT<sup>M</sup>KN<sup>T</sup>IT<sup>L</sup>PC<sup>R</sup>IK<sup>Q</sup>IIN<sup>M</sup>W<sup>Q</sup>RAG<sup>Q</sup>AM<sup>Y</sup>AP<sup>P</sup>IQ<sup>V</sup>  
 IR<sup>C</sup>ES<sup>N</sup>IT<sup>G</sup>LL<sup>L</sup>TRD<sup>G</sup>GN<sup>N</sup>TNET<sup>F</sup>RP<sup>G</sup>GG<sup>M</sup>RD<sup>N</sup>WR<sup>S</sup>E<sup>L</sup>Y<sup>K</sup>Y<sup>K</sup>V<sup>V</sup>K<sup>I</sup>E<sup>P</sup>LG<sup>V</sup>AP<sup>T</sup>RA<sup>K</sup>RR<sup>V</sup>VER<sup>E</sup>K<sup>R</sup>AV<sup>G</sup>IG<sup>A</sup>VF<sup>L</sup>GF<sup>L</sup>GA<sup>A</sup>GS  
QQHLLKLT<sup>V</sup>WGI<sup>K</sup>LQARVLAVERILKQQLLGIWCGSGKLICTTNVPWNSSWSNKSQNEIWDNMTWLQWDKEISNYTHIIYNLI  
EESQOQEKNEQDLLALDKWANLWNWFEDISNWLW\*  
 \*Amino acids seen in blue color is for easy identification of the junction of the  
 deleted fusion cleavage site.

Fig. 30C

**CODON-OPTIMIZED Con-A1-2003.seq**

**Nick name: 001 (1918 nt)**

TTCAGTCGACAGCCACCATGAGGGTGATGGGAATCCAACGGAACTGCCAGCATCTTCTCCGGTGGGAAACGATGATACTGGGAAT  
 GATAATAATCTGCTCTGCCCGTGAATAACCTCTGGGTACAGTGTACTACGGAGTGCCTGTATGGAAGACGCTGAAACCACTCTC  
 TTTTGTGCTTCCGATGCTAAAGCTACGAAACCGAGATGCACAATGTTTGGCCACCCACCCCTGCGTCCAACTGATCCTAATC  
 CACAAGAAATACATCTGGAGAATGTTACTGAGGAATTTAAACATGTGGAATAATAATGTTAGAGCAAAATGCACACTGACATCAT  
 TTCACTCTGGACCAATCACTCAAACCTCGTTAAACTTACCCCTCTGCGTGAACCTCAATTGTAGCAACGTCAACGTCACA  
 AATAATCAACCAACACTCACGAGGAGAAATTAATAATGCTCCTTTAATAATGACCACTGAACCTCGCGACAAAAACAATAAAG  
 TCTATTCACTGTTTTATAGGCTGGACGTCCTCAATCAACGAGAAACAATTTAAACAGTAGCTATCGACTTATCAATTGAATPAC  
 CTCTGCTATTACCCAAAGCTTGCTTAAGTCTCTTTGAACCAATCCCTATCCACTACTGTGCCCCAGCTGGATTCGCAATTCG  
 AAGTGCAAGGATAAGGAATCAACGGAACTGGCCCTTGCAAGAACGTTAGCACTGTCCAATGCACCTCACGGAAATCAAACCCAGTAG  
 TCAGCACTCAACTGCTCCTGAATGGCTCACTCGCCGAGAAAGGTTGATTAATCCGAAGCGAGAACATAACTAAACAATGCGAAGAC  
 AATAATGTTCAATTGACGAAACCCAGTGAAGATCAACTGTACTAGACCAATAACAACAAGAAATCTATCAGAATTGGCCCC  
 GGACAAGCCTTCTACGCAACAGGAGATATCATAGGTGACATCAGACAGGCCCAATTGCAACGTTTCAAGAAGCGAGTGAATAAAA  
 CACTCCAGAAAGTGGCAAGCAGCTGAGAAAATACTTTAAGAACAGACAATCATATTTACTAACTCCTCCGGAGGTGATCTCGA  
 AATAACCACTCATAGCTTTAATGTGGGGCGAATTTCTTACTGTAAACACATCTGGCCTCTTTAATCTACCTGGAATAACGGC  
 ACCATGAAAATAACTATCACCTCCCTTGCAGAAATTAAGCAATCATTAACATGTGCCAGAGCAGGACAGGCCATGTATGCC  
 CTCCCATCAAGGTGTGATTCGATGTGAAAGCAACATTAAGCAATTTACTGGACTTCTTGTGACCCGGATGGCGAAATAATAACCAATGA  
 GACATTCAGACCCGGCGCGGATA TCCGAGACAATGGCGAAGTGAACCTTTATAAAFACAAAAGTAGTTAAGATTGAGCCCTT  
 GGAGTTGCCCTACTAGAGCAAAAACATTTGACCCGTTCAAGCCAGGCAAGTGTCTCAGGAATCGTGCAGCAGCAAAAGTAACCTCC  
 TCCGAGCTATCGAGGCACAACAACATCTCTTGAAATTTGACCCGATGGGAAATCAAGCAATTCAGGGTAGGGTTTGGCTGTGGA  
 ACGCTATCTCAAGGATCAGCAGCTTCTGGGAATCTGGGATGCTCTGGGAAATTTGATATGTACTACAAACGTAACCCCTGGAACTCA  
 AGCTGGAGTAATAAAGCCAGAAACGAAATTTGGGATAAATGACCTGGCTGCAGTGGACAAAGAAATTTCTAATATATACTCATA  
 TCATATACAACTGATCGAAGAAATCACAGAACCCAGGAAAAGAAATGAGCAAGACCTTCTGGCCTTGGACAAGTGGGCTAACTT  
 GTGGAACCTGGTTGACATTAGCAACTGGCTGTGGTAAAGATCTTACAA

Fig. 31A

**CONSENSUS C-2003 (835 a.a)**  
 MRVRLRNCCQWIIWGIILGFWMMLMCNVVGNLWVTYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIIVL  
 ENVTENFNMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATMGEIKNCSEFNITTELRDCKKQKVYALFYRLDI  
 VPLNENNSYRLINCNNTSAITQACPVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGKPKVSTQLLLLNGSLAE  
 EEIIRSENLTNNAKTIIVHLNESVEIVCTRPNNTRKSIIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQKVSKKLKEHF  
 PNKTIKFEPSGGDLEITTHSFNCRGEEFFYCNTSKLFNSTYNSTNSTITLPCRKQIINMWQEVGRAMYAPPIAGNITCKSNITG  
 LLLTRDGGKNNTEFRPGGDMRDNRWSELYKYKVVVEIKPLGIAPTAKRRRVEREKRAVGI GAVFLGFLGAAGSTMGAASITLT  
 VQARQLLSGIVQQSNLLRAIEAQQHMLQLTVMGIKQLQTRVLAIERYLKDKQQLIGIWCSSGKLICTTAVPWNSSWSNKSQEDIW  
 DNMTWMQWDREISNYTDTIYRLLLEDSQOQEKNEKDLALDSWKNLWNWFEDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRV  
 RQGYSPLSFQTLTPNPRGPDRLGRIFEEGEGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDRFILLIARAARAVELLGRSSLRGL  
 QRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIEILIQICRAIRNIPRRIRQGFEEALQ  
 \*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF  
 design and the "W" underlined with red color is the last amino acid at the C  
 terminus, and all the remaining amino acids after the "W" will be deleted in 140CF  
 design..

Fig. 31B

**Con-C 2003 140CF.pep (619 a.a.)**  
**Nick name: 003**  
 MRVRLRNCCQWIIWGIILGFWMMLMCNVVGNLWVTYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIIVL  
 ENVTENFNMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATMGEIKNCSEFNITTELRDCKKQKVYALFYRLDI  
 VPLNENNSYRLINCNNTSAITQACPVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGKPKVSTQLLLLNGSLAE  
 EEIIRSENLTNNAKTIIVHLNESVEIVCTRPNNTRKSIIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQKVSKKLKEHF  
 PNKTIKFEPSGGDLEITTHSFNCRGEEFFYCNTSKLFNSTYNSTNSTITLPCRKQIINMWQEVGRAMYAPPIAGNITCKSNITG  
 LLLTRDGGKNNTEFRPGGDMRDNRWSELYKYKVVVEIKPLGIAPTAKKTLITVQARQLLSGIVQQSNLLRAIEAQQHMLQLTIVW  
**GIKQLQTRVLAIERYLKDKQQLIGIWCSSGKLICTTAVPWNSSWSNKSQEDIW**DNMTWMQWDREISNYTDTIYRLLLEDSQOQEK  
**EKDLLALDSWKNLWNWFEDITNWLW\***  
 \*Amino acids seen in blue color is for easy identification of the junction of the  
 deleted fusion cleavage site.



Fig. 31C

CODON-OPTIMIZED Con-C-2003 140CF (1,888 nt.)

Nick name:003

TTCAGTCGACAGCCACCATGCGAGTGAGAGCCATTCTGCGGAAATTGTCAGCAATGGTGGATCTGGGCATACTCGGATTCGGAT  
 GCTTATGATATGCAATGTTGTGGGAACCTGTGGGTACCCGTATACCTGATGAGTCTGGAGGAGGCTAAACAACCGCTG  
 TTCTGTGCAAGTGACGCCAAAGCTACGAGAAGAAGTGCACAACCTGTGGCTACCCACGCTTGTTCACACCGATCCAAACC  
 CCCAGGAAATCGTCTCGAGAACGTGACTGAAAACCTTAAACATGTGGAAGATGATGGTAGATCAGATGCACGAAAGATATCAT  
 TTCAATGTGGACCAATCATGAAACCATGCGTAAAACCTGACCCCTCTGCGTAAACCTTAACTGCACCAATGCAACTAATGCC  
 ACCAATACTATGGCGGAAATAAAAACTGTAGCTTTAAACATTAACAACGGAACCTCGGGATAAGAAACAAGGTTACCGGCTCT  
 TTACCGACTCGATATCGTCCCACTTAACGAGAAATAATAGTTACCGCTGATTAACCTGTAACACATCAGCCATTACGCAAGCTTG  
 CCCCAAAGTTTCTTTCGACCCCATCCCAATTCACATAATTGTCGCCCGCTGGATACGCTATACTTAAATGCAACAATAAAACATTT  
 AATGGAACCCGACCATGTAACAACGTACGTACCGTACAATGTACGACGGAATTAACCTGTTGTCTCAACCCAGCTTCTCCTTA  
 ACGGCTCATTTGGCGGAGGAAATAATTAATATCAGATCAGAAAACCTTGACCAACAATGCCAAAACCTATCGTGCACTCAATGA  
 ATCCGTGAAAATCGTGTGCACCCAGACCAATAACAATAACCGGAAATCAATCAGGATGGGCTGGCCAGACATTTTACGCTACA  
 GGTGATATAATTGGCGATATTAGACAAGCCCAATGCAACATATCAGAAGACAAGTGGATAAGACTCTGCAGAAAGTTTCTAAGA  
 AGCTGAAGGAACACTTTCCCAATAAAACGATTAAGTTCGAGCCCTTTCAGGAGGAGACCTTGAGATCAACAACACTCTTTTAA  
 TTGTAGAGGGGAGTTCTCTATTTGTAATACATCAAAGCTCTTAAACAGTACCTACAACCTCCACTAATAGTACCACTCACACTCCC  
 TGCAGAATAAAGCAATAATCAACATGTGGCAAGAAAGTGGCCGAGCAATGTAGCCCTCCCATCGCAGGCAACATTAACATGTA  
 AATCCAATATTACTGGCCTTTGTGTGACACGGGACGGCGGAAAGAAAGTAAAGTCTGTTGAAATCAAGCCACTGGGCATAGTCC  
 CGATAATTGGCGGAGCGACTCTACAAGTATAAAGTCTGTTGAAATCAAGCCACTGGGCATAGTCCCTACGAAAGCAAGACACTC  
 ACTGTTCAGGCTAGACAGCTGCTCTCCGGCATAGTGCACACAGCAATCCAATCTCCTGCGAGCTATCGAAGCCCAACAACATATGC  
 TCCAGTTACCGTCTGGGGAATCAACAATTGCAAAACACGAGTGTGGCGATAGAGAGATATTTGAAAGATCAGCAACTCCTGGG  
 GATTTGGGGCTGTTCAAGGTAAGCTCATCTGTACAACTGCGGTGCCGTGGAACCTCAAGCTGGAGTAAACAAGCCAAAGGATATA  
 TGGGACAACATGACTTGGATGCAGTGGGATCGAGAAATAAGCAACTATACAGATACCAATTTATCGGCTCCTGGAGGACTCACAGA  
 ACCAGCAGGAGAAAATAAGAAAAGATTGCTCGCGCTTGACAGTTGGAAGAAATTTGTGGAATTTGGTTCGACATTACAAAACCTGGCT  
 CTGGTAAAGATCTTACAA

Fig. 32A

**CONSENSUS G-2003 (842 a.a.)**  
 MRVKGIQRNWQHLLWKWGTLILGLVVICSSASNNLWTVVYGVVWEDADTFLFCASDAKAYSTERHNVWATHACVPTDPNPQEITL  
 ENVTFENFMWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVNVTNNTNNTKKEIKNCSFNITTEIRDKKKEYALFY  
 RLDVVPINDNGNSSIYRLINCNVSTIKQACPVTFDPIPIHYCAPAGFAILKCRDKKFNVTGPKCNVSTVQCTHGKPKVSTQLL  
 LNSLAEEEEIIIRSENITDNTKVIIVQLNETIEINCTRPNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCVSRTKWNEMLQKVK  
 AQLKKIFNKSITFNSSSGDLEITTHSFNCRGEFFYCNSTGLFNNLLNSTITLPCIKIQIVRMWQRVGQAMYAPPIAGNIT  
 CRSNITGLLTRDGGNNNTEFRPGGDMRDNRSELYKYIVKIKPLGVAPTRARRRVEREKRAVGLGALLGLFLGAAGSTMG  
AASTITVQVRQLLQARVLAVERYLKDQQLLGIWCCSGKLICTTNPWNTSWSN  
KSYNEIWDNMTWIEWEREISNYTQIYSLIEESONQOKEQDLALDKWASLWNWFDITKWLWYIKIFIMIVGGLIGLRIVFAV  
 LSIVNRVROGYSPLSFQTLTHHQREPPDRPERIEEGGEOQDKDRSIRLVSGFLALAWDDLRSLSCLFSYHRLRDFILIAARTVELLIG  
 RSSLKGRLGWGLKYLWNLWLLYWGQELKNSAINLLDTIAAVANWTDRIEVAQACRAILNIPRRIRQGLERALL  
 \*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF  
 design and the "W" underlined with red color is the last amino acid at the C  
 terminus, and all the remaining amino acids after the "W" will be deleted in 140CF  
 design.

Fig. 32B

**Con-G-2003 140CF (626 a.a.)**  
**Nick name: 007**  
 MRVKGIQRNWQHLLWKWGTLILGLVVICSSASNNLWTVVYGVVWEDADTFLFCASDAKAYSTERHNVWATHACVPTDPNPQEITL  
 ENVTFENFMWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVNVTNNTNNTKKEIKNCSFNITTEIRDKKKEYALFY  
 RLDVVPINDNGNSSIYRLINCNVSTIKQACPVTFDPIPIHYCAPAGFAILKCRDKKFNVTGPKCNVSTVQCTHGKPKVSTQLL  
 LNSLAEEEEIIIRSENITDNTKVIIVQLNETIEINCTRPNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCVSRTKWNEMLQKVK  
 AQLKKIFNKSITFNSSSGDLEITTHSFNCRGEFFYCNSTGLFNNLLNSTITLPCIKIQIVRMWQRVGQAMYAPPIAGNIT  
 CRSNITGLLTRDGGNNNTEFRPGGDMRDNRSELYKYIVKIKPLGVAPTRARRTLTVQVRQLLQARVLAVERYLKDQQLLGIWCCSGKLICTTNPWN  
TSWSNSYNEIWDNMTWIEWEREISNYTQIYSLIEESONQOKEQDLALDKWASLWNWFDITKWLW\*  
 \*Amino acids seen in blue color is for easy identification of the junction of the  
 deleted fusion cleavage site

Fig. 32C

CODON-OPTIMIZED Con-G-2003 140CF .seq

Nick name:007

TTCAGTCGACAGCCACCATGCGAGTGAAGGGAATCCAGAGAAAATTGGCAGCACCTTTGGAAGTGGGGCACACTCATCTCGGCCCT  
 TGTGATCATATGCTCGCCCTCAAATAACCTTTGGGTACAGTTTATTACGGCGTGCCCGTTTGGAGGACGCAGACACAACCTCTT  
 TTTTGTGCCAGCGTAAGCTTATTCAACAGAGAGCCATAACGTTTGGGTACACATGCATGCGTCCGACCCGATCCTAATC  
 CCCAGGAAATCACTTTGAGAAATGTTACAGAGAAATTTAATATGTGGAAGAACAAACATGGTTGAACAGATGCATGAAGACATAAT  
 TTCTCTCTGGGATGAATCTGAAAACCTTGGTGAAGCTTACACCACTGTGCGTTACCCCTGAATGCACTGACGTCAATGTCCACA  
 AATAATAATACCAACAATACAAAAAAGAAATCAAAAATTTGTTCTTTCAACATAAACCCGAGATACCGGATAAATAAAAAAGAAAAG  
 AATACGCCCTGTTCTACAGACTCGATGTGTCCTCCCAATTAATGACAACGGAAATTTCTCCATCTACCGACTTATCAATGTAAACGT  
 GTCTACAATCAACAGGCCCTGTCTAAAGTCACATTTGACCCCTATTCCCAATTCATTAAGTGTGCCCCCGTGGCTTCGCTATTCTT  
 AATGCCGAGACAAAATAATTAACGGAAACAGGCCATGCAAGAAATGTCTCAACAGTTCAATGCATCATGGAAATTAACCCAGTCCG  
 TTTCTACTCAACTCCTTCTCAATGGAGCCTGGCAGAGAGGAAATCATAATCCGCAAGCAAAACATAAACCCGACACACAAAAGT  
 AATCATCGTACAGTGAACGAGACCATTGAATAAATTTGACGAGACTAATAACACAAAGAAAAGCATACGCATCGGCCCC  
 GGACAGGCTTTCTACGCCACAGGACATTAATCGGAGATATCCGCCAGGCTCACTGTAATGTGTAGAACAAAATGGAACGAAA  
 TGCTTCAGAAGTCAAAGCTCAGCTCAAGAAAATATTCAACAAATCTATTACATCAACTCATCATCAGGGCCGATCTGGAGAT  
 AACAACTCATTCCTTCAACTGTCCGGGAGAAATTTTTTACTGTAACACGTCGGCCCTGTTCAACAATTCCTCTGAAATAGCACT  
 AACTCCACCATCACTTCCCATGTAAGATCAAACAATCGTCAAGAAATGTGGCAGCGAGTGGTCAAGCTATGTACGCCCTCCAA  
 TCGCCGGTAATATCACATGTAGAACCAATATCACAGGCTCTTGTGTTACAAGGACGGCGGAAACAACAACCCGAAACCTTCAG  
 ACCAGGAGGAGGACATCGGAGACAAATGGCCGGAGCGAGCTGTATAAATATAGATCGTAAAAATCAAACCAATTTGGGTGTAGCG  
 CCAACTAGAGCCCGAACACTGACCGTGCAGGTGAGGCAACTGCTGAGCGGCATGTCCAACAACAATCCAATCTTCTTAGAGCAA  
 TCGAGGCCAGCAGCATCTGCTCCAGCTTACTGTATGGGAAATCAAACAACCTGCAAGCAAGAGTATGGCAGTGGAGAGGTATCT  
 CAAGGACCAGCAGCTTCTGGGAAATTTGGGTTGCAGCGGAAAGCTCATATGTACAACCAATGTGCCCTGGAACACTAGTTGGAGT  
 AATAAGAGTTACAATGAAATCTGGGACAATATGACATGGATCGAATGGAGCGGAAAATATCCAACATACTCAGCAAAATCTATT  
 CCCTCATGGAAGAGTCAAGAACCCAGGAAAAGAAATGAGCAAGACCTCCTCGCCCTGGATAAATGGGCATCTCTGTGGAACCTG  
 GTTGCATAAATAAATGGTTGTGTAAGATCTTACAA

Fig. 33A

CONSENSUS 01 AE-2003 (854 a.a.)  
 MRVKETQMNWPNLWKWGTLLGLVICSASDNLWTVVYGVVWRDADTTLFCASDAKAHETEVEHNVWATHACVPTDPNPQEIHL  
 ENVTFENFMWKNMVEQMEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTVNNTVSNIIIGNITNEVRNCSFNMTTELDRDKK  
 QKVHALFYKLDIVQIEDNNSYRLINCVTSVIKQACPKISFDPIPIHYCTPAGYAILKCNCKNFNGTGPCKNVSSVQCTHGKIPVV  
 STQLLNGSLAEEEEIIRSENLTNNAKTIIVHLNKSVEINCTRPSNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEV  
 LKQVTEKLKEHFNNKTIIFQPPSGGDELEITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIIILPCKIKQIINMWQAGQA  
 MYAPPISGRINCVSNITGILLTRDGGANNETFRPGGNIKDNWRSELYKYKVQVIEPLGIAPTRAKRRVVEREKRAVIGIGAMI  
 EFGFLGAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSSGKIIC  
 TTAVPWNSTWSNRSEIWNMTWIEWEREISNYTNQIYEILTESQOQDRNEKDLLELDKWAASLWVWFDITNWLWYIKIFIMIV  
 GGLIGLRIIFAVLSIVNRVROQYSPLSFQPTHHQREPRPERIEEGGEGQGRDRSVRLVSGFLALAWDDLRSCLFSYHRLRDF  
 ILIAARTVELLGHSSLKGLRRGWEGLYLGNLLLYWGQELKISAIISLLDATAIAVAGWTDRIEVAQAGAWRAILHIPRRIRQGLE  
 RALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted as 140CF.

Fig. 33B

Con-AE01-2003 140CF.pep (638 a.a.)  
 Nick name: 008  
 MRVKETQMNWPNLWKWGTLLGLVICSASDNLWTVVYGVVWRDADTTLFCASDAKAHETEVEHNVWATHACVPTDPNPQEIHL  
 ENVTFENFMWKNMVEQMEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTVNNTVSNIIIGNITNEVRNCSFNMTTELDRDKK  
 QKVHALFYKLDIVQIEDNNSYRLINCVTSVIKQACPKISFDPIPIHYCTPAGYAILKCNCKNFNGTGPCKNVSSVQCTHGKIPVV  
 STQLLNGSLAEEEEIIRSENLTNNAKTIIVHLNKSVEINCTRPSNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEV  
 LKQVTEKLKEHFNNKTIIFQPPSGGDELEITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIIILPCKIKQIINMWQAGQA  
 MYAPPISGRINCVSNITGILLTRDGGANNETFRPGGNIKDNWRSELYKYKVQVIEPLGIAPTRAKR**TLTVQARQLLSGIVQQQ**  
**SNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSSGKIICTTAVPWNSTWSNRSEIWNMTWIEWEREISN**  
**YTNQIYEILTESQOQDRNEKDLLELDKWAASLWVWFDITNWLW\***

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

56/178

## Fig. 33C

CODON-OPTIMIZED Con-AE01-2003 140CF.seq (1945 nt.)

Nick name: 008

ttcagtcgacagccaccatgCGAGTCAAGGAAACACAAAATGAACTGGCCTAATCTGTGGAAGTGGGGCACCCTGATCCTGGGTTT  
 GGTCAATATTGCTCTCGGAGCGACAATCTCTGGTTACTGTCTATTACGGAGTCCCCGTTTGGAGAGATGCCGACACTACACTG  
 TTCTGGCCTCAGATGCCAAAGCTCATGAAGTGCATAATGTTTGGCAACCACCCCTGTGTTCTTACCACCCAAACC  
 CCCAAGAAATACACCTGGAAACGTCACCGAGAACTTTAATATGTGGAAGATAACATGGTTGAACAGATGCAAGAAAGCGTAAT  
 CAGCCTGTGGGATCAAGTCTGAAAACCTTGGTAAACTGACTCCACTTGGCTAACACTTAATGGACCAACCGGAACCTTGACA  
 AACGTTAACAACTCACTAACGTTCCAAACATCATCGGCAACATAACGAACGAAGTGAATAATGCAAGTTTCAATATGACTACAG  
 AGCTCCGGACAAGAACAGAGGTCCATGCTCTTTTACAACTCGACATCGTCCAGATCGAAGACAATAACAGCTACAGACT  
 TATAAATTGTAATACATCCGTGATTAACAAGCATGCCCAAAATAAGCTTCGATCTATTCCTATCCACTACTGTACTCCTGCC  
 GGCTATGCTATCTTGAAATGCAATGATAAGAACTTCAATGGACCGGACCTTGTAAAGAACGTTAGTGTGCAATGCACCTCAGG  
 GCATTAACCCAGTGGTAAGCACCAGTCTCTGAAACGGCTCTCTGGCAGAGGAAGATTAATTCGAAAGTGAACCTCAC  
 CAACAACTAGACTATCATCGTACATCTCAATAAATCAGTCGAAATTAATTCACCAGACCTCCAATAATACTAGAACTTCA  
 ATCACTATCGGCCAGGACAAAGTCTTTATAGAACAGGAGATATCATAGGAGATATCAGAAAAGGCATATTGGAGATAAAACGGGA  
 CAAAATGGAACGAAGTACTCAACAAAGTCAAGAGAGCTTAAGGAACATTTCAACAATAAAACCATTATTTTCAACCCCAAG  
 TGGCGGACCTCGAAATCACTATGCCACCTTCAACTGCCGGCGGAATTTTTTATTGCAATACCACATAAACTTTTCAACAAT  
 ACGTGCTCGGAAATGAGACCCATGGAGGCTGCAATGGAACAATCATACTCCCATGCAAGATAAAACAATCATTAACATGTGGC  
 AAGGTGCTGGACAAGCTATGATGCAACCCCAATATCCGGTAGAATTAATTCGTCAGCAACATCACTGGCATACTGCTCACFAG  
 AGACGGAGGACAAATAACAAATGAAACAATCCGACCAGGGCGCAACATTAAGGACAACTGGCGTCCGAACCTCTATAAG  
 TACAAAGTCGTACAGATCGAACCTCTTGGAAATAGCACCGACTCGCGTAAGACACTCACAGTACAGGCCCGACAACCTCTTTCTG  
 GAATCGTACAGCAGCAATCAACCTCTCCGGCAATCGAGGCCCAACAACATCTGCTTCCAGCTCACAGTTTGGGAATCAAGCA  
 GCTCCAGGCACGGTGTCCAGTGGAAAGATACCTGAAGGATCAGAAATTCCTTGGTCTCTGGGATGTTCTGGCAAAATAATC  
 TGCACACCGGGTTCCCTGGAATCAACATGGAGCAACCGGAGTTTGAAGAGATATGGAACAATATGACATGGATAGAGTGGG  
 AAAGGAAATTAGTAACTATACGAACTATACGAAATCTCACCGAAAGCCAAATCAGCAGGATCGCAACGAAAAGACCT  
 CCTCGAGCTTGATAAGTGGGCAATCCCTTTTGGAACTGGTTCCGACATCAAAAATTTGGCTCTGGTaaagatctttaa

57/178

Fig. 34A

Wild-type subtype A Env  
**00KE\_MSA4076-A (Subtype A, 891 a.a)**  
 MGAMGIQMNWQNLWRWGTMLGMLIICSVAEKSWVTVYGVVWRDAETTLFCASDAKAHDKEVHNVWATHACVPTDPNPQEMIL  
 ENVTEDFNMWKNSMVEQMHTDIIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTSNSTKDSATLDMKSEIQNCSEFNMTTELDRDK  
 KQKVSFLFYRLDVVQINENSSDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNCKKFNKGTGPTNVSTVQCTHGIKP  
 VVTQLLLLNGSLAEVEVMIRSENITENAKNIIVQFKEPVQIICIRPGNTRKSVHIGPGQAFYATGDIIGDIRQAHNCVNSRELWN  
 KTLQEVATQLRKHFRRNNTKIIFTNSSGGDVEITTHSFNCGGEFFYCDTSGLFNSSWTASNDMSQEAHSTESNITLQCRIKQIINN  
 WQRAGQAMYAPPPIPGIIRCESNITGLILTRDGGEGNNTNETFRPVGGNMRDNWRSELYKYVVKVEPLGVAPT~~KSRRRVRVEREK~~  
 RAVGLGAVFIGELGAAGSTMGAASMTLVQARQLLSGIVQQSNLLRAIEAQOHLKLTVWGIKQLQARVLAVERYLRDQQLLGI  
 WGCSSGLICTTNPWNSSWSNKSLSDEIWNMTWMQDKEVSNYTMIIYNLLEESONQOEKNEQELLALDKWANLWNNWFENISNWLW  
 YIKIFIMIVGGLIGLRIVFAVLSVINRVRQGYSPLSFQHTFPNPRGLDRPGRIFEEEGEQDRDRSIRLVSGFLALAWDDLRSLSLI  
 FSYHRLRDFILIAARTLELLGHNSLKGRLGWEGLYLWNLAYWGRELKISAI~~SLVDSIAIAVAGWTDRIIEIVQAI~~GRAILHI  
 PRRIRQGLERALI

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 34B

**00KE\_MSA4076-A 140CF.pep (647 a.a)**  
 Nick name: 011  
 MGAMGIQMNWQNLWRWGTMLGMLIICSVAEKSWVTVYGVVWRDAETTLFCASDAKAHDKEVHNVWATHACVPTDPNPQEMIL  
 ENVTEDFNMWKNSMVEQMHTDIIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTSNSTKDSATLDMKSEIQNCSEFNMTTELDRDK  
 KQKVSFLFYRLDVVQINENSSDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNCKKFNKGTGPTNVSTVQCTHGIKP  
 VVTQLLLLNGSLAEVEVMIRSENITENAKNIIVQFKEPVQIICIRPGNTRKSVHIGPGQAFYATGDIIGDIRQAHNCVNSRELWN  
 KTLQEVATQLRKHFRRNNTKIIFTNSSGGDVEITTHSFNCGGEFFYCDTSGLFNSSWTASNDMSQEAHSTESNITLQCRIKQIINN  
 WQRAGQAMYAPPPIPGIIRCESNITGLILTRDGGEGNNTNETFRPVGGNMRDNWRSELYKYVVKVEPLGVAPT~~KSRRRVRVEREK~~  
LLSGIVQQSNLLRAIEAQOHLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSSGLICTTNPWNSSWSNKSLSDEIWNMTW  
MQWDKEVSNYTMIIYNLLEESONQOEKNEQELLALDKWANLWNNWFENISNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

Fig. 34C

CODON-OPTIMIZED OOKA\_MSA4076-A 140CF.seq (1972 nt.)

Nick name: 011

ttcagtcgacagccaccATGGGGCAATGGGAATCCAGATGAACCTGGCAGAACCTCTGGCGATGGGCACAAATGATCCTGGGTAT  
GCTCATCTGCTGTGTCAGAAAAGTCAAGGTAACAGTCTACTACGGCGTACCAGTGTGGCGGACGCCGAAACCACTCTC  
TTCTGGCCTCCGATGCCAAAGCACACGATAAAGAGTCCACAATGTTTGGCTACCCATGCGTGGTCCAAACCGATCCTAAC  
CACAGAAATGATACTCGAAAACGTACTGAGACTTCAACATGTGGAAAATTTATGTTGAACAGATGCACACCGACATAAT  
ATCACTGTGGATCAGTCTCTCAAACCTGTGTCAAATTGACCCCTCTGCTTACACTGAACTGTTCCGACTCAAATATCACT  
TCTAAPTCAAAGGCAATAGTACGAAAGACTCCGCAACCTTGATATGAAAAGCGAAATACAGAACTGTTCAATTAATATGACCA  
CCGAACTGAGAGATAAAAAGCAGAAAGTTTATCTCTGTCTATCGATTGGACGTGGTTCAGATTAACGAAAATAGCAGCGATTA  
CCGACTCATTAAGTCAATACATCAGCAATCACACAGGCTTGCCCAAGGTAAACATTTGAGCCAAATCCCTATTCACCTACTGCGCC  
CCTGCAGGATTTGCCATCCTGAAATGCAACGATAAGAGTTTAAATGGGACAGGACCTGCACCAACCTCCACCCGTGCAATGCA  
CCCACGGCATAAAACCTGTTGTTACCACACAATGCTGCTCAATGGATCACTTGTGAAGAGGAACTCATGATTCGGTCTGAAA  
CATCACTGAAAATGCCAAAATATTAAGTTCAAGTTCAAAGAACCCGTCAGATCATTTGCATTCGCCCTGGTAACAACACTCGC  
AAGTCAGTGCACATTTGGCCCGCCAGGCTTCTATGCAACCCGGAGATTAATAGGCGACATCAGACAGGCACATTCGCAACGTC  
GCCGGAAATGTGGAACAAAACCTTTGCAGGAAAGTTGCTACTCAGCTGCGAAAACATTCAGAAAACAATACAAAAGATTTATTCAC  
TAATTCATCAGCGGTGACGTGGAGTCACTACCCATTCATTTAACTGTGGCGGAGAAATTTCTATTCGATACCTCTGGGCTC  
TTAAATCCCTCATGGACTGTAGCAACGATTCATGCAAGAACATGTACGCACCTCCATCCCGAATTAATCGATGTGAGTCTAAT  
AACAAATCATCAATATGTGGCAGCGGCGGTCAAGCAATGTACGCACCTCCATCCCGAATTAATCTACAACCGAGACTTCAGACCCGTAAGGCAATATGCCGA  
CACTGGCCCTCATCTGACCCGAGACGGTGGCGGAGTAATAATTTGAAGGTAGAACCTCTGGAGTGGCACCCCAAAATCACGAAACCTGA  
GACAAATGGCGATCCGAACTGTATAATAATAAAGTGGTGAAGGTAGAACCTCTGGAGTGGCACCCCAAAATCACGAAACCTGA  
CTGTGCAGGCACGCCAACTTTGTAGCGGAAATAGTCCAAACAGCAATCCAAATCTCTGAGAGCTATAGAAGCCCAACACCTGCT  
TAACTTACGGTGTGGGAAATCAACAAATTCAGGCAAGAGTGTGGCAGTGGAACTTCCCTGGAACTTCCCTGGTCAACCAAGTCTGGACGAAATAT  
ATCTGGGGATGTTCCGGTAAGTTGATTTGCACGCAAAACCTTCCCTGGAACTTCCCTGGTCAACCAAGTCTGGACGAAATAT  
GGGAAAATATGACATGGATGCAGTGGGACAAAGGAAAGTTAGCAACTATACACAGATGATCTACAACCTCCCTCGAAGAACTCAGAA  
TCAACAGGAAAACGAAACGAAACTGCTCGCCCTCGATAAGTGGGCTAACCTCTGGAACTGGTTAATAATTTCAAACCTGGTTG  
TGGtaagatcttataa

59/178

Fig. 35A

**Wild-type subtype B****QH0515.1g gp160 (861a.a)**

MRVKEIRRNQCQLRRWGTMLLGMMLICSATEQLWVTVYGVVWKEATTLFCASDAKAYVTEKHNWATHACVPTDPNPQEVVL  
 ENVTENFNWKNMNMVEQMHEDIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSGTNSSSWEKVQKEIKNCSEFNITTGIRGRVQ  
 EYSLFYKLDVIPIDSRNNSNNEFSSYRLISCNTSVITQACPKISFEPIPIHYCAPAGFAILKNDKFKFNGTGPCKNVSTVQCT  
 HGIKPVVSTQLLNGSLAEVEVIRSENFNTNPKSIIIVQLNKSVINCTRPNNTRKSIHIGAGKALYTGEIIGDIRQAHCNLSR  
 AQWNNTLKQIVIKLREQFGNKTIVFNQSSGGDVEIVMHSFNCGGEFFYCNSTQLENSTWNGNDTWNKTNDNITLPCRIKQ  
 IVNMWQVKGKAMYAPPPIRGQIRCSSKITGLILTRDGGTNGTNETETFRPGGNMKNRSELYKYVVKIEPLGIAPTAKARRV  
 QREKRAVGTIGAMFLGFLGAGSTMGAASLTLTVQARLLLSGIVQQNNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLRDQ  
 QLLGIWCSGRLICTINVPWNTSWSNRSLNIYIWDNMTWQWDREINNYTDYIYTLLEDAQNQQEKNEQELILELQKASLWVWEDI  
 TNWLIWIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPSLQTHLPARRGPDPRPEGIGEEGERDRSRVRLVHGFLALVWEDL  
 RSLCLFSYHRLRDLILLIVARTVEILGQGWALKYWNWNLILYWSLELKNSAVSLVDTFIAIAVAEGTDRIIEIARRIFRAFLHIPT  
 RIRQGLERALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design

Fig. 35B

**QH0515.1g 140CF (651a.a)****Nick name: 012**

MRVKEIRRNQCQLRRWGTMLLGMMLICSATEQLWVTVYGVVWKEATTLFCASDAKAYVTEKHNWATHACVPTDPNPQEVVL  
 ENVTENFNWKNMNMVEQMHEDIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSGTNSSSWEKVQKEIKNCSEFNITTGIRGRVQ  
 EYSLFYKLDVIPIDSRNNSNNEFSSYRLISCNTSVITQACPKISFEPIPIHYCAPAGFAILKNDKFKFNGTGPCKNVSTVQCT  
 HGIKPVVSTQLLNGSLAEVEVIRSENFNTNPKSIIIVQLNKSVINCTRPNNTRKSIHIGAGKALYTGEIIGDIRQAHCNLSR  
 AQWNNTLKQIVIKLREQFGNKTIVFNQSSGGDVEIVMHSFNCGGEFFYCNSTQLENSTWNGNDTWNKTNDNITLPCRIKQ  
 IVNMWQVKGKAMYAPPPIRGQIRCSSKITGLILTRDGGTNGTNETETFRPGGNMKNRSELYKYVVKIEPLGIAPTAKARRV  
 QARLLSGIVQQNNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLRDQQLLGIWCSGRLICTINVPWNTSWSNRSLNIYIWDNMTWQWDREINNYTDYIYTLLEDAQNQQEKNEQELILELQKASLWVWEDI  
 NMTWQWDREINNYTDYIYTLLEDAQNQQEKNEQELILELQKASLWVWEDI

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.



Fig. 35C

CODON-OPTIMIZED QH0515.1g 140CF.seq (1984 nt.)  
Nick name:012

ttcaagtgcagccaccattgagagtaaaagaaatcagacgcaactgtcagagggttgaggagatggggaacgatgctcctgggcatt  
gctgatgatattgcagtgccaccgaaacagctttgggtaacctgtgtaactatggtgtacctgtatgaaagaaagccactacaacccttg  
ttttggcggctccgacgcgcaaaagcctacgtaacagaaaagcacaacgtgtggccacacacatgcatgctgccaacagatccaatc  
ctcaggaagtcggtctggaaaatgtaacagaaaaattttaatatgtggaanaacaaatattgtagagcagatgcatgaaagatatcat  
ctcactgtgggaacaactccttgaaaacctgtgtcaaacctgaccccaacttgcgtaaacacttaactgtactgataagcttcgcaat  
gatacgtccggaaacaaattcaagcagctgggaaaagtgcaaaaggcgaaatcaaaaattggttcaatttaacatcactaccggta  
tcagagggcgggtacaggaatattctctttctacaactcgaactcccaatcgaactcccaatcgaactcccaatcgaactcccaat  
agaatttagtagttatcgcccttataagctgcaaacaccagcgtgattacacaagcgtgcccataaaatctctttgagcccattcct  
attcaactgcccaccagccggcttcgcccattcctcaaatgtaacgcaagaatgtaacggaacccggaccctgtgaaagattgtgt  
ccaccggtcaatgcaactcaagccctcggtttctacccaacttctcaatggtagccttggcggagggaagttgt  
gattcggctccgaaaattttacaaacacgtcaagtcattcaatcgctccagcttaataaatccgctcgttatttaattgtacaagacc  
aacaataacaccagaaaattccattcacatagggccgggaaagctctgtataccgggaaatgtaacggaacccggaccctgtgaaagattgt  
actgtaaacttgagtcggccagtggaacacacattgaaacagatcgtgatcaagctcagagagcagttcgggaataagactat  
cgtgtttaatcagagctccggcggctgaaacggcaatgacacatgcaactcttttaattgtggggtgaaatttttttactgcaattct  
acaaattgtttaacagcaccctgaaacggcaatgacacatgcaactgaaatgacacctggaaagatcgcacaaatgataatattactcttc  
cgtgcagaataaagcaaatcgtaaatgtggcaaaaagtggcaaggccatgtaccgccaccctataagaggacaaaattcggctg  
ttcttccaagatcacagctctgatactcacacgggacgggacgaaacgggacaaacggagaccctccgaccagggaggc  
ggcaacatgaaggataactggagaagtgaactttacaagtataaagtgtcaagattgagcctctgggtatcgccctactaagg  
ctaaaacactcaccgtgcaggctagattgctgcttttcagggatagtcacaacacagaaacaaacttcttagagccattgaaagcaca  
acacacttgctgcagttgacagttggggaattaaacagttgcagggccgggttctcgctgtgaaatcctcatggagatcag  
cagcttttgggtatctgggggtgttcagggccctcatatgcaaccacaaatgtcccttggaaatcctcatggagtaacaggtctc  
ttaattatatttgggacaaatgacatggatgcaatgggatagaaaatttaataactacaccgactacatctacacacttctgga  
ggacggccagaattcagcaggaagaaacggagcaggaactcctcgaattggataaagtgggcattcactgtggaattggttcgata  
actaatggcctttggtaaatcttaca

61/178

## Fig. 36A

Wild-type subtype C

DU123.6 gp160 (854 a.a)

MRVKGIQRNWPQWIIWIGLGFWMIIICRVVGNLWVTVYGVVWTEAKTTLFCASDAKAYEREVHNWVWATHACVPTDPNPQEIIVL  
 GNVTFENFMWKNDMVDQMHEDIISWDQSLKPCVKLTPLCVTLNCTDVKNATSNGTTYNNSIDSMNGEIKNCSEFNITTEIRDK  
 KQKVYALFYRPDVVPLNENSSYIILNCNTSTTQACPVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCHNVSTVQCTHGIKP  
 VVSTQLLNGSLAEEIIIRSENLTNNAKTIIVHLNESIEIVCTRPNNTRKSIIRIGPGQTVYATNDIIGDIRQAHCNISKTKWN  
 TTLEKVKELKEHFPSKAITFQPHSGGDLEVTTHSFNCRGGEFFYCDTTLKLFNESNLNTTNTTLLPCRIKQIVNMWQGVGRAMY  
 APPVEGNITCNSSITGLLLVRDGGNTSNSTPEIFRPGGMMKDNWRSELYKYKVVVEIKPLGVAPTAKARRVVEREKRAVIGAVL  
 EGFAGAAGSTMGAASITLTVQARQLLSGIVQQSNLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDDQLLGLWGCSSGKLLIC  
 PTTVPWSSWSNKSQTDIWDNMTWMQWDREISNYTGTIYKLLSESONQOEKNEKDLLALDSWKNLWSWFDITNWLWYIKIFIMIV  
 GGLIGLRIIFGVLSIVKRVROGYSPLSFQTLTPNPRGLDRLGRIEEEGEGEQDKDRSIRLVNGFLALAWDDLRSCLFSYHRLRDF  
 ILVAARAVELLGRSSLRGLQRGWEALKYLGNLVQYGGLELKRRAISLFDITAIATAVAEGTDRILEVILRIIRAIRNIPTRIRQGFEE  
 AALL

## Fig. 36B

DU123.6 140CF (638 a.a)

Nick name: 013

MRVKGIQRNWPQWIIWIGLGFWMIIICRVVGNLWVTVYGVVWTEAKTTLFCASDAKAYEREVHNWVWATHACVPTDPNPQEIIVL  
 GNVTFENFMWKNDMVDQMHEDIISWDQSLKPCVKLTPLCVTLNCTDVKNATSNGTTYNNSIDSMNGEIKNCSEFNITTEIRDK  
 KQKVYALFYRPDVVPLNENSSYIILNCNTSTTQACPVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCHNVSTVQCTHGIKP  
 VVSTQLLNGSLAEEIIIRSENLTNNAKTIIVHLNESIEIVCTRPNNTRKSIIRIGPGQTVYATNDIIGDIRQAHCNISKTKWN  
 TTLEKVKELKEHFPSKAITFQPHSGGDLEVTTHSFNCRGGEFFYCDTTLKLFNESNLNTTNTTLLPCRIKQIVNMWQGVGRAMY  
 APPVEGNITCNSSITGLLLVRDGGNTSNSTPEIFRPGGMMKDNWRSELYKYKVVVEIKPLGVAPTAKKTLTVQARQLLSGIVVQQQ  
 SNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDDQLLGLWGCSSGKLLICPTTVPWSSWSNKSQTDIWDNMTWMQWDREISN  
 YTGTYKLLSESONQOEKNEKDLLALDSWKNLWSWFDITNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the

deleted fusion cleavage site.

Fig. 36C

**CODON-OPTIMIZED DU123.6 140CF.seq (1945 nt.)**

**Nick name: 013**

ttcagtcgacagccaccattgcccgttaaagggaattcaagaaattggccgcaatggtggattttgggaaattctgggcttttggat  
gataattatattgcccgtttgctggaaattttgtgggtgactgtgfactacgggtgcccgtgtggactgagcctgagcctgagcctg  
ttctgtgtgtagcgatgccaaagcctatgaaacgggaagtgcacaattgtttggctactcatgcctgtgtccctaccgaccctaaacc  
ctcaggaaatagtgctcggcaatgtaacggaaacttcaacatgtggaataatgatgtgtggatcagatgcacgaaagacattat  
ctcaattctggaccaaagcctgaaacccctgctgcttcaactgactcctctctgctcaactcaattgcacagatgtcaaatgcaat  
gccacctcaaacggtacgacaacttacaacaaattctattgactctatgaacggcgaatcaaaaaattgcttcttaacatcacca  
ccgagatccggacaaaaagcagaaggtctatgcccctttttaccgccccgaaagttagtcccactcaacgagaaattccagctcata  
cattcctcaactgcaatcatcaactacacacagcatgccgaaagttagctttgatcccaattcctaatacattactgccc  
ccggccgttacgctatactgaaatgcaatgcaatgaaacttttaacgggaccggcccaattgcacaaactgtcaaccgtgcaatgca  
ctcatggcatcaagccccgtgtgtcaaccagctgctgtcaatggctcaattgcagaagaagaaattattatccgctctgagaa  
tcttactaacaatgcaaaaacgattatcgtgcaccttaatgaatcaatagaatcgtgtgactcggccccacaataatactaga  
aaaagcattcgcattcggacctggccagacagtttacgcaactaatgacatcatcgggacatccgacagggcccatgcaacatt  
ctaaaaccgaatggaatcaaacctggaaaaagtaaggaaaaacttaagaaactttcccctaaaggcgatcacgcttcaacc  
tcacagtgccggagacttggaagtcaaacacatcttttaactgcccgggagaattttttattgtgatacacaaaaactttt  
aatgaaatcaaatctcaaccacaactgacctgcccctgtagaatcaacaaatcgtaaacatgtggcaagggttgcagggg  
ttggaaggctatgtacctcccctcgaaaggaataaacgttaacagcagcatcaactgggctgcttcttctgtagacgg  
aggcaatcttaattcaactcctgaaatttttaggcctggcgtggcaatgaaagataaactggcctcagaactgtacaaa  
tacaaagtgttgaattaaagccctgggagtcgctccaaccaaagctaaaactcaactgcaagcagaagacagctcctttcag  
gcatcgtccagcaacagtcaaatctccttagagcaatcgaagccccaaacagcattgctccaactcacaactgagctgggggattaaaca  
gcttcaagccccgctgcttgcctatcgaaacgctatcttaagaccacaagcttcttggcctctggggtttagtggaataactcattc  
tgccccaccaccgctgcttggaaatagttcttggagtaataatcacagaccgafatttgggacaacatgacctggatgcaatggg  
atagggaaatttctaatatatactggcacaactctacaaaactcttggaaagaaagtcaaaatcagcaagaaaaaaacgaaaagaccct  
cctcgcccctggactcctggaaagaaatctttggagctggttcgacataacttaattggctgtggtaagaatccttataca

Fig. 37A

Wild-type subtype CRF01\_AE  
 97CNGX2F-AE (854 a.a.)

MRVKETQMNWPNLWKWGTLILGLVICSASDNLWVTVYGVVWRDADTTLFCASDAKAHETEVEHNWATHACVPTDPNPQEIHL  
 ENVTENFNWNRNMMVEQMEDVLSLWDQSLKPCVKLTPLCVTLNCTNANWNTNSNNTTNGPNKIGNITDEVKNCTFNMTTELKDKK  
 QKVHALFYKLDIVQINSSEYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNNDKNFNGTGPCKNVSSVQCTHGKIPVVS  
 TQLLLNGSLAEEIIRSENLTNNAKTIIVHLNKSVEINCRPSNNTRTSITMGPGQVFYRTGDIIGDIRKAYCEINGIKWNEVL  
 VQVTGKLKEHFNKTIIFQPPSGDLEIITHHFSCRGEFFYCNITKLFNNTCIGNTSMEGCNCNTIILPCKIKQIINMWQGVGQAMY  
 APPISGRINCVSNITGILLTRDGGADNNTTNETFRPGGNIKDNWRSELYKYVVEIEPLGIAPTRAKRRVVEREKRAVGIAMI  
FGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSCGKIIC  
TTAVPWNSSWSNKSFEIWDNMWIEWEREISNYTSQIYEILTESQNRNEKDLLELDKWAASLWNVFDITNWLWYIKIFIIIV  
 GSLIGLRIIFAVLSIVNRVRQGYSPLSFQPTTHHQREPRPEEIGEQQKDRSVRLVSGFLALAWDDLRSLCLFSYHLLRDF  
 ILIAARTVELLGHSSLKGLRRGWEGLKYLGNLLLYWGQEIKISALSLLNATAIAVAGWTRVIEVAQRARALLHIPRRIRQGLE  
 RALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 37B

97CNGX2F-AE 140CF.pep (629 a.a.)

Nick name: 018

MRVKETQMNWPNLWKWGTLILGLVICSASDNLWVTVYGVVWRDADTTLFCASDAKAHETEVEHNWATHACVPTDPNPQEIHL  
 ENVTENFNWNRNMMVEQMEDVLSLWDQSLKPCVKLTPLCVTLNCTNANWNTNSNNTTNGPNKIGNITDEVKNCTFNMTTELKDKK  
 QKVHALFYKLDIVQINSSEYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNNDKNFNGTGPCKNVSSVQCTHGKIPVVS  
 TQLLLNGSLAEEIIRSENLTNNAKTIIVHLNKSVEINCRPSNNTRTSITMGPGQVFYRTGDIIGDIRKAYCEINGIKWNEVL  
 VQVTGKLKEHFNKTIIFQPPSGDLEIITHHFSCRGEFFYCNITKLFNNTCIGNTSMEGCNCNTIILPCKIKQIINMWQGVGQAMY  
 APPISGRINCVSNITGILLTRDGGADNNTTNETFRPGGNIKDNWRSELYKYVVEIEPLGIAPTRARLTVQARQLLSGIVQQQ  
SNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSCGKIICTTAVPWNSSWSNKSFEIWDNMWIEWEREISN  
YTSQIYEILTESQNRNEKDLLELDKWAASLWNW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

64/178

**Fig. 37C****CODON-OPTIMIZED 97CNGX2F-AE 140CF.seq (1921 nt.)****Nick name: 018**

ttcagtcgacagccaccatgCGAGTAAAGAGACACAAATGAATGGCCCAATTTGTGAAAGTGGGGAACATTGATCCTGGGACT  
 GGTGATAATCTGTAGTGCATCCGACAAATCTCTGGGTGACCGTTTACTATGGTGTACCCAGTTTGGAGAGACGCTGATACCCACCCCTC  
 TTCTGTGCAAGCGACGCCAAAGCCCAAGTCCATAATGTATGGCCACCCACCGTGGTACCAACCGGACCCCTAATC  
 CCCAAGAGATCCACCTTGAGAAATGTAAGTGAATTTAAACATGTGGAGAAATAACATGGTGAACAAATGCCAGGAAGACGTTAT  
 TTCCCTGTGGACCAGAGCCTTAAACCCTGTGCAAAATGACTCCCTGTGTGACTCTCAATTTGTACAAACGCCAAATTTGGACC  
 AACAGCAACAACACTACCAACGGCCCTAACAAAATTTGGCAATATTAAGTGAAGTCAAGAACTGCACCTTTTAAACATGACAACAG  
 AACTGAAGGATAAGAAACAGAAAGTCCATGCTCTGTCTATAAGCTCGACATAGTACAAAATTAATAGCTCAGAAATATAGACTGAT  
 AAATGCAATACTCCGTTATCAAAACAGCCCTGTCCAAAGATAAGCTTCGATCCCATCCCTATTCACACTACTGCACACCCAGCCGGT  
 TACGCTATCCTGAAATGCAACGATAAAGAAATTTAAACGGCACAGGTCCTGCAAAACAGTTTCCCTGTCCAGTGTACACACCGGTA  
 TCAAGCCTGTAGTATCAACACAACTGCTCCTGAATGGCTCCTTGGCCGAAAGAGATCAATCAATTAGAAAGTGAACACCTGACCGAA  
 CAACGCCAAGACTATAATAGTGCACCTCAATAAATCTGTAGAAATCAACTGTACCCGACCCCTCAACACACACTCGAACAAAGTATA  
 ACAATGGGCCCTGGCCCAAGTTTAAACGGACCCGGCACATAATAGCCGATATCAGAAAGGCAATATTCAGGAGATCAATGGCATCA  
 AGTGGAAACGAAGTACTGGTTCAGTAACTGGAAAACTCAAAAGAACATTTTAAATAGAACCAATAATTCAGCCCTCCGAGTGGCGG  
 CGACCTCGAGATTATCACCCATCACTTTTCTGTAGAGGCGAATTTTAACTGTAAACAGCACCAAGCTCTTCAATAAACACGTG  
 ATCGGGAACACTTCTATGGAAGGATGTAATAATACCAATTAATCTGCTCCCTGTAAAGTCAAGCAGATTATCAACATGTGGCAGGGAG  
 TAGGTCAGGCAATGTACGCCACCAGATTCAGGACGGATCAATTCGCTATCAATAATCACCGGCATTCCTGTGACCCCGGACCG  
 AGCCGAGACAACAATACCACTAACGAGACATTTAGACCTGGAGCGGCAATAATAAGGATAAATGGAGAAAGTGGAGCTGTATAAA  
 TACAAAGTCGTAGAGATCGAACCCCTCGGCATTTGACCTCCAAACCCGGGACTCTCAACCGTACCAAGCTAGACAGCTGCTTTCTG  
 GCATAGTCCAACAGCAGTCAAACCTCCTCCGCGTATTTGAAGCACAACAACACCTGCTCCAGCTGACTGTGTGGGGAATCAAACA  
 ATTGCAAGCAAGAGTGTCCGCGTGGAAACGCTATTTGAAAGATCAGAAATTTCTTGGACTTTGGGGTGCAGCGGCAAAATTTAT  
 TGTACAAACAGCGGTGCCCTTGGAACTCATCCTGGAGTAAATAAAGCTTTGAAGAAATCTGGGACAATATGACATGGATTGAGTGGG  
 AGAGAGAGATTTCAAACATAACAAGCCAAATTTACGAAATACAGAAAGTCAAACCCAGCAGGACAGAAATGAGAAAGACCT  
 GCTCGAACTGGATAAAGTGGGCCCTCTTTGTGGAACTGGtaaatcttataca

Fig. 38A

**Wild-type DRCBL-G (854a.a.)**  
 MRVKGIQRNWQHLLWNWGIILGLVVICSAEKLWVTVYGVVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPOEINMR  
 NVTENFNMWKNMVEQMHEIIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFNMTTELDRDKKAEYALFYR  
 TDVVPINEMNNENGTSTWYRLTNCNVSTIKQACPVTFEPIPIHYCAPAGFAILKCVDKKFNFTGTCNNVSTVQCTHGKPKVV  
 STQLLNGSLAEKDIIISSENI SDNAKVIIVHLNRSVEINCTRPNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVSWTKWNET  
 LRDVQAKLQEIFINKSIEFNSSGGDLEITTHSFNCGGEFFYCNTSGLFNNSILKSNISENNDTITLNCIKQIVRMWQRVGOAM  
 YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGDMKNNRSELYKYKTVKIKSLGIAPTRARRRVVEREKRAVGVGAIF  
 LGFLGTAGSTMGAASITITVQVRQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLRARVLALELYLKDQQLLGIWGCSGKLIC  
TNVPWNTSWSNKSNEIWNMTWIEWEREIDNYTHIYSLIEQSIIQEKNEQDLLALDQWASLWSWFSISNWLWYIRIFVMIV  
 GGLIGLRIVFAVLSIVNRVROGYSPLSFQTLHHQREPD RPAGIEEGGEGQDRDRSIRLVSGFALAWDDLRSLCLFSYHRLRDF  
 ILIAARTVELLGRNSLKGLRGWEALKYLWNLLLYWARELKN SAINLLDTIAIAVANWTDRIEVAQRAGRAVLNIPRRIRQGLE  
 RALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 38B

**DRCBL-G 140CF.pep (630 a.a.)**  
**Nick name: 017**  
 MRVKGIQRNWQHLLWNWGIILGLVVICSAEKLWVTVYGVVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPOEINMR  
 NVTENFNMWKNMVEQMHEIIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFNMTTELDRDKKAEYALFYR  
 TDVVPINEMNNENGTSTWYRLTNCNVSTIKQACPVTFEPIPIHYCAPAGFAILKCVDKKFNFTGTCNNVSTVQCTHGKPKVV  
 STQLLNGSLAEKDIIISSENI SDNAKVIIVHLNRSVEINCTRPNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVSWTKWNET  
 LRDVQAKLQEIFINKSIEFNSSGGDLEITTHSFNCGGEFFYCNTSGLFNNSILKSNISENNDTITLNCIKQIVRMWQRVGOAM  
 YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGDMKNNRSELYKYKTVKIKSLGIAPTRARRRTLTVQVRQLLSGIVQOO  
SNLLRAIEAQHLLQLTVWGIKQLRARVLALELYLKDQQLLGIWGCSGKLICTNVPWNTSWSNKSNEIWNMTWIEWEREIDN  
YTYHIYSLIEQSIIQEKNEQDLLALDQWASLWSW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

Fig. 38C

CODON-OPTIMIZED DRCBL-G 140CF.seg (1921 nt.)

Nick name: 017

ttcagtcgacagccaccatgagagttaaaaggaatccaacggcaattggcaacacaccttggaaactggggcattattgatcttggact  
 ggtgataatttgtagcgctgaaaaacttggtaactgtgtacggcgtccctgttggaggatgccaacgccccctgttcc  
 tggcgaagtgatgcaaaaggctcacagcactgaaatctcaacaacttgggccaccacgcctgtgtgccaaaccgacctagtcctc  
 aggagatcaacatgagaaaacgttacccgaaattttaatatgtggaagaaataatggtggacaaaatgcacgaaagacataatttc  
 acttgggacgagctctgaaaccatgtgtgaaacttacccttaccctgtgcctgacctgaaactgtaccgaaatcaacaataactca  
 acgagaaatatcacagaagataccgaatgactaaactgttcccttaatatgacaaccgaaactggagacaaaagaggctgaat  
 acgcactttctaccgaaacagatgtgtaccaatcaacgagatgaaacaatgaaaacaatggaaccgaaacttacctactgcgcaccggcc  
 gacaactgtaacgttagcacaatcaagcagctgtgtaacgagtttaacggaaactgcaaccatgaaacaatggaaccgaaacttacctactgcgcaccggcc  
 ggattcggctattcttaagtccgtggataagagtttaacggaaactgcaaccatgaaacaatggaaccgaaacttacctactgcgcaccggcc  
 gaattaaagcctgtcgtttcaacccagttgctgactgaaatggatcactccagaaaaggtattattatctcaagccgaaaacataatc  
 tgataatgcaaaaggtcatcatcgtccactcaacccgtcagttgaaataaaactgcactcggcctaataataaacacaagacgctct  
 gtcgcaatcggccccaggacaagcttttacactaccgggaaagtattcggcgacataccgaaagcctcactgcaacgtagctgga  
 ccaagtggaatgaaacactgcccgatgttcaagcacaacttcaagaatactcataaaacaataatcaattgagttcaattctagctc  
 tggcggcaccctcgagattacaactcactcctttaaactggcgaggaaattcttttattgtaatactccggctcttcaacaac  
 tctatctcaaaagtacaatttctgaaaataatgacacaataatgcaagatcaagcagattgttaggatgtggcaac  
 gactcggacaagctatgtacggccccaccatcggcggaaataataacgtgtcgatcaaatcactggcctcatccttactagaga  
 tggcggagacataatagcaccagcgagataatcagaccagggcgagctatgaaaaacaactggaggctcagagctctacaag  
 tacaaaacagtcataaattaaaagcttggcattgtcccactcggccccgacactgactgtccaagtccgacagctcctgtccg  
 gaatcgtccacaacagtcctgctgcccgtatagaggctcaacaacatctccttcaactgactgtgtggggatcaaaa  
 attgagagcaagagtgctggcgtggaacggatcttaaggaccacaactcctgggcatatgggggtgttccggcacaactgatc  
 tgcacaaacaaatgtaccctggaacaccagctggtcaaaataaagttataatgagataatgggaaaacatgacatggattgaatggg  
 aaagggaattgacaattatacataccatatactctcctcaatcgaaacaatctcagatacaacaagaaaagaaatgaaacaagattt  
 gttggctcttgaccaaatgggcttcttgggattggtaaatcttcaaa

2003 Centralized HIV-1 Envelope Proteins and the Codon-Optimized Gene sequences

Fig. 39A

2003 Cons Env  
 MRVMGIQRNCQHLLWRWGILLFGMLLICSAAENLWTVVYGVVWKEANTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIIVLENVVTENF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTNNEEIKNCSFNIITEIRDKKQVYALFYKLDVVPIDDDNNSYRLI  
 NCNTSAITQACPVSFEPPIHYCAPAGFAILKCNDDKFKNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEIIIRSENIITNNAKTIIV  
 QLNESVEINCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHNCISRTKWNKTLQQVAKKLREHFNKTIIFNPSSGGDLEITTHSFNCGGE  
 FFYCNTELSEFNSTWNGTNETITLPCRIKQIINMWQVGVQAMYPPIEGKIRCTSNITGLLLTRDGGNNNTEFFRPGGDMRDNRWSELKYK  
 VVKIEPLGVAPTKAKRRVVEREKRAVGI GAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQAR  
 VLAVERYLKDQQLLGIWGCSSGKLICTTNVPWSSWSNKSQDEIWDNMTWMEWDKEINNYTDIYSLIEESQNOQEKNEQELLALDKWASLWN  
 WFDITNWLWYIKIFIMIVGGLIGLRIVFVLSIVNRVRQGYSPLSFQTLIPNPRGPDREGEIEEGEQDRDRSIRLVNGFLALAWDDLRSL  
 CLFSYHRLRDLILLIARTVELLGRRGWEALKYLWNLLOQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRCRAILLNIPRRIRQGFERAL  
 LLS

Fig. 40A

2003 M. Group .AnC. Env  
 MRVMGIQRNCQHLLWRWGILLIFGMLMICSAAENLWTVVYGVVWKEANTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIIVLENVVTENI  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTNNGEIKNCSFNIITEIRDKKQVYALFYRLDVVPINDNNSYRLI  
 NCNTSAITQACPVSFEPPIHYCAPAGFAILKCNDDKFKNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEIIIRSENIITDNAKTIIV  
 QLNESVEINCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHNCISGAENKTLQQVAAKLREHFNKTIIFKPSSGGDLEITTHSFNCGG  
 EFFYCNTEGLFNSTWNGTNETITLPCRIKQIIVNMWQVRVQAMYPPIAGNITCKSNITGLLLTRDGGTNNTEFFRPGGDMRDNRWSELKYK  
 KVKIEPLGVAPTKAKRRVVEREKRAVGI GAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQAR  
 RVLAVERYLKDQQLLGIWGCSSGKLICTTNVPWSSWSNKSQDEIWDNMTWMEWDKEINNYTDIYSLIEESQNOQEKNEQELLALDKWASLW  
 NWFDTNWLWYIKIFIMIVGGLIGLRIVFVLSIVNRVRQGYSPLSFQTLIPNPRGPDREGEIEEGEQDRDRSIRLVSGFLALAWDDLRSL  
 LCLFSYHRLRDFILLIARTVELLGRRGWEALKYLWNLLOQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRCRAILLNIPRRIRQGFERA  
 LLS

67/178



69/178

## Fig. 40B

**2003 M. Group .anc Env. seq. opt**  
 ATGCGGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCATCCTGATCTTCGGCATGCTGATGATCTGCTCCGCCCGCCGA  
 GAACCTGTGGTGAACCGTACTACGGCTGCCGTGGAAGAGGGCCAAACACCACCCCTGTTCTGGCCCTCCGACGCCAAGCCCTACGACA  
 CCGAGTGCACAACCGTGGCCACCCACCGTCCACCGACCCCAACCCAGGAGATCGTCTGGAGAACGTGACCCGAGAACTTC  
 AACATGTGAAGAACAACTGGTGGAGAGATGCACGAGGAGATCATCTCCCTGTGGGACAGTCCCTGAAGCCTGCGTGAAGCTGACCCC  
 CCTGTGCTGACCTGAACTGCACCCGAGTGAACCCACCAACAACTCCACCAACATGGGGAGATCAAGAACTGCTCCTTCAACATCACCA  
 CCGAGATCCGGACAAGAAAGTGAAGGTGACGCCCTGTTTACCGCTGACCGTGGTCCCATCAACGAAACAACTCCCTACCGCTGAT  
 AACTGCAACACTCCGCCATCACCCAGGCTGCCCAAGGTGCTTCGAGCCCATCCCATCACTACTGGCCCCCGGCTTCGCCAT  
 CCTGAAGTGAACGACAAGAGTTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTCCAGTGCAACCCAGGATCAAGCCCGTGGTGT  
 CCACCCAGTGTCTGAACGGCTCCCTGGCCGAGGAGAGATCATATCCGTCGAGAACATCACCGAACACCGCCAAAGACCATCATCGTG  
 CAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGCGAGCCCTTACGG  
 CACCGGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCGGCGCGAGTGAACAAGACCTTGACAGCAGGTGGCCGCCAAGC  
 TGCCGGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCCTCCGGCGGGGACTGGAGATCACCCACTCTCTTCAACTGCCGGCG  
 GAGTCTTCTACTGCAACACTCCGGCTGTTCAACTCCACTGGAAACCGCACCAACGAGACCATCACCTGCCCTGCCGCATCAAGCAGAT  
 CGTGAACATGTGCAGCGGCTGGCCAGGCCATGTAGCCCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCCGGCTGTGC  
 TGACCCCGACGGCGGCACCAACAACACCCGAGACTTCCGCCCGGGCGGGACATGCGGACAACCTGGCGTCCGAGCTGTACAAGTAC  
 AAGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGCCAAAGCCCGTGGTGGAGCGGAGAACGGCCGTGGCATCGGCG  
 CGTGTCTGGCTTCTGGCCCGCCGGTCCACCATGGCCCGCTCCATCACTGAGCCCTGACCGTGCAGCCCGCCAGCTGCTGCCGGC  
 TCGTGCAGCAGTCAACCTGCTGGCCCATCGAGCCACGACGACCTGCTGAGCTGACCGTGGGCGATCAAGCAGCTGCAGGC  
 CGCGTGGCGTGGAGCGTACCTGAAGGACCGAGCTGCTGGGCTGCTCGGCAAGCTGATCGCACCCCAACCGTGC  
 CTGGAACCTCTCTGGTCCAACAGTCCAGGACGAGATCTGGGACAACATGACCTGGATGCAGTGGGAGCGGAGATCTCCAACATCACCC  
 ACATCATCTACTCCCTGATCGAGGATCCCAAGAACCCAGGAGAAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGCCCTCCCTGTGG  
 AACTGGTTCGACATCAACTGCTGTGTATCATCAAGATCTTATCATGATCGTGGCGGCTGATCGCCCTGGGCACTGCTGTCCCGGT  
 GCTGTCCATCGTGAACCGCTGGCCAGGCTACTCCCCCTGCTCTCCAGACCCCTGATCCCCAACCCCGGACCCGACCCCGCCGCGG  
 GCATCGAGGAGGGCGGAGACCGGACCGCTCCATCCGCTGGTGTCCGGCTTCTGGCCCTGGCCCTGGACGACCTGGCTCC  
 CTGTGCTGTTCTCTTACCACCGCTCGGACTTATCTGATCGCCCGCCGACCGTGGAGCTGCTGGCCCGCCGGCTGGAGGCGCTCC  
 GAAGTACCTGTGGAACCTGCTGACGTACTGGGGCCAGGAGTGAAGAACTCCGCCATCTCCCTGCTGGACACCAACCGCCATCGCCGTGGCCG  
 AGGGCACCCGCGGTGATCGAGGTGGTGCAGCGCGCTGCCGGCCATCTCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC  
 CTGCTGTAA

70/178

Fig. 41A

**2003 CON A1 Env**

MRVMGIQRNCQHLLRWGTMILGMIIICSAAENLWVTVYYGVPVWKDAETILECASDAKAYETEMHNVWATHACVPTDPNPQEIHLNVTEEF  
 NMWKNMVEQMHADIISLWDQSLKPCVKLTPLCVTLNCSNVVNTNTTHEEEIKNCSEFMTTELRDCKKQKQVYSLFYRLDVVPIENNSNS  
 SYRLINCSAITOACPVSFEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGKIPVYSTQLLLNGSLAEFEVIRSENIITNNA  
 KTIIVQLTEPVKINCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHCVSRSEWNKTLOKVAKQLRKYFKNKTIIFNSSGGDEITTHS  
 FNCGGEFFYCNTSGLFNSTWNGTMKNTITLPCRKQIINMWQVQAGYAPPIQGVIRCESNITGLLFRDGGNNNTNETFRPGGGDMRDN  
 WRSELYKVKVVKIEPLGVAPTRAKRRVVEREKRAVIGAVELGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLKLTIV  
 WGIKQLQARVLAVERYLKDQQLLGIWGCCKLICTTNPVWSSWSNKSQNEIWDNMTWLQWDKEISNYTHIIYNLIEESQOQKNEQDLLA  
 LDKWANLWNWFEDISNWLWYIKIFIMIVGGGLIGLRIVFAVLSVINRVROGYSPLSFQHTPNRGLDRPGRIEEEGEGEQGRDRSIRLVSGLLA  
 LAWDDLRSCLFSYHRLRDFILIAARTVELLGHSSILKGLRLGWEGKYLWNLLLYWGRELKISAINLVDTIAIAVAGWTDREVIEIGQRIGRA  
 LLHIPRRIRQGLERALL\$

Fig. 42A

**2003 A1.Anc Env**

MRVMGIQRNCQHLLRWGTMIFGMIIICSAAENLWVTVYYGVPVWKDAETILECASDAKAYDTEVHNVWATHACVPTDPNPQEIHLNVTEEF  
 NMWKNMVEQMHADIISLWDQSLKPCVKLTPLCVTLNCSNVVNTNTTHEEEIKNCSEFMTTELRDCKKQKQVYSLFYRLDVVPIENNSNS  
 SYRLINCSAITOACPVSFEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGKIPVYSTQLLLNGSLAEFEVIRSENIITDNA  
 KTIIVQLTEPVKINCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHCVSRSEWNKTLOKVAALRKHFNKTIIFNSSGGDEITTHS  
 FNCGGEFFYCNTSGLFNSTWNGTMKNTITLPCRKQIINMWQVQAGYAPPIQGVIRCESNITGLLFRDGGNNNTNETFRPGGGDMRDN  
 WRSELYKVKVVKIEPLGVAPTRAKRRVVEREKRAVIGAVELGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLKLTIV  
 WGIKQLQARVLAVERYLKDQQLLGIWGCCKLICTTNPVWSSWSNKSQDEIWDNMTWLQWDKEISNYTHIIYNLIEESQOQKNEQDLLA  
 LDKWANLWNWFEDISNWLWYIKIFIMIVGGGLIGLRIVFAVLSVINRVROGYSPLSFOTLTPNPEGDRPGRIEEEGEGEQGRDRSIRLVSGLLA  
 LAWDDLRSCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGKYLWNLLLYWGRELKISAINLLDTIAIAVAGWTDREVIEIGQRICRA  
 LLNIPRRIRQGLERALL\$

Fig. 41B

2003 CON A1 Env. seq. opt  
ATGGCGTGTGGGCATCCAGCGCAACTGCCAGCACCTGCTGGCTGGGCAACCATGATCCTGGGCATGATCATCATCTGCTCCGGCCCGG  
GAACCTGTGGGTGACCGTGTACTACGGGTGCCCGTGGAAAGGACCGGAGACCACTGTTCTGGCCCTCCGACGCCAAGGCCTACGAGA  
CCGAGATGCACAACAGTGTGGCCACCCACCGCTGCGTGCCACCGACCCCAACCCAGGAGATCCACCTGGAGAACGTGACCGGAGGTTCC  
AACATGTGAAGAACAACATGTTGGAGCAGATGCACACCCACATCTCCCTGTGGACCCAGTCCCTGAAGCCCTGCGTGAAGCTGACCC  
CCTGTGGTGAACCTGAATGCCAAGCAACCTGACCCCAACACCAACCAACCCACGAGGAGGATCAAGAACACTGCTCTTCAACA  
TGACCACCGAGTGGCGACAAGAAGCAGAGGTACTCCCTGTTTACCGCTGGACGTGTTGCAGATCAACGAGAACAACCTCCAACTCC  
TCCTACCGCTGATCAACTGCAACCTCCGCCATCACCCAGGCTGCCCAAGGTGTCTCTTGGAGCCATCCCATCCACTACTGGCC  
CGCCGGCTTCGCCATCCTGAAGTGAAGCAAGGAGTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCCGTGACGTGACCCCGGCA  
TCAAGCCCGTGGTCCACCCAGCTGCTGTGAACGGTCCCTGGCCGAGGAGGAGTGTATCATCCGCTCCGAGAACATCACCAACAACGCC  
AAGACCATCATGTCAGCTGACCAAGCCCGTGAAGATCAACTGCACCCGCCCAACCAACAACACCCGCAAGTCCATCCGCATCGGCC  
CCAGGCTTCTACGCCACCGGACATCATCGCGACATCCGCCAGGCCACTGCAACGTGTCCCGTCCGAGTGAACAAGACCCCTGCAGA  
AGGTGGCCAAAGCTGGCAAGTACTTCAAGAACAAAGACCTCATCTTCAACCAACTCCTCCGGCGGACCTGGAGATCACCAACCCACTCC  
TTCAACTGGCCGATCAAGCAGATCAACATGTGGCAGCGGCCGCGGCAACCAACGAGACTTCCGCCCCCGGCGGCGGACATGGCCACTCC  
GCCCTGCCGATCAAGCAGATCAACATGTGGCAGCGGCCGCGGCAACCAACGAGACTTCCGCCCCCGGCGGCGGACATGGCCACTCC  
CCAAATCACCGGCTGCTGACCGCGGCAACCAACGAGACTTCCGCCCCCGGCGGCGGACATGGCCACTCC  
TGGCGTCCGAGCTGACAAAGTAAAGTGGTGAAGATCGAGCCCTGGCGTGGCCCGCCACCCGCGCAAGCGCGCGTGGTGGAGCGCGA  
GAAGCGCCCGTGGCATCGGCCCGTGTCTGGGTCTCTGGCGCCCGCGGCTCCACCATGGCGCCGCTCCATCACCCCTGACCCGTGC  
AGGCCCGCAGCTGTCCGGCATCGTGCAGCAGCTCAACCTGCTGGCGCCATCGAGGCCCAAGCAGCAGCTGCTGGCATCTGGGCTGCTCCGGCAA  
TGGGCTCAAGCAGCTGCAGGCCCGTGTGGCGTGGAGCGCTACCTGAAGGACCAAGCAGCTGCTGGCATCTGGGCTGCTCCGGCAA  
GCTGATCTGCACCAACAGTGCCTGGAATCCTCTGGTCCAAAGTCCAGAACGAGATCTGGGACAAACATGACCTGGCTGCAGTGGG  
ACAAAGGATCTCCAACTACACCCACATCATACAACTGATCGAGGAGTCCAGAACCCAGGAGGAAAGAACGAGCAGGACCTGCTGGCC  
CTGGACAAGTGGCCAACTGTGAACTGGTTCACATCTCCAACTGGCTGTGGTACATCAAGATCTTCCATCATGATCGTGGCGGCTGAT  
CGGCTGGCATCGTGTCCCGTGTCTCCGTGATCAACCGCTGGCCAGGGCTACTCCCGCTGTCTCCAGACCCACCCCGCTGAT  
CCCCGGCTGGACCGCCCGGCGCATCGAGGAGGCGGCGGAGCGGCGGACCGCTCCATCCGCTGATCGCCCGCCGACCGTCCCTGGCC  
CTGGCCCTGGACCGACTGGCTCCCTGTGCTGTCTCCTACCCCGCTGCGGACTTCATCTGATCGCCCGCCGACCGTGGAGCTGCT  
GGGCCACTCCTCCCTGAAGGCCCTGGCCCTGGCTGGAGGCGCTGAAGTACTGTGGAACTGTGTGTGTACTGGGCGCGGAGCTGAAGA  
TCTCCGCCATCAACCTGGTGGACACCATCGCCGTGGCCGCTGGACCGCGGCTGATCGAGATCGGCCAGCGCATCGGCCCGGCG  
ATCCTGCACATCCCCCGCATCCGCCAGGGCTGGAGCGCGCCCTGCTGTAA

Fig. 42B

2003 Al. anc Env. seq. opt

ATGGCGGTGATGGGCATCCAGCCAACTGCCAGCACCTGTGGCGCTGGGCACCAATGATCTTCCGCATGATCATCATCTGCTCCGCCGCCGA  
 GAACCTGTGGTGACCGTGTACTACGGCGTGCCTGTGGAGGACGCCAGACCACCTGTCTGCGCTCCGACGCCAAGCCCTACGACA  
 CCGAGGTGCACAACGTTGGGCCACCCACCGCTGCGTGGCCACCCCAACCCAGGAGATCGACTGGAGAACCTGACCCGAGGAGTTC  
 AACATGTGAAGAACAACATGGTGGAGCAGATGACCGCGACATCATCTCCTGTGGGACAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCTGAACTGCTCAACCTGAACTGACCAACAACACCAACCCAGGAGGAGATCAAGAACTGCTCCTTCAACA  
 TGACCACCGAGTGGCGACAAGAGCAGAGGTGTAACCTGTTTACCGCTGGAGCTGTTGCCATCAACGAGAACAACTCCAATCC  
 TCCTACCGCTGATCAACTGCAACACTCCGCCATCACCCAGGCTGCCCAAGGTGCTCCTGCAAGAACGTGTCACCCATCCACTACTGCGCCC  
 CGCCGGCTTCGCCATCCTGAAGTGAAGGACAAGGATTCACCGGACCGGCCCTGCAAGAACGTGATCCGCTCCGAGAACATCACCGACAACGCC  
 TCAAGCCCGTGGTCCACCCAGCTGCTGTGAACGGTCCCTGGCCGAGGAGGAGTGTGATGATCCGCTCCGAGAACATCACCGACAACGCC  
 AAGACCATCATCGTGCAGTACCGGACCGCTGAAGATCAACTGCAACCCGCCCAACCAACAACACCCGAACTCCATCCGATCGGCCCGG  
 CCAGCCCTTACGCCACCCGGGACATCATCGCGACATCCGCCAGGCCACTGCAACGTGTCCCGCACCCGAGTGGAAACAAGACCCCTGCAGA  
 AGGTGGCCCGCAGTCCGCAAGCACTTCAACAACAAGACCATCATCTTCAACTCCTCCCTGGCGGCGGACCTGGAGTCAACCCACTCC  
 TTCACTGCGGGCGGAGTCTTCTACTGCAACACTCCGGCTGTCAACTCCACTGGAACAACCGCACCATGAAGGACACCATCACCT  
 GCCCTGCCGATCAAGCAGATCATCAACATGTGGCAGCGGTGGCCAGGCCATGACGCCCCCTCCATCCAGGGCGTGTCCGCTGGAGT  
 CCAACATCACCGCCCTGCTGTACCCCGGACGGGCAACAACAACCAACGAGACTTCCGCCCCCGGCGGACATGCGCGGACATGCGCGACAAC  
 TGGCGTCCGAGCTACAAGTACAAGTGGTGAAGATCGAGCCCTGGCGTGGCCCAACCAACCAACCGCCGCAAGCCGCTGGTGGAGCGCGA  
 GAAGCGCCCGTGGCCCTGGCGCCGTTCTGGCTTCTGGCGCCCGGCTCCACATGGCGCCGCTCCATCACCCCTGACCCGCTG  
 AGCCCGCCAGCTGTCCGGCATCGTGCAGCAGTCAACCTGTGCGGCCATCGAGGCCAGCAGCCTGAGGCCAGCAGCCTGCTGAAGCTGACCCGTG  
 TGGGCATCAAGCAGTGCAGCCCGGCTGGCCGTGGAGGCTACCTGAAGGACCAAGCAGTGTGGGATCTGGGGCTGCTCCGGCAA  
 GCTGATGCAACCAACCGTGGACTCCTCCTGGTCCACAAGTCCAGGACGAGATCTGGGACAACATGACCTGGTGCAGTGGG  
 ACAAGGAGATCCCAACTACCCGACATCTACAACCTGATCGAGGAGTCCAGAACCAAGGAGGAGAAACGAGCAGACCTGCTGGCC  
 CTGGACAAAGTGGCCAACTGTGGAACCTGTTGACATCTCAAATGGCTGTGGTACATCAAGATCTTCAATCATGATCGTGGCGGCTGAT  
 CGCCTGGCATCGTGTGCGCGTGTGTCGATCAACCGGTGGCCAGGGTACTCCCTTGTCTTCCAGACCTGACCCCAACC  
 CCGAGGCCCGGACCCCGGCGATCGAGGAGGAGGCGGGGAGCAGGCCGCGACCGTCCATCCGCTGCTGGCTCCGCTGCTGGCC  
 CTGGCTGGGACGACCTGCGCTCCCTGTCCTTCTTCTACCCAGCTGGCGACTTCACTGTGATCGCCGCGGCTGCTGCTGCTGGGCGGAGTGAAGA  
 GGGCCGCTCCTCCCTGAAGGGCTTGGCTGGGCTGGAGGGCTGAAGTACCTGTGGAACCTGCTGCTGCTGCTGGGCGGCGGAGTGAAGA  
 TCTCCGCCATCAACTGTGGACACCATCGCCATCGCCGTGGCCGCTGGACCCGACCCGCTGATCGAGATCGGCCAGCGCATCTGCCGGCC  
 ATCTGAACATCCCCCGCCGATCCGCCAGGGCTGGAGCGGCCCTGCTGTAA

Fig. 39B

2003 CON-S Env. seq. opt

ATGGCGGTGATGGGCATCCAGCACTGCCAGCACCTGTGGCGCTGGGGCACTCTGATCTTTGGCATGTGATCATCTGCTCCGCCGCCGA  
 GAACCTGTGGGTGACCCGTACTACGGCGTGGCGGTGGAAAGGAGGCCAACACACCCCTGTCTGGCCCTCCGACGCCAAGGCCTACGACA  
 CCGAGGTGCACAACCTGTGGCCACCCACCGCTGGTGGCCACCCAGACCCCAACCCAGGAGATCGTGTGGAGAACGTTGACCCGAGAACTTC  
 AACATGTGGAAGAACAACTGGTGGAGCAGATGCACGGAGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCCGTGACCCCTGAACTGCCACCGAGCTGAACGCCCAACCAACACCCCAACAGGAGATCAAGAACTGCTCTTCAACATCACCA  
 CCGAGATCCGGCACAAAGAAAGGTGTACGCCCTGTTCTACAAGCTGGACGTGGTCCCATCGACGACAAACAATCCTTACCGCTGATC  
 AACTGCAACACCTCCGCCATCACCCAGGCTGCCCAAGGTGTCCTTCAGGCCATCCCATCTACTACTGCGCCCGCCCGCTTCGCCAT  
 CCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTCAAGTGCACCCACGGCATCAAGCCCGTGGTGT  
 CCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGATCATCTCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTG  
 CAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACCCGCAAGTCCATCCGCATCGGCCCGGCCAGGCTTCTACGC  
 CACCGGGACATCATCGGGACATCCGCCAGGCCACTGCAACATCTCCGCAACCAAGTGAACAAGACCCCTGCAGCAGGTGGCCAAAGAGC  
 TGGCGGACACTTCAACAAGACCATCATCTCAACCCCTCTCCGGCGGCACTGGAGATCACCAACCCACTCTTCAACTGCGGGCGGAG  
 TTCTTCTACTGCAACACCTCCGAGCTGTTCAACTCCACTGGAACGGCACCAACAACACCATCACCTGCCCTGCCCATCAAGCAGATCAI  
 CAACATGTGGCAGGGCGTGGCCAGGCCATGTACGCCCCCTCCATCGAGGGCAAGATCCGCTGCACCTCCAACATCACCGCCCTGCTGTGT  
 CCCGCCAGCGGGCAACAACAACACCGAGACTTCCGCCCCCGGGCGGACATGCGGACAACTGGCGTCCGAGCTGACAGCTGACAACTGTA  
 GTGGTGAAGATCGAGCCCTGGCGGTGGCCCAACCAAGCCAGCCGCGTGGTGGAGCGGAGAACGCCGCTGGCATCGGGCCATCGGGCCCGI  
 GTTCTGGGCTTCTGGCGCCCGGGCTCCACCATGGCGCCCTCCATCACCTGACCTGCAGGCCCGCCAGCTGTCCGGCATCC  
 TGACGACAGTCCAACCTGCTGGCGCCATCGAGGCCAGCAGCTGCTGGCATCTGGGGCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCA  
 GTGCTGGCCGTGGAGCGCTACCTGAAGGACCCAGCAGCTGCTGGCATCTGGGGCTGCTCCGGCAGCTGATCTGCACCAACCAACGTTGCCCTG  
 GAACTCCTCCTGTTCCAAAGTCCCAGGACGAGATCTGGACAACATGACCTGGATGGAGTGGACAAGGAGATCAACAACATCACCCGACA  
 TCATCTACTCCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAAGAACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGCCCTCCCTGTGGAA  
 TGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAATGATCGTGGCGGCCCTGATCGGCCCTGGCCATCGTGTTCGCCGTGCT  
 GTCCATCGTGAACCGGTGCCAGGGCTACTCCCCCTGTCTTCCAGACCTGATCCCCAACCCCGGCCCGCCGACCCCGAGGGCA  
 TCGAGGAGGAGGGCGGAGCAGGACCGGACCGCTCCATCCGCTGGTGAACGGCTTCTGGCCCTGGCTGGACGACCTGCGCTCCCTG  
 TGCCCTGTCTCTACCAACCGCTGCGGACCTGATCTGATCGCCCGCCGACCCGTTGGAGCTGCTGGCCCGCCGCGGCTGGAGGCCCTGAA  
 GTACCTGTGAACTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGACACCAACCCGCTCGCCGTGGCCGAGG  
 GCACCCGACCGGTGATCGAGGTGGTGCAGCGGCTGTGCCCGGCCATCTTGAACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGGCCCTG  
 CTGTAA

Fig. 43A

2003 CON A2 Env

MRVMGTORNYQHLLWRWGILILGMLIMCKATDLWVTVYYGVVWKDADTTILFCASDAKAYDTEVHNWVWATHACVPTDPNPQEVNLENVTEDFN  
 MWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSNANTNNTNSTMEEIKNCSYNTITTELKDKTKQVYSLFYKLDVVQLDESNKSEYYR  
 LINCNTSAITQACPKVSEFPIPIHYCAPAGFAILKCKDPRFNGTSCNNVSSVQCTHGKIPVASTQLLLNGSLAEGKVMIRSENITNNAKNI  
 IVQFNKPVPI TCIRPNNTRKSIRFGPGQAFYTNDIIGDIRQAHCNINKTKWNATLQKVAEQLRHFHPNKTIIFTNSSGGDLEITTHSFNCG  
 GEFYCNLTGLEFNSTWKNGTNNTTEQMITLPCRKQIINMWQRVGRAMYAPPIAGVIKCTSNITGIILTRDGGNNETETFRPGGGMDRDNWR  
 SELYKYKVKVIEPLGVAPTRAKRRVVEREKRAVGMGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQOQSNLLKAIQAQQHLLKLTVWG  
 IKQLQARVLALERYLDQQLLGIWGCSSGKLI CATTVPWNSWSNKTQEEIWNMMTWLQWDKEISNYTNI IYKLEESQNOQEKNEQDILLALD  
 KWANLWNWFENITNWLWYIRIFIMIVGGLIGLRIVIAIISVVNRVRQGYSPISFQIPTPNPEGLDRPGRIEEGGEGQGRDRSIRLVSGFLALA  
 WDDLRSCLFSYHRLRDCILIAARTVELLGHSSLKGLRLGWGLKYLWNLLLYWGRELKNSAISLLDITIAVAVAEWTDRVIEIGQRACRAIL  
 NIPRRIRQGFERALL\$

Fig. 44A

2003 CON B Env

MRVKGIRKNYQHLLWRWGTMLLGMLMICSAAEKLWVTVYYGVVWKEATTTILFCASDAKAYDTEVHNWVWATHACVPTDPNPQEVNLENVTEF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDEMNATNTNTI IYRWRGEIKNCSFNITTSIRDKVKQKEYALFYKLDVVPIDND  
 NTSYRLISCNTSVITQACPKVSEFPIPIHYCAPAGFAILKCNCKKFNCTGCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEVIRSENFTD  
 NAKTII VQLNESVEINCTRPNNTRKS IHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKOIVKKLREQFGNKTIVFNQSSGGDPEIVM  
 HSFNCGGEFFYCNTTQLEFNSTWNGTWNTEGNITLPCRKQIINMWQEVGKAMYAPP IRGQIRCSSNITGLLLITRDGGNNETEIFRPGGGDM  
 RDNWRSELYKYKVKVIEPLGVAPTKAKRRVVQREKRAVGI GAVFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQOQSNLLRAIEAQQHLLQ  
 LTVWGIKQLQARVLAVERYLKDQQLLGIWGCSSGKLI CTTAVPWNASWSNKSLEIWDNMTWMEWEREIDNYTSLIYTLIEESQNOQEKNEQE  
 LLELDKWAASLWNWFENITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPISFQTRLPAPRGPDRPEGIEEGGERDRDRSGRLVDG  
 FLALIWDLLRSCLFSYHRLRDLLEIIVTRIVELLGRRGMEVLKYWNLLQYWSQELKN SAVSILLNATAIAVAEGTDRVIEVQACRAILHI  
 PRRIHQGLERALL\$

Fig. 43B

2003 CON A2 Env. seq. opt

ATGCGCGTATGGGCACCCAGCGCAACTACCAGCACCTGTGGCGCTGGGGCATCCTGATCCTGGGCATGCTGATCATGTGCAAGGCCACCCGA  
 CCTGTGGGTGACCCGTGTACTACGGGTGCGGAGGACCCGACACCCCTGTTCGGCCCTCGACGCCAAGCCCTACGACACCCG  
 AGGTGCACAACGTTGGCCACCCACCGCTGCCACCCAGGAGTGAACCTGGAGAACCTGGAGAACCTGACCGGACTTCAAC  
 ATGTGGAAGAACAACTGGTGGAGCAGATGCACGAGACATCATCTCCCTGTGGGACCACTCCCTGAAGCCCTGCTGAAGCTGACCCCCCT  
 GTGGGTGACCCCTGAACTGCTCCAAACGCAACCACTCCACTGGAGGAGATCAAGAACTGCTCTACAAACATCACCAACCGAGC  
 TGCCGCAACAGACCCAGAGGTGTACTCCTGTTCTACAAGTGGAGTGGAGTCCAAACAGTCCAAACAGTCCGAGTACTACTACCGC  
 CTGATCAACTGCAACACCTCCGCCATCACCCAGGCTGCCCAAGGTGTCTTTCGAGCCCATCCCATCCACTACTGCGCCCTCCGCGGCTT  
 CGCCATCCTGAAGTGCAGGACCCCGCTTCAACGGCACCGGCTCCTGCAACAACAGTGTCTTCCGTGCACTGCAACCCACGGCATCAAGCCCG  
 TGGCCCTCCACCCAGTGTGTGAACGGTCCCTGGCCGAGGCAAGTGTGATCCGCTCCGAGAACATCACCAACAACCGCAAGAACATC  
 ATCGTGCAGTTCAACAAGCCCGTCCCATCACCTGCATCCGCCCAACAACAACACCCGAAAGTCCATCCGCTTCCGCCCCGGCCAGGCCCTT  
 CTACACCAACGACATCTCCCAACAAGACCATCCGCAAGACATCAACAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAGT  
 AGCTGGCGGAGCATCTCCCAACAAGACCATCCGCAAGACATCAACAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAGTCCAAAGT  
 GCGGAGTTCTTACTGCAACACCCAGCGGTGGCCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCAT  
 CTGCCGATCAAGCAGATCATCAACATGTGGCAGCGGTGGCCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCAT  
 ACATCACCGGATCATCTGACCCGCGACGGCGCAACAAGATCGAGCCCTGGCGTGGCCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCAT  
 TCCGAGTGTACAAGTACAAGTGTGAAGATCGAGCCCTGGCGTGGCCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCAT  
 CGCCGTGGGCATGGCGCGCTTCCGCGTCCGCGTGGCGTGGCCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCATGACGCGCGCAT  
 GCCAGTGTGTCGGCATCGTGCAGCAGTCAACCTGCTGAAGCCATCGAGGCCAGCAGCCTGCTGGAACATCTGGGCTGCTCCGGCAAGCTGAT  
 ATCAAGCAGTGCAGGCCCGGTGCTGGCCCTGGAGCGTACCTGCAGGACCCAGGAGATCTGGAACAACAATGACCTGGCTGCTCCGGCAAGCTGAT  
 CTGCGCCACCACTGGAACCTCCTCCTGGTCCAAACAGACCCAGGAGATCTGGAACAACAATGACCTGGCTGCTCCGGCAAGCTGAT  
 AGATCTCAAACFACACCAACATCATACAAGTGTGGAGAGTCCAGAACCCAGAGGAGAACAGAGCAGGACCTGTGGCTGGCCCTGGAC  
 AAGTGGCCAACTGTGGAACCTGTTCAACATCACCACTGGCTGCTCCCGCTGCTCCCGCTGCTCCCGCTGCTCCCGCTGCTCCCGCT  
 GCGCATCGTATCGCCATCATCTCCGTGGTGAACCGCGGCGAGCCGCGCATCCATCCCGCTGCTCCCGCTGCTCCCGCTGCTCCCGCT  
 GCGTGGACCCCGCGCGCATCGAGAGGCGCGGCGAGCCGCGCATCCATCCCGCTGCTCCCGCTGCTCCCGCTGCTCCCGCTGCTCCCGCT  
 TGGACGACCTGCGCTCCCTGTGCTTCTTCCACCGCGTGGCGACTGCATCCCTGATCGCCCGCGCATCCCTGCTGCTGCTGCTGCTGCT  
 CTCTCCCTGAAGGCGCTGCGCCCTGGGAGGCGCTGAAGTACCTGTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  
 CCATCTCCCTGTGACACCATCGCCCGTGGCCGAGTGGACCCGCGCATCGAGATCGGCGCGCATCGGCGCGCATCGGCGCGCATCGGCGCGCAT  
 AACATCCCCCGCGCATCCGCGCGCATCGAGGCGCGCTTCCGAGGCGCGCTTCCGAGGCGCGCTTCCGAGGCGCGCTTCCGAGGCGCGCTT

Fig. 44B

2003 CON B Env. seq. opt

ATGGCGGTGAAGGGCATCCCGCAAGAACTACCAGCACCTGTGGCGCTGGGGCACCATGCTGGGCAATGCTGATGATCTGCTCCGCCGCCGA  
 GAAGCTGTGGGTGACCCGTGTACTACGGCGTGCCTGTGAAAGGAGCCACACCCCTGTCTGGCCCTCCGACGCCAAGGCTTACGACA  
 CCGAGGTGCACAACAGTGTGGGCCACCCACGCTGCGTGCACCCAGGAGGTGCTGGAGAACGTGACCGAGAACTTC  
 AACATGTGGAAGAACAAACATGTTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCC  
 CCTGTGCGTGAACCTGAACCTGACCCGACCTGATGAACGCCACCAACACCAACCATCATCTACCGCTGGCGGGCGGAGATCAAGAAT  
 GCTCCTTCAACATCACACCTCCATCCGCGACAAGGTGCAGAAGGATACGCCCTGTCTACAAGTGGACGTGGTGGTCCCATCGACAAACGAC  
 AACACCTCCTACCGCCTGATCCTCTGCAACACCTCCGTGATCACCCAGGCTGCCCAAGGTGCTCTCGAGCCCATCCCCATCCACTACTG  
 CGCCCCGCGGCTTCGCCATCCTGAAGTGCACGACAAAGTTCACGGCACCGGCCCTGCACCAACGTGTCCAACCGTGCAGTGCACCC  
 ACGGCATCCGCCCGTGTCCACCCAGCTGCTGTAACGGCTCCCTGGCCGAGGAGGTGGTGTCCGCTCCGAGAACTTCAACCGAC  
 AAGCCAAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACATCGG  
 CCCCCGCGCCTTCTACACCAACCGCGAGATCATCGCGACATCCGCCAGGCCACTGCAACATCTCCCGGCCAAGTGGAAACAACACCC  
 TGAAGCAGATCGTGAAGAGCTGCGGAGCAGTTCGGCAACAGACCATCGTGTCAACCCAGTCCCTCCGGCGGACCCCGAGATCGTGATG  
 CACTCCTTCAACTGCGCGGAGTCTTCTACTGCAACACCCAGCTGTCAACTCCACTGGAAACGGCACCTGGAACAACACCCGAGGG  
 CAACATCACCTGCCCCGCGATCAAGCAGATCATCAACATGTGGCAGGAGTGGCAAGGCCATGTACGCCCCCAAGTGGAAACAACACCCGAG  
 TCCGCTGCTCCCAACATCACCGCTGTGTAAGTACAAGTACAAGTGAAGTGAAGTCCGAGCCCTGGGCGTGGCCCAAGGCCAAGCCCGTGGT  
 CGGACAACCTGCGCTCCGAGCTGTACAAGTACAAGTGAAGTGAAGTCCGAGCCCTGGGCGTGGCCCAACCAAGGCCAAGCCCGTGGT  
 GCAGCGGAGAAAGCGCCGTGGCATCGGCCATCGGCCATCGTCCGCGTTCCTGGGCGCCCGGCTCCACCATGGGCGCCGCTCCATGACCC  
 TGACCGTGCAGGCCCGCAGCTGCTGCCGATCGTGCAGCAGCAACAACCTGTGCGGCCATCGAGGCCAGCAGCTGTGGGCACTTGGGGCTG  
 CTGACCGTGTGGGATCAAGCAGCTGCAGGCCCGTGTGGCCGTGGAGCCTACCTGAAGAACAGCAGTGTGGGCAACATGACCTGGA  
 CTCGGCAAGCTGATCTGCACCAACCGCCGTGCCCTGGAACGCCCTCTGGTCCACAAGTCCCTGGACGAGATCTGGGACAACATGACCTGGA  
 TGGAGTGGAGCGCGAGATCGACAACCTACACCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAAGAACGAGCAGGAG  
 CTGCTGGAGCTGGACAAGTGGCCCTCCCTGTGGAACTGGTTCGACATCACCACTGGCTGTGGTACATCAAGATCTTCAATCATGATCGTGGG  
 CGCCCTGGTGGCCCTCGCATCGTGTTCGCCGTGTCCATCGTGAACCGCGTGCAGGCTACTCCCTGTCTTCCAGACCCCGC  
 TGCCCCCCCCCGGCCCGGACCGCCGAGGGCATCGAGGAGGAGGCGGCGGACCGGCTCCGGCCGCTGGTGGACGGC  
 TTCCCTGGCCCTGATCTGGACGACCTGCGCTCCCTGTGCTTCTCTACCCCTGCGGACCTGTGTGATCGTGCACCCGATCGT  
 GGAGCTGCTGGCCCGCGGCTGGAGGTGCTGAAGTACTGTTGGAACCTGCTGAGTACTGTTCCAGGAGCTGAAGAATCCCGCGTGT  
 CCTGCTGAACGCCACCGCCATCGCCGTGGCCGAGGCCACCGGCTGATCGAGGTGGTGCAGCGGCCCTGCCCGGCCATCTCTGCACATC  
 CCCCCCGCATCCGCCAGGGCTTGGAGCGGCCCTGCTGTAA



Fig. 45A

2003 B. anc Env

MRVKGIRKNCQHLLWRWGTMLLGLMLMICSAAENLWVTYYGVVWKEAATTLFCASDAKAYETEHNWVWATHACVPTDPNPQEVVLENTENF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDILLNATNTSTNMWRGEIKNCSENIITSIRDKMQKEYALFYKLDVVPIDNN  
 TSYRLINCNTSVITQACPVSFEPIPIHYCTPAGEAIIKCNCKFKNGTGPCKNVSTVQCTHGIRPVVSTQLLNGSLAAEEVVIRSENFTDN  
 AKTIIVQLNESVEINCTRPNNTRKSIHIGPGRAFYATGEIIGDIRQAHCNLSRAKWNNTLKQVVTKLREQFDNKTIVFNPSGGDPEIVMH  
 SFCGGEFFYCNTTQLEFNSTWNGTWNTEGNIITLPCRKQIINMWQEVGKAMYAPPPIRGQIRCSNITGLLLTRDGGNNETEIFRPGGDMR  
 DNWRSELYKYKVKVIEPLGVAPTAKRRVVOREKRAVIGAMFLGFLGAAGSTMGAASMTITVQARQLLSGIVQQNNLLRAIEAQHLLQL  
 TVWGIKQLQARVLAVERYLRDQQLLGIWCSGKLICTTVPWNASWSNKSLSDEIWNMTWMEWEREIDNYTGLIYTLIEESQOQEKNEQEL  
 LELDKWASLWNWFDTINWLWYIKIFIMIVGGLVGLRIVEAVLSIVNRVRQGYSPLSFQTRLPA PRGPDRPEGIEEGGERDRDRSRLVNGF  
 LALIWDLLRSLCLFSYHRLRDLLELIVARIVELLGRRGWEALKYWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQRACRAILHIP  
 RRIRQGLERALLS

Fig. 46A

2003 CON C Env

MRVRGILRNCCQWIIWGIIGFWMLMCNVVGNLWVTYYGVVWKEAKTTLFCASDAKAYEKEVHNWVWATHACVPTDPNPQEVVLENTENF  
 NMWKNMVDOMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNTMGEIKNCSENIITELRDKKQKVYALFYRLDIVPLNENNSYRLINC  
 NTSAITQACPVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGKIPVSTQLLNGSLAAEEIIRSENLTNNAKTIIIVHL  
 NESVEIVCTRPNNTRKSIHIGPGRQTFYATGDIIGDIRQAHCNISEDKWNKTILQKVSKLKEHFPNKTIKFEPPSSGGDLEITTHSFNCRGEF  
 FYCNTSKLFNSTYNSTNSTITLPCRKQIINMWQEVGRAMYAPPAGNIITCKSNITGLLLTRDGGKNNTEIFRPGGDMRDNWRSELYKYKV  
 VEIKPLGLAPTAKARRVVEREKRAVIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQNNLLRAIEAQHMLQLTVWGIKQLQTRV  
 LAIERYLKDQQLLGIWCSGKLICTTAVPWNSSWSNKSQEDIWDMNTWQWDREISNYTDTIYRLLEDSQOQEKNEKDLALLDSWKNLWNW  
 FDI TNWLWYIKIFIMIVGGLIGLRIFAVLSIVNRVRQGYSPLSFQTLTPNPRGPDRLGRIFEEEGGEQDRDRSIRLVSGFLALAWDDLRLSLC  
 LFSYHRLRDFILIAARAVELLGRSSLRGLQRGWEALKYLGSLVQYWGLELKKSAISLLDITIAIAVAEGTDRILELIQICRAIRNIPRRIRQ  
 GFEEALQ\$

77/178

## Fig. 45B

## 2003 B.anc Env. seq. opt

ATGGCGGTGAAGGCGCATCCGCAAGAACTGCCAGCACCTGTGGCGCTGGGGCACCACTGCTGGGCATGCTGATGATGATCTGCTCCGCCGCCGA  
 GAACCTGTGGGTGACCGGTGTACTACGGCGTGCCTGGAAGGAGGCCACACCACTTGTCTGGCTCCGACCGCAAGCCCTACGAGA  
 CCGAGGTGCACAACGTGTGGGCCACCCACCGCTGCGTCCCAACCCCAAGGAGTGTGTGGAGAACGTGACCGAGAACTTC  
 AACATGTGGAAGAACAACATGTTGGAGCAGATGCACGAGACATCATCTCCCTGTGGACCAAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCCCTGAACCTGCACCGACTGTGAAGCCCAACCACTCCACCAACATGTACCGCTGGCGGGCGGAGATCAAGAACT  
 GCTCCTTCAACATCACCACTCCATCCGGGACAAGATGCAGAGGATCGCCCTGTCTACAGCTGGACGTGGTGCCTCATCGACAACAAC  
 ACCTCCTACCGCTGATCAACTGCAACACTCCGTGATCACCCAGGCTGCCCAAGGTCTCTCGAGCCCATCCCATCCACTACTGCAC  
 CCCCCGGCTTCGCCATCTGAAGTGCAACGACAAGAAGTCAACGGCACCCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACG  
 GCATCCGCCCCGTGGTGTCCACCCAGCTGCTGAACGGTCCCTGGCCGAGGAGGTGGTGTCCGCTCCGAGAACTTCAACCGACAAC  
 GCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACCCGCAAGTCCATCCACATCGGCC  
 CGCCCGGCTTCTACGCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCACTGCAACCTGTCCCGCGCAAGTGGAAACAACACCCCTGA  
 AGCAGGTGGTGAACCAAGCTGCGGAGCAAGTTCGACAACAAGACCATCGTGTTCAAACCTCCTCCGGCGGACCCCGAGATCGTGATGCAC  
 TCCTTCAACTGCGGCGGAGTTCCTTACTGTCAACACCAACCCAGCTGTTCAACTCCACCTGGAACGGCACCTGGAACAACAACCCGAGGCAA  
 CATACCCCTGCTGCCGATCAAGCAGATCATCAACATGTGCGAGGAGTGGCAAGGCCATGTACGCCCCCTCCCATCCGCGGCCAGATCC  
 GCTGCTCTTCAACATCACCGCTGCTGTGACCCGACGGCGCAACAACGAGACCGAGATCTTCCGCCCCCAAGCCCAAGCCGCGGACATGCCG  
 GACAACTGGCGCTCCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCTGGCGGTGGCCCAACCAAGCCCAAGCCGCGGCGGCGGCGGCGG  
 GCGGAGAGCGCGCGTGGCATCGGCGCATGTTCTGGGCTTCTGGCGCCCGGCTCCACCATGGCGCCCAAGCCCTCCATGACCCCTGA  
 CCGTGCAAGCCCGCAGCTGTCCGGCATCGTGACGAGCAACAACCTGCTGCGGCCATCGAGGCCCAAGCAGCCTGTCGAGCTGCGAGCTG  
 ACCGTGTGGGCGATCAAGCAGCTGCAGGCCCGCTGCTGGCCGTGGAGCGCTACCTGCGCGACCAAGCAGCTGTTGGCAACAACATGACCTGGATGG  
 CGCAAGCTGATCTGCACCAACCTGCTGGTCCAAACAAGTCCCTGGACGAGATCTGGAAACAACAACATGACCTGGATGG  
 AGTGGGAGCGGAGATCGACAACCTGATCACCCCTGATCGAGGAGTCCAGAAACCAAGCAGGAGAGAAACGAGCAGGAGCTG  
 CTGGAGCTGGACAAGTGGCCCTCCCTGTGGAACCTGGTTCGACATCACCAACTGGCTGTGGTGTGATCAAGATCTTCATCATGATCGTGGGCGG  
 CCTGGTGGCCCTGCGCATCGTGTGCGCGTGTCCATCGTGAACCGCTGCGCCAGGGCTACTCCCCCTGTCCCTCCAGACCCCGCTGC  
 CCGCCCCCGGCCCGACCCCGAGGGCATCGAGAGGAGGGCGGAGCGGACCCGACCCGCTCCGGCCGCTGGTGAACCGCTTC  
 CTGGCCCTGATCTGGGACGACCTGCGTCCCTGTGCTTCTCCTAACCCGCTGCGGACCTGTGCTGATCGTGGCCCGCATCGTGGGA  
 GCTGCTGGCCCGCGGCTGGAGGCCCTGAAGTACTGGTGAACCTGCTGCAGTACTGGTCCCAGGAGCTGAAGAATCCCGCCGTGTCCC  
 TGTGAACGCCACCGCATCGCCGTGGCCGAGGCCACCGCCGCTGATCGAGGTGGTGCAGCGGCCCTGCCCGGCCATCTCTGCACATCCCC  
 CGCCGCATCCGCCAGGGCTGGAGCGGCCCTGCTGTAA

Fig. 46B

2003 CON C Env. seq. opt

ATGGCGTGGCGGATCCCTGCCAACTGCCAGCAGTGGTGGATCTGGGGCATCCCTGGGCTTCTGGATGCTGATGATCTGCAACCGTGGTGGG  
 CAACCTGTGGTGACCGGTGTAACGGCGTACCGGCTGTGGAAGGAGGCCAAGACCACCCCTGTCTGGCCCTCCGACGCCAAGGCCTACGAGA  
 AGAGGTGCACAACCGTGTGGCCACCCACCGCTGCGTGCACCCCAACCCAGGAGATCGTGTGGAGAACGTGACCGGAACTTC  
 AACATGTGGAAGAACGACATGGTGGACACAGATGACGAGGACATCATCTCCCTGTGGACCCAGTCCCTGAAGCCCTGCCGTGAAAGCTGACCCC  
 CCTGTGCCGTGACCCCTGAACTGCCACCAAGCCCAACGACCAACCATGGCGGAGATCAAGAACTGCTCTCTTCAAACATCAACCCAGC  
 TGGCGGACAAGAAGCAGAAGGTGTACGCCCTGTCTACCGCTGGACATCGTGCCTTGAACGAGAACAACCTCTACCGCTGATCAACTGC  
 AACACTCCGCCATCACCCAGGCTGCCCAAGGTGCTCTTGACCCCATCCCATCTACTGCGCCCCCGCGGTACGCCATCTCTGAA  
 GTGCAACAAGACCTTCAACGGCACCGGCCCTGCAACAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGTCTCCACCC  
 AGTGTCTGTAACGGCTCCCTGGCCGAGGAGATCATCTCGCTCCGAGAACCTGACCAACAAGCCAAAGACCATCATCTGTGCACTG  
 AACGAGTCCGTGGAGATCGTGTGCACCCCAACAACACCCGCAAGTCCATCCGCATCGGCCCGGCGAGACCTTCTAGCCACCGG  
 CGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCGAGGACAAGTGGAAACAAGACCTGCAAGAAGTGTCCAAGAAGCTGAAGG  
 AGCACTTCCCAACAAGACCATCAAGTTCGAGCCCTCTCCGGCGGACCTGGAGATCACCACTCCCTTCAACTGCCGCGGAGTTT  
 TTCTACTGCAACACTCCAAGCTGTTCAACTCAACTCAACTCCCAACTCCCAACTCCCAACTCCCAACTCCCAACTCCCAACTCCCAACTCCCA  
 CATGTGGCAGGAGTGGCCGGCCCATGTACGCCCCCATCGCGGCAACATCACCTGCAAGTCCAACATCACCCGGCTGCTGTGACCC  
 GCGACGGCGCAAGAACAACACCGAGACCTTCCGCCCGCGGCGGACATGCGCGACAACCTGGCGTCCGAGCTGTACAAGTACAAGGTG  
 GTGGAGATCAAGCCCTGGGCATCGCCCAAGGCCAAGCCCGCTGGTGGAGCGGAGAACCGCCGCGTGGGCATCGGCCCGTGT  
 CCTGGCTTCTGGCGCGCGCTCCACCATGGGCGCGCTCCATCACCCTGACCGTGCAGGCCCGCCAGCTGTCCGGCATCGTG  
 AGCAGCAGTCCAACTGCTGCGCGCCATCGAGGCCCAGCAGCATGCTGCAGTGCAGTGCAGTGGGCGATCAAGCAGTGCAGACCCCGGTG  
 CTGCCATCGAGCGCTACCTGAAGGACCGAGCTGCTGGCATCTGGGGTGTCTCCGGCAAGTGTGACCCACCCTGCTGCCCTGGAA  
 CTCCTCTGGTCCAAAGTCCAGGAGACATCTGGACAACATGACCTGGATGAGTGGACCGCGAGATCTCCAACACCGACACCA  
 TCTACCGCTGCTGGAGACTCCAGAACCCAGGAGACTGAGAGGACTGCTGGCCCTGGACTCCGACTCCGCAATCATCTCCGCGTGTCT  
 TCGACATCACCAACTGCTGTGGTACATCAAGATCTTCATCATGATCGTGGCGGCTGATCGGCCTGCGCATCATCTCCGCGTGTCT  
 CATCGTGAACCGGTGCCAGGGTACTCCCGCTGTCTCCAGACCTGACCCCAACCCCGCGGCTGATCGGCCTGCGCATCATCTCCGCGTGTCT  
 AGGAGGCGCGGAGCAGGACCGGACCGTCCATCCGCTGGTGTCCGGCTTCTGGCCCTGGCCCGGACTGCGCTCCCTGTCCCTGTGC  
 CTGTTCTCCACCGCTGCGGACTTCACTCCCTGATCGCCCGCGCTGGAGTGTGGCCCGCTCCCTCCCTGCGCGGCTGCAGCG  
 CGGCTGGAGGCCCTGAAGTACTTGGCTCCCTGGTGCAGTACTGGGCTTCCGAGTGAAGAAGTCCGCAATCTCCCTGCTGACACCATCG  
 CCATCGCGGTGGCCGAGGCCACCGCATCATCGAGTGCATCCAGCGGCTTGTCCCGGCGGCTTCCGCAACATCCCCCGCGCATCCGCGCAG  
 GGCTTCGAGGCCCGCCCTGCAGTAA

Fig. 47A

2003 C. anc Env  
 MRVMGILLRNCQQWIIWGIILGFWMMLICNVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEREVHNWATHACVPTDPNPQEMVLENTENF  
 NMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNTMGEMKNCSENIITELRDKKQKVYALFYRLDIVPLNDNNSYRLINC  
 NTSAITQACPVSFDPPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGKPVVSTQLLNGSLAEEEEIIIRSENLTDNAKTIIVHL  
 NESVEIVCTRPNNTRKSIIRIGPGQTFYATGDIIGDIRQAHNCNISEEKWNKTQORVGEKLEHFPNKTIKFAPSSGGDLLEITTHSFNCRGEF  
 FYCNTSRLEFNSTYNSKNSTITLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLTRDGGKNTTETFRPGGDMRDNRWRSSELYKYKV  
 VEIKPLGLAPTEAKRRVVEREKRAVIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHMLQLTVWGIKQLQTRV  
 LAIERYLKDQQLLGIWCSGKLLICTTAVPWNSSWSNKSEIWDNMTWQWDREISNYTDTIYRLEDSQKQEKNEQDLLALDSWENLWNV  
 FDITNWLWYIKIFIMIVGGLJGLRIIFAVLSIVNRVROGYSPLSFQTLTPNPRGPDRLGRIEEEEGEGQDRDRSIRLVSGFLALAWDDLRSLC  
 LFSYHRLRDFILLAAARAVELLGRSSLRGLQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQICRAIRNIPRRIRQ  
 GFEAALL\$

Fig. 48A

2003 CON D Env  
 MRVRGIQRNYQHLLWRWGIMLLGMLMICSVAENLWVTVYYGVPVWKEATTLFCASDAKASYKTEAHNIWATHACVPTDPNPQEIENVTENF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVKRNNTSNDTNEGEMKNCSENIITTEIRDKKQVHALFYKLDVVPIDDNNSNT  
 SYRLINCNTSAITQACPVTFEPIPIHYCAPAGFAILKCKDKKFKNGTGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEEEIIIRSENLTNNA  
 KIIIVQLNESVTINCTRPYNNTRQTPIGPGQALYTRIKGDIRQAHNCNISRAEWNKTLOQVAKKLGDLNKTIIIFKPSGGDPEITTHSF  
 NCGGEFFYCNTSRLEFNSTWNTKWNSTGKITLPCRIKQIINMWQGVGKAMYAPPIEGLIKCSSNITGLLTRDGGANNSHNETFRPGGDMR  
 DNWRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAIIGLGFELGAGSTMGAASMTLTVQARQLLSGIVQQSNLLRAIEAQHLLIQL  
 TVWGIKQLQARILAVERYLKDQQLLGIWCSGKHCITTVPWNSSWSNKSLDEIWNMTWMEWEREIDNYTGLIYSLIEESONQOQEKNEQEL  
 IELDKWASLWNVFSITQWLWYIKIFIMIVGGLJGLRIVFAVLSLVNRVROGYSPLSFQTLTPNPRGPDRLGRIEEEEGEGQDRGRSIRLVNGF  
 SALIWDLRLNLCLEFSYHRLRDLILLIAARIVELLGRRGWEALKYLWNLQYWIQELKNSAISLFDTTAIAVAEGTDRIEIVQACRAILNIP  
 TRIRQGLERALL\$

80/178

## Fig. 47B

## 2003 C. anc Env. seq. opt

ATGCCGTTGATGGGCATCCTGGGCACTGGGGCATCTGGGCTTCTGGATGCTGATGATCTGCAACAGTGGTGGG  
 CAACCTGTGGGTGACCGTGTACTACGGCGTGCCTGTGAAGGAGGCCAAGACCACCTGTCTGCGCCTCCGACGCCAAGCCTACGAGC  
 GCGAGGTGCACAACAGTGTGGCCACCCACCGCTGCGTGCACCGACCCCAACCCAGGAGATGGTGTGGAGAACGTGACCGGAACTTC  
 AACATGTGAAGAACGACATGGTGGACCAAGATGACCGAGGACATCATCTCCCTGTGGGACAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCCGTGACCCCTGAACTGCACCAACGCCACCAACACCATGGGCGAGATGAAGAATGCTCCTTCAACATCACCAACCGAGC  
 TGGCGGACAAGAAGCAGAAGTGTACGCCCTGTGACATCGTGCCTTGAACGACAACAACCTCTACCGCCTGATCAACTGC  
 AACACCTCCGCCATCACCCAGGCTGCCCAAAGTGTCTTCGACCCCATCCCATCCACTACTGCGCCCGCCGCTACGCCATCCTGAA  
 GTGCAACAACAAGACCTTCAACGGCACCGGCCCTGCAACAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGTCCACCC  
 AGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGATCATCCGCTCCGAGAACCTGACCGACAACGCCAAGACCATCATCGTGCACCTG  
 AACGAGTCCGTGGAGATCGTGTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCGATCGGCCCGCCGACAGACCTTCTACGCCACCCG  
 CGACATCATCGCGACATCCGCCAGGCCACTGCAACATCTCCGAGGAGAAGTGGAAACAAGACCTGACGCCCGCCGACAGACCTTCTACGCCACCCG  
 AGCACTTCCCAACAAGACCATCAAGTTCGCCCCCTCCTCCGGCGGACCTGGAGATCACCAACCCACTCCTTCAACTGCCCGGCGAGTTC  
 TTCTACTGCAACAACCTCCCGCTGTTCAACTCACCTACAACCTCCAAGAATCCACCATCACCTGCCCTGCCCTGCCGCATCAAGCAGATCATCAA  
 CATGTGGCAGGGCTGGCCGCGCATGTACGCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCGGCTGCTGCTGACCC  
 GCGACGGCGCAAGAACAACACCGAGACCTTCCGCCCGCGGCGGACATGCGGACAACCTGGCGTCCGAGCTTCAAGTACAAGTACAAGGTG  
 GTGGAGATCAAGCCCTGGGCATCGCCCAACCGAGGCCAAGCCCGCTCCATCCCTGACCGTGCAGGAGCGGAGAGCGCCGTGGGCATCGGCCGCTGT  
 CCTGGGCTTCTGGGCGCCGCGCTCCACCATGGGCGCCAGCAGCATGTGCAGCTGACCGTGTGGGCATCAAGCAGCTGCAGACCCCGCTG  
 AGCAGAGTCCAACCTGCTGCGGCCATCGAGGCCAGCAGCATGTGGGCTGCTCCGGCAAGCTGATCTGCACCAACCCGCTGCCCTGGAA  
 CTGGCCATCGAGCGCTACCTGAAGGACCCAGCAGCTGTGGCATCTGGGCTGCTCCGGCAAGCTGATCTGCACCAACCCGCTGCCCTGGAA  
 CTCCTCCTGGTCCAACAAGTCCAGGAGGATCTGGACAACATGACCTGGATGCAGTGGACCCGAGATCTCCAACATACACCGACACCA  
 TCTACCGCCTGCTGGAGGACTCCAGAACCCAGCAGGAGAAGAACGAGCAGGACCTGTGCCCTGGACTCCTGGGAGAACCCTGTGGAACCTGG  
 TTCGACATCAACCACTGGCTGTGTATACATCAAGATCTTCAATGATCGTGGGCGCTGATCGGCCCTGGCATCATCTTCGCCGTGCTGTC  
 CATCGTGAAACCGCTGCGCCAGGGCTACTCCCCCTGTCCCTCCAGACCTTGAACCCCAACCCCGGCCCGCCGACCCGCTGGGCCGATCG  
 AGGAGGAGGGCGGAGCAGGACCGGACCGCTCCATCCGCTGGTCCGGCTTCTGGCCCTGGCCCTGGGACGACCTGCCCTCCCTGTG  
 CTGTTCTCCTACCACCGCTGCGGACTTCACTGATCGCCCGCCCGCTGGAGTGTGGGCGCTCCCTCCCTGCGCGGCTGCAGCG  
 CGGCTGGGAGGCCCTGAAGTACCTGGGCTCCCTGGTGCAGTACTGGGCGCTGGAGCTGAAGAAGTCCGCCATCTCCCTGCTGGACACCATCG  
 CCATCGCCGTGGCCGAGGCCACCGACCCGCATCATCGAGCTGATCCAGCGCATCTGCCCGGCCATCCGCAACATCCCCCGCCGCTCCGCCAG  
 GGCTTCGAGGGCCCGCCCTGCTGTAA

81/178

## Fig. 48B

2003 CON D Env. seq. opt  
 ATGCGGTGGCGGCATCCAGCACTTGTGGGCTTGGGCAATCATGCTGGGCAATGCTGATGATGATGCTCCGTGGCCGA  
 GAACTGTGGGTGACCGTGTACTACGGCTGCCCGTGTGAAGGAGGCCACCACTTGTGCGCTCCGACGCCAAGTCTTACAAGA  
 CCGAGGCCACAACATCTGGCCACCACCGCTGCTGCCACCCAGGAGATCGAGCTGGAGAACGTGACCGGAACTTC  
 AACATGTGAAGAACAAATGTTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCTGAATGCACTGACCCGACCGGAGTGAAGCGCAACAACACTTCAACGACACCAACGAGGCGGAGATGAAGAACTGCTCTTCAACA  
 TCACCACCGAGATCCGCGACAAGAAGCAGGTGCACCCCTGTTTACAAGCTGGACGTGGTCCCATCGACGACAACAATCCAAACACC  
 TCCACCGCTGATCAACTGCAACCTCCGCCATCACCCAGGCTGCCCAAGTGACCTTCGAGCCCATCCCATCCACTACTGCGGCC  
 CGCCGGCTTCGCCATCCTGAAGTCAAGGACAAGAAGTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCA  
 TCCGCCCCGTGGTCCACCCAGTGTCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACCTGACCAACAACGCC  
 AAGATCATCTCGTGCAGTGAACGAGTCCGTGACCATCAACTGCACCCGCCCTACAACAACACCCGCCAGCGCACCCCATCGGCCCCGG  
 CCAGGCCCTGTACACCAACCCGATCAAGGCGGACATCCGCCAGGCCACTGAAACATCTCCCGCGCGGAGTGAACAAGACCCCTGCAGCAGG  
 TGGCCAAGAGCTGGCGGACCTGCTGAACAAGACCACTCATCTTCAAGCCCTCCTCCGGCGGACCCCGAGATCACCCACTCCTTC  
 AACTGGCGGGCGGAGTCTTCTACTGCAACACTCCCGCTGTCAACTCCACCTGGAAACAACCAAGTGAACCTCCACCGGCAAGATCAC  
 CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGGCGGTGGCAAGCCATGTACGCCCCCTCCCATCGAGGGCTGATCAAGTGTCT  
 CCTCCAACATCACCGGCTGCTGACCCCGGACGGCGCCAAACAATCCACAACGAGACTTCCGCCCCCGGCGGCGGACATGCGC  
 GACAACTGGCGCTCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTGGCGTGGCCCCACCCCGCCAAAGCCGCTCCATGACCCCTGA  
 GCGGAGAAAGCCGCTGCGCTGGCCATGTTCTGGGCTTCCCTGGCGCCCGCGGCTCCACCATGGGCGCCGCTCCATGACCCCTGA  
 CCGTGCAGGCCCGCAGTGTGTCCGGCATCGTGCAGCAGAACAACTGCTGGCGCCATCGAGGCCACAGCAGCTGCTGGGCTGTCAGTGC  
 ACCGTGTGGGCGATCAAGCAGTGCAGGCCCGCATCCTGGCCGTGGAGCGCTACTGGAAGTCCCTGGACGAGATCGAAACAACATGACCTGGATGG  
 CGCAAGCACATCTGCACCAACACCGTGCCTGGAACCTCCTGCTGCTCAACAAGTCCCTGGACGAGATCGAAACAACATGACCTGGATGG  
 AGTGGGAGCGCGAGATCGAACACTACACCGCCTGATCTACTCCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAAGAACGAGCAGGAGCTG  
 CTGGAGCTGGACAAGTGGCTCCCTGTGGAACCTGGTTCCTCATCACCCAGTGGTGTGGTACATCAAGATCTTCAATCATGATCGTGGGCGG  
 CCTGATCGGCCCTGGCATCGTGTCCCTGGTGAACCCGCTGGCCAGGCTACTCCCCCTGTCTCCAGACCCCTGGTGC  
 CGCCCCCGGCCCCGACCGCCGAGGGCATCGAGGAGGAGGCGGAGCAGGGCCCGGCTCCATCCGCTGGTGAACGGCTTC  
 TCCGCCCTGATCTGGACGACCTGCGCAACCTGTCCCTTCTTACCACCGCTGCGGACCTGATCTGATCGCCCGCCCGCATCGTGG  
 GCTGCTGGCCCGCGGCTGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACTCCGCCATCTCCC  
 TGTTCGACACCAACCGCCATCGCCGTGGCCGAGGCCACCGACCCCGTGTGATCGAGATCGTGCAGCGGCCCTGCCGCGCCATCTTGAACATCCCC  
 ACCCGCATCCGCCAGGGCTGGAGCGGCCCTGCTGTAA

Fig. 49A

2003 CON F1 Env  
 MRVGMQRN̄WQHLLGKWLFLGILIIICNAADNLWTVVYGVVWKEATTLFCASDAKSYEKEVHNWVWATHACVPTDPNPQEVVLENVTFENF  
 DMWKNMVEQMHTDIIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNDTNDNKTGAIQNCSENMTEVRDKKLVKHALFYKLDIVPISNNNSK  
 YRLINCENTSTITQACPKVSWDPIPIHYCAPAGYAILKCNDRKRENGTGPCKNVSTVQCTHGKIPVSTQLLLNGSLAEEDIIRSONISDNAK  
 TIVHLNESVQINCTRPNNTRKSIHLGPGQAFYATGEIIGDIRKAHCNISGTQWNKTLEQVKAKLKSHFPNKTIKFNSSGGDLEITMHSF  
 NCRGEFFYCNTSGLFNDTGSNGTITLPCRIKQIVNMWQEVGRAMYAAP IAGNITCNSNITGLLLTRDGGQNNTEFRPGGNNKMDNWRSELY  
 KYKVEIEPLGVAPTAKAKRQVVKRRRAVIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQNNLLRAIEAQOHLLOLTVWGIKQL  
 QARVLAVERYLKDQQLLGLWGCSSGKLICTTNVPWNSWSNKSQDEIWNMTWMEWEKEISNYSNIIYRLIEESQOQEKNEQELLALDKWAS  
 LWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRKGYSPLSLQTLIPSPREPDRPEGIEEGGEGQKDRSVRLVNGFLALVWDDL  
 RNLCLEFSYRHLRDFILIAARIVDRGLRRGWEALKYLGNLTQYWSQELKNSAISLLNTTAIVVAEGTDRVIEALQGRAGRAVLNIPRRIRQGLE  
 RALL\$

Fig. 50A

2003 CON F2 Env  
 MRVREMQRN̄WQHLLGKWLFLGILIIICNAADNLWTVVYGVVWKEATTLFCASDAKAYEREVHNWVWATYACVPTDPSQELVLENVTFENF  
 NMWKNMVDQMHTDIIISLWDQSLKPCVKLTPLCVTLNCTDVNVTINTNVTLGEIKNCSENFITTEIKDKKKKEYALFYRLDVVPINNSIVYR  
 LISCNTSTVTOACPKVSFEPIPIHYCAPAGFAILKCNDRKRENGTGLCRNVSTVQCTHGIRPVVSTQLLLNGSLAEEDIIRSENI SDNTKTI  
 IVQFNRSVEINCTRPNNTRKSI RIGPGRAFYATGDIIGDIRKAYCNINRTLWNETLKKVAEEFKNHENITVTFNPSGGDLEITTHSFNCR  
 GEFYCNTSGLFNDTGSNGTITLPCRIKQIVNMWQEVGRAMYAPPIAGQIQCNISNITGLLLTRDGGKNGSETLRPGGMDRDNWRSELYK  
 YKVKIEPLGVAPTAKAKRQVVKRRRAVIGAVLLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLKAI EAQOHLLOLTVWGIKQLQ  
 ARILAVERYLKDQQLLGLWGCSSGKLICTTNVPWNSWSNKSQDEIWDNMTWMEWEKEISNYTDTIYRLIEDAQOQEKNEQELLALDKWDNL  
 WSWFTITNWLWYIKIFIMIVGGLIGLRIVFAVLSVNVNRVROQGYSPLSLQTLIPNPRGPERPGGIEEGGEGQDRDRSIRIVSGFIALAWDDL  
 SLCLFSYRHLRDFILIAARTVDMGLKRGWEALKYLWNLPOYWGQELKNSAISLLDTTAIAVAEGTDRIIEVLQGRAGRAVLNIPRRIRQGFER  
 ALL\$

83/178

## Fig. 49B

**2003 con F1 Env. seq. opt**  
 ATGCCGCTGCCGGCATTGCAGCCGCAACTGGCAGCACCTGGGCAAGTGGGGCCTGCTGTTCCTGGGCATCCTGATCATCTGCAACGCCGCCGA  
 GAACCTGTGGGTGACCGTGTACTACGGGTGCCCGTGTGAAGAGGAGGCCACACACCCTGTGTTCGGCCTCCGACGCCAAGTCTCTACGAGA  
 AGGAGGTGCACAAACGTGTGGGCCACCCACCGCTGCGTGCACCCCAACCCAGGAGGTGGTGTGGAGAACGTGACCCGAGAACTTC  
 GACATGTGGAAGAACAAACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCTGACCCGTGAACCTGCACCGACGTGAACGCCACCAACGACACCAACGACACAAAGACCCGGCCCATCCAGAACTGCTCCTTCA  
 ACAIGACCACCGAGGTGCGGACAAAGAGTGAAGTGCACGCCCTGTCTACAAGTGGACATCGTGCCTATCTCCAACAACAACCTCCAAG  
 TACCGCTGATCAACTGCAACACTCACCATCCCAAGCTGCCCAAGTGTCTGGACCCCATCCCATCCACTACTGCGCCCCCGC  
 CGGTACGCCATCTCTGAAGTGCAACGACAAAGCGCTTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTCCAGAACATCTCCGACAACGCCAAG  
 AGCCCGTGTCCACCCAGCTGAACGAGTCCGTGACAGTCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACTGGGCCCGGCCA  
 ACCATCATCGTGCACTGAACGAGTCCGTGACAGTCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACTGGGCCCGGCCA  
 GGCCTTCTACGCCACCGCGGAGATCATCGGCGACATCCGCAAGGCCCACTGCAACATCTCCGGACCCAGTGGAAACAAGACCCCTGGAGCAGG  
 TGAAGGCCAAGCTGAAGTCCCACCTTCCCAAGACCATCAAGTCAACTCTCCCTCCGGCGGACCTGGAGATCACCATGCACTCCTTC  
 AACTGCCCGGGGAGTCTTCTACTGCAACACCTCCGGCCTGTCAACGACACCGGCTCCAACGGCACCCATCACCTGCCCTGCCGATCAA  
 GCAGATCGTGAACATGTGGCAGGAGGTGGCCCGCCATGTACCGCCCGCCCATCGCCGGCAACATCACCTGCAACTCCAACATCACCGGCC  
 TGCTGTGACCCCGGACGGCCAGAACAAACACCGAGACCTTCGCCCCCGGGCGCAACATGAAGGACAACTGGCGCTCCGAGCTGTAC  
 AAGTACAAGTGTGGAGTCGACCCCTGGCGTGGCCCCCAAGGCCAAGCGCCAGGTGGTGAAGCGCGAGCGCCGCGCGCTGGGCAT  
 CGGCCCGTGTCTCCGTGGCTTCTGGCGCCCGGCTCCACCATGGCGCCCGCTCCATCACCTGACCGTGCAGGCCCGCCAGCTGCTGT  
 CCGGCATCGTGACGACGAAACAACCTGCTGGCGCCATCGAGGCCACAGCACCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCTG  
 CAGGCCCGGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAAGCTGTGGCCCTGTGGGCTGCTCCGGCAAGTGTGACCCACCAA  
 CGTGCCCTGGAACCTCCTCTGTCCAACAGTCCAGGACGAGATCTGGAACAACATGACCTGGATGGAGTGGAGAGGATCTCCAAC  
 ACTCCAACATCATACCCGCTGATCGAGGATCCCAAGACCAAGAGGAGAACGAGCAGGAGTGTGGCCCTGGACAAAGTGGCCCTC  
 CTGTGGAACCTGTTCCACATCTCCAACCTGGTGTGATCAAGATCTTCAATCATGATCGTGGCGGCTGATCGGCCCTGCCCATCGTGT  
 CGCGTGTGTCATCGTGAACCGCTGCAAGGCTACTCCCCCTGTCCCTGCAGACCTGATCCCCCTCCCGGAGCCCGACCCGCC  
 CCGAGGCCATCGAGGAGGGCGGAGCAGGCCAGCCGCTCCGTGGCTGGTGAACGGCTTCTGGCCCTGGTGTGGACGACCTG  
 CGCAACTGTGCTGTCTCCTACCGCCACTGCGGACTTCACTCCTGATCGCCCGCCGATCGTGGACCGCGCTGCGCCGGCTGGGA  
 GGCCCTGAAGTACCTGGGCAACCTGACCCAGTACTGTTCCAGGAGTGAAGAACTCCGCCATCTCCCTGTGTAACAACACCCGCTG  
 TGGCCGAGGCCACCGCGTGTGATCGAGGCCCTGCAGGCCCGGCGGCTGCTGAACATCCCCCGCCGATCCGCCAGGCCCTGGAG  
 CGGCCCTGCTGTAA



Fig. 50B

2003 CON F2 Env. seq. opt

ATGCCGCTGGCGAGATGCAGCGCAACTGGCAGCACCTGGGCAAGTGGGCTGCTGTTCTTGGGCATCCTGATCATCTGCAACGCCGCCGGA  
 CAACCTGTGGGTGACCGGTGACTACGGCGTGCCTGGTGTGAAGGAGGCCAACACCCCTGTTCTGGCCCTCCGACGCCAAGCCCTACGAGC  
 GCGAGTGCACAACCGTGTGGCCACCTACCGCTGCCACCGACCCCTCCCGCCAGGAGCTGGTGTGGCAACCGTGGCCGAGAACTTC  
 AACATGTGAAGAAACAACATGGTGAACCGTGAACCGTGAACCGTGAACCGTGAACCGTGAACCGTGAACCGTGAACCGTGAACCGTGAACCGT  
 CCTGTGCTGACCCCTGAACCTGCACCGACGAAAGAGAGTACGCCCTGTTCTAACCGCTGGACGTGGTCCCATCAACAACTCCATCGTGTACCGC  
 CTGATCTCTGCAACACCTCCACCGTACCCAGGCTGCCCAAGTGTCTTCGAGCCCATCCCATCACTACTGCGCCCGCCCGCGGCTT  
 CGCCATCTGAAGTGCAACGACAAGAGTTCAACGGCACCGGCTGTGCCGCAACGTGTCCACCGTGCAGTGACCCACGGCATCCGCCCG  
 TGGTGTCAACCCAGCTGCTGAACCGCTCCCTGGCCGAGGACATCATCATCCGCTCCGAGAACATCTCCGACAAACCAAGACCATC  
 ATCGTGCAGTTCAACCGTCCGTGGAGATCAACTGCACCCCGCCCAACAACAACCCGCAAGTCCATCCGATCGGCCCGCCCGCGCCCTT  
 CTACGCCACCGGCACATCATCGGCGACATCCGCAAGGCTACTGTCAACATCAACCGCACCTGTGGAACGAGACCCCTGAAGAAGGTGGCCG  
 AGGAGTTCAAGAACCCTTCAACATCACCGTACCTTCAACCCCTCCCTCCGCGGCGACTGGAGATCAACCCACTCTTCAACTGCGCG  
 GCGGAGTTCTTCTACTGCAACACCTCCGACCTGTTCAACAACACCGAGTGAACAACAAGACCATCACCTGCCCTGCCGATCCGCCA  
 GTTCGTGAACATGTGGCAGCGCGTGGCCCGCCATGTACGCCCCCTCCATGCCGCGCAGATCCAGTGAACCTCAACATCAACGCGCTG  
 TGCTGACCCCGCACCGGCAAGAACGGCTCCGAGACCTCGCCCGCGGCGGACATGCGGCAACTGGCGCTCCGAGCTCCGAGCTGTACAAG  
 TACAAGTGGTGAAGATCGAGCCCTGGGCGTGGCCCAACCAAGCCAAAGCCAGGTGGTGCAGCGGAGAAAGCGCCGCTGGCATCGG  
 CGCCGTGCTGCTGGGCTTCTGGGCGCCGCTCCACCAATGGGCGCCCTCCATCACCTGACCTGACCGTGCAGGCTGGGCAAGCTGATCGC  
 GCATCGTGCAGCAGTCCAACTGTGAAGGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGGGCTGCTCCGGCAAGCTGATCGC  
 GCCCTGGAACCTCTCTGGTCCAACAAGTCCAGGACCGCAGAGCTGGGCAACATGACCTGGATGCAGTGGGAGAGGAGATCTCCAAC  
 CCGACACCATCTACCGCTGATCGAGACCGCCAGAACCCAGCAGGAGAAAGAACGAGCAGGACCTGTGGCCCTGGACAAAGTGGACAACTG  
 TGGTCCCTGGTTCACCATCAACCGTGCAGGCTACCTCCCGCTGTCAGACCTGATCCCAACCCCGCCCGCCCGCCGAGCCCGCCG  
 CGTGTGCTGGTGAACCGGTCAGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG  
 GCGGCATCGAGGAGGAGGCGGAGAGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG  
 TCCCTGTGCTGTTCTCTACCGCACTGCGGACTTCATCTGATCGCCGCECCGACCGTGGACATGGCCCTGAAGCGCGGCTGGGAGGC  
 CCTGAAGTACTGTGGAACCTGCCCGACTGGGCGCAGGAGTGAAGAACTCCGCCATCTCCCTGCTGGACACCAACCGCCATCGCCGTGG  
 CCGAGGGCACCGACCGCATCATCGAGGTGCTGCAGCGCGCCGCGGCTGCTGCACATCCCCCGCCCGCATCCCCAGGGCTTCGAGCGC  
 GCCCTGCTGTAA

Fig. 51A

2003 CON G Env

MRVKGIQRNWQHLLWKWGTLLILGLVICSASNNLWVTVYGVVWEDADTTLFCASDAKAYSTERHNWATHACVPTDPNPOEITLENVTENF  
 NMWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVNNTNNTKKEIKNCSENIITEIRDKKKEYALFYRLDVVPIINDNGNSS  
 IYRLINCNVSTIKQACPVTFDPIPIHYCAPAGEAIIKCRDKKFNSTVQCHGIPVSTQLLNGSLAEEIIRSENI TDNT  
 KVIIIVQLNETIEINCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHCVSRKWNEMLQVKAQKKIFNKSIITFNSSSGDLEITTHSE  
 NCRGEFFYCNTSGLFNSSLNSTITLPCKIKQIVRMQRVGOAMYAPPIAGNITCRSNI TGLLTRDGGNNNTETFRPGGDMRDNRWS  
 ELYKYIKVIKPLGVAPTRARRRVEREKRAVGLGAVLLGFLGAAGSTMGAASITLVQVRQLLSGIVQQSNLLRAIEAQQHLLQLITVWGI  
 KQLOARVLAVERYLKDQQLGIWCSGKLICTNVPWNTSWSNKSNEIWDNMTWIEWEREISNYTQQIYSLIEESQOQEKNEQDILLALDK  
 WASLWNWFDITKWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLTHHQREPRPERIEEGGEGQDKDRSIRLVSGFLALAW  
 DDLRSLCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGLYLWNNLLLYWGQELKNSAINLLDTIAIAVANWTDRIEVAQRACRAILN  
 IPRRIRQGLERALL\$

Fig. 52A

2003 CON H Env

TRVMEQORNYPSLWRWGTLLILGMLLCSAAGNLWVTVYGVVWKEAKTTLFCASDAKAYETEKHNVWATHACVPTDPNPQEMVLENVTENF  
 NMWENDMVEQMHTDIISLWDQSLKPCVKLTPLCVTLDCSNVNTNATNSRFNMQEELTNCSENFVTVIRDKQKQVHALFYRLDVVPIDDNNS  
 YQYRLINCNTSVITQACPVSFEPIPIHYCAPAGEAIIKCNKTFENGTPCTNVSTVQCHGIRPVSTQLLNGSLAEEQVIRSKNISDN  
 TKNIIIVQLNKPVEITCTRPNNTRKSIHILGPGQAFYATGDIIGDIRQAHCVSRKWNKTLHQVVTQLGKYFDNRTIIFKPHSGGDMVEVTH  
 SFNCRGEFFYCNTSGLFNSSWTNSTNDTKNIITLPCRIKQIVNMQRVGOAMYAPPIKGNITCVSNI TGLLITFDEGNNTVFRPGGDMRD  
 NWRSELYKYKVVKIEPLGVAPTEARRRVEREKRAVGMGAFFLGLGAAGSTMGAASITLVQARQLLSGIVQQSNLLRAIQAOQHMLQLI  
 VWGIKQLQARVLAVERYLKDQQLGIWCSGKLICTNVPWNTSWSNKSLEIWDNMTWMEWDKQINNYTEEIIYRLLVVSQTQOQEKNEQDLL  
 ALDKWASLWNWFSITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLIPNPRGDRPEGIEEGGEGQDRDRSVRLVNGFL  
 PLVWDDLRSLCLFSYRLLRDLILLIVVRTVELLGRGREALKYIWNLLLYWGQELKNSAINLLNTTIAIAVAEGTDRIIEIVQRAWRAILLHIPR  
 RIRQGFERTLL\$

Fig. 51B

2003 CON G Env. seq. opt

ATGCCGTTGAAGGCATCCAGCGCAACTGGCAGCACCTGTGAAAGTGGGGCACCCCTGATCCTGGGCCTGGTGATCATCTGCTCCGCCCTCCAA  
 CAACCTGTGGGTGACCGTGTACTACGGCGTGCCTGTGGAGGACCCGACACACCCTGTTCCTGGCCCTCCGACGCCAAGGCCACTACTCCA  
 CCGAGGCCACAACGTGTGGCCACCCACGCTGCGCCACCCAGGAGATCACCCCTGGAGAACCTGACCCGAGAACCTTC  
 AACATGTGGAAGAAACAATGGTGGAGCAGATGACGAGACATCATCTCCCTGTGGGACGAGTCCCTGAAAGCCCTGCGTGAAGCTGACCC  
 CCTGTGCGTGACCCCTGAACCTGACCCGACGTGAACGTGACCAACAACAACAACAAGAGATCAAGAATCAAGAACTGCTCTTCAACA  
 TCAACACCGAGATCCGCGACAAGAAGAGTAGCCCTGTTTACCGCTGGACGTTGGTCCCATCAACGACCAACGCAACTCCTCC  
 ATCTACCGCTGATCAACTGCAACGTGTCACCATCAAGCAGGCTGCCCAAGGTGACCTTCGAAGAACCTTCACCCCTGAGTGCACCCACGGCA  
 CGCCGGCTTCGCCATCCTGAAGTCCCGGACAAGAAGTCAACGGCACCGCCCTGCAAGAAGCTGTCCACCCATCCCATCTACTGCGCCCA  
 TCAAGCCCGTGGTGTCCACCAGCTGTGCTGAACGGTCCCTGSCCGAGGAGGATCATCATCCGCTCCGAGAACATCACCGACAACACC  
 AAGGTGATCATCGTGCAGCTGAACGAGACCATCGAGATCAACTGCACCCCGCCACAACAACAACAAGTCCATCCGCATCGGCCCGG  
 CCAGGCTTCTACGCCACCGGACATCATCGGACATCCGCCAGGCCACTGCAACGTGTCCCGCACCAAGTGAACGAGATGCTGCAGA  
 AGGTGAAGGCCAGTGAAGAAGATCTTCAACAAGTCCATCACTYCAACTCCTCCTCCGGGGACCTGGAGATCACCCACTCCTTC  
 AACTGCCCGGGAGTCTTCTACTGCAACACTCCGGCTGTTCAACAACCTCCCTGTGACTCACTCAACAATCCACCATCACCCCTGCCCTG  
 CAAGATCAAGCAGATCGTCCGATGTGGCAGCGGTGGCCAGGCCATGTACGCCCCCTCCATCGCCGGAACATCACCTGCCCTCCAACA  
 TCACCGCTGTCTGACCCCGCACCGGGCAACAACAACACCGAGACCTTCCGCCCGGGGGACATCGCGACAACTGGCGCTCC  
 GAGCTGACAAAGTACAAGATCGTGAAGTCAAGCCCTGGCGTGGCCCAACCGGACCTTCCGCCCGGGGGACATCGCGACAACTGGCGCTCC  
 CGTGGCCCTGGCGCGTGTGCTGGCTTCCCTGGCGCGCGCTCCACCATGGCGCGCTCCATCACCCCTGACCGTGCAGGTGCGCC  
 AGCTGCTGCCGATCGTGCAGCAGTCCAACTGCTGGCGCCATCGAGGCCAGCAGCATCTGTGCAGCTGACCGTGTGGGCATC  
 AAGCAGTGCAGGCCCGTGTGGCGTGGAGCGCTACCTGAAGGACCCAGCAGTGTGGCATCTGGGCATCTGGGCATCGAGTGGAGCGCGAGA  
 CACCACCAACGTGCCCTGGAACAACCTCCTGGTCCAAAGTCTTACAACAGATCTGGGACAACATGACCTGGATCGAGTGGAGCGCGAGA  
 TCTCCAACTACACCCAGCAGATCTACTCCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAAGAACGAGCAGGACCTGTGGCCCTGGACAAG  
 TGGCCCTCCTGTGGAACTGGTTCGACATCACCAAGTGGTGTGGTACATCAAGATCTTCAATCATGATCGTGGCGGCCCTGATCGGCCCTGG  
 CATCGTGTTCGCCGTGTCCATCGTGAACCGGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACCCCTGACCCACCCAGCGGAGC  
 CCGACCGCCCGAGCGCATCGAGGAGGCGGGGAGCAGGACCGCTCCATCCGCCCTGGTGTCCGGCTTCTGGCCCTGGCCCTGG  
 GACGACCTGCCCTCCCTGTGCCTGTTCTCCTACCCCGCTGCGGACTTCACTCTGATCGCCCGCCGACCGTGGAGCTGTGGCCCGCTC  
 CTCCCTGAAGGCCCTGGCCCTGGAGGCGCTGAAGTACCTGTGAACTGTGAACTGTACTGGGCGCAGGAGCTGAAGAATCCGCCCA  
 TCAACCTGCTGGACACCATCGCCATCGCCCTGGCCAACTGGACCGCCGCTGATCGAGGTGGCCAGCGCCCTGCCGCGCATCTTGAAC  
 ATCCCCCGCCGATCCGCCAGGGCTGGAGCGGCCCTGCTGTAA

87/178

## Fig. 52B

**2003 CON H Env. seq. opt**  
 ACCCGGTGATGGAGACCCAGCGCAACTACCCCTCCTGTGGCGTGGGCACCCCTGATCCTGGCATGCTGTGATCTGCTCCGCCGCCGG  
 CAACCTGTGGGTGACCGTGTACGGCGTGCCGTGTGAAGGAGGCCAAGACCACCCCTGTTCTGCGCCTCCGACCGCAAGCCTACGAGA  
 CCGAGAAGCACAACTGTGGCCACCCACCGCTGCGTGCACCCAGACCCCAACCCAGGAGATGTTGCTGGAGAACGTGACCCGAGAACTTC  
 AACATGTGGGAGAACGACATGTTGGAGCAGATGCACACCGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGAACCTGGACTGCTCAAACGTGAACACCAACGCCACCAACTCCCGCTTCAACATGCAGGAGGAGCTGACCAACTGCTCCT  
 TCAACGTGACCAACCGTGTATCCCGGACAAAGCAGAGGAGTGCACGCCCTGTTCTACCGCCTGGACGTGGTGCCTCATCGACGACAACTCC  
 TACCAGTACCGCCTGATCAACTGCAACACTCCGTGATCACCCAGCCTGCCCAAGTGTCTCTCGAGCCCATCCCATCCACTACTGCGC  
 CCCCAGGCTTCGCCATCCTGAAGTGAACAAAGACTTCAACGGCACCGGCCCTGCACCAACGTGTCCACCGTGCAGTGCACCCACG  
 GCATCCGCCCTGCTCAACCCAGCTGCTGTGAACGGCTCCCTGGCCGAGGAGCAGGTGATCATCCGCTCAAGAACATCTCCGACAA  
 ACCAAGAACATCATCGTGCAGCTGAACAAGCCCGTGGAGATCACCTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACCTGGGCC  
 CGCCAGGCTTACGCCACCGGACATCATCGCGACATCGCCAGGCCACTGCAACATCTCCGGCAAGAGTGAACAAGACCCCTGC  
 ACCAGTGTGACCCAGCTGGCAAGTACTTCGACAAACCGCACCATCTTCAAGCCCACTCCGGCGGACATGGAGTGAACCCAC  
 TCCTTCAACTGCCCGGGAGTTCTTCTACTGCAACACTCCGGCTGTCAACTCTCTGGACCAACTCCACCAACGACCAAGAACAT  
 CATCACCTGCCCTGCCGATCAAGCAGATCGTGAACATGTGGCAGCGGTGGCCAGGCCATGTACGCCCCCTCCATCAAGGCAACATCA  
 CCTGCGTCCAAACATCACCGCCTGATCCTGACCTTCGACGAGGCAACAACACCTGACCTTCCGCCCGGCGGACATGCGCGAC  
 AACTGGCGTCCGAGCTGACAAAGTACAAGTGGTGAAGATCGAGCCCTGGCGTGGCCCAACCGAGGCCCGCCCTCCATCACCTGACCG  
 CGAAGCCCGCGTGGCATGGCGCTTCTTCTGGCGTCCAACTGCTGGCGCCATCCAGCCCAAGCAGCAGCATGCTGCAAGTGAACC  
 TGCAGGCCCGCAGTGTCCGGCATCGTGCAGCAGTCCAACTGCTGGCGCCATCCAGCCCAAGCAGCAGTGTGGCATCTGGGGCTGCTCCGG  
 GTGTGGGCATCAAGCAGTGCAGGCCCGCTGCTGGCGTGGAGCGCTACCTGAAGGACCAAGTGTGGACAGATCTGGACAAACATGACCTGGATGGAGT  
 CAAGCTGATCTGCACCAACCTGCCCTGGAACTCCTGCTGCCAACAAGTCCCTGGACGAGATCTGGACAAACATGACCTGGATGGAGT  
 GGGACAAGCAGATCAACAACCTACCCGAGGATCTACCGCCTGCTGGAGTGTCCAGACCCAGCAGGAGAAAGACGAGCAGGACCTGCTG  
 GCCCTGGACAAGTGGCCTCCCTGTGAACTGGTTCCTCCATCAACCACTGGTGTGTGTACATCAAGATCTTCAATCATGATCGTGGCGGCCCT  
 GATCGCCTGCGCATCTTCGCCGTGTCCATCGTGAACCGGTGCCAGGGTACTCCCGCTGTCTCCAGACCCCTGATCCCCA  
 ACCCCGCGCCCGACCGCCGAGGCGATCGAGGAGGCGGCGAGCAGGACCGGACCGCTCCGTGGCCTGGTGAACGGCTTCTG  
 CCCCTGTTGGACGACCTGCGTCCCTGCTTCTCTACCGCCTGCTGCGGACCTGCTGTGATCGTGTGCGCACCGTGGAGCT  
 GCTGGCCCGCGCGGAGGCCCTGAAGTACCTGTGAACTGTCAGTACTGGGCGCAGGAGTGAAGAACTCCGCCATCAACTGC  
 TGAACACCAACCGCATCGCCGTGGCCGAGGCAACCGACCGCATCATCGAGATCGTGCAGCGGCCCTGGCGCCATCTCTGCACATCCCCCGC  
 CGCATCCGCCAGGGCTTCGAGCGCACCCCTGCTGTAA

Fig. 53A

2003 CON 01 AE Env  
 MRVKEQTQMNWPNLWKWGTLLILGLVICSASDNLWVTVYGVVWRDADTTLFCASDAKAHETE VHNWVWATHACVPTDPNPOEIHLENVTFENF  
 NMWKNMVEQMVEDVLSLWDQSLKPCVKLTPLCVTLNCTNANLTVNNTVSNIIIGNITNEVRNCSFNMTTELDRDKKQKVHALFYKLDIVQ  
 IEDNNSYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNCKNFNGTGPCKNVSSVQCTHGKPKPVVSTQLLLNGSLAEFEIIRSEN  
 LTNNAKTIIVHLNKSVEINCTRPSNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEVLKQVTEKLKEHFNNKTIIFQPPSGGDLE  
 ITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIIIPCKIKQIINMWQAGQAMYPPIISGRINCVSNIITGILLTRDGGANNNTNETFR  
 PGGNIKDNWRSELYKYKVQIEPLGAPTRAKRRVVEREKRAVIGAMIFGLGAAGSTMGAASITITVQARQLLSGIVQQSNLLRAIEA  
 QQHLLQLTVWGIKQOARVLAVERYLKDKQFLGLWGCSGKIICTTAVPNWSTWSNRSEFEIWNMTWIEWEREISNYTNIQIYEILTESQNQQ  
 DRNEKDLLLELDKWAASLWNWFEDITNWLWYIKIFIMIVGGLJGLRIIFAVLSIVNRVRQGYSPLSFQTPTHHQREPRPERIEEGGGEQGRDRS  
 VRLVSGFLALAWDDLRLSLCLFSYHRLRDFVLIARTVELLGHSSSLKGLRRGWEGLKYLIGNLLLYWGQELKISAIISLLDATAIAVAGWTD RVI  
 EVAQGAWRALIHIPRRIRQGLERALL\$

Fig. 54A

2003 CON 02 AG Env  
 MRVMGIQKNYP<sup>1</sup>LLWRWGMIIIFWIMIICNAENLWVTVYGVVWRDAETTLFCASDAKAYDTEVHNWVWATHACVPTDPNPOEIHLENVTFENF  
 MWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNANLTVNNTVSNIIIGNITNEVRNCSFNMTTELDRDKKQKVYALFYRLDVVQINKNNSQYR  
 LINCNTSAITQACP<sup>2</sup>VSFEPIPIHYCAPAGFALLKCNCKNFNGTGPCKNVSTVQCTHGKPKPVVSTQLLLNGSLAEFEIIRSENITNNAKTI  
 IVQLVKPVKINCTRPNNTRKSVRIGPGQTFYATGDIIGDIRQAHCVSRTKWNNTLQQVATQLRKYFNKTIIFANPSGGDLEITTHSFNCG  
 GEFFYCNTSELFNSTWNS<sup>3</sup>TWNNTEKCI TLQCR<sup>4</sup>IKQIVNMWQKVQAMYPPIQGVIRCESNITGLLLTRDGGNNNSTNETFRP<sup>5</sup>GGGDMRDNW  
 RSELYKYKVVKIEPLGVA<sup>6</sup>PTRAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITITVQARQLLSGIVQQSNLLRAIEAQHLLKLTWV  
 GIKQOARVLALERYLKDQQLLGIWGC<sup>7</sup>SGKLICTTTPWNSWSNKT<sup>8</sup>YNDIWDNMTWLQWDKEISNYTDIIYNLIEESQ<sup>9</sup>QKNEQD<sup>10</sup>LLAL  
 DKWASLWNWFEDITNWLWYIKIFIMIVGGLJGLRIVFAVLTIIINRVRQGYSPLSFQTLTHHQREPRPERIEEGGGEQDRDRSVRLVSGFLAL  
 AWDDLRLSLCLFSYHRLRDFVLIARTVELLGHSSSLKGLRLGWEALKYLIGNLLSYWGQELKNSAINLLDTIAIAVANWTD RVI  
 INIPRRIRQGLERALL\$

Fig. 53B

2003 CON 01 AE Env. seq. opt  
 ATGGCGGTAAAGGAGACCCAGATGAATGGCCCAACCTGTGGAAGTGGGGCACCCCTGATCCTGGGCCCTGGTGTATCATCTGCTCCGCCCTCCGA  
 CAACCTGTGGGTGACCCGTGTACTACGGCGTGTGGCCGACCGCCGACACCCCTGTCTGGCCCTCCGACCGCAAGGCCACGAGAGA  
 CCGAGGTGCACAACAGTGTGGCCACCCACGCTGCGTCCCAACCCAGGAGATCCACCTGGAGAACCTGGAGAACCTGACCGAGAACTTC  
 AACATGTGGAAGAACAACATGTTGGAGCAGATGTCAGAGGACGTGATCTCCCTGTGGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCC  
 CCTGTGCGTGAACCTGAACTGCAACACCGCCAACTGACCAACCTGAACAACATCAACAACCTGTCACACATCAATCGGCAACATCAACAACG  
 AGTGGCCAACTGCTCTCAACATGACCCAGCTGCGGCAAGAAGCAAGAGTGCACGCCCTGTTCTACAAGCTGGACATCGTGCAG  
 ATCAGGACAACAACCTCTACCGCTGATCAACTGCAACACTCCGCTGATCAAGCAGGCTGCCCAAGATCTCCTTGCACCCCATCCCAT  
 CCACTACTGCAACCCCGGCTACGCCATCCTGAAGTGCAACGACAAGAACTTCAACGGCACCGGCCCTGCAAGAACGTTCTCTCCGTGC  
 AGTGCAACCCACGGCATCAAGCCCGTGTGTCCACCCAGTGTGTAACGGCTCCCTGGCCGAGGAGATCATCATCCGCTCCGAGAAC  
 CTGACCAACAACGCCAAGACCATCATCGTGCACCTGAACAAGTCCGTGGAGATCAACTGCACCCGCCCTCAACAACAACCCGACCTCCAT  
 CACCATCGGCCCGGCAAGTGTCTACCGCACCGGACATCATCGGCGACATCCGCAAGGCTACTGCGAGATCAACGAGATCAACGGCACCAAGTGA  
 ACGAGGTGCTGAAGCAGGTGACCGAGAAGCTGAAGGAGCACTTCAACAACAAGACCATCATCTCCAGCCCCCTCCGGCGGACCTGGAG  
 ATCACCATGCACCACTTCAACTGCCGCGGAGTCTTCTACTGCAACACCAAGCTGTCAACAACACCTGCATCGGCAACGAGACCAT  
 GGAGGGTGCAAACGGCACCATCATCTGCCCTGCAAGATCAAGCAGATCATCAACATGTGGCAGGGCCGCGGCAACAACACCAACGAGACCTTCCGC  
 CCATCTCCGGCGCATCAACTGCGTGTCCAAACATCACCGGATCCTGCTGACCCCGACGGCGGCAACAACAACCAACGAGACCTTCCGC  
 CCGGGCGGCAACATCAAGGACAACTGGCGTCCGAGCTGTACAAGTACAAGTGGTGCAGATCGAGCCCCCTGGGATCGCCCCACCCG  
 CGCCAAAGCCCGGTGGTGGAGCGGAGAGCGCGCTGGGCATCGGCGCATGATCTTCGGCTTCTGGGCGCCGCGCTCCACCATGG  
 GCGCCGCTCCATCACCTGACCGTGCAGGCCCGCAGTGTCTCCGGCATCGTGCAGCAGCAGTCCAACTGCTGGCGCCATCGAGGCC  
 CAGCAGCACCTGCTGCAGTGAACCTGAGCAGCTCAGGCCCCGCTGGAACTCCACCTGGTCCAACTGAGCGCTACCTGAAGGACCAAGAT  
 CCTGGCCCTGTGGGCTGCTCCGGCAAGATCATCTGCACCAACCGCTGCCCTGGAACTCCACCTGGTCCAACTGAGTCCCTTCGAGGAGATCT  
 GGAACAACATGACCTGGATCGAGTGGAGCGCGAGATCTCCAACTACACCAACAGATCTACGAGATCTGACCCGAGTCCAGAACCCAGAT  
 GACCGCAACGAGAAGGACCTGCTGGAGCTGGACAAGTGGCCCTCCCTGTGGAACCTGTTGACATCAACAACCTGGTGGTACATCAAGAT  
 CTTTCATCATGATCGTGGCGGCTGATCGGCTGGCATCATCTTCGCCGTGCTGTCCATCGTGAACCGGCGGCGGAGCAGGCCGACCGCTCC  
 TGTCTTCCAGACCCCAACAGCGGAGCCCGACCGCCGAGCGCATCGAGGAGGGCGGCGGAGCAGGCCGCGGACCGCTCC  
 GTGCGCTGGTGTCCGGCTTCTGGCCCTGGACGACCTGCGTCCCTGTGCTGTTCTCCTAACCCGCTGCGGACTTCATCCT  
 GATCGCCCGCACCGTGGAGCTGCTGGCCACTCCTCCCTGAAGGCTCGCCCGGCTGGAGGGCTGAAGTACCTGGCAACCTGC  
 TGTGTACTGGGGCCAGGAGTGAAGATCTCCGCCATCTCCCTGCTGGAACCCCAACCGCATCGCCGTGGCCGGTGGACCGGACCGCGTGTATC  
 GAGGTGGCCCCAGGGGCTTGGCGGCTTGCACATCCCCCGGCGCATCCGCCAGGGCTGGAGCGGCCCCCTGTGTAA

Fig. 54B

2003 CON 02 AG Env. seq. opt  
 ATGGCGTGATGGGCATCCAGAAGAACTACCCCTGTGGCGCTGGGGCATGATCATTTCTGGATCATGATCATCTGCAACGCCCGAGAA  
 CCTGTGGGTGACCGTGTACTACGGCGTGTGGCGGACCGGAGACCCCTGTCTGGCCCTCCGACCCCAAGGCCATACGACACCCG  
 AGGTGCCACAACGTGTGGGCCACCCACCGCTGCGTGCACCCGACCCCAACCCAGAGATCCACCTGGAGAACCTGACCCGAGAACTTCAAC  
 ATGTGGAAGAACAACTGTTGGAGCAGATGCACGAGACATCATCTCCTGTGGACCACTCCTGAAGCCCTGCGTGAAGCTGACCCCTT  
 GTGGCTGACCTGGAGTGGCAACAACATCAACTCCAACACCACAACGCGGAGATCAAGAATGCTCTTCAACATGA  
 CCACCGAGCTGGCGGACAAGAGAGGTTACGCCCTGTTTACCGCTGGACGTGGTGCAGATCAACAAGAACAACTCCAGTACCGC  
 CTGATCAACTGCAACACTCCGCCATACCCAGCCTGCCCCAAGGTGCTTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCGGCTT  
 CGCCATCCTGAAGTGCACGACAAGGATTCACGGCACCGCCCTGCAAGAACGTGTCACCGTGCAGTGCACCCACGGCATCAAGCCCG  
 TGGTGTCCACCCAGCTGCTGTGAACGGCTCCCTGGCCGAGGAGATCGTGATCCGCTCCGAGAACATCACCAACAACGCAAGACCATC  
 ATCGTGCAGCTGGTGAAGCCCGTGAAGATCAACTGCACCCGCCCCAACAAACAACACCCGCAAGTCCGTGGCATCGGCCCGGACACCTT  
 CTACGCCACCGGACATCATGGCGACATCCGCCAGGCCCACTGCAACGTGTCGCCGACCAAGTGGAAACAACACCCTGCAGCAGGTGGCCA  
 CCCAGCTGGCAAGTACTCAACAAGACCATCATCTTCGCCAACCCCTCCGGCGGACCTGGAGATCACCAACCCTCCCTCAACTGCCG  
 GCGAGTTCTTACTGCAACACTCCGAGCTGTTCAACTCCACTGGAATCCACTGGAAACAACACCGAGAACTGCATCACCTGCAGTG  
 CCGCATCAAGCAGATCGTGAACATGTGGCAGAAGTGGCCAGGCCATGTACGCCCCCTCCATCCAGGCGTGCATCCGCTGCGAGTCCAACA  
 TCACCGCCTGCTGTGACCCCGACGGCGCAACAACACTCCACAAACGAGACCTTCGCCCCCGGCGCAAGCGCGGACATGCGGAGTCCAACA  
 CGTCCGAGCTGACAAGTACAAGTGGTGAAGATCGAGCCCTGGCGTGGCCCAACCCCGCCCAAGCGCCGCTCCATCGCGTGGAGCCGAGAA  
 GCGGCCGTGGCCCTGGCGCCGTTCCTGGGCTTCCTGGGCGCCCGGCTCCACCATGGCGCCAGCAGTGCATCACCTGACCCGTGCAGG  
 CCGCCAGCTGCTCCGGCATCGTGCAGCAGTCCAACCTGCTGCGGCCATCGAGGCCAGCAGCTGCTGGGCATCTGGGCATCTGGGCAAGCT  
 GGATCAAGCAGCTGCAGGCCCGCTGCTGGCTGGAACCTCCTGTTCCAAACAAGACCTACAACGACATCTGGACAACATGACCTGGCTGCAAGTGGGACA  
 AGGAGATCTCCAACTACACCGACATCATCTAACACCTGATCGAGGAGTCCAGAACCAGCAGGAGAAAGACGAGGACCTGCTGGCCCTG  
 GACAAGTGGCCCTCCCTGTGGAACCTGGTTCGACATACCAACTGGCTGGTACATCAAGATCTTTCATCATGATCGTGGCGGCCCTGATCGG  
 CCTGGCCATCGTGTCCCGTGTGACCATCATCAACCGCTGGCCAGGCTACTCCCCCTGTCTCCAGACCTGACCCACCCAGC  
 GCGAGCCCGACCGCCCGAGCGCATCGAGGCGGCGGAGCAGCAGCCGCTCCGTGGCCCTGGTGTCCGGCTTCCGGCTTCCGGCCCTG  
 GCGTGGGACGACCTGCGCTCCCTGTCCCTACCAACCGCTGCGGACTTCGTGTGATCGCCGCCCGCAGCAGCTGGAGTGTGGG  
 CCACTCCCTCCCTGAAGGCGCTGGGCTGGGACTGAACTGGCAACTGCTGTCTTACTGGGCCAGGAGCTGAAGAAT  
 CCGCCATCAACCTGCTGGACACCATCGCCATCGCCGTTGGCCAACTGGACCCGACCGGCTGATCGAGATCGGCCAGCGCGGCCGCCATC  
 CTGAACATCCCCCGCCCATCCGCCAGGGCTGGAGCGGCCCTGTCTGTAA

91/178

Fig. 55A

2003 CON 03 AB Env  
 MRVKEIRKHLRWRGTLFLGMLMICSAATENLWTVYGYVPVWKEATTLFCASDAKAYSKEVHNWATYACVPTDPSQEIPILENVVTENFNMG  
 KNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLKKNVTSTNTSSIKMMEMKNCSEFNITDLRDKVKKEYALFYKLDVVQIDNDSYRL  
 ISCNTSVTQACPKISFEPIPIHYCAPAGFAILKCNDDKFKFGTCTNVSTVQCTHGKIPVVSQLLNGLSLAEAEVVISRVNFTDNTKTII  
 VOLKEPVEINCTRPNNNTRKGIHIGPGRAFYATGDIIGDIRQAHCNISITKWNNTLKQIVIKLRKQFGNKTIVFNQSSGGDPEIVMHSFNCG  
 GEFFYCNTTKLFNSTWNGTEELNTEGDIVTLPCRIKQIINMWQEVGKAMYAPPIAGQIRCSSNITGLLLTRDGGNQSNVTEIFRPGGDMR  
 DNWRSELYKYKVKVIEPLGVAPTAKRRRVQREKRAVIGAVFLGFLGAAGSTMGAASITLVQARQLLSGIVQQNNLLRAIEAQHLLQI  
 TVWGIKQLQARVLAVERYLKDQQLLGIWCSGKLICTFAPWNTSWSNKSLSDEIWNNTWMEWEREINNYTGLIYNLIEESQNOQEKNEQEI  
 LALDKWASLWNWFDISKWLWYIKIFIMIVGGLVGLRIIFAVLSIVNRVRQGYSPLSFQTRLPTQRGPDREPGEIEEGGERDRDTSIRLVNGF  
 LALIWDDLRSLCLFIYHHLRDLLLLIAARIVELLGRRGWEALKYWWNLLQYWIQELKSSAINLIDTIAIAVAGWTDRIEIGQREFCRAIRNIP  
 RRIRQGAEKALQ\$

Fig. 56A

2003 CON 04 CPX Env  
 MRVMGIQRNYPHLWEWGTLILGLVICSASKNLWTVYGYVPVWRDAETTFPCASDAKAYDKEVHNWATHACVPTDPPNPQEIALKNVVTENF  
 NMWKNNMVEQMHEDIISLWDEGLKPCVKLTPLCVALNCSNATINNSTKTNSTEIEIKNCSEFNITTEIRDKVKKEYALFYRLDIVPINDSANN  
 SINSEYMLINCNASTIKQACPKVTFEPIPIHYCAPAGFAILKCNDDKFKFGTCTNVSTVQCTHGKIPVVSQLLNGLSLATEGVVIRSKNF  
 TDNTKNIIVQLAKAVKINCTRPNNNTRKSVHIGPGQTWATGEIIGDIRQAHCNISGNDWNETLQKIVEELRKHFPNKTIIFAPSAGDLEI  
 TTHSFNCGGEFFYCNTSELFNSTYMNSTNTINKTITLPCRIKQIVSMWQEVGQAMYAPPIAGSINCSSDITGIIILTRDGGNNNTNNEFR  
 PGGDMRDNRSELYKYKVKVIEPVGVAPTRARRRVQREKRAVIGAVFLGFLGAAGSTMGAASITLVQARQLLSGIVQQNNLLRAIEA  
 QQHLLRLTVWGIKQLQARVLALESYLKQQLLGIWCSGKLICTTNVPWNSWSNKSINDIWNMTWLQWDKEINNYTQIIYELLEESQNOQ  
 EKNEQDLLALDKWANLWNWFNINSWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSLQTLIPTTQRGPDREPGEIEEGGERDRSR  
 SIRLVNGFLPLIWDLLRNLCLEFSYRHLRNLLLLIVARTVELLGIWGWEALKYLWNLLLYWGQELRNSAINLLDTTIAIAVAEGTDRIIEAVQRA  
 CRAIRNIPRRIRQGLERALL\$



92/178

## Fig. 55B

2003 CON 03 AB Env. seq. opt  
 ATGGCGTGAAGGAGATCCGCAAGCACTGTGGCGTGGGACCCCTGTTCTTGGGCATGTGTGATGATCTGCTCCGCCACCGAGAACCTGTG  
 GGTGACCGTGTACTAGCGGTGCCCGTGTGAAGGAGGCCACCAACCCCTGTGTGGCCCTCCGACGCCAAGCCCTACTCCAAGGAGGTGC  
 ACAACGTGTGGCCACTACCGCTGCGTCCACCGACCCCTCCCGCCAGGAGATCCCGCTGGAGAACGTGACCAGGAACTTCAACATGGGC  
 AAGAACAACATGTTGGAGCAGATGACGAGGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGTGAACCCCTGTGCGGT  
 GACCTGAACCTGCACCGACCTGAAGAAGACGTGACCTCCACCAACCTCTCCATCAAGATGATGGAGATGAAGAACTGCTCCTTCAACA  
 TCACCAACCGACTGGCGGACAAAGTGAAGAAGGATACGCCCTGTTCTAAGTGGACGTGTGAGATCGAACACGACTCCTTACCGCCCTG  
 ATCTCTGCAACACTCCGTGTTGACCCAGCCCTGCCCAAGATCTCTTCCAGCCCATCCCATCACTACTGCGCCCCCGCCGCTTCGC  
 CATCTGAAGTGCAACGACAAGAGTTCAACGGCACCGCCCTGCAACACGTGTCCACCCTGTCAGTGCACCCACGGCATCAAGCCCGTGG  
 TGTCCACCCAGCTGCTGAAACGGTCTCCCTGGCCGAGGAGGTGTGATCCGCTCCGTGAACCTTACCAGAACACCAAGACCATCATC  
 GTGAGCTGAAGGAGCCCGTGGAGATCAACTGCAACCCGCCCCAACAAACACCGCAAGGCAATCAACATCGGCCCCCGCCGCTTCTA  
 CGCCACCGGGACATCATCGCGGACATCCCGCAGGCCACTGCAACATCTCGATCACCAAGTGAACAACACCCCTGAAGCAGATCGTGATCA  
 AGCTGGCAAGCAGTTCGGCAACAAGACCATCGTGTCAACCAAGTCTCCGGCGGACCCCGAGATCGTGATGCACTCCTTCAACTGCGGC  
 GCGAGTCTTACTGCAACACCAACCAAGTGTCAACTCCACTGGAACGGCACCGAGGAGTGAACAACACCGAGGGCGACATCGTGAC  
 CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGTGGCAAGGCCATGTACGCCCCCCCATCGCCGCCAGATCCGCTGCT  
 CCTCCAACATCACCGCTGCTGACCCGACGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA  
 GACAACTGGCGCTCCGAGCTGTACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA  
 GCGGAGAACCGCGTGGCATCGCGCCGTTCTGGGCTTCTGGGCGCCGCTCCACCATGGGCCCAAGCCCAAGCCCAAGCCGCGGTGGA  
 CCGTGCAGGCCCGCAGTGTGTCCGGCATCGTGCAGCAGCAACAACCTGTGGCGCCCATCGAGGCCCAAGCAGTGTGGCATCTGGGCTGCTC  
 ACCGTGTGGGCATCAAGCAGTGTGCAGGCCCGTGTGGCAACCTGTGGCGCCCATCGAGGCCCAAGCAGTGTGGCATCTGGGCTGCTC  
 CGCAAGCTGATCTGCACACCGCTGTGGCAACCTGTGGTCCAAAGTCCCTGGACGAGATCTGGAAACAACATGACCTGGATGG  
 AGTGGGAGCGGAGATCAACAACCTGATCGAGGAGTCCAGAACCCAGCAGGAGGAGAACCGAGCAGGAGATC  
 CTGGCCCTGGACAAGTGGCCCTCCCTGTGGAACCTGATCGACATCTCCAAGTGGCTGTGTTACATCAAGATCTTCAATCATGATCGTGGGG  
 CCTGGTGGCCCTGGCATCTCGCCGTGTCCATCGTGAACCGGTGCCCAGGGCTACTCCCGCTGTCTCCAGACCCCGCTGC  
 CCAACAGCGGCCCGACCGCCGAGGGCATCGAGGAGGGCGGAGCGGACCGGACACCTCCATCCGCTGGTGAACGGCTTC  
 CTGGCCCTGATCTGGACGACCTGGCTCCCTGTGCTTCACTACCAACCTGCTGGCGACCTGTGCTGATCGCCCGCCGATCGTGA  
 GCTGTGGCCCGCGGCTGGAGGCCCTGAAGTACTGGTGAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGTCTCCCGCCATCAAC  
 TGATCGACACCATCGCCATCGCCGTGGCCGGCTGGACCGGCTGATCGAGATCGGCCAGCGCTTCTGCCCGGCCATCCGCAACATCCCC  
 CGCCGCATCCGCCAGGGCGCCGAGAGGCCCTGCAGTAA

Fig. 56B

2003 CON 04 CPX Env. seq. opt

ATGGCCGTGATGGGCATCCAGCGCAACTACCCCCACCTGTGGAGTGGGGCACCCCTGATCCTGGGCCCTGGTGATCATCTGCTCCGGCCCTCCAA  
 GAACCTGTGGGTGACCCGTGTACTACGGCGTGGCCGACGCCGAGACCACCCCTTCTGGCCCTCCGACGCCCAAGCCCTACGACA  
 AGGAGGTGCACAACATCTGGGCCACCACCGCTGGTGGCCACCGACCCCAACCCCCAGGAGATCGCCCTGAGAAACGTGACCCGAGAACTTC  
 AACATGTGGAAGAACAAACATGTTGGAGCAGATGCACGAGGACATCATCTCCTGTGGACGAGGGCCCTGAAGCCCTGCGTGAAGCTGACCCCC  
 CCTGTGCGTGGCCCTGAACTGCTCCAACGCCACATCAACAACTCCACCAAGACCAACTCCACCGAGAGATCAAGAACTGCTCCTTCAACA  
 TCACCAACCGAGATCCGCGACAAAGAAGAGGAGTAGCCCTGTCTACCCGCTGGACATCGTGGCCATCGTGGCCATCAACGACTCCGCCAACAAACAAC  
 TCCATCAACTCCGAGTACATGCTGATCAACTGCAACGCCCTCCACCATCAAGCAGGCCCTGCCCAAGGTGACCTTCGAGCCCATCCCCATCCA  
 CTACTGGCCCCCGCCGGCTTCGCCATCCTGAAGTCAACGACAAGAACTTCAACCGGCTGGCCCTGCACCAACGTTGCTCCCTCCGTGCAGT  
 GCACCCACGGCATCAAGCCCGTGTCCACCCAGCTGCTGTAACGGCTCCCTGGCCACCGAGGGCTGGTGTGATCCGCTCAAGAACTTC  
 ACCGACAACACCAAGAACATCATCTGTCAGCTGGCCAAAGCCGTGAAGATCAACTGCAACCCGCCCAACAACAACACCCGAAAGTCCGTGCA  
 CATCGGCCCCCGCCAGACCTGGTACGCCACCAGGAGATCATCGGGACATCCGCCAGGCCACTGCAACATTCGGCAACGACTGGAACG  
 AGACCCCTGCAGAGATCGTGGAGGACTCGGCAAGCACTTCCCAACAAAGAACCATCATCTTCGCCCTCCGCCGGCGGACTGGAGATC  
 ACCACCCACTCCTTCAACTGCGGGCGGAGTCTTCTACTGCAACACTCCGAGCTGTTCAACTCCACTACATGAATCCACCAACTCCAC  
 CACCATCAACAAGACCATCACCCCTGCCATCAAGCAGATCGTGTCCATGTGGAGGAGGTGGCCAGGCCATGTACGCCCCCCCA  
 TCGCCGGCTCCATCAACTGCTCCTCCGACATCACCCGACATCACTGACCCCGACCGGCCAACAACAACAACGAGACCTTCCGC  
 CCGGGCGCGGACATGCGCGACAATGGCGTCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCGTGGCGGTGGCCCAACCCG  
 CGCCCGCCCGCGTGTGACGGCGGAGAGCCCGCCAGCTGTCCGGCATCGTGGCCCTCCCTGGGCTTCTGGGCGCCGCTCCACCATGG  
 GCGCCCTCCATCACCTGACCGTGCAGGCCCGCCAGCTGTCCGGCATCGTGCAGCAGAGTCCAACTGCTGCGCGCCATCGAGGCC  
 CAGCAGCACCTGCTGGCCCTGACCCGTGGGGCATCAAGCAGTGCAGGCCCGGCTGGCCCTGGAGTCCCTACCTGAAGGACCCAGCAGCT  
 GCTGGGCATCTGGGGTGTCCGGCAAGTGTGCAACCAACAGTGCCTGGAACTCCTCTGGTCCACAAGTCTTACAACGACATCT  
 GGGACAACATGACCTGGCTGCAGTGGGACAAGGAGATCAACAATAACCCAGATCATCTACGAGCTGTGGAGGAGTCCAGAACCCAGCAG  
 GAGAAGAACGACGAGACCTGTGGCCCTGGACAAGTGGCCCAACTGTGGAACTGGTCAACATCTCCAACCTGGCTGGTGTGGTACATCAAGAT  
 CTTCAATCATGATCGTGGCGGGCTGTGGCCCTGGCCATCATCTTCGCCGTGTCTCCATCGTGAACCCGGTGGCCAGGGCTACTCCCCC  
 TGTCCCTGCAGAACCTGATCCCCAACCCAGCGGCCCGCCGACCCCGAGGCCACCGAGGAGGAGGCCGCGGAGCAGGACCGCTCCCGC  
 TCCATCCGCCCTGGTGAACGGCTTCTGCCCTGTGTTGGACGACTGGCAACTGTGCTGTCTCTACCCGACCTGGCAACCTGCT  
 GCTGATCGTGGCCCGCACCGTGGAGCTGTGGGCATCCGCGGCTGGAGGCCCTGAAGTACCTGTGGAACCTGCTGTACTGGGGCCAGG  
 AGCTGGCAACTCCGCCATCAACTGCTGGACACCAACCCGATCGCCGTGGCCGAGGGCACCGACCCGATCATCGAGGCCGTGCAGCGGCC  
 TGCCCGGCCATCCGGCAACATCCCCCGCCGATCCGCCAGGGCCTGGAGCGCCCTGTGTAA

Fig. 57A

2003 CON 06 CPX Env  
 MRVKGIQKNWQHLLWKWGFLILGLVICSASNNMVTVYGYVPAWEDADTILFCASDAKAYSAEKNVWATHACVPTDPNPQEI ALENVTENF  
 NMWKNHMEQMHEDIISLWDESLKPCVKLTPLCVTLNCTNVTKNNNTKIMGREEEKNCSEFNVTEIRDKKKEVALFYRLDVVPIDDDNNSY  
 RLINCNASTIKQACPKVSEFEPIPIHYCAPAGFAILKCRDKNFNGTGPCKNVSTVQCTHGKPKVSTQLLNGSLAEELIIKSENLTDNTKT  
 IIVQLNKSVEIRCTRPNNNTRKSI SFGPGQAFYATGDIIGDIRQAHCVSRTDMNNMLQNVTAKLKELENKNI TFNSSAGGDLEITTHSFNC  
 GGEFFYCNTS QLFENSTRPNETNTITLPCIKIKQIVRMWQRVGQAMYAPPIAGNITCTSNITGLLLTRDGNNDSETFRPGGDMRDNRSELY  
 KYKVKIKPLGLAPTRARRRVVGREKRAVGLGAVFLGFTAGSTMGAASITLTVQVRQLLSGIVQQSNLLRAIEAQQHLIQLTVWGIKQI  
 QARVLAVERYLKDQQLLGIWCSGKLI CPTNVPWNASWSNKTINEIWNMTWIEWDREINNYTQOISLIEESONQOQEKNEQDLLALDKWAS  
 LWSFEDISNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSLQTLIPNPTGADRPGEIEEGGEGQGRTRSIRLVNGFLALAWDDL  
 RSLCLFSYHRLRDFVLLAARTVETLGHARGWEILLKYLGNLVCYWGQELKNSAISLLDITTAIAVANWTDRVIEVVQRFRAFLNIPRRIRQGF  
 RALL\$

Fig. 58A

2003 CON 08 BC Env  
 MRVGRTRRNYQQWIIWVGLGFWMLMVCNVEGNLWTVVYGYVPMWKEAKTTLFCASDAKAYETEVEHNWATHACVPTDPNPQEI VMENVTENF  
 NMWNNDMVNQMHEDEVISLWDQSLKPCVKLTPLCVTLNCTNVSSNGNGTYNETYNEVKEIKNCSEFNATLLRDRKKTYYALFYRLDIVPLND  
 ENSGKNSSEYYRLINCNTSAITQACPKVTFDPIPIHYCTPAGYAILKCNDDKKNFTGQCHNVSTVQCTHGKPKVSTQLLNGSLAEELII  
 RSENLTNNVKTIIVHLNQSVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCVSRTDMNNMLQNVTAKLKELENKNI TFNSSAGGDLEITTHSFNC  
 GDEITTHSFNCRGEEFFYCNTSGLFNGTYMNGTNNSSII TIPCRIKQIINMWQEVGRAMYAPPIEGNITCKSNITGLLLVRDGGRTESNNT  
 EIFRPGGDMRNNWRNELYKYKVEIKPLGVAPTAARRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLR  
 AIEAQQHMLQLTVWGIKQLQTRVLAIERYLKDQQLLGIWCSGKLICTTAVPWNSSWSNKSQQEIEWDNNMTWMQWQDKEISNYTNTIYRLLLED  
 QNQQRNEKDLLALDWSKNLWSWFDITNWLWYIKIFIMIVGGLIGLRIFAVLSIVNRVRQGYSPLSFQILTPNPGGPRGLGRIEEGGEGQD  
 KTRSIRLVNGFLALAWDDLRLNCLFSYHRLRDFILLTARGVELLGRNSLRGLQRGWEALKYLGSLVQYWGLELKKSTISLVDITIAIAVAE  
 DRIINIVQGICRAIHNI PRRIRQGFEEAALQ\$

Fig. 57B

2003 CON 06 CPX Env. seq. opt  
 ATGGCGGTGAAGGCGCATCCAGAAGAACTGGCAGCACCTGTGGAAGTGGGGCACCCCTGATCCTGGGCCCTGGTGATCATCTGCTCCGCCCTCCAA  
 CAACATGTGGTGGTACCCGTGTACTACGGCGTGGAGGACCGCCGACACCCATCCTGTCTGGCCCTCCGACGCAAGGCCACTACTCCG  
 CCGAAGCACAAACGTGTGGGCCACCCACCGCTGCGTGGCCACCCCAACCCAGGAGATCGCCCTGGAGAACGTGACCGAGAACTTC  
 AACATGTGGAAGAACCAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGACGAGTCCCTGAAGCCCTGCGTGAAGCTGACCC  
 CCTGTGCTGACCCCTGAACCTGCCAACCAACCGTGAACCAAGAACCAACCAAGATCATGGCCCGGAGGAGATCAAGAACTGCTCCTTCAACG  
 TGACCACCGAGATCCGGGACAAGAGAGAGTACGCCCTGTCTACCGCTGGACGTGGTGCCTATCGACGACAAACAACAACTCCCTAC  
 CGCCTGATCAACTGCAACGCCCTCACCATCAAGCAGGCCCTGCCCAAGGTGTCTTCCAGCCCATCCCATCCACTACTGCGCCCCCGCCGG  
 CTTCCGCATCCTGAAGTGCAGGACCAAGAACTTCAACGGCACCGCCCTGCAAGAACGTGTCCACCCTGTCAGTGCACCCACGGCATCAAGC  
 CCGTGTGTCCACCCAGCTGCTGTAACGGCTCCCTGGCCGAGGAGATCATCATCAAGTCCGAGAACCTGACCGACAAACACCAAGACC  
 ATCATCGTGCAGCTGAACAAGTCCGTGGAGATCCGTGCACCCGCCCAACAACAACACCCGCAAGTCCATCTCCTTCGGCCCGCCGAGCC  
 CTCTACGCCACCGCGACATCATCGCGACATCCGCCAGGCCACTGCAACGTGTCCGCACCGACTGGAACAACATGCTGCAGAACGTGA  
 CCGCAAGCTGAAGGAGCTGTTCAACAAGAACATCACCTTCAACTCCCGCCGGCGGACCTGGAGATCACCCACTCCCTTCAACTGC  
 GCGCGGAGTCTTCTACTGCAACACCTCCAGCTGTTCAACTCCACCCGCCCAACGAGACCAACACCATCACCTGCACCTCCAACATCACCGGCC  
 GCAGATCGTGGCATGTGGCAGCGCTGGCCAGGCCATGTACGCCCCCGCCCATCGCCGGCAACATCACCTGCACCTCCAACATCACCGGCC  
 TGCTGTGACCCCGACGCAACAACGACTCCGAGACCTTCCGCCCGCGGGCGACATGCGCGACAACTGGCGCTCCGAGCTCCGAGCTGTAC  
 AAGTACAAGGTGGTGAAGATCAAGCCCTGGGCATCGCCCCCAACCGCGCCCGCTCCATCACCTGACCCGTGCGAGGAGCGCCGTGGCCCT  
 GGGCCGCTGTCTGGCTTCCCTGGCCACCGCCGCTCCACCATGGCGCCCGCTCCATCACCTGACCCGTGACCGTGCAGGTGCGCCAGCTGCTGT  
 CCGGCATCGTGCAGCAGTCCAACTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCCGTGGGCGATGGGCGATCAAGCAGCTG  
 CAGGCCCGCTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGCTGCTCCGGCAAGCTGATCTGCCCCACCCAA  
 CGTGCCCTGGAACGCCCTCCTGGTCCAACAAGACCTACAACGAGATCTGGACAACATGACCTGGATCGAGTGGACCCCGGAGATCAACAAC  
 ACACCCAGCAGATCTACTCCCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAGAACGACGAGGACCTGTGGCCCTGGACAAGTGGGCCCTCC  
 CTGTGGTCCCTGGTTCGACATCTCAACTGGTGTGGTACATCAAGATCTTTCATCATGATCGTGGCGGCCCTGATCGGCCCTGGCATCGTGT  
 CGCCGTGCTCCATCGTGAACCGCTGCGCCAGGCTACTCCCCCTGTCCCTGCAGACCTGATCCCAACCCCAACCCGCGCCGACCCGCC  
 CCGCGAGATCGAGGAGGCGCGGAGCAGGCCCGCACCCGCTCCATCCGCTGGTGAACGGCTTCTTGGCCCTGGCCCTGGGACCGACTG  
 CGTCCCTGTGCTGTCTCCTACCCCGCTGCGGACTTCGTGTGATCGCCGCCCGCACCCGTGGAGACCTTGGGCCACCCGCGCTGGGA  
 GATCCTGAGTACCTGGCAACCTGTTGCTACTGGGCGCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAACCCGCTCGCCG  
 TGGCCAACTGGACCCGCGTGTGATCGAGGTGGTGCAGCGCTGTTCCGCGCCCTTCTCTGAACAATCCCCCGCCGCTCCGCGAGGCTTCGAG  
 CCGGCCCTGCTGTAA

Fig. 58B

2003 CON 08 BC Env seq. opt  
 ATGCGGTGGCGCACCCCGCAACTACCAGCAGTGGTGGATCTGGGGCGTCTGGATGCTGATGATCTGCAACCTGGAGGG  
 CAACCTGTGGGTGACCGTGTACTACGGCGTGTGGAAGGAGGCCAAGACCACCCCTGTTCTGCGCCTCCGACGCCAAGCCCTACGAGA  
 CCGAGGTGCACAACCGTGTGGGCCACCCACCGCTGCGTGGCCACCGACCCCAACCCAGGAGATCGTATGGAGAACGTGACCGAGAACTTC  
 AACATGTGGAACAACGACATGGTGAACCAAGATGCACGAGGACGTGATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGAACCTGGAGTGCACCAACCGTGTCCCAACGGCAACGGCACCTACAACGAGACTACAACGAGTCCGTGAAGGAGATCAAGA  
 ACTGCTCCTTCAACGGCCACCCCTGCTGCGCGACCGCAAGAACCGTGTACGCCCTGTTCTACCGCTGGACATCGTGCCCTTGAACGAC  
 GAGAACTCCGGCAAGAACTCCTCCGAGTACTACCGCTGATCAACTGCAACACCTCCGCCATCACCCAGSCCTGCCCAAGGTGACCTTCGA  
 CCCCATCCCCATCCACTACTGCACCCCGGCTACGCCATCCTGAAGTGCAACGACAGAAGTTCAACGGCACCGGCCAGTGCACAACG  
 TGTCCACCGTGCAGTCAACCCAGGCATCAAGCCGTTGTTCCACCCAGTGTGTAACGGTCCCTGGCCGAGCCGCGAGATCATCATC  
 CGTCCGAGAACCTTGACCACAACGTAAGACCCTTCTACGCCACCGCGGACATCATCGGGACATCCGCCAGGCCACTGCAACATCTCCA  
 CCGAAGTCCATCCGCATCGGCCCGGACAGCTTCAAGAGCTGCCGAGCATTTCCCAACAAGACCAATCAAGTTCGCCCTCCTCCCGGC  
 AGGACAAGTGGTACGAGACCTGACGCGGTGTCGAAGAGTGGCCGAGCATTTCTACTGCAACACCTCCGGCTGTCAACGGCACCTACATGAA  
 GCGACCTGGAGATCACCCACCTCCTTCAACTGCCCGCGGAGTTCTTACTGCAACACCTCCGGCTGTCAACGGCACCTACATGTA  
 CGGACCAACAATCCTCCTCCATCATCACCATCCCTGCCGATCAAGCAGATCATCAACATGTGGCAGGAGTGGCCGCGCCATGTACG  
 CCCCCCATCGAGGCAACATCACCTGCAAGTCCAACATCACCGGCTGTGCTGTGTCGCGACGGCGGCCACCGAGTCCAACAACAC  
 GAGATCTTCGCCCCGGCGGCGACATGGCCACAACATGGCCAACGAGCTGTACAAGTACAAGTGTGGAGTCAAGCCCCCTGGGCGT  
 GGCCCCAACCGCCGCAAGCGCGGTGGTGGAGCGGAGAGCGCGCTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCT  
 GCTCCACCATGGGCGCCCTCCATCACCTGACCGTGCAGCCGTCAGCCGTCAGTGCAGACCCCGCTGGAACTCCTCCCTGGTCCAACAAGTCCC  
 GCCATCGAGGCCAGCAGCAGTGTGAGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT  
 GGACCAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT  
 AGCAGGAGTCTGGGCAACATGACCTGGATGGATGGGCAAGGAGATCTCCAATACCAACACCACTACCGCTGTGGGCTGTGGGCTGTGGGCT  
 CAGAACGAGGAGGCAAGGAGGACCTGTGGCCCTGGACTCCTGGAAAGAACCTGTGGTCCCTGGTTCGACATCAACCACTGGCTGTG  
 GTACATCAAGATCTTCAATCATGATCGTGGCGGCTGATCGCCCTGGCATCATCTTCCCGTGTCCATCGTGAACCGGCTGGCCAGG  
 GCTACTCCCCCTGTCTCCAGATCCTGACCCCAACCCCGGCGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGC  
 AAGACCGTCCATCCGCTGGTGAACGGCTTCTGGCCCTGGCCCTGGCCGACCTGGCAACCTGTGCTGTCTTCTTACCCGCTGG  
 CGACTTCACTGTGACCGCCGCGGTGGAGCTGTGGCCGCAACTCCTTGGCGGCTGCAGCGGCTGGAGGCTGGAGGCTGGAGTACC  
 TGGGCTCCCTGGTGCAGTACTGGGCTGGAGCTGAAGAAGTCCACCATCTCCCTGGTGGACACCATCGCCATCGCCGTGGCCGAGGGCAC  
 GACCGCATCAACATCGTGCAGGGCATCTGCCGCGCATCCACAACATCCCCCGCGCATCCGCCAGGGCTTCGAGGCGCCCTGCAGTA

A

97/178

Fig. 59A

2003 CON 10 CD Env  
 MRVMGIQRNCOQWIIWGIIGFWMLMICNATGNLWVTVYYGVVWKEITTTLFCASDAKAYKAEAHNIWATHACVPTDPNPQEIIVLENTVENE  
 NMWKNMVDQMHEDIISLWDQGLKPCVKLTPLCVTLNCSVDNATNSATNTVVAGMKNCSFNITTEIRDKKKQYALFYKLDVVQIDGSNTSY  
 RLINCNTSAITQACPQVTFEPIPIHYCAPAGFAILKCNDDKFNGTGPKCNVSTVQCTHGIKPVVSTQLLNGSLAEIEIIRSENLTDNAKT  
 IIVQLNESVTINCTRPNNTRKSIIRIGPGQTFYATGDIIGNIRQAYCNIISGTEWNKTQQVAKKLGDLINKTTIIFKPSGGDPEITHTTFN  
 CGGEFFYCNTSKLFNSSWTSNNTGNTSTITLPCRIKQIINMWQGVGKAIYAPPIAGLINCSSNITGLLITRDGGANNSETFRPGGDMRDNW  
 RSELYKYKVVVIEPLGLAPTAKARRVVEREKRAIGLGAAGSTMGAASLTLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVW  
 GIKQLQARVLAVESYLKDQQLIGWCSGKHICTNVPWNSSWSNKSLEEIWDNMTWMEWEREIDNYTGLIYSLIEESQNKQEKNEQELLQL  
 DKWASLWNWFSITNWLWYIKIFIMIVGGLIGLRIVFAVLSLVNRVROGYSPLSFQTLPPAPRGPDRPEGIEEGEGEQGRGRSIRLVNGFSAL  
 IWDDLRLNCLFSYHRLRDLILLIATRIVEILLGRRGWEAIKYLWNLQYWIQELKNSAISLLDTTAIAVAEGTDRAIEIVQRAVRAVLNIPTRI  
 ROGLERALLS

Fig. 60A

2003 CON 11 CPX Env  
 MRVKETQRNWHNLWRWGLMIFGMLMICNATENLWVTVYYGVVWKKDADTTLFCASDAKAYSTEKHNWVWATHACVPTDPNPQEIIVLENTVENE  
 NMWKNMVEQMHEDIISLWDESILKPCVKLTPLCVTLNCTDVKNATNTTVEAAEIKNCSEFNITTEIKDKKKKEYALFYKLDVVPINDNNSIY  
 RLINCNVSTVKQACPQVTFEPIPIHYCAPAGFAILKCNDDKFNGTGPKCNVSTVQCTHGIKPVVSTQLLNGSLAEIEVIRSENFNNAKT  
 IIVQLNSSVRINCTRPNNTRKSIHIGPGQAFYATGDIIGDIRQAHANIISRAEWNNTLQQVAKQLRENFNKTIIFNNPSGGDLEITTHSFNC  
 GGEFFYCNTSRLENSTWNNDRNDTKQMHITLPCRIKQIIVNMWQRVGQAMYAPPIQKIRCNSTITGLLITRDGGNNNTNETFRPTGGDMRD  
 NWRSELYKYKVVIEKPLGVAPTRAKRRVVEREKRAVIGAVLLGFLGAAGSTMGAASITITVQARQLLSGIVQQNNLLKAEAQHLLKLT  
 VWGIKQLQARVLAVERYLKDQQLIGWCSGKLICTTNVPWNFSWSNKSDEIWDNMTWIEWEREINNYTQTIYTLLEESQNKQEKNEQDLLL  
 ALDKWASLWNWEDI SNWLWYIKIFIMIVGGLIGLRIFAVLSIVNRCRQGYSPLSFQTLTPNHKEADRPGGIEEGEGEQDRTRSIRLVSGFL  
 ALAWDDLRLNCLFSYHRLRDFILIAARIVETLGRRGWEILKYLGNLAQYWGQELKNSAISLLNATAIAVAEGTDRIIEVVHRVLRALLHIPR  
 RIRQGFERALLS

Fig. 59B

2003 CON 10 CD Env. seq. opt

ATGCCGCTGATGGGCATCCAGCGCAACTGCCAGCAGTGGTGGATCTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGCCACC

CAACCTGTGGTGGTACCCTGACTACGGCGTCCCCTGTGGAAGGAGACCACCAACCCCTGTTCTGGCCCTCCGACGCCAAGCCCTACAAGG

CCGAGGCCACAACATCTGGGCCACCACCGCTGCGTGGCCACCGACCCCAACCCAGGAGATCGTCTGGAGAACGTGACCCGAGAACTTC

AACATGTGAAGAACCGGCATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGACCAGGGCTGAAGCCCTGCGTGAAGCTGACCC

CCCTGTGGTACCCTGAACTGCTCCGACCGTGAACGCCACCAACTCCGCCACCAACACCCTGGTGGCCGGCATGAAGAACTGCTCTTCAACA

TCAACCCAGATCCCGCACAGAAGAAGCAGGAGTACGCCCTGTTTACAAGCTGGACGTGGTGCAGATCGACGGCTCAACACACTCTCTAC

CGCTGATCAACTGCAACACTCCGCCATCACCCAGGCTGCCCAAGTGCCTTTCGAGCCCATCCCATCCACTACTGCGCCCGCCGG

CTTCGCCATCCTGAAGTGAACGACAAGAGTTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGC

CCGTGGTGTCCACCCAGCTGAACCGTGAACGGCTCCCTGGCCGAGGAGATCATCATCCGCTCCGAGAACCTGACCAGCAACGCCAAGACC

ATCATGTGCAGCTGAACGAGTCCGTGACCATCACTGCACCCGCCCAACAACACCCGAAAGTCCATCCGATCCGATCGGCCCTGATCAACTGCTCCT

CTTACGCCACCCGGACATCATCGGCCAATCCGCCAGGCTACTGCAACATCTCCGGCACCGAGTGAACAAGACCCCTGCAGCAGGTGG

CCAAGAAGCTGGCGACCTGTGTAACAAGACCACTCATCTTCAAGCCCTCCTCCGGCGGACCCCGAGATCACCAACCCACACTTCAAC

TGCCGGCGGAGTCTTCTACTGCAACACTCCAAGTGTTCAACTCCTCTGGACCTCCAACAACCCGGCAACACTCCACCATCACCCCT

GCCCTGCCGATCAAGCAGATCATCAACATGTGGCAGGGCTGGCAAGGCCATCTACGCCCCCCCATCGCCGCTGATCAACTGCTCCT

CCAACATCACCGCTGTGCTGACCCGGACGGCGCCCAACAATCCGAGACCTTCGCCCCCGGGCGGCGACATGCGCGCAACTGG

CGCTCCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCTGGGCTGGCCCTCCACCAAGCCCAAGCCCGGCTGGTGGAGCCGAGAA

GGCCGCATCGCCCTGGCGCCGTGTTCTGGGCTTCTGGCGCCGCGCTCCACCATGGCGCCGCTCCCTGACCCCTGACCCCTGACCCGTGCAGG

CCCGCAGCTGCTGCCGATCGTGACGACGAGAACACTGCTGGCGCCCATCGAGGCCACGAGCCTGTCGAGCTGGCCATCTGGGCTGACCCGTGG

GGCATCAAGCAGCTGCAGGCCCGCTGCTGGCCGTGGAGTCCCTACCTGAAGGACCAGCAGCTGCTGGCATCTGGGCTGCTCCGGCAAGCA

CATCTGCACCAACCTACCCGGCTGATCTACTCCCTGATCGAGGAGTCCAGAACCCAGCAGGAGATCTGGACAACAATGACCTGGATGGAGTGGGAGC

GCGAGATCGACAACCTACCCGGCTGATCTACTCCATCACCACCTGGCTGTGTACTCAAGATCTTCAATCATGATCGTGGCGGCCCTGATCGG

GACAAGTGGCCCTCCCTGTGGAACCTGTTCTCCCTGGTGAACCGGCTGCCAGGGCTACTCCCCCTGCTCCAGACCCCTGCTGCCCGCCCTG

CCTGGCATCGTGTTCGCCGTGCTCCCTGGTGAACCGGCTGCCAGGGCTACTCCCCCTGCTCCAGACCCCTGCTGCCCGCCCTGCTCCG

GCGCCCGACCGCCCGAGGCATCGAGGAGGAGGCGGCGGAGCCCGCTCCATCCGCTGGTGAACGGCTTCTCCGCTGCTCCGCTG

ATCTGGACGACCTGGCAACCTGTGCTGTTCTCCTACCCGCTGCCGACCTGATCCTGATCGCCACCCGCATCGTGGAGTGTCTGG

CCGCCGGCTGGAGGCCATCAAGTACCTGTGGAACCTGTGGAACCTGATCCAGGAGCTGAAGAATCCGCCATCTCCCTGCTGGACA

CCACCGCCATCGCCGTGGCCGAGGCCACCGACCCGCTGCAGCGGCCCTGCTGAACATCCCCACCCCGCATC

CGCCAGGGCTGGAGCGGCCCTGCTGTAA

Fig. 60B

2003 CON\_11 CPX Env. seq. opt

ATGCGCGTGAAGGAGACCCAGCGCAACTGGCACAACCTGTGGCGCTGGGCGCTGATGATCTTCGGCATGCTGATGATCTGCAACGCCACCGA  
 GAACCTGTGGGTGACCGTGTACTACGGCGTGCAGGAGCGCCGACACCCCTGTTCTGGCCCTCCGACGCCAAGGCCCTACTCCA  
 CCGAGAAGCAACAACGTGTGGCCACCCACCGCTGCGTCCCAACCCAGGAGATCCCTGGAGAACGTGACCGGAACTTC  
 AACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGACGAGTCCCTGAAGCCCTGCGTGAAGCTGACCC  
 CCTGTGCGTGAACCTGAACCTGCACCGACGTGAAGACGCCACCAACACACCTGGAGGCCGCCGAGATCAAGAACTGCTCTTCAACATCA  
 CCACCGAGATCAAGGACAAAGAAGAGTAGCCCTGTTCTAAGCTGGACGTGGCCCATCAACGACAAACAACACTCCATCTAC  
 CGCCTGATCAACTGCAACGTGTCCACCGTGAAGCGCTGCCCAAGTGCCTTCGAGCCCATCCCATCCACTACTGCGCCCGCCCGG  
 CTTGCGCATCCACCCAGCTGCTGTGAACGGTCCCTGGCCGAGGCGAGGTGGCATCCGCTCCGAGAACTTCAACAAACGCCAAGAC  
 ATCATCGTGCACTGACTCCTCCGTGGCATCAACTGCACCGCCCAACAAACACCCGCAAGTCCATCCACATCGGCCCGCCAGG  
 CTTTACGCCACCGCGACATCATCGCGACATCCGCGAGGCCACTGAAACATCTCCCGCGCCGAGTGGAAACAACCCCTGCAGCAGGTGG  
 CCAAGCAGTGGCGGAACTTCAACAAGACCATCATCTTCAACCCCTCCGCGGCGACTGGAGATCACCACTCCCTTCAACTGC  
 GGCGGAGTCTTCTACTGCAACAACCTCCCGCTGTTCAACTCACTGGAACAACGACACCCGCAACGACCAAGCAGATGCACATCAC  
 CCTGCCCTGCCGATCAAGCAGATCGTGAACATGTGGCAGCGGTGGCCAGGCCATGTACGCCCTTCCGCCCCCATCCAGGCAAGATCCGCTGCA  
 ACTCCAAACATCACCGCTGTGTAACCCGACAAAGTGGTGGAGATCAAGCCCTGGCGTGGCCCCACCAACGAGACTTCCGCCCCACCGCGGACATGCGCGAC  
 AACTGGCGTCCGAGTGTACAAGTACAAGTGGTGGAGTCAAGCCCTGGCGTGGCCCCACCCGCGCAAGCCCTCCATCCAGGCAAGATGCGCGGAC  
 CGAGAAGCGCGCGTGGCATCGCGCGCTGTGCTGGCTTCTGGCGCGCCGCTCCACCATGGCGCCAGCAGTCCAGGCCCCAGGCGGAGTGGTGGAGCG  
 TGCAGGCCCGCAGTGTTCGGCATCGTGACAGTCCAACTGTGAAGGCCATCGAGGCCAGCAGTCTGGGCAACATCGGCGTGGTGAAGCTGACC  
 GTGTGGGCATCAAGCAGTGCAGGCCCGCTGGAACCTTCTCTGTCCAAACAAGTCCAGGAGTCCCAAGACGAGATCTGGGCAACATGACCTGGATCGAGT  
 CAAGTGTGACCAACCAACTACACCCAGCCTACACCCCTGTGGAGGAGTCCCAAGACTGGTGTGATACATCAAGATCTTCAATCATGATCGTGGCGGCGCT  
 GCGAGCGGAGATCAACAAGTGGCTCCCTGTGGAACCTGTGACTGATCCCAACTGGTGTGGTACATCAAGATCTTCAATCATGATCGTGGCGGCGCT  
 GCCCTGGACAAGTGGCCCTCCCTGTGGAACCTGTGACTGATCCCAACTGGTGTGGTACATCAAGATCTTCAATCATGATCGTGGCGGCGCT  
 GATCGGCTGGCCATCATCTTGGCGTGTCCATCGTGAACCGTGCAGGAGGCGCGGAGCAGGACCCGCTCCATCCGCTGGTCCAGACCTGACCCCA  
 ACCACAAGGAGGCCGACCCCGCGGCGCATCGAGGAGGCGCGGAGCAGGACCCGCTCCATCCGCTGGTCCGCTGGTCCGCTTCCG  
 GCCCTGGCTGGACGACCTGGCAACCTGTGCCTGTCTCTACCAACCGCTGCGCGACTTCACTGTGATCGCCGCGCGCATCGTGGAGAC  
 CCTGGCCCGCGGCTGGAGATCTGAAGTACCTGGCAACCTGGCCAGTACTGGGCGCAGGAGTGAAGAACTCCGCCATCTCCCTGC  
 TGAACGCCACCGCCATCGCCGTGGCCGAGGCAACCGCATCATCGAGGTGGTGCACCCGCTGCTGGCGCCATCTGCACATCCCCCGC  
 CGCATCCGCCAGGGCTTCGAGCGGCGCTGTGTA



Fig. 61A

2003 CON 12 BF Env

MRVRGMQRN̄WQH̄LGK̄WGL̄LFL̄GILL̄IIC̄NATENL̄WVT̄VYȲGVP̄VW̄KEAT̄TIL̄FCASD̄AKSȲERĒVHN̄VW̄ATHAC̄VPT̄DP̄NP̄QĒVDLĒNV̄TEN̄F  
DM̄WK̄NN̄MV̄EQ̄M̄HT̄DĪISL̄WD̄Q̄SL̄K̄PC̄VK̄L̄T̄PL̄CV̄TL̄N̄CT̄DAN̄ATAN̄AT̄KĒH̄PĒGR̄AGAĪQ̄NC̄SF̄NM̄T̄TĒVR̄DK̄Q̄MK̄V̄Q̄AL̄FȲRL̄DĪV̄P̄IS̄DN̄  
NS̄NEȲRL̄INC̄NT̄ST̄IT̄Q̄ACP̄KV̄SW̄D̄PĪPĪHȲCAP̄AḠYAĪLK̄ND̄KK̄F̄NḠT̄ḠP̄CK̄NV̄ST̄V̄Q̄CT̄HḠIK̄PV̄V̄ST̄Q̄LL̄NḠSL̄AĒEĒIĪIR̄SQ̄NIS̄  
DNĀKTĪIV̄HL̄NES̄V̄Q̄INC̄TR̄P̄NN̄TR̄KSĪHĪḠP̄GRĀFȲAT̄GDĪIḠDIR̄KĀHC̄NV̄SḠT̄Q̄WN̄KT̄LĒQ̄V̄KK̄L̄RS̄ȲF̄NT̄TĪKEN̄SS̄SḠGD̄PĒIT̄M̄  
HS̄F̄N̄C̄R̄GĒFF̄ȲC̄NT̄SK̄LF̄ND̄TV̄S̄ND̄TĪIL̄PC̄RĪK̄Q̄IV̄NM̄W̄Q̄EV̄GRAM̄YĀAPĪAḠNĪT̄CT̄SN̄IT̄GL̄LL̄TR̄D̄ḠGH̄NET̄N̄KT̄ET̄FR̄PḠGN̄MK̄DN̄  
WR̄SĒLȲKȲKV̄VĒIĒPL̄GV̄AP̄TR̄AK̄R̄Q̄V̄K̄RĒK̄R̄AV̄ḠIḠAL̄FL̄ḠFL̄GĀAḠST̄M̄GĀAS̄IT̄L̄TV̄Q̄AR̄QL̄LS̄GĪV̄Q̄Q̄SN̄LL̄RAĪEĀQ̄Q̄H̄LL̄QL̄TV̄  
WḠIK̄QL̄Q̄AR̄VL̄ĀVER̄YL̄KD̄Q̄QL̄GL̄WḠCS̄ḠK̄LICT̄TN̄VP̄WN̄SS̄WS̄N̄KS̄Q̄EĒĪW̄EN̄MT̄W̄MĒWĒKĒĪNN̄YS̄NĒĪȲRL̄IĒES̄QN̄Q̄Q̄EK̄NĒQ̄ELLĀ  
LD̄KW̄AS̄L̄WN̄WF̄DIS̄N̄WL̄WȲIR̄IF̄IMĪV̄ḠGL̄IḠL̄R̄IVĒFĀVLS̄IV̄NR̄VR̄K̄ḠYS̄PL̄SL̄Q̄THĪPS̄PRĒP̄DR̄PĒḠIĒEḠḠEQ̄ḠK̄DR̄SV̄RL̄V̄NḠFLĀ  
LĪW̄DD̄L̄RS̄L̄CL̄FS̄ȲH̄RL̄R̄D̄LL̄LĪV̄TR̄IV̄ELL̄ḠRR̄ḠWĒVL̄KȲW̄WN̄LL̄Q̄ȲWS̄Q̄EL̄K̄NS̄AIS̄L̄INT̄TAĪV̄VĀEḠT̄DR̄V̄IĒAL̄Q̄RV̄GRĀILL̄N̄IP̄RR̄  
IR̄Q̄GL̄ER̄ALL̄S̄

Fig. 62A

2003 CON 14 BG Env

MK̄AK̄GT̄QR̄N̄W̄Q̄SL̄WK̄WḠTL̄IL̄GL̄VĪICS̄AS̄ND̄L̄W̄VT̄VYȲGVP̄VW̄KEAT̄TIL̄FCASD̄AKAȲDĀEV̄HN̄VW̄ATHAC̄VPT̄DP̄NP̄QĒVALĒNV̄TEN̄F  
NM̄WEN̄NM̄VD̄Q̄M̄Q̄EDĪISL̄WD̄Q̄SL̄K̄PC̄VEL̄T̄PL̄CV̄TL̄N̄CT̄DF̄N̄NT̄N̄NT̄N̄TR̄ND̄ḠĒGĒIK̄NC̄SF̄NĪT̄SL̄RD̄K̄IK̄KEȲAL̄FȲNL̄DV̄V̄Q̄MD̄ND̄  
NSS̄ȲRL̄T̄SC̄NT̄S̄ĪIT̄Q̄ACP̄KV̄S̄FT̄PĪPĪHȲCAP̄AḠF̄V̄IL̄K̄CN̄N̄KT̄F̄NḠT̄P̄CT̄NV̄ST̄V̄Q̄CT̄HḠIR̄PV̄V̄ST̄Q̄LL̄NḠSL̄AĒEĒĪV̄IR̄SK̄N̄FT̄D̄  
NĀKTĪIV̄OL̄K̄D̄PĪĒINC̄TR̄P̄NN̄TR̄KRĪTM̄ḠP̄GR̄V̄LȲTT̄ḠQ̄ĪIḠDIR̄KĀHC̄NIS̄KT̄K̄WN̄NT̄LḠQ̄IV̄KK̄L̄RĒQ̄FM̄N̄KT̄IV̄F̄QR̄SS̄GḠD̄PĒIV̄M̄  
HS̄F̄N̄C̄R̄GĒFF̄ȲC̄NT̄QL̄FN̄ST̄WR̄SN̄ST̄W̄ND̄T̄T̄ET̄N̄TD̄LĪT̄L̄PC̄RĪK̄Q̄IV̄NM̄W̄Q̄VK̄K̄AM̄YĀP̄PĪSḠQ̄IR̄C̄SS̄NĪT̄GL̄LL̄IRD̄ḠGS̄N̄NT̄ET̄F̄  
R̄P̄GḠGN̄MK̄DN̄WR̄SĒLȲKȲKV̄VK̄IĒPL̄GV̄AP̄TR̄AK̄RR̄V̄V̄Q̄RĒK̄R̄AV̄ḠIḠAL̄IF̄ḠFL̄GĀAḠST̄M̄GĀAS̄MT̄L̄TV̄Q̄AR̄QL̄LS̄GĪV̄Q̄Q̄NN̄LL̄RAĪĒ  
AQ̄QH̄ML̄QL̄TV̄WḠIK̄QL̄Q̄AR̄VL̄ĀVER̄YL̄KD̄Q̄QL̄GĪWḠCS̄ḠK̄LICT̄T̄VP̄W̄NAS̄WS̄N̄KS̄L̄DD̄ĪW̄NN̄MT̄W̄MĒWĒRĒID̄NȲT̄GL̄ĪȲTL̄IĒQS̄ON̄Q̄  
Q̄ER̄NĒQ̄EL̄LĒLD̄KW̄AS̄L̄WN̄WF̄NĪTN̄WL̄WȲIK̄IF̄IMĪIḠGL̄IḠL̄R̄IVĒFĀVLS̄ĪNR̄VR̄K̄ḠYS̄PL̄SF̄Q̄TL̄TH̄HQ̄RĒP̄DR̄P̄ḠRĪĒĒEḠḠEQ̄ḠK̄DR̄  
S̄IR̄LV̄SḠFL̄AL̄AW̄DD̄L̄RS̄L̄CL̄FS̄ȲH̄RL̄R̄DF̄IL̄IĀART̄VEL̄L̄GR̄SS̄IK̄GL̄RL̄IḠWĒGL̄K̄YL̄WN̄LL̄LȲWḠRĒL̄K̄NS̄AIN̄LL̄DT̄VAĪAV̄AN̄WT̄DRĀ  
IĒV̄V̄Q̄RV̄GRĀVL̄N̄IP̄VR̄IR̄Q̄GL̄ER̄ALL̄S̄

Fig. 61B

2003 con 12 BF Env. seq. opt  
ATGGCGGTGGCGGATGCAGCGCAACTGGCAGCACCTGGCAAGTGGGCCCTGCTGTTCCTGGGCATCCTGATCATCTGCAACGCCACCGA  
GAACCTGTGGGTGACCGTGTACTAGCGGTGCCCGTGTGGAAGGAGGCCACCAACCCCTGTGTGCGCTCCGACGCCAAAGTCTCTACGAGC  
GCGAGGTGCACAACAGTGTGGCCACCCACCGCTGCGTGGCCACCGACCCCAACCCCAAGGAGGTGGACCTGGAGAACGTTGACCCGAGAACTTC  
GACATGTGGAAGAACAACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCC  
CCTGTGCGTGAACCTGAACCTGACCGCAACCGCCAAAGGAGCACCCGAGGCCCTGTCTACCCGCTGGACATCGTGGACCCCATCCAGAACT  
GCTCCTTCAACATGACCAACCGAGGTGGCGGACAAAGCAGATGAAGGTGACGGCCCTGTCTACCCGCTGGACATCGTGGACCCCATCCCACTA  
AACTCCAACGAGTACCGCTGATCAACTGCAACACTCCACCATCACCCAGGCTGCCCAAGTGTCTGGACCCCATCCCATCCACTA  
CTGGCCCCCGGCTACGCCATCTGAAGTGAACGACAAGAAGTCAACGGCACCCGGCCCTGCAAGAAGAGTGTCCACCGTGCAGTGA  
CCCACGGCATCAAGCCCGTGTCCACCCAGCTGTGTAACGGTCCCTGSCCGAGGAGGATCATCATCCGCTCCAGAACATCTCC  
GACAACGCCAAGACCATCATCGTGCACTGAACGAGTCCGTGCAGATCAACTGCACCCGCCCAACAACAACCCGCAAGTCCATCCCAT  
CGCCCCGGCGGCTTCTACGCCACCGCGGACATCATCGGGACATCCGCAAGGCCACTGCAACGTTCCGGCACCCAGTGAACAAGA  
CCCTGGAGCAGGTGAAGAAGACTCGCTCTACTTCAACACCACTCAAGTCAACTCCTCCGCGGGGACCCCGAGATCACCATG  
CACTCCTCAACTGCCCGGGAGTCTTCTACTGCAACACTCCAACTGTTCAACGACACCGTGTCCAAACGACACCATCATCTGCCCCG  
CCGATCAAGCAGATCGTGAACATGTGGCAGGAGTGGCCGCGCATGTACCGCCGCCCCATCGCCGGCAACATCACCTGCACCTCCAACA  
TCACCGCCTGCTGTGACCCCGGACCGCGGCAACGAGACCAACAGACCCGAGACCTTCCGCCCGGGGGCAACATGAAGGACAAC  
TGGCGTCCGAGTGTACAAGTACAAGTGAAGTGGAGATCGAGCCCTGGCGTGGCCCGCCACCCGCGCAAGCCAGGTGGTGAAGCGCGA  
GAAGCGCCCGTGGCATCGGCGCTTCTGGGCTTCTGGCGCCCGCCCTCCACCATGGCGCCGCTCCATCACCTGACCGTGC  
AGGCCCGCAGTGTCTGGCATCGTGCAGCAGTCCAACTGTGGCGCCATCGAGGCCAGCAGCACTGTGCAGTGCAGTGCAGTGC  
TGGGGCATCAAGCAGTGCAGGCCCGCTGCTGGCCGTGGAGCGTACTGAAGGACCCAGCAGTGTGGCCCTGTGGGCTGCTCCGGCAA  
GCTGATCTGCACCAACCTGCAACTCTCCTCTGGTCCAAAGTCCAGGAGGATCTGGGAGAACATGACCTGGATGGAGTGG  
AGAAGGATCAACAATACTCAACGAGATCTACCGCTGATCGAGGATCCAGAACCCAGCAGGAGAGAACAGCAGGAGTGTGGCC  
CTGGACAAATGGCCCTCCCTGTGGAATCGATCGACATCTCAACTGGCTGTGTACTCCGCTACTCCCTGTCCCTGACAGCCCATCCCTCC  
CCCGAGCCCGACCCCGAGGGCATCGAGGAGGGCGGGCAGGCAAGGACCGTCCGTGGCCCTGGTGAACGGCTTCTCTGGCC  
CTGATCTGGACGACCTGCGCTCCCTGTGCTTCTCTACCCGCTGCGGACCTGTGTGTGATCTGTGACCCGATCGTGGAGTGT  
GGCCCGCGGCTGGAGGTGCTGAAGTACTGTGGAACCTGTGCACTGCTCCAGGAGTGAAGAATCCGCCATCTCCCTGTGA  
ACACCAACCGCATCTGTGGTGGCCGAGGCCACCGACCGCTGATCGAGGCCCTGCGAGCGGCTGGCCGCGCCATCTGAAACATCCCGCGCGC  
ATCCGCCAGGGCTGGAGCGGCCCTGTGTAA

102/178

Fig. 62B

**2003 CON 14 BG Env. seq. opt**  
 ATGAAGCCAAAGGCAACCAGCGCAATGGCAAGTGGGGCAACCTGATCCTGGGCCCTGGTGATCATCTGCTCCGCCCTCCAA  
 CGACCTGTGGGTGACCGGTACTACGGCGTGGAAAGGAGGCCACCAACACCTGTCTGGCCTCCGACGCCAAGGCTACGACG  
 CCGAGGTGCACAACGTGTGGCCACCCACCGCTGCGTCCCAACCCGACCCCAACCCAGGAGGTGGCCCTGGAGAACGTGACCGAGAATT  
 AACATGTGGAGAAACAACATGGTGGACCAGATGACGAGGACATCATCTCCTGTGGACCAGTCCCTGAAGCCCTGCTGGAGCTGACCC  
 CCTGTGCTGACCCCTGAACCTGCACCGACTTCAACAACACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  
 GCTCCTTCAACATCACCACTCCCTGGCGGACAAAGATCAAGAAAGGAGTACGCCCTGTCTACAACCTGGACGTGGTGCAGATGGACAACGAC  
 AACTCCTTACCGCTGACCTCCTGCAACACCTCCATCACCCAGGCTGCCCAAGGTGTCTTCAACCCATCCCCATCCACTACTG  
 CGCCCCCGCGCTTGGTGTCCACCCAGTGAAGGACCCCATCGAGATCAACTGCAACCCGCCCCCAACAACAACCCGCAAGGCATCACCATGGG  
 AACGCCAAGACCATCATCGTGCAGCTGAAGGACCCCATCGAGATCAACTGCAACCCGCCCCCAACAACAACCCGCAAGGCATCACCATGGG  
 CCGCGCGCTGTGTACACCCAGGACAGATCATCGGCGACATCCGCAAGGCCACTGCAACATCTCCAAGACCAAGTGGAAACAACCC  
 TGGCCAGATCGTGAAGAAGTGGCGAGCAGTTTCTACTGCAACACCCAGCTGTCAACTCCACTGGCGTCCAACTCCACTGGAACGACAC  
 CACTCCTTCAACTGCGGCGGAGTCTTCTACTGCAACACCCAGCTGTCAACTCCACTGGCGTCCAACTCCACTGGAACGACAC  
 CACGAGACCAACAACACCCAGCTGATACCTGCTCCCAACATCACCGCTGTCAAGCAGATCGTGAACATGTGGCAGAAAGTGGCAAGCCATGTACG  
 CCCCCCATCTCCGGCCAGATCCGCTGCTCCCAACATCACCGCTGTCAAGTCCGCGACGGCGCTCCAAACAACCGGACCTTCC  
 CGCCCCGGCGCAACATGAAGGACAACTGGCGTCCGAGCTGTAAGTACAAGTGGTGAAGATCGAGCCCTCGAGCCCTGGCCCTG  
 CCGGCCAAGCGCGCTGGTGCAGCGGAGAGCGCCGTCAGGATCGGCGCTGTGTTCCGCTTCCCTGGCGCCCGCCGCTCCACCA  
 TGGCGCCGCTCCATGACCTGACCGTGCAGGCGCGGAGAGCGCCGTCAGTGTCCGGCATCGTGCAGCAGCAGAACACTGCTGCCGCCATCGAG  
 GCCAGCAGCAGTGCAGCTGACCGTGTGGGATCAAGCAGTGCAGGCCCGCTGCTGGCCGTGGAGCGCTACCTGAAGGACCAAGCA  
 GTGCTGGGCATCTGGGCTGCTCCGGCAGCTGATCTGCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  
 TCTGGAACAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  
 CAGGAGCGCAACGAGGAGTGTGGAGTGGACAAGTGGCCCTCCCTGTGGAACTGGTTCACCAATCACCAACTGATCGAGCAGTCCCAAGAACCA  
 GATCTTATCATGATCATCGGCGCTGATCGGCTGCGCATCGTGTTCGCCGTGCTGCCATCATCAACCGGTGGCAAGGGTACTCC  
 CCGTGTCTCCAGACCTGACCCACCAACGAGCGGAGCCCGCCCGGCGCATCGAGGAGGAGGGCGGAGCAGGACAAGGACCGG  
 TCCATCCGCTGGTCCGGCTTCCCTGGCCCTGGACGACCTGGCTCCCTGTGCTGTCTCTACCACCGCTGCGGACTTCAT  
 CCTGATCGCCCGCCAGCTGGAGTGTGGCCCTGCTCCCTGAAGGCGCTGGCCCTGGGCTGGAGGCGCTGAAGTACCTGTGGAAAC  
 TGTGCTGTACTGGGCGCGAGTGAAGAACTCCGCCATCAACTGCTGGACACCGTGGCCATCGCCGTGGCCAACTGGACCGGACCGCGC  
 ATCGAGGTGGTGCAGCGGCTGGGCCGCTGTAACATCCCCGTGGGCATCCGCCAGGGCCTGGAGCGGCCCTGCTGTAA

Centralized HIV-1 gag/nef/pol Protein and the Codon-optimized Gene Sequences

Fig. 63A

1. 2003\_CON S gag.PEP  
 MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETSEGCQIIEQLPALQIGSEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEEEQNKSQKTKQAAADTGNSSKVSQNYPIVQNLQGMVHQAI SPRTLNAWKVVEEKAFSPEVIPMFSALSEGATPQDL  
 NMLNTVGGHQAAAMQMLKDTINEEAAEWDRLLHPVHAGPIPPGQWREPRGSDIAGTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSIIDIRQPKPEFRDYVDRFFKTLRAEQATQDVKNWMTDILLVQANPNPDKTILKALPGATLEEMMTACQGVGSPSHKARVLAEAMS  
 QVINTTIIIMQRGNFKGQKRIIKCFNCGKEGHLARNCRAPRKKGCKGKEGHQMKDCTERQANFLKWIWPSNKRPGNLFQSRPEPTAPPAAE  
 SFGFGEIITPSPKQEPKDKELYPLASLKSIFGNDPLISQ\$

Fig. 63B

2003\_CON S gag.OPT  
 ATGGCGC<sup>~</sup>CGCGCCTCCGCTGTGTCCGGCGGCAAGCTGGACCGCCTGGGAGAAGATCCGCCCTGGCCCCCGGGCGGCAAGAAGTACCGCCT  
 GAGCACCTGGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCTCCGAGGGCTGCCAGAGATCATCG  
 AGCAGCTGCAGCCCGCTGCAGACCGGCTCCGAGGAGCTGGCTCCCTGTACAACACCGTGGCCACCTGTACTGCGTGCCACCGAGCGCATC  
 GAGGTGAAGGACACCAAGGAGCCCTGGACAAGATCGAGGAGGAGCAGAACAGTCCAAGCAGAGAACCCAGAGCCCGCCCGCCGACACCGG  
 CAACTCCTCAAGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCAACAGGCCATCTCCCCCGCACCTGAAACG  
 CCTGGGTGAAGTGGTGGAGGAGGCCCTTC<sup>~</sup>CCCCAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC<sup>~</sup>CCAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACAGGCCCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGGAGTGGACCGCCT  
 GCACCCCGTGACGCGGCCCATCCCCCGCCAGATGCGCGAGCCCGGCTCCGACATCGCCGACCACTCCACCTGCGAGGAGC  
 AGATCGGCTGGATGACCTCAACCCCTCCCGTGGCGGAGATCTACAAGCGCTGGATCATCCTGGGCTGAAACAAGATCGTGGCGCATG  
 TACTCCCCGTTCATCCTGGACATCCGCCAGGGCCCAAGGACCTTCCGGACTACGTGGACCGCTTCTTCAAGACCTGCGGCGCGGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCTTGTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGAAGGCCCTGG  
 GCCCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCTGGCGGCCCTCCCAAGGCCCGCTGCTGGCCGAGGCCATGTCC  
 CAGGTGACCAACACCATCATGATGCAGCGCGCAACTTCAAGGCCAGAACGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCCA  
 CATCGCCCGCAACTGCGCGGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGACTGCAACCGAGCGCCAGG  
 CCAACTTCTGGGCAAGATCTGGCCCTCCAACAAGGCCCGCCCGCAACTTCTGTGAGTCCCGCCCGAGCCCAACCGCCCGCCCGCGGAG  
 TCCTTCGGCTTCGGCGGAGGAGATCACCCCTCCCAAGCAGGAGGCCCAAGGACAAAGGAGCTGTACCC<sup>~</sup>CCCTGGCCTCCCTGAAAGTCCCCTGTT  
 CGGCAACGACCCCTGTCCCAGTAA

Fig. 64A

2. 2003 M. GROUP .anc gag .PEP  
 MGARASVLSGGKLD AWEKIRLRPGKKKYRLKHLVWASRELERFALNPGLLETAEGCQQIMGQLQPALQGTGTEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEEEQNKSQKTQAAADKGDSSQVSNYPIVQNLQGMVHQAI SPRILNAWVKVVEEKAFSPEVI PMFSALSEGATPQDL  
 NTMLNTVGGHQAMQLKDTINEEA EWDRLHPVHAGPIPPQOMREPRGSDIAGTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSIIDIRQGPKEPFDRDYDRFFKTLRAEQATQDVKNWMTDLLVQANPDKTILKALPGGATLEEMMTACQGVGGPGHKARVLAEAMS  
 QVTNANIMMQRGNFKPRRIVKCFNCGKEGHIARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWPSNKGKRPNGFLQSRPEPTAPPAAE  
 SFGFGEIITPSPKQEPKDKELYPLASLKS LFGSDPLSQ\$

Fig. 64B

2003 M. GROUP .anc gag .OPT  
 ATGGCGCCCGCCCTCCGTGCTCCGGGGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGGCCCGGGCAAGAAGTACCGCCCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTGCCCTGAACCCGGCTGTGGAGACCGCCGAGGCTGCCAGCAGATCATGG  
 GCCAGTGCAGCCCGCCCTGCAGACCGGCAACCGAGAGCTGCCCTCCCTGTACAACACCGTGGCCACCCCTGTACTGCGTGCAACCGCATC  
 GAGGTGAAGGACACCAAGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAGCAGAAGACCCAGCAGGCCCGCCGACCAAGG  
 CGACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAAGGCCATCTCCCCCGCACCCCTGAACG  
 CCTGGTGAAGGTGGTGGAGAGAAAGGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGGACCCCTG  
 AACACCATGCTGAACACCGTGGCGGCCACACGCGCCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCCGAGTGGGACCCGCTG  
 GCACCCCGTGCACGCCGCCCATCCCCCGCCAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACCCCTCCACCCCTGCAGGAGC  
 AGATCGGCTGGATGACCTCCAACCCCATCCCGTGGCGAGATCAACAGCGCTGGATCATCTGGGCCCTGAACAAGATCGTGGCATG  
 TACTCCCCGTGCTCCATCCTGGACATCCGACATCCGACAGGCCCCAAGGACCCCTCCCGGACTACGFGGACCGCTTCTTCAAGACCCCTGCCGCCGA  
 GCAGGCCACCCAGGACGTGAGAACTGGATGACCGACACCCCTGTGGTGCAAGAACGCCAACCCCGACTGCAAGACCATCTGAAGCCCTGG  
 GCCCGGCCACCCCTGGAGGAGATGATGACCCCTGCCAGGGCTGGCGGCCCGCCGACAAAGGCCCGTGTGCCGAGGCCCATCTGAAGCCCTGG  
 CAGTGAACCAACGCCAACATCATGATGCAGCGCGCAACTTCAAGGCCCGCCCGCCGATCGTGAAGTCTTCAACTGCGGCAAGGAGGCCA  
 CATCGCCGCAACTGCCCGCCCGCCCGCAAGAGGCTGTGGAAGTGGGCAAGGAGGCCACAGATGAAGACTGCACCGAGCCGACAGG  
 CCAACTTCTGGCAAGATCTGGCCCTCCAACAAGGCCCGCCCGCAACTTCTTGCAGTCCCGCCCGGACCCCGCCCGCCCGCCGAG  
 TCCTTCGGCTTCGGCGAGGAGATCACCCCTCCCCCAAGCAGGAGCCCAAGGACTGTACCCCTGGCCCTCCCTGAAAGTCCCTGTT  
 CGGCTCCGACCCCTGTCCCAGTAA

104/178

Fig. 65A

3. 2003 CON A1 gag.PEP  
 MGARASVLSGGKLDAWEKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETTEGCQQIMEQLQPALKTGTEELRSLYNTVATLYCVHQRI  
 DVKDTKEALDKIEEIQNKSKQKTQAAADTGNSSKVSQNYPIVQNAQGMVHQSLSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL  
 NMMNLNVGGHQAAOMLKDITNEEAAEWRDLHPVHAGPIPPGQMREPRGSDIAGTTSTPQEQIGWMTGNPPIPVGDIYKRWIILGLNKIVRM  
 YSPVSILDIKQGPKEFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQANANPDCKSILRALPGGATLEEMMTACQVGGPQGHKARVLAEAMS  
 QVQHTNIMMQRGNFRGQKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSKGRPGNFPQSRPEPTAPPAEI  
 FGMGEEITSPPKQEQKDREQDPLVSLKSLFGNDPLISQ\$

Fig. 65B

3. 2003 CON A1 gag.OPT  
 ATGGCGCCCGCGCCCTCCGTGCTGTCCGGGGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGCGCCCGGGGCAAGAAGTACCGCCT  
 GAAGCACTGCTGTGGCCCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCACCGAGGCTGCCAGAGATCATGG  
 AGCAGCTGACGCCCGCTGAAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCCGTGGCCACCCCTGTACTGCGTGCACAGCGCATC  
 GACGTGAAGGACACCAAGGAGGCGCTGGACAAGATCGAGGAGATCCAGAACAAAGTCCAAGCAGAACCCAGAGGCCCGCCGACACCGG  
 CAACTCCTCCAAGGTGCCAGAACTACCCCATCGTGCAGAACGCCAGGCCAGATGGTGCAACAGTCCCTGTCCCGCCGACCCCTGAACG  
 CCTGGTGAAGTGATCGAGGAGAAGGCTTCTCCCGAGGTGATCCCATGTCTCCCGCCCTGTCCGAGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGGCCGCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGCCGCGAGTGGGACCGCCT  
 GCACCCCGTGCACGCCGCGCCATCCCGCCGCGCAGATGCGCGAGCCCGCGCTCCGACATCGCCGGCACACCTCCACCCCGCAGGAGC  
 AGATCGGCTGGATGACCGCAACCCCGCATCCCGTGGCGACATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCGTGTCCATCCTGGACATCAAGCAGGCCCCAAGGAGCCCTTCCCGGACTACGTGGACCGCTTCTTCAAGACCCCTGCGGCGCGA  
 GCAGGCCACCCAGGAGTGAAGAACTGGATGACCCGAGACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGGCGCCTGG  
 GCCCGGCGCCACCCCTGGAGGAGATGATGACCGCTGCCAGGCGGTGGCGGCCCCCGCCACAAGGCCCGGCTGTGCTGGCCGAGCCATGTCC  
 CAGGTGCAGCACACCAACATCATGATGCAGCGCGGCACTTCCCGGCGCAGAAAGCGCATCAAAGTCTTCAACTGCGGCAAGGAGGCCACCT  
 GGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGTGTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGGACTGCAACCGAGCGCCAGGCCA  
 ACTTCTGGGCAAGATCTGGCCCTCTTCCAAGGCGCGCCCGCAACTTCCCGCAGTCCCGCCCGGAGCCACCGCCCCCGCCGAGATC  
 TTCGGCATGGCGAGGAGATCACCTCCCGCCCGCAAGCAGGAGCAGAACCCCGGAGCAGGAGCCCGCCCTGTGTCCCTGAAAGTCCCTGTT  
 CGGCAACGACCCCTGTCCAGTAA

Fig. 65C

4. 2003 A1.anc gag. PEP

MGARASVLSGGKLDLDAWEKIRLRPGGKKKYLKHLVWASRELERFALNPGLETAEGCQQIMGQLQPALKKTGTEELRSLYNTVATLYCVHQRI  
EVKDTKEALDKIEEIQNKSKQKTQAAADTGNSSKVSQNYPIVQNAQGQMVHQSLSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL  
NMMLNIVGGHQAAMQLKDTINEEAAEWDRLHPVHAGPIPPGQMPREPRGSDIAGTSTLQEQIGWMTGNPPIPVGDIYKRWIILGLNKIVRM  
YSPVSLDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQANANPDCKSILRALPGATLEEMMTACQGVGGPGHKARVLAEAMS  
VQONTDIMMQRGNFRGPKRIKCFNCGKEGHLARNCRAPRKKGCWKCKGEGHMKDCTERQANFLGIKWIWSSKGRFNGFPQSRPEPTAPPAEN  
FGMGEEMISSPKQEQKDRREQYPLVSLKSLFGNDPLSQ\$

Fig. 65D

2003 A1.anc gag. OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGGGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGCGCCCGGGGCAAGAAGTACCGCCT  
GAAGCACCTGGTGTGGCCTCCCGGAGCTGGAGCGCTTCCGCTGAACCCCGGCTGTGGAGACCGCCGAGGCTGCCAGCAGATCATGG  
GCCAGCTCAGCCCGCCTGAAGACCGGACCGAGCTGCCCTCCCTGTACAACACCCGTGGCCACCCCTGTACTGCGTGCACCCAGCGCATC  
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAAGTCCAAGCAGAAGACCCAGCAGGCCCGCCGACACCCGG  
CAACTCCTCAAGGTGCCAGAACTACCCCATCGTGCAAGACGCCAGGCCAGATGGTGCACCAAGTCCCTGTCCCGCCGACCCCTGAACG  
CCTGGTGAAGGTGATCGAGGAGAAGCCTTCTCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGCGCCACCCCCAGGACCTG  
AACATGATGCTGAACATCGTGGGGGCCACCGAGCCGATGCAGATGCTGAAGGACACCATCAACAGGAGGCCCGCAGTGGGACCGCCT  
GCACCCGTGCACGCGGCCCATCCCGCCGAGATGGCGAGCCCGCGGCTCCGACATCGCCGGCACCACTCCACCCCTGCAGGAGC  
AGATCGGCTGGATGACCGGCAACCCCATCCCGTGGCGACATCTACAAGCGTGGATCATCTGGGCTGAACAAGATCGTGGCGCATG  
TACTCCCGTGTCCATCCTGGACATCCGCAAGGCCCCAAGGAGCCCTTCCCGACTACGTGGACCGCTTCAAGACCCCTGCGGCGCGA  
GCAGGCCACCCAGGAGTGAAGAACTGGATGACCCGAGACCCCTGCTGTGCAAGAACGCCAACCCGACTGAAAGTCCATCTGCGGCGCCTGG  
GCCCGGCGCCACCCCTGGAGGAGATGATGACCGCCTGCCAGGCGTGGCGGCCCGCCGCAAGGCCCGCGTGTGGCCGAGGCCATGTCC  
CAGGTGCAGAACACCGACATCATGATGCAGCGCGGCAACTTCCCGGCCCCCAAGCGCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACCT  
GGCCCGCAACTGCCGCGCCCGCCGCAAGAGGCTGTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGGACTGCACCGAGCGCCAGGCCA  
ACTTCTGGGCAAGATCTGGCCCTCTTCCAAAGGCGCCCGCAACTTCCCGCAGTCCCGCCCGGAGCCACCGCCCGCCCGCCGAGAAC  
TTCGGCATGGCGGAGGAGATGATCTCTCTCCCAAGCAGGAGCAGAGGACCGGAGCAGTACCCCGCCCTGGTGTCCCTGAAGTCCCTGT  
CGGCAACGACCCCTGTCTCCAGTAA

107/178

Fig. 66A

**5. 2003 CON A2 gag . PEP**

MGARASILSGGKLDWEKIRLRPGGKKYRLKHLVWASRELEKFSINPSLLETSEGCRRQIIRQLQPALQGTTEELKSLYNTVAVLYCVHQRI  
 DVKDTKEALDKIEEEQNCKKQKTQHAADTGNSSSSQNYPIVQNAQGMVHQAI SPRTLNAWKVVEKAFSPEVIPMFTALSEGATPQDL  
 NTMLNTVGGHQAAQMQLKDTINEEAAEWRDLHPVHAGPIPPGOMREPRGSDIAGTTSLOEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSLDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTDILLVQANPDKSILRALPGATLEEMMTACQGVGGPSHKARVLAFAEMS  
 QVQNTNIMMQRGNFRGQKRILKFCNCGKEGHLARNCRAPRKKKGCWKCKEGHQMKDCTERQANFLGKIWPNSNKGPRGNFPQSRTEPTAPPA  
 ENLRMGEIITSSLKQELKTRPYNPAISLKSFLGNDPLSQ\$

Fig. 66B

**2003 CON A2 gag . OPT**

ATGGGCGCCGGCCCTCCATCCTGTCCGGGGCAAGCTGGACGCCCTGGAGAAAGATCCGCCCTGCGCCCGGCGGCAAGAAGTACCCGCCCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGAAGTCTCCATCAACCCCTCCCTGTGGAGACCTCCGAGGGCTGCCGCCAGATCATCC  
 GCCAGCTGCAGCCCGCCCTGCAGACCGGACCGGACCGAGGAGCTGAACTCCCTGTACAACACCGTGGCCGTGTACTGCGTGCAACCCAGCGCATC  
 GACGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAAGTCAAGTCAAGCAGAGACCCAGCACCGCCGCGGACACCCGG  
 CAATCCTCCTCCTCCAGAACTACCCATCGTGCAGAACGCCAGGGCCAGATGGTGCACAGGCCATCTCCCCCGCACCCCTGAACG  
 CCTGGGTGAAGTGGTGGAGGAGAGGCTTCTCCCCGAGGTGATCCCATGTTCACCGCCCTGTCCGAGGGCGCCACCCCGGACCCCTGAACG  
 AACACCATGCTGAACACCGTGGCGGCCACAGGCCCATGCAAGTGTGAAGGACACCATCAACGAGGAGGCGCGGACCCCTCCAGGACCTG  
 GCACCCCGTGCACGCGGCCCATCCCCCGGCGAGATGCGCGAGCTCCGACATCGCCGGCACCATCCACCCCTGCAGGAGC  
 AGATCGGCTGGATGACCTCAACCCCATCCCCGCGGAGATCAAGCGCTGGATCATCCTGGCCCTGAACAAGATCGTGGCCATG  
 TACTCCCCGTGCTCCATCCTGGACATCCGCCAGGCCCAAGGAGCCCTTCCCGGACTACGTGGACCCGCTTCTCAAGACCCCTGCCGCCGA  
 GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCCGCCCTGG  
 GCCCGGCGCCACCTGGAGGAGATGATGACCGCCTGCCAGGCGTGGCGGCCCTCCACAAAGGCCCGGTGCTGGCCGAGGCCATGTCC  
 CAGGTGCAGAACACCAACACCAACATCATGATGCAGCGCGGCAACTTCCGCGGCCAGAACCGCATCAAGTGTCAACTGCGGCAAGGAGGG  
 CCACCTGGCCCGCAACTGCCCGCCCCCGCAAGAGGGTGTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGACTGCACCGAGCGCC  
 AGGCCAACTTCCCTGGGCAAGATCTGGCCCTCCAAACAAGGGCCCGCCCGCAACTTCCCCAGTCCCCGACCGAGCCACCGCCCCCGCC  
 GAGAACCTGGCGATGGCGGAGGATCACCTCCTCCCTGAAGCAGGAGCTGAAGACCCCGGAGCCCTACAACCCCGCCATCTCCCTGAAGTC  
 CCTGTTCCGCAACGACCCCTGTCCCAGTAA



## Fig. 67A

## 6. 2003 CON B gag . PEP

MGARASVLSGGELDRWEKIRLRPGGKKKYLKHIWVASRELERFAVNPGLLETSEGCROILGQLPQLSQTGSEELRSLYNTVATLYCVHQRI  
 EVKDTKEALEKIEEENKSKKKAQAAAADTGNSSQVSONYPIVONLQGMVHQAI SPRLLNAWVKVVEEKAFSPVIFMFSALSEGATPODL  
 NMLNTVGGHQAAMQMLKETINEEAAEWDRLLHPVHAGPIAPGQMPREPRGSDIAGTSTLQEQIGWMTNNPPIPVGEIYKRWII LGLNKIIVRM  
 YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANANPDCKTILKALGPAATLEEMMTACOGVGGPGHKARVLAEAMS  
 QVTNSATIMMQRGNFRNQRKTVKFCNCGKEGHIAKNCRAPRKKGWCKGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPE  
 ESFRFGEETTPSQKQEPIDKELYPLAS\$

## Fig. 67B

## 2003 CON B gag . OPT

ATGGCGCCCGCGCCTCCGTGCTGTCCGGCGGAGCTGGACCGCTGGGAGAAAGATCCGCCCTGCGCCCGGCGGCAAGAAGTACAAGCT  
 GAAGCACATCGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCGTGAACCCCGGCTGTGGAGACCTCCGAGGGCTCCGCGCAGATCCTGG  
 GCCAGCTGCAGCCTCCCTGCAGACCGGCTCCGAGGAGCTCGCTCCCTGTACAACACCGTGGCCACCTGTACTGCGTGCACCCAGCGCATC  
 GAGTGAAGGACACCAAGGAGCCTTGGAGAAGATCGAGGAGGAGCAAGTCCAAGAAGAGGCCAGAGGCCCGCCGCGCACACCCG  
 CAACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGCCAGATGTTGTCACAGGCCATCTCCCCCGCACCCCTGACCG  
 CCTGGGTGAAGTGGTGGAGGAGGAGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGCAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACAGGCCGCTGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCCGCCGAGTGGGACCCGCT  
 GCACCCCGTGCACGCCGCCCCATCGCCCCCGGAGATGCGCGGAGCCCGCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCAACAAACCCCGTGGCGGAGATCTACAAGCGTGGATCATCCTGGGCTGAACAAGATCGTGGCGCATG  
 TACTCCCCACCTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCTTCCCGGACTACGTGGACCGCTTCTACAAGACCTTCCGCGCCGA  
 GCAGGCTCCAGGAGGTGAAGAACTGGATGACCCGCTGCCAGGCGTGGCGGCCCGCCACAAAGGCCCGTGTGCGGAGGCCATGTCC  
 GCCCGCGCCACCTGGAGGAGATGATGACCCGCTGCCAGGCGTGGCGGCCCGCCACAAAGGCCCGTGTGCGGAGGCCATGTCC  
 CAGGTGACCAACTCCGCCACCATCATGATGACGCGCGGCAACTTCCGCAACCAAGCCAAAGCCGTGAAGTCTCAACTGCCGGCAAGGAGGG  
 CCACATCGCCAAAGAACTGCCCGCCCCCGCAAGAGGGCTGTGGAAGTGCAGCAAGGAGGCCACCAAGATGAAGGACTGCACCCGAGCGCC  
 AGGCCAACTTCTTGGCAAGATCTGGCCCTCCACAAAGGCCCGCCCGCAACTTCTGTGAGTCCCGCCCGAGCCCAAGCCCGCCCGGAG  
 GAGTCTCTCCGCTTCCGGCGGAGAGACCAACCCCTTCCAGAAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCTGGCCTCCTTAA

108/178

Fig. 67C

7. 2003 B. anc gag. PEP

MGARASVLSGGKLDKWEKIRLRPGKKKYLKHIWASRELERFAVNPGLLETSEGRQILQLPALQGTSEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEEEQNKSKKKAQAAADTGNSSQVSONYPIVQNLQGMVHQAI SPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL  
 NTMLNTVGGHQAAQMLKETINEEAEDWDLHPVHAGPIAFQGMREPRGSDIAGTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM  
 YSPISILDIRQPKPEFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQANPDKTKILLKALPAATLEEMMTACQVGGPGHKARVLAEAMS  
 QVTNSTTIMQRGNFRDQKIVKCFNCGKEGHIARNCRAPRKKGCKWCKEGHQMKDCTERQANFLGKIWPSHKGRPNFLQSRPEPTAPE  
 ESFRFGEETTTPSQQEPIDKELYPLASLKSFLGNDPSSQ\$

Fig. 67D

2003 B. anc gag. OPT

ATGGCGCCCGCGCCTCCGTGTCCGGCGCAAGCTGGACAAAGTGGGAGAAGATCCGCCCTGGCCCGCGGCAAGAAGTACAAGCT  
 GAAGCACATCGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCGTGAACCCCGCCTGTGGAGACCTCCGAGGGCTGCCGACAGATCCTGG  
 GCCAGCTGACGCCCGCCTGCAGACCGGCTCCGAGGAGCTGCCCTCCCTGTAAACACCGTGGCCACCTGTACTGCGTGCACACGCGCATC  
 GAGGTGAAGGACACCAAGGAGCCCTGGACAAGATCGAGGAGGAGCAAGTCCAAGAAGAAGCCAGCAGGCCGCGCCGACACCCGG  
 CAATCCTCCAGGTGTCCAGAACTACCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCCACAGGCCATCTCCCGCCACCCCTGAACG  
 CCTGGTGAAGTGGTGGAGGAGAGGCTTCTCCCGAGGTGATCCCATGTCTCCCGCCTGTCCGAGGGCCACCCCGAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACAGGCCGCCATGCAGATGCTGAAGGAGACCATCAACAGGAGGCCCGCGAGTGGACCCCTG  
 GCACCCCGTGCACGCCGCGCCCATCGCCCGGCGAGATCGCGGCTCCGACATCGCCGACATCGCCGACCACTCCACCCCTGCAGGAG  
 AGATCGGCTGGATGACCAACACCCCGGCGAGATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCATG  
 TAGTCCCGCATCTCCATCTGGACATCCGCAAGGCCCTCCGCGACTACGTGGACCGCTTCTACAAGACCTGCGCGCCGA  
 GCAGGCTCCAGGACGTGAAGACTGGATGACCGGACCCCTGTGTCAGAACGCCAACCCCGACTCAAGACCATCTGAAGGCCCTGG  
 GCCCGCCGACCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGCGGCCGCCAAGGCCCGCGTGGCCGAGGCCATGTCC  
 CAGGTGACCAACTCCACCACTCAIATGATGACGCGCGCAACTCCGCGACAGCGCAAGATCGTGAAGTCTCAACTGCGGCAAGGAGGG  
 CCACATGCGCCGCAACTGCCGCGCCCGCAAGAGGGTGTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGGACTGCACCGAGCGCC  
 AGGCCAACTTCTGGGCAAGATCTGGCCCTCCACAGGGCCCGCCCGCAACTTCTGCAATCCCGCCCGAGCCACCGCCCGCCGAG  
 GAGTCTTCCGCTTCGGCGGAGGACCAACCCCTCCAGAGGAGGCCCATCGACAAGGAGCTGTACCCCTGGCCCTCCCTGAAGTC  
 CCTGTTCGGCAACGACCCCTCCTCCAGTAA

Fig. 68A

8. 2003 CON C gag . PEP  
 MGARASILRGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLLETSEGCKQIIKQLQPALQGTTEELRSLYNTVATLYCVHEKI  
 EVRDTKEALDKIEEONKSQQKTQAKAADGKVSQNYPIVQNLQGMVHQAISPRILNAWVKVIEEKAFSPVIMFTALSEGATPQDLNTM  
 LNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIAPGQMRPRGSDIAGTSTLQEQIAWMTSNPPIPVGDIYKRWIIILGLNKIVRMYSP  
 VSILDIKQGPKEFRDYDRFFKTLRAEQATQDVKNWMTDLLVQANANPDCKTILRALPGATLEEMMTACQGVGPGPSHKARVLAEAMSQAN  
 NTNIMMQRSNFKPKRIVKCFNCGKEGHIARNCRAPRKKGCWKCKEGHQMKDCTERQANFLGKIWPSHKGRPGNFIQNRPEPTAPPAESFR  
 FEETTPAPKQEPKDREPLTSLKSLFGSDPLSQ\$

Fig. 68B

2003 CON C gag . OPT  
 ATGGCGCCCGCCCTCCATCCTGTGGCGGGCAAGCTGGACAAGTGGGAGAAGATCCGCCCTGGCCCCGGGCAAGAAGCCACTACATGCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGCTGCAAGCAGATCATCA  
 AGCAGCTGACGCCCGCCCTGCAGACCGGACCCGAGGAGCTGCGTCCCTGTACAACACCGTGGCCACCTGTACTGCGTGCACGAGAAGATC  
 GAGTGGCGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGACAAGTCCAGCAGAAGACCCAGCAGGCCAAGGCCCGCCGACCG  
 CAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAAGGCCATTCGCCCGACCCCTGAACGCCCTGGGTGA  
 AGGTGATCGAGGAGAAGGCCCTTCGCCCGAGGTGATCCCATGTTCACCCGCTGTCCGAGGGCCACCCCCAGGACCTGAACACCCATG  
 CTGAACACCGTGGCGGCCACCAAGGCCCATGTCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGACCGCTGCACCCCGT  
 GCACGGCCGCCCATCGCCCGGAGATGCGCGGACATCGCGGCTCCGACATCGCCGGCACCATCCACCTGCAGGAGCAGATCGCCT  
 GGATGACCTCCAACCCCTTCCCGGACTACAGCGCTGGATCATCCTGGCCCTGAACAAGATCGTGCATGTACTCCCC  
 GTGTCCATCCTGGACATCAAGCAGGCCCCAAGGACCTCCCGGACTACGTGGACCGTTCCTCAAGACCTGCGCGCCGAGCAGGCCAC  
 CCAGGACGTGAAGAACTGGATGACCGCTCCAGGCTGGCGGCCCTCCCAAGCCCGACTGCAAGACCATCTGCGCGCCCTGGGCCCAAC  
 CCACCTGGAGGAGATGACCGCTCCAGGCTGGCGGCCCTCCCAAGCCCGACTGCAAGACCATCTGCGCGCCCTGGGCCCAAC  
 AACACCAACATCATGATGACGCTCCAACCTCAAGGGCCCAAGCGCATCGTGAAGTGTCAACTGCGGCAAGGAGGCCACATCGCCCG  
 CAACTGCCGCGCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGGACTGCACCGAGGCCAGGCCAACTCC  
 TGGCAAGATCTGGCCCTCCCAAGGCGCCCGCAACTTCCTGCAGAACCGCCCGAGCCACCGCCCGCGGAGTCCCTCCCG  
 TTCGAGGAGACCAACCCCGCCCAAGCAGGAGCCCAAGGACCGGAGCCCTGACCTCCCTGAAGTCCCTGTTCCGGCTCCGACCCCTGTC

## Fig. 68C

9. 2003 C. *anc. gag*. PEP

MGARASILRGKLDITWEKIRLRPGGKHYMIKHLVWASRELERFALNPGLLETSEGCKQIMKQLPALQITGTEELRSLYNTVATLYCVHERI  
 EVRDTKEALDKIEEEQNKSQKTQAEAADGNDGKVSQNYPIVONLQGMVHQAI SPRLLNAWKVVEKAFSP EVI PMFTALSEGATPQDL  
 NTMLNTVGHQAAMQMLKDTINEEAAEWDRLLHPVHAGPVAPGOMREPRGSDIAGTTSFLOEQIAWMTSNPPI PVGDIYKRWI ILLGNKIVRM  
 YSPVSIIDIKQGPKEFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNPANPDCKTILRALGPGATLEFMMTACQGVGGPGHKARVIAFAMS  
 QANNTNIMQRSNFKGPKRIVKCFNCGKEGHLARNCRAPRKKGCKGKEGHQMKDCTERQANFLGIWPSHKGRPGNFIQSRPEPTAPPAE  
 SFRFEETTPAPKQEPKDRPLETSLKSLFGSDPLSQ\$

## Fig. 68D

2003 C. *anc. gag*. OPT

ATGGCGCCCGCGCTCCATCCTGCGGGGCAAGCTGGACACCTGGGAGAAAGATCCGCCCTGCGCCCCGGGGCAAGAAGCACTACATGAT  
 CAAGCACCTGGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGECTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATGA  
 AGCAGCTGCAGCCCGCTGCAGACCGGACCGAGGAGCTCGCTCCCTGTACAACACCGTGGCCACCTGTACTGCGTGCACGAGCGCATC  
 GAGGTGCGGACACCAAGGAGCCCTGGACAAGATCGAGGAGGACAAAGTCCCAGCAGAGAGACCCAGCAGGGCCGAGGCCCGCCGACGG  
 CGAACCGCAAGGTGTCCAGAACTACCCCATCGTGCAGAACTCGAGGGCCAGATGGTGCACCAGGCCATCTCCCCCGCACCCCTGAAACG  
 CCTGGGTGAAGGTGGTGGAGGAGAGGCTTCTCCCCGAGGTGATCCCCATGTTTCCCGCCCTGTCGAGGGCCCAACCCCAAGGACCTG  
 AACACCATGCTGAACACCCGTGGCGGCCACAGGCCCATGAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCCGAGTGGACCCGCT  
 GCACCCCGTGCACGCCCGCTGGCCCGCCCGGAGCCCGGCTCCGACATCGCCGGCACCATCCACCCCTGCAGGAGC  
 AGATCGCCTGGATGACCTCCAACCCCATCCCGTGGCGACATCTACAAGCGCTGGATCATCCTGGGCCCTGAACAAGATCGTCCGATG  
 TACTCCCCGTTCATCCTGGACATCAAGCAGGGCCCAAGGAGCCCTTCGGCGACTACGTGGACCGCTTCTTCAAGACCTTGGCGCCGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACCCCTGTGTGCAGAACGCCAACCCCGACTGCAAGACCATCTTGGCGCCCTGG  
 GCCCGGCCACCCCTGGAGGAGATGATGACCGCTGCCAGGGCTGCCAAGGCGCCCAAGGCCCGCTGCTGGCCGAGGCCATGTCC  
 CAGGCCAACAAACACATCATGATGCAGCGCTCCAACCTTCAAGGGCCCAAGCGCATCGTGAAGTCTCAACTGCGGCAAGGAGGCCA  
 CATCGCCGCAACTGCCCGGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGGACTGCACCGAGGCCAGG  
 CCAACTTCTGGGCAAGATCTGGCCCTCCCAAGGGCCCGCCGCAACTTCTGCAAGTCCCGCCGAGCCACCGCCCGCCCGCCGAG  
 TCCTTCCGCTTCGAGGAGACCAACCCCGCCCGCCCAAGGAGCCCGGAGCCCTGACCTCCCTGAAGTCCCTGTTCGGCTCCGA  
 CCCCCCTGTCCCAGTAA

111/178

Fig. 69A

10. 2003\_CON D gag .PEP  
 MGARASVLSGGKLD<sup>1</sup>DAWEKIRLRPGGKKYRLKHI<sup>2</sup>VWASRELERFALNPG<sup>3</sup>LLETSEGCKQIIGQLQPAIQ<sup>4</sup>TSEELRSLYNTVATLYCV<sup>5</sup>HERI  
 EVKDTKEALEKIEEEQNKS<sup>6</sup>KKKAQQAAA<sup>7</sup>DTGNS<sup>8</sup>SQV<sup>9</sup>SONYPIV<sup>10</sup>QNLQGMVHQ<sup>11</sup>AIS<sup>12</sup>PR<sup>13</sup>TINAWK<sup>14</sup>VEEKAF<sup>15</sup>SPEV<sup>16</sup>IPMFSAL<sup>17</sup>SEGAT<sup>18</sup>PQDL  
 N<sup>19</sup>MLNTV<sup>20</sup>GGHOAA<sup>21</sup>QMLKETINEEAAEW<sup>22</sup>DR<sup>23</sup>LHPV<sup>24</sup>HAGPV<sup>25</sup>APGQ<sup>26</sup>MR<sup>27</sup>PRGSDI<sup>28</sup>AGT<sup>29</sup>STLQEQIG<sup>30</sup>WMT<sup>31</sup>SN<sup>32</sup>PP<sup>33</sup>IPVGEI<sup>34</sup>YKR<sup>35</sup>WILLGLN<sup>36</sup>KI<sup>37</sup>VRM  
 YSPV<sup>38</sup>SILDIR<sup>39</sup>QGPKE<sup>40</sup>PF<sup>41</sup>RDY<sup>42</sup>DRFY<sup>43</sup>KTLRAEQAS<sup>44</sup>QDV<sup>45</sup>KNWMT<sup>46</sup>ETLLV<sup>47</sup>Q<sup>48</sup>NANP<sup>49</sup>DCKTIL<sup>50</sup>KALGPEATLEEM<sup>51</sup>MTAC<sup>52</sup>QGV<sup>53</sup>GGP<sup>54</sup>SHKAR<sup>55</sup>VLA<sup>56</sup>EAMS  
 QATNSAAVM<sup>57</sup>QRGN<sup>58</sup>FK<sup>59</sup>PRKI<sup>60</sup>IKCFNCGKEGHIA<sup>61</sup>KNCRAP<sup>62</sup>RKKG<sup>63</sup>CKG<sup>64</sup>KEGHQ<sup>65</sup>MKDC<sup>66</sup>TERQAN<sup>67</sup>FLGKI<sup>68</sup>WPSHKGR<sup>69</sup>PN<sup>70</sup>FLQSR<sup>71</sup>PEPT<sup>72</sup>APPA  
 ESFGFEEITPSQ<sup>73</sup>KEQK<sup>74</sup>DELY<sup>75</sup>PLTSLKSL<sup>76</sup>FGND<sup>77</sup>PLISQ<sup>78</sup>

Fig. 69B

2003\_CON D gag .OPT  
 ATGGCGCCGCGCCTCCGTG<sup>1</sup>TCCGGCGGCAAG<sup>2</sup>TGGACCGCCTGGGAGAAG<sup>3</sup>ATCCGCCCTGCGCCCCGGCGGCAAGAA<sup>4</sup>AGTACCGCCT  
 GAAGCACATCGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCG  
 GCCAGCTGCAGCCCGCATCCAGACCGGCTCCGAGAGTCCGCTCCCTGTACAAACACCCGTGGCCACCTGTACTGCTGACGAGCGCATC  
 GAGTGAAGGACACCAAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAAGTCCAAGAAGAAAGGCCACAGAGCCCGCCGCGGACACCCGG  
 CAATCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAAGGCCAT<sup>1</sup>TCCCCCGCACCCCTGAAACG  
 CCTGGTGAAGGTGATCGAGGAGAGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC<sup>2</sup>CAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCCACAGCCCGCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCCGCCGAGTGGGACCGCCT  
 GCACCCCGTGCACCGCCGCCCCGTGGCCCCCGGACAGATGCCGAGCCCGGCTCCGACATCGCCGGCACCACTCACCTGCAGGAGC  
 AGATCGGCTGGATGACCTCAACCCCGCATCCCGTGGCCGAGATCAACAAGCGCTGGATCATCTGGCCCTGAACAAGATCGTGGCGCATG  
 TACTCCCCGTGTCATCTGGACATCCGCCAGGCCCCAAGGAGCCCTCCGCCACTACGTGGACCGCTTCTACAAGACCTGCGCGCCGA  
 GCAGGCTCCAGGACGTGAAGAACTGGATGACCGAGACCCCTGCTGGTGCAGAACGCCCAACCCCGACTGCAAGACCATCTGAAGGCCCTGG  
 GCCCGAGGCCACCTGGAGGAGATGATGACCCGCTGCCAGGGCTGGGGGCCCTCCCAAGGCCCGCGTGTGGCCGAGGCCATGTCC  
 CAGCCACCAACTCCGCCCGCGTGTGATGATGACGCGCGCAACTCAAGGGCCCCCGCAAGATCATCAAGTGTCAACTGCGGCAAGGAGGG  
 CCACATCGCCAAGAACTGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCAAGGACCAGATGAAGGACTGCACCCGAGCGCC  
 AGGCCAACTTCTGGGCAAGATCTGGCCCTCCCAAGGGCCGCCCGGCAACTTCTCTGCAGTCCCGCCCGAGCCACCGCCCCCCCCGCC  
 GAGTCTTGGCTTCGGCGAGGAGATCACCCCTCCAGAAAGCAGGAGCAGAAAGGAGTGTACCCCTGACCTCCCTGAAGTCCCT  
 GTTCGGCAACGACCCCTGTCCCAAGTAA

Fig. 70A

11. 2003 CON F gag. PEP  
 MGARASVLSGGKLD<sup>1</sup>DAWEKIRLRPGGKKYRMKHLVWASRELERFALDPG<sup>2</sup>LLETSEGCQKIIIGLQ<sup>3</sup>PSLIQTGSEELRSLYNTVAVLYCVHQK<sup>4</sup>  
 EVKDTKEALEKLEEEQNKSQ<sup>5</sup>QKTQAAADKGVSONYPIVQNLQGMVHQ<sup>6</sup>AISPRTLN<sup>7</sup>AWKVIEEKAFSPEVIMPFSALSEGATPQDINTML<sup>8</sup>  
 NTVGGHQAA<sup>9</sup>MQLKDTINEEAAEWDRLLHPVHAGPIPPGQMREPRGSDIAGTSTLQEQIQWMTSNPPVPVGD<sup>10</sup>IYKRWIIILGLNKIVRMYSPV<sup>11</sup>  
 SILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDLLVQ<sup>12</sup>NANPDCKTILKALGPGATLEEMMTACQGVGPGHKARVLAEAMSQATN<sup>13</sup>  
 TAIMMQKSNFKGQRRIVKFCNCGKEGHLAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSNKGRPGN<sup>14</sup>FLOQRPEPTAPPAESFGF<sup>15</sup>  
 REEITPSPKQEQKDEGLYPPLASLKSIFGNDP\$

Fig. 70B

2003 CON F gag. OPT  
 ATGGCGCCCGGCGCTCCGTGCTGTCGGCGGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGGCCCCGGCGGCAAGAAGTACCCGCAT  
 GAAGCACCTGGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGGACCCCGGCTGCTGGAGACCTCCGAGGGCTGCCAGAAGATCATCG  
 GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTCGCTCCCTGTACAACACCCGTGGCCGTGTACTGCGTGCACCCAGAAGGTG  
 GAGGTGAAGGACACCAAGGAGGCCCTGGAGAAGCTGGAGGAGGAGCAACAAGTCCCAGCAGAAACCCAGCAGGCCCGCCGCGCAAGGG  
 CGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCCAGGCCATCTCCCCCGCACCTGAACGCTGGTGAAGG  
 TGATCGAGGAGAAGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGAGGACCTGAACACCATGCTG  
 AACCCGTGGCGGCCACCAAGGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGCCCGCAGTGGGACCCGCTGCACCCCGTGCA  
 CGCCGGCCCATCCCCCGGCCAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACACCTCCACCTGCAGGAGCAGATCCAGTGGGA  
 TGACCTCAACCCCGTGGCCGACATCTACAAGCGCTGGATCATCTGGGCCCTGAACAAGATCGTGGCATGTACTCCCCCGTGTG  
 TCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGGACTACGTGGACCCGCTTCTTCAAGACCTGCGCCGAGCAGCCACCCA  
 GGAGTGAAGGGCTGGATGACCGACACCCCTGCTGTGAGAACGCCAACCCGACTGCAAGACCATCTGAAGGCCCTGGGCCCGGGCCCA  
 CCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGCGGCCACCAAGGCCCGCTGCTGGCCGAGGCCATGTCCCAGGCCACCAAC  
 ACCGCCATCATGATGCAGAA<sup>1</sup>TCCAACTTCAAGGGCCAGCGCCGCATCGTGAAGTGTCAACTGCGGCAAGGAGGCCACATCGCCAAGAA  
 CTGCCGGCCCCCGAAGAGGGCTGTGGAAGTGGCGCGGAGGGCCACAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCCTGG  
 GCAAAGATCTGGCCCTCCAACAAGGGCCGCCCGGCAACTTCTGTGAGTCCCGCCCGAGCCACCGCCCCCGCGGAGTCCCTTCGGCTTC  
 CGCGAGGAGATCACCCCTCCCCCAAGCAGGAGCAGAGGGCCTGTACCCCCCTGGCCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCTAA

113/178

Fig. 71A

12. 2003 CON G gag. PEP  
 MGARASVLSGGKLD<sup>1</sup>DAWEKIRLRPGGKKYRMKHLVWASRELERFALNPDLLETAEGCQIMQLOPALQ<sup>2</sup>TGTEELRSLFNTVATLYCVHQRI  
 EVKDTKEALEEVEFKIQKKSQK<sup>3</sup>TQQAAMDEGNSSQVSQNYPIVQNAQGMVHQAI<sup>4</sup>SPRTLNAWVKVVEEKAFSP<sup>5</sup>EVI<sup>6</sup>PMFSALSEGATPQDL  
 NTMLNTVGGHQAA<sup>7</sup>MLKDTINEEAAEWDRMHPQ<sup>8</sup>QAGPIPPGQIREPRGSDIAGTSTLQEQIRWMTSNPPI<sup>9</sup>PVGEIYKRWILLGLNKIVRM  
 YSPVSI<sup>10</sup>LDIRQGPKEFRDYVDRFFKTLRAEQATQEVKGMWTD<sup>11</sup>LLVQNPANPDCKTILRALGPGATLEEMMTACQ<sup>12</sup>GVGGPSHKARVLA<sup>13</sup>EAMS  
 QASGAAA<sup>14</sup>IMMQKSNFKGPRRTIKCFNGKEGHLARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKI<sup>15</sup>WPSNKG<sup>16</sup>RPGNFI<sup>17</sup>QNRPEPTAPP  
 AESFGFEEIAPSPKQEQKEKELYPLASLKSIFGSDP\$

Fig. 71B

2003 CON G gag. OPT  
 ATGGCGCCGGCC<sup>1</sup>TCCGTGCTGTCCGGGGCAAGCTGGACGCCTGGGAGAAGATCCGCCCTGGCCCCGGGGCAAGAAGTACCGGCAT  
 GAAGCACCTGGTGTGGCCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCCGACCTGCTGGAGACCCGCGAGGCTGCCAGCAGATCATGG  
 GCCAGCTGCAGCCCGCCTGCAGACCCGAGGAGTGGCTCCCTGTTCAACACCCGTGGCCACCTGTACTGTGTCACCCAGCGCATC  
 GAGGTGAAGCACCAAGGAGCCCTGGAGGAGTGGAGAAGATCCAGAGAAGTCCAGCAGAGACCCAGAGGCCGCCATGGACGAGGG  
 CAATCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCCAGGGCCAGATGGTGCAACAGGCCATTC<sup>2</sup>CCCCCGCACCCCTGAACG  
 CCTGGGTGAGGTGGTGGAGGAAAGCCTTC<sup>3</sup>CCCCAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCC<sup>4</sup>CCAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACAGCCCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGGACCCGCAT  
 GCACCCCAAGCAGCGGCCCATCCCGAGCCAGATCCCGAGCCCGGCTCCGACATCGCCGGCACCACTCCACCCTGCAGGAGC  
 AGATCCGCTGGATGACCTCCAACCCCCCATCCCGTGGCGAGATCTACAAGCGCTGGATCATCCTGGCCCTGAACAAGATCGTGGCATG  
 TACTCCCGGTGTCATCTGGACATCCGCCAGGCCCAAGGAGCCCTTC<sup>5</sup>CGCGACTACGTGGACCCCAACCCCGACTGCAAGACCATCTTCAAGACCTGGCGCCCGA  
 GCAGGCCAACCCAGGAGTGAAGGCTGGATGACCGACACCCCTGCTGGTGCAGAACCCCAACCCCGACTGCAAGACCATCTTGGCGCCCTGG  
 GCCCGGCCCA<sup>6</sup>CCCTGGAGGAGTGTACCCGCTGCCAGGGCTGGCGGCCCTCCCAAGGCCCGCGTGTGGCCGAGGCCATGTCC  
 CAGGCCTCCGGCCCGCCGATCATGATGCAGAACTCAAGGCCCCCGCCACCATCAAGTGTCAACTCGGCAAGGA  
 GGGCACCTGGCCCGCAACTGCCCGCCCGCAAGAGGCTGCTGGAAGTGCAGGCAAGGAGGCCAC<sup>7</sup>CAGATGAAGACTGCACCCGAGC  
 GCCAGGCCAACTTCCCTGGCAAGATCTGGCCCTCCAACAAGGCCCGCCCGCAACTTCTGCAAGAACCGCCCGAGCCACCGCC<sup>8</sup>CCCC  
 GCCGAGTCTTTCGGCTTCGGCGAGGAGATCGCC<sup>9</sup>CCCTCCCCCAAGCAGGAGCAGAAAGGAGGCTGTACCCCTGGCCTCCCTGAAGTC  
 CCTGTTCCGGCTCCGACCCCTAA

Fig. 72A

13. 2003 CON H gag . PEP  
 MGARASVLSGGKIDAWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLLETAEGCLOIIIEQLQPAIKTGTTELQSLFNTVAVLYCVHQRI  
 DVKDTKEALGKIEEIQNKSQKTQOAAADKEKDNKVSQNYPIVQNAOQMVOHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL  
 NAMLNTVGGHQAAMQLKDTINEEAAEWDRLLHPVHAGPIPPGQMRPRGSDIAGTSTLQEQIAWMTGNPPIPVGDIYKRWIIILGLNKIVRM  
 YSPVSLDIKQGPKEPFRDYVDRFETLRAEQATQDVKNWMTDLLVQANANPDCKTILRALGQASIEEMMTACQGVGGP SHKARVLAEAMS  
 QVTNANAAIMQKGNFKGPRKI VKCFNCGKEGHIARNCRAPRKKGCKWCKGREGHQMKDCTERQANFLKGIWPSKGRPGNFIQSRPEPTAPP  
 AESFGFEEMTPSPKQELKDKPEPLASLRSLFGNDPLSQS

Fig. 72B

2003 CON H gag . OPT  
 ATGGGCGCCGCGCCTCCGTGCTCGGTGCTCGGGGCAAGCTGGACGCTGGGAGAAGATCCGCCCTCGGCCCGGGCAAGAAGTACCGCCT  
 GAAGCACCTGGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGCCTGCTGGAGACCCGCGAGGGCTGCCTGCAGATCATCG  
 AGCAGCTGACGCCCGCCATCAAGACCGGACCGGAGCTGCAGTCCCTGTTCAACACCCGTGGCCGTGCTGTAAGTGCACCCAGCCATC  
 GACGTGAAGGACACCAAGGAGGCGCTGGGCAAGATCGAGGAGATCCAGAACAAAGTCCAGCAGAAAGACCCAGCCAGCCGCGCCGACAAAGGA  
 GAAGGACAACAAGGTGTCCAGAACTACCCCATCGTCAGAACGCCCGGCGAGATGTCACCCAGCCATCTCCCCCGCACCTGAAACG  
 CCTGGGTGAAGGTGGTGGAGAGAAGCCTTCTCCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGGACCTGAAACG  
 AACGCCATGCTGAACACCGTGGCGGCCACCCAGCCGCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGCCCGGAGTGGGACCGCCT  
 GCACCCGCTGCACGCCGCCCATCCCCCGGCCAGATGCGGAGCCCGGCTCCGACATCGCCGGCACCATCCACCTGCAGGAGC  
 AGATCGCCTGGATGACCGGCAACCCCATCCCCGTGGCGACATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCATG  
 TACTCCCCGTGTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCCGGACTACGTGGACCCGTTCTTCAAGACCCCTGCGGCCCGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCCTGCTGTGTCAGAACGCCAACCCGACTGCAAGACCATCTGCGGCCCTGG  
 GCCAGGGCCTCCATCGAGGAGATGATGACCGCCTGCCAGGGCTGGCGGCCCTCCACAAAGCCCGGTGCTGGCCGAGGCCATGTCC  
 CAGGTGACCAACGCCAACCGCCATCATGATGCAAGGGCAACTTCAAGGGCCCCCGCAAGATCGTGAAGTGTCAACTGCGGCAAGGA  
 GGGCCACATCGCCCGCAACTGCCCGCCCCCGCAAGAGGCTGCTGGAAGTGGCGCCGAGGGCCACAGATGAAAGGACTGCACCCGAGC  
 GCCAGGCCAACTTCTGGGCAAGATCTGGCCCTCTCCAAAGGGCCCCCGCAACTTCTGCAAGTCCCGCCCGGAGCCACCCGCCCCC  
 GCCGAGTCTTCCGCTTCGGCGAGGAGATGACCCCTTCCCCCAAGCAGGAGCTGAAGGACAAGGAGCCCCCTGGCCTCCCTGGCGTCCCT  
 GTTCCGGCAACGACCCCTGTCCCAGTAA



Fig. 73A

14. 2003 CON K gag . PEP  
 MGARASVLSGGKLDITWEKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETTEGCRQIIRQLQPSLQTSSEELKSLFNTVATLYCVHQRI  
 EVRDTKEALDKLEEEONKSQOKTQOETADKGVSONYPIVONLQGMVHQALSPRTLNAWKVIEEKAFSPEVIPMFSAISEGATPQDLNITML  
 NTVGGHQAAAMQLKDTINEEAAEWDRLLHPVHAGPIPPGOMREPRGSDIAGTTSLQEQITWMTSNPPVPGEIYKRWII LGLNKI VRMYSPI  
 SILDIRQGPKEPFDRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQANANPDCKTILKALPGGASLEEMMTACQGVGPGHKARILAEAMSQVTN  
 TAVMMQRGNFKGQRKI IKCFNCGKEGHIARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWPSNKGPRGNFLOSRPEPTAPPAESFFG  
 GEEITPSRQETKDKEQGPPLTSLKSLFGNDPLSQ\$

Fig. 73B

2003 CON K gag . OPT  
 ATGGCGCCCGCGCTCCGTGTCGCGCGCAAGCTGGACACCTGGGAGAAGATCCGCCCTGCGCCCCGGCGGAAGAAGTACCGCCT  
 GAAGCACCTGGTGTGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCACCGAGGGCTGCCGCCAGATCATCC  
 GCCAGCTGCAGCCCTCCCGAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCCGTGGCCACCCCTGTACTGCGTGCACCCAGCGCATC  
 GAGGTGCGGCACACCAAGGAGCCCTGGACAAGCTGGAGGAGCAGACAAAGTCCAGCAGAAAGACCCAGCAGAGAGACCCGCCGACAAGG  
 CGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGCCAGATGGTGCACCCAGGCCCTGTCCCCCGACCCCTGAACGCCCTGGGTGAAG  
 TGATCGAGGAGAAGGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGAGGACCTGAACACCATGCTG  
 AACCCGTGGCGGCCACAGGCCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGACCCGCTGCACCCCGTGC  
 CGCCGCCCATCCCCCGGCGAGATGCGCGAGCCCGCGCTCCGACATCCCGGACCCACCTCCACCCCTGCAGGAGAGATCACCTGGA  
 TGACCTCCAAACCCCGTGGCGGAGATCTACAAGCGTGGATCATCTGGGCCCTGAACAAGATCGTGGCATGTACTCCCCCGTG  
 TCCATCTGGACATCCGCCAGGCCCAAGGAGCCCTTCCCGACTACGTGGACCGCTTCTTCAAGACCCCTGCGGCCGAGCAGGCCACCCA  
 GGAGGTGAAGAACTGGATGACCGACACCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCGCGCCT  
 CCTGGAGGAGATGATGACCGCTGCGAGGCGTGGCGGCCCGGCCACAAGCCCGCATCTGGCCGAGGCCATGTCCAGGTGACCAAC  
 ACCGCCGTGATGATGACGCGGCAACTTCAAGGCCAGCGCAAGATCATCAAGTCTTCAACTGCGGCAAGGAGGCCACATCGCCCCGCAA  
 CTGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCAAGGAGGCCACCCAGATGAAGGACTGCACCGAGCCAGGCCAACCTTCTCTGG  
 GCAAGATCTGGCCCTCCAACAAGGCCCGCCCGGCAACTTCTGCACTCCCGCCCGAGCCACCGCCCCCGCGGAGTCTTCCGGCTTC  
 GCGGAGGAGATCACCCCTTCCCCCGCCAGGAGACCAAGGACAAAGGAGCAGGGCCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCCCTGTCCCCAGTAA

Fig. 74A

15. 2003 CON 01 AE gag. PEP  
 MGARASVLSGGKLD~~AW~~EKIRLRPGGKKKRYMKHLVWASRELERFALNPGLLLETAEGCQQIIIEQLQSTLKTGSEELKSLFN~~TV~~ATLWCVHQRI  
 EVKDTKEALDKIEEVONKSQOKTQOAAAAGTSSSKVSONYPIVONAQQMVHQPLSPRTLNAWVKVVEEKGFNPEVIMPFSALSEGATPQDL  
 NMMLNIVGHHQAAMQMLKETINEEAAEWD~~RV~~HPVHAGPIPPGQMRPRGSDIAGTSTLQEQIGWMTNNPPIPVGDIYKRWIIILGLNKIVRM  
 YSPVSI~~LD~~IRQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQANANPDCKSILKALGTGATLEEMMTACQGVGGPSSHKARVLAEAMS  
 QAQHANIMMQRGNFKGQKRIKCFNCGKEGHLARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWP~~SN~~KGRPGNFPQSRPEPTAPPAEN  
 WGMGEEITSLPKQEQDKKEHPPPLVLSKLSLFGNDPLSQ\$

Fig. 74B

2003 CON 01 AE gag. OPT  
 ATGGCGCCCGGCGCTCCGTGCTGTCCGGCGGCAAGCTGGACCGCTGGGAGAAGATCCGCCCTGCGCCCGGGCAAGAAGTACCGCAT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGAGGGCTGCCAGCAGATCATCG  
 AGCAGTGCAGTCCACCCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCGTGGCCACCCCTGTGGTGGTGCACCCAGCGCATC  
 GAGTGAAGGACACCAAGGAGCCCTGGACAAAGATCGAGGAGGTGCAGAACAAAGTCCAGCAGAGAGACCCAGCAGGCCCGCCCGGACCCGG  
 CTCTCTCCAAAGGTGTCCAGAACTACCCCATCGTGCAGAACGCCAGGGCCAGATGGTGACACAGCCCTGTCCCCCGCACCTGAACG  
 CCTGGGTGAAGGTGGAGGAGAAGGCTTCAACCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGCAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGGCCGCGATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCCGCCGAGTGGGACCGCGT  
 GCACCCCGTGCACGGCCGCGCCATCCCGCGGCGAGATGCGGCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCAACAACCCCGTGGCGACATCAAAAGCGTGGATCATCTGGGCC~~TA~~GAACAAGATCGTGGCGCATG  
 TACTCCCCGTGTCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCCGGACTACGTGGACCGCTTACAAGACCCCTGCGCGCCGA  
 GCAGGCCACCCAGGAGGTGAAGACTGGATGACCGAGACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGAAGGCCCTGG  
 GCACCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCACAAAGGCCCGCTGCTGGCCGAGGCCATGTCC  
 CAGGCCAAGCACGCCAACAATCATGATGCAGCGCGCAACTCAAGGCCAAGAACCGCATCAAGTCTCAACTGCGGCAAGGAGGCCACCT  
 GGCCCGCAACTGCCCGGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCAAGGAGGCCACCAAGATGAAGGACTGCACCCGAGCCAGGCCA  
 ACTTCTGGGCAAGATCTGGCCCTCCAAACAAGGGCCCGCCCGCAACTTCCCCAGTCCCCCGCCGAGCCACCGCCCCCGCCGAGAC  
 TGGGCAATGGCGGAGGAGATCACCTCCTGCCCAAGCAGGAGCAAGGAGCACCCCCCCCCCTGGTGTCCCTGAAGTCCCTGTT  
 CGGCAACGACCCCTGTCCAGTAA

Fig. 75A

16. 2003 CON 02 AG gag .PEP  
 MGARASVL<sup>1</sup>SGK<sup>2</sup>L<sup>3</sup>DAWE<sup>4</sup>K<sup>5</sup>IRLRPGGKKYR<sup>6</sup>LKHLVWASRELERFALNPGLLETAEGCQQIMEQLQSAIRTGSEELKSLYNTVATLWCVHQRI  
 DIKDFKEALDKIEEVONKSKQKTQAAAATGSSSSQNYPIVQNAQGMTHQSMSPRTLNAWKVIEEKAFSPSEVI PMFSALSEGATPQDLNMM  
 LNI VGGHOAA<sup>10</sup>MQMLKDTINEEAAEWD<sup>11</sup>RVHPVHAGPIPPGQMRPRGSDIAGTSTLQEQIGWMTSNPPI PVGEIYKRWIVLGLNKIVRMYSY  
 VSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQ<sup>15</sup>NANPDCKSILRALPGATIEEMMTACQGVGPGHKARVLAEAMSQVQ  
 QSNIMMORGNFRGQRTIKCFNCGKEGHLARNCKAPRKKGCKGKEGHQMKDCTERQANFLGKIWPSSKGRPNFPQSRPEPTAPPAESFGM  
 GEEITSSPKQEPDRDKGLYPLPLTSLKSLFGNDP\$

Fig. 75B

2003 CON 02 AG gag .OPT  
 ATGGCGCCCGGCTCCGTGTCGGGGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGGCCCGGGCGGCAAGAAGTACCCGCTT  
 GAAGCACCTGGTGTGGCCCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG  
 AGCAGCTGCAGTCCGCCCTGGCACCCTGGAGAGCTGAAGTCCCTGTACAACACCCGTGGCCACCCCTGTGGTGCACCCAGCGCATC  
 GACATCAAGGACACCAAGGAGGCCCTGGACAAAGATCGAGGAGGTGCAGAACAAAGTCCAAGCAGAAGACCCAGCAGGCCCGCCGCCACCCGG  
 CTCCCTCCAGAACTACCCCATCGTGCAGAACGCCAGGCCAGATGACCCACAGTCCATGTCCCCCGCACCCCTGAACGCCCTGGGTGA  
 AGTGATCGAGGAGAGGCCCTTCTCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCCAGGACCTGAACATGATG  
 CTGAACATCGTGGCGGGCCACAGGCCGCTGAGATGCTGAAGGACACCATCAACGAGGAGGCCGCGAGTGGGACCCGCTGCACCCCGT  
 GCACGCCGCCCATCCCCCGGCCAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACACCTCCACCTGCAGGAGCAGATCGGCT  
 GGATGACCTCCAACCCCCCATCCCCGTGGCGGAGATCTACAAGCGCTGGATCGTGTGGCCCTGAACAAGATCGTGCATGTACTCCCC  
 GTGTCCATCTGGACATCCGCCAGGCCCAAGGAGCCCTTCCCGGACTACGTGGACCCGCTTCTTCAAGACCTTGGCGCCGAGCAGGCCAC  
 CCAGGAGGTGAGAACTGGATGACCGGACCCCTGCTGGTGCAGAACGCCCAACCCCGACTGCAAGTCCATCTGCGGCCCTGGGCCCGGCG  
 CCACCTGGAGGAGATGATGACCCCTGCCAGGGCGTGGCGGCCCAAGGCCCGGTGCTGGCCGAGGCCATGTCCCAGGTGCAG  
 CAGTCCAAACATCATGATGACGCGCGCAACTTCCGCGCCAGCCACCATCAAGTGTTCAACTGCGGCAAGGAGGCCACCTGGCCCCGCAA  
 CTGCAAGGCCCCCGCAAGAGGGCTGCTGGAAAGTGGGCAAGGAGGCCACAGATGAAGGACTGCAACCGAGCGCCAGGCCAACTTCTTGG  
 GCAAGATCTGGCCCTCCTCCAAGGGCCGCCCGCAACTTCCCGCCAGTCCCGCCAGCCACCGCCCGGCGGAGTCTTCGGCATG  
 GCGGAGGAGATCACCTCCTCCCGCAAGCAGGAGCCCGCGCACAAAGGGCCCTGTACCCCGCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCCTAA

Fig. 76A

17. 2003 CON 03 ABG gag. PEP  
 MGARASVLSGGKLD~~AW~~EKIRLRPGKKKRYIKHLVWASRELEREFALNPSLLETSEGCQQILEQLQPTLKTGSEELKSLYNTVATLYCVHQRI  
 EIKDTKEALDKIEEI QNKSQKTQAAATGSSSKVSNYPIVQNAQGMTHQSMSPTFLNAWVKVIEEKAFSPEVI PMFSALSEGATPQDL  
 NMMLNIVGGHQAAMQMLKDTINEEAAEWDRLHPAQAGFPFGQMPREPRGSDIAGTTSTLQEQIGWMTSNPPI PVGDIYKRWI IGLNKIVRM  
 YSPVSI LDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTETLLVQNPANPDCKTILRALGSGATLEEMMTACQGVGGPGHKARVLAEAMS  
 QVQANIMMQSNFRGPKRIKCFNCGKDGHLARNCRAPRKKGCKWCKGKEGHQMKDCTERQANFLGRIPWSSKGRPGNFQSRPEPSAPPAEN  
 FGMGEEITPSLKQEQKDREOHPPSISLKSLFGNDPLSQ\$

Fig. 76B

2003 CON 03 ABG gag. OPT  
 ATGGGGCCCGGCGCTCCGTGCTGTCCGGGGCAAGCTGGACGCTGGGAGAAGATCCGCCCTCGGCCCGGGCAAGAAGTACCCGCAT  
 CAAGCACCTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCTCCGAGGGCTCCAGCAGATCCTGG  
 AGCAGCTGACGCCACCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTACAACACCGTGGCCACCTGTACTGCGTGCACCCAGCGCATC  
 GAGATCAAGGACACCAAGGAGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAAGCAGAAGACCCAGAGCCCGCCACCCGGCACCCG  
 CTCTCTCCAAAGGTGTCCCAAGACTACCCCATCGTGCAGAACGCCAGGGCCAGATGACCCACCATGTTCTCCGCCCTGTCCGAGGGCCACCC  
 CCTGGGTGAAGGTGATCGAGGAGAAGGCTTCTCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGCCAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGGCCCGCATGCAAGTGTGAAGGACACCATCAACGAGGAGCCCGGAGTGGGACCCGCT  
 GCACCCGCCAGGCGGCCCTTCCCGCCCGCCAGATGCGGAGCCCGCGGCTCCGACATCGCCGACCCACCTCCACCTGCAAGGAGC  
 AGATCGGCTGGATGACCTCAACCTCGACATCCCGCAGGCCCAAGGACCTTCCCGGACTACGTGGACCGCTTCTTCAAGACCTGCGGCCCGA  
 TACTCCCGCTGTCCATCCTGGACATCCCGCAGGCCCAAGGACCTTCCCGGACTACGTGGACCGCTTCTTCAAGACCTGCGGCCCGA  
 GCAGGCCACCCAGGACGTGAGAACTGGATGACCCGAGACCTTGTGTGAGAACGCCAACCCCGACTGCAAGACCAATCCTGCGGCCCTGG  
 GCTCCGGCCACCTGGAGGAGATGATGACCCGCTGCCAGGGCGTGGCGGCCCGCCCAAGGCCCGCGTGTGCGCGAGGCCATGTCC  
 CAGGTGCAGAACGCCAACATCATGATGACAGAGTCCAACTTCCCGGCCCAAGCGCATCAAGTGTCAACTGCGGCAAGGACGGCCACCT  
 GGCCCGCAACTGCCCGCCCGCCCGCAAGAGGGCTGTGGAAGTGGCAAGGAGGCCCCACAGATGAAGGACTGCAACGAGGCCAGGCCA  
 ACTTCCCTGGCCGATCTGGCCCTCCCTCAAGGGCCCGCCCGCAACTTCCCGGAGTCCCGCCCGGAGCCCTCCGCCCCCGCGGAGAAC  
 TTCGGCATGGCGGAGGAGATCACCCCTCCCTGAAAGCAGGAGCAGAGGACCCCGGAGCAGCACCCCGCTCCATCTCCCTGAAAGTCCCTGTT  
 CGGCAACGACCCCTGTCCAGTAA

Fig. 77A

18. 2003 CON 04 CFX gag . PEP  
 MGARASVLGGKLD~~AW~~ERIRLRPGKKYRLKHLVWASRELERFALNPLLETAEGCQQLMEQLQSTLKTGSEELKSLFNTIATLWCVHQRI  
 DVKDTKEALDKVEEMONKSQKTKQAAADTGGSSNVSONYPIVQNAOQOMVHOSISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL  
 NMMLNIVGGHQAAQMMLKDTINEEAAEWDRAHPVHAGPIPPGOMREPRGSDIAGTSTLQEQIGWMTSNPPPIPVGEIYKRWIIILGLNKIVRM  
 YSPVSILDIRQGPKEPFRDYVDRFFKCLRAEQATQEVKNWMTETLLVQANPDCKSILKALGTGATLEEMMTACQVGGPSSHKARVLAEAMS  
 QASNAAAAIMMOKSNEFKQRRRIIKCFNCGKEGHLARNCRAPRKKGCKGKEGHQMKDCTERQANFLGRMWPPSSKGRPNFLQSRPEPTAPP  
 AESLEMKEETTSPPKQEPDRDKELYPLTSLKSLFGSDPLSQS.

Fig. 77B

2003 CON 04 CFX gag . OPT  
 ATGGGGCCCGGCTCCGTGCTCGGGGCAAGCTGGAGCGCATCCGCCCTGGCCCGGGGCAAGAAGTACCGCCT  
 GAAGCACTGGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCGGCTGTGAGACCGCGAGGGCTGCCAGCAGCTGATGG  
 AGCAGTGCAGTCCACCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTCAACACCATCGCCACCTGTGGTGCCTGCACCCAGCGCATC  
 GACGTGAAGGACACCAAGGAGGCCCTGGACAAGTGGAGGAGATGCAGAACAAAGTCCAAGCAGAAGACCCAGCAGGCGCCCGCACACCGG  
 CGGCTCCTCAACGTGTCCAGAACTACCCCATCGTGCAGAACGCCAGGCGCAGATGGTGCACCATCCATCTCCCCCGCACCCCTGAACG  
 CCTGGTGAAGGTGATCGAGGAGAAGGCTTCTCCCCGAGTGTCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGCGCCATGCAGATGCTAAGGACACCATCAACGAGGAGGCCCGCGAGTGGGACCGCGC  
 CCACCCGTGCACCGCGCCCATCCCCCGGCGAGATGCGGAGCCCGGCTCCGACATCGCCGGCACCATCCACCTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCTCAACCCCATCCCCGTGGCGAGATCTACAAGCGTGGATCATCTGGGCTGAACAAGATCGTGGCATG  
 TACTCCCCGTTCATCTGGACATCCGCCAGGCCCTCCGGGACTACGTGGACCGCTTCTCAAGTGCCTGGCGGCCGA  
 GCAGGCCACCCAGGAGTGAAGAACTGGATGACCGAGACCTGTGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGAAGGCCCTGG  
 GCACCGGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCCAACAAGCCCGCTGTGGCCGAGGCCATGTCC  
 CAGGCTCCAAACGCCCGCCGATCATGATGCAAGTCCAATTCAGGGCCAGCGCCGATCATCAAGTGTCAACTGGCGCAAGGA  
 GGGCCACTGGCCCGCAACTGCCCGCCCGCAAGAGGGCTGTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGACTGCACCCGAGC  
 GCCAGGCCAACTTCTGGGCCGATGTGGCCCTCTCCAAAGGCCCGCCCGCAACTTCTGTGAGTCCCGCCCGAGCCACCCCGCCCGC  
 GCCAGTCCCTGGAGATGAAGGAGGAGACCACTCTCCCCCAAGCAGGAGCCCGGCAAGGAGTGTACCCCTGACCTCCCTGAAAGTC  
 CCTGTTCGGCTCCGACCCCTGTCCCCAGTAA

Fig. 78A

19. 2003 CON 06 CPX gag . PEP  
 MGARASVLSGGKLD~~EW~~EKIRLRPGKKKYRLKHLVWASRELERFALNPLLETAEGCQOIIEQLQSALKTGSEELKSLYNTVATLYCVHQRI  
 KVTDTKEALDKIEEIQNKSKOKAQAAAAATGNSSNLSONYPIVQNAQGMVHQAI SPRTLN~~AW~~VKVEEKAFSEV~~IP~~MF~~S~~ALSEGATPQDL  
 NMMLNIVGGHQAAQM~~L~~KDTINEEAAEWD~~R~~VHPVHAGPIPPGQMRPRGSDIAGTSTLQEQIGWMTSNPP~~IP~~VGEIYKRWIIILGLNKIVRM  
 YSPVSILDIRQGPKEPERDYVDRFFKTLRAEQATQEVKNWMTDLLVQ~~NAN~~PDCKTILKALGPGATLEEMMTACQGVGGP~~G~~HKARVLAEAMS  
 QASGTEAAIMMQS~~N~~FKGPKRSIKCFNCGKEGHLARNCRAPRKKGWKCKGEGHQMKDCTERQANFLGIWPSNKG~~R~~PNFLQNRPEPTAPP  
 AESFGFEETA~~PS~~PKQEPKEKELYPLASLKS~~L~~FGNDP\$

Fig. 78B

2003 CON 06 CPX gag . OPT  
 ATGGGCGCCGGCCTCCGTGCTGTCCGGGGCAAGCTGGACGAGTGGAGAAAGATCCGCCCTCGGCCCGGGGCAAGAAGATACCGCCT  
 GAAGCACCTGGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCGCCGAGGGCTGCCAGCAGATCATCG  
 AGCAGTGCAGTCCGCCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTACAACACCCGTGGCCACCCCTGFACTGCGTGCACCCAGCGCATC  
 AAGGTGACCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGCAGAAGGCCCAAGGCCAGCGCCCGCCGACCCGG  
 CAACTCCTCCAACTGTCCCAAGAACTACCCCATCGTGCAGAACGCCAGGCCAGATGGTGCACCAAGGCCATCTCCCCCGCACCTGAACG  
 CCTGGTGAAGTGATCGAGGAGAAGGCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGCCGATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCCGAGTGGGACCCGCGT  
 GCACCCGTCACCGGCCCATCCCCCGGCCAGATGCGGAGCCCGCGGCTCCGACATCGCCGGACCCACCTCCACCCCTGCAGGAGC  
 AGATCGGCTGGATGACTCCAAACCCCATCCCCGTGGCGAGATCTACAAGCGCTGGATCATCTGGGCCCTGAACAAGATCGTGGCCATG  
 TACTCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCAAGGAGCCCTCCGGGACTACGTGGACCGCTTCTTCAAGACCCCTGCCGCCCGA  
 GCAGGCCACCCAGGAGTGAAGAATGGATGACCGACACCCCTGCTGTGTGAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG  
 GCCCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCAAGGCCCGCTGTGCAAGACCATCCTGAAGGCCCTGTCC  
 CAGGCTCCGGCACCGAGGCCCATCATGATGCAAGTCCAATCAAGGGCCCCAAGCGTCCATCAAGTGTCAACTGCGGCAAGGA  
 GGGCCACTGGCCCGAACTGCCGCCCCCGCAAGAAGGCTGTGGAAGTGGCAAGGAGGCCCAACAGATGAAGGACTGCACCCGAGC  
 GCCAGGCCAACTTCTTGGCAAGATCTGGCCCTCCAAACAAGGCCCGCCCGGCAACTTCTTGCAGAAACCGCCCGGACCCGCCCCCCCC  
 GCCGAGTCCCTTCGGCTTCGGCGGAGAGACCGGCCCTTCCCCCAAGCAGGAGCCCAAGGAGACTGTACCCCTGGCCCTCCCTGAAGTC  
 CCTGTTCGGCAACGACCCCTAA

121/178

Fig. 79A

20. 2003 CON 07 BC gag . PEP  
 MGARASILRGGKLDKWEKIRLRPGGKHYMLKHLVWASRELERFALNPLLELSETSEGCKQIIKQLQPALQGTGTEELRSLFNTVATLYCVHTEI  
 DVRDTKEALDKIEEFQNKIQKTQAKEADGKVSQNYPIVONLQGMVHQPISPRTLNAWKVVEKAFSEV I PMFSALSEGATPQDLNMTM  
 LNTVGGHQAAAMQILKDTINEEAAEWDRLHPVHAGPIAPQMREPRGSDIAGTTSNLOEQJAWMTSNPPVPVGDIIYKRWII LGLNKKIVRMYS  
 TSILLDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDILLVQANPDCKTILRALGPGASIEEMMTACQGVGSPSHKARVLAEAMSQTN  
 STILLMQRSNFKSKRIVKCFNCGKEGHIARNCRAPRKKGWCKGKEGHQMKDCTERQANFLGKIWPSHKGRFNGNLFQSRPEPTAPPEESFRF  
 GEETTPSQQEPIDKELYPLTSLKSLFGNDPSSQ\$

Fig. 79B

2003 CON 07 BC gag . OPT  
 ATGGCGCCCGCCCTCCATCCTTGGCGGGGCAAGCTGGACAAGTGGGAGAAGATCCGCCGTGGCCCCGGCGGCAAGAAGCACTACATGCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA  
 AGCAGCTGACGCCCGCCCTGCAGACCGGCAACCGGAGAGCTGCCCTCCCTGTTCAACACCCGTGGCCACCCTGTACTGCGTGACACCCGAGATC  
 GACGTGGCGCACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGAAACAAGATCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGACCGG  
 CAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACAGCCCATTCGCCCCCGCACCCCTGAACGCCTGGGTGA  
 AGTGGTGGAGGAGAGGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGGCCACCCCGAGTGGACCCGCTGCACCCCGT  
 CTGAACACCGTGGCGGCCACCAAGCCGATGCAGATCCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGACCCGCTGCACCCCGT  
 GCACGCCGCCCATCGCCCCCGGCGAGTGGCGGACATACAAAGCCGTGGATCATCTGGCCCTGAACAAGATCGTGCATGTACTCCCC  
 GGATGACCTCCAACCCCGTGGCGGACATCAAAAGCCGTGGATCATCTGGCCCTGAACAAGATCGTGCATGTACTCCCC  
 ACCTCCATCTGGACATCAAGCAGGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACCCCTGCGCGCCGAGCAGGCCAC  
 CCAGGACGTGAGAACTGGATGACCCGACACCTGCTGGTGCAGAACGCCAAACCCCGACTGCAAGACCATCTGCGGCCCTGGGCCCGGGG  
 CCTCCATCGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCCAAGGCCCGGTGCGCGAGGCCATGTCAGACCCAAAC  
 TCCACCATCCTGATGACGCTCAAACCTCAAGGCTCCAAAGCCATCGTGAAGTGTCAACTCGGCAAGGAGGCCACATCGCCCCGCAA  
 CTGCCGCCCCCGCAAGAGGCTGTGGAAGTCCGGCAAGGAGGCCACAGATGAAGGACTGCACCGAGCCGAGGCCAATTCTCTGG  
 GCAAGATCTGGCCCTCCACACAAGGCCGCCCCGCAACTTCTGTGAGTCCCGCCCGAGCCACCGCCCCCGGAGGAGTCTTCCGCTTC  
 GCGGAGGAGACCACCCCTCCAGAGGAGGCCCATCGACAAGGAGCTGTACCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCTCCTCCAGTAA

122/178

Fig. 80A

21. 2003 CON 08 BC gag. PEP  
 MGARASILRGGKLDKWEKIRLRPGGKHHYMLKHLVWASRELERFALNPLGELLESEGCKQIKIQLPALQOTGTEELRSLFNTVATLYCVHAEI  
 EVRDTKEALDKIEEEQNKIQOKTQAKEADEKVSQNYPIVONLQGMVHQPLSPRTLNAWVKVVEEKAFSPPEVI PMFTALSEGATPQDLNMT  
 LNTVGGHOAAMQMLKDTINEEAAEWRDLHPVHAGVPAGOMREPRGSDIAGTTSTLQEQIGWMTNPPPIPVGEIYKRWII LGLNKIVRMYSP  
 TSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALPGASLEEMMTACQGVGGPSHKARVLAEAMSQTN  
 NTLIMORSNFKGSKRIVKCFNCGKEGHI AKNRAPRKKGCKWKGKEGHQMKDCTERQANFLGKIWP SHKGRPNFLQSRPEPTAPPAESFRF  
 EETTPAPKQEPKDRREPLTSLRSLFGSDPLISQ\$

Fig. 80B

2003 CON 08 BC gag. OPT  
 ATGGCGCCCGGCTCCATCCTGCGGGCGCAAGCTGGACAAGTGGAGAAGATCCGCCCTGCGCCCGGGCAAGAAGCACTACATGCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA  
 AGCAGTGCAGCCCGCTGCAGACCGGACCGGACCGGAGCTGCGCTCCCTGTTCACACCCGTGGCCACCTGTACTGCGTGCAACCGGAGATC  
 GAGGTGCGGACACCAAGGAGCCCTGGACAAGATCGAGGAGGAGAGAACAGATCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGACGA  
 GAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCCAGCCCTGTCCCCCGCACCTGAACCGCCTGGGTGA  
 AGGTGGTGGAGGAGAAAGGCTTCTCCCCGAGGTGATCCCCATGTTACCCGCTGTCCGAGGGCCACCCCGAGGACCTGAACACCCATG  
 CTGAACACCGTGGCGGCCACACGCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGGAGTGGACCCGCTGCACCCCGT  
 GCACGCGCCCGTGGCCCGGCGAGATGCGCGAGCCCGGCTCCGACATCGCCGGCACCACTCCACCCCTGCAGGAGCAGATCGGCT  
 GGATGACCAAGAACCCCCCATCCCCGTGGCGAGATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCATGTACTCCCC  
 ACCTCCATCCTGGACATCAAGCAGGCCCAAGGACCCCTCCCGACTACGTGGACCGCTTCTTCAAGACCCCTGCGCCCGGAGCAGGCCAC  
 CCAGGACGTGAAGAACTGGATGACCGACACCCCTGTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCCCTGGGCCCGGCG  
 CCTCCCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCCAACAAGGCCCGTGTGCCGAGGCCATGTCCCAGACCAAC  
 AACACCATCCTGATGCAGCGCTCCAACTTCAAGGGCTCCAAGCGCATCGTGAAGTCTTCAACTGCGGCAAGGAGGCCACATCGCCAAAGAA  
 CTGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGGAGGCCACCCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCCTGG  
 GCAAGATCTGGCCCTCCCAACAAGGGCCCGCCCGGCAACTTCTGCAATCCCGCCCGAGCCCAAGCCCGCCCGGAGTCTTCCGCTTC  
 GAGGAGACCACCCCGCCCAAGCAGGAGCCCAAGGACCGGAGCCCTGACCTCCCTGCGCTCCCTGTTCCGGCTCCGACCCCGCTGTCCCA  
 GTAA



Fig. 81A

22. 2003 CON 10 CD gag . PEP  
 MGARASVLSGGKLD~~EW~~EKIRLRPGGKKYRLKHLVWASRELERFALNPNLLETSEGCKQIIQLQPAIQTSGEEIKSLYNTVATLYCVHERI  
 KVTDTKEALDKIEEEQTKSKKAQOATADTGNSSQVSONYPIVQNLQGMVHQPLSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL  
 NTMLNTVGGHOAMQMLKETINEEAEDWDLRHPVQAGPVAPGQIREPRGSDIAGTSTLQEQIRWMTSNPPIPVGEIYKRWIIIGLNKIVRM  
 YSPVSI~~LD~~IRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMTACQGVGGPSHKARVLAEAMS  
 QATSGNAIMMQRGNFKGPKKIIKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLKGIWPSNKGKRPNGFLQSRPEPTAPPA  
 ESFGFGEIITPSQKQEQKDKELHPLASLKSIFGNDPLSQS

Fig. 81B

2003 CON 10 CD gag . OPT  
 ATGGGCGCCCGGCTCCGTGCTGTCGGGGCAAGCTGGACGAGTGGGAGAAGATCCGGCTCGGCCCGGGCAAGAAGAAGTACCCGCT  
 GAAGCACCTGGTGTGGGCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGGCTGCAAGCAGATCATCG  
 GCCAGCTGCAGCCCGCATCCAGACCGGCTCCGAGGAGATCAAGTCCCTGTACAACACCCGTGGCCACCTGTACTGCGTGCACGAGCGCATC  
 AAGGTGACCCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGACCAAGTCCAAGAAGAAGGCCCAAGGCCCAGAGCCACCGCCGACACCCGG  
 CAACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAAGCCCCGTGTCCCCCGCACCCCTGAACG  
 CCTGGGTGAAGGTGATCGAGGAGAAGGCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCCAACCCCGCAGGACCTG  
 AACACCATGCTGAACACCCGTGGCGGCCACAGCCCGCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCCGGAGTGGGACCCGCT  
 GCACCCGTGCAGGCCGCCCCGTGGCCCCCGGATCCCGGCTCCGACATCGCCGGCACCATCCTGGCCCTGAACAAGATCGTGGCATG  
 AGATCCGCTGGATGACCTCCAACCCCATCCCGTGGCGGAGATCTACAAGCGTGGATCATCCTGGCCCTGAACAAGATCGTGGCATG  
 TACTCCCCGTGTCATCTGGACATCCCGCAGGCCCAAGGACCCCTTCCCGACTACGTGGACCCGTTCTACAAGACCTTCCGCGCCGA  
 GCAGGCCITCCAGGACGTGAAGAACTGGATGACCGAGACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGAAGGCCCTGG  
 GCCCGCCACCCCTGGAGGAGATGATGACCCGCTGCCAGGGCGTGGCGGCCCTCCCAAGGCCCGCGTGTGCCGAGGCCATGTCC  
 CAGGCCACCTCCGGCAACCCCATCATGATGCAGCGCGGCAACTTCAAGGGCCCCAAGAAGATCATCAAGTGTCTCAACTGCGGCAAGGAGGG  
 CCACATGCCCAAGAACTGCCGCCCCCGCAAGAAAGGCTGCTGGAAGTGCAGGCGCCGAGGCCACACAGATGAAGGACTGCACCGAGCGCC  
 AGGCCAACTTCCITGGCAAGATCTGGCCCTCCAACAAGGCCCGCCGCAACTTCTGCAAGTCCCGCCCCGAGCCACCGCCCCCGCC  
 GAGTCTTCGGCTTCGGCGAGGAGATCACCCCTCCACAGAAAGCAGGAGCAGAAGGAGCTGCACCCCTGGCCTCCCTGAAGTCCCT  
 GTTCGGCAACGACCCCTGTCCAGTAA

124/178

Fig. 82A

23. 2003\_CON\_11\_CPX\_gag.gag.PEP  
**gag . PEP**MGARASVLSGGKLDWAEKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETAEGCQQLPALGTGTELRSLYNTVATL  
 YCVHHRIEVKDTKEALDKIEEIQNKSQKQAAADTGNSSKVSQNYPIVQNAQGMVHQAI SPRTLNAWKVVEEKAFSP EVI PMFSALSE  
 GATPQDLNMMNLNIVGGHQAAAMQMLKDTINEEAAEWDRVHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIGWMTGNPPVPVGEIYRRWIIIG  
 LNKIVRMYSPVSIIDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKSWMTETLLIQNANPDCKSILRALGPGATLEEMMTACQGVGPGPHKAR  
 VLAEAMSQVQQTINIMQRSNFKGQKRIKFCNCGKEGHLARNCRAPRKKGCWKCKEGHQMKDCTERQANFLGKIWPSSKGRPGNFIQSRPEP  
 TAPPAESFGFEEIAPSPKQEPKEKELYPLTSLKSLFGSDPLSQ\$

Fig. 82B

2003\_CON\_11\_CPX\_gag.OPT  
 ATGGCGCCCGGCTCCGTGCTCCGGCGGCAAGCTGGAGCGCTGGGAGAAGATCCGCCCTGGCCCGCGGCAAGAAGTACCGCCT  
 GAAGCACCTGGTGTGGCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG  
 GCCAGCTGACGCCCGCTGGCACCGGACCGGAGCTGGCTCCCTGTACAACCCGTGGCCACCCTGTACTGCGTGCACCCACCGCATC  
 GAGTGAAGCACCAAGGAGGCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGCAGAAGAGCAGAGGCGCCCGCCGACCCCG  
 CAACTCCTCAAAGTGTCCAGAACTACCCATCGTGCAAGACGCCAGGCCAGATGGTGCACCCAGGCCATCTCCCCCGCACCCCTGAACG  
 CCTGGTGAAGTGTGGAGGAGAAAGCCCTTCTCCCCAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCCAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGGCCCGCATGCAGATGCTGAAGGACACCATCAACGAGGCGCCGCGAGTGGGACCCGGT  
 GCACCCGTGCACGCCGCCCATCCCCCGGCCAGATGCCGAGCCCGCGGCTCCGACATCGCCGGCACCATCCACCTGCAAGGAGC  
 AGATCGGCTGGATGACCGGCAACCCCGTGGCGAGATCTACCGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCGCATG  
 TACTCCCCGTGCTCCATCCTGGACATCCGCCAGGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACCTGCGGCCGA  
 GCAGGCCACCCAGGAGGTGAAGTCTGGATGACCGAGACCTTGTGTATCCAGAACCACCCGACTGCAAGTCCATCTGCGGCCCTGG  
 GCGCGGCCACCTGGAGGAGATGATGACCGCTGCCAGGCGTGGCGGCCCAAGGCCCGCGTGTGGCCGAGGCCATGTCC  
 CAGGTGCAGCAGACCAACATCATGATGACCGCTCCAACCTCAAGGCCAGAACCGCATCAAGTGTTCAACTGCGCAAGGAGGCCACCT  
 GGCCCGCAACTGCCCGGCCCGCCGCAAGAGGCTGCTGGAAGTGGCAAGGAGGCCACCCAGATGAAGGACTGACCCGAGCCAGGCCA  
 ACTTCTGGCAAGATCTGGCCCTCTCCAAGGCCCGCCCGCAACTTCTGAGTCCGCCCCGAGCCACCGCCCGCCCGCGAGTCC  
 TTCGGCTTCGGCGAGGAGATCGCCCTCCCAAGCAGGAGCCCAAGGAGGAGTGTACCCCTGACCTCCCTGAAGTCCCTGTTCGG  
 CTCCGACCCCTGTCCCAAGTAA

125/178

Fig. 83A

24. 2003 CON 12 BF.gag.PEP  
 MGARASVLSGGEIDRWEKIRLRPGGKKYRLKHIIVWASRELEFAVNPGLLETSEGCRKIIIGQLPSLQGTSEELRSLYNTIAVLYFVHQKV  
 EVKDTKEALDKLEEFQNKSQKTQAAADKGVSNQYPIVQNLQGMVHQALSPTLNAWKVVEEKAFSPEVI PMFSALSEGATPODLNLTML  
 NTVGGHQAAAMQLKDTINEEAAEWDRLHPVHAGPIPPGQMPREPRGSDIAGTSTLQEQIQWMTSNPPVPVGEIYKRWII LGLNKIVRMYSPV  
 SILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQANANPDCKTILKALGPGATLEEMMTACQGVGPGHKARVLAEMSQVFN  
 TTVMMQKSNFKGQRRIVKCFNCGKEGHI AKNCRAPRKKGCKGREGHQMKDCTERQANFLGKIWPSNKGPRGNFLQNRPEPTAPPAESFFG  
 GEEITPSPKQEQKDEGLYPPLASLKSIFGNDP\$

Fig. 83B

2003 CON 12 BF.gag.OPT  
 ATGGCGCCCGGCTCCGTGCTGTCCGGCGCGAGCTGGACCGCTGGAGAAAGATCCGCCCTGGCCCGGGCGCAAGAAGTACCGCCT  
 GAAGCACATCGTGTGGGCTCCCGAGCTGGAGCGCTTCGCCGTGAACCCGGCTGCTGGAGACCTCCGAGGGTCCGCAAGATCATCG  
 GCCAGTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTCGCTGCTACAAACCATCGCCGTGTACTTCTGTGCACCCAGAAAGTG  
 GAGGTGAAGGACACCAAGGAGCCCTGGACAAGCTGGAGGAGGAGCAAGTCCAGCAAGAACCCAGAGCCCGCCGCGCAAGG  
 CGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAAGCTGTCCCGCCCTGTCCCGCCCTGAAACCATGCTG  
 TGGTGGAGGAGAAGCCCTTCTCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCAGGACCTGAACCATGCTG  
 AACACCGTGGCGGCCACAGCCAGATGCAGATGCTGAAGGACACCATCAACGAGGAGCGCCGAGTGGACCGCTGCACCCCGCTGCA  
 CGCCGGCCCATCCCCCGCCAGATGCAGGAGCCCGGCTCCGACATCGCCGACACCATCCACCTGCAGGAGCAGATCCAGTGGA  
 TGACCTCAACCCCGTGCAGGAGATCTACAGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCATGFACTCCCGCTG  
 TCCATCTGGACATCCGCCAGGCCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACCTGGCGCCGAGCAGGCCACCCA  
 GGAGGTGAAGGGCTGGATGACCGACACCCCTGCTGTTGCAGAACCCCGACTGCAAGACCATCTGAAGGCCCTGGGCCCGGCGCCA  
 CCCGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGCCAGCCCGCATCGTGAAGTGTCAACTGCGCAAGGAGGCGCCACATCGCCAAGAA  
 ACCACCGTATGATGCAGAACTCCAACTCAAGGGCCAGCCCGCATCGTGAAGTGTCAACTGCGCAAGGAGGCGCCACATCGCCAAGAA  
 CTGCCGCCCCCGCAAGAGGCTGCTGGAAAGTGGCCCGGAGGGCCACAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGG  
 GCAAGATCTGGCCCTCCAAACAAGGGCCCGCCCGCAACTTCTTGCAGAACCGCCCGAGCCCAAGGCGCCCGCCCGGAGTCTTCGGCTTC  
 GCGGAGGAGATCACCCCTTCCCCCAAGCAGGAGCAGAAAGGAGGGCCCTGTACCCCCCTGGCCCTCCCTGAAGTCCCTGTTCCGGCAACGA  
 CCCCCTAA

126/178

Fig. 84A

25. 2003 CON 14 BG gag. PEP  
 MGARASVLSGGKLDWEEKIRLRPGGKKYRMKHLVWASRELERFALNPDLLLETAEGCQQIMGQLQPALQGTGTEEIRSLFNTVATLYCVHQKI  
 EVKDTKEALEEVEKAQKKSQKKQQAAMDEGNNSQASQNYPIVQNAQGMVHQAI SPRILNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLN  
 TMLNTVGGHQAAOMQLKDTINEEAEWDRMHPQQAQPIPPGQIREPRGSDIAGTSTLQEQIRWMTSNPPIPVGEIYKRWIILGLNKIVRMY  
 SPVSIIDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGMWTDLLVQANPDCKTILRALPGATLEEMMTACQGVGGPSHKARVLAEMSQ  
 ASGATIMMQKSNFKGPRRNKICFNCGKEGHLARNCRAPRKKKGCWKCKEGHQMKDCTESKANFLGKIWPSNKGPRPGNFLQNRPEPTAPPAES  
 FGFEEIAPSPKQEPKEKEIYPLASLKSFLGSDP\$SQ\$

Fig. 84B

2003 CON 14 BG gag. OPT  
 ATGGCGCCCGGCGCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCCTGGGAGAAAGATCCGGCTCGCGCCCGGGCGCAAGAAGTACCGCAT  
 GAGCACCTGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGACCTGTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG  
 GCCAGTGCAGCCCGCTGCAGACCGGACCCGAGGAGATCCGCTCCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCCACAGAAAGATC  
 GAGGTGAGGACACCAAGGAGGCCCTGGAGGAGTGGAGAGCCAGAAAGTCCCAGAAAGACAGCAGGCCCGCCATGACGAGGCGCA  
 CAACTCCAGGCTCCAGAACTACCCCATCGTGCAGAACGCCAGGGCCAGATGGTGACCAAGGCCATCTCCCCCGCACCCAGGACCTGAACGCT  
 GGTGAAGGTGGTGGAGAGAGGCCCTTCTCCCCGAGGTGATCCCAATGTTCTCCGCTGTCCGAGGGCCACCCCGAGACCTGAAC  
 ACCATGCTGAACACCGTGGCGGCCACAGGCCGCTGCAGATGCTGAAGGACACCAATCAACGAGGAGGCCCGCCGAGTGGGACCCGATGCA  
 CCCCAGAGGCCGCCCATCCCCCGCCAGATCCGCGAGCCCGCGGCTCCGACATCCCGGCCACCACTCCACCTCCAGGAGCAGA  
 TCCGCTGGATGACCTCCAACCCCTCCCGTGGCGGATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGCCGATGTAC  
 TCCCCGTGTCCATCTGGACATCCGCCAGGCCCAAGGAGCCCTCCGCGACTACGTGGACCCTTCTCAAGACCTCGCGCCCGGAGCA  
 GGCAACCAAGGAGTGAAGGCTGGATGACCCGACACCCCTGTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCCCTGGGCC  
 CCGCGCCACCTGGAGGAGATGATGACCCGCTGCCAGGCGTGGCGGCCCTCCCAAGGCCCGCGTGTGGCCGAGGCCATGTCCCAG  
 GCCTCCGGCCACCATCATGATGCAAGTCCAATCAAGGCCCCCGCCGCAACATCAAGTCTTCAACTGCGGCAAGGAGGCCACCT  
 GGCCCGCACTGCCCGGCCCCCGCAAGAGGCTGCTGGAAGTGGCAAGGAGGCCACAGATGAAGACTGCACCGAGTCCAAGGCCA  
 ACTTCTGGGCAAGATCTGGCCCTCCAACAAGGCCCGCCCGCAACTCTGCAGAACCGCCCGAGCCACCGCCCGCCGAGTCC  
 TTCGGCTTCGGCGGAGGAGATCGCCCCCTCCCCCAAGCAGGAGCCCAAGGAGAGATCTACCCCTGGCCCTCCCTGAAGTCCCTGTTCGG  
 CTCGACCCCTAATCCCCAGTAA

128/178

Fig. 85A

31. 2003 CONS nef.PEP  
 MGKWSKSSIVGWPVAVRERIRRTTPAAEGVAVSQDLDDKHGAI TSSNTAATNADCAWLEAQEEEEVGFVPRPQVPLRPMTYKGAFDLSHFLK  
 EKGGLDGLIYSKKRQEI LLDLWVYHTQGYFPDWQNYTPGPIRYPLTFGWCFLKLVDPDEEVEEANEGENNCLLHPMCQHGMEDEREVLMMWK  
 FDSRLALRHIARELHPEFYKDC\$

Fig. 85B

2003 CONS nef.OPT  
 ATGGCGGCAAGTGGTCCAAGTCCCTCCATCGTGGGCTGGCCCGCGGTGCGGAGGCGCATCCGCCACCCCGCCCGGAGGGCGGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCACGGGCCATCACCTCCTCAACACCGCCGACCAACCGCGACTGCGCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGGCCCTTGGACCTGTCCCACTTCCTGAAG  
 GAGAAGGCGGCTGGACGGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCCCGCATCCGCTACCCCTGACCTCGGCTGGTGTCAAGCTGGTGGCCCGTGGACCCCGAGGAGGTGG  
 AGGAGGCCAACGAGGGCGAGAACAACCTGCTGCACCCCATGTGCCAGCACCGGCATGGAGGACCGGAGGCTGCTGATGTGGAAG  
 TTCGACTCCCGCCTGGCCCTGGCCACATCGCCCGGAGCTCACCCGAGTCTACAAGGACTGCTAA

Fig. 86A

32. 2003 M. GROUP.anc nef.PEP  
 MGKWSKSSIVGWPVAVRERMRRTAPAAEGVAVSQDLDDKHGAI TSSNTAATNADCAWLEAQEEEEVGFVPRPQVPLRPMTYKAAFDLSHFLK  
 EKGGLDGLIYSKKRQEI LLDLWVYHTQGYFPDWQNYTPGPIRYPLTFGWCFLKLVDPDEEVEEANEGENNCLLHPMCQHGMEDEREVLMMWK  
 FDSRLALRHIARELHPEFYKDC\$

Fig. 86B

2003 M. GROUP.anc nef.OPT  
 ATGGCGGCAAGTGGTCCAAGTCCCTCCATCGTGGGCTGGCCCGCGGTGCGGAGGCGCATGCGCCGCAACCGCCCGCCCGGAGGGCGGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCACGGGCCATCACCTCCTCAACACCGCCGCAACCGCCGACTGCGCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGGCCCTTGCACCTGTCCCACTTCCTGAAG  
 GAGAAGGCGGCTGGACGGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCCCGCATCCGCTACCCCTGACCTCGGCTGGTGTCAAGCTGGTGGCCCGTGGACCCCGAGGAGGTGG  
 AGGAGGCCAACGAGGGCGAGAACAACCTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACCGGAGGCTGCTGATGTGGAAG  
 TTCGACTCCCGCCTGGCCCTGGCCACATCGCCCGGAGCTCACCCGAGTCTACAAGGACTGCTAA

129/178

**Fig. 87A****33. 2003 CON A nef .PEP**

MGKWSKSI<sup>1</sup>VGWPD<sup>2</sup>IRIR<sup>3</sup>RR<sup>4</sup>PPA<sup>5</sup>AK<sup>6</sup>VGAV<sup>7</sup>SQD<sup>8</sup>LD<sup>9</sup>KYGA<sup>10</sup>VT<sup>11</sup>INNTA<sup>12</sup>AT<sup>13</sup>QASCAWLEA<sup>14</sup>QEEEE<sup>15</sup>EVGF<sup>16</sup>PVRP<sup>17</sup>QVPL<sup>18</sup>RPMT<sup>19</sup>FKGAFD<sup>20</sup>LSFFL<sup>21</sup>  
 KEKGLDGLI<sup>22</sup>YSQKRQE<sup>23</sup>ILDLWVY<sup>24</sup>NTQGYFPD<sup>25</sup>WQNY<sup>26</sup>TPGPT<sup>27</sup>RFPL<sup>28</sup>TFGWCF<sup>29</sup>KLVPVD<sup>30</sup>PEVEEA<sup>31</sup>TEGENN<sup>32</sup>CLLHP<sup>33</sup>ICQHGMD<sup>34</sup>DEEKEV<sup>35</sup>LMW<sup>36</sup>  
 KFD<sup>37</sup>SRLARR<sup>38</sup>HIALEM<sup>39</sup>HPEFY<sup>40</sup>KDC<sup>41</sup>\$

**Fig. 87B****2003 CON A nef .OPT**

ATGGCGGCAAGTGGTCCAAAGTCCATCGTGGGCTGGCCCGACATCCGGGAGCGCATCCGCCGACCC<sup>1</sup>CC<sup>2</sup>CC<sup>3</sup>CC<sup>4</sup>CC<sup>5</sup>CC<sup>6</sup>CC<sup>7</sup>CC<sup>8</sup>CC<sup>9</sup>CC<sup>10</sup>CC<sup>11</sup>CC<sup>12</sup>CC<sup>13</sup>CC<sup>14</sup>CC<sup>15</sup>CC<sup>16</sup>CC<sup>17</sup>CC<sup>18</sup>CC<sup>19</sup>CC<sup>20</sup>CC<sup>21</sup>CC<sup>22</sup>CC<sup>23</sup>CC<sup>24</sup>CC<sup>25</sup>CC<sup>26</sup>CC<sup>27</sup>CC<sup>28</sup>CC<sup>29</sup>CC<sup>30</sup>CC<sup>31</sup>CC<sup>32</sup>CC<sup>33</sup>CC<sup>34</sup>CC<sup>35</sup>CC<sup>36</sup>CC<sup>37</sup>CC<sup>38</sup>CC<sup>39</sup>CC<sup>40</sup>CC<sup>41</sup>CC<sup>42</sup>CC<sup>43</sup>CC<sup>44</sup>CC<sup>45</sup>CC<sup>46</sup>CC<sup>47</sup>CC<sup>48</sup>CC<sup>49</sup>CC<sup>50</sup>CC<sup>51</sup>CC<sup>52</sup>CC<sup>53</sup>CC<sup>54</sup>CC<sup>55</sup>CC<sup>56</sup>CC<sup>57</sup>CC<sup>58</sup>CC<sup>59</sup>CC<sup>60</sup>CC<sup>61</sup>CC<sup>62</sup>CC<sup>63</sup>CC<sup>64</sup>CC<sup>65</sup>CC<sup>66</sup>CC<sup>67</sup>CC<sup>68</sup>CC<sup>69</sup>CC<sup>70</sup>CC<sup>71</sup>CC<sup>72</sup>CC<sup>73</sup>CC<sup>74</sup>CC<sup>75</sup>CC<sup>76</sup>CC<sup>77</sup>CC<sup>78</sup>CC<sup>79</sup>CC<sup>80</sup>CC<sup>81</sup>CC<sup>82</sup>CC<sup>83</sup>CC<sup>84</sup>CC<sup>85</sup>CC<sup>86</sup>CC<sup>87</sup>CC<sup>88</sup>CC<sup>89</sup>CC<sup>90</sup>CC<sup>91</sup>CC<sup>92</sup>CC<sup>93</sup>CC<sup>94</sup>CC<sup>95</sup>CC<sup>96</sup>CC<sup>97</sup>CC<sup>98</sup>CC<sup>99</sup>CC<sup>100</sup>CC<sup>101</sup>CC<sup>102</sup>CC<sup>103</sup>CC<sup>104</sup>CC<sup>105</sup>CC<sup>106</sup>CC<sup>107</sup>CC<sup>108</sup>CC<sup>109</sup>CC<sup>110</sup>CC<sup>111</sup>CC<sup>112</sup>CC<sup>113</sup>CC<sup>114</sup>CC<sup>115</sup>CC<sup>116</sup>CC<sup>117</sup>CC<sup>118</sup>CC<sup>119</sup>CC<sup>120</sup>CC<sup>121</sup>CC<sup>122</sup>CC<sup>123</sup>CC<sup>124</sup>CC<sup>125</sup>CC<sup>126</sup>CC<sup>127</sup>CC<sup>128</sup>CC<sup>129</sup>CC<sup>130</sup>CC<sup>131</sup>CC<sup>132</sup>CC<sup>133</sup>CC<sup>134</sup>CC<sup>135</sup>CC<sup>136</sup>CC<sup>137</sup>CC<sup>138</sup>CC<sup>139</sup>CC<sup>140</sup>CC<sup>141</sup>CC<sup>142</sup>CC<sup>143</sup>CC<sup>144</sup>CC<sup>145</sup>CC<sup>146</sup>CC<sup>147</sup>CC<sup>148</sup>CC<sup>149</sup>CC<sup>150</sup>CC<sup>151</sup>CC<sup>152</sup>CC<sup>153</sup>CC<sup>154</sup>CC<sup>155</sup>CC<sup>156</sup>CC<sup>157</sup>CC<sup>158</sup>CC<sup>159</sup>CC<sup>160</sup>CC<sup>161</sup>CC<sup>162</sup>CC<sup>163</sup>CC<sup>164</sup>CC<sup>165</sup>CC<sup>166</sup>CC<sup>167</sup>CC<sup>168</sup>CC<sup>169</sup>CC<sup>170</sup>CC<sup>171</sup>CC<sup>172</sup>CC<sup>173</sup>CC<sup>174</sup>CC<sup>175</sup>CC<sup>176</sup>CC<sup>177</sup>CC<sup>178</sup>CC<sup>179</sup>CC<sup>180</sup>CC<sup>181</sup>CC<sup>182</sup>CC<sup>183</sup>CC<sup>184</sup>CC<sup>185</sup>CC<sup>186</sup>CC<sup>187</sup>CC<sup>188</sup>CC<sup>189</sup>CC<sup>190</sup>CC<sup>191</sup>CC<sup>192</sup>CC<sup>193</sup>CC<sup>194</sup>CC<sup>195</sup>CC<sup>196</sup>CC<sup>197</sup>CC<sup>198</sup>CC<sup>199</sup>CC<sup>200</sup>CC<sup>201</sup>CC<sup>202</sup>CC<sup>203</sup>CC<sup>204</sup>CC<sup>205</sup>CC<sup>206</sup>CC<sup>207</sup>CC<sup>208</sup>CC<sup>209</sup>CC<sup>210</sup>CC<sup>211</sup>CC<sup>212</sup>CC<sup>213</sup>CC<sup>214</sup>CC<sup>215</sup>CC<sup>216</sup>CC<sup>217</sup>CC<sup>218</sup>CC<sup>219</sup>CC<sup>220</sup>CC<sup>221</sup>CC<sup>222</sup>CC<sup>223</sup>CC<sup>224</sup>CC<sup>225</sup>CC<sup>226</sup>CC<sup>227</sup>CC<sup>228</sup>CC<sup>229</sup>CC<sup>230</sup>CC<sup>231</sup>CC<sup>232</sup>CC<sup>233</sup>CC<sup>234</sup>CC<sup>235</sup>CC<sup>236</sup>CC<sup>237</sup>CC<sup>238</sup>CC<sup>239</sup>CC<sup>240</sup>CC<sup>241</sup>CC<sup>242</sup>CC<sup>243</sup>CC<sup>244</sup>CC<sup>245</sup>CC<sup>246</sup>CC<sup>247</sup>CC<sup>248</sup>CC<sup>249</sup>CC<sup>250</sup>CC<sup>251</sup>CC<sup>252</sup>CC<sup>253</sup>CC<sup>254</sup>CC<sup>255</sup>CC<sup>256</sup>CC<sup>257</sup>CC<sup>258</sup>CC<sup>259</sup>CC<sup>260</sup>CC<sup>261</sup>CC<sup>262</sup>CC<sup>263</sup>CC<sup>264</sup>CC<sup>265</sup>CC<sup>266</sup>CC<sup>267</sup>CC<sup>268</sup>CC<sup>269</sup>CC<sup>270</sup>CC<sup>271</sup>CC<sup>272</sup>CC<sup>273</sup>CC<sup>274</sup>CC<sup>275</sup>CC<sup>276</sup>CC<sup>277</sup>CC<sup>278</sup>CC<sup>279</sup>CC<sup>280</sup>CC<sup>281</sup>CC<sup>282</sup>CC<sup>283</sup>CC<sup>284</sup>CC<sup>285</sup>CC<sup>286</sup>CC<sup>287</sup>CC<sup>288</sup>CC<sup>289</sup>CC<sup>290</sup>CC<sup>291</sup>CC<sup>292</sup>CC<sup>293</sup>CC<sup>294</sup>CC<sup>295</sup>CC<sup>296</sup>CC<sup>297</sup>CC<sup>298</sup>CC<sup>299</sup>CC<sup>300</sup>CC<sup>301</sup>CC<sup>302</sup>CC<sup>303</sup>CC<sup>304</sup>CC<sup>305</sup>CC<sup>306</sup>CC<sup>307</sup>CC<sup>308</sup>CC<sup>309</sup>CC<sup>310</sup>CC<sup>311</sup>CC<sup>312</sup>CC<sup>313</sup>CC<sup>314</sup>CC<sup>315</sup>CC<sup>316</sup>CC<sup>317</sup>CC<sup>318</sup>CC<sup>319</sup>CC<sup>320</sup>CC<sup>321</sup>CC<sup>322</sup>CC<sup>323</sup>CC<sup>324</sup>CC<sup>325</sup>CC<sup>326</sup>CC<sup>327</sup>CC<sup>328</sup>CC<sup>329</sup>CC<sup>330</sup>CC<sup>331</sup>CC<sup>332</sup>CC<sup>333</sup>CC<sup>334</sup>CC<sup>335</sup>CC<sup>336</sup>CC<sup>337</sup>CC<sup>338</sup>CC<sup>339</sup>CC<sup>340</sup>CC<sup>341</sup>CC<sup>342</sup>CC<sup>343</sup>CC<sup>344</sup>CC<sup>345</sup>CC<sup>346</sup>CC<sup>347</sup>CC<sup>348</sup>CC<sup>349</sup>CC<sup>350</sup>CC<sup>351</sup>CC<sup>352</sup>CC<sup>353</sup>CC<sup>354</sup>CC<sup>355</sup>CC<sup>356</sup>CC<sup>357</sup>CC<sup>358</sup>CC<sup>359</sup>CC<sup>360</sup>CC<sup>361</sup>CC<sup>362</sup>CC<sup>363</sup>CC<sup>364</sup>CC<sup>365</sup>CC<sup>366</sup>CC<sup>367</sup>CC<sup>368</sup>CC<sup>369</sup>CC<sup>370</sup>CC<sup>371</sup>CC<sup>372</sup>CC<sup>373</sup>CC<sup>374</sup>CC<sup>375</sup>CC<sup>376</sup>CC<sup>377</sup>CC<sup>378</sup>CC<sup>379</sup>CC<sup>380</sup>CC<sup>381</sup>CC<sup>382</sup>CC<sup>383</sup>CC<sup>384</sup>CC<sup>385</sup>CC<sup>386</sup>CC<sup>387</sup>CC<sup>388</sup>CC<sup>389</sup>CC<sup>390</sup>CC<sup>391</sup>CC<sup>392</sup>CC<sup>393</sup>CC<sup>394</sup>CC<sup>395</sup>CC<sup>396</sup>CC<sup>397</sup>CC<sup>398</sup>CC<sup>399</sup>CC<sup>400</sup>CC<sup>401</sup>CC<sup>402</sup>CC<sup>403</sup>CC<sup>404</sup>CC<sup>405</sup>CC<sup>406</sup>CC<sup>407</sup>CC<sup>408</sup>CC<sup>409</sup>CC<sup>410</sup>CC<sup>411</sup>CC<sup>412</sup>CC<sup>413</sup>CC<sup>414</sup>CC<sup>415</sup>CC<sup>416</sup>CC<sup>417</sup>CC<sup>418</sup>CC<sup>419</sup>CC<sup>420</sup>CC<sup>421</sup>CC<sup>422</sup>CC<sup>423</sup>CC<sup>424</sup>CC<sup>425</sup>CC<sup>426</sup>CC<sup>427</sup>CC<sup>428</sup>CC<sup>429</sup>CC<sup>430</sup>CC<sup>431</sup>CC<sup>432</sup>CC<sup>433</sup>CC<sup>434</sup>CC<sup>435</sup>CC<sup>436</sup>CC<sup>437</sup>CC<sup>438</sup>CC<sup>439</sup>CC<sup>440</sup>CC<sup>441</sup>CC<sup>442</sup>CC<sup>443</sup>CC<sup>444</sup>CC<sup>445</sup>CC<sup>446</sup>CC<sup>447</sup>CC<sup>448</sup>CC<sup>449</sup>CC<sup>450</sup>CC<sup>451</sup>CC<sup>452</sup>CC<sup>453</sup>CC<sup>454</sup>CC<sup>455</sup>CC<sup>456</sup>CC<sup>457</sup>CC<sup>458</sup>CC<sup>459</sup>CC<sup>460</sup>CC<sup>461</sup>CC<sup>462</sup>CC<sup>463</sup>CC<sup>464</sup>CC<sup>465</sup>CC<sup>466</sup>CC<sup>467</sup>CC<sup>468</sup>CC<sup>469</sup>CC<sup>470</sup>CC<sup>471</sup>CC<sup>472</sup>CC<sup>473</sup>CC<sup>474</sup>CC<sup>475</sup>CC<sup>476</sup>CC<sup>477</sup>CC<sup>478</sup>CC<sup>479</sup>CC<sup>480</sup>CC<sup>481</sup>CC<sup>482</sup>CC<sup>483</sup>CC<sup>484</sup>CC<sup>485</sup>CC<sup>486</sup>CC<sup>487</sup>CC<sup>488</sup>CC<sup>489</sup>CC<sup>490</sup>CC<sup>491</sup>CC<sup>492</sup>CC<sup>493</sup>CC<sup>494</sup>CC<sup>495</sup>CC<sup>496</sup>CC<sup>497</sup>CC<sup>498</sup>CC<sup>499</sup>CC<sup>500</sup>CC<sup>501</sup>CC<sup>502</sup>CC<sup>503</sup>CC<sup>504</sup>CC<sup>505</sup>CC<sup>506</sup>CC<sup>507</sup>CC<sup>508</sup>CC<sup>509</sup>CC<sup>510</sup>CC<sup>511</sup>CC<sup>512</sup>CC<sup>513</sup>CC<sup>514</sup>CC<sup>515</sup>CC<sup>516</sup>CC<sup>517</sup>CC<sup>518</sup>CC<sup>519</sup>CC<sup>520</sup>CC<sup>521</sup>CC<sup>522</sup>CC<sup>523</sup>CC<sup>524</sup>CC<sup>525</sup>CC<sup>526</sup>CC<sup>527</sup>CC<sup>528</sup>CC<sup>529</sup>CC<sup>530</sup>CC<sup>531</sup>CC<sup>532</sup>CC<sup>533</sup>CC<sup>534</sup>CC<sup>535</sup>CC<sup>536</sup>CC<sup>537</sup>CC<sup>538</sup>CC<sup>539</sup>CC<sup>540</sup>CC<sup>541</sup>CC<sup>542</sup>CC<sup>543</sup>CC<sup>544</sup>CC<sup>545</sup>CC<sup>546</sup>CC<sup>547</sup>CC<sup>548</sup>CC<sup>549</sup>CC<sup>550</sup>CC<sup>551</sup>CC<sup>552</sup>CC<sup>553</sup>CC<sup>554</sup>CC<sup>555</sup>CC<sup>556</sup>CC<sup>557</sup>CC<sup>558</sup>CC<sup>559</sup>CC<sup>560</sup>CC<sup>561</sup>CC<sup>562</sup>CC<sup>563</sup>CC<sup>564</sup>CC<sup>565</sup>CC<sup>566</sup>CC<sup>567</sup>CC<sup>568</sup>CC<sup>569</sup>CC<sup>570</sup>CC<sup>571</sup>CC<sup>572</sup>CC<sup>573</sup>CC<sup>574</sup>CC<sup>575</sup>CC<sup>576</sup>CC<sup>577</sup>CC<sup>578</sup>CC<sup>579</sup>CC<sup>580</sup>CC<sup>581</sup>CC<sup>582</sup>CC<sup>583</sup>CC<sup>584</sup>CC<sup>585</sup>CC<sup>586</sup>CC<sup>587</sup>CC<sup>588</sup>CC<sup>589</sup>CC<sup>590</sup>CC<sup>591</sup>CC<sup>592</sup>CC<sup>593</sup>CC<sup>594</sup>CC<sup>595</sup>CC<sup>596</sup>CC<sup>597</sup>CC<sup>598</sup>CC<sup>599</sup>CC<sup>600</sup>CC<sup>601</sup>CC<sup>602</sup>CC<sup>603</sup>CC<sup>604</sup>CC<sup>605</sup>CC<sup>606</sup>CC<sup>607</sup>CC<sup>608</sup>CC<sup>609</sup>CC<sup>610</sup>CC<sup>611</sup>CC<sup>612</sup>CC<sup>613</sup>CC<sup>614</sup>CC<sup>615</sup>CC<sup>616</sup>CC<sup>617</sup>CC<sup>618</sup>CC<sup>619</sup>CC<sup>620</sup>CC<sup>621</sup>CC<sup>622</sup>CC<sup>623</sup>CC<sup>624</sup>CC<sup>625</sup>CC<sup>626</sup>CC<sup>627</sup>CC<sup>628</sup>CC<sup>629</sup>CC<sup>630</sup>CC<sup>631</sup>CC<sup>632</sup>CC<sup>633</sup>CC<sup>634</sup>CC<sup>635</sup>CC<sup>636</sup>CC<sup>637</sup>CC<sup>638</sup>CC<sup>639</sup>CC<sup>640</sup>CC<sup>641</sup>CC<sup>642</sup>CC<sup>643</sup>CC<sup>644</sup>CC<sup>645</sup>CC<sup>646</sup>CC<sup>647</sup>CC<sup>648</sup>CC<sup>649</sup>CC<sup>650</sup>CC<sup>651</sup>CC<sup>652</sup>CC<sup>653</sup>CC<sup>654</sup>CC<sup>655</sup>CC<sup>656</sup>CC<sup>657</sup>CC<sup>658</sup>CC<sup>659</sup>CC<sup>660</sup>CC<sup>661</sup>CC<sup>662</sup>CC<sup>663</sup>CC<sup>664</sup>CC<sup>665</sup>CC<sup>666</sup>CC<sup>667</sup>CC<sup>668</sup>CC<sup>669</sup>CC<sup>670</sup>CC<sup>671</sup>CC<sup>672</sup>CC<sup>673</sup>CC<sup>674</sup>CC<sup>675</sup>CC<sup>676</sup>CC<sup>677</sup>CC<sup>678</sup>CC<sup>679</sup>CC<sup>680</sup>CC<sup>681</sup>CC<sup>682</sup>CC<sup>683</sup>CC<sup>684</sup>CC<sup>685</sup>CC<sup>686</sup>CC<sup>687</sup>CC<sup>688</sup>CC<sup>689</sup>CC<sup>690</sup>CC<sup>691</sup>CC<sup>692</sup>CC<sup>693</sup>CC<sup>694</sup>CC<sup>695</sup>CC<sup>696</sup>CC<sup>697</sup>CC<sup>698</sup>CC<sup>699</sup>CC<sup>700</sup>CC<sup>701</sup>CC<sup>702</sup>CC<sup>703</sup>CC<sup>704</sup>CC<sup>705</sup>CC<sup>706</sup>CC<sup>707</sup>CC<sup>708</sup>CC<sup>709</sup>CC<sup>710</sup>CC<sup>711</sup>CC<sup>712</sup>CC<sup>713</sup>CC<sup>714</sup>CC<sup>715</sup>CC<sup>716</sup>CC<sup>717</sup>CC<sup>718</sup>CC<sup>719</sup>CC<sup>720</sup>CC<sup>721</sup>CC<sup>722</sup>CC<sup>723</sup>CC<sup>724</sup>CC<sup>725</sup>CC<sup>726</sup>CC<sup>727</sup>CC<sup>728</sup>CC<sup>729</sup>CC<sup>730</sup>CC<sup>731</sup>CC<sup>732</sup>CC<sup>733</sup>CC<sup>734</sup>CC<sup>735</sup>CC<sup>736</sup>CC<sup>737</sup>CC<sup>738</sup>CC<sup>739</sup>CC<sup>740</sup>CC<sup>741</sup>CC<sup>742</sup>CC<sup>743</sup>CC<sup>744</sup>CC<sup>745</sup>CC<sup>746</sup>CC<sup>747</sup>CC<sup>748</sup>CC<sup>749</sup>CC<sup>750</sup>CC<sup>751</sup>CC<sup>752</sup>CC<sup>753</sup>CC<sup>754</sup>CC<sup>755</sup>CC<sup>756</sup>CC<sup>757</sup>CC<sup>758</sup>CC<sup>759</sup>CC<sup>760</sup>CC<sup>761</sup>CC<sup>762</sup>CC<sup>763</sup>CC<sup>764</sup>CC<sup>765</sup>CC<sup>766</sup>CC<sup>767</sup>CC<sup>768</sup>CC<sup>769</sup>CC<sup>770</sup>CC<sup>771</sup>CC<sup>772</sup>CC<sup>773</sup>CC<sup>774</sup>CC<sup>775</sup>CC<sup>776</sup>CC<sup>777</sup>CC<sup>778</sup>CC<sup>779</sup>CC<sup>780</sup>CC<sup>781</sup>CC<sup>782</sup>CC<sup>783</sup>CC<sup>784</sup>CC<sup>785</sup>CC<sup>786</sup>CC<sup>787</sup>CC<sup>788</sup>CC<sup>789</sup>CC<sup>790</sup>CC<sup>791</sup>CC<sup>792</sup>CC<sup>793</sup>CC<sup>794</sup>CC<sup>795</sup>CC<sup>796</sup>CC<sup>797</sup>CC<sup>798</sup>CC<sup>799</sup>CC<sup>800</sup>CC<sup>801</sup>CC<sup>802</sup>CC<sup>803</sup>CC<sup>804</sup>CC<sup>805</sup>CC<sup>806</sup>CC<sup>807</sup>CC<sup>808</sup>CC<sup>809</sup>CC<sup>810</sup>CC<sup>811</sup>CC<sup>812</sup>CC<sup>813</sup>CC<sup>814</sup>CC<sup>815</sup>CC<sup>816</sup>CC<sup>817</sup>CC<sup>818</sup>CC<sup>819</sup>CC<sup>820</sup>CC<sup>821</sup>CC<sup>822</sup>CC<sup>823</sup>CC<sup>824</sup>CC<sup>825</sup>CC<sup>826</sup>CC<sup>827</sup>CC<sup>828</sup>CC<sup>829</sup>CC<sup>830</sup>CC<sup>831</sup>CC<sup>832</sup>CC<sup>833</sup>CC<sup>834</sup>CC<sup>835</sup>CC<sup>836</sup>CC<sup>837</sup>CC<sup>838</sup>CC<sup>839</sup>CC<sup>840</sup>CC<sup>841</sup>CC<sup>842</sup>CC<sup>843</sup>CC<sup>844</sup>CC<sup>845</sup>CC<sup>846</sup>CC<sup>847</sup>CC<sup>848</sup>CC<sup>849</sup>CC<sup>850</sup>CC<sup>851</sup>CC<sup>852</sup>CC<sup>853</sup>CC<sup>854</sup>CC<sup>855</sup>CC<sup>856</sup>CC<sup>857</sup>CC<sup>858</sup>CC<sup>859</sup>CC<sup>860</sup>CC<sup>861</sup>CC<sup>862</sup>CC<sup>863</sup>CC<sup>864</sup>CC<sup>865</sup>CC<sup>866</sup>CC<sup>867</sup>CC<sup>868</sup>CC<sup>869</sup>CC<sup>870</sup>CC<sup>871</sup>CC<sup>872</sup>CC<sup>873</sup>CC<sup>874</sup>CC<sup>875</sup>CC<sup>876</sup>CC<sup>877</sup>CC<sup>878</sup>CC<sup>879</sup>CC<sup>880</sup>CC<sup>881</sup>CC<sup>882</sup>CC<sup>883</sup>CC<sup>884</sup>CC<sup>885</sup>CC<sup>886</sup>CC<sup>887</sup>CC<sup>888</sup>CC<sup>889</sup>CC<sup>890</sup>CC<sup>891</sup>CC<sup>892</sup>CC<sup>893</sup>CC<sup>894</sup>CC<sup>895</sup>CC<sup>896</sup>CC<sup>897</sup>CC<sup>898</sup>CC<sup>899</sup>CC<sup>900</sup>CC<sup>901</sup>CC<sup>902</sup>CC<sup>903</sup>CC<sup>904</sup>CC<sup>905</sup>CC<sup>906</sup>CC<sup>907</sup>CC<sup>908</sup>CC<sup>909</sup>CC<sup>910</sup>CC<sup>911</sup>CC<sup>912</sup>CC<sup>913</sup>CC<sup>914</sup>CC<sup>915</sup>CC<sup>916</sup>CC<sup>917</sup>CC<sup>918</sup>CC<sup>919</sup>CC<sup>920</sup>CC<sup>921</sup>CC<sup>922</sup>CC<sup>923</sup>CC<sup>924</sup>CC<sup>925</sup>CC<sup>926</sup>CC<sup>927</sup>CC<sup>928</sup>CC<sup>929</sup>CC<sup>930</sup>CC<sup>931</sup>CC<sup>932</sup>CC<sup>933</sup>CC<sup>934</sup>CC<sup>935</sup>CC<sup>936</sup>CC<sup>937</sup>CC<sup>938</sup>CC<sup>939</sup>CC<sup>940</sup>CC<sup>941</sup>CC<sup>942</sup>CC<sup>943</sup>CC<sup>944</sup>CC<sup>945</sup>CC<sup>946</sup>CC<sup>947</sup>CC<sup>948</sup>CC<sup>949</sup>CC<sup>950</sup>CC<sup>951</sup>CC<sup>952</sup>CC<sup>953</sup>CC<sup>954</sup>CC<sup>955</sup>CC<sup>956</sup>CC<sup>957</sup>CC<sup>958</sup>CC<sup>959</sup>CC<sup>960</sup>CC<sup>961</sup>CC<sup>962</sup>CC<sup>963</sup>CC<sup>964</sup>CC<sup>965</sup>CC<sup>966</sup>CC<sup>967</sup>CC<sup>968</sup>CC<sup>969</sup>CC<sup>970</sup>CC<sup>971</sup>CC<sup>972</sup>CC<sup>973</sup>CC<sup>974</sup>CC<sup>975</sup>CC<sup>976</sup>CC<sup>977</sup>CC<sup>978</sup>CC<sup>979</sup>CC<sup>980</sup>CC<sup>981</sup>CC<sup>982</sup>CC<sup>983</sup>CC<sup>984</sup>CC<sup>985</sup>CC<sup>986</sup>CC<sup>987</sup>CC<sup>988</sup>CC<sup>989</sup>CC<sup>990</sup>CC<sup>991</sup>CC<sup>992</sup>CC<sup>993</sup>CC<sup>994</sup>CC<sup>995</sup>CC<sup>996</sup>CC<sup>997</sup>CC<sup>998</sup>CC<sup>999</sup>CC<sup>1000</sup>CC<sup>1001</sup>CC<sup>1002</sup>CC<sup>1003</sup>CC<sup>1004</sup>CC<sup>1005</sup>CC<sup>1006</sup>CC<sup>1007</sup>CC<sup>1008</sup>CC<sup>1009</sup>CC<sup>1010</sup>CC<sup>1011</sup>CC<sup>1012</sup>CC<sup>1013</sup>CC<sup>1014</sup>CC<sup>1015</sup>CC<sup>1016</sup>CC<sup>1017</sup>CC<sup>1018</sup>CC<sup>1019</sup>CC<sup>1020</sup>CC<sup>1021</sup>CC<sup>1022</sup>CC<sup>1023</sup>CC<sup>1024</sup>CC<sup>1025</sup>CC<sup>1026</sup>CC<sup>1027</sup>CC<sup>1028</sup>CC<sup>1029</sup>CC<sup>1030</sup>CC<sup>1031</sup>CC<sup>1032</sup>CC<sup>1033</sup>CC<sup>1034</sup>CC<sup>1035</sup>CC<sup>1036</sup>CC<sup>1037</sup>CC<sup>1038</sup>CC<sup>1039</sup>CC<sup>1040</sup>CC<sup>1041</sup>CC<sup>1042</sup>CC<sup>1043</sup>CC<sup>1044</sup>CC<sup>1045</sup>CC<sup>1046</sup>CC<sup>1047</sup>CC<sup>1048</sup>CC<sup>1049</sup>CC<sup>1050</sup>CC<sup>1051</sup>CC<sup>1052</sup>CC<sup>1053</sup>CC<sup>1054</sup>CC<sup>1055</sup>CC<sup>1056</sup>CC<sup>1057</sup>CC<sup>1058</sup>CC<sup>1059</sup>CC<sup>1060</sup>CC<sup>1061</sup>CC<sup>1062</sup>CC<sup>1063</sup>CC<sup>1064</sup>CC<sup>1065</sup>CC<sup>1066</sup>CC<sup>1067</sup>CC<sup>1068</sup>CC<sup>1069</sup>CC<sup>1070</sup>CC<sup>1071</sup>CC<sup>1072</sup>CC<sup>1073</sup>CC<sup>1074</sup>CC<sup>1075</sup>CC<sup>1076</sup>CC<sup>1077</sup>CC<sup>1078</sup>CC<sup>1079</sup>CC<sup>1080</sup>CC<sup>1081</sup>CC<sup>1082</sup>CC<sup>1083</sup>CC<sup>1084</sup>CC<sup>1085</sup>CC<sup>1086</sup>CC<sup>1087</sup>CC<sup>1088</sup>CC<sup>1089</sup>CC<sup>1090</sup>CC<sup>1091</sup>CC<sup>1092</sup>CC<sup>1093</sup>CC<sup>1094</sup>CC<sup>1095</sup>CC<sup>1096</sup>CC<sup>1097</sup>CC<sup>1098</sup>CC<sup>1099</sup>CC<sup>1100</sup>CC<sup>1101</sup>CC<sup>1102</sup>CC<sup>1103</sup>CC<sup>1104</sup>CC<sup>1105</sup>CC<sup>1106</sup>CC<sup>1107</sup>CC<sup>1108</sup>CC<sup>1109</sup>CC<sup>1110</sup>CC<sup>1111</sup>CC<sup>1112</sup>CC<sup>1113</sup>CC<sup>1114</sup>CC<sup>1115</sup>CC<sup>1116</sup>CC<sup>1117</sup>CC<sup>1118</sup>CC<sup>1119</sup>CC<sup>1120</sup>CC<sup>1121</sup>CC<sup>1122</sup>CC<sup>1123</sup>CC<sup>1124</sup>CC<sup>1125</sup>CC<sup>1126</sup>CC<sup>1127</sup>CC<sup>1128</sup>CC<sup>1129</sup>CC<sup>1130</sup>CC<sup>1131</sup>CC<sup>1132</sup>CC<sup>1133</sup>CC<sup>1134</sup>CC<sup>1135</sup>CC<sup>1136</sup>CC<sup>1137</sup>CC<sup>1138</sup>CC<sup>1139</sup>CC<sup>1140</sup>CC<sup>1141</sup>CC<sup>1142</sup>CC<sup>1143</sup>CC<sup>1144</sup>CC<sup>1145</sup>

130/178

Fig. 88C

**35. 2003 A1.anc nef. PEP**

MGGKWSKSSIVGWPEVRRMRRTPPAAKGVGAVSQDLDLKHGAVTSSNTAANNPGCAWLEAQEEEEVGFVVRPQVPLRPMTYKGAFDLISHFLK  
 EKGGLDGLIYSKKRQEIILDLWVYHTQGYFPDWNQYTPGPGIRYPLTFGWCFKLVVDPVPAEVEEATEGENNSLLHPICOHGMDDDEEREVILMWK  
 FDSRLALKHRARELHPEFYKDC\$

Fig. 88D

**2003 A1.anc nef. OPT**

ATGGCGGGCAAGTGGTCCAAGTCCCTCCATCGTGGGCTGGCCCGAGGTGCGGAGCGCATGCGCCGCACCCCGCCCGCCCGCCCAAGGCGGTGGG  
 CGCCGTGTCCCAGGACCTGGACAAGCACGGCCCGTGTACTCTCCAACACCGCCGCAACACCCGGCTGCGCCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGGTGGCTTCCCCTGGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGGCCCTTCGACCTGTCCACTTCCTGAAG  
 GAGAAGGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGGCGCCGGCATCCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTCCCGTGGACCCCGCCGAGGTGG  
 AGGAGGCCACCGAGGGCGAGAACAACCTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGACGAGGAGCGCGGAGGTGCTGATGTGGAAAG  
 TTCGACTCCCCTGGCCCTGAAGCACCGGCGGAGCTGCACCCCGAGTCTACAAGGACTGCTAA

Fig. 89A

**36. 2003 CON A2 nef. PEP**

MGGKWSKSSIVGWP<sup>~</sup>PAIRERMRKRTPPAAEGVAVSQDLATRGAVTSSNTAATNPDCAWLEAQEEEEVGFVVRPQVPLRPMTFKGAFDLISHFL  
 KEKGGLDGLIYSOKRQDILLDLWVYHTQGYFPDWNQYTPGPGTRYPLTFGWCFKLVVDPSEVEEATEGENNSLLHPICOHGIEDPEREVLRW  
 KFDSRLALRHRARELHPEFYKDC\$

Fig. 89B

**2003 CON A2 nef. OPT**

ATGGCGGGCAAGTGGTCCAAGTCCCTCCATCGTGGGCTGGCCCGCCCATCCGCGGAGCGCATGCGCAAGCGCACCCCGCCCGCCCGGAGGCGGT  
 GGGCGCGGTGTCCCAGGACCTGGCCACCCGGCGGCGCGTGACCTCTCCAACACCCGCGCCACCAACCCGACTGCGCTGGCTGGAGGCCCC  
 AGGAGGAGGAGGAGGTGGCTTCCCCTGGCCCGCCAGGTGCCCTTCGACCTGACCTTCAAGGGCGCTTCGACCTGTCCACTTCCTG  
 AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCCAGAAAGCGCCAGGACATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCC  
 CGACTGGCAGAACTACACCCCGGCGCCGGCATCCCTGACCTTCGGCTGGTGTCAAGCTGGTCCCGTGGACCCCGCCGAGGTGGCTGGCTGG  
 TGGAGGAGGCCACCGAGGGCGGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATCGAGGACCCCGGAGCGCGGAGGTGCTGGCTGG  
 AAGTTCGACTCCCCTGGCCCTGCGCCACCCGGCGGAGTGCACCCCGAGTCTACAAGGACTGCTAA

**Fig. 90A**

**37. 2003 CON\_B nef .PEP**  
 MGKWSKRSVVGWPTVRRMRRAEPAADGVAVSRDLKHKGAISSNTAANNADCAWLEAQEEEEVGFVPRPQVPLRPMTYK GALDLSHFLK  
 EKGGLEGLIYSQKRQDILLDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPEPEKVEEANEGENNSLLHHPMSLHGMDDPEREVLVWK  
 FDSRLAFHMHMARELHPEYYKDC\$

**Fig. 90B**

**2003 CON-B nef .OPT**  
 ATGGCGGCAAGTGGTCCAAGCGCTCCGTGGTGGCGGAGCGCATGGCGCGGAGCCCGCCGCGCCGCGGACGGCGTGGG  
 CGCCGTGTCCCGGACCTGGAGAAGCACGGCGCATACCTCCAAACACCGCCCAACAACGCCGACTGGCCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCCCAGGTGCCCCATGACCTACAAGGGCCCCGTGACCTGTCCACTTCCCTGAAG  
 GAGAAGGCGGCTGGAGGCGCTGATCTACTCCAGAAGCGCCAGGACATCTGGACCTGTGGGTGTACCAACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGGATCCGTAACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGGCCGTGGAGCCCGAGAAGGTGG  
 AGGAGGCCAACGAGGGCGGAGAACAATCCCTGTGCACCCCATGTCCCTGCACGGCATGGACGACCCCGAGCGCGGAGGTGCTGGTGTGGAAG  
 TTCGACTCCCGCCTGGCCTTCCACCACATGGCCCCGAGCTGCACCCGAGTACTACAAGGACTGCTAA

**Fig. 90C**

**38. 2003 B.anc nef .PEP**  
 MGKWSKSSMGGWPAVRRMRRAEPAADGVAVSRDLKHKGAISSNTAATNADCAWLEAQEEEEVGFVPRPQVPLRPMTYKAALDLSHFLK  
 EKGGLEGLIYSQKRQDILLDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPEPEKVEEATEGENNSLLHHPMCQHGMDDPEKEVLVWK  
 FDSRLAFHMHMARELHPEYYKDC\$

**Fig. 90D**

**2003 B.anc nef .OPT**  
 ATGGCGGCAAGTGGTCCAAGTCCCTCCATGGGCGGCTGGCCCCCGCCGTGGCGGAGCGCATGAAGCGCGCCGAGCCCGCCGCGGACGGCGTGGG  
 CGCCGTGTCCCGGACCTGGAGAAGCACGGGCCATACCTCCAAACACCGCCCAACAACGCCGACTGGCCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCCCAGGTGCCCCATGACCTACAAGGGCCCCGTGACCTGTCCACTTCCCTGAAG  
 GAGAAGGCGGCTGGAGGCGCTGATCTACTCCAGAAGCGCCAGGACATCTGGACCTGTGGGTGTACCAACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGGATCCGTAACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGGCCGTGGAGCCCGAGAAGGTGG  
 AGGAGGCCAACGAGGGCGGAGAACAATCCCTGTGCACCCCATGTCCAGCACGGCATGGACGACCCCGAGGAGGAGGTGCTGGTGTGGAAG  
 TTCGACTCCCGCCTGGCCTTCCACCACATGGCCCCGAGCTGCACCCGAGTACTACAAGGACTGCTAA

131/178



132/178

**Fig. 91A**

**39. 2003 CON 02 AG nef . PEP**  
 MGKWSKSSIVGWP<sup>~</sup>PKVRERIRQT<sup>~</sup>PPAATGVGAASQDLDRHGAI<sup>~</sup>TSNSTAATNADCAWLEAQEEEEVGFV<sup>~</sup>RPQVPLRPMTYKAAVDL<sup>~</sup>SHFLK  
 EKGGLEGLIYSKKRQEI<sup>~</sup>LDLWYHTQGF<sup>~</sup>FDWQNYTPGPGTRFPLTFGWC<sup>~</sup>FKLVPM<sup>~</sup>DPAEVEEANE<sup>~</sup>GENNSLLHPI<sup>~</sup>CQHGMEDE<sup>~</sup>DREVLVWR  
 FDSLAFK<sup>~</sup>HRARELHPEFYKDC\$

**Fig. 91B**

**2003 CON 02 AG nef . OPT**  
 ATGGCGGCAAGTGGTCCAAGTCCATCGTGGCTGGCCCAAGGTGCGGAGCGCATCCGCCAGACCC<sup>~</sup>CCCGCCACCCGGCGTGGG  
 CGCCGCTCCAGGACCTGGACCGGCCATCACCTCCTCAACACCCGCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCC<sup>~</sup>CCAGGTGCCCCATGACCTACAAGSCCGCGTGACCTGTCCCACTTCCCTGAAG  
 GAGAAGGGCCCTGGAGGCTGATCTACTCCAAGAAGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTTCTTCC<sup>~</sup>CCGA  
 CTGGCAGAACTACACCCCGGCCACCCGCTTCC<sup>~</sup>CCCTGACCTTCGGCTGGTCTCAAGCTGGTGCCTATGGACCCCGCGAGGTGG  
 AGGAGCCAAACGAGGGCGAGAACTCCCTGTCTGCACCCCATCTGCCAGACCGGCATGGAGACCGCGAGGTGCTGTTGTGGCGC  
 TTCGACTCCTCCCTGGCCCTCAAGCACCGCGCCCGGAGCTGCACCCCGAGTCTACAAGGACTGCTAA

**Fig. 92A**

**40. 2003 CON C nef . PEP**  
 MGKWSKSSIVGWP<sup>~</sup>PAVRERIRRTEPAAEGVGAASQDLDKHGAL<sup>~</sup>TSSNTATNNADCAWLEAQEEEEVGFV<sup>~</sup>RPQVPLRPMTYKAAFDL<sup>~</sup>SFFL  
 KEKGGLEGLIYSKKRQEI<sup>~</sup>LDLWYHTQGY<sup>~</sup>FPDWQNYTPGPGVRYPLTFGWC<sup>~</sup>FKLV<sup>~</sup>VPDPREVEEANE<sup>~</sup>GENNCLLHPMSQH<sup>~</sup>GMEDE<sup>~</sup>DREVLK  
 KFD<sup>~</sup>SHLARRRH<sup>~</sup>MARELHPEYKDC\$

**Fig. 92B**

**2003 CON C nef . OPT**  
 ATGGCGGCAAGTGGTCCAAGTCCATCGTGGCTGGCCCGCGGAGCGCATCCGCCGCA<sup>~</sup>CCGAGCCCGCCCGCGGCGGTGGG  
 CGCCGCTCCAGGACCTGGACAGCACCGGCCCTGACCTCTCAACACCCGCCACCAACACCGGACTGCGCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCC<sup>~</sup>CCAGGTGCCCCATGACCTACAAGSCCGCGTTCCGACCTGTCCCTTCTTCCCTG  
 AAGGAGAAGGGCGGCTGGAGGCTGATCTACTCCAAGAAGCCAGGAGATCCTGGA<sup>~</sup>CCTGTGGGTGTACCACACCCAGGGCTACTCC  
 CGACTGGCAGAACTACACCCCGGCCGTGCGTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGCCTGGACCCCGCGAGG  
 TGGAGGAGGCCAAACGAGGGCGAGAACTGCCTGTGCACCCCATGTCCACACCGGCATGGAGGACCGCGAGGTGCTGAAGTGG  
 AAGTTCGACTCCCACCTGGCCCGCCACATGGCCCGCGGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

133/178

**Fig. 92C****41. 2003 C. anc nef . PEP**

MGGKWSKSSIVGWPAVRRERMRRTTEPAAEGVGAASQDLDLKHGALTSNTAANNADCAWLEAQEEEEVGFVVRPQVPLRPMTYKAAFDLSFFL  
 KEKGGLDGLIYSKKRQEIILDLWVYHTQGYFPDWNQYTPGPGVRYPLTFGWCFKLVVDPREVEEANEENCLLHPMSQHGMEDDEDREVLKW  
 KFDSLARRHMRARELHPEYKDC\$

**Fig. 92D****2003 C. anc nef .OPT**

ATGGCGGCAAGTGGTCCAAGTCCATCGTGGGTGGCCCGCCGTGCGGAGCGCATGGCCCGCACCGAGCCCGCCCGGAGGGCGGTGGG  
 CGCCGCTCCAGACTGGACACTGACCTCCTCAACACCGCCGCAACAACGCGACTGCGCCTGGCTGGAGGCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCCCGCCAGTGCCTGACCTACAAGGCCCTCGACCTGTCCCTTCTTCTG  
 AAGGAAAGGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCAGAGATCCTGGACCTGTGGGTGTACCACCCAGGCTACTTCC  
 CGACTGGCAGAACTACACCCCGCCCGGCTGCGTACCCCTGACCTTCGGCTGGTCAAGTGGTGGACCCCGCGGAGG  
 TGGAGAGGCCAACGAGGCGGAGAACACTGCTGTCACCCCATGTCCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTGG  
 AAGTTCGACTCCACCTGGCCCGCCACATGGCCCGCCGACTGACCCCGAGTACTACAAGGACTGCTAA

**Fig. 93A****42. 2003 CON D nef .PEP**

MGGKWSKSSIVGWPAIRERIRRTTEPAADGVGAVSRDLEKHGALTSNTAATNADCAWLEAQEEEEVGFVVRPQVPLRPMTYKAAFDLSHFL  
 KEKGGLEGLVWSQKRQEIILDLWVYNTQGFDPDWNQYTPGPGIRYPLTFGWCFELVPDPEEVEEATEGENNCLLHPMCQHGMEDPEREVLW  
 RFNSRLAFEHKARVLLHPEFYKDC\$

**Fig. 93B****2003 CON D nef .OPT**

ATGGCGGCAAGTGGTCCAAGTCCATCGTGGGTGGCCCGCCATCGCGAGCGCATCCGCCGACCGAGCCCGCCCGGACGGCGGTGGG  
 CGCCGTGCCCGACTGGAGAAGCAGGCGCCATCACCCTCAACAACCGCCGCAACGCGACTGCGCCTGGCTGGAGGCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCCCGCCAGTGCCTGACCTACAAGGCCCGCCCTGGACCTGTCCACTTCCCTG  
 AAGGAAAGGGCGGCTGGAGGGCTTGGTCCAGAAAGCGCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGCTTCTTCCC  
 CGACTGGCAGAACTACACCCCGCCCGGCTGCGTACCCCTGACCTTCGGCTGGTGTGAGCTGGTCCCGTGGACCCCGAGGAGG  
 TGGAGGAGGCCAACGAGGCGGAGAACACTGCTGTCACCCCATGTGCCAGCACGGCATGGAGGACCCCGAGCGGAGGTGCTGATGTGG  
 CGCTCAACTCCCGCCCTGGAGCACAAAGCCCGCTGCTGACCCCGAGTCTACAAGGACTGCTAA

134/178

**Fig. 94A****43. 2003 CON F1 nef .PEP**

MGGKWSKSSIVG $\bar{W}$ PAVRERMRPTPPAAEGVAVSQDLERRGAISSNTGATNPDLAWLEAQEEEEVGFVPRPQVPLRPMTYKGAVDLSHFLK  
 EKGGLEGLIYSKKRQEIILDLWVYHTQGYFPDWQNYTPGIRYPLTFGWCFKLVVDPEEVEKANEENCLLHPMSQHGMEDREVLIIWK  
 FDSRLALRHIARERHPEFYQD\$

**Fig. 94B****2003 CON F1 nef .OPT**

ATGGCGGCAAGTGGTCCAAGTCCATCGTGGGCTGGCCCGCGGTGCGGAGCGCATGCGCCCAACCCCGCCCGCCCGCCCGCCCGCGGTGGG  
 CGCCGTGTCCAGGACCTGGAGCGCGGCGCATCACCTCCCAACACCCGCGCCACCAACCCCGACCTGGCCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGTGGCTTCCCGGTGCGCCCAAGTGCACCTAACAAGGGCCCGTGGACCTGTCCCACTTCCTGAAG  
 GAGAAGGCGGCTGGAGGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCCCGGCATCCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGCCTGGACCCCGGAGGAGGTGG  
 AGAAGGCCAACGAGGGCGGAGAACACTGCCTGTGCACCCCATGTCCAGCACGGCATGGAGGACGAGGACCCCGGAGGTGCTGATCTGGGAAG  
 TTCGACTCCCGCTGGCCCTGGCCACATCGCCCGGAGCGCCACCCCGAGTCTACCAGGACTAA

**Fig. 95A****44. 2003 CON F2 nef .PEP**

MGGKWSKSSIVG $\bar{W}$ PTIRERIRRTIPVAEGVAVSQDLDKHGAISSNTRATNADLAWLEAQEEDVEVGFVPRPQVPLRPMTYKAAFDLSHFLK  
 EKGGLEGLIYSKKRQEIILDLWVYHTQGYFPDWQNYTPGTRYPPLTFGWCFKLVVDPEEVEKANEENCLLHPMSLHGMEDREVLKWK  
 FDSRLALRHIARERHPEYYKD\$

**Fig. 95B****2003 CON F2 nef .OPT**

ATGGCGGCAAGTGGTCCAAGTCCATCGTGGGCTGGCCCAACCAATCCGGGAGCGCATCCGCCCAACCCCGCCCGCCCGCCCGCGGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCAGCGGCCATCACCTCCCAACACCCGCGCCACCAACCCCGACCTGGCCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGTGGCTTCCCGGTGCGCCCAAGTGCACCTAACAAGGGCCCGTGGACCTGTCCCACTTCCTGAAG  
 GAGAAGGCGGCTGGAGGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCCCGGCATCCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGCCTGGACCCCGGAGGAGGTGG  
 AGAAGGCCAACGAGGGCGGAGAACAACTGCCTGTGCACCCCATGTCCAGCACGGCATGGAGGACGAGGACCCCGGAGGTGCTGAAGTGGGAAG  
 TTCGACTCCCGCTGGCCCTGGCCACATCGCCCGGAGCGCCACCCCGAGTACTACAAGGACTAA

135/178

Fig. 96A

45. 2003 CON G nef .PEP  
 MGGKWSKSSIVG<sup>W</sup>PEVRERIRQTPPAEAGVAVSQDLARHGAI<sup>T</sup>SSNTAANNPDCAWLEAQEEDSEVGFVVRPQVPLRPMTYKGAFDLSEFFL  
 KEKGGLDGLIYSKKRQDILD<sup>LW</sup>VYNTQGFEPDQWNYTPGPTFRPLTFGWCFKLVPM<sup>D</sup>PAEVEEANKGENNSLLHPICQHGMEDEDEREVLVW  
 RFDSSLARRHRIARELHPEYKDC\$

Fig. 96B

2003 CON G nef .OPT  
 ATGGCGGCAAGTGGTCCAAAGTCCATCGTGGCTGGCCGAGGTGGGAGCGGCATCCGCCAGACCCCGCCCGCCGAGGGCGTGGG  
 CGCCGTGCCAGGACCTGGCCCGCACGGCCATCACCTCCTCAACACACGCGCCCAACCCGACTGCGCTGGCTGGAGGCCCCAGG  
 AGGAGACTCCGAGGTGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGGCGCTTCGACCTGCTCCTTCTTCCCTG  
 AAGGAGAAGGGCGCTGGACGGCTGATCTACTCCAAGAAGCGCCAGGACATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCC  
 CGACTGGCAGAACTACACCCCGGCCACCCGCTTCCCTGACCTCGGTGGTCAAGCTGGTGCCCATGGACCCCGCCGAGG  
 TGGAGGAGGCCAACAAAGGGCGAGAACACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGACGAGGACCGCGAGGTGCTGGTGTGG  
 CGCTTCGACTCCTCCCTGGCCCGCCACATCGCCCGGAGTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 97A

46. 2003 CON H nef .PEP  
 MGGKWSKSSIGG<sup>W</sup>PAIREIRRAEPAEAGVAVSRDLDRRGAVTINN<sup>T</sup>ASTNPDSAMLEAQEEEEVEVGFVVRPQVPLRPMTYKGAFDLSEFFL  
 KEKGGLEGLIYSKKRQEILD<sup>LW</sup>VYNTQGYEPDQWNYTPGPGERYPLTFGWC<sup>F</sup>KLV<sup>P</sup>VD<sup>P</sup>Q<sup>E</sup>VEKANEKANE<sup>N</sup>NSLLHPICQHGMEDEEREVLMW  
 KFDSRLAFRHHRIARELHPEFYKDC\$

Fig. 97B

2003 CON H nef .OPT  
 ATGGCGGCAAGTGGTCCAAAGTCCATCGGCGGCTGGCCCGCCATCCGGGAGCGGCATCCGCCCGCCCGAGCCCGCCGAGGGCGTGGG  
 CGCCGTGCCCGGACCTGGACCGCGGCGCTGACCATCAACAACACCCCTCCACCAACCCGACTCCGCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGGCGCTTCGACCTGCTCCTTCTTCCCTG  
 AAGGAGAAGGGCGCTGGAGGGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTACTTCCC  
 CGACTGGCAGAACTACACCCCGGCCACCCCTGACCTTCCGCTGGTCAAGCTGGTGCCCGTGGACCCCGCCGAGG  
 TGGAGAGGCCAACGAGGCGAGAACACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGACGAGGACCGCGAGGTGCTGATGTGG  
 AAGTTCGACTCCCGCTGGCCCTCCGCCACATCGCCCGGAGCTGCACCCCGAGTTCACAAGGACTGCTAA

136/178

**Fig. 98A****47. 2003\_CON 01\_AE nef .PEP**

MGGKWSKSSIVGWPQVRERIKQTPPATEGVGVAVSQDLDKHGAVTSSNMNADCVWLRAQEEEEVGFVPRPQVPLRPMTYKGAFDLSFFLKEK  
 GGLDGLIYSKKRQEIILDLWVYNTQGFPPDWQNYTPPGPIRYPLCFGWCFKLVVPDPREVEEDNKGENNCLLHPMSQHGIEDEEREVLWVKFD  
 SALARKHIIARELHPEYKDC\$

**Fig. 98B****2003\_CON 01\_AE nef .OPT**

ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGGCTGGCCCCAGGTGGCGAGCGCATCAAGCAGACCCCCCCCCCGCCACCAGGGCGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCAGGCGCGGTGACCTCTCAACATGAACAACCGCCGACTGCGTGTGGCTGCGCCGCCAGGAGG  
 AGGAGGTGGGCTTCCCGGTGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGCGCCTTCGACCTGTCTTCTTCTTGAAGGAGAAG  
 GGCGCCTGGACGGCCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCCCGACTGGCA  
 GAACTACACCCCGCGCCCGCATCCGCTACCCCTGTGCTCGGCTGCTCAAGCTGGTGGACCCCGGAGCCCGGAGGTGGAGGAGG  
 ACAACAAGGCGAGAACAACCTGCTGCACCCCATGTCCAGCACGGCATCGAGGACGAGGAGCGCGGAGGTGCTGATGTGGAAGTTCGAC  
 TCCGCGCTGGCCCCAAGCACATCGCCCCGAGCTGCACCCCGACTACAAGGACTGCTAA

**Fig. 99A****48. 2003\_CON 03\_AE nef .PEP**

MGGKWSKSSIVGWPQVRERIRRAPAPAARGVPSQDLDKYGAVTSSNTAANNADCAWLEAQKEEEVGFVPRPQVPLRPMTYKGAFDLSHFL  
 KEKGLDGLIYSKKRQEIILDLWVYHTQGYFPDWQNYTPGPIRFPLTFGWICYKLVVPDPDEVEEATEGENNSLLHPICOHGMDDDEEKEVLMW  
 KFDSRLALTHRARELHPEFYKDC\$

**Fig. 99B****2003\_CON 03\_AE nef .OPT**

ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGGCTGGCCCCAGGTGGCGAGCGCATCCGCGCGCCCGCCCGCCCGCCCGCGCGCGT  
 GGGCCCCGTGTCCCAGGACCTGGACAAGTACGGCGCGGTGACCTCTCAACACCGCCCAACACCGCGACTGCGCTGGCTGGAGGCC  
 AGAAGGAGGAGGAGGTGGGCTTCCCGGTGGCCCCAGGTGCCCTGCGCCCATGACCTACAAGGCGCCTTCGACCTGTCCCACTTCCCTG  
 AAGGAGAAGGCGGCTGGACGGCCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCC  
 CGACTGGCAGAACTACACCCCGGCCCGCATCCGCTTCCCGCTGACCTTCCGCTGGTGTACAGCTGGTCCCGTGGACCCCGACGAGG  
 TGGAGGAGGCCACCGAGGCGGAGAACAACTCCCTGTGCACCCCATCTGCCAGCACGGCATGGACGACGAGGAGAAGGAGGTGCTGATGTGG  
 AAGTTCGACTCCCGCTGGCCCTGACCCACCGCGCGGAGCTGCACCCCGAGTCTACAAGGACTGCTAA

137/178

Fig. 100A

49. 2003 CON 04 CFX nef .PEP  
 MGGKWSKSSIVGWP<sup>AI</sup>RRMRQRGPAQAEPAAGVAVSODLDKHGAI<sup>TS</sup>SN<sup>TA</sup>ATNDKAWLEAQEEEEVEGFPV<sup>RP</sup>QVPLRPMTFKAAALD  
 LSHFLKEKGLDGLIYSKKRQEI<sup>LD</sup>LWVYHTQGYFPD<sup>WQ</sup>NYTPGPFERFPLCFGWCKLVPVDPQEV<sup>EE</sup>ATEGENNCLLHPISQHGMEDEER  
 EVLKWKFD<sup>SR</sup>LAYK<sup>HI</sup>ARELHPEFYKDC\$

Fig. 100B

2003 CON 04 CFX nef .OPT  
 ATGGCGGCAAGTGGTCCAAAGTCCCTCCATCGTGGGCTGGCCCGCCATCCGGAGCGCATGCGCCAGCGGGCCCGCCAGGCCGAGCCCGC  
 CGCCGGCGGTGGCGCGGTGTCCAGGACCTGGACAAGCAGGGCCATCACCCTCCAACACCGCCGACCAACCCGACAAAGGCCT  
 GGCTGGAGGCCAGGAGGAGGAGGTGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTCAAGGCCCGCCCTGGAC  
 CTGTCCACTTCTGAAGGAGAAGGCGGCTGATCTACTCCAAGAAGCCAGGATCCTGGACCTGTGGGTGTACAACAC  
 CCAGGGCTACTTCCCGACTGGCAGAACTACACCCCGCGGAGCGCTTCCCGCTGTGCTCGGCTGGTCAAGTGGTGC  
 TGGACCCCGAGGAGTGGAGGCCACCGAGGGCAGAACTGCTGTGACCCCATCTCCAGCACGGCATGGAGGACGAGGAGCGC  
 GAGGTGCTGAAGTGAAGTTCGACTCCCGCCTGACACATCGCCCGGAGCTGCACCCCGAGTCTACAAGGACTGCTAA

Fig. 101A

50. 2003 CON 06 CFX nef .PEP  
 MGGKWSKSSIVGWP<sup>QV</sup>RRMRNPTEGAAEGVAVSODLDKHGAI<sup>TS</sup>SN<sup>TA</sup>TATNAACAWLEAQTEDEVGFPV<sup>RP</sup>QVPLRPMTYKGAFLDLSFF  
 LKEKGLDGLIYSKKRQEI<sup>LD</sup>LWVYHTQGF<sup>FP</sup>DWQNYTPGPGIRYPLTFGW<sup>CY</sup>KLV<sup>VP</sup>DPKEVEEDTKGENNCLLHPMCQHGVEDEEREVLM  
 WKFDSS<sup>L</sup>ARRH<sup>I</sup>AREMHPEFYKDC\$

Fig. 101B

2003 CON 06 CFX nef .OPT  
 ATGGCGGCAAGTGGTCCAAAGTCCCTCCATCGTGGGCTGGCCCGCCAGGTGCGGAGCGCATGCGCAACCCCGCCAGGGCCCGCCGAGGG  
 CGTGGCGCCGTGTCCAGGACCTGGACAAGCAGCGCCATCACCTCTCCAACACCGCCACCAACCGCCCTGCGCTGGCTGGAGG  
 CCCAGACCGAGGACGAGTGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTACAAGGGCGCTTCGACCTGTCTCTTC  
 CTGAAGGAGAAGGCGGCTGGACGGCTGATCTACTCCAAGAAGCCAGGATCCTGGACCTGTGGGTGTACAACACCGAGGCTTCTT  
 CCCCAGTGGCAGAACTACACCCCGCCCGCATCCGCTACCCCTGACCTTCGGCTGGTGTACAAGTGGTGCCTGGACCCCAAGG  
 AGGTGGAGGAGGACACCAAGGCGGAGAACAACTGCTGTGACCCCATGTGCCAGACGGCGTGGAGGACGAGGCGGAGGTGCTGATG  
 TGGAGTTCGACTCCTCCCTGGCCCGCCACATCGCCCGGAGATGACCCCGAGTCTACAAGGACTGCTAA

138/178

**Fig. 102A**

51. 2003 CON 08 BC nef. PEP  
 MGKWSKSSIVG<sup>W</sup>PA<sup>I</sup>R<sup>R</sup>IR<sup>R</sup>TE<sup>P</sup>AA<sup>D</sup>GV<sup>G</sup>AV<sup>S</sup>RD<sup>L</sup>E<sup>K</sup>H<sup>G</sup>A<sup>I</sup>T<sup>S</sup>NT<sup>A</sup>D<sup>T</sup>N<sup>A</sup>D<sup>C</sup>A<sup>W</sup>L<sup>E</sup>T<sup>O</sup>E<sup>E</sup>E<sup>E</sup>E<sup>V</sup>G<sup>F</sup>P<sup>R</sup>P<sup>Q</sup>V<sup>P</sup>L<sup>R</sup>P<sup>M</sup>T<sup>F</sup>K<sup>G</sup>A<sup>L</sup>D<sup>L</sup>S<sup>F</sup>F<sup>L</sup>K  
 EKGGLEGLIYSKRRQ<sup>E</sup>ILDLWVYHTQGYFPDWHNYTPG<sup>V</sup>R<sup>F</sup>PLTFG<sup>W</sup>C<sup>F</sup>K<sup>L</sup>V<sup>P</sup>VD<sup>P</sup>RE<sup>V</sup>E<sup>E</sup>A<sup>N</sup>E<sup>G</sup>E<sup>D</sup>N<sup>C</sup>L<sup>L</sup>H<sup>P</sup>V<sup>C</sup>Q<sup>H</sup>M<sup>E</sup>D<sup>E</sup>H<sup>R</sup>E<sup>V</sup>L<sup>K</sup>W<sup>K</sup>  
 FDSQLAHRH<sup>R</sup>ARELH<sup>P</sup>E<sup>F</sup>Y<sup>K</sup>DC<sup>§</sup>

**Fig. 102B**

2003 CON 08 BC nef. OPT  
 ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGCTGGCCGCCCATCCGGAGCGCATCCGCCGACCCGAGCCCGCCGACCGCGTGGG  
 CGCCGTGTCCCGGACCTGGAGAAGCACGGGCCATCACCTCCAACACCGCCGACACCAACGCGACTGCGCTGGCTGGAGACCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCCCTCAGGTGCCCTGCGCCCATGACCTCAAGGGCCCTGGACCTGTCTTCTCCTGAAG  
 GAGAAGGGCCCTGGAGGCTGATCTACTCAAGAAGCGCAGGAGATCTGACCTGTGGTGTACACACCCAGGCTACTTCCCCGA  
 CTGGCAAACTACACCCCGCCCGGCTGCGCTTCCCGTGGCTTCAAGTGGTGGCCCGGAGCCCGGAGGTGG  
 AGGAGGCCAACGAGGCGGAGACAACCTGCTGCACCCCGTGCACCGCATGGAGGACGAGCACCCGCGAGGTGCTGAAAGTGGAA  
 TTCCGACTCCAGCTGGCCCAACCGCCACCGGAGTGCACCCCGAGTCTACAAGGACTGCTAA

**Fig. 103A**

52. 2003 CON 10 CD nef. PEP  
 MGKWSKSSIVG<sup>W</sup>PA<sup>V</sup>R<sup>R</sup>IR<sup>R</sup>TD<sup>P</sup>AAE<sup>G</sup>V<sup>G</sup>AA<sup>S</sup>RD<sup>L</sup>E<sup>K</sup>Y<sup>G</sup>A<sup>I</sup>T<sup>S</sup>NT<sup>A</sup>Q<sup>T</sup>N<sup>P</sup>D<sup>C</sup>A<sup>W</sup>L<sup>E</sup>A<sup>Q</sup>E<sup>E</sup>E<sup>E</sup>E<sup>V</sup>G<sup>F</sup>P<sup>R</sup>P<sup>Q</sup>V<sup>P</sup>L<sup>R</sup>P<sup>M</sup>T<sup>Y</sup>K<sup>G</sup>A<sup>F</sup>D<sup>L</sup>S<sup>F</sup>F<sup>L</sup>  
 KEKGGLEGLIYSKRRQ<sup>E</sup>ILDLWVYNTQGF<sup>F</sup>PD<sup>W</sup>Q<sup>N</sup>Y<sup>T</sup>PG<sup>I</sup>R<sup>Y</sup>PLTFG<sup>W</sup>C<sup>Y</sup>K<sup>L</sup>V<sup>P</sup>VD<sup>P</sup>RE<sup>V</sup>E<sup>E</sup>A<sup>N</sup>E<sup>G</sup>E<sup>N</sup>S<sup>L</sup>L<sup>H</sup>P<sup>M</sup>S<sup>L</sup>H<sup>G</sup>M<sup>E</sup>D<sup>P</sup>H<sup>G</sup>E<sup>V</sup>I<sup>M</sup>W  
 KFDSNLAHKH<sup>M</sup>ARELH<sup>P</sup>E<sup>F</sup>Y<sup>K</sup>DC<sup>§</sup>

**Fig. 103B**

2003 CON 10 CD nef. OPT  
 ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGCTGGCCGCCCATCCGGAGCGCATCCGCCGACCCGAGCCCGCCGACCGCGTGGG  
 CGCCGCTCCCGGACCTGGAGAAGTACGGGCCATCACCTCCAACACCGCCGACACCCGACTGCGCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGGTGGCTTCCCGTGGCCCTCAGGTGCCCTGCGCCCATGACCTACAAGGGCCCTTCGACCTGTCTTCTCCTG  
 AAGGAAGGGCCCTGGAGGCTGATCTACTCCAAGCGCCGAGGACATCTGGACCTGTGGTGTACAACACCCAGGCTTCTTCCC  
 CGACTGGCAGAACTACACCCCGCCCGGATCCGCTACCCCTGACCTTCGGCTGGTGTACAAGTGGTGGCCCGTGGACCCCGGAGG  
 TGGAGGAGGCCAACGAGGCGGAGAACACTCCCTGCTGCACCCCATGTCCCTGCACGGCATGGAGGACCCCAACGGCGAGGTGCTGATGTGG  
 AAGTTCGACTCCAACCTGGCCCAACAGCACATGGCCCGGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

139/178

## Fig. 104A

## 53. 2003 CON 11 CFX nef .PEP

MGGKWSKSSIVGWP<sup>~</sup>PEIRERLRRTPPTAAAEVGVAVSKDLEKHGAVTSNTAQTNAACAWLEAQEEEEVGFVPRPQVPLRPMTYKGAFDLGEF  
LKEKGLDGLIYSKKRQEI<sup>~</sup>LDLWVYHTQGYFPDWQNYTPGPIRYPLCFGWCFKLVPEPREVEEANEGENNCLLHPMSQHGMDDDEEREVL<sup>~</sup>M  
WKFDSSLARRRHARELHPDFYKDC\$

## Fig. 104B

## 2003 CON 11 CFX nef .OPT

ATGGCGGCAAGTGGTCCAAGTCCCTCCATCGTGGGTGGCCCGAGATCCGGAGCGCCTGGCGCCACCCCGCCACCCCGCCGCGGAGGG  
CGTGGCGCCCGTGTCCAAGACCTGGAGAACGACGGCGCCGTGACCTCCTCAACACCCCGCCAGACCAACCGCCCTGGCCTGGCTGGAGG  
CCAGAGGAGGAGGAGGTGGGTTC<sup>~</sup>CCCGTGGCCCGCCAGGTGCCCTGGCCCATGACCTACAAGGGCCCTTCGACCTGGGCTTC  
CTGAAGGAGAGGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTT  
CCCCGACTGGCAGAACTACACCCCGGCTACCCCTGTGCTCGGCTGGCTCAAGCTGGTGC<sup>~</sup>CCCGTGGAGCCCCCGG  
AGGTGGAGGAGGCCAACGAGGGCGAGAACAACTGCCCTGTGACCCCATGTCCAGACGGCATGGACGACGAGGCGCGAGGTGCTGATG  
TGGAA<sup>~</sup>GTTCGACTCCTCCCTGGCCCGCCACATCGCCCGGAGCTGCACCCCGACTTCTACAAGGACTGCTAA

## Fig. 105A

## 54. 2003 CON 12 BF nef .PEP

MGGKWSKSSIVGWP<sup>~</sup>DIRERMRRAPPAAEGVAVSQDLENRGAITSSNTRANNPDLAWLEAQEEEEVGFVPRPQVPLRPMTYKALDLSHFLK  
EKGGLEGLIYSKKRQEI<sup>~</sup>LDLWVYHTQGYFPDWQNYTPGPIRYPLTFGWCFKLVVDPEEVEKANEGENNCLLHPMSQHGMEDREVL<sup>~</sup>MWK  
FDSRLALRHHAREKHPEFYQDC\$

## Fig. 105B

## 2003 CON 12 BF nef .OPT

ATGGCGGCAAGTGGTCCAAGTCCCTCCATCGTGGGTGGCCCGACATCCGGAGCGCCTGGCGCCCGCCCGCCCGCCCGCGGCGGTGGG  
CGCCGTGTCCAGGACCTGGAGAACCGCGGCCATCACTCCTCAACACCCCGCCCAACACCCCGACCTGGCCTGGCTGGAGGCCCAGG  
AGGAGGAGGAGGTGGGTTC<sup>~</sup>CCCGTGGCCCGCCAGGTGCCCTGACCTACAGGGCCCTGGA<sup>~</sup>CCTGTCCACTTCTTGAAG  
GAGAGGGCGGCTGGAGGCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCGGA  
CTGGCAGAACTACACCCCGCCCGCATCCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGC<sup>~</sup>CCCGTGGACCCCGAGGAGGTGG  
AGAA<sup>~</sup>GCCCAACGAGGGCGAGAACAACTGCCCTGTGCA<sup>~</sup>CCCATGTCCAGCACGGCATGGAGGACGAGGACCCCGGAGGTGCTGATGTGGAAG  
TTCGACTCCCGCCTGGCCCGCCACATCGCCCGGAGAACCCCGAGTTCTACCAGGACTGCTAA



Fig. 106A

55. 2003 CON 14 BG nef. PEP

MGKWSKCSI VGVPEVRRIRRTPPAAVGVGAVSQDLAKHGAITSNTAANNPDCAWLEAQEEDSEVGFVRRPQVPLRPMTYKGAFDL SFFL  
KEKGLDGLIYSKQRQDILDLLWVYNTQGFPPDWQNYTPGPTRYPLTFGWCFLKLEPVDPAEVEEATKGENNSLLHPI CQHGMEDADNEVLIW  
RFDSSLARRHRIARELHPDFYKDC\$

Fig. 106B

2003 CON 14 BG nef. OPT

ATGGGGGCAAGTGGTCCAAGTGTCCATCGTGGGCTGGCCGAGGTGGCGGAGCGCATCCGCGGCACACCCCGCCGCGCGCGGTGGCGGTGG  
CGCCGTGTCCCAGGACCTGGCCAGCAGCGGCCATCACCTCCCAACACCCGCGCAACAACCCCGACTGCGCTGGCTGGAGGCCCAGG  
AGGAGGACTCCGAGGTGGCTTCCCCTGGCCCGCCAGGTGCCCTGACCTACAGGGCCCTTCGACTGTCTTCTTCCCTG  
AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCAAGCAGCGCCAGGACATCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCC  
CGACTGGCAGAACTACACCCCGGCCCCGACCCGCTACCCCTGACCTTCCGCTGGTGTCAAGCTGGAGCCCGTGGACCCCGCCGAGG  
TGGAGGAGGCCACCAAGGGCAGAACTCCCTGTGCACCCCATCTGCCAGCACGGCATGGAGACGCCGACAACGAGGTGCTGATCTGG  
CGCTTCGACTCCTCCCTGGCCCGGCGCATCGCCCGGAGCTGCACCCCGACTTCTACAAGGACTGCTAA

Fig. 107A

61. 2003 CON S pol. PEP

FFRENLAFOGEAREFSSEQTRANSPTRSRELRVRGGDNP LSEAGAEQGTVLSLFPQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEEIN  
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNI IGRNMLTQIGCTLNFPI SPIETVPVKLKPMDGPKVKQWPLTEEK  
IKALTEICTEMEKEKISKIGPENPYNTPIFAIKKDS TKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTLVDVGDAYFSPVPLDE  
DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS PAIFQSSMTKILEPFRTONPEI VIQYMDLIVGSDLEIGQHRTKIEELREHLLRWGF  
TTPDKKHQKEPFFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEAELELAEN  
REILKEPVHGVYDPSKDLIAEIQKQGDQWTYQIYQEPFKNLKTGKYAKMRSHTNDVKQLTEAVQKIATESI VIWGTKPKFRLPIQKETW  
ETWWTEYQATWIPWEFVNTPLVLKLYQLEKEPIVGAETFFVDGAANRETKLGAGYVTDGRQKVVSLFETTNOKTELQAIHLALQDSG  
SEVNI VTD SQYALGI IQAOPDKSESELVNIQIEQLIKKEKVVLSWPAHKGIGGNEQVDKLVSTGIRKVLFLDGDKAQEEHEKHYHSNWRAM  
ASDFNLPPIVAKEI VASCDKQLKGEAMHGQVDCSPGIWQDCTHLEGI I LVAVHVASGYIEAEVI PAETGOETAYFILKLAGRWPVKVIH  
TDNGSNFTSAVKAACWWWAGIQEFGIPYNPQSQGVESMNKELKII GQVRDQAEHLKTA VQMAVFIHNFKRKGGIGGYSAGERI IDIIAT  
DIQTKELQKQITKI QNFRVYRDSRDP IWKGPALLWKGE GAVVIQDNSEIKVVP RRKAKI IRDYGKOMAGDDCVAGRQDED\$

Fig. 107B

2003\_con\_s pol. OPT

TTCTTCCGCGAGAACCTGGCCCTCCAGCAGGGGAGGCCGGGAGTTCTCTCCGAGCAGACCCGCGCCAACTCCCCACCTCCCGCGAGCTGCCGCGTGCCG  
CGCGCGGACAACCCCTGTCCGAGGCCGGCCGAGCCAGGGCACCTGTCCCTGTCTTCCCCAGATCACCCCTGTGGCAGCGCCCTTGGTGACCCG  
TGAAGATCGCGCGCCAGCTGAAGGAGGCCCTGTGGACACCGCGCCGACACCCGTGTGGAGAGATCAACCTGCCCGCAAGTGAAGCCCAAGATG  
ATCGCGCGCATCGCGCGCTTCAATCAAGGTGCGCCAGTACGACCAAGATCTGATCGAGATCTGCGCAAGAGCCCATCGGCACCCGTGTGGTGGCCCCAC  
CCCCGTGAACATCATCGCGCCGCAACATGTGACCCAGATCGGCTGACCCCTGAACCTTCCCATCTCCCCCATCGAGACCGTGCCTGAAGCTGAAGCCCCG  
GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCGTGACCGAGGAGAGATCAAGGCCCTGACCGGATCTGCAACGAGATGGCGCAAGTGGTGGACTTCCGCGAGCTGAACAA  
AAGATCGCCCCCGAGAACCCCTACAACAACCCCATCTTCCCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGTGGTGGACTTCCGCGAGCTGAACAA  
GGCACCCAGGACTTCTGGAGGTGCAGTGGCATCCCCACCCCGCGCTGAAGAAGAAGTCCGTGACCCGTGGTGGACGTGGCGGACGCCCTACT  
TCTCCGTGCCCTGGACGAGACTTCCGCAAGTACCCGCTTCAACCTCCCTCATCAACAGAGACCCCGCATCCGCTACCGTACCAACAGCTGCTG  
CCCCAGGCTGGAAGGCTCCCCGCCATCTCCAGTCTCCATGACCAAGATCTTGAGCCCTTCCGCAACCCAGAACCCCGAGATCGTGATCTACCAGTA  
CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCACCAAGATCGAGGAGCTGGCGGAGCACCTGCTGCGTGGGCTTCAACACCC  
CCGACAAGAAGCACCAAGAGGCCCTTCTGTGGATGGGTACGAGTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGAGAAGGAC  
TCCCTGGACCGTGAACGACATCCAGAGCTGGTGGCAAGTGAACCTGGCTCCAGATCTACCCCGCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGG  
CGCGCCAAGGCCCTGACCGACATCGTGTGCCCTGACCGAGGAGGCGAGTGGAGTGGCCGAGAACCCCGAGATCTGAAGGAGCCCGTGCACGGCGGTG  
ACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGCCAGGACCAAGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACC  
GGCAAGTACGCCAAGATCGCTCCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAAGATCGCCACCGAGTCCATCGTGATCTGGGGCAA  
GACCCCAAGTCCGCTGCCATCCAGAGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGATCCCGAGTGGGAGTTCGTGAACA  
CCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGACCCATCGTGGCGCCGAGACCTTCTACGTGGACGGCCCGCCCAACCCGAGACCAAGCTG  
GGCAAGGCCGGCTACGTGACCCGACCCGCGCCAGAAAGTGGTCCCTGACCGAGACCCCAACAGAAAGACCGAGCTGCAGGCCATCCACCTGGCCCT  
GCAGGACTCCGGCTCCGAGGTGAACATCGTGAACCGACTCCAGTACGCCCTGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGCTGGTGAACC  
AGATCATCGAGCAGCTGATCAAGAAGGAAAGTGTACTGTCTGGGTGCCCGCCACAAGGCCATCGGGCGCAACGAGCAGGTGGACAAGTGGTGTCC  
ACCGGCATCCGCAAGGTGTCTCCGACGGCATCGACAAGGCCAGGAGACGAGAAATACCACTCCAACCTGGCGGCCATGGCCCTCCGACTTCAA  
CCTGCCCCCATCGTGGCCAAAGGAGATCGTGGCTCCTGCGACAAGTGGCCAGTGAAGGGCGAGGCCATGCACGGCCAGTGGACTGCTCCCGCCGATCT  
GGCAGCTGGACTGCACCCACTGGAGGGCAAGATCATCTGGTGGCCGTGACCTGGCTCCGGTACATCGAGGCCGAGGTGATCCCGCCGAGACCCGGC  
CAGGAGACCGCTACTTCTTCACTTGAAGCTGGCCGGCCGTGGCCGTGAAGTGTATCCACACCGACAACGGCTCCAACCTCACCTCCGCGCCGTGAAGGC  
CGCTGTGTGGCCGGCATCCAGCAGGAGTTCGGCATCCCTACAACCCAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
TCGGCCAGGTGCGGACAGGCCGAGCACCTGAAGACCCGCGTGCAGATGGCCGTCTTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGCGGCTACTCC  
GCCGCGAGCGCATCATCGACATCATCGCCACCGACATCCAGGAGCTGCAGAAGCAGATCAACCAAGATCCAGAACTTCCCGCTGTACTACCCGGA  
CTCCCGGACCCCATCTGGAAGGGCCCGCCAAAGCTGTGTGGAAGGGCGAGGCCCGCTGGTATCCAGGACAACCTCCGAGATCAAGTGGTGGTCCCGCCG  
GCAAGGCCAAGATCATCCGCGACTACGGCAAGGAGATGGCCGGCGCCCGCCAGGACGAGGACTAA

Fig. 108A

**62. 2003 M GROUP anc pol. PEP**

FFRENLAFOQGEAREFSSEQTRANSPTSRELVRGGDNPLSEAGAERQGTVSFSPQITLWQRPLVTIKIGQLREALLDTGADDTVLEEIN  
 LPGKWKPKMIGGIGGFKVRQYDQILIEICGKKAIGTVLVGPTPVIIGRNMLTQIGCTLNFPIETVPVKLPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKBDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDE  
 DFRKYTAFTIPSTNETPGIRYQYNVLPQGWKSPAIFQSSMTKILEPFRTKNPEIYIYQYMDLLYVGSDEIGQHRAKIEELREHLLRWGF  
 TTPDKKHQKEPFLWNGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAEELELAEN  
 REILKEPVHGYYDPSKDLIAEIQKQGDQWTVQIYQEPFKNLKTGKYAKMRSHTNDVKQTEAVQKIATESIWIWPKTPKFRPLPIQKETW  
 ETWWEYQATWIPWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRQKVVSLTETTNQKTELQAIHLALQDGS  
 SEVNI VTD SQYALGIIQAQPKSESELVNOIIEQLIKKEKVVLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYSNWRAM  
 ASDENLPPVVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEKVVILVAVHVASGYIEAEVIPAETGQETAYFILLKLAGRWPVKVIH  
 TDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAT  
 DIQTKELQKQITTKIQNFRVYRDSRDP IWKGPAKLLWKGEGAVVIQDNSEIKVPPRRKAKIIRDYGKQMGAGDDCVAGRQDED\$

**63. 2003 CON A1 pol. PEP**

FFRENLAFOQGEARKFSSEQTGANSPTSRLDWDGGRDLSLPEAGAERQGTGPTFFFPQITLWQRPLVTIVRIGGQLKEALLDTGADDTVLEDI  
 NLPKWKPKMIGGIGGFKVRQYDQILIEICGKKAIGTVLVGPTPVIIGRNMLTQIGCTLNFPIETVPVKLPGMDGPKVKQWPLTEEK  
 KIKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKBDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD  
 EFRKYTAFTIPSTNETPGIRYQYNVLPQGWKSPAIFQSSMTKILEPFRSKNPEIYIYQYMDLLYVGSDEIGQHRTKIEELRAHLLSWG  
 FTTPDKKHQKEPFLWNGYELHPDKWTVQPIELPEKESWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEAEELELAE  
 NREILKDPVHGYYDPSKDLIAEIQKQGDQWTVQIYQEPFKNLKTGKYARKRSHTNDVKQLAEVVQKVVMEIWIWPKTPKFRPLPIQKET  
 WETWMDYQATWIPWEFVNTPPLVKLWYQLEKDPVGAETFYVDGAANRETKLGKAGYVTDGRQKVVSLTETTNQKTELHAIHLALQDS  
 GSEVNI VTD SQYALGIIQAQPRSESELVNOIIEKLIKGDVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEHYHSNWR  
 MASDENLPPVVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEKVVILVAVHVASGYIEAEVIPAETGQETAYFILLKLAGRWPVKV  
 HTDNGSNFTSAAVKAACWWANIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA  
 TDIQTKELQKQITTKIQNFRVYRDSRDP IWKGPAKLLWKGEGAVVIQDNSEIKVPPRRKAKIIRDYGKQMGAGDDCVAGRQDED\$

142/178

Fig. 109A

## Fig. 108B

## 2003\_M\_GROUP\_anc\_pol.1.OP1

TTCTCCGGGAGAACTGGCCCTCCAGCAGGGGAGGCCGGAGTTCTCCTCCGAGAGACCCGGCCCAACTCCCCACCTCCCGCGAGCTGCGCGTGG  
 CGGGGGGACAACCCCTGTCCGAGCGCGCCGAGCGCCAGGGCACCGTGTCTTCTCCTTCCCCCAGATCACCTGTGGCAGCGCCCTGGTGTGACCA  
 TCAAGATCGGCGCCAGCTGCGGAGGCCCTGCTGGACACCGCCGACACCGTGTGGAGGAGATCAACCTGCCCGGCAAGTGAAGCCCAAGATG  
 ATCGGGGCATCGGGGCTTCAACAAGTGGCCAGTACGACCAAGATCCTGTATCGAGATCTGGCGCAAGAGGCCATCGGCACCGTGTGGTGGGCCCCAC  
 CCCCCTGAACATCATCGGGCCGAACATGCTGACCCAGATCGGCTGCACTTCCCACTTCCCACTCGAGACCGTGCCTGTAAGCTGAAGCCCG  
 GCATGGACGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGAGATCGACCCGAGATGGAGAAGGAGGCAAGATCTCC  
 AAGATCGGCCCGAACCCTAACACCCCCCTGTTCCCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGTGGTGGACTTCCCGGAGCTGAACAA  
 GCGZACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCCCACCCCGCCCTGAAGAAGAAGTCCGTGACCGTGGACCGTGGCGACCGCTACT  
 TCTCCGTGCCCTGGACCGAGGACTTCCGCAAGTACACCGCTTACCAATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAGTACAACCGTGTG  
 CCCCAGGCTGGAAGGCTCCCCCGCATCTCCAGTCCATGACCAAGATCCTGGAGCCCTTCCGCACCAAGAACCCCGAGATCGTGTATCTACCAGTA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGCCGTGGGCTTCAACCCC  
 CCGACAAGAAGCACCAAGAGGCCCTTCTGTGGATGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGTGCCTCCGAGAAGGAC  
 TCTTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAATGGCCCTCCAGATCTACCCGGCATCAAGGTGAAGCAGTGTCAAGTGTGCG  
 CGGGCCAAAGCCCTGACCCGACATCGTGCCCTGACCGAGGAGCCGAGTGGAGTGGCCGAGAACCAGGAGATCCTGAAGGAGCCCGTGCACGCGCTGT  
 ACTACGACCCCTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGACCCAGTGGACCTACAGATCTACCAGGAGCCCTCAAGAACCCTGAAGACC  
 GGAAATACGCCAAGATGCGCTCCGCCACACCAACGACCTGAAGCAGTGCAGGCCCTGACAGAGATCGCAAGAGATCGCCACCGAGTCCATCGTGTGGGCAA  
 GACCCCAAGTTCGGCTGCCCATCCAGAAGGACCTGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCTGGTGAAGCTGTGGTACCAAGCTGGAGAAGGAGCCCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCGCCCAACCGCGAGACCAAGCTG  
 GGCAAGCCGGCTACGTGACCCGCGCCAGAAAGTGGTGTCCCTGACCCGAGACCAACCAAGAGACCGGAGCTGCAGGCCATCCACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGAACCGACTCCCAAGTACGCCCTGGGCATCATCCAGGCCACGCCCAAGTCCGAGTCCGAGTGGTGAACC  
 AGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACTGTCTGGTGGCCCGCCACAGGGCATCGGGGCAACGAGAGGTGGACAAGCTGGTGTCC  
 TCCGGCATCCGCAAGTGTCTTCTGGACGGCATCGCAAGGCCAGGAGGACGAGAAATCCACTCCAACCTGGCGGCCATGGCCCTCCGACTTCAA  
 CCTGCCCCCTGGTGGCCAAAGGAGATCGTGGCTTCTGGACAAGTGCAGCTGAAGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCGCGCATCT  
 GCGAGTGGACTGACCCACTGGAGGCAAGGTGATCCTGTTGGCCGTGCACGTGGCTCCCGGTACATCGAGGCCGAGGTGATCCCCCGGAGACCGGC  
 CAGGAGACCGCTACTTCAATCCTGAAGTGGCCGGCCGTGGCCGTGAAGTGTACACACCGACAACCGCTCCAACCTCACCTCCGCGCCGTGAAGGC  
 CGCTGCTGGTGGCCGGCATCCAGCAGGAGTTCGGCATCCCCCTAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCGGACCAAGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTCAATCCACAACCTCAAGCGCAAGGCGGCATCGGCGGCTACTCC  
 GCCGCGAGGCGCATCATCGACATCATCGCACCGGACATCCAGACCAAGGAGTGCAGAAGCAGATCACCAAGATCCAGAACTTCCCGGTGTACTACCCGGA  
 CTCCCGGACCCCATCTGGAAGGGCCCCGCAAGCTGCTGTGGAAGGGCGGCGCCGTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGGCCCCGGC  
 GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCCGACGACTGGCTGGCCGGCCGACGAGGAGGACTAA

143/178

Fig. 109B

2003\_con a1 pol.1.OPT

TTCTTCGGGAGAACCTGGCCCTTCCAGCAGGGGAGGCCGCAAGTTCTCCTCCGAGCAGACCGGCGCCAACTCCCCACCTCCCGGACCTGTGGGACGG  
CGGCCGACTCCCTGCCCTCCGAGGCCGCGCCGAGCGCCAGGGCACCGGCCACCTTCTCCTTCCCCAGATCACCCCTGTGGCAGCGCCCTGGTGA  
CCGTGGCATTGGCGCCAGTGAAGGAGGCCCTGTGGACACCGCGCCGACGACACCGTGTGGAGGACATCAACCTGCCCGCAAGTGAAGCCCAAG  
ATGATCGCGCCGATCGCGCCGCTTCAAGGTGAAGCAGTACGACCCAGATCCTGTATCGAGATCTGGGGCAAGAGGCCATCGGCACCGTGTGGTGGGCC  
CACCCCGTGAACATCATCGCGCCGCAACATGCTGACCCAGATCGGCTGCACCTGAACTTCCCCATCTCCCCATCGAGACCGTGGCCCGTGAAGCTGAAGC  
CCGGCATGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGAGATCTGCACCCGAGATGGAGAGGAGGGCAAGATC  
TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCCATCTTCGCCATCAAGAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAA  
CAAGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCGCTGAAGAAGAAGTCCGTTGACCGTGTGGACGTGGCGGACCGCCT  
ACTTCTCCGTGCCCTGGACGAGTCTTCCGCAAGTACACCGCCTTCAACATCCCTCCACCAACGAGACCCCGGCATCCGCTACCAGTACAACAGTG  
CTGCCACAGGGTGAAGGGTCCCGCCATCTTCCAGTCTCCATGACCAAGATCTTGGAGCCCTTCCGCTCCAAGAACCCCGAGATCATCATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGGATCGGCCAGCACCGCAACAGATCGAGGAGTGGCGCCACCTGCTGTCTCTGGGGCTTTCACCA  
CCCCGACAAGAAGCACAGAAGGAGCCCCCTTCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAG  
GAGTCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAATGGCCCTCCAGATCTACCGCGCATCAAGGTGAAGCAGCTGTGCAAGCTGTCT  
GCGGGCCCAAGGCCCTGACCGACATCGTGAACCTGACCGAGGAGCCGAGCTGGAGTGGCCGAGAACCAGGAGATCTTGAAGGACCCCGTGCACCGGG  
TGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGACCACTGGACCTACAGATCTACCCAGGAGCCCTTCAAGAACCTGAAG  
ACCGCAAGTACGCGCAAGCGCTCCGCCACACCAACGACCTGAAGCAGTGGCCGAGTGGTGCAGAAAGTGGTGTGAGTCCATCGTATCTGGGG  
CAAGACCCCAAGTCAAGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGATGGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGA  
ACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGACCCCATCGTGGCGCCGAGACCTTCTACGTGGACGGCCGCGCAACCCGAGACCAAG  
CTGGCAAGCGCGCTACGTGACCGACCCGGCCGAGAAAGTGTGTCCCTGACCGGAGACCAACCAAGACCCGAGCTGCACGCCATCCACCTGGC  
CCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGAACCGACTCCAGTACCGCCCTGGGCATCATCCAGGCCCGACCCCGCTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGAAGTGTGATCGGCAAGGACCAAGGTGTACCTGTCTGGTGGCCGCCCCACAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTG  
TCCTCCGGCATCCGCAAGGTGTCTCCGACCGCATCGACAAGGCCAGGAGGACCGAGGCTACCACTCCAACCTGGCGGCCCATGGCCCTCCGACTT  
CAACTGCCCCCATCGTGGCCAAAGGAGATCGTGGCCCTCTGCGACAAGTGCAGTGAAGGGGAGGCCCATGACGGCCAGGTGACTGCTCCCGCCGCA  
TCTGGCAGCTGGACTCACCCACTGGAGGCAAGTGTATCCCTGGTGGCCGCTGACCTGGCCCTCCGGCTACATCGAGGCCGAGGTGATCCCGCCGAGACCC  
GGCCAGGACCCGCTACTTCTGCTGAAGCTGGCCGCGCTGGCCGCTGAAGTGGTGCACACCGACAACCGCTCCAACCTCACTCCCGCCGCGCTGAA  
GGCCGCTGTGGTGGCCACATCCAGCAGGAGTTCGGCATCCCTTACACCCCGCTCCAGGGCGTGGTGGAGTCCATGACAAGGAGCTGAAGAAGA  
TCATCGGCCAGGTGCGGAGCAGGCCGAGACCTGAAGACCGCCGTGAGATGGCCGTGTTCATCCAACACTCAAGCGCAAGGGCGGCATCGGCGGCTAC  
TCCGCGGCGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCCAGAACTTCCGCGTGTACTACCG  
CGACTCCCGGACCCCATCTGGAAGGGCCCGCCCAAGTGTCTGTGGAAGGGCGGAGGGCCCGTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGCCCC  
GCCGCAAGGCCCAAGATCATCCCGGACTACGGCAAGCAGATGGCCCGGCGGACCGACTGCGTGGCCCGGCGCCAGGACGAGGACTAA

145/178

## Fig. 109C

## 64. 2003 A1.anc pol. PEP

FFRENLA FQGEARKF SSEQTRANSPTSRELWDGGRDLSLSEAGAERQGTVPFSFPQITLWQRPLVTVKIGGQLKEALLDTGADDDTVLEDI  
 NLPKWKPKMIGGIGGFIKVRQYDQILJIEICGKKAIGTVLVGPTPVNI IGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEE  
 KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD  
 EFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGPSAIFQSSMTKILEPFRSKNPEI VIYQYMDDL YVGSDDLEIGQHRAKIEELRAHLLSWG  
 FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEAELAE  
 NREILKDPVHGYYDPSKDLVAEIQKQGDQWTYQIYQEPFKNLKTGKYAKKRSHTNDVKQLTEVVQKVATESI VIWKT PKFRLP IQKET  
 WETWMEYWQATWIPWEFEVNTPPLVKLWYQLEKEPIAGAEYFVVDGAANRET KLGKAGYVTDGRQKVVSLETETTNQKTELHAIHLALQDS  
 GSEVNI VTD SQYALGI IQAQPDRSESELVNI IEKLEKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
 MASDFNLPPIVAKEIVASCDCQKQKGEAMHGQVDCSPGIWQLDCTHLEKVIILVAVHVASGYIEAEVI PAETGQETAYFLKLAGRWPVKV  
 HTDNGSNFTSAAVKAACWWANIQQEFGI PYNPQSQGVVESMNKELKKIIGQVREQAHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIIA  
 TDIQTKELQKQITKIQNFVRYRDSRDP IWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKA KIRDYGKQMAAGDDCVAGRQDED\$

## Fig. 109D

2003\_A1.anc pol.1.OPT

TTCTCCGGAGAACCTGGCCCTCCAGCAGGGGGCCGCAAGTTCTCTCCGAGCAGACCCGGCCAAACTCCCCACCTCCCGGAGCTGTGGGACGG  
 CGCCCGGACTCCCTGTGTCCGAGGCCGGCCGAGCCAGGGCACCGTGCCTCTCTCTCCCTCCCCAGATCACCTGTGGCAGCGCCCTGGTGA  
 CCGTGAAGATCGCGGCCAGCTGAAGGAGGCCCTGTGGACACCGGCCGACGACACCGTGTGGAGGACATCAACCTGCCCGCAAGTGAAGCCCCAAG  
 ATGATCGCGGCCATCGCGGCCATCAAGGTGGCCAGTACGACCCAGATCTGTATCGAGATCTGGGCAAGAAGGCCATCGGCACCGTGTGGTGGGCC  
 CACCCCTGTGAACATCATCGCGCCGAACATGTGACCCAGATCGGCTGACCCCTGAACCTCCCATCTCCCCATCGAGACCGTGCCTGAAAGCTGAAGC  
 CCGCATGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGAGATCTGCACCGAGATGGAGAAGGAGGCCAAGATC  
 TCCAAGATCGGCCCGAGAACCCCTACAACACCCCTGTTCGCCATCAAGAGAAGGACTCCACCAAGTGGCGCAAGTGGTGAATCCGCGAGCTGAA  
 CAAGCCACCCAGGACTCTGGGAGGTGACGTGGGCATCCCCACCCCGGCCCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGCGGACCGCT  
 ACTTCTCCGTGCCCTGGACGAGTCTCCGCAAGTACACCGCCCTCACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAAGTACAACGTG  
 CTGCCCCAGGGCTGGAAGGGTCCCTCCAGTCTCCATGACCAAGATCTGGAGCCCTTCGGTCCAAGAACCCTCGAGATCGTGTATCTATCCA  
 GTACATGGAACGACTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGAGCTGGCCGCCACCTGTCTCTGGGGCTTCACCA  
 CCCCCAAGAAGCACAGAGAGGCCCTCCCTTCTGTGGATGGCTACGAGTGCACCCCGACAAAGTGGACCGTGCAGCCCATCAAGCTGCCCGAGAAG  
 GACTCTGGAACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAACCTGGCCCTCCAGATCTACGGCGCATCAAGGTGAAGCAGCTGTGCAAGCTGT  
 GCGGGCCCAAGCCCTGACCGACATCGTGAACCTGACCGAGGAGGCCGAGCTGGAGTGGCCGAGAACCGCGAGATCTTGAAGAACCCCTGCACGGGG  
 TGTACTAGAACCCCTCAAGGACCTGTGGCCGAGATCCAGAAGCAGGGCCAGGACCAAGTGAACCTACAGATCTACCAGGAGCCCTTCAAGAACCCTGAAG  
 ACCGGCAAGTACGCCAAGAAGCGCTCCGCCACACCAACGACGTGAAGCAGCTGACCCGAGGTGTGCAGAAGTGGCCACCGAGTCCATCGTGTATCTGGGG  
 CAAGACCCCAAGTCCGCCCTGCCATCCAGAAGGACCTGGGAGACCTGGTGGATGGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGA  
 ACACCCCTCCCTGGTGAAGTGTGGTACCAGCTGGAGAAGGACCCATCGCCGGCCGAGACCTTCTACGTGGACGGCCGCGCCACCCGCGGAGACCAAG  
 CTGGCAAGCCGGCTACGTGACCGACCGCCGCGCAGAAAGTGGTGTCCCTGACCCGAGACCAACCAAGAGACCCGAGCTCACGCCATCCACCTGGC  
 CCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCACGCCGACCGCTCCGAGTCCGAGCTGGTGA  
 ACCAGATCATCGAGAAGCTGATCGAGAAGGAGAGGTGTACCTGTCTGGTGGCCGCCACAAAGGCCATCGCGCGCAACGAGCAGGTGGACAAGCTGGTG  
 TCCCTCCGGCATCCGCAAGTGTCTTCTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCCTCAACTGGCGCGCCATGGCCTCCGACTT  
 CAACCTGCCCCCATCGTGGCCAAAGGATCGTGGCCCTCTGGACAAGTGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
 TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGGTGATCTTGTGGCCGTGCACGTGGCCCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACC  
 GGCCAGGACCGCCCTACTTCTGTGAAGTGGCCCGCCGTGGCCCGTGAAGTGGTGCACACCGCAACGGCTCCAACTCACCTCCGCGCCCGCTGAA  
 GGCCGCTGTGGTGGCCCAACATCCAGCAGGAGTTCGGCATCCCTTACAACCCCAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
 TCATCGGCCAGGTGGCGGAGCAGGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTCTATCCACAACCTCAAGCGCAAGGGCGGCATCGGGCGCTAC  
 TCCGCCGGGAGCGCATCATCGACATCATCGCCACCCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCAACAAGATCCAGAACTCCGCGTGTACTACC  
 CGACTCCCGGACCCCATCTGGAAGGGCCCGCCCAAGTGTGTGGAAGGGCGAGGGCCCGCTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGCCCC  
 GCCGCAAGCCAAAGATCATCCGGGACTACGGCAAGCAGATGGCCGGCCGCGCAGGACGAGACTAA

146/178

147/178

Fig. 110A

65. 2003 CON A2 pol. PEP

FFRENLA FQRE ARKFSSE QNRANSPTSRELNRNGGRDNNLSEAGAE EQGTVHSCNFPQITLWQRPLVTVKIEGQLRÉALLDTGADDTVLEDI  
NLPGRWPKMIGGIGGFIKVRQYDQIAIEICGKRAIGTVLVGPTPVNIIGRNMLVQLGCTLNFFPISPIETVPVKLPGMDGPKVKQWPLTEE  
KIKALTEICKEMEKEGKISKIGPENPYNTPVFAIKKDDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLH  
EDFRKYTAFTIP SINNETPGIRYQYNVLPQGKGS PAIFQSSMTKILEPFRSKNPEMVIYQYMDLLYVGSDDLEIGQHRAKIEELRAHLLRWG  
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPKEDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTDIVTLTKEAELELEE  
NREILKNPVHGVYDPSKDLIAEIQKQGQDQWYQIYQEPFKNLKTGKYAKRKSTHTNDVKLQTEAVQKIAIESIVIWGKTPKFRPLPIQKET  
WETWTEYWOATWIPWEFVNTPPLVKLWYQLETEPIAGAEIFYVDGAANRETKLGKAGYVTDGRQKIVSLTETTNQKTELHAIYLALQDS  
GLEVNI VTD SQYALGIIQAOPDRSESELVNQIIEKLI EKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEHYHSNWRA  
MAHDFNLPPIVAKEIVASCDCQKGEAMHGQVDCSPGIWQDCTHLEGKVIILVAVHVASGYIEAEVI PAETGQETAYFILKLAGRWPVKVI  
HTDNGPNFTSATVKAACWWAGVQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA  
TDIQTKELQKQI IKIQNFRVYRDSRDP IWKGPÄKLLWKGE GAVVIQDN SDIKV VPRRKAKIIRDY GKQ MAGDDCVAGRQDED\$

Fig. 111A

66. 2003 CON B pol. PEP

FFREDLAF POGKAREFSSE QTRANSPTRRELQVWGRDNNLSSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQKÉALLDTGADDTVLEEM  
NLPGRWPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFFPISPIETVPVKLPGMDGPKVKQWPLTEE  
KIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD  
KDFRKYTAFTIP SINNETPGIRYQYNVLPQGKGS PAIFQSSMTKILEPFRKQNPDIYIYQYMDLLYVGSDDLEIGQHRTKIEELRQHLLRWG  
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVILPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVIPLTEAELELAE  
NREILKEPVHGVYDPSKDLIAEIQKQGQWYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIVIWGKTPKFRPLPIQKET  
WEAWTEYWOATWIPWEFVNTPPLVKLWYQLEKEPIVGAETIFYVDGAANRETKLGKAGYVTDGRQKVVSLTDTTNQKTELQAIHLALQDS  
GLEVNI VTD SQYALGIIQAOPDKSESELVSQIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGDIDKAQEEHEHYHSNWRA  
MASDFNLPPVVAKEIVASCDCQKGEAMHGQVDCSPGIWQDCTHLEGKIIILVAVHVASGYIEAEVI PAETGQETAYFILKLAGRWPVKTI  
HTDNGSNFTSTTVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIA  
TDIQTKELQKQI IKIQNFRVYRDSRDP IWKGPÄKLLWKGE GAVVIQDN SDIKV VPRRKAKIIRDY GKQ MAGDDCVAGRQDED\$



Fig. 110B

2003\_CON\_A2\_pol1.OPT

TTCTCCGGAGAACCTGGCCCTCCAGAGCGGAGGCCCGAAGTTCTCTCCGAGCAGAACCCGCCCAACTCCCCACCTCCCGCGAGCTGCGCAACGG  
CGCCCGGACAACTGCTGTCCGAGCCGGCCGAGAGCAGGGCACCGTGCCTCTTGCAACTTCCCCAGATCACCTTGTGGCAGCCGCCCTGGTGA  
CCGTGAAGATCGAGGCCAGCTGGCAGGCCCTGTGGACACCGGCCGACACCCGTGTGGAGACATCAACCTGCCCGCAAGTGGAAAGCCCCAAG  
ATGATCGCGGGCATCGCGGGCTTCAAGGTGCGCCAGTACGACCAAGTCCGATCGAGATCTGGGCAAGCGGCCCATCGGCACCGTGTGGTGGGCC  
CACCCCGTGAACATCATCGGCCGCAACATGTGGTGCAGTGGCTGCACCTGAACTTCCCATCTCCCCATCGAGACCGTGCCTGAAAGCTGAAGC  
CCGGCATGGAACGGCCCAAGGTGAAGCAGTGGCCCTGACCCGAGGAGAAAGTCAAGGCCCTGACCCGAGATCTGCAAGGAGATGGAGAAGGAGGCCAAGATC  
TCCAAGATCGGCCCGGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGGAGCTGAA  
CAAGCCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCACCCCGCGCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGCGGACGCCCT  
ACTTCTCCGTGCCCTGCACGAGGACTTCCGCAAGTACACCCGCTTACCATCCCTCCATCAACAACGAGACCCCGCCGATCCGCTACCAGTACAACAGTG  
CTGCCCCAGGGTGAAGGGTCCCCCGCATCTCCAGTCTCCATGACCAAGATCTTGAGCCCTTCCGCTCAAGAACCCCGAGATGGTATCTACCA  
GTACATGGACGACTGTACGTGGCTCCGACCTGGAGATCGGCCAGACCCGGCCCAAGATCGAGGAGCTGGCGCCACCTGCTGGCTGGGCTTCACCA  
CCCCGACAAAGAAGCACCAGAAGGAGCCCCCTTCTGTGGTGGCTACGAGTGCACCCGCAAGTGGACCGTGCAGCCATCAAGCTGCCCGAGAAG  
GACTCTGGACCTGAAAGACATCCAGAAGCTGGTGGCAAGTGAACCTCCAGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCT  
GCGGGCACCAAGGCCCTGACCGCATCTGTGACCTGACCAAGGAGCCGAGCTGGAGTGGAGGAGAACCGCGAGATCTTGAAGAACCCTGTCACCGG  
TGTAFTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGCCAGGACCAAGTGGACCTACAGATCTACCAGGAGCCCTTCAAGAACCCTGAAG  
ACCGGCAAGTACGCCAAGCAGTCCACCACACCAAGCTGAAGCAGCTGACCGAGCCCTGCAGAAAGTCCGATCGAGTCCATCGTATCTGGG  
CAAGACCCCAAGTTCGGCTGCCATCCAGAAGGAGACTGGGAGACTGGTGGACCGACTTGGACCGCCACTGGATCCCGAGTGGGAGTTCGTGA  
ACACCCCTCCCTGGTGAAGCTGTGGTACAGCTGGAGACCGAGCCCATCGCCGGCCCGAGACCTTCTACGTGGACGGCCGCGCAACCGCGAGACCAAG  
CTGGCAAGGCCGCTACGTGACCGACCCGGCCGCAAGAGTCTGTCCCTGACCGAGACCAACCAAGAGACCGAGCTGCACGCCATCTACCTGGC  
CCTGCAGGACTCCGGCTGGAGGTGAACATCGTACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACCGCTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGAAGTGTATCGAGAAGGAGCGGTGTACTCTCTGGTGGCCGCCCCAAGGGCATCGGCCGCAACGAGAGGTGGACAAAGTGGTG  
TCCTCCGGCATCCGCAAGTGTCTCTGGACGGCATCGACAAGGCCCAGGAGGACACGAGCTTCCAACTGGCGGCCCATGGCCACGACTT  
CAACTGCCCCCATCGTGSCCAAGGAGATCGTGGCTCCTCGGACAAAGTGCAGCTGAAGGGGAGGCCATGCACGGCCAGGTGGACTGCTCCCCGGCA  
TCTGGCAGCTGGACTGCACCCACTGGAGGGCAAGTGTCTGTGGCTGCACTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCGCCGAGACC  
GGCCAGGACCCGCTACTTCACTCTGAACTGGCCGCGCTGGCCGTGAAGGTGATCCACACCGACAAAGGCCCAACTTCACTCCGCCACCGTGA  
GGCCGCTGTGGTGGCCGGCTGCAGCAGGAGTTCGGCATCCCTTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
TCATCGGCCAGGTGGCGACCCAGCCGAGCACCTGAAGACCCCGCTGAGATGGCCGTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGCGGGCTAC  
TCCGCCGGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCCAGAACTTCCCGCGTGTACTACCG  
CGACTCCCGGACCCCATCTGGAAGGGCCCCCGCAAGCTGTGTGGAAGGGCGGCGCCGCTGGTGTATCCAGACAACTCCGACATCAAGGTGGTGGCCCC  
GCCGCAAGCCAAAGATCATCCGGACTACGGCAAGCAGATGGCCGGCGGACGACTGCTGGCCGGCCCGCAGGACGGACTAA

148/178

Fig. 111B

2003\_CON\_B\_pol.OPT

TTTTCGGGAGGACCTGGCCCTCCCAAGGCAAGGCCCGGAGTTCTCTCCGAGCAGACCCGGCCCAACTCCCCACCCCGCGGAGCTGCAGGTGTG  
GGCCCGGACAACAACACTCCCTGTCGAGGCGCGGCGGACCCGTCCTTCTCCTTCCCCAGATCACCTGTGGCAGGCCCCCTGGTGA  
CCATCAAGATCGGCGGCGGAGTGAAGGAGCCCTGCTGGACACCGGCGGCGGACGACACCGTGTGGAGGAGATGAACCTGCCGCGCTGGAAGCCCAAG  
ATGATCGGCGGCTATCAAGGTGCGGCTGAGTACGACACAGATCTGATCGAGATCTGGGCCACAAGGCCATCGGCACCGTGTGGTGGGCC  
CACCCCGTGAACATCATGGCCGCAACTGCTGACCCAGATCGGCTGCACCCCTGAACTTCCCATCTCCCCATCGAGACCGTGCCTGTAAGCTGAAGC  
CCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCCGAGATGGAGAAGGAGGCAAGATC  
TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAA  
CAAGGCACCCAGGACTTCTGGGAGGTGACAGTGGGCATCCCCACCCCGCGGCTGAAGAAGAAGACTCCGTGACCCGTGGACCGTGGCGGACCGCT  
ACTTCTCCGTGCCCTGGACAAGGACTTCGGCAAGTACACCGCCTTCCCATCCCATCAACAACGAGACCCCGGCGCATCCGCTACCCAGTACAACCGTG  
CTGCCCCAGGGCTGGAAGGGCTCCCCCGCATCTTCCAGTCTCATGACCAAGATCTTGAGCCCTTCCGCAAGCAGAACCCCGACATCGTGATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCCGACCAAGATCGAGGAGCTGGCCAGCACCTGCTGCCGTGGGCTTCACCA  
CCCCGACAAGAAGCACCAAGAGGAGCCCCCTTCTGTGGATGGCTACGAGTGCACCCCGCAAGTGGACCGTGCAGCCCATCGTGTGCCCGAGAAG  
GACTCTGGAACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAACCTGGCCCTCCAGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCT  
GGCGGCACCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAGCCGAGCTGGAGTGGCCGAGAACCGGAGATCTGAAGGAGCCCGTGCACCGGG  
TGTAACGACCCCTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGCCAGTGGACCTACAGATCTACCAGGAGCCCTTCAAGAACCTGAAG  
ACCGCAAGTACGCCCGCATGCGCGGGCCCCACACGACGTGAAGCAGTGAACCGGCGGCTGCAGAAAGATCGCCACCGAGTCCATCGTGATCTGGGG  
CAAGACCCCAAGTCAAGCTGCCATCCAGAAGGAGACTGGGAGCCCTGGTGGACCGACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCGTGA  
ACACCCCCCTGGTGAAGCTGTGTTACAGCTGGAGAAGGAGCCCATCGTGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCCGCGAGACCAAG  
CTGGCAAGGCCGCTACGTGACCGACCGGCGCCAGAAAGTGGTGTCCCTGACCCGACACCAACAGAACCCGAGCTGCAGGCCATCCACCTGGC  
CCTGAGGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCCAAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGCTGGTGT  
CCCAGATCATCGAGCAGTGTCAAGAAGGAGAGGTGTACTGGCTGGGTGCCCGCCACAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTG  
TCCGCGGCATCCGCAAGGTGCTGTTCTTGGACGGCATCGACAAGGCCAGGAGCACGAGAGTACCCTCAACTGGCGGCCATGGCCTCCGACTT  
CAACTGCCCCCGTGGTGGCAAGGAGATCGTGGCTCTTGGACAAGTCCAGCTGAAGGCGAGGCCATGACGGCCAGGTGGACTGCTCCCCCGGCA  
TCTGGCAGCTGGACTGACCCACTGGAGGGCAAGATCATCTGGTGGCTGGACCTGGCTTCCGGCTACATCGAGGCCGAGGTGATCCCCCGGAGACC  
GGCCAGGACCCGCTACTTCTGTGTAAGCTGGCCGCGCTGGCCGCTGAAGCCATCCACAACCGGCTCCAACTTCACTCCACCCAGCCGTAAG  
GGCCGCTGCTGGTGGCCGGCATCAAGCAGAGTTCGGCATCCCTTACAACCCAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
TCATCGGCCAGGTGCGCGACCCAGCCGAGCACCTGAAGACCGCCGTGAGATGGCCGTGTTCAATCCACAACCTCAAGCGCAAGGCGGCATCGGCGGCTAC  
TCCGCGGCGAGGCGCATCGTGGACATCATGCCACCGACATCCAGACCAAGGAGCTGCAGAAAGCAGATCCCAAGACTTCCGCGTGTACTACCG  
CGACTCCCGGACCCCTGTGGAAGGGCCCCCGCAAGCTGTGTGGAAGGGCGGCGGCTGGTGTATCCAGGACACTCCGACATCAAGGTGGTGGCCCC  
GCCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGGACTGGCTGGCCTCCCGCAGGACGAGGACTAA

150/178

## Fig. 111C

## 67. 2003 B. anc pol. PEP

FFRENLAFPQKAREFSSEQTRANSPTRRELQVWGRDNNPLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEM  
 NLPKWKPKMIGGIGGFIVKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEE  
 KIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD  
 KDFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPETVIYQYMDLLYVGSDEIGQHRTKIEELREHLLRWG  
 FTTDPKKHQKEPPFLWNGYELHPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPVPLTEEALELAE  
 NREILKEPVHGVYDPSKDLIAEIQKQGGQWTVQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIWIWGKTPKFKLPIQKET  
 WEAWWTEYQATWIPWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRQKVVSLDTTNNQKTELQAIHLALQDS  
 GLEVNIVTDSQYALGIIQAQPDKSESELVSQIEQLIKKEKVVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
 MASDFNLPVVAKEIVASCDKCOLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI  
 HTDNGSNFTSTTVKAACWWAGIKQEFGIPYNPQSQGVVESMKNKELKIIIGQVRDQAEHLKTAVQMAVFIHNEFKRKGIGGYSAGERIVDIIA  
 TTDIQTKELOKQITKIQNFRVYRDRSRLWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIRDYKGQMGAGDDCVASRQDEDŞ

Fig. 111D

2003\_B\_anc pol.1.OPT

TTCTTCGGGAGAAACTGGCCCTTCCCCAGGGCAAGGCGCGAGTTCTCTCCGAGCAGACCCGGCCAACTCCCCACCCGGCCGGAGCTGCAGGTGTG  
GGGCCGGACAACAACCCCTGTCCGAGGCCGGCCGACCGCCAGGGCACCGTTCCTTCTCCCCAGATCACCCCTGTGGCAGCCGCCCTTGGTGA  
CCATCAAGATCGGGCCAGCTGAAGGAGGCCCTGTCTGGACACCGGGCCGACACCGTGTGGAGAGATGAACCTGCCCGCAAGTGGAAAGCCCAAG  
ATGATCGGGCCGATCGGGCCCTTCAATCAAGGTGGCCAGTACGACAGATCCTGATCGAGATCTGGGCCACAAGGCCATCGGCACCCGTCGTGGTGGGCC  
CACCCCGTGAACATCATCGGCCGAACCTGTGACCCAGATCGGTGCACCTGAACTTCCCCATCTCCCCATCGAGACCTGCCGTGAAGCTGAAGC  
CCGGATGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGTGGAGATCTGCACCCGAGATGGAGAGGGGCAAGATC  
TCCAAAGATCGGGCCCGAGAACCTTACAACACCCCGTGTTCGCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAA  
CAAGCCACCCAGGACTTCTGGAGTGCAGCTGGGCATCCCCACCCCGGCCCTGAAGAAGAAGTCCGTGACCGTGGCTGGACGTGGCGGACGCCCT  
ACTTCTCCGTGCCCCGGACAAGGACTTCCGCAAGTACACCGCTTACCATCCCTTCCATCAACAACGAGACCCCGGCATCCGCTACCAAGTACAACGCTG  
CTGCCCAAGGCTGGAAGGCTCCCCCGCCATCTTCCAGTCTTCCATGACCAAGATCTGGAGCCCTTCCGCAAGCAGAACCCCGAGATCGTGTATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGCCAGCACCCCAAGATCGAGGAGCTGCCGGAGCACCTGTGGCTGGGGCTTACCA  
CCCCGACAAGAAGCACCAAGAGGAGCCCCCTTCTGTGGATGGCTACGAGTGCACCCCGACAAGTGGACCGTGCAGCCCATCGTGTGCCCGAGAAAG  
GACTCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAATGGCCCTCCAGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGTCT  
GCGGCAACCAAGGCCCTGACCGAGTGGTGCCTGACCCGAGGAGCCGAGCTGGAGTGGCCGAGAACCCCGAGATCTTGAAGGAGCCCGTGCACGGCCG  
TGTACTACGACCCCTCCAAGGACCTGTATCGCCGAGATCCAGAAGCAGGGCCAGGCTGAGCTTACAGATCTACAGGAGCCCTTCAAGAACCCTGAAG  
ACCGCAAGTACCGCCGATGCGCGGCCCAACCAACGACGTGAAGCAGCTGACCGAGCCGTGCAGAGATCGCCACCGAGTCCATCGTGTATCTGGGG  
CAAGACCCCAAGTCAAGTGCCTTCCAGAGGACTGGAGGCTGGTGGACCGAGTCTGGCAGGCCACTGGATCCCGAGTGGAGTTCGTGA  
ACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCGTGGCGCCGAGACTTCTACGTGGACGGCCCGCCCAACCGCGAGACCAAG  
CTGGCAAGGCCGGCTACGTGACCCGACCCGGCCAGAAAGTGGTGTCCCTGACCGACACCAACAGAACCCGAGCTGCAGGCCATCCACCTGGC  
CTTGAGGACTCCGGCTGGAGGTGAACATCGTACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAAGTCCGAGTCCGAGCTGGTGT  
CCAGATCATCGACAGCTGATCAAGAGGAGAAGTGTACTTGGCTGGGTGCCCGCCACAAGGCCATCGGGCCGCAACGAGCAGGTGGACAAAGTGGTG  
TCCGGCCGATCCGCAAGTGTGTCTTGGACGGCATCGAACGGCCAGGAGGAGCAGAGAGTACCACTCCAACCTGGCGGCCATGGCCCTCCGACTT  
CAACTGCCCGCTGGTGGCCAAAGGAGATCGTGGCTCTCTGCAAAAGTGCAGCTGAAGGGGAGGCCATGCAAGGCCAGGTGGACTGCTCCCCCGGCA  
TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACC  
GGCCAGGAGACCGCTACTTCTCTGAAGTGGCCCGCTGGCCGTGAAGGTGATCCACACCGACAAAGGCTCCAACCTCAACCTCCACCAACCGTGA  
GGCCCGCTGTGGTGGCCGGCATCAAGCAGGAGTTCGGCATCCCTTACAACCCAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
TCACTGGCCAGGTGGCGACAGGCCGAGCACCTGAAGACCCCGCTGCAGATGGCCGTGTTCATCCACAACCTCAAGCCGCAAGGGCCGATCGGGCCGTAC  
TCCGCCGGGAGCGCATCGTGGACATCATGCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCCACAAGTCCAGAACTTCCCGCTGTACTACCG  
CGACTCCCGGACCCCTGTGGAAGGGCCCGCCGCTGTGATCCAGGACAACTCCGACATCAAGGTGGTGGCC  
GCCCAAGGCCAAGATCATCCCGGACTACGGCAAGCAGATGGCCCTCCCGCCAGGACGAGACTAA

152/178

## Fig. 112A

## 68. 2003 CON C pol. PEP

FFRENLAFFQGEAREFFPSEQTRANSPTSRELQVRGDNPRSEAGAERQGTLNFPQITLWQRPLVSIKVGQIKEALLDTGADDTVLEEINLPG  
 KWKPKMIGGIGGFVKVRQYDQILLIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKA  
 LTAICEEMEKEGKITKIGPENPYNTPVFAIKKSDTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTLDVGDAYFVPLDEGFR  
 KYTAFTIPSIINNETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFRQNPVIYQYMDLLYVGSDDLEIGQHRAKIEELREHLLKWGFTTP  
 DKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEALELAENREI  
 LKEPVHGYYDPSKDLIAEIQKQGHQWTVYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFRLP IQKETWETW  
 WTDYWQATWIPWEFVNTTPLLKLVYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTDGRQKIVSLTETTNQKTELQAIQLALQDSGSEV  
 NIVTDSQYALGI IQAQPDKSESELVNIIEQLIKKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASE  
 FNLPPIVAKEIVASCDCQLKGEAIGHQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYYIILKLAGRWPVKVIHTDN  
 GSNFTSAAVKAACWWAGIQQEFFGIPYNPQSQGVVESMKNELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQ  
 TKELQKQIIKIQNERVYYRDSRDP IWKGPAKLLWKGEAVVIQDNSDIKVVPRRKA KIKDYGKQOMAGADCVAGRQDED\$

Fig. 112B

2003\_CON\_C\_pol1.OPT

TTCTTCGGGAGAACCTGGCCCTCCCCAGGGCCGAGGTCGCCCTCCGAGCAGACCCGGCCCAACTCCCCACCTCCCGCGAGCTGCAGGTGG  
CGGACAAACCCCGCTCCGAGGCCGGCCGAGCCAGGCCACCTGAACTTCCCCAGATCACCTGTGGCAGCCCGCCCTGGTGTCCATCAAGGTGG  
GCGCCAGATCAAGGAGGCCCTGTGGACACCGCCCGCAGACACCGTGTGGAGGATCAACTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGC  
ATCGGCGGCTTCATCAAGGTGCGCCAGTACGACAGATCTGTGATCGGCAAGAGGCCATCGGCACCGTGTGGTGGCCCAACCCCGCTGAA  
CATCATCGGCGCAACATGTGACCCAGCTGGGCTGACCCCTGAACTTCCCCATCTCCCCATCGAGACCGTGTCCCGTGAAGCTGAAGCCCGGCATGGACG  
GCCCCAAGGTGAAGCAAGTGGCCCTGACCCGAGGAGAAGATCAAGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGSCAAGATCACCAAGATCGGC  
CCCGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACTCCACCAGTGGCGCAAGTGGTGGACTTCCCGGAGCTGAACAAGCGCACCCCA  
GGACTTCTGGGAGGTGCAGCTGGGCACTCCCCACCCCGCCCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGCGACCGCTACTTCTCCGTGC  
CCCTGGACGAGGGCTTCGCAAGTACACCGCTTACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAGTACAACCGTGTGCCCCAGGGC  
TGGAAAGGCTCCCGCCATCTCCAGTCTCCATGACCAAGATCTTGGAGCCCTCCCGCCAGAACCCCGAGATCGTGATCTACCAGTACATGGACGA  
CCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCCCAAGATCGAGGAGCTGGCGGAGCACCTGCTGAAGTGGGCTTCAACCCCGACAAGA  
AGCACAGAAAGGACCCCGCTTCTGTGGATGGCTACGAGCTGACCCCGACAAGTGGACCGTGCAGCCATCCAGTGGCCGAGAGGACTCCTGGACC  
GTGAACGACATCCAGAAGTGGTGGCAAGCTGAACTGGCCCTCCAGATCTACCCGGCATCAAGGTGGCCAGCTGTGCAAGTGTGCCGGGCGCCAA  
GGCCCTGACCGACATCGTGCCCTGACCGAGGAGGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCCGTGCACGGCGTGTACTACGACC  
CCTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGCCACGACAGTGAACCTACCAGATCTACCAGGAGCCCTCAAGAACCTGAAGACCCGGCAAGTAC  
GCCAAGATGCCACCCGCCACCAACGACCTGAACGAGCTGACCGAGCCGTGCAGAAGATCGCCATGGATCCATCGTGTGGGCAAGACCCCGCAA  
GTTCGGCTGCCATCCAGAAGGAGACCTGGTGGACCGACTACTGGCAGGCCACTGGATCCCGAGTGGGAGTTCGTGAACACCCCGCCCGC  
TGGTGAAGCTGTGGTACCAAGCTGGAGAGGACCCCATCGCCGGCGCCGAGACCTTACGTGGACGGCGCCCAACCGCGAGACCAAGATCGGCAAGGCC  
GGCTACGTGACCGCGCCGAGAGATCGTGTCCCTGACCGGAGACCCCAACCAAGAACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACTC  
CGGCTCCGAGGTGAACATCGTGACCGACTCCCAAGTACCGCCCTGGCATCATCCAGGCCCAAGTCCCGACAAGTCCGAGTGGTGAACCAAGATCATCG  
AGCAGTGAACAAGAGGCGCGTGTACCTGTCTGGTGGCCCGCCCAAGGGCATCGGCGGCAACGAGAGGTGGACAAGTGGTGTCTCCCGGCATC  
CGCAAGTGTGTCTGGACGGCATCGACAAGGCCCAAGGAGGACGAGAAAGTACCACTCCAAGTGGCGGCCATGGCCCTCCAGTTCAAACCTGCCCGC  
CATCGTGGCCAAAGAGATCGTGGCTCTCGCAAAAGTCCAGCTGAAGGGGAGGCCATCCACGGCCAGGTGGACTGTCCCGCCGATCTGGCAGCTGG  
ACTGACCCACCTGGAGGSCAAGATCATCTGTGGTGGCTGCACGTGGCTCCGGTACATCGAGGCCGAGGTGATCCCGCCGAGACCGGCCAGGAGACC  
GCCTACTACATCTGAAGTGGCCCGTGGCCGTGAAGGTGATCCACACCGACAACGGCTCCAACCTCCCGCCCGCTGAAGCCCGCTGCTG  
GTGGCCCGGCATCCAGCAGGAGTTCGGCATCCCTACAACCCAGTCCAGGGCGTGGTGGATCCATGAACAAGGAGTGAAGAAGATCATCGGCCAGG  
TGGCGACCCAGGCCGAGACCTGAAGACCCCGTGCAGATGGCCGTGTTCATCCCAACTCAAGCGCAAGGGCGGCATCGCGGCTACTCCCGCCGGCGAG  
CGCATCATCGACATCATCGCCACCGACATCCAGACCAAGAGCTGCAGAAGCAGATCATCAAGATCCAGAACTTCCCGGTGTACTACCGCGACTCCCGCGA  
CCCCATCTGGAAGGCCCGCCCAAGCTGTCTGTGGAAGGGCGGCGCGTGGTGTCCAGACAACCTCCAGGACAACCTCCAGGCTGGTGGTCCCGCCCGCAAGGCCA  
AGATCATCAAGGACTACGGCAAGCAGATGGCCGGCCGACTGCTGTGGCCGGCCAGGACGAGGACTAA

154/178

## Fig. 112C

## 69. 2003 C.anc pol.PEP

FFRENLAFFQGEAREFPSEQTRANSPTSRELOVGRDNPRSEAGAEERQGTLLTNFPQITLWQRPLVSIKVGGOIKEALLDGTGADDDTVLEEINL  
 PGKWKPKMIGGIGGFIVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKI  
 KALTAICEEMEKEGKTTKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFVSPLDEG  
 FRKYTAFI PSINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRAQNPEIYIYQYMDLLYVGSdleIGQHRAKIEELREHLLKKGFT  
 TPDKKHQKEPPFLWNGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVROCKLLRGAKALTDIVPLTEEALELELAENR  
 EILKEPVHGVYDPSKDLIAEIQKQGHQWYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIYVWGTPKFRLP IQKETWE  
 TWWTDYWQATWIPWEFEVNTPLVKLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTDGRQKIVSLTETTNQKTELOAIQLALQDSDGS  
 EVNIVTDSQYALGIIQAQPKSESELVNIIEQLIKKEVYLSWVPAHKGIGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMA  
 SEFNLPPIVAKELIVASCDCQKLGKGEAMHGQVDCSPGIWOLDCTHLEGKILLVAVHVASGYIEAEVI PAETGQETAYFILLKLAGRWPVKVIHT  
 DNGSNFTSAAVKAACWWAGIQQEFGIYINPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDI IATD  
 IQTKELQKQIIKIQNFRVYRDSRDP IWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIRDYGKQOMAGADCVAGRQDED\$

Fig. 112D

2003\_c.anc pol.1.OPT

TTCTCCGGAGAACCTGGCCCTCCCCAGGGCCGAGTTCCCTCCGAGCAGACCGGGCCAACTCCCCACCTCCCGCGAGCTGCAGGTGGG  
CCGGACAACCCCGCTCCGAGGCGGCGGAGCCAGGCACTGACCTGAACCTCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGTCCATCA  
AGTGGGGCCAGATCAAGGAGGCCCTGTGGACACCGGGCCGACACCGTGTGGAGGAGATCAACCTGCCGCAAGTGAAGCCCAAGATGATC  
GGCGCATCGGCGGCTTCAACAAGGTGGCCAGTACGACCAAGTCTGATCGAGATCTGGGCAAGAGGCCATCGGCACCGTGTGGTGGCCCAACCC  
CGTGAACATCATCGGCCCAACATGTCAGTGGGTGCACCCCTGAACCTCCCATCTCCCATCGAGACCGTGGCCCTGAAGTGAAGCCCGGCA  
TGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGGAGGAGATGGAGAAGGAGGCCAAGATCACCAAG  
ATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACTCCACCAAGTGGCGAAGCTGGTGGACTTCCGCGAGCTGAACAAGCG  
CACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCCCTGAAGAAGAAGTCCGTGACCCGTGTGACCGTGGCGGACCGCCTACTTCT  
CCGTGCCCTGGACGAGGCTTCCGCAAGTACACCGCCTTACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAGTACAACCTGTGCC  
CAGGCTGGAAGGGCTCCCGCCCATCTTCCAGTCTTCCATGACCAAGATCTGGAGCCCTTCCGCGCCAGAACCCCGAGATCGTGATCTACCAGTACAT  
GGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGCTGGCGAGCACCTGCTGAAGTGGGCTTACCACCCCGG  
ACAAGAAGCACCAAGAGGAGCCCTTCTGTGGATGGCTACGAGTGCACCCGCAAGTGGACCCATCCAGCTGCCAGCTGCCGAGAAAGGACTCC  
TGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAACCTGGCCCTCCAGATCTACCCCGGCATCAAGGTGGCGCAGCTGTCAAAGCTGTGCCGG  
CGCAAGCCCTGACCGACATCGTGCCTGACCGAGGCGGAGCTGGAGTGGCCGAGAACCGGAGATCTTACCCCGGCATCAAGGTGGCGCAGCTGTCAAAGCTGTGCCGG  
AGCACCCCTCAAGGACCTGATCGCCGATCCAGAGCAGGCGCACACGACCAAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCCTGAAGACCCGGC  
AAGTACGCCAAGATGCGCACCGCCACACCAAGCTGAAGCAGCTGACCCGAGCCGTGCAGAAGATCGCCATGGAGTCCATCGTGATCTGGGGCAAGAC  
CCCCAAGTTCGGCCTGCCCATCCAGAAGGACCTGGGAGACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCGTGAACACCC  
CCCCCTGGTGAAGCTGTGGTACAGCTGGAGAAGGAGGCCCATCGCCGGCCGAGACCTTCTACGTGGACGGCGCCCAACCCCGAGACCAAGATCGGC  
AAGGCCGGCTACGTGACCGACCGCGCCCGCAGAAAGATCGTGTCCCTGACCCGAGACCAACCAAGAGACCCGAGCTGCAGGCCATCCAGTGGCCCTGCA  
GGACTCCGGCTCCGAGTGAACATCGTGACCGACTCCAGTACGCCCCGAGTACCTGTCTGGTGGTGGCCGCCCCAAGGGCATCGGCGCAAGCTGGTGTCC  
TCACTGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGTCTGGTGGCCGCCCCAAGGGCATCGGCGCAAGCTGGGCAAGCTGGTGTCC  
GGCATCCGCAAGTGTGTCTTCCCTGGACGGCATCGACAAGGCCAGGAGGACGAGAAGTACCCTCAACTGGCGGCCCATGGCCCTCCGAGTTCACCT  
GCCCCCATCGTGGCCAAAGGATCGTGGCCCTCTGCGAACAAGTCCAGTGAAGGGCGGAGGCCATGACGGCCAGGTGGACTGCTCCCCGGCATCTGGC  
AGTGGACTGCACCCACTGGAGGCAAGATCATCTGGTGGCCGTGCACGTGGCTCCGGTACATCGAGGCCGAGGTGATCCCCCGGAGACCGGCCAG  
GAGACCGCCTACTTCACTCTGAAGTGGCCCGCCGTGGCCCGTGAAGTGTACACCCGCAACGGCTCCAACTTCCCTCCCGCCCGCTGAAGGCCG  
CTGTGTGGCCGGCATCCAGCAGGAGTTCGCCATCCCTACAACCCAGTCCAGGGCCGTGGTGGATCCATGAACAAGAGCTGAAGAAGATCATCG  
GCCAGTGGCGGACAGGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACCTCAAGCGCAAGGGCGCATCGGCCGCTACTCCGCC  
GGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGTGCAGAAGCAGATCATCAAGATCCAGAACTTCCGGTGTACTACCGGACTC  
CCGGACCCCATCTGGAAGGGCCCCCAAGCTGTGGAAAGGGCGGAGGCCGTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGCCCGCCGCA  
AGGCCAAGATCATCCGGACTACGGCAAGCAGATGGCCGGCCGCGACTGGTGGCCGGCCCGCAGGACGAGGACTAA



156/178

## Fig. 113A

## 70. 2003 CON D pol.PEP

EFRENLAFOGKAGELSSSEQTRANSPASRELRVQGDNDPLSETGAERQGTVSFNFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEIN  
 LPGWKPKMIGGIGGFIVKVRQYDQILIEICGHKAIGTVLGGPTPVNIIGRNLLTQIGCTLNFPIETVPVKLKPMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISRIGPENPYNTPIFAIAKKKDDSTKWRKLVDFRELNKRRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDE  
 DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPETIYQYMDLQVGSDELEIGQHRTKIEELREHLLRWGF  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLEPEKESWTVNDIQKLVGKLNWASQIYPGIKVRQCLLGRGTKALTEVIPLEEAELELAEN  
 REILKEPVHGVYDPSKDLIAEIQKQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIAIESIYIWKTKPFRLP IQKETW  
 ETWTEYWQATWIPWEFEVNTPPLVKLWYQLEKEPIIGAEYFYVDGAANRETKLGKAGYVTDGRQKVVPLTDTTNQKTELQAINLALQDSG  
 LEVNIVTDSQYALGIIQAQPKSESELVSIIEQLIKKEKVVYLAWVPAHKGIGNEQVDKLVSNIGIRKVLFLDGDIDKAQEEHEKYHNNWRAM  
 ASDENLPPVVAKEIVASCDCQKLGKGEAMHGQVDCSPGIWQLDCTHLEKVIILVAVHVASGYIEAEVI PAETGQETAYFLLKLAGRWPVKVH  
 TDNGSNFTSAAVKAACWWAGIKQEFFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAT  
 DIQTKELQKQIIKIQNFVYRDSRDPIMWGPAPKLLWKGEGAVVIQDNSDIKVVPRRKVKIIRDYKQMGAGDDCVASRQDED\$

## Fig. 114A

## 71. 2003 CON F1 pol.PEP

EFRENLAFOGGEARKFPSEQTRANSPASRELRVQGDNDPLSEAGAERRGTVPSSLFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEDI  
 NLPGWKPKMIGGIGGFIVKVRQYDQILIEICGHKAIGTVLGGPTPVNIIGRNLLTQIGCTLNFPIETVPVKLKPMDGPKVKQWPLTEEK  
 KIKALTEICTEMEKEGKISRIGPENPYNTPVFAIAKKKDDSTKWRKLVDFRELNKRRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD  
 KDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPIAFQCSMTKILEPERTKNPDIYQYMDLQVGSDELEIGQHRTKIEELREHLLRWGF  
 FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQIPDKDSWTVNDIQKLVGKLNWASQIYPGIKVRQCLLGRGAKALTDIVPLTAEAELELAE  
 NREILKEPVHGVYDPSKDLIAEIQKQGQWTYQIYQEPFKNLKTGKYAKMRSHTNDVKQLTEAVQKIAIESIYIWKTKPFRLPILKET  
 WDTWTDYWQATWIPWEFEVNTPPLVKLWYQLETEPIVGAETFYVDGASNRETKKAGYVTDGRQKVVPLTDTTNQKAELOAIHLALQDS  
 GSEVNIVTDSQYALGIIQAQPKSESELVNIIEQLIQKEKVVYLSWVPAHKGIGNEQVDKLVSNIGIRKVLFLDGDIDKAQEEHEKYHNNWRRA  
 MASDFNLPPVVAKEIVASCDCQKLGKGEAMHGQVDCSPGIWQLDCTHLEKVIILVAVHVASGYIEAEVI PAETGQETAYFLLKLAGRWPVKII  
 HTDNGSNFTSAAVKAACWWAGIQQEFFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA  
 TDIQTRRELQKQITKIQNFVYRDSRDPVWKGPAKLLWKGEGAVVIQDENSEIKVVPRRKAKIIRDYKQMGAGDDCVAGRQDED\$.

Fig. 113B

2003\_con\_d pol:OPT

TTCTCCGGGAGAACCTGGCCTTCCCCAGGGCAAGGGCGGAGCTCTCCGAGCAGACCCGGCCAACTCCCCACCTCCCGGAGCTCCGGCGTGTG  
GGCGGGACAAACCCCTGTCCGAGACCGGCGAGCCAGGGCACCGTGTCTCAACTTCCCCAGATCACCTGTGGCAGCGCCCTGGTGAACA  
TCAAGATCGGGCCAGCTGAAGGAGGCCCTGTGGACACCGGGCCGACACACCGTGTGGAGAGATCAACTGCCCGGCAAGTGAAGCCCAAGATG  
ATCGGGCGCATCGGGCGCTTCAACAAGTGGCCAGTACGACAGATCCTGATCGAGATCTGGCCACAAGGCCATCGGCACCGTGTGGTGGGCCCCAC  
CCCCGTGAACATCATCGGCCGAACCTGTGACCCAGATCGGCTGACCTGAACTTCCCCATCTCCCCATCGAGACCGTGTCCCGTGAAGCTGAAGCCCG  
GCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCGTGACCGAGGAGAAGATCAAGCCCTGACCCGAGATCTGCACCGAGATGGAGAAGGAGGCAAGATCTCC  
CGCATCGGCCCGAGAACCCCTACAACACCCCATCTTCCCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGTGGTGGACTTCCGGAGCTGAACAA  
GGCACCCAGGACTTCTGGGAGTGCAGCTGGCATCCCCACCCCGCCCTGAAGAAGAAGTCCGTGTGACCTGGACGTGGCGACGCCCTACT  
TCTCCGTGCCCTGGACGAGACTTCCGCAAGTACACCGCTTCACTCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCACTACCAACCGTCTG  
CCCCAGGCTGGAAGGCTCCCCCGCATCTCCAGTCCCTCATGACCAAGATCTTGGAGCCCTTCCGCAAGCAGAACCCCGAGATCGTGTATCTACCAGTA  
CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGGCGAGCACCTGCTGCCGTGGGCTTCACCCACC  
CCGACAAGAAGCACCAAGAGGACCCCTTCTGTGGATGGCTACGAGCTGACCCCGACAAGTGGACCGTGCAGCCCATCAAGCTGCCCGAGAAGGAG  
TCTTGGACCGTGAACGACATCCAGAAGTGGTGGCAAGTGAACCTGGCCCTCCAGATCTACCCCGCATCAAGGTGGCCAGCTGTCAAGCTGCTGCG  
CGCACCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGGCCGAGCTGGAGTGGCCGAGAACCCTGGAGATCTTGAAGGAGCCCTGCACGGCGTGT  
ACTACGACCCCTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACCTACAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACC  
GGCAAGTACGGCCGATGCGCGGCGCCACCAACGACCTGAAGCAGTGAACCGGCGTGCAGAGATCGCCATCGAGTCCATCGATCTGGGGCAA  
GACCCCAAGTTCGGCTGCCATCCAGAAGGACCTGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACA  
CCCCCCCCTGGTGAAGCTGTGTACCAAGTGGAGAAAGGACCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCAACCGGAGACCAAGCTG  
GGCAAGGCCGGCTACGTGACCCGACCCGGCCGAGAGGTGTGCCCTGACCCGACACCCACCAAGAGACCGAGCTGCAGGCCATCAACCTGGCCCT  
GCAGGACTCCGGCTGGAGGTGAACATCGTGAACCGACTCCAGTACGCCCTGGCATCATCCAGGCCACCAAGGCAAGTGCAGGCCATCAACCTGGCCCT  
AGATCATCGAGCAGTGAACAAGGAAAGTGTACCTGGCCCTGGTGGCCGCCCCACAGGGCATCGGGGCAACGAGCAGGTGGACAAGCTGGTGTCC  
AACGGCATCCGCAAGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGGACACGAGAACTCCACAACAACTGGCGGCCATGGCTCCGACTTCAA  
CCTGCCCCCGTGGTGGCCAAAGGAGATCGTGGCCCTCTGGCAAAAGTGGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCGGCATCT  
GGCAGCTGGACTGCACCCACTGGAGGSCAAGTGTATCTGGTGGCCGTGCACCTGGCCCTCCGGTACATCGAGGCCGAGGTGATCCCCCGGAGACCGGC  
CAGGAGACCGCCCTACTTCCCTGTAAAGCTGGCCCGCCGCTGAAAGTGTGGACCCGCAACCGGCTCCAACCTTCACTCCCGCCGCGGTGAAGGC  
CGCTGTGTGGTGGCCCGGATCAAGCAGGAGTTCGGCATCCCTACAACCCAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
TCGGCCAGGTGCGGACACAGCCGAGCACCTGAAGACCCCGTGCAGATGGCCGTGTTCATCCACAACCTCAAGCCGCAAGGGCGGCATCGCGGCTACTCC  
GCCGGGAGCGCATCATCGACATCATGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCCAGAACTTCCCGGTGTACTACCGGA  
CTCCCCGACCCCATCTGGAAGGGCCCCCAAGCTGCTGTGGAAGGGCGGCGCTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGTCCCCCGCC  
GCAAGGTGAAGATCATCCGCGCATACGGCAAGCAGATGGCCGGCCGACGACTGCTGGCTCCCGCCAGGACGAGGACTAA

Fig. 114B

2003\_CON\_F1\_pol1.OPT

TTCTCCGGAGAACCTGGCCCTCCAGAGGGCGAGGCCCGCAAGTTCCCTCCGAGCAGACCCGGCCCAACTCCCCCGCCTCCCCGGAGCTGGCGGTGCA  
GGCGGGGACAACCCCTGTCCGAGGGCCGGCCGAGCGCCCGGCAACCGTCCCTCCTGTCTCCCCAGATCACCCCTGTGGCAGCGCCCTTGGTGA  
CCATCAAGATCGGCGGCCAGCTGAAGGAGGCCCTGTGGACACCGGGCCGACACCGTGTGGAGACATCAACCTGCCCGGCAAGTGAAGCCCCAAG  
ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGAAGCAGTACGACACATCTGATCGAGATCTGGGCCACAAGGCCATCGGCACCGTGTGGTGGCCCC  
CACCCCTGAACATCATCGGCCGCAACATGTGACCCAGATCGGTGACCCCTGAACCTTCCCATCTCCCATCGAGACCGTCCCGCTGAAGCTGAAGC  
CCGGCATGGACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAAGTCAAGCCCTGACCGGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC  
TCCAAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACTCCACCAAGTGGCCCAAGCTGGTGGACTCCGCGAGCTGAA  
CAAGCGCACCCAGGACTTCTGGGAGGTGCAAGTGGGCATCCCAACCCCGCGCTGAAGAAGAAGTCCGTGACCCCTGGACCTGGACCTGGCCGACGCCCT  
ACTTCTCCGTGCCCTGGACAAGGACTTCCGCAAGTACACCCGCTTACCATCCCTCCGTGAACAACGAGACCCCGGCATCCGCTACCACTACAACGTG  
CTGCCCAAGGCTGGAAGGCTCCCCGCTTCCAGTCTCCATGACCAAGATCTTGGAGCCCTTCCGCACCAAGAACCCTCCGACATCGTGTACTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCACCAAGATCGAGGAGCTGGCGGAGCACCTGCTGAAGTGGGCTTCAACA  
CCCCGACAAGAACAACGAGGAGCCCCCTTCTGTGGATGGCTACGAGCTGACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGACAAG  
GACTCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGCCCTCCAGATCTACCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCT  
GGCGGGCCCAAGGCTGACCGGACATCGTGTCCCTGACCCCGAGGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCTTGAAGGAGCCCGTGCACGGCG  
TGTAACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGCGCAGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCCTGAAG  
ACCGCAAGTACGCCAAGATGGCTCCGCCACACCAACGAGCTGAAGCAGCTGACCGAGCCCTGCAAGAAGTCCGCTGGAGTCCATCGTGTCTGGGG  
CAAGACCCCAAGTTCGGCTGCCATCCTGAAGGAGACTGGGACACTGGTGGACCCACTTGGCAGGCCACTTGGATCCCGGAGTGGGAGTTCGTGA  
ACACCCCTTGTGAAGCTGTGTTACCAGCTGGAGACCGGACCCATCGTGGCGCCGAGACCTTCTACGTGGACGGCCCTCCAACCGGAGACCAAG  
AAGGGCAAGGCGGCTACGTGACCGGCGGCGCCAGAGGTGGTGTCCCTGACCGAGACCAACAACAGAGGCGGAGCTGCAGGCCATCCACCTGGC  
CCTGAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGCTGGTGA  
ACAGATCATCGAGCAGCTGATCCAGAAGGAAAGTGTACTGTCTGGTCCCGCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAAGCTGGT  
TCCCGCGCATCCGCAAGATCCTGTCTCTGGACGGCATCGACAAGGCCAGGAGGAGCACGAGAAGTACCAACAACACTGGCGGCCATGGCTCCGACTT  
CAACTGCCCCCGTGGTGGCCAAAGGATCGTGGCTCCGCAAGTCCAGTCCAGTGAAGGGCGGCGCATGCAAGGCGGAGTGGACTGCTCCCCCGGCA  
TCTGGCAGCTGGACTGCACCCACTGGAGGGCAAGATCATCTGGTGGCCGTGACCTGGCTCCGGCTACATCGAGGCGGAGGTGATCCCCGCGGAGACC  
GGCCAGGACCGCCTACTTCTGAGCTGGCCCGGCTGGCCCGTGAAGATCATCCACACCGCAACCGGCTCCAACTTCACTCCGCGCCCGTGA  
GGCCGCTGTGGTGGCCGGCATCCAGCAGGAGTTCGGCATCCCTACAACCCCAAGTCCAGGGCGGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
TCATCGGCCAGGTGGCGACCGAGCCAGCCTGAAGACCGCGTGCAGATGGCCGTTCATCCACAACCTTCAAGCGCAAGGGCGGCGCATCGGCGGCTAC  
TCCGCCGGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCCCGAGCTGCAGAAGCAGATCAACAAGATCCAGAACTTCCGCGTGTACTACCG  
CGACTCCCGGACCCCGTGTGAAGGGCCCGCCAAAGTGTGTGAAGGGCGGAGGCGGCTGGTGTCCAGGACAACCTCCAGATCAAGGTGGTGGCCCC  
GGCCGAAGGCCAAGATCATCCGGACTACGGCAAGCAGATGGCCGGCGGCGGCTGGTGGCCGAGGACTAA

Fig. 115A

72. 2003 CON F2 pol. PEP

FFRENLAFOOGEARKEFSSEQTRANSPASRELVRRRGDNSLPEAGAERQGTSSLDFFPQITLWQRPLVTIKVGGQLREALLDTGADDDTVLEEDI  
NLPGKWKPKMIGGIGGFIVKVRQYDQIPIEICGQKAI GTVLVGPFPVNI IGRNMLTQIGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTTEE  
KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKRTOQDFWEVQLGIPHPAGLKKKKSVTVLVDVGDAYFSVPLD  
KEFRKYTAFTIP SINNETPGIRYQYNVLPQGWKSPALFQSSMTKILEPFRANKNPEI VIYQYMDLLYVGSDELEIGQHRTKIEELREHLLRWG  
FTTPDKKHQKEPPFLWMGYELHPDKWTVQAIQLPDKSSWTVNDIQKLVGKLNWASQIYPGIRVVKHLCKLLRGAKALTDVPLTAAEAELELAE  
NREILKEPVHGVYDPSKDLIAEIQKQHDQWYQIYQEPHKNLKTGKYARRKSAHTNDVKQLTEVVQKIATEGIVIWGKVPKFRLP IQKET  
WEIWWTEYWQATWIPWEFVNTPPLVKLWYQLETEPIVGAETFYVDGAANRET KLGKAGYVTDGRQKVVPLTETTNQKTELQAIHLALQDS  
GSEVNI VTD SQYALGIIQAHPDKSESELVNOIIEQLIQKERVYLSWVPAHKGIGNEQVDKLVSTGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
MASDENLPPVVAKEIVASCDCQKLGGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVI PAETGQETA YFILLKLAGRWPVKII  
HTDNGSNFTSTVVKAAACWWAGIQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTA VQMAVFIHNEFRKKGIGGYSAGERIIDIIA  
TDIQTKELQKQITKIQNFRVYFRDSRDPVWKGPAKLLWKGE GAVVIQDNNEIKVVPRRKAKIIRDYKGQ MAGDDCVAGRQDED\$

Fig. 116A

73. 2003 CON G pol. PEP

FFRENLAFOOGEAREFFSSEQARANSPTRRRELVRRRGDSPLPEAGAEGKGAISLSFFPQITLWQRPLVTIKVGGQLLEALLDTGADDDTVLEEIIN  
LPGKWKPKMIGGIGGFIVKVRQYDQIILIEISGKKAIGTVLVGPTPINI IGRNMLTQIGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTTEEK  
IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKDKSTKWRKLVDFRELNKRTOQDFWEVQLGIPHPAGLKKKKSVTVLVDVGDAYFSVPLDE  
NFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKSPALFQSSMTKILEPFRANKNPEI VIYQYMDLLYVGSDELEIGQHRAKIEELREHLLRWGE  
FTTPDKKHQKEPPFLWMGYELHPDKWTVQAIQLPDKESWTVNDIQKLVGKLNWASQIYPGIRKVKQLCKLLRGAKALTDI VPLTAAEAELELAEIN  
REILKEPVHGVYDPSKELIAEVQKQGLDQWYQIYQEPYKNLKTGKYAKRGSHTNDVKQLTEVVQKIATESI VIWGKTPKFKLPIRKETW  
EVWTEYWQATWIPWEFVNTPPLVKLWYRLETEPIPGAETIYVDGAANRET KLGKAGYVTDKQKQKIITLLETNQAELQAIHLALQDSG  
SEVNI VTD SQYALGIIQAQDRSESELVNOIIEQLIKKEKVVYLSWVPAHKGIGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRRAM  
ASDFNLPPVVAKEIVASCDCQKLGGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVI PAETGQETA YFILLKLAGRWPVKVIH  
TDNGSNFTSAAVKAAACWWANITQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTA VQMAVFIHNEFRKKGIGGYSAGERIIDIIAS  
TDIQTKELQKQITKIQNFRVYFRDSRDPVWKGPAKLLWKGE GAVVIQDNNEIKVVPRRKAKIIRDYKGQ MAGDDCVAGRQDED\$



## Fig. 116B

161/178

2003\_con\_g\_pol.OPT

TTCTTCGGGAGAACCTGGCCCTCCAGAGGGCGAGCCCGGAGTTCTCCTCCGAGCAGGCCCGCCCAACTCCCCACCCGCGCGAGCTGCGCGTGG  
 CCGCGGACTCCCCCTGCCGAGCGCGGCAAGGCGCCATCTCCCTGTCCCTCCCCAGATCACCTTGTGGCAGCGCCCTGGTGGTACCG  
 TGAAGATCGGCGCCAGCTGATCGAGGCCCTGTGACACCGGCGCGGACCGTGTGGAGGAGATCAACCTGCCCGCAAGTGAAGCCCAAGATG  
 ATCGGCGGCTTCAACAAGTGGCCAGTACGACAGATCTCGGCAAGAAGCCATCGGCACCGTGTGGTGGGCCCCAC  
 CCCCATAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACCTGAACTTCCCCTCCGATCGAGACCGTGGCCGTGAAGTGAAGCCCG  
 GCATGGACGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAAGATCAAGGCCCTGACCCGAGATTCGACCCGAGATGGAGAAGGAGGGCAAGATCTCC  
 AAGATCGGCCCGAGAACCCCTAACACACCCCATTTCCGATCAAGAAGAAGACTCCACCAAGTGGGCAAGTGGTGGACTTCGCGAGCTGAACAA  
 GGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCACCCCGCGCTGAAGAAGAAGTCCGTGACCGTGGACGTGGCGACGCGCTACT  
 TCTCCGTGCCCTGGACGAGAACTTCCGCAAGTACACCGCTTACCATCCCCCTCACCAACAGAGACCCCGGCATCCGCTACCAGTACAACGCTGTG  
 CCCCAGGCTGAAAGGCTCCCCCGCATCTCCAGTCCATGACCAAGATCTTGGAGCCCTTCCGCACCAAGAACCCCGAGATCGTGATCTACCAGTA  
 CATGGACGACTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCGCCAAAGATCGAGGAGCTGCCGAGCACCTGTGCCGTGGGCTTCACCCACC  
 CCGACAAGAAGCACCAAGAGGCCCCCTTCTGTGGATGGCTACGAGCTGACCCCGCAAGTGGACCGTGCAGCCCATCCAGCTGCCCGACAAGGAG  
 TCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACTGGGCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGTGCG  
 CGGCCCAAGCCCTGACCCGACATCGTGCCCTGACCCGAGCGGAGCTGGAGTGGCCGAGAACCCCGAGATCTTGAAGGAGCCCGTGCACGGCGTGT  
 ACTACGACCCCTCCAAAGGAGTGCAGGAGGTCAGAAAGCAGGCTGGACCTGACCTACAGATCTACCAGGACCCCTACAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGCGGGCTCCGCCACACCAAGCTGAAGCAGTGAAGGAGTGGTGCAGAAGATCGCCACCGAGTCCATCGTGATCTGGGGCAA  
 GACCCCAAGTCAAGCTGCCATCCGCAAGGAGACTGGGAGTGTGGTGGACCGAGTACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCTGGTGAAGCTGTGTACCGCTGGAGACCGAGCCCATCCCCGGCGCGAGACTACTACGTGGACGGCGCCCAACCCGCGAGACCAAGCTG  
 GGCAAGGCGGCTACGTGACCGACAAGGCAAGCAGAGATCATCCCTGACCGAGACCAACCAAGAGCCGAGCTGCAGGCCATCCACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACCGCTCCGAGTCCGAGTGGTGAACC  
 AGATCATCGAGCAGTGCATCAAGAAGGAGAGTGTACTGTCTGGTGGCCCGCCACAGGGCATCGGCGGCAACGAGCAGGTGGACAAGTGGTGTCC  
 TCCGGCATCCGCAAGTGTCTTGGACGGCATCGACAAGGCCAGAGGAGCACGAGCGTACCCTCAACTGGCGGCCATGGCTCCGACTTCAA  
 CCTGCCCCCATCGTGGCCAAAGGAGATCGTGGCTCTCGGACAAGTCCAGCTGAAGGGGAGGCCATGCACGGCCAGGTGGACTGTCCCGCGCATCT  
 GGAGCTGACTGACCCACTGGAGGCAAGATCATCTGTGGTGGCTGACCGTGGCTCCGGTACATCGAGGCCGAGGTGATCCCCCGGAGACCCGGC  
 CAGGAGACCGCTACTTCACTCTGAAGTGGCGCGGCTGGCCGTGAAGTGCATCACACCGACAACGGCTCCAATTCACCTCCCGCGCGTGAAGGC  
 CGCTGTGTGGTGGCCAAACATCACCCAGGAGTTCGGCATCCCCCTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGTGAAGAAGATCA  
 TCGGCCAGGTGCGGACCCAGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTATCCACAACCTCAAGCGCAAGGGCGGCATCGGCGGCTACTCC  
 GCCGGGAGCGCATCATCGACATCATCGCCCTCCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGTCCAGAACTTCCCGCTGTACTACCCGGA  
 CTCCCCGACCCCATCTGGAAGGGCCCCCAAGCTGTGTGGAGGGCGGCGGCTGGTGTATCCAGGACAACAACGAGATCAAGGTGGTGGTCCCCCGCC  
 GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGGCGGCTGGCCAGGACGAGGACTAA

Fig. 117A

74. 2003 CON H pol. PEP  
 FFRENLA<sup>Q</sup>QREAR<sup>K</sup>ESPEQARANSPTSRELVRRGDDPLSEAGAEQGTSLSPFPQITLWQRPLVTVKIEGQLREALLDTGADDTVLEEINL  
 PGKWKPKMIGGIGGFIVRQYEQVAIEICGKKAIGTVLVGPTPVNIIGRNILTIQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKI  
 KALTEICIEMEKEGKISKIGPENPYNTPIFAIAKKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVSVDVGDAYFSVPLDKD  
 FRKYTFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPENMIYQYMDDLVYVGSDDLEIGQHRAKIEELRAHLRWRGFT  
 TPDKKHQKEPPFLWGYELHPDKWTVPV<sup>K</sup>LPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQCKLIRGAKALTDIVPLTKEAELELEAENR  
 EILREPVHGVYDDPSKDLIAEIQKQPDQWYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIATESIVIWGKIPKFR<sup>L</sup>PIQKETWE  
 TWWTEHWQATWIPWEFVNTPHLVKLVYQLETEPIAGAEY<sup>Y</sup>VDGAANRETKIGKAGYVTD<sup>R</sup>GKQKVVS<sup>L</sup>TETTNQKTELQAIY<sup>L</sup>ALQDSGL  
 EVNIVTDSQYALGIIQAQPKSESELVNIIEELIKKEKVVLSWVPAHKGIGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEHERYHNNWRAMA  
 SDFNLPPIVAKEIVASCDCQKLG<sup>E</sup>AMHGQVDCSPGIWQ<sup>L</sup>DCTHLEGVILVAHVVASGYIEAEVI<sup>P</sup>AETGQETAYFILLKLAGRWPVKMIHT  
 DNGSNFTSAAVKAACWWADIQQEFGIPYNPQSQGVVESMNKELKIIQVVRDQAEHLRTAVQMAVFIHNFKRKGGIGGYSAGERIIDI<sup>I</sup>ATD  
 IQTKELQKQISKIQKFRVY<sup>R</sup>DRSDRPIWKGPAKLLWKGE<sup>G</sup>AVVIQDNSEIKV<sup>V</sup>PRR<sup>K</sup>KAKIIRDYGKQ<sup>M</sup>AGD<sup>D</sup>CVAGRQ<sup>D</sup>ED\$

Fig. 118A

75. 2003 CON 01 AE pol. PEP  
 FFRENLA<sup>Q</sup>QKAG<sup>E</sup>FSSEQTRANSPTS<sup>R</sup>KLGDGGRDNLLEAAGAERQGTSSSFSFPQITLWQRPLVTVKIEGQLREALLDTGADDTVLEEDI  
 NLP<sup>G</sup>KWKPKMIGGIGGFIVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNILTIQIGCTLNFPISPIDTVPVTLKPGMDGPKVKQWPLTEE  
 KIKALTEICKEMEEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVDVGDAYFSVPLD  
 ESFRKYTFTIPSINNETPGIRYQYNVLPQGWKGS<sup>P</sup>AI<sup>F</sup>QSSMTKILEPFRK<sup>N</sup>PEMVIYQYMDDLVYVGSDDLEIGQHR<sup>T</sup>KIEELRAHL<sup>L</sup>SWG  
 FTTPDKKHQKEPPFLWGYELHPDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQCKLIRGAKALTDIVPLTKEAELELAE  
 NREILKTPVHGYYDPSKDLVAEVQKQGDQWYQIYQEPFKNLKTGKYARKRS<sup>A</sup>HTNDV<sup>R</sup>QLTEV<sup>V</sup>QKIATESIVIWGKTPKFR<sup>L</sup>PIQRET  
 WETWMEY<sup>W</sup>QATWIPWEFVNTPHLVKLVYQLEKDPYVGAETFYVDGAASRETKL<sup>G</sup>KAGYVTD<sup>R</sup>GKQKVVS<sup>L</sup>TETTNQKTELHAIHLALQDS  
 GSEVNI<sup>V</sup>TD<sup>S</sup>OYALGIIQAQPKSESEV<sup>N</sup>OIIIEELIKKEKVVLSWVPAHKGIGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEHERYHNSWRT  
 MASDFNLPPIVAKEIVANCDKQKLG<sup>E</sup>AMHGQVDCSPGIWQ<sup>L</sup>DCTHLEGVILVAHVVASGYIEAEVI<sup>P</sup>AETGQETAYFILLKLAGRWPVKVI  
 HTDNGSNFTSAAVKAACWWANVRQEFGIPYNPQSQGVVESMNKELKIIQVVRDQAEHLK<sup>T</sup>AVQMAVFIHNFKRKGGIGGYSAGERIIDI<sup>I</sup>IA  
 TDIQTKELQKQITKIQFRVY<sup>R</sup>DRSDRPIWKGPAKLLWKGE<sup>G</sup>AVVIQDN<sup>S</sup>DIK<sup>V</sup>VPRR<sup>K</sup>KAKIIRDYGKQ<sup>M</sup>AGD<sup>D</sup>CVAGRQ<sup>D</sup>ED\$

Fig. 117B

2003\_CON\_H\_pol.1.OPT

TTCTCCGGAGAACTGGCCCTCCAGCAGCGGAGGCCCGCAAGTTCTCCCGGAGAGGCCCGCCGCAACTCCCCACCTCCCGGAGCTGCGCGTGG  
CCGGCGGACGACCCCTGTCCGAGCCGGCGGAGGGCCAGGGCACCTCCCTGTCTTCCCGCAGATCACCTGTGGCAGCGCCCTGGTGACCCGTGA  
AGATCAGGGCCAGCTGCGCGAGGCCCTGTGGACACCGCGCCGACACCGTCTGGAGGAGATCAACCTGCCCGGCAAGTGAAGCCCAAGATGATC  
GGCGGCATCGCGGCTTCATCAAGGTGCGCCAGTACGAGCAGGTGCCATCGAGATCTGCGGCAAGAGCCATCGGCACCGTGTGGTGGCCCGCCACCC  
CGTGAAACATCATCGGCCCGCAACATCTGTACCCAGATCGGCTGACCCCTGAACTTCCCAATCTCCCGCATCGAGACCGTCCCGTGAAGTGAAGCCCGGCA  
TGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCATCGAGATGGAGAAGGAGGCAAGATCTCCAAG  
ATCGGCCCGGAAACCCCTACAACACCCCATCTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAACAAGCG  
CACCCAGGACTTCTGGAGGTGCAGTGGGCATCCCCACCCCGCCCTGAAGAAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAACAAGCG  
CCGTGCCCTGGACAAAGACTTCCGCAAGTACACCGCTTACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAGTACAACGTGTGCCCC  
CAGGCTGGAAGGGCTCCCGCCGCTTCCAGTCTCCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAACCCCGAGATGATCATCTACCAGTACAT  
GGACGACTGTACGTGGCTCCGACTGGAGATCGGCCAGCACCGCCCAAGATCGAGGAGCTGGCGCCACCTGCTGCGCTGGGGCTTACCACCCCGG  
ACRAGAAGCACAGAAAGGCCCTTCCCTGTGGATGGGTACGAGTGCACCCCGACAAAGTGGACCGTGCAGCCCGTGAAGCTGCCGAGAAGACTCC  
TGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACTGGCCCTCCAGATCTACCCCGCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGGG  
CGCAAGCCCTGACCGACATCGTGGCCCTGACCAAGAGGCGAGCTGGAGCTGGCCGAGAACCCCGGAGATCTCTGCGGAGCCCGTGCACGGCGTACT  
ACGACCCCTCCAAGACTGTATCGCCGAGATCCAGAAGCAGGCCCCGACCAAGTGGACTTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGAACCGGC  
AAGTACGCCAAGATGCGCACCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGTCCATCGTGGACGGCCGCAACCGGAGACCAAGATCGGC  
CCCCAAGTCCCGCTGCCATCCAGAAGGAGACCTGGAGACCTGGTGAACCGGACCTGACCGGAGCCACCAACCAAGAACCGAGCTGCAGGCCATCTACCTGGCCCTGCA  
AAGCCCGCTACGTGACCGACCGGCAAGCAGAAGGTGGTGTCCCTGACCGGAGCCACCAACCGAGACCGAGCTGCAGGCCATCTACCTGGCCCTGCA  
GGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCCAAGTACGCCCTGGSCATCATCCAGGCCAGCCGACAAAGTCCGAGTCCGAGTGGTGAACCCAGA  
TCATCGAGGAGCTGATCAAGAAGGAGAAAGTGTACTTCTCTGGTGGTGGCCCAAGGGCATCGGGCGCAACCGAGCAGGTGGTGAACCCAG  
GGCATCCGCAAGGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGGACGAGCTACCAACAACCTGGCGGCCATGGCCTCCGACTTCAACCT  
GCCCCCATCGTGGCCAAAGGAGATCGTGGCCCTCTGGACAAGTCCAGTGAAGGGCGAGGCCATGCACGGCCAGTGGACTGCTCCCCCGGCATCTGGC  
AGCTGGACTGCACCCACTGGAGGGCAAGTGTATCTGGTGGCCGTGACCTGGCCCTCCGGTACATCGAGGCCGAGGTGATCCCCCGGACACCGGCCAG  
GAGACCGCTACTTTCATCTGAAGCTGGCCCGTGGCCCTGACCTGGCCCTCCGGTACATCGAGGCCGAGGTGACTGCTCCCCCGGCATCTGGC  
CTGTGGTGGCCGACATCCAGCAGGAGTTCGGCATCCCCTACAACCCCGAGTCCCAAGTCCAGAACCTTCAAGCGCAAGGGCGGCATCGGGCGCTACTCCGCC  
GCCAGTGGCCGACCCAGCCGAGCCTGGCACCCCGCTGAGATGGCCGTGTTTCATCCACAACCTTCAAGCGCAAGGGCGGCATCGGGCGCTACTCCGCC  
GGCGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGTGCAGAAGCAGATCTCCAAGATCCAGAAGTTCGGCGTGTACTACCGCGACTC  
CCCGACCCCATCTGGAAGGGCCCCCAAGCTGCTGGAAGGGCGAGGCCCGCTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGTCCCGCCCGCA  
AGGCCAAGATCATCCGGACTACCGCAAGCAGATGGCCCGCCGACGACTCGGTGGCCCGCCAGGACGAGGACTAA





165/178

Fig. 119A

76. 2003 CON 02 AG pol.PEP  
 FFRENLAFOQGEARKEFSSEQTGNTSPTRELWDGGRDNLSEAGTEGQTISSFNFPQITLWQRPLVTVRIGGQLIEALLDTGADDTVLEEI  
 NLP GKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTEE  
 KIKALTDICTEMEKEGKISKIGPENPYNTPVFAIKKSDTKWRKLVDFRELNKRTODFWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVPLD  
 KDFRKYTAFITPSVNNETPGIRYQYNVLPQGWKSPAI FQASMTKILEPFRTKNPEI VI YQYMDL YVGS DLEIGQHRAKIEELREHLLRWG  
 FTT PDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLRGAKALTDIVTLTEEALELAE  
 NRELLKEPVHGVYDPTKDLIAEIQKQGDQWTYQIYQEPFKNLKTGKYAKMRS AHTNDVKQTEVQKVATESI VIWGTKPKFRLPIQRET  
 WEAWMEYQATWIPWEFVNTPLVKLWYQLEKDPVGAETFYVDGAANRETKLGKAGYVDRGRQKVVSLTETTNQKTELHAIHLALQDS  
 GSEVNI VTD SQYALGIIQAQDRSESELVNIIEKLI EKDKVYLSWVPAHKGIGGNEQVDKLVSN GIRKVLFDGIDKAQEEHERYHSNWRA  
 MASDFNLPPIVAKEIVASCDCQKLG EAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVI PAETGQETAYFILLKLAGRWPVKVI  
 HTDNGSNFTSAAVKAACWVANVTQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNEFRKKGIGGYSAGERIIDIIA  
 SDIQTKELQKQITKIQNFVYYRDSRDP IWKGP AKLLWKGE GAVVIQDN SDIKV VPRRKAKIIRDY GKQ MAGD DDCVAGRQDED\$

Fig. 120A

77. 2003 CON 03 AB pol.PEP  
 FFRENLAFOQGEARKEFSSEQTRAI SPTSRLKLDGGRDNLPEPETGTERQGTASSFNFPQITLWQRPLVTVRIGGQLKEALLDTGADDTVLEDI  
 NLP GKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTEE  
 KIKALTDICTEMEKEGKISKIGPENPYNTPVFAIKKSDTKWRKLVDFRELNKRTODFWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVPLD  
 ODFRKYTAFITPSVNNETPGIRYQYNVLPQGWKSPAI FQSSMTKILEPFRKQNP EIVIYQYMDL YVGS DLEIGQHRTKIEELREHLLRWG  
 FTT PDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLRGAKALTEVIPLTAEALELAE  
 NRELLKEPVHGVYDPTSKDLVAEIQKQGGQWTYQIYQEPFKNLKTGKYARLRGAHTNDVKQTEAVQKIATESI VIWGTKPKFRLPIQKET  
 WETWTEYWOATWIPWEFVNTPLVKLWYQLEKEPIVGAETFYVDGAANRETKSGKAGYVDRGRQKVVSLTDTTNQKTELQAIHLALQDS  
 GLEVNI VTD SQYALGIIQAQDKSESELVSIIEQLIKKEKVIYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFDGIDKAQEAHEKYHSNWRA  
 MASDFNLPPVVAKEIVASCDCQKLG EAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVI PAETGQETAYFVLKLAGRWPVKII  
 HTDNGSNFISTAVKAACWVAGIKQEFGIPYNPQSQGVVESMNKQLKQIIGQVRDQAEHLKTAVQMAVFIHNEFRKKGIGGYSAGERIIDIIA  
 TDIQTKELQKQIKI QNFVYYRDSRDP IWKGP AKLLWKGE GAVVIQDN NDIKV VPRRKAKIIRDY GKQ MAGD DDCVAGRQDED\$

Fig. 119B

2003\_CON\_02\_AG\_pol.1.OPT

TTCTCCGGAGAACCTGGCCTCCAGCAGGGGCGGCAAGTTCTCTCCGAGCAGACCGGCACCAACTCCCCACCTCCCGGAGCTGTGGGACGG  
CGGCCGGACAACCTGCTGTCCGAGCCGGCACCGAGGGCCAGGGACCACTCTCTTCAACTTCCCAGATCACCTGTGGCAGGCCCCCTGGTGA  
CCGTGGCATCGGCGGCAGTGTGAGGCCCTGCTGGACACCGGCGCCGACACCCGTCTGGAGGATCAACTGCCCGCAAGAGCCATCGGCACCGTGTGGTGGCC  
ATGATCGGCGGCATCGGCGGCTCATCAAGGTGCCAGTACGACCCAGATCTGATCGAGATCTGGCAAGAGCCATCGGCACCGTGTGGTGGCC  
CACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGATCGGCTGCACCCCTGAACTTCCCCATCGAGACCGTGCCTGTAAGCTGAAGC  
CCGCATGGACGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGACATCTGACCGAGATGGAGAAGAGGGCAAGATC  
TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAA  
CAAGCGCACCCAGGACTTCTGGAGTGCAGTGGGCATCCCCACCCCGCCGCTGAAGAAGAAGTCCGTGACCGTGGACGTGGCGGACGCCCT  
ACTTCTCCGTGCCCTGGACAAGGACTTCCGCAAGTACACCGCCTTCAACATCCCCCTCCGTGAACAACGAGACCCCGGCATCCGTACCAGTACAACGTTG  
CTGCCCCAGGGCTGGAAGGCTCCCCCGCATCTCCAGGCCCTCCATGACCAAGATCTTGGAGCCCTTCCGCAACAAGACCCCGAGATCGTGTATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCCCAAGATCGAGGAGCTGGCGGAGCACCTGTGCGCTGGGCTTCAACA  
CCCCGACAAGAAGCACCAAGAAGGAGCCCCCTTCTGTGATGGCTACGAGTGCACCCCGACAAGTGGACCTGCAGCCCATCCAGTGCSCCGAGAAG  
GACTCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAACCTGGCCCTCCAGATCTACGCCGCATCAAGTGAAGCAGCTGTGCAAGCTGCT  
GGCGGCGCCAAAGCCCTGACCGACATCGTGAACCTGACCGAGGAGCCGAGCTGGAGCTGGCCGAGAACCGGAGATCTTGAAGGAGCCCGTGCACGGCG  
TGTACTACGACCCCAAGGACCTGTATCGCCGAGATCCAGAAGCAGGGCCAGGACCAAGTGGACCTACAGATCTACAGAGCCCTTCAAGAACCTGAAG  
ACCGCAAGTACGCCAAGATGCGCTCCGCCACACCAACGACTGAAGCAGTGAAGGCTGGTGCAGAGGTGGCCACCGAGTCCATCGTGTATCTGGGG  
CAAGACCCCAAGTCCCGCTGCCATCCAGCGGAGACTGGAGGCCCTGGTGGATGAGTACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCGTGA  
ACACCCCGCTGGTGAAGCTGTGTACCAGCTGGAGAAGGACCCCATCGTGGCGCCGAGACCTTCTACGTGGACGGCCCGCCAAACCGCGAGACCAAG  
CTGGGCAAGCCCGCTACGTGACCGACCGCGCCGCAAGGTTGGTCCCTGACCGAGACCCAAACAGAGACCGGAGCTGCACGCCATCCACCTGGC  
CCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGAACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACCGCTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGAAGCTGATCGAGAAGGACAAGTGTACTGTCTGGGTGCCCGCCACAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTG  
TCCAACGGCATCCGCAAGTGTCTTCTGGACGGCATCGACAAGCCAGGAGGACACGAGCCCTACCACTCCAATGGCGGCCCATGGCCCTCCGACTT  
CAACTGCCCCCATCGTGGCCAAAGGATCGTGGCTCTCGACAAAGTGCAGCTGAAGGGCGAGCCCATGCAAGGCTGGACTGTCTCCCCCGGCA  
TCTGGAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGGTGGCCCTGCACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCGCGGAGACC  
GGCCAGGAGACCGCTACTTCTGAAAGTGGCCGCGCTGGCCCTGAAAGTGTATCCACACCGACAACGGCTCCAACCTTCACTCCGCGCCGCTGAA  
GGCCGCTGTGGTGGCCCAAGTGAACCCAGGAGTCCGGCATCCCCTACAACCCAGTCCCAGGGCTGGTGGATCCATGAACAAGGAGCTGAAGAAGA  
TCATCGGCCAGGTGGCCGACCCAGGACCTGAAGACCGCCGCTGACAGTGGCCGTGTTCATCCACAACCTCAAGGCAAGGGCGGCATCGGCGGCTAC  
TCCGCGGGGAGCGCATCATCGACATCATCGCTCCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCAACAAGATCCAGAACCTTCCGCGTGTACTACCG  
CGACTCCCGGACCCCATCTGGAAGGGCCCGCCCAAGCTGTGTGGAAGGGCGAGGGCCCGCTGGTGTATCCAGGACAACCTCCGACATCAAGGTGGTGGCCCC  
GCCGCAAGGCCAAGATCATCCGGACTACGGCAAGCAGATGGCCCGGCGGACGACTGCGTGGCCCGCCCGCCAGGACCGAGGACTAA

167/178

## Fig. 120B

2003\_CON\_03\_AB\_pol.OPT  
 TTCTCCGGAGAACCTGGCCCTTCAGCAGCGGAGGCCCGCAAGTTCTCTCCGAGCAGACCCGGCCCATCTCCCCACCTCCCGCAAGCTGTGGGACGG  
 CGCCCGGACAAACCCCTGCCCGAGACCGGCACCGAGCCCGAGGCAACCGCTCTCTCAACTTCCCCAGATCACCCCTGTGGCAGCGCCCTGGTGA  
 CCGTGGCATCGGGGCCAGCTGAAGGAGGCCCTGTGGACACCGGGCCGACGACACCGTGTGGAGGACATCAACCTGCCCGCAAGTGAAGCCCAAG  
 ATGATCGGGGCATCGGGCTTTCATCAAGGTGGCCAGTACGACAGATCTGTGATCGAGATCTGGGGCAAGAGGCCATCGGCACCGTGTGGTGGCC  
 CACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGTGGCTGCAACCTGAATCTCCCAATCTCCCCATCGAGACCGTGGCCCTGACCCCTGAAGC  
 CCGGCATGGACGGCCCCAAAGTGAAGCAGTGGCCCTGACCGAGGAGAAAGATCAAGGCCCTGACCCGACATCTGCAAGGAGATGGAGAGGAGGGCAAGATC  
 TCCAAGATCGGGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAGAAAGACTCCACCAAGTGGCCAAAGTGGTGGACTTCCCGAGCTGAA  
 CAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCACCCCGGCTGAAGAAAGAAAGTCCGTGACCGTCCGTGGACCGTGGCGGACGCCCT  
 ACTTCTCCGTGCCCTGGACAGGACTTCCGCAAGTACACCGCCTTACCAATCCCTCCACCAAGATCTGGAGCCCTTCCGCAAGCAGCAACCCCGGATCGT  
 CTGCCCCAGGGTGAAGGCTCCCCCGCATCTCCAGTCTCCATGACCAAGATCTGGAGCCCTTCCGCAAGCAGCAACCCCGGATCGTGTATCTACCA  
 GTACATGGACGACTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGGCGGAGCACCTGCTGGCTGGGGCTTACCCA  
 CCCCCAAGAAGCACCAAGAGGCCCTTCTGTGGATGGCTACGAGTGCACCCCGCAAGTGGACCGTGCAGCCATCAAGTGCAGTGGTGGTGGTGGTGG  
 GACTCTGGACCTGAACGACATCCAGAAGCTGTGGCAAGCTGGAGCTGGAGCTGGCCAGGCTGACAGTCCAGATCTACAGGATCTACAGGAGCCCTTCAAGAACCTGAAG  
 GCGGGCCCAAGCCCTCAAGGACTGTGGTGGCCGAGATCCAGAAGCAGGGCCAGGCTGACAGTCCAGATCTACAGGATCTACAGGAGCCCTTCAAGAACCTGAAG  
 TGTACTAGACCCCTCAAGGACTGTGGTGGCCGAGATCCAGAAGCAGGGCCAGGCTGACAGTCCAGATCTACAGGATCTACAGGAGCCCTTCAAGAACCTGAAG  
 ACCGGCAAGTACGCCCGCTGGCGGCCCAACAGACTGAAGCAGCTGAAGCAGCTGACCGAGCCGTGCAGAAATCGCCACCGAGTCCAGGAGCCCTTCAAGAACCTGAAG  
 CAAGACCCCAAGTCAAGCTGCCATCCAGAAGGAGACTGGGAGACTGGGAGCTGGAGCTGGCCAGGCTGACAGTCCAGGATCTACAGGAGCCCTTCAAGAACCTGAAG  
 ACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGCTGCTCCCTGACCCGACACCAACCAAGAGTCCAGGAGCCCTTCAAGAACCTGAAG  
 TCCGGCAAGGCCGCTACGTGACCGACCGCGGCCCGCAGAGGTGGTGTCCCTGACCCGACACCAACCAAGAGTCCAGGAGCCCTTCAAGAACCTGAAG  
 CCTGAGGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCAGTACCGCTGGCATCATCCAGGCCCGAGCCGACAAAGTCCGAGTCCGAGTGGTGT  
 CCCAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGACTGGCTGGGTGCCCGCCCAAGGGCATCGGGCGCAACGAGCAGGTGGACAAGTGGTGGT  
 TCCGGCCGCATCCGCAAGGTGCTTCTCCGACGGCATCGACAAGGCCAGGAGGCCCAAGAGTACCACTCCAATGGCGGCCATGGCCCTCCGACTT  
 CAACCTGCCCGCTGGTGGCCAGGAGATCGTGGCTCTCTGCGACAAGTCCAGTCCAGTGAAGGGCGAGGCCATGCAAGGCAAGTGGACTGCTCCCCCGGCA  
 TCTGGCAGCTGGACTGCACCCACTGGAGGGCAAGATCATCTGTGGTGGCCGTGACCTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCGCGGAGACC  
 GGCCAGGAGACCGCTACTTCTGTGCTGAAGCTGGCCCGCTGGCCATCCCTACAACCCAGTCCAGTGGTGGAGTCCATGAACAAGCAGCTGAAGCAGA  
 GGCCCGCTGTGGTGGCCCGCATCAAGCAGGAGTTCGGCATCCCTACAACCCAGTCCAGTGGTGGAGTCCATGAACAAGCAGCTGAAGCAGA  
 TCATCGGCCAGGTGGCCGACCGAGCACCTGAAGACCGCTGCAGATGGCCGTGTTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGGGCGCTAC  
 TCCGGCCGAGCGCATCATCGACATCATCGCCACCGGATCCAGACCAAGGAGTGCAGAAAGCAGATCATCAAGATCCAGAACTTCCGCGTGTACTACCG  
 CGACTCCCGGACCCCTCTGGAAAGGCCCGCCAAAGCTGTGGAAAGGCCGAGGGCCGCTGGTATCCAGGACCAACAGCAGATCAAGTGGTGGTGGCC  
 GCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGCGAGCTGCGTGGCTCCCCCGGACGAGGACTAA

168/178

Fig. 121A

78. 2003 CON 04 CPX po1.PEP  
 FFRENVAFQOGEAREKFSSEQARANSPTRELRVRRGDSPLPEAGAEGQGAISISFPQITLWQRPVTVIRIGGQIEALLDTGADDTVLEEDIN  
 LPGKWKPKMIGGIGGFIKVRQYDQIPFIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGTLNFPISPIETVPVKLKPMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKNDSTRWRKLVDFRELNKRTODEFWEVQLGIPHPAGLKKKSVTVLDVGDAYFVSVPDLP  
 EFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSFAIFQCSMTKILEPFRTKNPEIIVIQYMDLTVGSDLEIGQHRAKIEELREHLLRWGF  
 STPDKKHQKEPPFLWMGYELHPDKWTVQPIQLAEKDSWTVNDIQKLVGKLNWASQIYPGKVKQLCKLLRGAKALTDIVPLTTEAELELAEN  
 REILKEPVHGAYYDPSKDLIAEIQKQGQWTYQIYQEPYKNLKTGKYAKTRSAHTNDVROLTAEVQKIAMECIVIWGKTPKFRPLPIQKETW  
 DTWTEYWQATWIPWEFVNTPLVKKLWQLETDPIAGAETFYVDGAASRETKQKAGYVTDGRQKVVSLSETTNQKTELQAIYLALQDSDG  
 SEVNI VTSQYALGIIQAQPDRESIDLNVQIEQLIQKDKVYLSWVPAHKGIGGNEQVDKLVNSGIRKVLFLDGDIDKAQEEHEKYHNNWRAM  
 ASDFNLPVVAKEIVASCNKQKLGKGEAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKIIH  
 TDNGSNFTSAAVKAACWWADIQEEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAS  
 DIQTKELQKQITKIQFRVYRDSRDP IWKGPALLWKGEAVVIQDNSDIKVVPRRKAKI IRDYGKQ MAGDDCVAGRQDED\$

Fig. 122A

79. 2003 CON 06 CPX po1.PEP  
 FFRENLAFOQGEAREKFSSEQARANSPTRELRVRRGDSPLPEAGAEGQGAISISFPQITLWQRPVTVIRIGGQIEALLDTGADDTVLEEDIN  
 LPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGTLNFPISPIETVPVKLKPMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKNDSTRWRKLVDFRELNKRTODEFWEVQLGIPHPAGLKKKSVTVLDVGDAYFVSVPDLP  
 DFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSFAIFQCSMTKILEPFRTKNPEIIVIQYMDLTVGSDLEIGQHRAKIEELREHLLRWGF  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPKDQSWTVNDIQKLVGKLNWASQIYPGKVKQLCKLLRGAKALTDIVPLTAEAELELAEN  
 REILKEPVHGAYYDPSKDLIAEIQKQGQWTYQIYQEPHKNLKTGKYARIKSAHTNDVKQTEAVQKIALESIVIWGKTPKFRPLPIQKETW  
 ETWTEYWQATWIPWEFVNTPLVKKLWQLETEPIVGAETFYVDGAANRETKKAGYVTDGRQKVVSLSETTNQKTELQAINLALQDSDG  
 SEVNI VTSQYALGIIQAQPDKSESELVQIEQLIKKEKVVLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGDIDKAQEDHERYHNSWRAM  
 ASDFNLPPIVAKEIVASCNKQKLGKGEAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH  
 TDNGSNFTSAAVKAACWWANITQEEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAS  
 DIQTKELQKQITKIQNERVYRDSRDP IWKGPALLWKGEAVVIQDENSEIKVVPRRKAKI IRDYGKQ MAGDDCVAGRQDED\$

169/178

## Fig. 121B

2003\_con\_04\_cpx\_pol.0pt  
 TTCTCCGGAGAACGTGGCCTTCCAGCAGCGGAGGCCCGCAAGTTCTCTCCGAGCAGGCCCGGCAACTCCCCCGCCCGCGAGCTGCCGGACGA  
 GCGGGGACAACTGTCTCCGAGCCCGGACCGAGGGCCAGGACCACTCTCTTCAACTTCCCCAGATCACCTGTGGCAGGCCCTGGTGACCA  
 TCAAGATCGGGCCAGATCCGGAGGCCCTGTGGACACCGGGCCGACGACACCGTGTGGAGAGATCAACTGCCCCGCAAGTGAAGCCCAAGATG  
 ATCGGGCCATCGGGCTTCAVCAAGGTGGCCAGTACGACAGATCCCAATCGAGATCTGGGGCAAGAGGCCCATCGGCACCGTGTGGTGGCCCCAC  
 CCCCCTGAACATCATCGGGCCCAACTGTACCCAGTGGCCCTGACCCCTGAACCTTCCCATCTCCCATCGAGACCGTGCCGTGAAGCTGAAGCCCCG  
 GCATGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGAGATCTGACCGAGATGGAGAGGAGGCCAAGATCTCC  
 AAGATCGGCCCGAGAACCCCTAACACACCCCATCTTCCCATCAAGAAGAACTCCACCCGTGGCGCAAGCTGGTGGACTTCCCGCGAGCTGAACAA  
 GCGCACCCAGGACTTGGGAGGTGCAGTGGGCATCCCCACCCCGCCCTGAAGAAGAAAGTCCGTGACCCGTGTGGACGTGGCGGACCGCCTACT  
 TCTCCGTGCCCCCTGGACCCCGAGTCCGCAAGTACACCGCCTTACCATCCCTCCACCAACAGAGACCCCGGCATCCGCTACCAAGTACAACGCTGCTG  
 CCCAGGGCTGGAAGGCTCCCCCGCATTTCCAGTGTCCATGACCAAGATCTTGGAGCCCTTCCGCAACCAAGAACCCCGAGATCGTGATCTACCAGTA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGGCAGCACCGCCCAAGATCGAGGAGCTGGCGAGCACCTGCTGGCTGGGGCTTCTCCACCC  
 CCGACAAAGACACCAAGAGGCCCTTCTCTGTGGATGGCTACGAGTGCACCCGACAAAGTGGACCGTGCAGCCCATCCAGCTGGCGGAGAGGAC  
 TCCTGGACCCGTGAACGACATCCAGAAGCTGGTGGCAAGTGAACCTGGGCTCCAGATCTACCCGGCATCAAGGTGAAGCAGCTGTCAAGCTGCTGG  
 CGGGCCAAGGCCCTGACCGACATCGTCCCCCTGACCAACGAGCCGAGCTGGAGCTGGCCAGATCTACCCGGCATCAAGGTGAAGCAGCTGTCAAGCTG  
 ACTACGACCCCTCAAGGACCTGATCGCCGAGATCCAGAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTACAAGAACCTGAAGACC  
 GGAAAGTACGCCAAGACCCGCTCCGCCATCCAGAAGGAGACCTGGGACACTGGTGGACCGAGTGCAGAAAGATCGCCATGGAGTGCATCGTGATCTGGGGCAA  
 GACCCCAAGTCCGCTGCCCCATCCAGAAGGAGACCTGGGACACTGGTGGACCGAGTACCTGGTGGACCGGACCTTCTACGTGGACGGCCCGCTCCCGGAGACCAAGCAG  
 CCCCCCTGGTGAAGCTGTGTACAGCTGGAGACCGACCCCATCCGCGGGCCGAGACCTTCTACGTGGACGGCCCGCTCCCGGAGACCAAGCAG  
 GGCAAGCCGGCTACGTGACCGACCGGGCCAGAAAGTGTCTCTGTCCGAGACCAACCAAGAGACCGGAGCTGCAGGCCATTAACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGAACCGACTCCAGTACGCCATCGGCATCAACAGGCCAGCCCGACCCGCTCCGAGTCCGACCTGGTGAACC  
 AGATCATCGAGCAGCTGATCCAGAAGGACAAAGTGTACTGTCTGGACGGCATCGACAAGGCCCAGAGGAGCACGAGAAGTACCACAACACTGGCCGCCATGGCCTCCGACTCAA  
 AACGGCATCCGCAAGTGTCTTCTGGACGGCATCGACAAGGCCCAGAGGAGCACGAGAAGTACCACAACACTGGCCGCCATGGCCTCCGACTCAA  
 CCTGCCCCCGTGGTGGCCAAAGGATCGTGGCTCTTGCAACAAGTCCAGTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCT  
 GGCAGCTGGACTGCAACCCACTGGAGGGCAAGATCATCTGGTGGCCGTGCACTGGCTCCCGGTACATCGAGGCCGAGGTGATCCCGCCGAGACCCGGC  
 CAGGAGACCCCTACTTCACTTCAAGCTGGCCGGCCGCTGGCCCGTGAAGATCATCCACCCGACAAAGGCCCAACTCACTCCGCCCGCTGAAGGC  
 CGCCTGCTGGTGGCCGACATCCAGCAGGAGTTCGGCATCCCTTACACCCCGAGTCCAGGGCTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCCGACCGGAGCACCTGAAGACCCCGTGCAGATGGCCGTTCATCCACAACCTCAAGCGCAAGGGCCGATCGGGGGCTACTCC  
 GCCGGCAGCGCATCATGCATCGACTCCGACATCCAGACCAAGGAGCTGCAGAGCAGATCCCAAGATCCAGAACTTCCGCGTGTACTACCCCGGA  
 CTCCCCGACCCCATCTGGAAGGGCCCGCCAAAGTGTGTGGAAGGGCAGGGCCCGTGGTGTCCAGGACAACTCCGACATCAAGGTGGTGGTCCCCCGCC  
 GCAAGGCCAAGATCATCCCGACTACGGCAAGCAGATGGCCCGGACGACTGCTGGCCGGCCCGCAGGACGAGACTAA

Fig. 122B

2003\_CON\_06\_CPX\_pol.1.OPT

TTCCTCCGGAGAACCTGGCCCTCCAGCAGGGCGAGGCGCCGGAGTTCTCCTCCGAGCAGGCCCGCCCAACTCCCCACCCGCGCGAGCTGCGCGTGGCG  
 CCGCGGCACTCCCCCTGCCCGAGGGCCGAGGGCCAGGGCCGCAATCTCCCTGTCCCTCCCCAGATFACCCCTGTGGCAGCGCCCTGGTGGTACCCG  
 TGGCATTCCGGCGCCAGCTGATCGAGGCCCTGTGGACACCGGGCCGACGACACCGTGTGGAGGACATCAACTGCCCCGCAAGTGAAGCCCAAGATG  
 ATCGCGGCATCGCGGCTTTCATCAAGGTGCGCCAGTACGACAGATCCTGATCGAGATCTGGGCAAGAGGCCATCGGCACCGTGTGGTGGGCCCCAC  
 CCCCCTGAACATCATCGGCGCAACATGCTGACCCAGATCGGTCGACCCCTGAACCTCCCATCTCCCATCGAGACCGTGGCCCTGAAGCTGAAGCCCGG  
 GCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCCGAGGAAAGATCAAGGCCCTGACCCGAGATCGACCCGAGATGGACCCGAGAGGGCAAGATCTCC  
 AAGATCGGCCCCGAGAACCCCTACAACACCCCTATCTCGCCATCAAGAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGGAGCTGAACAA  
 GCGCACCCAGGACTTCTGGGAGGTGCAAGTGGGCATCCCCACCCCGCGCCTGAAGAAGAAAGTCCGTGACCCGTGGACCGTGGCGGACCGCCTACT  
 TCTCCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTACCAATCCCTCCATCAACAACGAGACCCCGCGCATCCCGTACCAGTACAACCGTGTG  
 CCCCAGGGCTGAAAGGCTCCCCCGCATCTCCAGTCTCATGATCAAGATCTTGAGCCCTTCCGCATCAAGAACCCCGAGATCGTGATCTACCAGTA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCCCAAGATCGAGGAGCTGGCGAGCACCTGTGAAGTGGGCTTCAACCCACCC  
 CCGACAAGAGCACCAAGAGGACCCCTTCTGTGGATGGCTACGAGCTGACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGACAAGGAC  
 TCCCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGCCCTCCAGATCTACCCCGCATCAAGGTGAAGCAGCTGTCAAGCTGTGCGG  
 CCGCGCAAGGCCCTGACCGCATCTGTGCCCTGACCCCGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCCGTGCACGGCGTGT  
 ACTACGACCCCTCAAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCAAGAACCTGAAGACC  
 GGCAAGTACGCCCGCATCAAGTCCGCCACACCAACGACGTGAAGCAGCTGACCGAGCCGTGCAGAAGATCGCCCTGGAGTCCATCGTATCTGGGGCAA  
 GACCCCAAGTTCGGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGGATCTGGCAGGCCACTTGGATCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGACCGGACCCATCGTGGCGCCGAGACCTTACGTGGACGGCGCCCAACCGCGAGACCAAGAAAG  
 GGCAAGGCCGGTACGTGACCGACCGCGGCCGAGAGGTGGTTCCTTGACCGAGACCCCAACAGAGACCGAGCTGCAGGCCCATCAACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCCGACTCCAGTACGCCCTGGGCATCATCCAGGCCACAGCCCGACAAGTCCGAGTCCGAGTGGTGAACC  
 AGATCATCGAGCAGTGTCAAGAAGGAGAGGTGTACCTGTCTGGTGGCCGCCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGTGGTGTCC  
 ACCGGCATCCGCAAGGTGTCTCTGGACGGCATCGACAAGGCCAGGAGGACCAAGGCTACCACTCCCAACTGGCGGCCCATGGCCCTCCGACTTCAA  
 CCTGGCCCCCATCGTGGCCAAAGGATCGTGGCCCTCCCTGGACAAGTGGCAGCTGAAGGGGAGGCCCATGCACGGCCAGGTGGACTGCTCCCCGGCATCT  
 GCGAGTGGACTGCACCCACTGGAGGGCAAGATCATCTGTGGCTGCACGTGGCCCTCCGCTACATCGAGGCCGAGGTGATCCCCGGCAGACCCGGC  
 CAGGAGACCGCCTACTTTCATCTGAAGTGGCCCGCCGCTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAACCTCCACCTCCGCGCCGCTGAAGGC  
 CGCCTGTGTGGCCCAACATCACCCAGGAGTCCGCATCCCTACAACCCCAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCGGACCCAGCCGAGCACCTGAAGACCCCGCTGAGATGGCCGTGTTCATCCACAACCTCAAGCGCAAGGGCCGATCGCGGGCTACTCC  
 GCGGGCAGGCGCATCATCGACATCATCGCCTCCGACATCCAGAACCAAGGAGCTGCAGAGCAGATCACCAAGATCCAGAACTTCCCGCTGTACTACCCGGA  
 CTCCCGGACCCCATCTGGAAGGGCCCCGCAAGCTGTGTGGAAGGGCGGCGCCGCTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGGCCCCGCGC  
 GCAAGGCCAAGATCATCCGGGACTACGGCAAGCAGATGGCCGGCGGACGACTCGTGGCCGGCCCGCAGGACCGGACTAA

171/178

## Fig. 123A

80. 2003 CON 08 BC pol.PEP  
 FFREILAFQGEAREFPPEQTRANSPTSRELVQRGDNPSSEAGTERQGTINFPQITLWQRPLVSIKVGQIKEALLDTGADDTVLEEVNLPG  
 KWPKMIGGIGGFIKVRQYEQIPIECGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTEEKIKA  
 LTAICDEMEKEGKITKIGPDNYPNTPIFAIRKKDSSKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDKDFR  
 KYTAF TIPSVNNETPGIRYQYNVLPQGWKGS PAIFQCSMTKILEPFRKQNPDI VIYQYMDL YVGS DLEIGQHR TKIEELREHLLKKGFTTP  
 DKKHQKEPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEALELELAENREI  
 LKEPVHGAYYDPSKELIAEIQKQGQDQWTYQIYQEPFNKLTGKYAKMRTAHTNDVKQLTEAVQKIAMESI VIWGKIPKFRLP IQKETWETW  
 WTDYQATWIPWEFVNTPLVCLWQLEKDP IAGVETFYVDGAANRETKIGKAGYVDRGRKKIVSLTDTTNQKTELQAIYIALQDSGSEV  
 NIVTDSQYALGIIQAQPKSESELVNIIEQLIKKERVYLSWVPAHKGIGGNEQVDKLVSNGIRKVLFLDGDKAQEEHEKHYHSNWRAMASD  
 ENLPIVAKELIVASCDCQCKLGEAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVI PAETGQETAYFILKLAGRWPVKVIHTDN  
 GSNFTSAAVKAACWAGIQQEFGI PYNPQSQGVVESMNKELKLIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIVDIIATDIQ  
 TRELQKQIIKIIONFRVYYRDSRDP IWKGP AKLLWKGE GAVVIQD NSDIKVVPRRKAKI IKDYGKQ MAGADCVAGRQDEDQ

## Fig. 124A

81. 2003 CON 10 CD pol.PEP  
 FFRENLAFOQRKARELPSEQTRANSPTSRELRVWGGDNTLSETGAERQGA VSLSPFQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEEMN  
 LPGKWKPMIGGIGGFIKVRQYDQILIEICGYKAI GTVLVGP TPVNI IGRNLLTQIGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISRIGPENYPNTPIFAIKKSDTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLYE  
 DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS PAIFQCSMTKILEPFRKQNPDI VIYQYMDL YVGS DLEIGQHR TKIEELRHLKKGFT  
 TTPDKKHQKEPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEALELELAEN  
 REILKEPVHGVYDPSKDLIAEIQKQGQDQWTYQIYQEPHNKLTGKYAKRRTAHTNDVKQLTEAVQKIAQESIVIWGKTPKFRLP IQKETW  
 ETWWTDYWQATWIPWEFVNTPLVCLWQLEKDP IAGVETFYVDGAANRETKIGKAGYVDRGRQKVISITD TTNQKTELQAINLALQDSG  
 SEVNI V TDSQYALGIIQAQPKSESELVNIIEQLIKKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDKAQEEHEKHYHNNWRAM  
 ASDFNLPVVAKEIVASCDCQCKLGEALHGQVDCSPGIWQDCTHLEGKIVLVAVHVASGYIEAEVI PAETGQETAYFLLKLAGRWPVKVHVH  
 TDNGSNFTSAAVKAACWAGIQQEFGI PYNPQSQGVVESMNKELKLIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIAT  
 DIQTKELQKQIIKIIONFRVYYRDSRDP IWKGP AKLLWKGE GAVVIQD NSDIKVVPRRKAKI IKDYGKQ MAGADCVASRQDEDQ



Fig. 123B

2003\_CON\_08\_BC\_pol.1.OPT  
TTCTTCCGGAGATCCTGGCCCTTCCCCAGGGGAGCCCGGAGTTCCCCCGGAGACCCGGCCAACTCCCCACCTCCCCGGAGCTGCAGGTGG  
CGGGACAACCCCTCCTCCGAGCCGGCAGCCAGCCCTGAACTTCCCCAGATCACTTGGCAGCGCCCTGGTGTCCATCAAGGTGG  
GGCCAGATCAAGGAGGCCCTGTGGACACCGGGCCGACACCGTGTGGAGAGGTGAACCTGCCCGCAAGTGAAGCCCAAGATGATCGGCGGC  
ATCGGGGGTTTCAACAAGTGGCCAGTACGAGCAGATCCCCATCGAGATCTGGCAAGAGCCATCGGCACCGTGAAGCTGAAGCCCGGCATGGACG  
CATCATCGGCCGAACATGTGACCCAGTGGCTGCACCTGAACTTCCCATCTCCCATCGAGACCGTCCCGTGAAGCTGAAGCTGAAGCCCGGCATGGC  
GCCCAAGGTGAAGCAGTGGCCCTGACCCGAGAGAAATCAAGGCCCTGACCCCATCTGCCAGCAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGC  
CCCGACAACCCCTACAACACCCCCATTTCCGCTCCGCAAGAACTCCTCCAAGTGGCGCAAGCTGTTGGACTTCCGCGAGCTGAACAAGCGCACCCCA  
GGACTTCTGGAGGTGCAGTGGGCTTCCCCACCCCGCCCTGAAGAAGAAGTCCGTGACCCGTGGACCTCCGCTGGACCTGGGCGACCCCTACTTCTCCGTGC  
CCCTGGACAAGGACTTCCGCAAGTACACCCGCTTACCATCCCCCTCCGTGAACAACGAGACCCCGGCATCCGCTACCCAGTACAACCTGCTGCCCCAGGGC  
TGAAGGGCTCCCCGCCATCTTCCAGTCTCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACCCCGACATCGTGTACTACAGTACATGGACGA  
CCTGTACGTGGCTCCGACCTGGAGTGGCTACGAGTGCACCCCGCAAGTGAACCGTCCAGCCCATCCAGTCCCGGAGAGGACTCCTGGACC  
AGCACAGAAGGAGCCCCCTTCTGTGATGGCTACGAGTGCACCCCGCAAGTGAACCGGAGATCTGAAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTAC  
GTGAACGACATCCAGAAGTGGTGGCAAGTGAACCTCCAGATCTACCCCGGCATCAAGTGGCCAGCTGTGCAAGCTGTGCAAGCTGTGCCGGCGCCAA  
GGCCCTGACCGACATCGTGGCTTGAACCGGAGGAGCCAGTGCACCCCGCAAGTGAACCGGAGATCTGAAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTAC  
CCTCCAAGGAGTGAATCGCCGAGATCCAAACGACGTGAAGCAGTGCACCGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACCCCAA  
GCCAAGATGGCACCCGCAAGGAGACTGGGAGACTGGTGGACCGACTTGGTGGACCGGAGTGCACCGAGTCCATCGTGTGATCTGGGCAAGATCCCCAA  
GTTCCGCTGCCCATCCAGAAGGAGACTGGGAGACTGGTGGACCGACTTGGTGGACCGGAGTGCACCGGAGTGCACCGGAGTGCACCGGAGTGCACCGGAGTGC  
TGGTGAAGCTGTGGTACCAGTGGAGAAGGCCCATCCCATCCCGGCTGGAGACTTCTACTGTGGACGGCCCGCCAAACCGCGAGACCAAGATCGGCAAGGCC  
GGCTACGTGACCGACCGGGCCGCAAGAGATCGTGTCCCTGACCGACACCAACCAAGAGACCGAGTGCAGGCCATCTACATCGCCCTGCAGGACTC  
CGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACCCCTGGGATCCAGGCCAGCCCGCAAGTCCGAGTCCGAGTGGTGAACCAAGTGCATC  
AGCAGTGAACAAGAGGCGGCTGTACTGTCTCTGGTGTCCCTGGTGCACCAAGGCAATCGGGGCAACGAGCAGGTGGACAAGTGGTGTCCAACCGGCATC  
CGCAAGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGCAGAGAAATACCACTCCAATGGCCGCCATGGCCCTCCGACTTCAACCTGCCCTCC  
CATCGTGGCCAAGGATCGTGGCCCTCTGGACCAAGTCCAGCTGAAGGCGGAGGCCATGACCGGCCAGGTGGACTGTCCCCGGCATCTGGCAGCTGG  
ACTGCACCCACCTGGAGGGCAAGATCATCTGTGGTGGCTGACCGTCCGCTACATCGAGCCGAGGTGATCCCCCGCGAGACCGGCCAGGAGACC  
GCCACTTCACTCTGAAGTGGCCGCGCTGGCCCTGAAGTGTCCACACCGACAACCGCTCCAACCTTCACTCCGCGCGGTGAAGCCCGCTGCTG  
GTGGCCGGCATCCAGCAGGAGTTCGGCATCCCCAACCCTCAGGCGTGGTGAAGTCCATGAACAAGGAGCTGAAGAAGTGTGATCGGCCAGG  
TGGCGACCAAGCCGAGCACTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGCATCGCGGCTACTCCGCCGGCGGAG  
CGCATCGTGGACATCATCGCCACCGACATCCAGACCCCGGAGTGCAGAAGCAGATCATCAAGATCCAGAATCTCCGCGTGTACTACCGGACTCCCGGGA  
CCCCATCTGGAAGGGCCCCGCAAGCTGTGTGGAAGGGCGAGGGCCCGTGGTGTATCCAGGACAACCTCCGACATCAAGGTGGTGGCCCGCGCAAGGCCA  
AGATCATCAAGGACTACGGCAAGCAGATGGCCCGCCGACTGGCTGGCCGGCCGACGAGGACTAA

Fig. 124B

2003\_con\_10\_cd\_pol.1.opt  
TTCTTCCGGAGAACCTGGCCCTCCAGCAGCGCAAGGCCCGGAGCTGCCCTCCGAGCAGACCCGGCCAAACTCCCCACCTCCCGGAGCTGGCGGTGTG  
GGGGGGGCAACAACCCCTGTCCGAGACCGGGCCGAGCCGAGGGCCCGTGTCCCTGTCTTCCCCCAGATCACCCCTGTGGAGCGCCCTGGTGACCG  
TGAAGATCGGGCCAGCTGAAGGAGGCCCTGTGGACACCGGGCCGACGACACCCGTGTGGAGGAGATGAACCTGCCCGCAAGTGAAGCCCAAGATG  
ATCGGGGGCATCGGGGGTTCATCAAGGTGGCCAGTACGACCCAGATCTGTATCGAGATCTGCGGGTACAAGGCCATCGGCACCCGTGTGGTGGCCCCAC  
CCCCGTGAACATCATCGGGCCCAACTGTGACCCAGATCGGCTGACCCCTGAACCTCCCAATCTCCCCATCGAGACCCGTGCCCGTGAAGCTGAAGCCCG  
GCATGGACGGCCCAAGTGAAGTGGCCCTGACCCGAGGAGATCAAGGCCCTGACCCGAGATCTGCACCCGAGATGGAGAAGGAGGGCAAGATCTCC  
CGCATCGSCCCCGAGAACCCCTACAACACCCCATCTTGCCATCAAGAAAGAGACTCCACCAAGTGGCGCAAGTGGTGAACCTCCGCGAGCTGAACAA  
GGCACCCAGGACTTCTGGAGGTGCAGTGGGCATCCCCACCCCGGGCTGAAGAAAGAAAGTCCGTGACCCGTGTGGACCTCCGCTACCAACCTGCTG  
TCTCCGTGCCCTGTACGAGGACTTCGCAAGTACACCCGCTTACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCAAGTGGTGAATCTACCA  
CCCCAGGCTGGAAGGCTCCCCCGCCATCTTCCAGTCCCTGACCAAGATCTGGAGCCCTTCGCAAGCAGAACCCCGAGATGGTGAATCTACCAAGT  
CATGGACGACTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCCGATCAAGATCGAGGAGCTGGCGGCCACCTGCTGAAGTGGGCTTCACCAACCC  
CCGACAAGAACCCAGAGGACCCCTTCTGTGGATGGCTACGAGTGCACCCGACAAAGTGGACCTAACCCCGGCATCAAGGTGGCCAGCTGCACGGCGTGT  
TCTTGGACCCGTGAACGACATCCAGAACTGTGGCAAGTGGCTGAACCTCCAGATCTACCCCGGCATCAAGGTGGCCAGCTCTGAAGGAGCCCGTGCACGGCGTGT  
CGGCGCAAGCCCTGACCCGACATCGTGGCCCTGACCCGAGGCGGAGCTGGAGTGGCCGAGAACCCGAGATCTGAAGGAGCCCGTGCACGGCGTGT  
ACTACGACCCCTCAAGGACTGTGCGCGAGATCCAGAACGAGCTGACCGAGCCGTGCAAGATCGCCGACCTCAAGGTGGCCAGCTGTGCAAGCTGTGG  
GGCAAGTACGCCAAGCGCCGACCCGACCAACGACCTGGAGACTTGGTGGACCCGACTACTGGCAGGCCACTTACTGACCGCCGCAACCCGAGACCCAAAGCTG  
GACCCCAAGTCCCGCTGCCATCCAGAGGAGACTGGGAGACTGGTGGACCCGACTACTGGCAGGCCACTTACTGACCGCCGCAACCCGAGACCCAAAGCTG  
CCCCCCCCGTGAAAGCTGTGTACAGCTGGAGAGGAGCCCATCGTGGGCGCCGAGACCTTACTGACCGCCGCAACCCGAGACCCGAGACCCAAAGCTG  
GGCAAGCCGGCTACGTGACCGACTCCAGTACCCCTGGCATCATCCAGGCCAGCCCAACAGAGACCCGAGCTGCAGGCCATCAACCTGGCCCT  
GCAGGACTCCGGCTCCGAGGTGAACATCGTGAACCGACTCCAGTACCCCTGGCATCATCCAGGCCAGCCCAACAGAGCTGCAGGCCATCAACCTGGCCCT  
AGATCATCGAGCAGCTGATCAAGAAAGGAGAGGTGTACCTGTCTTGGGTGCCCGCCCAAGGCCAGGAGGACGAGAAAGTACCAACAACCTGGCCGCACTCAA  
TCCGGCATCCGCAAGGTGTCTTGGACGGCATCGACAAGGCCAGGAGGACGAGAAAGTACCAACAACCTGGCCGCACTGGCCGCACTGGCCGCACTCAA  
CCTGCCCCCGTGGTGGCCAAAGGATCGTGGCTTCTTGGACAAAGTGGCCAGTGAAGGGCGAGGCCCTGCACGGCCAGGTGGACTGCTCCCCGGCATCT  
GGCAGCTGGACTGCACCCACTGGAGGGCAAGTGTATCTTGGTGGCCGTGCACCTGGCCGCACTGCAGGGCCGAGGTGATCCCCCGGAGACCCGGC  
CAGGAGACCCGCTACTTCTTCTGTGAAGCTGGCCCGCTGGCCGCTGAAAGGTGGTGCACACCGCAACGGCTCAAACCTCCGCGCCGCTGAAGGC  
CGCCTGTGTGGCCCGCATCAAGCAGGAGTTCGGCATCCCTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
TCGGCCAGGTGGCCGACCGCCGAGCCTGAAGACCCCGCTGCAGATGGCCGTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGGCGGCTACTCC  
GCCGGCAGCGCATCATCGACATCATCGCCACCGCATCCAGACCAAGGAGCTGCAGAACGATCATCAAGATCCAGAACTTCCGCGTGTACTACCCGGA  
CTCCCCGACCCCATCTGGAAGGGCCCGCCAAAGCTGTGTGGAAGGGCGAGGGCGCCGTGGTGTATCCAGGACAACTCCGACACTCAAGGTGGTGGCCCCGGC  
GCAAGGTGAAGATCATCAAGGACTACGGCAAGCAGATGGCCCGCCGACTGCGTGGCCCTCCCGCAGGACGAGGACCCAG

174/178

Fig. 125A

**82. 2003 CON 11 CPX pol. PEP**

FFRENLAFOQGEAREFPEQARANSPTRELRVRGGDPLPETGAEGEGAISFNFPQITLWQRPLVTIKVAGQLKEALLDTGADDTVLEED  
 LPRGWKPKMIGGIGGFIKVRQYEIIIEIEGKKAIGTVLGGPTPVNIIGRNMLTOIGCTLNFPISPIDTVPVKLKPMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE  
 SFRKYTAFTIP SINNETPGIRYQYNVLPQGWKSPALFQSSMTKILEPFRTONPEIIVIQYMDLLYVGSDELEIGQHREKVEELRKHLLKKGFE  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKCEWTVNDIQKLVGKLNWASQIYPGKVKQLCKLLRGTKALTDIVPLTAAEALELAEN  
 REILKEPVHGVYDPSKDLIAEVQKGLDQWTYQIYQEPFNKLTGKYAKRRRTAHTNDVRQLAEVVQKISMESIVIMGKIPKFRPLIQRETW  
 ETWTDYWQATWIPWEFEVNTPLVKLWYQLEKEPIIGAETFYVDGAANRETKLGKAGYVTDKGRQKVVVTLTETTNQKTELEAHLALQDSG  
 LEVNI VTD SQYALGIIQAQPKSESELVSQIIEQLIKKEKVVLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHERYHSNWRAM  
 ASDFNLPPIVAKAIVASCDCQKKEAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILLKLAGRWPVKVIH  
 TDNGSNFTSAAVKAACWAWANIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAHLKTAVQMAVFIHNEKRKGGIGGYSAGERIIVDIAT  
 DLQTKELQKQITIKIQNFRVYRDSRDP IWKGPALLWKGE GAVVIQDNSDIKVVP RRKAKIIRDYGKQ MAGDDCCVAGRQDED\$

Fig. 126A

**83. 2003 CON 12 BF pol. PEP**

FFRENLAFOQGEARKFPSEQARANS PASRELWVRRGDNPLSEAGAERRGTVP SLSFPQITLWQRPLVTIKVGGQLKEALLDTGADDTVLEED  
 NLPGWKPKMIGGIGGFIKVKQYDNILIEICGHKAI GTVLVGGPTPVNIIGRNLLTQLGCTLNFPIETVPVKLKPMDGPKVKQWPLTEE  
 KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD  
 KDFRKYTAFTIPSVNNETPGIRYQYNVLPQGWKSPALFQSSMTKILEPFRKQNPDIIVIQYMDLLYVGSDELEIGQHRTKIEELRQHLLRWG  
 FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPKEDSWTVNDIQKLVGKLNWASQIYPGKVKQLCRLLRGTKALTEVPLTKEAELELAE  
 NREILKEPVHGVYDPSKDLIAEQKQGGQWTYQIYQEPFNKLTGKYARMRGAHTNDVKQLTEAVQKITTESIVIMGKTPKFRPLILKET  
 WDTWWTEYWOATWIPWEFEVNTPLVKLWYQLETEPIAGAETFYVDGASNRETKKAGYVTDGRQKAVSLTETTNQKAEHLAIQLALQDS  
 GSEVNI VTD SQYALGIIQAQPKSESELVNIIEQLIKKEKVVLSWVPAHKGIGGNEQVDKLVSA GIRKILFLDGDIDKAQEEHEKYHNNWRA  
 MASDFNLPVVAKEIVASCDCQKKEAMHGQVDCSPGIWQDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILLKLAGRWPVKTI  
 HTDNGPNFSSAAVKAACWAWAGIQEFGIPYNPQSQGVVESMNKELKKIIRQVRDQAEHLKTAVQMAVFIHNEKRKGGIGGYSAGERIIDIIS  
 TDIQTRELQKQI IKIQNFRVYRDSRDPVWKGPALLWKGE GAVVIQDNSEIKVVP RRKAKIIRDYGKQ MAGDDCCVAGRQDED\$

Fig. 125B

2003\_con\_11\_cpx\_pol.OPT

TTCTTCCGCGAAGAACTGGCCCTCCAGCAGGGCGAGGCTTCCCCCGAGCAGGCCCGCGCCAACTCCCCACCTCCCGGAGCTGGCGGTGGG  
 CGGGCGGACTCCCCCTGGCCGAGACCGGGCCGAGGGCGGCCATCTCTTCAACTTCCCCCAGATCACCTGTGGCAGCGCCCTGGTGACCA  
 TCAAGGTGGCCGCGCAGCTGAAGGAGCCCTGTGGACACCGGGCCGACCGACACCGTGTGGAGGAGATCGACCTGCCGGCCGCTGAAAGCCCAAGATG  
 ATCGCGGCATCGCGGCTTTCATCAAGGTGCGCCAGTACGAGGAGATCATCATCGAGATCGAGGGCAAGAGCCATCGGCACCTGGTGGGCCCCAC  
 CCCCCTGAACATCATCGCGCCGCAACATGCTEACCCAGATCGGCTGCACCTGAACCTTCCCATCTCCCATCGACACCTGCCCCGTGAAGCCCG  
 GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCGACCCGAGGAGAAAGATCAAGGCCCTGACCCGAGATCGACCCGAGATGGAGAAAGGAGGCAAGATCTCC  
 AAGATCGGCCCGGAAACCCCTACAACACCCCGTGTTCGCCATCAAGAAAGGACTCCACCAAGTGGCGAAGCTGGTGGACTCCGCGAGCTGAACAA  
 GCGACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCACCCCGCCCTGAAGAAGAAAGTCCGTGACCCGTGGACCTGGACCTGGCGACCGCTACT  
 TCTCCGTGCCCTGGACGAGTCTTCCGCAAGTACACCGCTTCCATCCCTCCATCAACAAACGAGACCCCGGCATCCGCTACAGTACACAGCTGTG  
 CCCAGGGCTGGAAGGCTCCCCCGCCATCTTCCAGTCTTCCATGACCAAGATCTTGGAGCCCTTCCGACCCAGAACCCCGAGATCGTGTATCTACCAGTA  
 CATGGACGACTGTACGTGGGCTCCGACTGGAGATCGGCCAGCACCCGAGAAAGTGGAGAGCTGGCAAGCACCTGTGAAGTGGGCTTCACCCACCC  
 CCGACAAGAACCCAGAAAGGAGCCCTTCTGTGGATGGCTACGAGTGCACCCGACAAAGTGGACCCGTGCAGCCCATCCAGCTGCCCGACAAGGAG  
 TGCTGACCCGTGAACGACATCCAGAAAGTGGTGGCAAGCTGAACTGGGCCCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGTGCG  
 CGCACCAAGGCCCTGACCGACATCGTCCCCCTGACCCCGAGCCGAGCTGGAGCTGGCCGAGAACCCGAGATCTTGAAGGAGCCCGTGCACGGCGTGT  
 ACTACGACCCCTCAAGGACTGTATCGCCGAGTGCAGAGCAGGGCTGGACCACTGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGCCGACCCGACCAACGACGAGTGGCCGAGGTGGTGCAGAAAGATCTCCATGGAGTCCATCGTGTGATCTGGGGCAA  
 GATCCCAAGTTCGGCTGCCCCATCCAGCGGAGACTGGGAGACTGGTGGACCCGACTACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCTGGTGAAGCTGTGTACAGCTGGAGAAAGGCCCATCATCGCGCCCGAGACCTTCTACGTGGACGGCCCGCCAAACCCGAGACCAAGCTG  
 GGCAAGCCGGCTACGTGACCCGACAAAGGCCCGCAGAAAGTGTGACCTGACCCGAGACCAACCAAGAGACCCGAGCTGGAGGCCATCCACCTGGCCCT  
 GCAGGACTCCGGCTGGAGGTGAACATCGTGAACCGACTCCAGTGGCCCGCCACAAAGGCCATCGCGGCAACGAGCAGGTGGACAAGTGGTGTCC  
 AGATCATCGAGCAGTGCATCAAGAAGGAGAGGTGTACTGTCTGGTGGCCCGCAGAGGCTACCACTCCAACTGGCGGCCATGGCCTCCGACTTCAA  
 TCCGGCATCCGCAAGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGCAGCGCTACCACTCCAACTGGCGGCCATGGCCTCCGACTTCAA  
 CCTGCCCCCATCGTGGCCAAAGGAGATGTTGGCTCTTGGCAAGTGCACAGTGAAGGCCGAGGCCATGCACGGCCAGGTGGACTGTCCCCCGGCATCT  
 GGAGCTGGACTGCACCCACTGGAGGGCAAGATCATCTGGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACCGGC  
 CAGGAGACCGCTACTTTCATCTGAAGCTGGCCGGCCGCTGGCCCGTGAAGTGTCCACACCGCAACCGCTCCAACTCCGCTCCGCGCCGTGAAGGC  
 CGCCTGCTGGTGGCCAAACATCCAGCAGGAGTTCGGCATCCCCCTACAACCCCGAGTCCAGGGCCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCGAGCAGGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTTCATCCACAACCTCAAGCGCAAGGGCCGCATCGGGCGGTACTCC  
 GCGCGGAGCGCATCGTGGACATCGCCACCGACTGCAGACCAAGGAGCTGCAGAAAGCAGATCAACAAGTCCAGAACTTCCCGCTGTACTACCCGGA  
 CTCCCGGACCCCATCTGGAAAGGCCCGCCAAAGCTGTGTGAAAGGGCGAGGCCCGCTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGCCCGGC  
 GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCCGCAGCGACTGCGTGGCCCGCCAGGACGAGGACTAA

175/178

Fig. 126B

2003\_CON\_12\_BF\_pol.OPT

TTCTCCGGAGAACCTGGCCCTCCAGCAGGGCGAGGCCCGCAAGTTCCCTCCGAGCAGGCCCGCCCAACTCCCCCGCCTCCCGGAGCTGTGGGTGCG  
CCGGGGGACAAACCCCTGTCCGAGCCCGCGAGCCCGCCACCGTGCCTCCCTGTCTCCCCAGATCACCTGTGGCAGGCCCCCTGGTGA  
CCATCAAGGTGGCGGCCAGCTGAAGGAGCCCTGTCTGGACACCGGGCCGACACCGTGTGGAGACATCAACTGCCCGCAAGTGGAAAGCCCAAG  
ATGATCGCGGCATCGCGGCTTCAITCAAGGTGAAGCAGTACGACAACTCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGTGGTGGGCC  
CACCCCGTGAACATATCGCCCGAACCTGTGACCCAGTGGCTGCACCTGAATTCCTCCATTCCTCCCATCGAGCCGTGCCGTGAAGCTGAAGC  
CCGGATGGACGGCCCCAAGTGAAGCAGTGGCCCCGTGACCCGAGGAGAATCAAGCCCTGACCGGATCTGCACCGAGATGGAGAAGSAGGGCAAGATC  
TCCAAAGATCGGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCTCCGAGCTGAA  
CAAGCGCACCCAGGACTTCTGGAGGTGAGTGGCATCCCCACCCCGCCGCTGAAGAAGAAGTCCGTGACCGTGGCTGGACCGTGGCGGACGCCCT  
ACTTCTCCGTGCCCTGGACAAGGACTTCCGCAAGTACACCGCTTCACCATCCCTCCGTGAACAACGAGACCCCGGCATCCGCTACAGTACAACGTTG  
CTGCCCCAGGGTGGAAAGGCTCCCCCGCATCTCCAGTCTCCATGACCAAGATCTTGGAGCCCTTCCGCAAGCAGAACCCCGACATCGTGTATACCA  
GTACATGGACGACCTGTACTGTGGCTCCGACTGGAGATCGGCCAGCACCAAGTCCGAGGAGCTGGCCAGCACCCTGTCCGTGGGGCTTACCA  
CCCCGACAAGAAGCACCAAGAAGGCCCCCTTCTGTGGATGGCTACGAGCTGCACCCGACAAAGTGGACCGTGCAGCCCATCGTGTGCCCGGAGAAG  
GACTCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAATGGCTCCAGATCTACCCCGCATCAAGGTGAAGCAGCTGTGCCCGCTGT  
GGCGGCACCAAGGCCCTGACCGAGTGTATCCCCGTGACCAAGAGGCCGAGCTGGAGTGGCCGAGAACCCGCGAGATCTTGAAGGAGCCCGTGCACGGCG  
TGTAACGACCCCTCCAAGGACTGTATCGCCGAGATCCAGAAGCAGGGCCAGGCTGAGCTTACAGATCTACAGGAGCCCTTCAAGAACCTGAAG  
ACCGCAAGTACGCCCGCATCGCGGCCACACCAACGACTGAAGCAGTGAAGCAGTGAAGGCTGCAGAGATCACCAAGTCCATCGTGTATCTGGGG  
CAAGACCCCAAGTTCGCCCTGCCATCTTGAAGGAGCCTGGACACCTGTGGACCGACTGGCAGGCCACCTGGATCCCCGAGTGGAGTTCGTGA  
ACACCCCGCTGGTGAAGCTGTGTACCAGCTGGAGACCGCCCATCGCCCGCCGAGACCTTACGTGGACGGCCCTCAACCGCGAGACCAAG  
AAGGCAAGGCCGCTACGTGACCGCCCGCCAGAGCCCGTGTCCCTGACCGAGACCAACCAAGAGGCCGAGCTGCACGCCATCCAGCTGGC  
CCTGCAAGACTCCGGCTCCGAGGTGAACATCGTGACCCGACTCCAGTACGCCCTGGGCAATCATCCAGGCCAGCCCGACAAAGTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGCAGTGTCAAGAAGGAGAGTGTACTGTCTGGGTGCCCGCCACAAGGGCATCGCGGCAACGAGCAGGTGGACAAGCTGGTG  
TCCGCGGCATCCGCAAGATCTGTCTGGACGGCATCGACAAGGCCAGGAGGACGAGAAAGTACCAACAACCTGGCGGCCATGGCCCTCCGACTT  
CAACTGCCCCCGTGGTGGCCAAAGGATCTGTGGCTCTCGACAAAGTCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
TCTGGCAGCTGGACTGCACCCCAAGATCATCTGTGGCCGTGCACGTGGCTCCGGCTACCTGGAGGCCGAGGTGATCCCCGGCCGAGACC  
GGCCAGGAGACCGCTACTTCACTGAAAGTGGCCCGCTGGCCCGTGAAGACCAATCCACACCGACAAAGGCCCAACTTCTCTCCGCGCCGCTGAA  
GGCCGCTGTGTGGCCGCGCATCCAGCAGGATTCGGCATCCCTACAACCCCGAGTCCCAAGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
TCATCCGCCAGGTGGCCAGCCAGCCGACCTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCCAACTTCAAGCGCAAGGGCGGCATCGGCCGGCTAC  
TCCGCCGGGAGGCGCATCATCGACATCATCTCCACCGCATCCAGACCCCGGAGCTGCAGAGCAGATCATCAAGATCCAGAACTTCCGCGTGTACTACCG  
CGACTCCCGGACCCCGTGTGGAAGGCCCCCGCAAGCTGTGTGGAAGGCCGAGGGCCCGTGGTGTATCCAGGACAACTCCGAGATCAAGGTGGTGGCCCC  
GCCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGGACGACTGCGTGGCCGGCCCGCAGGACGAGGACTAA

177/178

## Fig. 127A

84. 2003 CON 14 BG pol.PEP  
 FFRENLAFOQGEAREFEPEQARANSPTRRRELWVRGRDSPPEARAEGKGDIPISLPQITLWQRPLVTVRIGGQLIEALLDTGADDDTVLEDIN  
 LPGKWKPKMIGGIGGFVKVRQYDQILLIEICGKKAIGTVLVGPTPINIIGRNMLTQIGCTLNFEISPIETVPVKLKPMDGPKVKQWPLTEEK  
 IKALTDICTEMEREGKISKIGPENPYNTPIFAIKKSDTKWRKLVDFRELNKRTODEFWEVQLGIPHPISGLKPKKSVTVL.DVGDAYFVSVPLDE  
 SFRKYTAFITPSTNNETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFRKINPEIYIYQYMDL.VVGSDEIGQHRAKIEELRKHLLSWGE  
 TTPDKKHQKEPPFLMMGYELHPDKWTVQPIQLPDKESWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTAAEALELAEN  
 REILKEPVHGVYEPKELIAEVQKGLDQWTYQIYQEPYKNLKTGKYAKRGSHTNDVKQLTEVVQKIATESI.VIWKTPKFKLPPIRKETW  
 EVWTEYWQATWIPDWEFVNTPLVKLWYRLETEPIAGAETYVVDGAANRET.KLGKAGYVTDKQKQKIIITLTTETNQKAELOAIHIALQDSDG  
 SEVNI.VTDSQYALGIIQAQDRSESEVVNQIIEQLIKKEKVVLSWVPAHKGIGNEQVDKLVSSGIRKVFLELDGIDKAQEEHEKHYHSNWRAM  
 ASDENLPPVVAKEIVASCDCQQLKGEAMHGQVDCSPGIWQLDCTHLEKIIILVAVHVASGYIEAEVI.PAETGQETA.YFILLKLAGRWPVKIIH  
 TDNGSNFT.SAAVKAACWWANITQEFFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTA.VQMAVFIHNFRRKGGIGGYSAGERIIDIIAS  
 DIQTKELQKQITKIQNFRVYFRDSRDP.IWKGPAKLLWKGEAVVIQDNNEIKVVPRRKAKIIRDYKQMA.GDDDCVAGRQDED\$

Fig. 127B

2003\_con\_14\_BG pol.1.OPT

TTCTCCGCGAGAACCTGGCCCTCCAGCAGGGGAGGCCCGGAGTTCTCCCGAGCAGGGCCCGCCAACTCCCCACCCCGCGAGCTGTGGGTGGC  
CCGCGGCGACTCCCGCTGCCCGAGGCCCGGCGGCAAGGGCGACATCCCGCTGTCCCTGCCCCAGATCACCTGTGGCAGCGCCCGCTGGTGACCCG  
TGCGCATCGCGGGCCAGCTGATCGAGGCCCTGCTGGACACCGCGCCGACACCGTGTGGAGGACATCAACTGCCCGGCAAGTGGAAAGCCCAAGATG  
ATCGCGGCATCGCGGGCTTCATCAAGGTGCGCCAGTACGACCAAGTCTGATCGAGATCTGGGCAAGAAAGGCCATCGGCACCGTGTGGTGGCCCGCAC  
CCCCATCAACATCATCGCGCCGCAACATGTCACCCAGATCGGCTGCACCTGAATTCCTCCCATCGAGACCGTCCCGTGAAGCTGAAGCCCG  
GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCGTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGCACCGAGATGGAGCGGCAAGATCTCC  
AAGATCGGGCCCCGAGAACCCCTACAACACCCCTCTTCGCCATCAAGAAGAAGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCGCGGAGCTGAACAA  
GGCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCCGCCTGAGGAAAGAAAGTCCGCTGGACCGTGGCGGACGCGCTACT  
TCTCCGTGCCCTGGACGAGTCTTCCGCAAGTACACCCGCTTACCATCCCTCCACCAACAACGAGACCCCGGCTCCGCTACCAATCAACCGTCTGT  
CCCCAGGGTGGAAAGGCTCCCGCCATCTCCAGTCTCCATGACCAAGATCTTGGAGCCCTCCCGCATCAAGAACCCCGAGATCGTGTATCTACCCAGTA  
CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCACCGCCCAAGATCGAGGAGTGCAGGACCTGTGTCTGGGGCTTCACCCACC  
CCGACAAAGAAGCACCAAGAGGCCCTTCTGTGGATGGGTACGAGCTGCACCCCGCAAGTGGACCGTGCAGCCCATCCAGCTGCCGACAAAGGAG  
TCTTGGACCGTGAACGATCCAGAAGCTGGTGGCAAGTGAATGGGCTCCAGATCTACCCCGCATCAAGGTGAAGCAGCTGTGCAAGCTGTGCG  
CGCGCCAAAGGCCCTGACCGACATCGTGCCTTGACCGCCGAGGCTGGAGTGGCCGAGAACCCGAGATCTTGAAGGAGCCCGTGCACGCGGTGT  
ACTACGAGCCCTCCAAAGGAGTGTATCGCCGAGGTGCAGAAAGCAGGCCCTGGACCACTACAGATCTACAGGAGCCCTACAAGAACCTGAAGACC  
GGCAAAGTACGCCAAAGCGGGCTCCGCCCCACCAACGACGTGAAGCAGCTGACCGAGGTGGTGCAGAAAGTCCGACCCGAGTCCATCGTGTCTGGGGCAA  
GACCCCAAGTTCAGCTGCCCATCCGCAAGGAGACCTGGAGGTGTGGTGGACCGGACTGGACGCCACTCCCGACTGGGAGTTCGTGAACA  
CCCCCCCCCTGGTGAAGCTGTGTAACCGCTGGAGACCGGCTCATCGCCCGCCGAGACCTACTACGTGGACGGCCCGCCCAACCCGAGACCAAGCTG  
GGCAAGGCCGCTACGTGACCGACCAAGGGCAAGCAGAAAGTCAACCCCTGACCGAGACCAACCAAGAGGCCGAGCTGCAGGCCATCCACATCGCCCT  
GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCCGGCATCATCCAGGCCCGAGCCCGTCCGAGTCCGAGGTGGTGAACC  
AGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACTGTCTGGTGCCTGGCCCAAGGGCATGGCGGCAACGAGCAGGTGGACAAAGTGGTGTCC  
TCCGGCATCCGCAAGGTGTGTTCTGGACGGCATCGACAAGGCCCCAGGAGGAGCACGAGAAGTACCACTCCAATGGCGGCCATGGCCCTCCGACTCAA  
CCTGCCCCCGTGGTGGCCAAAGGAGATCGTGGCTTCTCGGACAAAGTCCAGTCCAGGCGGAGCCATGCACGGCCAGGTGGACTGTCCCCCGGCATCT  
GGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCGCGGAGACCCGGC  
CAGGAGACCGCCCTACTTCACTCTGAAGCTGGCCCGCCGTGGCCGTGAAGATCATCCACACCGACAAAGGCTCAACTTCACTCCGCGCCGCTGAAGGC  
CGCTGTGGTGGCCCAACATCACCCAGGAGTTCGGCATCCCTTACACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
TCGGCCAGGTGGCGACAGGCCGAGCACCTGAAGACCCCGCTGCAGATGGCCGTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGCGGGCTACTCC  
GCCGCGAGCGCATCATCGACATCATCGCTCCGACATCCAGACCAAGGAGCTGCAGAGCAGATCACCAAGATCCAGAACTTCCCGCTGTACTTCCCGGA  
CTCCCCGCAACCCCATCTGGAAGGGCCCCCAAGCTGTGTGGAAGGGCGAGGCCCGCTGGTGCATCCAGGACAAACAGGATCAAGGTGGTGGCCCGGCC  
GCAAGGCCAAGATCATCCCGGACTACGGCAAGCAGATGGCCGGCCGCGGACGACTGCGTGGCCGGCCCGCAGGACGAGGACTAA