

 (19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2012-0086379 (43) 공개일자 2012년08월02일
(51) 국제특허분류(Int. Cl.) C07K 14/155 (2006.01) C12N 7/00 (2006.01) A61K 39/21 (2006.01) (21) 출원번호 10-2012-7018893(분할) (22) 출원일자(국제) 2004년09월17일 심사청구일자 2012년07월18일 (62) 원출원 특허 10-2006-7007686 원출원일자(국제) 2004년09월17일 심사청구일자 2009년09월09일 (85) 번역문제출일자 2012년07월18일 (86) 국제출원번호 PCT/US2004/030397 (87) 국제공개번호 WO 2005/028625 국제공개일자 2005년03월31일 (30) 우선권주장 60/503,460 2003년09월17일 미국(US) 60/604,722 2004년08월27일 미국(US)	(71) 출원인 듀크 유니버시티 미합 노쓰 캐롤라이나 27708-0083 더럼 박스 90083 리서치 드라이브 노쓰 빌딩 230 더 유니버시티 오브 앨라배마 앳 버밍햄 리서치 파운데이션 미합중국, 앨라배마주 35294-0111, 버밍햄, 사우스 #1120지 에이비, 701 20티에이치 스트리트 더 리젠츠 오브 더 유니버시티 오브 캘리포니아 미합중국, 94607-5200 캘리포니아, 오클랜드, 프랭클린 스트리트 1111, 5th 플로어 (72) 발명자 코버, 베테, 티. 미합중국, 뉴멕시코주 87544, 로스 알라모스, 빅락 루프 1290 한, 베아트리카, 에이치. 미합중국, 앨라배마주 35223, 버밍햄, 락힐 로드 3571 (뒷면에 계속) (74) 대리인 특허법인세신

전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **컨센서스 / 원형 면역원**

(57) 요약

본 발명은 면역원에 관한 것으로서, 특히 HIV 일차 아이솔레이트의 스펙트럼을 중화시키는 항체를 유도하는 면역원, 또는 T세포 면역반응을 유도하는 면역원에 관한 것이다. 본 발명은 상기한 면역원을 이용하여 항-HIV 항체를 유도하는 방법 및/또는 T세포 면역반응을 유도하는 방법에 관한 것이다. 본 발명은 또한 본원 면역원을 인코딩하는 핵산 서열에 관한 것이다.

대표도

MRVMGIQRNCQHLWRWGMTMILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNWVAT
HACVPTDPNPQEIVLENVTENFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVRNVSSNG
V1
TETDNEEIKNCSEFNITTELROKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYRLINCNTSAITQACP
V2
KVSFEPIPIHYCAPAGFAILKCNDDKKFNGTGPKCNVSTVQCTHGIKPVVSTQLLNGSLAEEEEIIRSEN
V3
ITNNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGQAFYATGEIIGDIRQAHONISRTKWNKTLOQVAK
V4
KLREHFNNKTIIFKPSSGGDLEITTHSFNCGGEFFYNTSGLFNSTWMFNNGTYMNGTKDNSETITLPCR
V5
IKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNSNKNKTETFRPGGGMDRDNWRSELYKYK
VVKIEPLGVAPTAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLR
AIEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGC SGKLICTTNVPWNSSWSNKSODEIWDNMT
WMEWEREISNYTDIIYRLIEESQNOQEKNEQELLALDKWASLWNWFDITNLWYIKIFIMIVGGLIGLRI
VFAVLSIVNRVRQGYSPLSFQTLIPNPRGPD RPEGIEEEGGEQGRDRSIRLVNGFLALAWDDLRLCLFS
YHRLRDFILIAARTVELLGRSLRGLOKGWEALKYLGNLLQYWGOELKNSAISLLDTTAIAVAEGTDRVI
EIVQRACRAILNIPRRIRQGLERALL

(72) 발명자

쇼우, 조지, 엠.

미합중국, 앨라배마주 35223, 버밍햄, 락힐 로드 3571

코테, 데니스.

미합중국, 앨라배마주 35222, 버밍햄, 39티에이치 스트리트 734

리, 잉 잉

미합중국, 앨라배마주 35266, 후버, 마운틴 코브 2529

데커, 줄리

미합중국, 앨라배마주 35007, 알라바스터, 콜로니얼 드라이브 1116

헤이네스, 바톤, 에프.

미합중국 노스캐롤라이나주 27707, 더람, 웬트워쓰 스트리트 3923

가오, 펑

미합중국 노스캐롤라이나주 27712, 더람, 컨트리 클럽 드라이브702

리아오, 후아-진

미합중국 노스캐롤라이나주 27514, 샤펀 힐, 페블 스프링 로드200

특허청구의 범위

청구항 1

서열번호 1(도 1a), 서열번호 5(도 6c), 서열번호 6(도 6d), 서열번호 28(도 20c), 서열번호 29(도 20d), 서열번호 30(도 26a), 서열번호 33(도 28b), 서열번호 39(도 30b), 서열번호 42(도 31b), 서열번호 45(도 32b), 서열번호 48(도 33b), 서열번호 51(도 34b), 서열번호 54(도 35b), 서열번호 57(도 36b), 서열번호 60(도 37b), 또는 서열번호 63(도 38b)에 개시된 아미노산 서열을 포함하는 분리된 단백질.

청구항 2

제 1항의 단백질을 인코딩하는 뉴클레오티드 서열을 포함하는 핵산.

청구항 3

제 2항에 있어서, 상기 단백질은 서열번호 30(도 26a), 서열번호 33(도 28b), 서열번호 39(도 30b), 서열번호 42(도 31b), 서열번호 45(도 32b), 서열번호 48(도 33b), 서열번호 51(도 34b), 서열번호 54(도 35b), 서열번호 57(도 36b), 서열번호 60(도 37b), 또는 서열번호 63(도 38b)에 개시된 아미노산 서열을 포함하는 것인 핵산.

청구항 4

제 3항에 있어서, 상기 핵산은 서열번호 31(도 26b), 서열번호 34(도 28c), 서열번호 40(도 30c), 서열번호 43(도 31c), 서열번호 46(도 32c), 서열번호 49(도 33c), 서열번호 52(도 34c), 서열번호 55(도 35c), 서열번호 58(도 36c), 서열번호 61(도 37c), 또는 서열번호 64(도 38c)에 개시된 뉴클레오티드 서열을 포함하는 것인 핵산.

청구항 5

하기를 인코딩하는 뉴클레오티드 서열을 포함하는 핵산으로, 상기 뉴클레오티드 서열은 인간 세포에서 발현되기 위해 최적화된 코돈을 포함하는 것인 핵산:

CON6 HIV gp160 단백질,

서브타입 C 원형(ancestral) HIV 외피(envelope) 단백질,

서브타입 C 컨센서스(consensus) HIV 외피 단백질,

서브타입 C 컨센서스 HIV gag 단백질,

서브타입 C 컨센서스 HIV nef 단백질,

그룹 M 컨센서스 HIV 외피 단백질,

서브타입 A 컨센서스 HIV 외피 단백질,

그룹 M 컨센서스 HIV gag 단백질,

그룹 M 컨센서스 HIV pol 단백질,

그룹 M 컨센서스 HIV nef 단백질,

서브타입 C 컨센서스 HIV pol 단백질,

서브타입 B 컨센서스 HIV gag 단백질, 또는

서브타입 B 컨센서스 HIV 외피 단백질.

청구항 6

제 5항에 있어서, 상기 핵산은 서열번호 2(도 1d), 서열번호 3(도 6a), 서열번호 4(도 6b), 서열번호 11(도 13e), 서열번호 12(도 13f), 서열번호 14(도 14b), 서열번호 16(도 18b), 서열번호 17(도 19a), 서열번호 18

(도 19b), 서열번호 19(도 19c), 서열번호 20(도 19d), 서열번호 25(도 20a), 또는 서열번호 26(도 20b)에 개시된 뉴클레오타이드 서열을 포함하는 것인 핵산.

청구항 7

서열번호 65 (도 39a), 서열번호 66 (도 40a), 서열번호 68 (도 41a), 서열번호 69 (도 42a), 서열번호 73 (도 43a), 서열번호 74 (도 44a), 서열번호 77 (도 45a), 서열번호 78 (도 46a), 서열번호 81 (도 47a), 서열번호 82 (도 48a), 서열번호 85 (서열번호 49a), 서열번호 86 (도 50a), 서열번호 89 (도 51a), 서열번호 90 (도 52a), 서열번호 93 (도 53a), 서열번호 94 (도 54a), 서열번호 97 (도 55a), 서열번호 98 (도 56a), 서열번호 101 (도 57a), 서열번호 102 (도 58a), 서열번호 105 (도 59a), 서열번호 106 (도 60a), 서열번호 109 (도 61a), 서열번호 110 (도 62a), 서열번호 113 (도 63a), 서열번호 115 (도 64a), 서열번호 117 (도 65a), 서열번호 119 (도 65c), 서열번호 121 (도 66a), 서열번호 123 (도 67a), 서열번호 125 (도 67c), 서열번호 127 (도 68a), 서열번호 129 (도 68c), 서열번호 131 (도 69a), 서열번호 133 (도 70a), 서열번호 135 (도 71a), 서열번호 137 (도 72a), 서열번호 139 (도 73a), 서열번호 141 (도 74a), 서열번호 143 (도 75a), 서열번호 145 (도 76a), 서열번호 147 (도 77a), 서열번호 150 (도 78a), 서열번호 152 (도 79a), 서열번호 154 (도 80a), 서열번호 156 (도 81a), 서열번호 158 (도 82a), 서열번호 160 (도 83a), 서열번호 162 (도 84a), 서열번호 164 (도 85a), 서열번호 166 (도 86a), 서열번호 168 (도 87a), 서열번호 170 (도 88a), 서열번호 172 (도 88c), 서열번호 174 (도 89a), 서열번호 176 (도 90a), 서열번호 178 (도 90c), 서열번호 180 (도 91a), 서열번호 182 (도 92a), 서열번호 184 (도 92c), 서열번호 186 (도 93a), 서열번호 188 (도 94a), 서열번호 190 (도 95a), 서열번호 192 (도 96a), 서열번호 194 (도 97a), 서열번호 196 (도 98a), 서열번호 198 (도 99a), 서열번호 200 (도 100a), 서열번호 202 (도 101a), 서열번호 204 (도 102a), 서열번호 206 (도 103a), 서열번호 208 (도 104a), 서열번호 210 (도 105a), 서열번호 212 (도 106a), 서열번호 214 (도 107a), 서열번호 216 (도 108a), 서열번호 217 (도 109a), 서열번호 220 (도 109c), 서열번호 222 (도 110a), 서열번호 223 (도 111a), 서열번호 226 (도 111c), 서열번호 228 (도 112a), 서열번호 230 (도 112c), 서열번호 232 (도 113a), 서열번호 233 (도 114a), 서열번호 236 (도 115a), 서열번호 237 (도 116a), 서열번호 240 (도 117a), 서열번호 241 (도 118a), 서열번호 244 (도 119a), 서열번호 245 (도 120a), 서열번호 248 (도 121a), 서열번호 249 (도 122a), 서열번호 252 (도 123a), 서열번호 253 (도 124a), 서열번호 256 (도 125a), 서열번호 257 (도 126a), 및 서열번호 260 (도 127a) 중 어느 하나에 개시된 아미노산 서열의 CF, 또는 CFI 형태를 포함하는 분리된 단백질.

청구항 8

서열번호 72 (도 39b), 서열번호 67 (도 40b), 서열번호 70 (도 41b), 서열번호 71 (도 42b), 서열번호 75 (도 43b), 서열번호 76 (도 44b), 서열번호 79 (도 45b), 서열번호 80 (도 46b), 서열번호 83 (도 47b), 서열번호 84 (도 48b), 서열번호 87 (도 49b), 서열번호 88 (도 50b), 서열번호 91 (도 51b), 서열번호 92 (도 52b), 서열번호 95 (도 53b), 서열번호 96 (도 54b), 서열번호 99 (도 55b), 서열번호 100 (도 56b), 서열번호 103 (도 57b), 서열번호 104 (도 58b), 서열번호 107 (도 59b), 서열번호 108 (도 60b), 서열번호 111 (도 61b), 서열번호 112 (도 62b), 서열번호 114 (도 63b), 서열번호 116 (도 64b), 서열번호 118 (도 65b), 서열번호 122 (도 66b), 서열번호 124 (도 67b), 서열번호 128 (도 68b), 서열번호 132 (도 69b), 서열번호 134 (도 70b), 서열번호 136 (도 71b), 서열번호 138 (도 72b), 서열번호 140 (도 73b), 서열번호 142 (도 74b), 서열번호 144 (도 75b), 서열번호 146 (도 76b), 서열번호 148 (도 77b), 서열번호 151 (도 78b), 서열번호 153 (도 79b), 서열번호 155 (도 80b), 서열번호 157 (도 81b), 서열번호 159 (도 82b), 서열번호 161 (도 83b), 서열번호 163 (도 84b), 서열번호 120 (도 65d), 서열번호 126 (도 67d), 서열번호 130 (도 68d), 서열번호 165 (도 85b), 서열번호 167 (도 86b), 서열번호 169 (도 87b), 서열번호 171 (도 88b), 서열번호 175 (도 89b), 서열번호 177 (도 90b), 서열번호 181 (도 91b), 서열번호 183 (도 92b), 서열번호 187 (도 93b), 서열번호 189 (도 94b), 서열번호 191 (도 95b), 서열번호 193 (도 96b), 서열번호 195 (도 97b), 서열번호 197 (도 98b), 서열번호 199 (도 99b), 서열번호 201 (도 100b), 서열번호 203 (도 101b), 서열번호 205 (도 102b), 서열번호 207 (도 103b), 서열번호 209 (도 104b), 서열번호 211 (도 105b), 서열번호 213 (도 106b), 서열번호 173 (도 88d), 서열번호 179 (도 90d), 서열번호 185 (도 92d), 서열번호 215 (도 107b), 서열번호 218 (도 108b), 서열번호 219 (도 109b), 서열번호 224 (도 110b), 서열번호 225 (도 111b), 서열번호 229 (도 112b), 서열번호 234 (도 113b), 서열번호 235 (도 114b), 서열번호 238 (도 115b), 서열번호 239 (도 116b), 서열번호 242 (도 117b), 서열번호 243 (도 118b), 서열번호 246 (도 119b), 서열번호 247 (도 120b), 서열번호 250 (도 121b), 서열번호 251 (도 122b), 서열번호 254 (도 123b), 서열번호 255 (도 124b), 서열번호 258 (도

125b), 서열번호 259 (도 126b), 서열번호 261 (도 127b), 서열번호 221 (도 109d), 서열번호 227 (도 111d), 및 서열번호 231 (도 112d) 중 어느 하나에 개시된 뉴클레오타이드 서열을 포함하는 핵산.

청구항 9

제 2 항, 제 5항 및 제 8항 중 어느 한 항의 핵산을 포함하는 벡터.

청구항 10

제 1항, 제 2항, 제 5항, 제 7항 및 제 8항 중 어느 한 항의 적어도 하나 이상의 단백질 또는 핵산과 담체(carrier)를 포함하는 조성물.

청구항 11

제 1항, 제 2항, 제 5항, 제 7항 및 제 8항 중 어느 한 항의 적어도 하나 이상의 단백질 또는 핵산을, 면역반응을 유도할 정도의 양으로 인간을 제외한 포유동물에 투여하는 단계를 포함하는, 인간을 제외한 포유동물에서의 면역 반응 유도 방법.

청구항 12

서열번호 35(도 29a)의 단백질 또는 상기 단백질의 면역원 부분(immunogenic portion)을 인코딩하는 뉴클레오타이드 서열을 포함하는 핵산.

청구항 13

제 12항에 있어서, 상기 부분은 상기 서열번호 35(도 29a)의 단백질의 gp120 부분을 포함하는 것인 핵산.

청구항 14

제 12항에 있어서, 상기 부분은 상기 서열번호 35(도 29a)의 단백질의 gp41 부분을 포함하는 것인 핵산.

청구항 15

제 12항에 있어서, 상기 핵산은 서열번호 37(도 29c)에 기재된 뉴클레오타이드 서열, 또는 상기 서열번호 35(도 29a)의 단백질의 면역원 부분을 인코딩하는 이들의 분획인 것인 핵산.

청구항 16

서열번호 35(도 29a)의 단백질, 또는 gp140, gp140CF 또는 gp140CFI 형태를 인코딩하는 뉴클레오타이드 서열을 포함하는 핵산.

청구항 17

제 12항 또는 제 16항에 있어서, 상기 뉴클레오타이드 서열은 최적화된 코돈인 것인 핵산.

청구항 18

제 12항 또는 제 16항의 핵산을 포함하는 벡터.

청구항 19

제 12항 또는 제 16항의 핵산 및 담체(carrier)를 포함하는 조성물.

청구항 20

제 12항 또는 제 16항의 핵산을, 면역반응을 유도할 정도의 양으로 인간을 제외한 포유동물에 투여하는 단계를 포함하는, 인간을 제외한 포유동물에서의 면역 반응 유도 방법.

명세서

기술분야

[0001] 본 출원은 2003년 9월 17일에 출원된 이전 출원 제60/503,460호 및 2004년 8월 27일에 출원된 이전 출원 제60/604,722호를 우선권으로 청구하며, 이들의 전체 내용은 참조로서 본원에 삽입되어 있다.

[0002] 본 발명은 일반적으로 면역원에 관한 것이고, 구체적으로는 광범위한 스펙트럼의 HIV 제1 분리물(isolate)을 중화시키는 항체를 유도하기 위한 면역원 및/또는 T 세포 면역 반응을 유도하는 면역원에 관한 것이다. 또한, 본 발명은 이러한 면역원을 이용한, 항-HIV 항체의 유도 방법 및/또는 T 세포 면역 반응 유도 방법에 관한 것이다. 더불어, 본 발명은 본 발명의 면역원을 인코딩하는 핵산 서열에 관한 것이다.

배경 기술

[0003] HIV-1의 높은 레벨의 유전적 변이성은 AIDS 백신 개발의 주요 장애물로 여겨졌다. HIV 그룹 M, N 및 O 간의 유전적 차이는 gag 및 env 유전자 각각에서 30% 내지 50%로 광범위하였다 (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)). 그룹 M 내의 바이러스들은 추가로 9개의 유전적으로 차이 나는 서브타입(A-D, F-H, J 및 K)으로 분류된다 (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), 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 acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (2000)). HIV-1 서브타입 내의 env 유전자의 유전적 변이가 30% 만큼 높아짐에 따라, 모든 HIV-1 서브타입에 대한 교차(cross)-서브타입 T 및 B 세포 면역 반응을 계속적으로 유도하는 것이 어려워졌다. 또한, HIV-1은 다른 서브타입 내에서 종종 재조합함으로써, 순환 재조합 형태(circulating recombinant forms, CRFs)를 만들어낸다 (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 acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (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)). 20% 이상의 HIV-1 분리물은, 복합 서브타입이 공통적인 지형 지역에서 재조합체이며 (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)), 재조합 바이러스의 높은 유병률(prevalence rate)은 추가로 실험적 HIV-1 면역원을 디자인하는 것을 복잡하게 할 수 있다.

[0004] AIDS 백신 개발에 있어서의 이러한 과제를 극복하기 위해, 3개의 컴퓨터 모델(트리(tree)의 컨센서스, 원형 및 중심)이 센트럴화된 HIV-1 유전자를 생성하는데 사용되었다 (Gaschen et al, Science 296:2354-2360 (2002), Gao et al, Science 299:1517-1518, Nickle et al, Science 299:1515-1517 (2003), Novitsky et al, J. Virol. 76:5435-5451 (2002), Ellenberger et al, Virology 302:155-163 (2002), Korber et al, Science 288:1789-1796 (2000)). HIV 생물학은 성-형(star-like)의 계통발생(phylogenies)을 유발하며, 이 결과로서, 3 종류의 서열이 서로서로 2-5% 다르다 (Gao et al, Science 299:1517-1518 (2003)). 3개의 센트럴화된 유전자 전략의 어느 것이든 면역원 및 공지된(filed) 바이러스 균주 사이의 단백질 거리를 줄여줄 것이다. 컨센서스 서열은 얼라인먼트의 각 포지션에 있는 가장 공통적인 아미노산을 기초로 하여 인공 서열을 제작함으로써, 백신 균주 및 동시 순환(contemporary circulating) 바이러스 사이의 서열 비유사성 정도를 최소화시킨다 (Gaschen et al, Science 296:2354-2360 (2002)). 원형 서열들은 컨센서스 서열과 유사하지만 최대-가능성(maximum-likelihood) 계통발생학적 분석 방법을 이용하여 생성된다 (Gaschen et al, Science 296:2354-2360 (2002), Nickel et al, Science 299:1515-1517 (2003)). 이렇게 하는데 있어서, 이 방법은 분석된 최근의 야생형 서열의 이론적 원형 서열을 재현한다 (도 26). Nickle 등은 센트럴화된 HIV-1 서열, 트리의 중심(center of the tree, COT)-원형 서열과 유사하지만 아우트라이어(outlier)에 의해 영향을 적게 받음-을 형성하는 다른 방법을 제안하였다 (Science 299:1515-1517 (2003)).

[0005] 본 발명은 만약 센트럴화된 면역원이 동물에서 T 세포 및 B 세포 면역 반응을 모두 유도할 수 있는지 확인하기 위해, 디자인된 연구의 결과로부터 최소 일부분의 결과를 유발한다. 이 연구들은 인공 그룹 M 컨센서스 env 유

전자 (CON6)의 제조 및 DNA 플라스미드 및 가용성 gp120 및 gp140CF 단백질로서 CON6 외피를 발현하는 재조합 백신 바이러스의 컨스트럭션과 관련이 있다. 이 결과들은 CON6 Env 단백질이 생물학적으로 기능적이고, 야생형 HIV-1의 선형, 형태학적 및 글리칸-의존성 에피토프를 가지며, HIV 서브타입 B 및 C 모두의 T 세포 에피토프를 인식하는 사이토카인-생성 T 세포를 유도함을 설명해준다. 중요하게는, CON6 gp120 및 gp140CF 단백질이 서브타입 B 및 C HIV-1의 제1 분리물의 서브셋(subset)을 중화하는 항체를 유도한다.

[0006] 중심화된 HIV-1 유전자 어프로치 연구의 반복되는 특징은 HIV-1 서열의 빠르게 확장하는 진화 및 HIV 서열 데이터베이스 (즉, Los Alamos National Database)에 수집된 서열이 매년 신규한 서열로서 꾸준히 업데이트되고 있다는 사실로부터 유래된다. CON6 gp120 외피 유전자는 1999년도의 Los Alamos 국립 데이터베이스 서열로부터 유래되며, Con-S는 2000년도의 Los Alamos 국립 데이터베이스 서열로부터 유래된다. 추가로, CON6는 중국의 서브타입 C V1, V2, V4 및 V5 Env 서열을 가지는 반면, Con-S는 모든 그룹 M 컨센서스 Env 일치 및 변형 부위-최소-길이 변형 루프까지 짧아짐-를 가진다. 2003년도 그룹 M의 시리즈에 대해 코돈-최적화된 유전자 및 서브타입 컨센서스 서열이, HIV의 제1 분리물에 반응하는 광범위한 반응성 T 및 B 세포의 비교 및 이들 세포의 유도에 사용하기 위한 야생형 HIV-1 Env 유전자의 대응하는 시리즈를 가지는 것으로서 디자인되었다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 HIV의 제1 분리물의 광범위한 스펙트럼을 중화시키는 항체를 유도하기 위한 면역원 및/또는 T 세포 면역 반응을 유도하는 면역원, 그리고 이를 인코딩하는 핵산 서열에 관한 것이다. 또한, 본 발명은 항-HIV 항체의 유도 방법, 및/또는 이러한 면역원을 이용한 T 세포 면역 반응의 유도 방법에 관한 것이다.

[0008] 본 발명의 목적 및 이점은 하기의 기재에 의해 명확해질 것이다.

과제의 해결 수단

[0009] 본 발명은 광범위한 스펙트럼의 인간 면역결핍 바이러스 (human immunodeficiency virus, HIV) 제1 분리물을 중화시키는 항체를 유도하는 면역원 및/또는 T 세포 반응을 유도하는 면역원에 관한 것이다. 면역원은 1 이상의 컨센서스(consensus) 또는 원형(ancestral) 면역원 (예, Env, Gag, Nef 또는 Pol) 또는 이의 일부 또는 변이체를 포함한다. 또한, 본 발명은 컨센서스 또는 원형 면역원, 또는 이의 일부 또는 변이체를 인코딩하는 핵산 서열에 관한 것이다. 더불어 본 발명은 상기 면역원 및 이를 인코딩하는 서열 모두를 이용하는 방법에 관한 것이다.

발명의 효과

[0010] 본 발명이 특정 컨센서스 및 원형 면역원에 대한 참조문헌 (예, 그룹 M 컨센서스 Env에 대해)과 함께 자세하게 기재되어 있지만, 본원에 개시된 시도가 다양한 컨센서스 또는 원형 면역원 (예, 다른 HIV-1 그룹 (예, N 및 O)에 대한 외피)을 제조하는데 사용될 수 있는 것이 바람직할 것이다.

도면의 간단한 설명

[0011] 도 1a - 1d: 그룹 M 컨센서스 env 유전자 (CON6)의 제조 및 발현. CON6 gp160의 완전 아미노산 서열이 표시되어 있다.

(도 1a)(서열번호 1) 야생형 CRF08_BC (98CN006) env 유전자의 5 부위가 밑줄친 글자로 표시되어 있다. 다양한 부위가 서열 위에 중괄호로 표시되어 있다. 잠재적 N-링크된 글리코실화 위치가 굵은 글씨로 강조되어 있다. (도 1b) CON6 gp120 및 gp140CF의 컨스트럭트. CON6 gp120 및 gp140CF 플라스미드는 중지 코돈을 gp120 절단 위치 뒤 또는 막투과 도메인의 앞에 각각 도입시킴으로써 제조하였다. gp120/gp41 절단 위치 및 gp41의 융합 도메인을 gp140CF 단백질에서 절단시켰다. (도 1c) CON6 gp120 및 gp140CF의 발현. *갈란투스 니발리스 아가로스* (*galanthus Nivalis agarose*) 렉틴 컬럼을 이용하여 rVV-감염된 293T 세포의 세포배양 상등액으로부터 CON6 gp120 및 gp140CF를 정제하였다. gp120 및 gp140CF 모두를 10% SDS-폴리아크릴아미드 겔에서 분리하였으며 코마시 블루로 염색하였다. (도 1d)(서열번호 2) 고발현된 인간 유전자용 코돈 사용에 기초하여 최적화된 CON6 env 유전자.

도 2a - 2e. 가용성 CD4 (sCD4) 및 항-Env mAbs에 대한 CON6 gp120 gp140 CF의 결합. (도 2a - 2b) 표시된 mabs 및 sCD4 각각이 CM5 센서 칩 (BIAcore)에 공유결합으로 고정되어 있으며, CON6 gp120 (도 2a) 또는

gp140CF (도 2b) (각각 100 $\mu\text{g}/\text{ml}$ 및 300 $\mu\text{g}/\text{ml}$)이 각 표면 상에 주입되어 있다. 17b mab를 제외하고, 항-gp120 mabs와 반응한 gp120 및 gp140CF 단백질 모두를 테스트하였으며, 이들은 CON6 gp120 및 gp140CF 모두와 미약하게 결합하는 것으로 나타났다. CON6 gp120 및 gp140CF에 결합하는 17b mab의 유도를 확인하기 위해, CON6 gp120 (도 2c) 또는 gp140CF (도 2d) 단백질을 sCD4 또는 mabs A32 또는 T8로 고정된 각각의 유동(flow) 세포 상에 캡처시켰다 (400 - 580 RU). 각 표면을 안정화시킨 다음, mAb 17b를 주입한 뒤 고정된 각 유동 세포 상에 플로우시켰다. 표면 각각을 안정화시키고 난 뒤, mAb 17b를 주입하고 고정된 유동 세포 각각에 흘려보냈다. 커브의 오버레이에서는, CON6 Env 단백질에 대한 mab 17b의 결합력이 sCD4 및 mab A32 표면 모두에서 두드러지게 향상되나 T8 표면에서는 그렇지 않음을 나타내어 준다 (도 2c - 2d). 인간 mabs에 대한 CON6 gp120 및 gp140CF의 결합력을 ELISA에서 검증하기 위해, mabs 447, F39F, A32, IgG1b12 및 2F5의 20 $\mu\text{g}/\text{ml}$ 농축 용액을 CON6 gp120 및 gp140CF에 대해 적정하였다 (도 2e). Mabs 447 (V3), F39F (V3), A32 (gp120) 및 IgG1b12 (CD4 결합 위치) 각각은 CON6 gp120 및 140 웰에 모두 잘 결합하였으나, 2F5 (항-gp41 ELDKWS)(서열번호 321)는 gp140CF에만 결합하였다. gp120의 mab 447 및 F39F 결합력에 대한 엔드포인트 역가 농도는 각각 $<0.003 \mu\text{g}/\text{ml}$ 및 $0.006 \mu\text{g}/\text{ml}$ 이었으며; mab A32에 대해서는 $<0.125 \mu\text{g}/\text{ml}$ 이고; IgG1b12에 대해서는 $<0.002 \mu\text{g}/\text{ml}$ 이며; 2F5에 대해서는 $0.016 \mu\text{g}/\text{ml}$ 이었다.

도 3a 및 3b. CON6 외피의 감염도 및 공-수용체(coreceptor) 사용. (도 3a) CON6 및 대조군 env 플라스미드를 HIV-SG3 Δ env 구조체와 함께 인간 293T 세포에 동시형질감염시킴으로써 Env-유사비리온(pseudovirions)을 제조하였다. 동량의 유사비리온 각각 (5 ng p24)을 JC53-BL 세포를 감염시키는데 사용하였다. 감염된 세포를 β -gal 발현시키기 위해 염색하고 난 뒤, 유사 비리온 p24 마이크로그램 당 푸른색 세포의 갯수 (감염 단위, IU)를 계수함으로써, 감염도(IU/ μg p24)를 측정하였다. (도 3b) AMD3100 및/또는 TAK-799를 1시간 동안 (37°C) JC53BL 세포에 처리한 뒤, 동량의 각 Env-유사비리온 p24 (5 ng)로 감염시킴으로써, CON6 env 유전자의 공-수용체 사용을 확인하였다. 대조군 그룹 (저해제 없음)에서의 감염도를 100%로 정하였다. 저해(blocking) 효율은, 저해제가 없는 대조군 배양물의 저해 실험에 대해 비교한 저해 실험의 IU 백분율로서 나타내었다. 데이터는 평균 $\pm$ SD로 표시하였다.

도 4. 다중 서브타입 항혈청에 대한 다중 서브타입 Env 단백질의 웨스턴 블랏 분석. 동량의 Env 단백질 (100 ng)을 10% SDS-폴리아크릴아미드 젤 상에서 분리하였다. 전기영동을 한 후에, 단백질을 Hybond ECL 니트로셀룰로오즈 막에 트랜스퍼하고 HIV-1 감염된 환자의 혈청 (1:1,000) 또는 CON6 gp120 DNA 프라이머로 면역화되고 rVV로 추가접종시킨 (1:1,000) 기니아 피그의 혈청으로 반응시켰다. 단백질-결합된 항체를 형광-표지된 제2 항체로 탐침시킨 뒤, 이미지를 스캔하고 적외선 이미지 기기(infrared imager Odyssey)(Li-Cor, Lincoln, NE)에 기록하였다. 서브 타입은 Env 단백질 및 혈청 IDs의 뒤에 단문자(single-letter)로 표시하였다. 4 내지 6개의 혈청을 각 서브 타입에 대해 테스트하였으며, 반응 패턴은 동일 서브 타입의 모든 혈청 사이에서 유사하였다. 각 서브 타입 혈청에 대한 대표적인 하나의 결과가 표시되어 있다.

도 5. 마우스에 있는 CON6 Env 면역원에 의해 유도된 T 세포 면역 반응. 비장 세포를 면역화된 마우스 각각 (5 마우스/그룹)으로부터 분리하였다. CON6 (검은색 컬럼), 서브타입 B (빛금친 컬럼), 서브타입 C (흰색 컬럼) 및 배지 (펩타이드 없음; 회색 컬럼)의 중복되는 Env 펩타이드 풀(pool)을 이용하여 인 비트로에서 비장 세포를 자극시키고 난 뒤, ELISPOT 분석을 이용하여 INF- γ 생산 세포를 확인하였다. CON6 gp120 또는 gp140CF에 의해 유도된 T 세포 INF- γ 반응을, 서브타입 특이 Env 면역원 (JRFL 및 96ZM651)에 의해 유도된 것과 비교하였다. 각 외피 펩타이드 풀에 대한 전체 반응을 백만개 비장세포 당 SFCs로 표시하였다. 각 컬럼에 대한 값은 INF- γ SFCs (n=5 마우스/그룹)의 평균 $\pm$ SEM이다.

도 6a - 6e. 코돈 사용 최적화된 서브타입 C 원형 및 컨센서스 외피 유전자의 (도 6a 및 6b 각각, 서열번호 3 및 4) 컨스트럭션. 고발현된 인간 유전자의 코돈 사용을 반영하기 위해, 원형 및 컨센서스 아미노산 서열 (도 6c 및 6d 각각, 서열번호 5 및 6)을 전사하였다. 20 bp로 중첩되는 쌍으로 된(paired) 올리고뉴클레오타이드 (80-머)를, 제한효소 위치 EcoRI, BbsI, Bam HI 및 BsmBI를 포함하는 5' 불변 서열을 포함하도록 디자인하였다. BbsI 및 BsmBI는 이들의 인식 서열의 바깥을 절단하는 타입 II 제한 효소이다. 쌍으로된 올리고머를, PCR 및 18 bp 불변 서열에 대한 컴플리먼트 프라이머를 이용하여 단계별로 각각 연결시켰으며, 140 bp PCR 산물을 생산하였다. 이들을 pGEM-T에 서브클론하고 서열분석하여, 우연한 돌연변이/절단의 부재를 확인하였다. 적당한 인서트를 포함하는 4개의 pEGM-T 서브클론 각각을 절단한 뒤 pcDNA3.1에 함께 라이게이션하였다. 전체 유전자가 pcDNA3.1에 재컨스트럭트될 때까지, 다중-단편 라이게이션을 유전자의 5'으로부터 3' 말단까지 단편 그룹 간에 단계별로 반복적으로 발생시켰다 (도 6e에 있는 개념도 참조).

도 7. JC53-BL 세포는 CD4 및 HIV-1 공-수용체인 CCR5 및 CXCR4를 높은 레벨로 발현하는 HeLa 세포의 유도체

이다. 또한, 이들은 HIV-1 LTR로부터 각각 발현되는 루시퍼라아제 및 β -갈락토시다아제의 리포터 카세트도 포함한다. 리포터 유전자의 발현은 HIV-1 Tat의 생산에 의존한다. 보다 상세하게는, 세포를 24웰 또는 96웰 플레이트에 접종하여 37°C에서 24시간 동안 배양한 뒤, DEAE-텍스트란을 37°C에서 30분간 처리하였다. 바이러스를 1% DMEM으로 순차적으로 희석한 뒤, DEAE-텍스트란에서 배양되는 세포에 첨가하여 3시간 동안 37°C에서 반응시키고, 각 웰에 추가로 세포 미디어를 첨가하였다. 37°C에서 최종적으로 48시간 동안 배양하고 난 뒤, 세포를 고정하고, β -갈락토시다아제 발현 푸른색 초점들 (foci)을 시각화시키기 위해 X-gal을 이용해 염색하거나, 루시퍼라아제 활성을 측정하기 위해 동결-해동을 3회 반복하였다.

도 8. 서브타입 C 원형 및 컨센서스 env 유전자의 서열 열라인먼트. 서브타입 C 원형 env 서열 (하단 라인, 서열번호 8) 및 컨센서스 env 서열 (상단 라인, 서열번호 7)의 열라인먼트에서는 95.5% 서열 상동성을 나타내었으며; 아미노산 서열 차이가 표시되어 있다. 표시된 차이 하나는 V1 루프의 염기에서의 C 원형 env 유전자에 있는 글리코실화 위치의 추가이다. 플러스 표시는 표시된 위치에서의 아미노산의 내부-클래스 차이를 나타내고; 막대는 아미노산의 클래스에서의 변화를 나타낸다. 잠재적 N-글리코실화 위치가 푸른색으로 표시되어 있다. gp140 유전자에 대한 절단 위치도 표시되어 있다.

도 9. 293T 세포에 있는 서브타입 C 원형 및 컨센서스 외피의 발현. 코돈-최적화된 gp160, gp140 또는 gp120 서브타입 C 원형 및 컨센서스 유전자를 포함하는 플라스미드를 293T 세포에 형질감염시켰으며, 웨스턴 블랏 분석을 이용해 세포 분해물로부터 단백질 발현을 측정하였다. 형질감염후 48시간에, 세포 분해물을 수득하였으며, BCA 단백질 분석을 이용해 전체 단백질 양을 측정하여, 전체 단백질의 2 μ g을 4-20% SDS-PAGE 젤 상의 각 레인에 로딩하였다. 단백질을 PVDF 막으로 트랜스퍼하고 서브타입 C 감염된 환자로의 HIV-1 혈장으로 탐침시켰다.

도 10a 및 10b. 도 10a. 코돈-외적화된 서브타입 C 원형 및 컨센서스 gp160 및 gp140을 이용한 env-결핍 HIV-1의 트랜스 상보성. 코돈-최적화된 서브타입 C 원형 또는 컨센서스 gp160 또는 gp140 유전자를 포함하는 플라스미드를 HIV-1/SG3 Δ env 프로바이러스와 함께 293T 세포로 동시-형질감염시켰다. 형질감염후 48시간에, 유사타입 바이러스를 포함하는 세포 상등액을 수득하여, 원심분리로 청정화시키고, 0.2 μ M 필터를 통해 필터한 뒤, 20% 수크로즈 쿠션(cushion)을 통해 펠렛화시켰다. 각 바이러스 펠렛에 있는 p24의 정량화는 쿨터(Coulter) HIV-1 p24 항원 분석을 이용하여 확인하였으며; 25 ng의 p24를 4-20% SDS-PAGE 젤 상의 코돈-최적화된 외피를 포함하는 입자에 대한 각 레인에 로딩하였다. 250 ng의 p24를 rev-의존성 야생형 서브타입 C 96ZAM651env 유전자의 동시-형질감염에 의해 생성된 입자에 대한 각 레인에 로딩하였다. 각 레인 당 로딩된 p24의 양에 있어서의 차이점은 웨스턴 블랏을 이용하여 rev-의존성 외피의 시각화를 가능하게 하는데 필요하였다. 단백질을 PVDF 막에 트랜스퍼한 뒤, HIV-1 서브타입 B 및 서브타입 C 감염된 환자로부터 모은 혈장으로 탐침시켰다. 도 10b. 서브타입 C 원형 및 컨센서스 외피 당단백질을 포함하는 바이러스 입자의 감염도. 원형 또는 컨센서스 gp160 또는 gp140 외피를 포함하는 유사타입 바이러스의 감염도를 JC53-BL 분석법을 이용하여 측정하였다. 수크로즈 쿠션 정제된 바이러스 입자를 쿨터 p24 항원 분석법으로 분석하고, 각 펠렛의 5배 연속 희석액을 DEAE-텍스트란 처리된 JC53-BL 세포와 반응시켰다. 48시간의 배양 기간 후에, 세포를 고정한 뒤 염색하여 β -갈락토시다아제 발현 세포를 시각화하였다. 감염도를 p24 ng 당 감염성 단위로서 표시함으로써, 주입 유사비리온의 농도 차이를 정량화하였다.

도 11. 서브타입 C 원형 및 컨센서스 외피의 공-수용체 사용. 원형 또는 컨센서스 외피를 포함하는 유사타입 입자를 AMD3100 (CXCR4의 특이 저해제), TAK779 (CCR5의 특이 저해제), 또는 AMD3000+TAK779의 존재하에서 DEAE-텍스트란 처리된 JC53-BL 세포로 반응시켜 공-수용체 사용을 확인하였다. CXCR4를 사용하는 것으로 알려진 분리물인 NL4.3 및 CCR-5-사용 분리물로 알려진 YU-2를 대조군으로서 포함시켰다.

도 12A - 12C. 서브타입 C 원형 및 컨센서스 외피 당단백질의 중화 민감도. 원형, 컨센서스 또는 96ZAM651 gp160 외피를 포함하는 등가량의 유사비리온 (1,500 감염 단위)를 HIV-1 서브타입 C 감염된 환자의 혈장 샘플의 패널과 전-반응시킨 뒤, 96웰 플레이트에 있는 JC53-BL 세포 단일층에 첨가하였다. 플레이트를 2일 동안 배양시킨 뒤 바이러스 감염도의 표지자로서 루시퍼라아제 활성을 측정하였다. 바이러스 감염도는 각 항체 농도에서 생산된 루시퍼라아제 단위 (luciferase units, LU)를 대조군 감염에 의해 생성된 LU로 나눔으로써 계산하였다. 평균 50% 저해 농도 (IC50) 및 각 항체 희석액에서의 실제 % 중화를 각 바이러스에 대해 계산하였다. 모든 루시퍼라아제 실험에서의 결과는, 병행하는 감염에 있는 푸른색 초점을 직접 계수함으로써 확인하였다.

도 13a - 13f. 293T 세포로의 형질감염 이후의 컨센서스 서브타입 C Gag (도 13a) 및 Nef (도 13b)의 단백질 발현. 컨센서스 서브타입 C Gag 및 Nef 아미노산 서열이 도 13c 및 13d (서열번호 9 및 10)에 각각 표시되어

있으며, 이를 인코딩하는 서열이 도 13e 및 13f (서열번호 11 및 12)에 각각 표시되어 있다.

도 14a - 14c. 도 14a 및 14b (서열번호 13 및 14)는 Con-S Env 아미노산 서열 및 이를 코딩하는 서열을 각각 나타낸다. 도 14c는 인 비트로 전사 및 번역 시스템을 이용한 그룹 M 컨센서스 Con-S Env 단백질의 발현을 나타낸다.

도 15a 및 15b. 포유동물 세포에서의 Con-S env 유전자의 발현. (도 15a- 세포 분해물, 도 15b-상등액)

도 16a 및 16b. CON6 및 Con-S env 유전자의 감염도 (도 16a) 및 공-수용체 사용 (도 16b).

도 17a - 17c. CON6 및 Con-S Env-유사비리온에서의 Env 단백질 결합 (도 17a - 분해물, 도 17b - 상등액, 도 17c - 펠렛).

도 18a - 18d. 도 18a 및 18b (서열번호 15 및 16)는 서브타입 A 컨센서스 Env 아미노산 서열 및 이를 인코딩하는 핵산 서열을 각각 나타낸다. 도 18c 및 18d는 포유동물 세포에서의 A.con env 유전자의 발현을 나타낸다 (도 18c - 세포 분해물, 도 18d - 상등액).

도 19a - 19h. M.con.gag (도 19a, 서열번호 17), M.con.pol (도 19b, 서열번호 18), M.con.nef (도 19c, 서열번호 19) 및 C.con.pol (도 19d, 서열번호 20) 핵산 서열 및 이에 대응되는 인코딩된 아미노산 서열 (도 19e - 19h, 각각)(서열번호 21 내지 24).

도 20a - 20d. 서브타입 B 컨센서스 gag (도 20a, 서열번호 25) 및 env (도 20b, 서열번호 26) 유전자. 대응하는 아미노산 서열이 도 20c 및 20d (서열번호 28 및 29)에 표시되어 있다.

도 21. 293T 세포에서의 서브타입 B 컨센서스 env 및 gag 유전자의 발현. 코돈-최적화된 서브타입 B 컨센서스 gp160, gp140 및 gag 유전자를 포함하는 플라스미드를 293T 세포에 형질감염시켰으며, 웨스턴 블랏 분석을 이용하여 세포 분해물로부터 단백질 발현을 측정하였다. 형질감염후 48시간에, 세포 분해물을 수득한 뒤, 전체 단백질 양을 BCA 단백질 분석법으로 측정하고, 전체 단백질 중 2 μ g을 4-20% SDS-PAGE 젤의 각 레인에 로딩하였다. 단백질을 PVDF 막에 트랜스퍼한 뒤 HIV-1 서브타입 B 감염된 환자의 혈청으로 탐침시켰다.

도 22. 서브타입 B 컨센서스 외피의 공-수용체 사용. 서브타입 B 컨센서스 gp160 Env를 포함하는 유사타입 입자를 AMD3100 (CXCR4의 특이 저해제), TAK779 (CCR5의 특이 저해제) 및 AMD3000+TAK779의 존재하에서 DEAE-텍스트란 처리된 JC53-BL 세포로 반응시켜 공-수용체 사용을 측정하였다. CXCR4를 이용하는 것으로 알려진 분해물인 NL4.3 및 CCR5-이용 분해물로 알려진 YU-2를 대조군으로 포함시켰다.

도 23a 및 23b. env-결핍 HIV-1의 코돈-최적화된 서브타입 B 컨센서스 gp160 및 gp140 유전자와의 트랜스 상보성. 코돈-최적화된 서브타입 B 컨센서스 gp160 또는 gp140 유전자를 포함하는 플라스미드를 HIV-1/SG3 Δ env 프로바이러스(provirus)와 함께 293T 세포에 동시-형질감염시켰다. 형질감염후 48시간에 유사타입 바이러스를 포함하는 세포 상등액을 수득한뒤, 원심분리로 청정화하고, 0.2 μ M 필터를 통해 필터한후, 펠렛을 20% 수크로즈 쿠션으로 통과시켰다. 각 바이러스 펠렛에서의 p24의 정량은 쿨터 HIV-1 p24 항원 분석법을 이용하여 확인하였으며; 25 ng의 p24를 4-20% SDS-PAGE 젤 상의 각 레인에 로딩시켰다. 단백질을 PVDF 막에 트랜스퍼하였으며, 감염된 HIV-1 서브타입 B 환자 혈청의 항-HIV-1 항체로 탐침시켰다. rev-의존 NL4.3 env를 이용한 트랜스 상보성을 대조군으로 포함시켰다. 도 23b. 서브타입 B 컨센서스 외피를 포함하는 바이러스 입자의 감염도. 컨센서스 B gp160 또는 gp140을 포함하는 유사타입 바이러스는 JC53-BL 분석법을 이용하여 확인하였다. 수크로즈 쿠션 정제된 바이러스 입자들을 쿨터 p24 항원 분석법으로 분석하였으며, 각 펠렛의 5배 연속 희석물을 DEAE-텍스트란 처리된 JC53-BL 세포와 반응시켰다. 48시간의 반응기간 후에, 세포를 고정된 뒤 염색하여 β -갈락토시 다아제 발현 세포를 시각화하였다. 감염도를 p24 ng 당 감염 단위로서 표현하였다.

도 24a - 24d. 서브타입 B 컨센서스 gp160 외피를 포함하는 비리온의 중화 민감도. 서브타입 B 컨센서스 또는 NL4.3 Env (gp160)를 포함하는 유사비리온 등가량 (1,500 감염 단위)을 3개의 다른 모노클로날 중화 항체 및 HIV-1 서브타입 B 감염된 환자의 혈장 샘플 패널과 전반응시켰으며, 이를 96웰 플레이트에 있는 JC53-BL 세포 단일층에 첨가하였다. 상기 플레이트를 2일 동안 배양한 뒤 바이러스 감염도의 표지자로서 루시퍼라아제 활성을 측정하였다. 바이러스 감염도는 각 항체 농도에서 생산된 루시퍼라아제 단위 (LU)를 대조군 감염에서 생성된 LU로 나눔으로써 계산하였다. 평균 50% 저해 농도 (IC₅₀) 및 각 항체 희석액에서의 실제 % 중화를 각 바이러스에 대해 계산하였다. 모든 루시퍼라아제 실험에서의 결과는, 병행하는 감염에 있는 푸른색 초점을 직접 계수함으로써 확인하였다. 도 24a. 서브타입 B 컨센서스 Env (gp160)를 포함하는 유사비리온의 중화. 도 24b. NL4.3 Env (gp160)을 포함하는 유사비리온의 중화. 도 24c. 서브타입 B 컨센서스 Env (gp160)를 포함하는 유

사비리온의 중화. 도 24d. NL4.3 Env (gp160)을 포함하는 유사비리온의 중화.

도 25a 및 25b. 도 25a. 수크로즈 농도구배 분획의 농도 및 p24 분석. 0.5 ml 분획을 20-60% 농도구배 수크로즈로부터 수득하였다. 분획 번호 1은 농도구배 튜브의 밑바닥으로부터 얻은 가장 농도가 짙은 분획을 나타낸다. 농도는 굴절계 (refractometer)를 이용하여 측정하였으며, 각 분획에서의 p24의 양은 쿨터 24 항원 분석법으로 확인하였다. 분획 6-9, 10-15, 16-21 및 22-25를 함께 모았으며, 웨스턴 블랏으로 분석하였다. 예 측한 대로, 비리온은 1.16-1.18 g/ml 농도에서 침전되었다. 도 25b. 서브타입 B 컨센서스 gag 및 env 유전자의 동시-형질감염에 의한 VLP 생산. 293T 세포를 서브타입 B 컨센서스 gag 및 env 유전자로 동시-형질감염시켰다. 형질감염후 48시간에 세포 상등액을 수득하였으며, 20% 수크로즈 쿠션을 통해 청정화하고, 20-60% 농도구배의 수크로즈를 통해 추가로 정제하였다. 농도구배로부터 선택된 분획을 모아, 20 ml의 PBS에 첨가한 뒤, 100,000 ×g에서 밤새 원심분리하였다. 재현탁된 펠렛을 4-20% SDS-PAGE 젤 상에 로딩하였으며, 단백질을 PVDF 막에 트랜스퍼하고, HIV-1 서브타입 B 감염된 환자로의 혈장으로 탐침시켰다.

도 26a 및 26b. 도 26a. 2000 Con-S 140CFI.ENV (서열번호 30). 도 26b. 코돈-최적화된 2000년도 Con-S 140CFI.seq (서열번호 31).

도 27. HIV-1 외피 펩타이드에 반응하는 각 C57BL/6 마우스 T세포. CON6 gp140CFI 및 Con-S gp140CFI의 C57BL/6 마우스에서의 비교 면역원성. 마우스를 DNA 프라임에 있는 HIV5305 (서브타입 A), 2801 (서브타입 B), CON6 또는 Con-S 외피 유전자로 그룹당 5마리씩 면역반응시키고, rVV 추가접종시켰다. rVV 추가접종(boost) 후 10일에, HIV-1 UG37 (A), MN (B), Ch19 (C), 89.6 (B), SF162 (B) 또는 펩타이드가 없는 음성 대조군의 Envs로부터의 중첩되는 펩타이드 혼합물을 이용하여, 비장 세포를 IFN- γ 스팟-형성 세포에 대하여 분석하였다.

도 28a - 28c. 도 28a. Con-B 2003 Env. 펩타이드 (841개 아미노산) (서열번호 32). 밑줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밑줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 28b. Con-B-140CF.펩타이드 (632개 아미노산) (서열번호 33). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계(junction)를 나타낸다. 도 28c. 코돈-최적화된 Con-B 140CF. 서열 (1927개 뉴클레오티드) (서열번호 34).

도 29a - 29c. 도 29a. CON_OF_CONS-2003 (829개 아미노산) (서열번호 35). 밑줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밑줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 29b. ConS-2003 140CF.펩타이드 (620개 아미노산) (서열번호 36). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 29c. 코돈-최적화된 ConS-2003 140CF. 서열 (1891개 뉴클레오티드) (서열번호 37).

도 30a - 30c. 도 30a. CONSESUS_A1-2003 (845개 아미노산) (서열번호 38). 밑줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밑줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 30b. Con-A1-2003 140CF.펩타이드 (629개 아미노산) (서열번호 39). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 30c. 코돈-최적화된 Con-A1-2003 서열 (서열번호 40).

도 31a - 31c. 도 31a. CONSESUS_C-2003 (835개 아미노산) (서열번호 41). 밑줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밑줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 31b. Con-C-2003 140CF.펩타이드 (619개 아미노산) (서열번호 42). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 31c. 코돈-최적화된 Con-C-2003 140 CF (1,888개 뉴클레오티드) (서열번호 43).

도 32a - 32c. 도 32a. CONSESUS_G-2003 (842개 아미노산) (서열번호 44). 밑줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밑줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 32b. Con-G-2003 140CF.펩타이드 (626개 아미노산) (서열번호 45). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 32c. 코돈-최적화된 Con-G-2003 서열 (서열번호 46).

도 33a - 33c. 도 33a. CONSESUS_01_AE-2003 (854개 아미노산) (서열번호 47). 밑줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밑줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 33b. Con-AE01-2003 140CF.펩타이드 (638개 아미노산) (서열번호 48). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 33c. 코돈-최적화된 Con-AE01-

2003 서열 (1945개 뉴클레오타이드) (서열번호 49).

도 34a - 34c. 도 34a. 야생형 서브타입 A Env. 00KE_MSA4076-A (서브타입 A, 891개 아미노산) (서열번호 50). 밀줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밀줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 34b. 00KE_MSA4076-A 140CF. 펩타이드 (647개 아미노산) (서열번호 51). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 34c. 코돈-최적화된 00KE_MSA4076-A 140CF 서열 (1972개 뉴클레오타이드) (서열번호 52).

도 35a - 35c. 도 35a. 야생형 서브타입 B. QH0515.1g gp160 (861개 아미노산) (서열번호 53). 밀줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밀줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 35b. QH0515.1g 140CF (651개 아미노산) (서열번호 54). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 35c. 코돈-최적화된 QH0515.1g 140CF 서열 (1984개 뉴클레오타이드) (서열번호 55).

도 36a - 36c. 도 36a. 야생형 서브타입 C. DU123.6 gp160 (854개 아미노산) (서열번호 56). 밀줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밀줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 36b. DU123.6 140CF (638개 아미노산) (서열번호 57). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 36c. 코돈-최적화된 DU123.6 140CF 서열 (1945개 뉴클레오타이드) (서열번호 58).

도 37a - 37c. 도 37a. 야생형 서브타입 CRF01_AE. 97CNGX2F-AE (854개 아미노산) (서열번호 59). 밀줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밀줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 37b. 97CNGX2F-AE 140CF 펩타이드 (629개 아미노산) (서열번호 60). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 37c. 코돈-최적화된 97CNGX2F-AE 140CF 서열 (1921개 뉴클레오타이드) (서열번호 61).

도 38a - 38c. 도 38a. 야생형 DRCBL-G (854개 아미노산) (서열번호 62). 밀줄친 아미노산 서열은 140CF 디자인에서 결실된 융합 도메인이고, 밀줄친 "W"는 C-말단의 마지막 아미노산이며, "W" 이후의 모든 아미노산은 140CF 디자인에 결실되어 있다. 도 38b. DRCBL-G 140CF 펩타이드 (630개 아미노산) (서열번호 63). 굵게 표시된 아미노산은 결실된 융합 절단 위치의 경계를 나타낸다. 도 38c. 코돈-최적화된 DRCBL-G 140CF 서열 (1921개 뉴클레오타이드) (서열번호 64).

도 39a 및 39b. 도 39a. 2003 Con-S Env (서열번호 65). 도 39b. 2003 Con-S Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 72).

도 40a 및 40b. 도 40a. 2003 M.Group.Anc Env (서열번호 66). 도 40b. 2003 M. Group.anc Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 67)

도 41a 및 41b. 도 41a. 2003 CON_A1 Env (서열번호 68). 도 41b. 2003 CON_A1 Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 70)

도 42a 및 42b. 도 42a. 2003 A1.Anc Env (서열번호 69). 도 42b. 2003 A1.anc Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 71)

도 43a 및 43b. 도 43a. 2003 CON_A2. Env (서열번호 73). 도 43b. 2003 CON_A2. Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 75)

도 44a 및 44b. 도 44a. 2003 CON_B Env (서열번호 74). 도 44b. 2003 CON_B Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 76)

도 45a 및 45b. 도 45a. 2003 B.anc Env (서열번호 77). 도 45b. 2003 B.anc Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 79)

도 46a 및 46b. 도 46a. 2003 CON_C Env (서열번호 78). 도 46b. 2003 CON_C Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 80)

도 47a 및 47b. 도 47a. 2003 C.anc Env (서열번호 81). 도 47b. 2003 C.anc Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 83)

도 48a 및 48b. 도 48a. 2003 CON_D Env (서열번호 82). 도 48b. 2003 CON_D Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 83)

도 49a 및 49b. 도 49a. 2003 CON_F1 Env (서열번호 85). 도 49b. 2003 CON_F1 Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 87)

도 50a 및 50b. 도 50a. 2003 CON_F2 Env (서열번호 86). 도 50b. 2003 CON_F2 Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 88)

도 51a 및 51b. 도 51a. 2003 CON_G Env (서열번호 89). 도 51b. 2003 CON_G Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 91)

도 52a 및 52b. 도 52a. 2003 CON_H Env (서열번호 90). 도 52b. 2003 CON_H Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 92)

도 53a 및 53b. 도 53a. 2003 CON_01_AE Env (서열번호 93). 도 53b. 2003 CON_01_AE Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 95)

도 54a 및 54b. 도 54a. 2003 CON_02_AG Env (서열번호 94). 도 54b. 2003 CON_02_AG Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 96)

도 55a 및 55b. 도 55a. 2003 CON_03_AB Env (서열번호 97). 도 55b. 2003 CON_03_AB Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 99)

도 56a 및 56b. 도 56a. 2003 CON_04_CPX Env (서열번호 98). 도 56b. 2003 CON_04_CPX Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 100)

도 57a 및 57b. 도 57a. 2003 CON_06_CPX Env (서열번호 101). 도 57b. 2003 CON_06_CPX Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 103)

도 58a 및 58b. 도 58a. 2003 CON_08_BC Env (서열번호 102). 도 58b. 2003 CON_08_BC Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 104)

도 59a 및 59b. 도 59a. 2003 CON_10_CD Env (서열번호 105). 도 59b. 2003 CON_10_CD Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 107)

도 60a 및 60b. 도 60a. 2003 CON_11_CPX Env (서열번호 106). 도 60b. 2003 CON_11_CPX Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 108)

도 61a 및 61b. 도 61a. 2003 CON_12_BF Env (서열번호 109). 도 61b. 2003 CON_12_BF Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 111)

도 62a 및 62b. 도 62a. 2003 CON_14_BG Env (서열번호 110). 도 62b. 2003 CON_14_BG Env.seq.opt. (Seq.opt = 코돈 최적화된 인코딩 서열) (서열번호 112)

도 63a 및 63b. 도 63a. 2003_CON_S gag.PEP (서열번호 113). 도 63b. 2003_CON_S gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 114)

도 64a 및 64b. 도 64a. 2003_M.GROUP.anc gag.PEP (서열번호 115). 도 64b. 2003_M.GROUP.anc gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 116)

도 65a - 65d. 도 65a. 2003_CON_A1 gag.PEP (서열번호 117). 도 65b. 2003_CON_A1 gag.OPT (서열번호 118). 도 65c. 2003_A1.anc gag.PEP (서열번호 119). 도 65d. 2003_A1.anc gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 120)

도 66a 및 66b. 도 66a. 2003_CON_A2 gag.PEP (서열번호 121). 도 66b. 2003_CON_A2 gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 122)

도 67a - 67d. 도 67a. 2003_CON_B gag.PEP (서열번호 123). 도 67b. 2003_CON_B gag.OPT (서열번호 124). 도 67c. 2003_B.anc gag.PEP (서열번호 125). 도 67d. 2003_B.anc gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 126)

도 68a - 68d. 도 68a. 2003_CON_C gag.PEP (서열번호 127). 도 68b. 2003_CON_C gag.OPT (서열번호 128). 도 68c. 2003_C.anc gag.PEP (서열번호 129). 도 68d. 2003_C.anc gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 130)

도 69a 및 69b. 도 69a. 2003_CON_D gag.PEP (서열번호 131). 도 69b. 2003_CON_D gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 132)

도 70a 및 70b. 도 70a. 2003_CON_F gag.PEP (서열번호 133). 도 70b. 2003_CON_F gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 134)

도 71a 및 71b. 도 71A. 2003_CON_G gag.PEP (서열번호 135). 도 71b. 2003_CON_G gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 136)

도 72a 및 72b. 도 72A. 2003_CON_H gag.PEP (서열번호 137). 도 72b. 2003_CON_H gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 138)

도 73a 및 73b. 도 73a. 2003_CON_K gag.PEP (서열번호 139). 도 73b. 2003_CON_K gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 140)

도 74a 및 74b. 도 74a. 2003_CON_01_AE gag.PEP (서열번호 141). 도 74b. 2003_CON_01_AE gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 142)

도 75a 및 75b. 도 75a. 2003_CON_02_AG gag.PEP (서열번호 143). 도 75b. 2003_CON_02_AG gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 144)

도 76a 및 76b. 도 76a. 2003_CON_03_ABG gag.PEP (서열번호 145). 도 76b. 2003_CON_03_ABG gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 146)

도 77a 및 77b. 도 77a. 2003_CON_04_CFX gag.PEP (서열번호 147). 도 77b. 2003_CON_04_CFX gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 148)

도 78a 및 78b. 도 78a. 2003_CON_06_CPX gag.PEP (서열번호 150). 도 78b. 2003_CON_06_CPX gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 151)

도 79a 및 79b. 도 79a. 2003_CON_07_BC gag.PEP (서열번호 152). 도 79b. 2003_CON_07_BC gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 153)

도 80a 및 80b. 도 80a. 2003_CON_08_BC gag.PEP (서열번호 154). 도 80b. 2003_CON_08_BC gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 155)

도 81a 및 81b. 도 81a. 2003_CON_10_CD gag.PEP (서열번호 156). 도 81b. 2003_CON_10_CD gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 157)

도 82a 및 82b. 도 82a. 2003_CON_11_CPX gag.PEP (서열번호 158). 도 82b. 2003_CON_11_CPX gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 159)

도 83a 및 83b. 도 83a. 2003_CON_12_BF gag.PEP (서열번호 160). 도 83b. 2003_CON_12_BF gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 161)

도 84a 및 84b. 도 84a. 2003_CON_14_BG gag.PEP (서열번호 162). 도 84b. 2003_CON_14_BG gag.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 163)

도 85a 및 85b. 도 85a. 2003_CONS nef.PEP (서열번호 164). 도 85b. 2003_CONS nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 165)

도 86a 및 86b. 도 86a. 2003_M GROUP.anc nef.PEP (서열번호 166). 도 86b. 2003_M GROUP.anc.nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 167)

도 87a 및 87b. 도 87a. 2003_CON_A nef.PEP (서열번호 168). 도 87b. 2003_CON_A nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 169)

도 88a - 88d. 도 88a. 2003_CON_A1 nef.PEP (서열번호 170). 도 88b. 2003_CON_A1 nef.OPT (서열번호 171).

도 88c. 2003_A1.anc nef.PEP (서열번호 172). 도 88d. 2003_A1.anc nef.OPT. (OPT = 코돈 최적화된 인코딩 서열) (서열번호 173)

도 89a 및 89b. 도 89a. 2003_CON_A2 nef.PEP (서열번호 174). 도 89b. 2003_CON_A2 nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 175)

도 90a - 90d. 도 90a. 2003_CON_B nef.PEP (서열번호 176). 도 90b. 2003_CON_B nef.OPT (서열번호 177). 도 90c. 2003_B.anc nef.PEP (서열번호 178). 도 90d. 2003_B.anc nef.OPT. (OPT = 코돈 최적화된 인코딩 서열) (서열번호 179)

도 91a 및 91b. 도 91a. 2003_CON_02_AG nef.PEP (서열번호 180). 도 91b. 2003_CON_02_AG nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 181)

도 92a - 92d. 도 92a. 2003_CON_C nef.PEP (서열번호 182). 도 92b. 2003_CON_C nef.OPT (서열번호 183). 도 92c. 2003_C.anc nef.PEP (서열번호 184). 도 92d. 2003_C.anc nef.OPT. (OPT = 코돈 최적화된 인코딩 서열) (서열번호 185)

도 93a 및 93b. 도 93a. 2003_CON_D nef.PEP (서열번호 186). 도 93b. 2003_CON_D nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 187)

도 94a 및 94b. 도 94a. 2003_CON_F1 nef.PEP (서열번호 188). 도 94b. 2003_CON_F1 nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 189)

도 95a 및 95b. 도 95a. 2003_CON_F2 nef.PEP (서열번호 190). 도 95b. 2003_CON_F2 nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 191)

도 96a 및 96b. 도 96a. 2003_CON_G nef.PEP (서열번호 192). 도 96b. 2003_CON_G nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 193)

도 97a 및 97b. 도 97a. 2003_CON_H nef.PEP (서열번호 194). 도 97b. 2003_CON_H nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 195)

도 98a 및 98b. 도 98a. 2003_CON_01_AE nef.PEP (서열번호 196). 도 98b. 2003_CON_01_AE nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 197)

도 99a 및 99b. 도 99a. 2003_CON_03_AE nef.PEP (서열번호 198). 도 99b. 2003_CON_03_AE nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 199)

도 100a 및 100b. 도 100a. 2003_CON_04_CFX nef.PEP (서열번호 200). 도 100b. 2003_CON_04_CFX nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 201)

도 101a 및 101b. 도 101a. 2003_CON_06_CFX nef.PEP (서열번호 202). 도 101b. 2003_CON_06_CFX nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 203)

도 102a 및 102b. 도 102a. 2003_CON_08_BC nef.PEP (서열번호 204). 도 102b. 2003_CON_08_BC nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 205)

도 103a 및 103b. 도 103a. 2003_CON_10_CD nef.PEP (서열번호 206). 도 103b. 2003_CON_10_CD nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 207)

도 104a 및 104b. 도 104a. 2003_CON_11_CFX nef.PEP (서열번호 208). 도 104b. 2003_CON_11_CFX nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 209)

도 105a 및 105b. 도 105a. 2003_CON_12_BF nef.PEP (서열번호 210). 도 105b. 2003_CON_12_BF nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 211)

도 106a 및 106b. 도 106a. 2003_CON_14_BG nef.PEP (서열번호 212). 도 106b. 2003_CON_14_BG nef.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 213)

도 107a 및 107b. 도 107a. 2003_CON_S pol.PEP (서열번호 214). 도 107b. 2003_CON_S pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 215)

도 108a 및 108b. 도 108a. 2003_M GROUP anc pol.PEP (서열번호 216). 도 108b. 2003_M.GROUP anc pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 218)

도 109a - 109d. 도 109a. 2003_CON_A1 pol.PEP (서열번호 217). 도 109b. 2003_CON_A1 pol.OPT (서열번호 219). 도 109c. 2003_A1.anc pol.PEP (서열번호 220). 도 109d. 2003_A1.anc pol.OPT. (OPT = 코돈 최적화된 인코딩 서열) (서열번호 221)

도 110a 및 110b. 도 110a. 2003_CON_A2 pol.PEP (서열번호 222). 도 110b. 2003_CON_A2 pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 224)

도 111a - 111d. 도 111a. 2003_CON_B pol.PEP (서열번호 223). 도 111b. 2003_CON_B pol.OPT (서열번호 225). 도 111c. 2003_B.anc pol.PEP (서열번호 226). 도 111d. 2003_B.anc pol.OPT. (OPT = 코돈 최적화된 인코딩 서열) (서열번호 227)

도 112a - 112d. 도 112a. 2003_CON_C pol.PEP (서열번호 228). 도 112b. 2003_CON_C pol.OPT (서열번호 229). 도 112c. 2003_C.anc pol.PEP (서열번호 230). 도 112d. 2003_C.anc pol.OPT. (OPT = 코돈 최적화된 인코딩 서열) (서열번호 231)

도 113a 및 113b. 도 113A. 2003_CON_D pol.PEP (서열번호 232). 도 113b. 2003_CON_D pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 234)

도 114a 및 114b. 도 114A. 2003_CON_F1 pol.PEP (서열번호 233). 도 114b. 2003_CON_F1 pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 235)

도 115a 및 115b. 도 115A. 2003_CON_F2 pol.PEP (서열번호 236). 도 115b. 2003_CON_F2 pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 238)

도 116a 및 116b. 도 116a. 2003_CON_G pol.PEP (서열번호 237). 도 116b. 2003_CON_G pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 239)

도 117a 및 117b. 도 117a. 2003_CON_H pol.PEP (서열번호 240). 도 117b. 2003_CON_H pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 242)

도 118a 및 118b. 도 118a. 2003_CON_01_AE pol.PEP (서열번호 241). 도 118b. 2003_CON_01_AE pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 243)

도 119a 및 119b. 도 119a. 2003_CON_02_AG pol.PEP (서열번호 244). 도 119b. 2003_CON_02_AG pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 246)

도 120a 및 120b. 도 120a. 2003_CON_03_AB pol.PEP (서열번호 245). 도 120b. 2003_CON_03_AB pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 247)

도 121a 및 121b. 도 121a. 2003_CON_04_CPX pol.PEP (서열번호 248). 도 121b. 2003_CON_04_CPX pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 250)

도 122A 및 122B. 도 122A. 2003_CON_06_CPX pol.PEP (서열번호 249). 도 122B. 2003_CON_06_CPX pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 251)

도 123A 및 123B. 도 123A. 2003_CON_08_BC pol.PEP (서열번호 252). 도 123B. 2003_CON_08_BC pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 254)

도 124A 및 124B. 도 124A. 2003_CON_10_CD pol.PEP (서열번호 253). 도 124B. 2003_CON_10_CD pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 255)

도 125A 및 125B. 도 125A. 2003_CON_11_CPX pol.PEP (서열번호 256). 도 125B. 2003_CON_11_CPX pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 258)

도 126A 및 126B. 도 126A. 2003_CON_12_BF pol.PEP (서열번호 257). 도 126B. 2003_CON_12_BF pol.OPT (OPT = 코돈 최적화된 인코딩 서열) (서열번호 259)

도 127a 및 127b. 도 127a. 2003_CON_14_BG pol.PEP (서열번호 260). 도 127B. 2003_CON_14_BG pol.OPT (OPT

= 코돈 최적화된 인코딩 서열) (서열번호 261)

발명을 실시하기 위한 구체적인 내용

- [0012] 본 발명의 일 구현예에 따르면, 컨센서스 env 유전자는 특정 HIV-1 그룹 (그룹 M은 서브타입 A-D, F-H, J 및 K로 분류됨)의 서브타입 각각에 대한 env 유전자의 컨센서스 서열을 제조함으로써, 예를 들어 Los Alamos HIV 서열 데이터베이스에 있는 서열로부터 (예컨대 MASE(Multiple Aligned Sequence Editor)를 이용하여) 컨스트럭트될 수 있다. 그 후, 대량으로 서열분석된 서브타입을 피하기 위해, 모든 서브타입 컨센서스들의 컨센서스 서열이 제조될 수 있다 (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). 실시예 1에 개시된 그룹 M 컨센서스 env 유전자 (CON6로 표시됨)의 경우, CRF08_BC 제조함 균주 (98CN006)의 5개 고변이성 부위 (V1, V2, V4, V5 및 gp41에 있는 세포질 도메인에 있는 부위)가 서열에 있는 결실 부위를 채우는데 사용된다 (참조, Con-S에 대한 대응되는 부위). 높은 레벨의 발현을 위해, 고발현된 인간 유전자에 대한 코돈 사용에 기초하여 컨센서스 또는 원형 유전자의 코돈이 최적화될 수 있다 (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)).
- [0013] 1999년도 컨센서스 그룹 M env 유전자와 함께, CON6는, ELISPOT γ -인터페론 비장 스팟 형성 세포의 수 및 2개의 마우스 균주에서 인식된 에피토프의 수를 이용함으로써, CON6 대(versus) 야생형 B 및 C env에 의한 상위 (superior) T 세포 반응의 유도를 설명하는 것이 가능하였다 (표 1 및 2는 BALB/c 마우스의 데이터를 보여준다). HIV-1 제1 분리물에 대한 항체 중화를 유도하는 CON6 Env 단백질의 능력을 수개의 서브타입 B Env와 비교하였다. CON6에 의해 유도된 중화 항체의 표적은 수개의 비-B HIV-1 균주를 포함한다.

[표 1]

BALB/c 마우스에서의 CON6, JRFL 및 96ZM651 및 Env 면역원의
T-세포 에피토프 매핑
(표 1의 서열은 각각 서열번호 262 내지 287를 순서대로 나타냄)

펩타이드	면역원			T-세포 반응
	CON6	JRFL(B)	96ZM651(C)	
CON6(그룹M컨센서스)				
16 DTEVHNWVWTHACVP	+		+	CD4
48 KNSSEYYRLINCNTS	+		+	CD4
49 EYYRLINCNTSAITQ				
53 CPKVSFEPIPIHYCA	+			CD4
54 SFEPIPIHYCAPAGF				
62 NVSTVQCTHGKIPVV	+			CD4
104 ETITLPCRIGKIINM	+			CD8
105 LPCRIKIINMWQGV				
130 GIVQQQSNLLRAIEA	+			CD4
131 VQQSNLLRAIEAQQHL				
134 AQQHLQLTLVWGKQLQ	+			CD4
135 LQLTVWGKQLQARVL				
서브타입B (MN)				
6223 AKAYDTEVHNWVWATQ	+			CD4
6224 DTEVHNWVWATQACVP				
6261 ACPKISFEPIPIHYC	+			CD4
6262 ISFEPIPIHYCAPAG				
6286 RKRIHIGPGRAFYTT		+		CD8
6287 HIGPGRAFYTTKNII				
6346 IVQQQNNLLRAIEAQ	+			CD4
6347 QNNLLRAIEAQQHML				
서브타입C(chnl9)				
4834 VPVWKEAKTTLFCASDAKSY			+	CD4
4836 GKEVHNWVWTHACVPTDPNP	+		+	CD4
4848 SSENSSEYYRLINCNTSAIT	+		+	CD4
4854 STVQCTHGKIPVYSTQLLLN	+			CD4
4884 QQSNLLRAIEAQGHLLQLTV	+			CD4
4885 AQGHLLQLTVWGKQLQTRV	+			CD4

[0015]

[0016] [표 2]

C57BL/6 마우스에서의 CON6, gp120 면역원의 T 세포 에피토프 매핑
(표 2의 서열은 각각 서열번호 288 내지 304를 순서대로 나타냄)

펩타이드	펩타이드서열	T 세포 반응
CON 6 (컨센서스)	2 GIQRNCQHLWRWGT	CD8
	3 NCQHLWRWGTMLGM	
	16 DTEVHNWATHACVP	CD4
	53 CPKVSFEPIPIHYCA	CD4
	97 FYCNTSGLFNSTWMF	CD8
	99 FNSTWMFNGTYMFNG	CD8
서브타입 B (MN)	6210 GIRRN'YQHW'WGWT	CD8
	6211 NYQHW'WGWTMLLGL	
	6232 NMWKNNMVEQMHEDI	CD4
	6262 ISFEPIPIHYCAPAG	CD4
	6290 NIIGTIRQAHCNISR	CD4
	6291 TIRQAHCNISRAKWN	
서브타입 C (Chn 19)	4830 MRVTGIRKNYQHLWRWGT ML	CD8
	5446 RWGTMLLGMLMCSAAEN	CD8
	4836 GKEVHNWATHACVPTDP NP	CD4
	4862 GDIRQAHCNISKDKWNETL Q	CD4
	4888 LLGIWGCSGKLICTTTVPWN	CD8

[0017]

[0018]

2000년도 컨센서스 그룹 M env 유전자인, Con-S에 대해, 2개의 마우스 균주에서의 T 세포 γ 인터페론 ELISPOT 분석에서 Con-S 외피가 CON6 외피 유전자와 같은 면역원으로 밝혀졌다 (C57BL/6에 대한 데이터가 도 27에 나타나 있다). 더불어, 기니아 피그에 있는 항체에 대한 단백질 면역원으로서 CON6 및 Con-S gp140 Envs의 비교에서 (표 3), gp140 Envs 모두는 중화된 서브타입 B 제1 분리물인 항체를 유도하는 것으로 밝혀졌다. 그러나, Con-S gp140도 서브타입 C 분리물 TV-1 및 DU123 뿐만 아니라 하나의 서브타입 A HIV-1 제1 분리물의 강한 (robust) 중화를 유도하였으나, CON6는 그렇지 않았다.

[0019] [표 3]

그룹 M 컨센서스 CON6 및 Con-S Envs의 HIV-1 제1 부리온의 종화를 유도하려는 능력

HIV-1 부리온 (Subtype)	CON6 gp140CF							CON6 gp140 CFI						
	CON6 gp140CF							CON6 gp140 CFI						
	770	771	772	775	781	783	784	788	776	777	778	780	782	784
EX08(E)	520	257	428	189	218	164	>540	199	>540	>640	>540	>540	>540	>540
QH0892 (B)	46	55	58	77	<20	91	100	76	109	<20	<20	<20	<20	<20
SS1198(E)	386	306	284	222	431	242	>540	351	>540	296	>540	>540	>540	>540
JFLFL(B)	<20	<20	<20	<20	<20	169	<20	<20	<20	<20	<20	<20	<20	<20
BG1168(F)	<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
6-01(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
TY-1(C)	<20	<20	<20	<20	<20	<20	<20	<20	356	439	>540	>540	>540	>540
DJ123(C)	<20	<20	71	74	<20	72	<20	<20	176	329	387	376	376	376
DUI72(C)	<20	<20	96	84	<20	<20	<20	<20	<20	235	<20	<20	<20	<20
ZM18108(S(C))	ND	ND	ND	ND	<20	<20	<20	<20	84	61	86	43	43	43

[0020]

[illegible]

[0021]

[0022]

[0023]

[0024]

[0025]

‡ 4회 및 5회 면역화 후의 50% 중화 역가

2000년도 Con-S 140CFI.ENV 서열이 도 26A에 표시되어 있다. Gp140 CFI는 막투과 및 세포질 도메인의 결실에 추가로, 절단-위치가 결실되고 (c), 융합-위치가 결실되며 (F) gp41 면역우세(immunodominant) 부위가 결실되는 (I) HIV-1 외피 디자인을 지칭한다. 코돈-최적화된 2000년도 Con-S 140 CFI 서열이 도 26B에 표시되어 있다.

컨센서스 번역원의 다음 반복으로서, 그리고 실제의 HIV-1 번역원이 수개의 서브타입 컨센서스 유전자의 다가 혼합물이 될 수 있다는 사실의 인지에서, 서브타입 및 컨센서스 유전자의 혼합물 또는 센트럴화된 유전자 및 야생형 유전자의 혼합물, 11개 서브타입 컨센서스의 시리즈 및 야생형 유전자들이, 2003년도 Los Alamos 국립 데이터베이스 서열로부터의 그룹 M 컨센서스 유전자 뿐만 아니라 서브타입 A, B, C, CRF AE01 및 G로부터 디자인되었다. 야생형 서열들은, 이들이 초기 유전된(transmitted) HIV-1 군주 (이들 군주들은 백신에 의해 대항하여 보호되어야 하는 필요성의 가능성이 가장 큼)로부터 유발되는 것으로 알려져 있기 때문이나, 이 서브타입의 데이터베이스에 있는 가장 최근에 기탁된 군주이기 때문에 선택되었다. 이들 뉴클레오티드 및 아미노산 서열들이 도 28-38에 개시되어 있다 (표시된 모든 140CF 디자인에 대해, 140CF 유전자는 코작(Kozak) 서열 (GCCACCATGG/A) (서열번호 306) 및 SalI 위치를 포함하는 5' 서열 "TTCAGTCGACGCCACC" (서열번호 305) 및 중지 코돈 및 BglII 위치를 포함하는 3' 서열 TAAAGATCTTACAA (서열번호 307)로 플랜킹될 수 있다). 도 39 - 62에 표시된 것은 2003 센트럴화된 (컨센서스 및 원형) HIV-1 외피 단백질 및 코돈 최적화된 유전자 서열이다.

CON6 gp140 (이는 비-클레이드 (clade) B HIV 균주를 중화시키지 않음) 및 Con-S gp140 (이는 비-클레이드 B HIV 균주를 중화시키는 항체를 유도함) 사이의 주요한 차이가 Con-S V1, V2, V4 및 V5 부위에 있다. 클레이드

B 군주에 대해, V3 부위의 펩타이드들이 중화 항체를 유도할 수 있다 (Haynes et al, J. Immunol. 151:1646-1653 (1993)). 그래서, Th-V1, Th-V2, Th-V4 및 Th-V5 펩타이드의 컨스트럭션은 광범위하게 활성을 띄는 요구되는 항-비-클레이드 B 중화 항체를 유발하는 것으로 추측될 수 있다. 따라서, 표 4에 표시되어 있는 Th-V 펩타이드들은, Con-S gp140으로부터 유래되는 펩타이드 면역원(들)으로서 사용하기 위해 고찰될 수 있다. gag Th 결정인자(determinant) (GTH, 표 4) 또는 다른 HIV 군주에 있는 모든 상동성 GTH 서열이, 면역원성을 촉진시키기 위해 사용될 수 있으며, HIV gp120의 C4 부위는 well (KQIINMQVVGKAMYA) (서열번호 308) 또는 다른 HIV 군주로부터의 상동성 C4 서열로서 사용될 수 있다 (Haynes et al, J. Immunol. 151:1646-1653 (1993)). N-말단 헬퍼 결정인자를 가진 Con-S V1, V2, V4 및 V5 펩타이드들은 Corixa's RC529 (Baldridge et al, J. Endotoxin Res. 8:453-458 (2002))와 같은 적합한 보체 (adjuvant)에서 제형화될 때 단독으로 또는 함께 사용될 수 있으며, 광범위하게 교차 반응하는 중화 항체를 비-클레이드 B 분리물로 유도하는데 사용될 수 있다. 표 4에서 각 서열은 순서대로 서열번호 309 내지 318을 나타낸다.

[표 4]

1)	GTH Con-S V1 132-150	YKRWILGLNKIVRMYTNVNVNTNTNTEEKGEIKN
2)	GTH Con-S V2 157-189	YKRWILGLNKIVRMYTEIRDKKQKVYALFYRLDVVPIDNNNNSSNYR
3)	GTH Con-S V3 294-315	YKRWILGLNKIVRMYTRPNNNTRKSIHGPGQAFYAT
4)	GTH Con-S V4 381-408	YKRWILGLNKIVRMYNTSGLFNSTWIGNGTKNNNTNDITLP
5)	GTH Con-S V5 447-466	YKRWILGLNKIVRMYRDGGNNNTNTEIFRPGGGD
6)	GTH Con-6 V1 132-150	YKRWILGLNKIVRMYNVRNVSSNGTETDNEEIKN
7)	GTH Con-6 V2 157-196	YKRWILGLNKIVRMYTELDDKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYR
8)	GTH-Con6 V3 301-322	YKRWILGLNKIVRMYTRPNNNTRKSIHGPGQAFYAT
9)	GTH Con-6 V4 388-418	YKRWILGLNKIVRMYNTSGLFNSTWFMNGTYMFMNGTKDNETITLP
10)	GTH Con 6 V5 457-477	YKRWILGLNKIVRMYRDGGNNNSKNKTETFRPGGGD

본 발명은 본원에 구체적으로 개시된 서열의 부분 및 변이체를 포함하는 것이 바람직할 것이다. 예를 들어, 코돈 최적화된 컨센서스 인코딩 서열의 형태(forms)는, 절단되거나 비-절단된 gp120/41과 함께 gp140CF, gp140 CFI, gp120 또는 gp160 형태로서 컨스트럭트될 수 있다. 예를 들어, 그리고 컨센서스 및 원형 외피 서열에 관해서는, 본 발명은 V3이 전혀 없는 외피 서열을 포함한다. 선택적으로, V3 서열은 바람직한 서열, 예컨대 미합중국 특허 제10/431,596호 및 미합중국 우선특허 제60/471,327호에 개시된 바와 같은 서열로부터 선택될 수 있다. 게다가, 반응의 범위에 대한 최적의 면역원은 그룹 M 컨센서스 gag, pol, nef 및 env 인코딩 서열의 혼합을 포함할 수 있으며, 뿐만 아니라 gag, pol, nef 및 env HIV 유전자에 대한 서브타입 컨센서스 또는 원형 인코딩 서열의 혼합을 포함할 수 있다. 바이러스 군주에 있는 부위적 차이를 이용해 다루기 위해, 유효한 혼합물은 컨센서스/원형 및 야생형 인코딩 서열의 혼합을 포함할 수 있다.

본 발명의 컨센서스 또는 원형 외피는, 정상적으로는 HIV 비리온의 표면에 일시적으로만 노출되거나 거의 노출되지 않는 중화 에피토프의 중간 형태 (intermediate conformation)를 노출시키기 위해 "활성화" 될 수 있다. 상기 면역원은 B 림프구에 대해 자극(presentation)하기 위한 이용가능한 특이 에피토프를 만드는 컨센서스 또는 원형 외피의 "동결(frozen)" 유도된 형태가 될 수 있다. 이에피토프 자극의 결과는 광범위하게 HIV를 중화하는 항체의 생산이다. (주의점이 WO 02/024149 및 본원에 개시된 활성화된/유발된 외피에 개시되어 있다)

융합 중간체 면역원의 개념은, gp41 HR-2 부위 펩타이드인 DP178이 gp41의 꼬이지 않은 형태를 캡처할 수 있고 (Furata et al, Nature Struct. Biol. 5:276 (1998)), 포르말린-고정된 HIV-감염 세포가 광범위하게 중화 항체를 생산할 수 있다는 (LaCasse et al, Science 283:357 (1997)) 결과와 일치한다. 최근 꼬인-코일 (coiled-coil) 부위에 대한 모노클로날 항체는, 꼬인-코일 gp41 구조의 HR1 및 HR2 부위에 있는 gp41의 형태적 결정인자에 결합하였으나 HIV를 중화시키지는 않았다 (Jiang et al, J. Virol. 10213 (1998)). 그러나, 후자의 연구에서는 꼬인-코일 부위가, 정확한 항체가 생성되는 경우에 결합하는 항체에 대해 이용가능한 것으로 증명되었다.

본 발명의 일 목적의 면역원은, 가용성 형태이거나 예를 들어 세포 소포에 또는 트랜스지질 이중층 외피를 포함하는 리포솜에 고정된 컨센서스 또는 원형 외피를 포함한다. 훨씬 더 천연의 외피를 만들기 위해, gp140 또는 gp160 컨센서스 또는 원형 서열이, 천연의 3중 외피를 형성하기 위해 지질 이중층에 배열될 수 있다. 선택적으로는 알드리티오 (aldrithio) 1-2 불활성화된 HIV-1 비리온에 있는 유발된 gp160이 면역원으로서 사용될 수 있

다. 또한, gp160은 gp160 또는 gp140 (gp140은 막투과 부위를 가지고 가능하게는 다른 gp41 부위가 결실된 gp160임) 중 어느 재조합 단백질로서도 존재할 수 있다. gp160 또는 gp140에 결합하는 것은, 재조합 CCR5 또는 CXCR4 공-수용체 단백질 (또는 이들의 세포외 도메인 펩타이드 또는 단백질 단편), 또는 gp120 및/또는 가용성 CD4 상의 CXCR4 또는 CCR5 결합 위치에 결합하는 항체 또는 CD4의 결합 작용을 모방하는 항체 또는 다른 리간드가 될 수 있다. 선택적으로는 CD4, CCR5 (또는 CXCR4) 또는 가용성 CD4를 포함하는 소포 또는 리포솜, 및 CCR5 또는 CXCR4 gp120 결합 위치의 반사 펩타이드이다. 선택적으로, 최적 CCR5 펩타이드 리간드는, 특정 티로신이 황산화된 CCR5의 N-말단 펩타이드가 될 수 있다 (Bormier et al, Proc. Natl. Acad. Sci. USA 97:5762 (2001)). 유발된 면역원은 막에 결합할 필요가 없을 것이지만, 용액 내에 존재하고 용액에서 유발될 수 있을 것이다. 선택적으로, 가용성 CD4 (sCD4)는, CD4 펩타이드 미메토프 (mimetopes)에 의해 유발된 외피 (gp140 또는 gp160)에 의해 치환될 수 있다 (Vitra et al, Proc. Natl. Acad. Sci. USA 96:1301 (1999)). gp160 또는 gp140을, 세포 융합을 유도하는 gp160의 구조와 관련하여 변화하도록 "유발시키는", 다른 HIV 공-수용체 분자도 사용될 수 있다. 가용성 CD4 (sCD4)를 이용한 가용성 HIV gp140 제1 분리물 HIV 89.6 외피의 라이게이션은 gp41에서 형태적 변화를 유발하였다.

[0032] 일 구현예에 따르면, 본 발명은 CD4 결합 위치가 차단된 CD4-연결된 단백질과는 달리, CCR5 결합 위치가 노출된 수용체 (CD4)-연결된 컨센서스 또는 원형 외피의 특성을 가진 면역원에 관한 것이며, 이 면역원은 CD4 결합 위치가 노출되어 있다 (오픈). 더불어, 이 면역원은 숙주 CD4가 전혀 없을 수 있으며, 숙주에 투여되는 동안 잠재적으로 위대한 항-CD4 항체의 생산을 피한다.

[0033] 면역원은 A32 모노클로날 항체 (mab)에 의해 인식되는 gp120 상의 위치에 결합하는 리간드와 연결된 컨센서스 또는 원형 외피를 포함할 수 있다 (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)). 한 A32 mab는 CD4를 모방하고, gp120에 결합할 때는, CCR5 결합 위치를 상향조절 (노출)하는 것으로 밝혀졌다 (Wyatt et al, J. Virol. 69:5723 (1995)). 또한, 이러한 리간드를 이용한 gp120의 라이게이션은 CD4 결합 위치를 상향조절하지만 gp120에 결합하는 CD4를 제어하지는 못한다. 선택적으로, 이러한 리간드들은 절단된 gp120, 비-절단된 gp140 및 절단된 gp41에 결합한 gp41의 HR-2 결합 위치도 상향조절하며, 이에 의해 이들 단백질들-이들 각각은 항-HIV 중화 항체에 대한 잠재적 표적임-상의 HR-2 결합 위치를 노출시킨다.

[0034] 본 구현예의 특정 목적에서, 면역원은 완전한 A32 mab, A32 mab의 Fab2 단편, 또는 A32 mab의 Fab 단편 중의 하나로 라이게이션된 가용성 HIV 컨센서스 또는 원형 gp120 외피를 포함하며, 그 결과, 컨센서스 또는 원형 외피 상의 CD4 결합 위치, CCR5 결합 위치 및 HR-2 결합 위치가 노출되고/상향조절된다. 면역원은 A32 mab (또는 이의 단편)가 결합된 컨센서스 또는 원형 외피를 포함하거나, 0.3% 포르말데히드 또는 DTSSP (Pierce Chemical Company)와 같은 헤테로이중작용 교차-링커와 같은 교차-링커로 교차-링크되고 결합된 A32 mab (또는 이의 단편)를 가진 컨센서스 또는 원형 외피를 포함할 수 있다. 또한, 면역원은 비-절단된 컨센서스 또는 원형 gp140 또는 비-절단된 gp140, 절단된 gp41 및 절단된 gp120의 혼합물을 포함할 수 있다. 컨센서스 또는 원형 gp140 및/또는 gp120에 결합되거나, gp41에 비-공유결합으로 결합한 gp120에 결합된 A32 mab (또는 이의 단편)는, gp41, gp120 및 비-절단된 gp140에 있는 HR-2 결합 위치의 상향조절 (노출)을 유발한다. 또한, gp120 또는 gp140에 대한 A32 mab (또는 이의 단편)의 결합은 CD4 결합 위치 및 CCR5 결합 위치의 상향조절을 유발한다. gp120이 복합체 (complex)를 포함함에 따라, 비-절단된 gp140 및 A32 mab (또는 이의 단편)를 포함하는 복합체는 0.3% 포르말데히드 또는 DTSSP와 같은 교차-링커로 교차-링크되거나 교차-링크되지 않는 면역원으로서 사용될 수 있다. 일 구현예에서, 본 발명은 Fab 단편 또는 전체 A32 mab에 결합되고 교차 링크되며, 선택적으로는 HR-2 결합 단백질에 결합되고 교차-링크된, 가용성 비-절단된 컨센서스 또는 원형 gp140을 포함하는 면역원에 관한 것이다.

[0035] gp120 상의 A32 mab 결합 위치에 결합하는 리간드로 유발된 컨센서스 또는 원형 외피 단백질은, mab 17b에 의해 인식되는 CCR5 결합 위치와 같은 A32 mab 결합 위치와는 다른 위치에 결합하는 리간드에 의해 유발되는, 제2 외피를 포함하는 최소 1개의 제2 면역원과 혼합하여 투여될 수 있다. 17b mab (Kwong et al, Nature 393:648 (1998) available from the AIDS Reference Repository, NIAID, NIH)는 gp120에 대한 sCD4의 결합을 증가시킨다. 예를 들어, 이 제2 면역원은 (이는 단독으로 또는 상기에 기술된 것과는 다른 유발된 면역원과 조합하여 사용될 수 있음), 전체 17b mab, 17b mab의 Fab2 단편 또는 17b mab의 Fab 단편 중의 하나와 연결되는 가용성 HIV 컨센서스 또는 원형 외피를 포함할 수 있다. 다른 항체 (또는 이의 단편)를 포함하는 다른 CCR5 리간드가, 17b mab의 장소(lieu)에 사용될 수 있는 노출되는 CD4 결합 위치를 유발하는 것이 바람직할 것이다. 이러한 추

가의 면역원은 결합된 17b mab 또는 이의 단편 (또는 상기에 표시된 것과 같은 다른 CCR5 리간드)을 가진 gp120을 포함하거나, 0.3% 포르말데히드 또는 DTSSP (Pierce Chemical Company)와 같은 헤테로이중기능성 교차-링커와 같은 시약으로 교차-링크되고 결합된 17b mab 또는 이의 단편 (또는 상기에 표시된 것과 같은 다른 CCR5 리간드)을 가진 gp120을 포함할 수 있다. 선택적으로, 이 면역원은 추가로 비-절단된 gp140를 단독으로 포함하거나, 절단된 gp41 및 절단된 gp120의 혼합물을 포함할 수 있다. 이러한 혼합물에 있는 gp140 및/또는 gp120에 결합한 Mab 17b 또는 이의 단편 (또는 상기에 표시된 것과 같은 다른 CCR5 리간드)는 CD4 결합 부위의 노출을 유발한다. 17b mab, 또는 이의 단편 (또는 상기에 표시된 것과 같은 다른 CCR5 리간드), gp140 복합체는 0.3% 포르말데히드 또는 DTSSP와 같은 제제로 교차-링크되거나 비교차-링크되어 존재할 수 있다.

[0036] T649Q26L 및 DP178과 같은 가용성 HR-2 펩타이드는 상술한 복합체에 첨가됨으로써, 비-절단된 컨센서스 gp140 분자 뿐만 아니라 컨센서스 gp120 및 gp41 상의 에피토프를 안정화시킬 수 있으며, 복합체로 교차-링크되거나 비교차-링크되어 투여될 수 있다.

[0037] 모노클로날 항체 (mabs) 시리즈는, 상기에 기재한 17b mab에 추가로, gp120 상의 CD4 결합 위치에 결합하는 mab IgG1b12 (Roben et al, J. Virol. 68:482 (1994), Mo et al, J. Virol. 71:6869 (1997), gp120 상의 형태학적 결정인자에 결합하는 mab 2G12 (Trkola et al, J. Virol. 70:1100 (1996)) 및 gp41의 막 근위 부위에 결합하는 mab 2F5 (Muster et al, J. Virol. 68:4031 (1994))를 포함하는 많은 HIV 제1 분리물을 중화시키는 것으로 밝혀졌다.

[0038] 상기에 표시한 바와 같이, 본 발명에 따른 퓨소제닉 (fusogenic) 에피토프를 "동결"하는데 다양한 시도가 사용될 수 있다. 예를 들어, "동결"은 꼬인 코일 부위의 부분을 나타내는 DP-178 또는 T649Q26L 펩타이드의 첨가에 의해 영향을 받을 수 있으며, CD4-유발된 컨센서스 또는 원형 외피에 추가될 때, 융합을 방해한다 (Rimsky et al, J. Virol. 72:986-993 (1998)). HR-2 펩타이드 결합된 컨센서스 또는 원형 gp120, gp140 또는 gp160은 면역원으로서 사용되거나 DTSSP 또는 DSP (Pierce Co.)와 같은 제제, 포르말데히드 또는 유사한 효과를 가진 다른 교차링크 제제에 의해 교차링크될 수 있다.

[0039] 또한, "동결"은 0.1% 내지 3%의 포르말데히드 또는 파라포르말데히드, 단백질 교차-링크 제제 모두를 복합체에 첨가시킴으로써 영향을 받을 수 있으며, 이에 의해 CD4, CCR5 또는 CXCR4, HR-2 펩타이드 gp160 복합체를 안정화시키거나, "유발된" gp41 분자를 안정화시키거나, 또는 이 모두를 할 수 있다 (LaCasse et al, Science 283:357-362 (1999)).

[0040] 더불어, 컨센서스 또는 원형 gp41 또는 gp120 융합 중간체의 "동결"은, 아미노 그룹과 반응하는 2개의 NHS 에스테르를 이용하는 DSP (dithiobis[succinimidylpropionate]) (Pierce Co. Rockford, ILL., No. 22585ZZ) 또는 수용성 DTSSP (Pierce Co.)와 같은 헤테로이중기능성 제제의 첨가에 의해 영향을 받을 수 있으며, 이에 의해 CD4, CCR5 또는 CXCR4, HR-2 펩타이드 gp160 복합체를 교차 링크시키고 안정화하거나 또는 "유발된" gp41 분자를 안정화하거나 또는 이 작용을 모두 한다.

[0041] 면역화되거나 백신화된 동물 및 인간에서의 T 세포 면역 반응의 분석에서는, 외피 단백질이 비록 중화 항체를 유도하는 유일한 유전자더라도, 정상적으로는 T 세포 면역 반응에 대한 주요한 표적이 아님을 보여준다. HIV-1 Gag, Pol 및 Nef 단백질은 잠재적 T 세포 면역 반응을 유도한다. 따라서, 본 발명은 체액성 및 세포성 면역 반응을 모두 유도할 수 있는 컨센서스 또는 원형 면역원의 반복을 포함한다. 컨센서스 또는 원형 서열의 서브단위는 T 또는 B 세포 면역원으로서 사용될 수 있다 (실시예 6 및 7, 및 이에 참조된 도면 및 도 63-127 참조).

[0042] 본 발명의 면역원은 당업계에 공지된 기술을 이용하여, 약제학적으로 허용가능한 담체 및/또는 보체 (알룸 (alum)과 같은)를 사용하여 제형화될 수 있다. 본 발명의 면역원의 적합한 투여 경로는 전신 (예, 근육내 또는 피하)을 포함한다. 면역 반응이 점막 (mucosal) 면역 시스템 (예, 비강내)에서 유발되는 경우 선택적인 경로가 이용될 수 있다.

[0043] 본 발명의 면역원은 당업자에게 공지된 방법을 이용해 화학적으로 합성하고 정제될 수 있다. 또한, 면역원은 공지된 재조합 DNA 기술에 의해 합성될 수 있다. 본 발명의 면역원을 인코딩하는 핵산은 예를 들어 인코딩 서열이 네이키드(naked) DNA로서 투여되는 DNA 백신의 성분으로서 사용될 수 있거나, 예를 들어 면역원을 인코딩하는 미니유전자가 바이러스 벡터에 존재할 수 있다. 상기 인코딩 서열은 예를 들어 복제 또는 비-복제 아데노 바이러스 벡터, 아데노-연관 바이러스 벡터, 감쇠된 마이코박테리움 투베르쿨로시스 (mycobacterium tuberculosis) 벡터, 바실러스 칼메트 귀린 (Bacillus Calmette Guerin, BCG) 벡터, 우두 또는 변형된 우두 안

카라 (Modified Vaccinia Ankara, MVA) 벡터, 다른 폭스 바이러스 벡터, 재조합 폴리오 (polio) 및 다른 장관 (enteric) 바이러스 벡터, 살모넬라 중 박테리아 벡터, 시겔라 (Shigella) 중 박테리아 벡터, 베네주엘리안 에 퀴 엔세팔리티스 바이러스 (Venezuelean Equine Encephalitis virus, VEE) 벡터, 쉴리키 포레스트 (Semliki Forest) 바이러스 벡터, 또는 담배 모자이크 바이러스 벡터에 존재할 수 있다. 인코딩 서열은 예를 들어 CMV 프로모터와 같은 활성 프로모터를 가진 DNA 플라스미드로서 발현될 수 있다. 또한, 다른 생 (live) 벡터가 본 발명의 서열을 발현하는데 사용될 수 있다. 본 발명의 면역원의 발현은, 면역원을 인코딩하는 핵산, 바람직하게는 인간 세포에서 발현을 최적화하는 코돈 및 프로모터를 이용해 이들 세포에 도입시킴으로써, 환자 자신의 세포에서 유도될 수 있다. DNA 백신을 제조하고 이용하는 방법의 예가 미합중국 특허 제5,580,859호, 제 5,589,466호 및 제5,703,055호에 개시되어 있다.

[0044] 본 발명의 조성물은 약제학적으로 허용가능한 전달 시스템에 있는 본원 발명의 면역원 또는 이를 인코딩하는 핵산 서열을 면역학적 유효량으로 포함한다. 상기 조성물은 면역결핍 바이러스 감염의 예방 및/또는 치료에 사용될 수 있다. 본 발명의 조성물은 보체, 유화제, 약제학적으로 허용가능한 담체 또는 백신 조성물에 통상적으로 사용되는 다른 성분을 이용해 제형화될 수 있다. 최적의 제형은 당업자에 의해 용이하게 디자인될 수 있으며, 즉시 분비 및/또는 지연된 분비를 위해, 그리고 전신성 면역의 유도 및/또는 국소적 점막 면역의 유도 (예, 제형은 비강내 투여용으로 디자인될 수 있다)를 위한 제형을 포함할 수 있다. 본 발명의 조성물은 피하, 비강내, 경구, 근육내, 또는 다른 비경구 또는 소화관 경로를 포함하는 모든 통상적인 경로에 의해 투여될 수 있다. 면역원은 단일 투여량 또는 복합 투여량으로 투여될 수 있다. 최적의 면역화 스케줄은 당업자에 의해 용이하게 결정될 수 있으며, 환자, 조성물 및 예상되는 효과에 따라 달라질 수 있다.

[0045] 본 발명은 본 발명의 면역원 및/또는 이를 인코딩하는 핵산 및/또는 상기에 표시된 벡터에 있는 미니유전자로서 발현된 면역원 모두의 직접 사용을 고려한다. 예를 들어, 면역원을 인코딩하는 미니유전자는 프라임 (prime) 및/또는 추가접종 (boost)으로 사용될 수 있다.

[0046] 본 발명은 본원에 개시된 어떠한 모든 아미노산 서열 및 적용가능한 것으로서, 이의 CF 및 CFI 형태 뿐만 아니라 이를 인코딩하는 핵산 서열 (및 이러한 인코딩 서열에 상보적인 핵산)도 포함한다.

[0047] 본 발명의 특정 요지는 하기의 비-제한적인 실시예에 훨씬 자세히 개시될 수 있다.

[0048] 실시예 1

[0049] 인공 HIV-1 그룹 M 컨센서스 외피

[0050] 구체적 실험

[0051] *재조합 우두 바이러스 (VV)에서의 CON6 gp120 및 gp140 단백질의 발현.*

[0052] HIV-1 CON6 외피 단백질의 분비 형을 발현하고 정제하기 위해, 중지 코돈을 gp120 절단 위치 (REKR) (서열번호 319)의 뒤 및 막투과 도메인 (YIKIFIMIVGGLIGLRIVFAVLSIVN) (서열번호 320)의 앞에 각각 도입시킴으로써 CON6 gp120 및 gp140CF 플라스미드를 컨스트럭트하였다. gp120/gp41 절단 위치 및 gp41의 융합 도메인을 gp140CF 단백질에서 결실시켰다. CON6 gp120 및 gp140CF DNA 컨스트럭트 모두를 pSC65 벡터 (Bernard Moss, NIH, Bethesda, MD로부터)의 SalI 및 KpnI 제한효소 위치에 클로닝하였다. 이 벡터는 p7.5 프로모터에 의해 조절되는 lacZ 유전자를 포함한다. 등을 맞댄 (back-to-back) P E/L 프로모터를 CON6 env 유전자를 발현하는데 사용하였다. BSC-1 세포를 6웰 플레이트의 각 웰에 2×10^5 농도로 접종하고, 0.1 pfu/세포의 MOI로 야생형 우두 바이러스 (WR)를 감염시켰으며, 감염 2시간 후에, CON6 env 유전자를 포함하는 pSC65-유도된 플라스미드를 VV-감염된 세포에 형질감염시키고, 공지된 바에 따라 재조합 (r) VV를 선별하였다 (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)). CON6 env 유전자를 포함하는 재조합 VV를 PCR 및 서열 분석으로 확인하였다. CON6 외피 단백질의 발현은 SDS-PAGE 및 웨스턴 블랏 분석법으로 확인하였다. 재조합 CON6 gp120 및 gp140CF는 아가로즈 갈란투스 니발리스 (galanthus Nivalis) 렉틴 비드 (Vector Labs, Burlingame, CA)를 이용하여 정제하였으며, 사용할 때까지 -70°C에 보관하였다. 재조합 VV 발현 JRFL (vCB-28) 또는 96ZM651 (vT241R) gp160은 NIH AIDS Research and Reference Reagent Program (Bethesda, MD)로부터 얻었다.

- [0053] 모노클로날 항체 및 gp120 야생형 외피.
- [0054] gp120 (A32), gp120 V3 루프 (F39F) 및 CCR5 결합 위치 (17b) 상의 형태적 결정인자에 대한 인간 mabs는 제임스 로빈슨 (Tulane Medical School, New Orleans, LA) (Wyatt et al, Nature 393:705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995))으로부터 받았다. Mabs 2F5, 447, b12, 2G12 및 가용성 CD4는 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 10:1651-1658 (1994), Trkola et al, J. Virol. 70:1100-1108 (1996))으로부터 얻었다. T8은 gp120 C1 부위 (P. Earl, NIH, Bethesda, MD로부터 받음)에 위치하는 쥐의 mab이다. BaL (서브타입 B), 96ZM651 (서브타입 C) 및 93TH975 (서브타입 E) gp120s는 QBI, Inc. 및 AIDS, NIH 부서로부터 제공받았다. 92U037 (서브타입 A) 및 93BR029 (서브타입 F) gp140 (절단 및 비절단됨)을 발현하는 CHO 세포주는 NICBS, England로부터 획득하였다.
- [0055] 표면 플라스몬 공명 바이오센서 (Surface Plasmon Resonance Biosensor, SPR) 측정 및 ELISA.
- [0056] SPR 바이오센서 측정은 BIAcore 3000 기기 (BIAcore Inc., Uppsala, Sweden) 상에서 측정하였으며, 데이터 분석은 BIAevaluation 3.0 소프트웨어 (BIAcore Inc, Uppsala, Sweden)를 이용하여 수행하였다. 10 mM의 Na-아세테이트 버퍼 (pH 4.5)에 있는 항-gp120 mabs (T8, A32, 17b, 2G12) 또는 sCD4를, 단백질을 고정하기 위한 표준 아민 커플링 프로토콜을 사용하여 CM5 센서 칩에 직접 고정시켰다. FPLC 정제된 CON6 gp120 모노머 또는 gp140CF 올리고머 재조합 단백질을 100 및 300 $\mu\text{g}/\text{ml}$ 의 농도로 각각 CM5 센서 칩 위에 흘려보냈다. 검은색 인라인(in-line) 참조 표면 (아민 커플링에 대해 활성화 및 비-활성화됨) 또는 비-결합 mab 대조군은 비-특이적 또는 벌크 반응을 빼는데 사용하였다. 가용성 89.6 gp120 및 부적절한(irrelevant) IgG는, 각각 양성 대조군 및 음성 대조군으로서 사용하였으며, CON6 Env 단백질을 주입하기 전에 각 mab 표면을 활성화시키기 위해 사용하였다. CON6 외피 단백질의 결합은 10-30 $\mu\text{l}/\text{ml}$ 의 속도로 연속적으로 흘려보내는 PBS (150 mM NaCl, 0.005% 계면활성제 P20, pH 7.4)를 이용해 25°C에서 실시간으로 모니터링하였다. 결합한 단백질을 제거한 뒤, 재생 용액 (10 mM 글리신-HCl, pH 2.9)을 5-10 μl 펄스로 단독 또는 복수로 펄스시키는 각 결합 사이클을 수행함으로써 센서 표면을 재생시켰다. CON6 gp120 및 gp140CF 단백질에 대한 다양한 mabs의 반응성을 확인하기 위해, 개시된 바에 따라 (Haynes et al, AIDS Res. Hum. Retroviruses 11:211-221 (1995)) ELISA를 수행하였다. rgp120 또는 gp140 단백질에 대한 인간 mab 결합력을 분석하기 위해, 말단-지점 역가를 가장 높은 mab 역가 (20 $\mu\text{g}/\text{ml}$ 에 서부터 시작)로 정했으며, 여기서 mab 결합한 CON6 gp120 및 gp140CF Env 단백질은 백그라운드 대조군 (비-결합 인간 mab)보다 ≥ 3 배이었다.
- [0057] 감염도 및 공-수용체 사용 분석.
- [0058] HIV-1/SG3 Δenv 및 CON6 또는 대조군 env 플라스미드를 인간 293T 세포로 동시 형질감염시켰다. 유사타입 바이러스를 획득하고, 여과한 뒤 p24 농도를 정량하였다 (Dupont/NEN Life Sciences, Boston, MA). 각 유사 비리온에 대해 동량의 p24 (5 ng)를 JC53-BL 세포를 감염시키는데 사용하고 감염도를 측정하였다 (Derdeyn et al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). JC53-BL 세포는 CD4, CCR5 및 CXCR4 수용체를 발현하며, HIV-1 긴 말단 반복 (long terminal repeat, LTR)의 전사 조절 하에서 안정하게 삽입된 β -갈락토시다아제 (β -gal) 유전자를 포함한다. 이들 세포들은, β -gal 발현에 대해 염색하고 p24 유사비리온 마이크로그램 당 푸른색 세포 (감염 단위)의 수 (IU/ μg p24)를 계수함으로써, 유사비리온 농축물 (stocks)의 감염 역가를 정량하는데 사용될 수 있다 (Derdeyn et al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). CON6 env 유전자의 공-수용체 사용을 확인하기 위해, JC53BL 세포에 1.2 μM AMD3100 및 4 μM TAK-799를 1시간 동안 37°C에서 처리한 뒤, 동량의 각 Env 유사타입 바이러스 p24 (5 ng)로 감염시켰다. 차단 효율은 저해 실험으로부터 얻은 감염 단위를 저해제가 없는 대조군 배양과 비교한 퍼센트로서 표시하였다. 대조군 (차단제가 없는)의 감염도를 임의로 100%로 정하였다.
- [0059] 면역화 (immunization).
- [0060] 모든 동물들은 듀크 대학 동물 사용 및 사육 위원회에 의해 승인된 동물 사용 프로토콜과 AALAC 지침하에서 듀

크 대학 동물 사육소에서 사육하였다. 판매처 (Sigma Chemical Co., St. Louis, MO)에서 제공되는 프로토콜에 기초하여, 재조합 CON6 gp120 및 gp140CF 당단백질을 RIBI-CWS 보체와 함께 안정한 에멀전으로 제형화하였다. 항-외피 항체를 유도시키기 위해, 4마리의 무작위 교배계 기니아 피그 (Harlan Sprague, Inc., Chicago, IL) 각각에 100 μ g의 정제된 CON6 gp120 또는 gp140CF를 3주마다 피하에 주입하였다 (전체 5회의 면역화). 혈청 샘플을 열-불활성화시켰으며 (56°C, 1시간), 사용할 때까지 -20°C에 보관하였다.

[0061] 항-외피 T 세포 반응을 유도하기 위해, 6-8 주령의 암컷 BALB/c 마우스 (Frederick Cancer Research and Developmental Center, NCI, Frederick, MD)에 50 μ g의 플라스미드 DNA를 무릎(quadriceps)에 있는 근육내 (i. m.)로 3주의 간격으로 3회로 면역화시켰다. 최종 DNA 면역화로부터 3주 후에, Env 단백질을 발현하는 rVV 107 PFU를 마우스에 추가접종 (boost)시켰다. 추가접종 2주 후에, 모든 마우스들을 안락사시킨 뒤 비장세포 분리를 위해 비장을 적출하였다.

[0062] 중화 분석 (Neutralization assay).

[0063] 중화 분석은, Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000)에 개시된 바와 같은 MT-2 분석법, HIV-1 제1 분리물의 패널을 이용한 5.25.GFP.Luc.M7 세포에서의 루시페라아제-기초 다중 복제 사이클 HIV-1 감염 분석법 (Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000), Bures et al, J. Virol. 76:2233-2244 (2002)), 또는 불활성화된 HIV-1 비리온을 이용한 융합체 (아무것도 없는 것으로부터 융합) 저해 분석법 (Rossio et al, J. Virol. 72:7992-8001 (1998)) 중의 하나를 이용해 수행하였다. 루시페라아제-기초 분석법에서, 중화 항체는, Nathaniel R. Landau, Salk Institute, La Jolla, CA (Brandt et al, J. Biol. Chem. 277:17291-17299 (2002))에 의해 제공되는 5.25.EGFP.Luc.M7 세포에서의 루시페라아제 활성의 감소 작용으로서 측정하였다. 조직 배양 감염 용량 50 (TCID₅₀)의 무-세포 바이러스 500개를, 96웰 평편한-바닥 배양 플레이트에 3개씩 있는 표시된 혈청 희석액 150 μ l와 반응시켰다. 5.25.EGFP.Luc.M7 세포를 DEAE 텍스트란 (10 μ g/ml)을 포함하는 배지에 5×10^5 /ml 농도로 현탁시켰다. 세포 (100 μ l)를 첨가하여, 대조군 웰 (혈청 샘플을 테스트하지 않음)에 있는 세포의 10%가, 형광 현미경에 의한 GFP 발현에 대해 양성이 될때까지 첨가하였다. 이때, 배지 부피의 반을 제거함으로써 세포를 2배로 농축시켰다. Wallac 1420 멀티라벨 계수기 (PerkinElmer Life Sciences, Boston, MA) 상에 있는 Bright-Glo™ 기질 (Promega, Madison, MI)을 이용하여 루시페라아제 활성을 측정하기 위해, 50 μ l의 현탁 세포를 96웰 흰색 고체 플레이트 (Costar, Cambridge, MA)에 옮겼다. MT-2 및 루시페라아제 분석법에서의 중화 역가는 바이러스 감염의 $\geq 50\%$ 이 억제되는 것이었다. 역가가 1:20 이상 (즉, $>1:30$)인 값만이 유의적으로 양성이라고 간주하였다. 융합체 저해 "아무것도 없는 것으로부터의 융합" 분석은, SupT1 세포로 첨가된 HIV-1 서브타입 B 균주 ADA 및 AD8 (Larry Arthur and Jeffrey Lifson, Frederick Research Cancer Facility, Frederick, AD로부터 받음)의 HIV-1 알드리티올-2 (aldrithiol-2, AT-2) 불활성화된 비리온을 이용하였으며, 융합체 저해 역가는 전-채혈된 (prebleed) 혈청과 비교해 융합체의 $\geq 90\%$ 이 저해되는 역가로 결정하였다.

[0064] 효소 링크된 면역 스팟 (Enzyme linked immune spot, ELISPOT) 분석.

[0065] 면역화된 마우스 각각의 비장세포의 단일-세포 현탁액은, 잘게 다지고 70 μ m 나일론 세포 여과기 (BD Labware, Franklin Lakes, NJ)를 통과시킴으로써 준비하였다. CON6 gp140의 중복 Env 펩타이드들 (159개 펩타이드, 11개에 의해 중복되는 15머(mers))은 Boston Bioscience, Inc (Royal Oak, MI)로부터 구입하였다. MN gp140의 중복 Env 펩타이드들 (서브타입 B; 170개 펩타이드, 11개에 의해 중복되는 15머) 및 Chn19 gp140의 중복 Env 펩타이드들 (서브타입 C; 69개 펩타이드, 10개에 의해 중복되는 20머)은 NIH AIDS Research and Reference Reagent Program (Bethesda, MD)로부터 수득하였다. 각 마우스의 비장세포 (5 생쥐/그룹)를, CON6, 서브타입 B 및 서브타입 C Env 단백질의 중복 Env 펩타이드 풀(pool)을 사용하여 인 비트로에서 자극시켰다. 96웰 PVDF 플레이트 (MultiScreen-IP, Millipore, Billerica, MA)를 항-IFN- γ mab (5 μ g/ml, AN18; Mabtech, Stockholm, Sweden)로 코팅하였다. HEPES 완충된 완전 RPMI 배지를 이용해 37°C에서 2시간 동안 플레이트를 블로킹한 후, 모아진 중복 외피 펩타이드 50 μ l (13 CON6 및 MN 풀, 각 풀에 13-14 펩타이드; 9 Chn19 풀, 각 풀에 7-8 펩타이드)를 각각 최종 농도가 5 μ g/ml이 되도록 플레이트에 첨가하였다. 그리고 난 후, 1.0×10^7 /ml 농도의 비장 세포 50 μ l를 2개씩 웰에 첨가한 뒤 16시간 동안 37°C에서 5% CO₂에서 반응시켰다. 상기 플레이트를 1:1000배

로 회색한 스트렙타비딘 알칼라인 포스파타아제 (streptavidin alkaline phosphatase) (Mabtech, Stockholm, Sweden) 100 μ l로 반응시켰으며, BCIP/NBT (플러스) 알칼라인 포스파타아제 기질 (Moss, Pasadena, MD) 100 μ l를 사용해 보라색 스팟을 현상하였다. 면역스팟 계수 시스템 (CTL Analyzers, Cleveland, OH)을 이용해 스팟 형성 세포 (spot forming cells, SFC)를 측정하였다. 각 외피 캡타이드 풀에 대한 전체 반응도는 106 비장세포 당 SFCs로서 표시하였다.

[0066] 결과

[0067] *CON6 외피 유전자 디자인, 컨스트럭션 및 발현.*

[0068] Los Alamos HIV 서열 데이터베이스에 있는 서열의 각 HIV-1 서브타입에 대한 env 유전자의 컨센서스 서열을 제조하고, 대량으로 서열분석된 서브타입을 피하기 위해 모든 서브타입 컨센서스의 컨센서스 서열을 제조함으로써, 인공 그룹 M 컨센서스 env 유전자 (CON6)를 컨스트럭트하였다 (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). 그리고 난 뒤, CRF08_BC 재조합 균주 (98CN006)의 5개의 매우 다양한 부위 (V1, V2, V4, V5 및gp41의 세포질 도메인에 있는 부위)를 CON6 서열에 있는 결실된 부위를 채우는데 사용하였다. CON6 V3 부위는 그룹 M 컨센서스이다 (도 1a). 높은 레벨의 발현을 위해, CON6 env 유전자의 코돈을 고발현된 인간 유전자에 대한 코돈 사용에 기초하여 최적화하였다 (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)). (도 1d 참조) 코돈 최적화된 CON6 env 유전자를 컨스트럭트하여 pcDNA3.1 DNA의 EcoRI 및 BamHI 위치에 서브클로닝하였다 (Gao et al, AIDS Res. Hum. Retroviruses, 19:817-823 (2003)). 높은 레벨의 단백질 발현은, 293T 세포에 형질감염시키고 난 뒤 웨스턴-블랏 분석으로 확인하였다. 특성화하고 면역원으로서 사용하기 위한 재조합 CON6 Env 단백질을 얻기 위해, 분비된 gp120 및 미-절단된 gp140CF를 발현하도록 rVV를 제조하였다 (도 1b). 각 단백질에 대한 순도는 환원 조건하에서 코마시 블루 젤에서 확인했을 때 $\geq 90\%$ 이었다 (도 1c).

[0069] *CD4 결합 도메인 및 다른 야생형 HIV-1 에피토프는 CON6 단백질 상에 보존된다.*

[0070] CON6 단백질이 CD4에 결합하고 다른 야생형 HIV-1 에피토프를 발현하는지 확인하기 위해, CON6 gp120 및 gp140CF의 가용성 CD4에 결합하는 능력, 수개의 잘-특성화된 항-gp120 mabs에 결합하는 능력 및 CD4-유도된 형태학적 변화를 수반하는 능력을 분석하였다. 첫번째로, BIAcore CM5 센서 칩을 다른 sCD4 또는 mabs로 코팅하여 CON6 Env 단백질에 결합하는 능력을 모니터링하였다. 모노머 CON6 gp120 및 올리고머 gp140CF 모두는 sCD4에 효율적으로 결합하였으나, 항-gp120 mabs T8, 2G12 및 A32는, gp120의 CCR5 결합 위치에 있는 CD4 유도성 에피토프를 인식하는 mab 17b에 지속적으로 결합하지 못하였다 (도 2a 및 2b). sCD4 및 A32 모두는 야생형 gp120에 결합하고 난 뒤에 17b 결합 에피토프를 노출시킬 수 있다 (Wyatt et al, Nature 393: 705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). sCD4 또는 A32에 의해 17b 에피토프가 CON6 Envs 상에서 유도되는지 확인하기 위해, sCD4, A32 및 T8를 센서 칩 상에 코팅시키고, 그리고 난 후 CON6 gp120 또는 gp140CF를 캡쳐하였으며, mab 17b 결합 활성을 모니터링하였다. sCD4 또는 mab A32를 결합시키고 난 뒤에, CON6 gp120 및 gp140CF를 형태학적 변화를 수반하고 mab 17b에 결합하도록 유발하였다 (도 2c 및 2d). 반대로, mab T8을 결합시키고 난 뒤, 17b 에피토프는 노출되지 않았다 (도 2c 및 2d). 그 후, ELISA는, gp120 V3 루프 (447, F39F), CD4 결합 위치 (b12) 및 CON6 gp120 및 gp140CF의 gp41 중화 결정인자 (2F5)에 대한 인간 mabs의 패널의 반응도를 결정하기 위해 사용하였다 (도 2e). CON6 rgp120 및 rgp140CF 단백질 모두는 중화 V3 mabs 447 및 F39F, 그리고 잠재적 중화 CD4 결합 위치 mab b12에 잘 결합하였다. 또한, C-말단 gp41 에피토프에 결합함으로써 HIV-1 제1 분리물을 중화시키는 Mab 2F5도 CON6 gp140CF에 잘 결합하였다 (도 2e).

[0071] *CON6 env 유전자는 생물학적으로 기능적이며, 이의 공-수용체로서 CCR5를 이용한다.*

[0072] CON6 외피 유전자가 생물학적으로 기능을 하는지 여부를 확인하기 위해, 이를 env-결핍성 SG3 프로바이러스 클론과 함께 293T 세포에 동시-형질감염시켰다. 유사타입 바이러스를 회수한 뒤 JC53BL 세포를 감염시켰다. CON6 Env 유사비리온으로 감염된 JC53-BL 세포에서 푸른색 세포가 검출되었으며, 이는 CON6 Env 단백질이 생물학적으로 기능을 함을 제시해 준다 (도 3a). 그러나, 감염 역가는 YU2 또는 NL4-3 야생형 HIV-1 외피를 가지는 유사비리온보다 1-2 로그 (log)가 더 낮았다.

- [0073] CON6 env 유전자에 대한 공-수용체 사용은 다음에 확인하였다. CXCR4 저해제 AMD3100을 이용해 처리했을 때, NL4-3 Env-유사비리온의 감염도가 차단되었으나, YU2 또는 CON6 Env-유사비리온의 감염도는 저해되지 않았다 (도 3b). 반대로, CCR5 저해제 TAK-779로 처리했을 때에는, NL4-3 Env-유사비리온의 감염도는 영향을 받지 않았으나, YU2 또는 CON6 Env-유사비리온의 감염도는 저해되었다. 상기 저해제 모두를 처리했을 때에는, 모든 유사비리온의 감염도가 저해되었다. 종합하면, 이들 데이터는 CON6 외피가 표적 세포로 유입하기 위해 CCR5 공-수용체를 이용함을 보여준다.
- [0074] *CON6 gp120의 다른 서브타입 혈청과의 반응.*
- [0075] 다중 서브타입 선형 에피토프가 CON6 gp120 상에 보존되는지 확인하기 위해, 재조합 Env 단백질 패널 (gp120 및 gp140)을 제조하였다. 동량의 각 Env 단백질 (100 ng)을 SDS-폴리아크릴아마이드 겔에 로딩하고, 니트로셀룰로오스에 트랜스퍼한 뒤, 웨스턴 블랏 분석에서 항-CON6 gp120 기니아 피그 혈청 (1:1,000 희석) 뿐만 아니라 서브타입 A 내지 G 환자 혈청으로 반응시켰다. 각 HIV-1 서브타입에 대하여, 4 내지 6개의 환자 혈청을 테스트하였다. 각 서브타입에 대해 대표적인 1개의 혈청이 도 4에 표시되어 있다.
- [0076] 테스트된 모든 서브타입 혈청이 패널에 있는 Envs에서 다양한 반응도를 나타내었으나, 모든 그룹 M 서브타입 환자 혈청은 CON6 gp120 Env 단백질과 동일하게 반응하는 것으로 확인되었으며, 이는 환자 혈청에 의해 인식되는 야생형 HIV-1 Env 에피토프가 CON6 Env 단백질에 잘 보존되어 있음을 설명해준다. 기니아 피그에서 유발된 CON6 gp120 항혈청이 다른 서브타입 Env 단백질에 반응할 수 있는지 확인하는 실험을 다음에 수행하였다. CON6 혈청은 서브타입 A Env 단백질을 제외한, 그 자신 및 다른 서브타입 Env 단백질과 동일하게 잘 반응하는 것으로 확인되었다 (도 4).
- [0077] *CON6, 서브타입 B 및 서브타입 C 외피 중첩 펩타이드에 대한 T 세포 반응의 유도.*
- [0078] CON6 Env 면역원에 의해 유도되는 T 세포 면역 반응을 서브타입 특이 면역원에 의해 유도되는 것과 비교하기 위해, 추가로 2 그룹의 생쥐를 서브타입 B 또는 서브타입 C DNAs 및 서브타입 B 또는 C 외피 단백질을 발현하는 대응되는 rVV로 면역화시켰다. 서브타입 B (JRFL) 또는 서브타입 C (96ZM651) Env 면역원으로 면역화된 마우스는 제1 서브타입-특이 T 세포 반응을 갖는다 (도 5). JRFL (서브타입 B) 면역원으로 면역화된 마우스의 IFN- γ SFCs는, 서브타입 B (MN) 펩타이드 풀로 자극시키고 난 뒤에 검출되었으나, 서브타입 C (Chn19) 또는 CON6 펩타이드 풀로는 검출되지 않았다. 96ZM651 (서브타입 C) 면역원으로 면역화된 마우스의 IFN- γ SFCs는, 서브타입 C (Chn19) 및 CON6 펩타이드 풀로 자극시키고 난 뒤에 검출되었으나, 서브타입 B (MN) 펩타이드 풀로는 검출되지 않았다. 반대로, IFN- γ SFCs는, 서브타입 B 또는 C 펩타이드 풀 뿐만 아니라 CON6 펩타이드 풀로 자극될 때, CON6 Env 면역원으로 면역화된 마우스로부터 확인되었다 (도 5). CON6 gp140에 의해 유도된 T 세포 면역 반응은 CON6 gp120에 의해 유도된 것보다 훨씬 강력한 것으로 보였다. 종합하자면, 이들 데이터는 CON6 gp120 및 gp140CF 면역원이, 야생형 서브타입 B 및 C 외피의 T 세포 에피토프를 인식한 T 세포 반응을 유도하는 것이 가능함을 설명해준다.
- [0079] *HIV-1 서브타입 B 및 C 제1 분리물을 중화시키는 재조합 CON6 gp120 및 gp140CF 외피에 의한 항체의 유도.*
- [0080] CON6 외피 면역원이 HIV-1 제1 분리물을 중화시키는 항체를 유도할 수 있는지 확인하기 위해, 기니아 피그를 CON6 gp120 또는 gp140CF 단백질로 면역화시켰다. 4 또는 5회의 면역화 후에 수집한 혈청은, 중화 분석 및 대응되는 전-채혈된 혈청과 비교하기 위해 사용하였다. 2개의 AT-2 불활성화된 HIV-1 분리물 (ADA 및 AD8)을 융합체 저해 분석법에서 테스트하였다 (표 5A). 2개의 서브타입 B SHIV 분리물, 8개의 서브타입 B 제1 분리물, 4개의 서브타입 C 및 각 서브타입 A, D 및 E 제1 분리물 1개를 MT-2 또는 루시퍼라아제-기초 분석에서 테스트하였다 (표 5B). 융합체 저해 분석에서는, CON6 gp120 및 gp140CF 단백질에 의해 유도된 항체가 AT-2 불활성화된 ADA 및 AD8-유도된 융합체를 강력하게 저해시키는 것으로 확인되었다 (표 5A). MT-2 분석에서는, 2개의 gp120 및 1개의 gp140CF 혈청에 의해 SHIV 분리물 2개 중 1개 (SHIV SF162P3)에서 약한 중화가 확인되었다 (표 5B). 루시퍼라아제-기초 분석에서는, 모든 gp120 및 gp140CF 혈청에 의해 서브타입 B 제1 분리물 8개 중 4개 (BX08, SF162, SS1196 및 BAL)에서 강한 중화가 확인되었으며, gp120 및 gp140CF 혈청에 의해 서브타입 B 분리물 8개 중 2개 (6101, 0692)에서 약한 중화가 확인되었다. HIV-1 PAVO에 대해서는 중화가 검출되지 않았다 (표 5B).

다음으로, 4개의 서브타입 C HIV-1 분리물에 대해 CON6 항-gp120 및 gp140CF 혈청이 테스트되었으며, 항-CON6 gp120 혈청에 의해 일차적으로 분리물 4개 중 3개 (DU179, DU368 및 S080)가 확인되었다. 1개의 gp140CF 혈청, 제653는 DU179를 강력하게 중화시켰으며 S080은 약하게 중화시켰다 (표 5B). 최종적으로, 항-CON6 Env 혈청은, 서브타입 D 분리물 (93ZR001)을 강력하게 중화시키고, 서브타입 E (CM244) 분리물을 약하게 중화시키며, 서브타입 A (92RW020) 분리물을 중화시키지 않았다.

[표 5A]

HIV-1 그룹 M 컨센서스 외피 CON6 단백질의 용해 억제 항체를 유도하려는 능력				
기니아 피그 번호	면역원	용해제 지해 항체 역가 ¹		
		AD8	ADA	
646	gp120	270	270	
647	gp120	90	90	
648	gp120	90	270	
649	gp120	90	90	
기하적 평균 역가		119	156	
650	gp140	270	270	
651	gp140	90	90	
652	gp140	270	810	
653	gp140	270	90	
기하적 평균 역가		270	207	

1 Sup T1 세포의 HIV-유도된 융합체가 전-면역 혈청에 비해 >90% 저해되는 역의(reciprocal) 혈청 희석. 모든 전-채혈된 혈청은 음성이었다 (역가<10).

[표 5B]

그룹 M 컨센서스 HIV-1 외피 CON6 gp120 및 gp140CF 단백질의 HIV 제1 분리물 중화 항체를 유도하는 능력														
HIV 분리물 (서브타입)	CON6 gp120 단백질 기니아 피그 번호					CON6 gp140CF 단백질 기니아 피그 번호					HIV: Serum			
	646	647	648	619	GM1	650	651	652	653	GM2	UnMab+	CD4:µg/ml	HIV: Serum	
	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	NT
SHV 89.5P(B)	270	30	48	270	270	270	270	270	270	270	NT	0.2µg/ml	NT	NT
SHV SF162(B)	270	183	254	55	102	199	64	229	150	187	0.7µg/ml	NT	NT	2354
HX106(B)	270	38	35	270	270	270	90	72	73	39	1.1µg/ml	NT	NT	NT
6101(B)	270	270	270	270	270	270	270	270	270	270	2.7µg/ml	NT	NT	NT
B-311(S)(B)	31	32	34	270	24	28	33	30	45	33	0.8µg/ml	NT	NT	769
0692(B)	270	270	270	270	270	270	270	270	270	270	2.9µg/ml	NT	NT	NT
PAY07(B)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	NT
SH162(B)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
SS119(B)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
BAL(B)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
92RW020(A)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
DU179(C)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
DU368(C)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
S021(C)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
S080(C)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
93ZR001(D)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270
CM244(E)	270	270	270	270	270	270	270	270	270	270	NT	NT	NT	270

- [0086] * MT-2 분석; 다른 HIV 분리물 모두를 M7-루시퍼라아제 분석에서 테스트하였다.
- [0087] HIV-1 분리물 QH0692, SS1196, SF162, 6101, BX08, BG1168 및 BAL은 후-주입 5개 혈청으로 분석하였으며; 다른 HIV-1 분리물은 후-주입 4개 혈청으로 분석하였다. ND = 수행하지 않음.
- [0088] HIV + 혈청은, HIV-1 + HIV-1 분리물 SS1196에 대해 공지된 중화 활성을 갖는 인간 혈청 (LEH3) 또는 항-gp120 기니아 피그 혈청(#)이다. GMT = 그룹 당 4마리 동물의 기하 평균 역가. 보고된 중화 역가는 전-채혈 혈청에 있는 모든 백그라운드 중화를 제외하고 난 뒤의 것이다. ‡TriMab2 = 인간 mabs 2F5, b12 및 2G12의 혼합.
- [0089] 결론
- [0090] 기능적 Env 단백질을 인코딩하는 것으로 개시된 인공 HIV-1 그룹 M 컨센서스 env 유전자 (인코딩 서열) (CON6 및 Con-S)의 생산은, 바이러스의 삽입을 유발하기 위해 CCR5 공-수용체를 이용할 수 있다. 중요하게는, 이들 그룹 M 컨센서스 외피 유전자는 T 및 B 세포 반응을 유발할 수 있으며, 서브타입 B 및 C HIV-1 제1 분리물의 에피토프를 인식하였다. 추가로, Con-S는 서브타입-C 및 A HIV-1 균주를 강력하게 중화시키는 항체를 유도한다 (표 3 참조).
- [0091] HIV-1에 대한 보호의 관련성은 결론적으로 공지되어 있지 않다. 동물 모델로부터의 주목할 만한 데이터 및 HIV-1-감염된 환자의 연구에서는, HIV-1 백신 개발의 목표가 광범위한-반응성의 CD4+ 및 CD8+ 항-HIV-1 T 세포 반응 (Letvin et al, Annu. Rev. Immunol. 20:73-99 (2002)) 및 다중 서브타입의 HIV-1 제1 분리물을 중화시키는 높은 레벨의 항체 (Mascola et al, J. Virol. 73:4009-4018 (1999), Mascola et al, Nat. Med. 6:270-210 (2000))를 유도해야 함을 제시한다.
- [0092] HIV-1의 높은 수준의 유전자 변이성은, 임상적으로 유용하여야 하는 충분한 범위의 면역 반응을 유도하는 것이 가능한 면역원의 디자인을 어렵게 한다. T 및 B 세포 반응을 위한 에피토프 기초의 백신 (McMichael et al, Vaccine 20:1918-1921 (2002), Sbai et al, Curr. Drug Targets Infect. Disord. 1:303-313 (2001), Haynes, Lancet 348:933-937 (1996)), 융합 중간체에 반사적인 구속된 (constrained) 외피 (Fouts et al, Proc. Natl. Acad. Sci. USA 99:11842-22847 (2002)) 뿐만 아니라 항-HIV-1 중화 항체를 유도하기 위한 보존된 고-순위 (high-order) 구조의 노출이, HIV-1 변이성을 극복하기 위해 제안되었다 (Roben et al, J. Virol. 68:4821-4828 (1994), Sapphire et al, Science 293:1155-1159 (2001)). 다양성이 끊임없이 증가하고 HIV-1이 급속히 진화함에 따라, 바이러스는 복합 표적으로 빠르게 이동하며, HIV-1 변이 복잡성의 나머지는 이들 시도 모두가 문제가 있게 한다. HIV-1 면역원 디자인에 대한 최근의 가장 공통적인 시도는, 백신이 테스트되어야 하는 부위로부터 존재하거나 존재하지 않는 야생형 분야 HIV-1 분리물을 선택하는 것이다. 다가 외피 면역원은 다중 외피 면역원을 결합하도록 디자인되었다 (Bartlett et al, AIDS 12:1291-1300 (1998), Cho et al, J. Virol. 75:2224-2234 (2001)).
- [0093] 상술한 연구에서는, 후보 면역원 및 야생형 분야 바이러스 균주 사이의 유전적 거리가 감소된 그룹 M 컨센서스 env 유전자 (CON6)를 제조함으로써, HIV-1 면역원 디자인에 대한 새로운 전략을 테스트한다. CON6 env 유전자는 대부분의 위치에 있는 가장 공통적인 아미노산을 선택하여 모든 서브타입에 대해 제조되었다 (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). 가장 공통적인 아미노산이 사용되었기 때문에, 다수의 항체 및 T 세포 에피토프가 잘 보존되었다. 중요하게는, 그룹 M 컨센서스 env 서열 및 모든 서브타입 env 서열 사이의 유전적 거리가 약 15%이었으며, 이는 야생형 서브타입 사이의 거리 (30%)의 절반이다 (Gaschen et al, Science 296:2354-2360 (2002)). 이 거리는 동일한 서브타입 내의 바이러스에서와 거의 동일하다. 더불어, CRFs가 서브타입 중에서 전체적으로 유전적 차이를 증가시키지 않았기 때문에, 그룹 M 컨센서스 env 유전자는 모든 재조합 바이러스 env 유전자와 약 15% 달랐다.
- [0094] CON6-Env 유사비리온의 감염도는, 인공 외피가 "최적의" 기능적 형태에 있지 않지만 바이러스의 삽입을 유도할 수 있음을 나타내는, 감염도가 포함된 싱글-라운드 감염 시스템을 이용해 확인하였다. 주요한 HIV-1 감염된 환자가 R 바이러스로 최초에 감염되었기 때문에, 이의 공-수용체로서 CON6 외피 사용된 CCR5 (R5)가 중요하다.
- [0095] BIAcore 분석에서는 CON6 gp120 및 gp140CF 모두가 결합한 sCD4 및 야생형 HIV-1 Env 단백질에 결합하는 mabs의 수를 보여준다. 야생형 HIV-1 외피에 대해 동일하게 항원성인 CON6 gp120 및 140CF 단백질의 발현은, HIV-1 면역원 개발에 있어서 중요한 단계이다. 그러나, 많은 야생형 외피 단백질은, 잠재적 중화 인간 mabs가 결합하

는 에피토프를 발현하나, 그 자체의 면역원으로서 사용될 때에는 중화 인간 mabs 특이성의 항-HIV-1 항체를 광범위하게 중화하지 않는다.

[0096] 중화 항체 연구는, CON6 gp120, CON6 gp140CF 및 Con-S gp140CFI 유도된 중화된 항체가, 비-서브타입 B 제1 HIV 분리물의 가장 강력한 중화를 유도하는 Con-S gp140CF를 이용하여 서브타입 B, C 및 D HIV-1 제1 분리물을 선별하는 것을 촉진시킨다. 그러나, 가장 중화시키기 어려운 제1 분리물 (PAVO, 6101, BG1168, 92RW020, CM244)이 항-CON6 gp120 또는 gp140 혈청에 의해 단지 약하게 중화되거나 중화되지 않음이 확실하다 (표 4b). 그럼에도 불구하고, 중화 항체를 유도하기 위한 Con-S 외피 면역원성이 기대되고 있으며, 반응의 범위는 비-서브타입 B HIV 분리물에 대한 Con-S 서브유닛 gp140CFI 외피 단백질을 이용해 형성되는 것으로 제시된다. gp120 및 gp160을 발현하는 폭스바이러스 컨스트럭트를 이용한 기존의 연구에서는 높은 레벨의 중화 항체를 제조하지 못했다 (Evans et al, J. Infect. Dis. 180:290-298 (1999), Polacino et al, J. Virol. 73:618-630 (1999), Ourmanov et al, J. Virol. 76:292-302 (2002), Excler and Plotkin, AIDS 11(Suppl A):S127-137 (1997)). 분비된 CON6 gp120 및 gp140을 발현하는 rVV가 컨스트럭트되었으며, HIV-1 제1 분리물을 중화시키는 항체가 유도되었다. HIV 중화 항체 면역원은, Con-S gp140CF 또는 이의 서브유닛과 서브타입 B 분리물을 가장 잘 중화시키는 면역원의 조합이 될 수 있다.

[0097] 올리고머 gp140 단백질의 구조는 단백질 면역원성을 평가하는데 있어서 중요하다. 이점에 있어서, 빠른 성능 액체 크로마토그래피 (fast performance liquid chromatography, FPLC) 및 분석적 초원심분리 (analytical ultracentrifugation)에 의한 정제된 CON6 gp140CF 단백질의 연구에서는, 정제된 gp140 피크가 작은 다imer 성분을 가지는 우세한 트리머(trimer)를 포함함을 설명하였다.

[0098] 따라서, CON6, Con-S 또는 본원에 개시된 2003 그룹 M 또는 서브타입 컨센서스 또는 원형 인코딩 서열은, CD4-Env 복합체를 포함하는 다양한 잠재적으로 "향상된" 외피 면역원, 경직된 외피 구조, 및 트리머형 (trimeric) 올리고머 형태를 제조하기 위한 관심을 갖게 하는 후보이다. HIV-1 감염 및/또는 SHIV 챌린지 모델에서의 질병에 대해 방어하고자 하는 CON6-유도된 T 및 B 세포 반응의 능력은 비-인간 영장류에서 연구될 것이다.

[0099] 상기 연구에서는, 그룹 M 컨센서스 env 유전자 (CON6) 및 Con-s와 같은 인공 센트럴화된 HIV-1 유전자가, 야생형 서브타입 B 및 C Env 단백질에 있는 T 세포 에피토프 뿐만 아니라 그룹 M 컨센서스 Env 단백질에 대해 T 세포 반응을 유도한다고 설명하였다 (도 5). DNA 프라임 및 CON6 gp140CF 면역원을 이용한 rVV 추가접종 요법이 서브타입 B 및 C 에피토프를 인식하는 IFN- γ 생산 T 세포를 명확하게 유도하지만, CON6 외피에서 발견되는 것과 같은 센트럴화된 서열이 야생형 HIV-1 유전자보다 교차-클레이드 T 세포 반응을 훨씬 더 유도하는지를 확인하기 위해 추가의 연구가 필요했다 (Ferrari et al, Proc. Natl. Acad. Sci. USA 94:1396-1401 (1997), Ferrari et al, AIDS Res. Hum. Retroviruses 16:1433-1443 (2000)). 그러나, CON6 (및 Con-S env 인코딩 서열) 프라임 및 추가접종된 비장세포 T 세포가 HIV-1 서브타입 B 및 C T 세포 에피토프를 인식한다는 사실은, CON6 (및 Con-S)가 T 세포 반응을 유도할 수 있어 임상적으로 유용할 수 있음을 설명해 주는데 있어서 중요한 단계가 된다.

[0100] 3가지의 컴퓨터 모델 (컨센서스, 원형 및 트리의 중심 (COT))이, 센트럴화된 HIV-1 유전자를 제조하기 위해 제안되었다 (Gaschen et al, Science 296:2354-2360 (2002), Gao et al, Science 299:1517-1518 (2003), Nickle et al, Science 299:1515-1517 (2003), Korber et al, Science 288:1789-1796 (2000)). 이들 모두는 서브타입 내 또는 서브타입간의 대부분의 HIV-1 서열에 대한 성형(star-like)의 계통발생학 트리의 뿌리 부분에 위치하는 경향이 있다. 실험적 백신으로서, 이들 모두는 면역원 및 필드 바이러스 균주 사이의 유전자 거리를 줄여 줄 수 있다. 그러나, 컨센서스, 원형 및 COT 서열 각각은 이점과 단점이 있다 (Gaschen et al, Science 296:2354-2360 (2002), Gao et al, Science 299:1517-1518 (2003), Nickle et al, Science 299:1515-1517 (2003)). 컨센서스 및 COT는 표본추출된 최근의 야생형 바이러스에 있는 서열 또는 에피토프를 나타내며, 아우트라이어(outlier) HIV-1 서열에 의해 거의 영향을 받지 않지만, 원형 서열을 나타내는 원형은 아우트라이어 서열에 의해 상당히 영향을 받는다. 그러나, 최근에는, 센트럴화된 서열이, 다양한 HIV-1 균주에 대한 광범위한 면역 반응을 유발하기 위한 최상의 면역원으로서 제공될 수 있음이 공지되어 있지 않으며, 이러한 다른 계획을 테스트하기 위해 연구가 진행되고 있다.

[0101] 이와 함께, 상기 데이터는, HIV-1 인공 CON6 및 Con-S 외피가 야생형 HIV-1 에피토프에 대해 T 세포 반응을 유도할 수 있으며, HIV-1 제1 분리물을 중화하는 항체를 유도할 수 있음을 보여주며, HIV-1 백신 디자인에 있는 인공 센트럴화된 HIV-1 서열 이용의 실현가능성 및 가망을 설명해준다.

[0102] **실시예 2**

[0103] HIV-1 서브타입 C 원형 및 컨센서스 외피 당단백질

[0104] 구체적 실험

[0105] HIV-1 서브타입 C 원형 및 컨센서스 env 유전자는 Los Alamos HIV 분자 면역학 데이터베이스 (<http://hiv-web.lanl.gov/immunology>)에서 입수하였으며, 코돈-사용을 포유동물 세포 발현에 최적화시키고 합성하였다 (도 6). 최적의 발현이 가능하게 하기 위해, 코작 (Kozak) 서열 (GCCGCCGCC)을 개시 코돈의 바로 업스트림에 삽입하였다. 전장 유전자에 추가로, gp41 막-스패닝 (spanning) 도메인 (IVNR) 및 gp120/gp41 절단 위치 (REKR)의 바로 뒤에 중지 코돈을 삽입하여 2개의 절단된 env 유전자를 제조함으로써, gp140 및 gp120의 당단백질 형을 각각 제조하였다 (도 8).

[0106] 유전자를 인 비트로 전사/번역 시스템에서의 통합성 (integrity)에 대해 테스트하였으며, 포유동물 세포에서 발현시켰다. 원형 및 컨센서스 서브타입 C 외피가 융합 및 삽입의 유도를 가능하게 하는지 확인하기 위해, gp160 및 gp140 유전자를 HIV-1/SG3Δenv 프로바이러스와 함께 동시-형질감염시켰으며, 제조된 유사비리온을 JV53-BL 세포 분석법을 이용해 감염도를 테스트하였다 (도 7). 또한, 공-수용체 사용 및 외피 중화 감염도는 JC53-BL 분석법을 약간 변형시킨 것으로 확인하였다. 코돈-사용 최적화되고 rev-의존성인 96ZAM651 env 유전자를 등가의(contemporary) 서브타입 C 대조군으로서 사용하였다.

[0107] 결과

[0108] 코돈-최적화된 서브타입 C 원형 및 컨센서스 외피 유전자 (gp160, gp140, gp120)는 포유동물 세포에서 env 당단백질을 높은 레벨로 발현한다 (도 9).

[0109] 코돈-최적화된 서브타입 C gp160 및 gp140 당단백질은 바이러스 입자에 효율적으로 결합된다. 수크로즈-정제된 유사비리온의 웨스턴 블랏 분석에서는, rev-의존성 등가의 외피 대조군의 비리온 결합과 비교해 코돈-최적화된 외피의 비리온 결합이 10배 높음을 보여준다 (도 10A).

[0110] 서브타입 C 컨센서스 gp160 또는 gp140 외피로 유사타입된 비리온은 대응되는 gp160 및 gp140 원형 외피를 포함하는 유사비리온에 비해 훨씬 더 감염성이 있었다. 더불어, gp160 외피들은 이들 각각의 gp140 등가물에 비해 훨씬 더 지속적으로 감염성이 있었다 (도 10b).

[0111] 서브타입 C 원형 및 컨센서스 외피 모두는 바이러스 삽입을 유발하는 공-수용체로서 CCR5를 이용한다 (도 11).

[0112] 유사비리온을 포함하는 서브타입 C 원형 및 컨센서스 gp160의 감염도는 서브타입 C 감염된 환자의 혈장에 의해 중화되었다. 이는, 이러한 인공 외피가 천연 HIV-1 env 당단백질의 구조와 유사한 구조를 가지고, 공통의 중화 에피토프가 보존되어 있음을 제시한다. 중화 포텐셜에서는 어떠한 유의적인 차이도, 서브타입 C 원형 및 컨센서스 env 당단백질 (gp160) 사이에서 확인되지 않았다 (도 12).

[0113] 결론

[0114] HIV-1 서브타입 C 바이러스는 가장 많이 발병하는 순환 분리물 중의 하나이며, 전세계의 신규 감염 중에서 약 50%를 차지한다. 광역의 순환 HIV-1 균주 사이의 유전자 다양성은 백신 디자인에 대한 시도를 유발한다. 비록 HIV-1 Env 단백질이 가변성이 높지만, 감염된 숙주에서 체액성 및 세포성 면역 반응 모두를 유도할 수 있다. 70개의 HIV-1 완전 서브타입 C env 서열을 분석함으로써, 컨센서스 및 원형 서브타입 C env 유전자가 생성되었다. 이들 서열은 등가의 서브타입 C 균주로부터 대략 등거리에 있으며, 교차-방어적 면역을 훨씬 더 유도하는 것으로 예측되었다. 재-컨스트럭트된 원형 또는 컨센서스 서열 유도의-면역원은, 백신 후보자 및 등가의 분리물 사이의 유전적 차이의 확장을 최소화한다. 그러나, 컨센서스 및 원형 서브타입 C env 유전자는 아미노산 서열이 5% 다르다. 컨센서스 및 원형 서열 모두는 분석하기 위해 합성되었다. 코돈-최적화된 서브타입 C 원형 및 컨센서스 외피 유전자가 컨스트럭트되었으며, 발현된 당단백질의 인 비트로 생물학적 특성이 확인되었다. 합성 서브타입 C 컨센서스 및 원형 env 유전자는 등가의 서브타입 C 야생형 외피 당단백질과 유사한 구조, 기능 및 유사한 항원성을 가진 당단백질을 발현한다.

[0115] 실시예 3

[0116] 서브타입 C gag 및 nef 유전자의 컨센서스의 코돈-사용 최적화 (C.con.gag 및 C.con.nef)

[0117] 서브타입 C 바이러스는 전세계에 있는 그룹 M 바이러스의 모든 서브타입 중에서 가장 많이 발병하는 바이러스가 되었다. 최근, HIV-1 감염된 사람의 50% 이상이 HIV-1 서브타입 C 바이러스를 가지고 있다. 또한, 상당한 내부(intra)-서브타입 C 변이성이 있다: 다른 서브타입 C 바이러스는 이의 Gag, Pol, Env 및 Nef 단백질과 각각 10%, 6%, 17% 및 16% 만큼 많이 차이날 수 있다. 가장 중요하게는, 어느 한 나라로부터의 서브타입 C 바이러스가 전세계의 다른 나라로부터 분리된 바이러스만큼 다양할 수 있다는 것이다. 단지 예외는, 서브타입 C가 훨씬 더 최근에 도입된 인도/중국, 브라질 및 에티오피아/지부티로부터의 HIV-1 균주이다. 심지어 한 나라 내에서도 서브타입 C 바이러스의 유전적 다양성이 높기 때문에, 단일 바이러스 분리물 기초의 면역원이 동일 지역 내에서 순환하는 다른 분리물에 대해 방어적 면역성을 유발하지 않을 수도 있다.

[0118] 따라서, 서브타입 C 바이러스의 gag 및 nef 유전자 서열은, 50% 컨센서스 역치를 이용함으로써 상기 유전자 모두에 대한 컨센서스 서열을 제조하는데 사용되었다. 원조 (founder) 바이러스로의 잠재적 치우침을 피하기 위해, 오직 1개의 서열을 인도/중국, 브라질 및 에티오피아/지부티 각각으로부터 사용하여, 서브타입 C 컨센서스 서열을 제조하였다 (C.con.gag 및 C.con.nef). C.con.gag 및 C.con.nef 유전자 모두의 코돈은 고발현된 인간 유전자의 코돈 사용에 기초하여 최적화하였다. 293T 세포의 형질감염 후의 단백질 발현은 도 13에 표시되어 있다. 보여지는 바와 같이, 컨센서스 서브타입 C Gag 및 Nef 단백질 모두는, 효율적으로 발현되었으며, Gag- 및 Nef-특이 항체에 의해 인지되었다. C.con.gag 및 C.con.nef 유전자 모두의 단백질 발현 레벨은, 천연 서브타입 env 유전자 (96ZM651)의 발현과 비교할 만하였다.

[0119] 실시예 4

[0120] 전장 "컨센서스 변이 위치를 가진 컨센서스 env 유전자의 컨센서스" (CON-S)의 합성

[0121] 합성된 "컨센서스 env 유전자의 컨센서스 (CON-S)"에서, 다양한 부위를 등가의 서브타입 C 바이러스 (98CN006)의 대응 부위로 치환하였다. 또한, 디자인된 con/con 유전자도 컨센서스 변이 부위 (CON-s)를 추가로 포함한다. Con-S env 유전자의 코돈을 고발현된 인간 유전자의 코돈 사용에 기초하여 최적화하였다. (아미노산 서열 및 핵산 서열 각각에 대해 도 14a 및 14b 참조)

[0122] 3' 말단에 20 bp가 중복되고, 그의 5' 및 3' 말단에 제한효소 위치인 EcoRI 및 BbsI 뿐만 아니라 BsmBI 및 BamHI를 각각 포함하는 불변 서열을 포함하는 쌍을 이루는 올리고뉴클레오타이드 (80머)를 디자인하였다. BbsI 및 BamHI는 그의 인식 서열의 바깥을 절단하는 타입 II 제한효소이다. 상기 제한효소들은 이러한 방법에서 올리모머 내에 위치하며, 18 bp의 불변 위치에 인접한 4개의 잔기를 먼저 절단하고, 이후의 라이게이션 단계를 위해 각 단편의 말단에 있는 4 염기 5' 오버행을 남긴다. 18 bp 불변 서열에 상보적인 프라이머와 PCR을 이용해 26개의 쌍을 이루는 올리고머를 각각 연결시켰다. T/A 클로닝 방법을 이용해 각 쌍을 pGEM-T (Promega)에 클로닝시켰으며, 서열 분석하여 우연한 돌연변이/절단의 부재를 확인하였다. 그리고 난 뒤, 적당한 인서트를 포함하는 pGEM-T 서브클론을 분해한 뒤, 1% 아가로스 젤에서 이동시키고, 젤을 정제하였다 (Qiagen). 4개의 108-머 각각을 pcDNA3.1 (Invitrogen)에 다중-단편 라이게이션 반응으로 라이게이션시켰다. 유전자의 5' 말단에서부터 3' 말단까지 단계별로 4회의 라이게이션을, 단편 그룹에서 유발시켰다. 이 과정은 전체 유전자가 pcDNA3.1 벡터에 재컨스트럭트될 때까지 반복하였다.

[0123] 코돈 사용 최적화된 올리고 쌍을 함께 라이게이션시킴으로써, 완전한 Con-S 유전자를 컨스트럭트하였다. 이의 개방 판독 프레임 (open reading frame)을 확인하기 위해, 인 비트로 전사 및 번역 분석을 수행하였다. 단백질 생성물은, 번역 단계 동안 S35-메티오닌으로 표지시키고, 10% SDS-PAGE에서 분리시킨 후 자가방사기록법 (radioautography)으로 판독하였다. 발현된 Con-S gp160의 예측되는 사이즈를 7개의 클론 중 4개에서 확인하였다 (도 14c).

[0124] 293T 세포로 형질감염되고 난 뒤 포유동물 세포에서의 CONs Env 단백질 발현을 웨스턴 블랏 분석법으로 확인하였다 (도 15). Con-S Env 단백질의 발현 레벨은, 컨센서스 보존적 부위 및 98CN006 바이러스 분리물의 다양한 루프를 포함하는 기존의 CON6 env 클론으로부터 관찰된 것과 매우 유사하였다.

[0125] Con-S env 클론 및 env-결핍 SG3 프로바이러스 클론을 293T 세포로 동시형질감염시킴으로써, Env-유사비리온을

생산하였다. 형질감염 2일 후에, 유사비리온을 수득하여 JC53BL-13 세포에 감염시켰다. 감염 단위(infectious units, IU)는, X-gal로 염색하고 난 뒤 푸른색 세포를 계수하는 3회의 단독 실험으로 측정하였다. CON-6 env 클론과 비교했을 때, Con-S env 클론은 JC53BL-13 세포에서 유사한 수의 IU를 생산하였다 (도 16). 상기 2개 모두에 대한 IU 역가는 SG3 골격 클론 대조군 (Env 없음)에 비해 약 3 log가 더 높다. 그러나, 역가는 양성 대조군 (천연 HIV-1 env 유전자, NL4-3 또는 YU2)에 비해 약 2 log가 더 낮다. 이들 데이터는 컨센서스 그룹 M env 클론 모두가 생물학적으로 기능을 함을 제시해 준다. 그러나, 이들의 기능성(functionality)은 절충되었다. 기능적 컨센서스 env 유전자는, 이들 Env 단백질이 정확하게 접히고, 천연 Env 단백질의 기본 형태를 보존하며, 통상적인 Env 면역원으로서 개발될 수 있음을 보여준다.

[0126] 다음으로, JC53-BL 세포로 삽입되기 위한 공-수용체 Con-S Env의 사용을 확인하였다. CXCR4 저해제 AMD3100으로 처리될 때에는, NL4-3 Env-유사비리온의 감염도가 저해되었으나, YU2, Con-S 또는 CON6 Env-유사비리온의 감염도는 저해되지 않았다. 반대로, CCR5 저해제 TAK779로 처리될 때에는, NL4-3 Env-유사비리온의 감염도는 영향을 받지 않았으나, YU2, Con-S 또는 CON6 Env-유사비리온의 감염도는 저해되었다. 상기 저해제 모두로 처리될 때에는, 모든 유사비리온의 감염도가 저해되었다. 종합하자면, 이들 데이터는 Con-S 뿐만 아니라 CON6 외피가, 표적 세포로 삽입되기 위해 CCR5를 이용하나 CXCR4 공-수용체는 그렇지 않음을 보여준다.

[0127] 다음으로, CON6 또는 Con-S Env 단백질이 유사비리온으로 동등하게 효과적으로 삽입될 수 있는지 여부를 확인하였다. 얼마나 많은 Env 단백질이 유사비리온으로 삽입되는지 정확하게 비교하기 위해, 각 유사비리온을 동일한 농도로 SDS-PAGE에 로딩하였다: 세포 분해물에 대해 총 단백질 5 μ g, 세포 배양 상등액에 대해 p24 25 ng, 또는 정제된 바이러스 농축물 (고속으로 원심분리하고 나서 농축된 유사비리온)에 대해 p24 150 ng. 모든 표본 (세포 분해물, 세포 배양 상등액 또는 정제된 바이러스 농축물)에서, CON6 또는 Con-S Env-유사비리온에 결합된 Env 단백질의 양은 차이가 없었다 (도 17).

[0128] 실시예 5

[0129] 컨센서스 서브타입 A 전장 env (A.con.env) 유전자의 합성

[0130] 서브타입 A 바이러스는, HIV-1 감염의 70% 이상이 보고된 아프리카 대륙에서 2번째로 가장 많이 발병하는 HIV-1이다. 아프리카 및 전세계에서 가장 많이 발병하는 바이러스인 서브타입 C 바이러스에 대한 컨센서스 gag, env 및 nef 유전자는 상기에서 제조하였다. 서브타입 A 및 C 바이러스 간의 유전자 거리가 env 유전자의 30% 만큼 높기 때문에, 상기 서브타입 모두에서의 교차 반응성 및 저해도는 최적화되지 않을 것이다. 또한, 모든 서브타입에 대한 2개의 그룹 M 컨센서스 env 유전자도 제조하였다. 그러나, 모든 각각의 서브타입 바이러스들을 표적으로 하기 위해서는, 서브타입 컨센서스 유전자 및 동일한 서브타입의 필드 바이러스 간의 유전자 거리가 그룹 M 컨센서스 유전자와 이들 동일한 바이러스 간의 거리보다 작을 것이기 때문에, 서브타입 특이 컨센서스 유전자가 훨씬 더 효과적일 것이다. 따라서, 컨센서스 유전자는, 서브타입 A 특이 면역원의 개발을 위해 제조될 필요가 있다. A.con.env 유전자의 코돈은 고발현된 인간 유전자의 코돈 사용에 기초하여 최적화시켰다. (아미노산 및 핵산 서열에 대해 각각 도 18a 및 18b 참조)

[0131] 올리고의 각 쌍을 증폭하고, 클로닝하며, 라이게이션하고 서열분석하였다. A.con env 유전자의 개방 판독 프레임 인 비트로 전사 및 번역 시스템으로 확인하고 난 뒤, A.con env 유전자를 293T 세포로 형질감염시키고, 단백질 발현 및 특이성을 웨스턴 블랏 분석으로 확인하였다 (도 18). 그리고 난 뒤, A.con 외피가 생물학적으로 기능적인지 여부를 확인하였다. 이를 env-결합 SG3 프로바이러스 클론과 함께 293T 세포에 동시-형질감염시켰다. 유사타입 바이러스를 수득한 뒤 JC53BL 세포를 감염시키는데 사용하였다. A.con Env-유사비리온으로 감염된 JC53-BL 세포에서 푸른색 세포가 검출되었으며, 이는 A.con Env 단백질이 생물학적으로 기능을 함을 제시해 준다 (표 6). 그러나, A.con Env-유사비리온의 감염 역가는 야생형 서브타입 C 외피를 가진 유사비리온의 감염 역가보다 약 7배 낮았다 (표 6). 종합하자면, 생물학적 기능성 A.con Env 단백질은, 정확하게 폴딩하고, Env 면역원으로서 사용되는 경우 선형(linear) 및 형태적 T 및 B 세포 에피토프를 유도할 수 있음을 제시한다.

[0132] [표 6]

		JC53BL13 (IU/ul)		
		3/31/03	4/7/03	4/25/03
		비여과된 상등액	0.22μm 여과된 것	0.22μm 여과된 것
A.con	+SG3	4	8.5	15.3
96ZM651	+SG3	87	133	104
SG3	골격	0	0.07	0.03
음성 대조군		0	0.007	0

A.con env 유전자를 이용한 유사비리온의 감염도

[0133]

[0134] 실시예 6

[0135] 전장 "컨센서스 gag, pol 및 nef 유전자의 컨센서스" (M.con.gag, M.con.pol 및 M.con.nef) 및 서브타입 C 컨센서스 pol 유전자 (C.con.pol)의 디자인

[0136] 그룹 M 컨센서스 유전자에 대해, 2개의 다른 env 유전자를, 1개는 바이러스 특이 변이 부위 (CON6)를 이용하고 다른 1개는 컨센서스 변이 부위(Con-S)를 이용해 컨스트럭트하였다. 그러나, 면역화되고 백신화된 동물에서의 T 세포 면역 반응의 분석에서는, env 유전자가 중화 항체를 유도하는 단 하나의 유전자라고 하더라도 정상적으로는 T 세포 면역 반응에 대한 주요한 표적이 아님을 보여준다. 대신에, HIV-1 Gag, Pol 및 Nef 단백질이 잠재적 T 세포 면역 반응을 유도하는데 있어서 중요한 것으로 밝혀졌다. 모든 서브타입에 대해 광범위한 체액성 및 세포성 면역 반응을 모두 유도할 수 있는 다양한 면역원을 제조하기 위해, env 유전자만이 아닌 다른 그룹 M 컨센서스 유전자를 컨스트럭트할 필요가 있을 것이다. "컨센서스 gag, pol 및 nef 유전자의 컨센서스" (M.con.gag, M.con.pol 및 M.con.nef)를 디자인하였다. 서브타입 컨센서스 pol 유전자를 제조하기 위해, 서브타입 C 컨센서스 pol 유전자 (C.con.pol)도 디자인하였다. M.con.gag, M.con.pol, M.con.nef 및 C.con.pol 유전자의 코돈은 고발현된 인간 유전자의 코돈 사용에 기초하여 최적화하였다 (핵산 및 아미노산 서열에 대해서도 19 참조).

[0137] 실시예 7

[0138] 합성 서브타입 B 컨센서스 gag 및 env 유전자

[0139] 구체적인 실험

[0140] 서브타입 B 컨센서스 gag 및 env 서열은 37개 및 137개의 등가의 HIV-1 균주로부터 각각 유도하였으며, 포유동물 세포 발현에 대해 코돈-사용을 최적화하고, 합성하였다 (도 20a 및 20b). 최적의 발현을 보증하기 위해, 코작 서열 (GCCGCCGCC)을 개시 코돈의 바로 업스트림에 삽입하였다. 전장 env 유전자에 추가로, gp41 막-스패닝 도메인 (IVNR)의 바로 뒤에 중지 코돈을 삽입함으로써 절단된 env 유전자를 제조하였으며, 이에 의해 gp145 유전자가 제조되었다. 인 비트로 전사/번역 시스템에서, 유전자를 삽입성 (integrity)에 대해 테스트하였으며, 포유동물 세포에서 발현시켰다. (서브타입 B 컨센서스 Gag 및 Env 서열이 도 20c 및 20d에 각각 표시되어 있다)

[0141] 서브타입 B 컨센서스 외피가 융합과 삽입을 유도하는 것이 가능한지 여부를 확인하기 위해, gp160 및 gp145 유전자를 HIV-1/SG3 Δenv 프로바이러스로 동시-형질감염시켰으며, JC53-BL 세포 분석을 이용해 제조된 유사비리온의 감염도를 측정하였다. JC53-BL 세포는 CD4 및 HIV-1 공-수용체 CCR5 및 CXCR4를 높은 레벨로 발현하는 HeLa 세포의 유도체이다. 또한, 이들은 HIV-1 LTR로부터 발현되는 루시페라아제 및 β-갈락토시다아제의 리포터 카세트 포함한다. 리포터 유전자의 발현은 HIV-Tat의 생산에 의존한다. 상세하게는, 세포를 24-웰 플레이트에 접종하여, 37℃에서 24시간 동안 배양한 뒤 DEAE-덱스트란을 처리해 37℃에서 30분 동안 반응시켰다. 바이러스를 1% DMEM에 순차적으로 희석한 뒤, DEAE-덱스트란에서 배양되고 있는 세포에 첨가하고, 추가의 세포 배지 500 μl를 각 웰에 첨가하고 난 뒤 37℃에서 3시간 동안 반응시켰다. 37℃에서 마지막 48시간 반응시키고 난 뒤, 세포를 고정시키고, 푸른색 초점들 (foci)을 현미경으로 계수하기 위해 PBS로 덮었다. 백그라운드를 확인하기 위해 사용된, 모의-감염된 웰에 대한 계수를, 각 샘플 웰의 계수로부터 뺐다. 또한, JC53-BL 분석을 약

간 변형시켜 공-수용체 사용 및 외피 중화 민감도를 확인하였다.

[0142] 서브타입 B 컨센서스 Gag 단백질이 Env 당단백질에 결합하는 바이러스-유사 입자 (virus-like particles, VLPs)를 생산할 수 있는지 확인하기 위해, 293T 세포를 서브타입 B 컨센서스 gag 및 env 유전자로 동시-형질감염시켰다. 형질감염 후 48시간에, VLPs를 포함하는 세포 상등액을 수집하였으며, 테이블탑 원심분리기로澄清화하고, 0.2 mM 필터로 필터하였으며, 20% 수크로즈 쿠션을 통해 펠렛화하였다. 상기 VLP 펠렛을 PBS로 재현탁하고 20-60% 연속 농도구배의 수크로즈에 트랜스퍼하였다. 100,000 ×g의 속도로 밤새 원심분리하고 난 뒤, 0.5 ml 분획을 수득하여 p24 함량을 분석하였다. 또한, 각 분획의 굴절률을 측정하였다. VLPs에 대한 농도가 정확하고 높은 레벨의 p24를 포함하는 분획을 모든 뒤 마지막으로 펠렛화하였다. VLP-포함 펠렛을 PBS에 재현탁하고 4-20% SDS-PAGE 젤에 로딩하였다. 단백질을 PVDF 막에 트랜스퍼한 뒤 서브타입 B HIV-1 감염된 환자의 혈청으로 탐침시켰다.

[0143] 결과

[0144] 코돈-사용 최적화된, 서브타입 B 컨센서스 외피 (gp160, gp145) 및 gag 유전자는 포유동물 세포에서 높은 레벨의 당단백질을 발현한다 (도 21).

[0145] 서브타입 B gp160 및 gp145 당단백질은 바이러스 입자에 효율적으로 결합되었다. 수크로즈-정제된 유사비리온의 웨스턴 블랏 분석에서는, 컨센서스 B 외피 삽입 레벨이 rev-의존성 증가 외피의 삽입에 비해 최소한 5배 높음을 보여주었다 (도 23a). 서브타입 B 컨센서스 gp160 또는 gp145 외피로 유사타입된 비리온은 rev-의존성 증가 외피를 포함하는 유사비리온에 비해 훨씬 더 감염성이 있었다 (도 23b).

[0146] 서브타입 B 컨센서스 외피는 CD4 포함 표적 세포로 삽입되기 위해 공-수용체로서 CCR5를 이용한다 (도 22).

[0147] 서브타입 B 컨센서스 gp160 외피를 포함하는 유사비리온의 감염도는, HIV-1 서브타입 B 감염된 환자의 혈장에 의해 중화되었으며 (도 24c), 모노클로날 항체를 중화시켰다 (도 24a). 이는, 서브타입 B 합성 컨센서스 B 외피의 전체적인 구조가 천연 HIV-1 Env 당단백질과 비슷하고, 공통적인 중화 에피토프가 완전한 상태로 남아 있음을 제시해 준다. 도 24b 및 24d에서는 서브타입 B 대조군 외피의 중화 프로파일 (NL4.3 Env)을 보여준다.

[0148] 서브타입 B 컨센서스 Gag 단백질은 세포막으로부터 발아 (bud)할 수 있으며, 바이러스-유사 입자를 형성할 수 있다 (도 25a). 코돈-최적화된 서브타입 B 컨센서스 gag 및 gp160 유전자의 동시-형질감염에서는 결합된 외피를 가지는 VLPs를 생산한다 (도 25b).

[0149] 결론

[0150] 합성 서브타입 B 컨센서스 env 및 gag 유전자는 증가의 서브타입 B Env 및 Gag 단백질에 대한 구조, 기능 및 항원성이 유사한 바이러스 단백질을 발현한다. 서브타입 B 컨센서스 유전자에 기초한 면역원이, CTL을 유발하고, 광범위한 HIV-1 분리물 세트에 대해 방어적인 면역 반응을 중화할 것임이 예견된다.

[0151] * * *

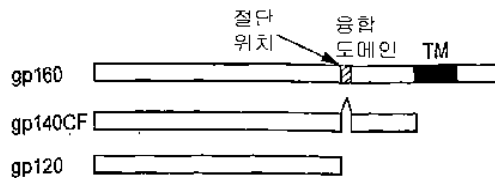
[0152] 모든 문헌 및 상기에 기재된 다른 정보 원천은 그 전체가 참조로서 본원에 삽입된다. 또한 삽입되는 참조문헌은 Liao et al, J. Virol. 78:5270 (2004)이다.

도면

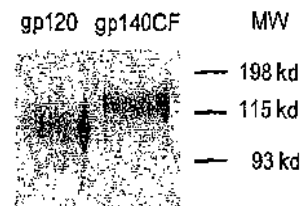
도면1a

MRVMGIORNCQHLLRWGTMILGMLMICSAAENLWTVYVGVVWKEANTLLFCASDAKAYDTEVHNVWAT
HACVPTDPNPQEIIVLENVTEFNFMWKNMVEQNHEDIISLWDQSLAPCVKLTPLCVTLNCTNVRVSSNG
V1
TETDNEETKNCSEFNITTELDRKKQVYALFYRLDVPIDDKNSSEISGKNSSEYRNLNCTNSAITQACP
V2
KVSFEPIPIHYCAPAGFAILKCNCKKFNCTGCPCKWVSTVQCTHGKIPWVSTQLLINGSLAEEIIIRSEN
V3
ITNNAKTIIVQINESVEINCTRPNNTRKSIHIGPGQAFYATGETIGDIRQAHONISRTKWNKTLLQQVAK
V4
KLREHFNKTIIFKPSSGGDLEITTHSENCGEFFYCNSTGLFNSTMMFNCTYMENGTCKDNSETITLPCR
V5
IKQILNMWQGVGOAMYAPPIEGKITCKSNITGLLLLTRDGGNNSNKKKTETTFPGGGMDRDNWRSELYKYK
VVKLEPLGVAPTAKARRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVOARQLLSGIVQQQSNLLR
AIEAQHLLQLTVWGIKQLOARVLAVERYLKDQQLLGWCCSGKLICTTNVPWSSWSNKSQDEIWDNMT
WMEWEREISNYTDIIYRLIEESQOQEKNEQELLALDKWASLWNVFDTITNWLWYIKIFINIVGGLIGIRI
VFAVLSIVNRVRQGYSPLSFOTLIIPNPRGPDPRPEGIEEGEGEQGRDRSIRLVNGFLALAWDDLRSICLFS
YHRLRDFILIAARTVELIGRRSLRGLOKWEALKYLGNLLQYWQQLKNSAISLLDTTATAVAEGTDRVI
EIVQACRAILNIPRIRQGLERALL

도면1b



도면1c



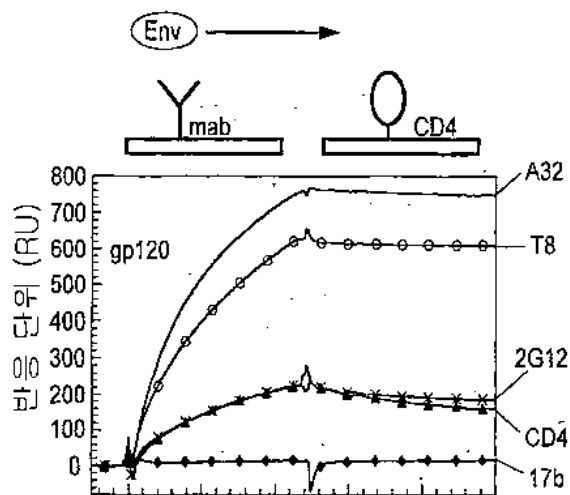
도면1d

CON6.env (그룹 M env 컨세서스. 이는 98CN006 바이러스의 env 유전자에 있는 5개의 변이 부위를 포함하나 공유 도메인에 있는 부위는 아직 포함하지 않음)

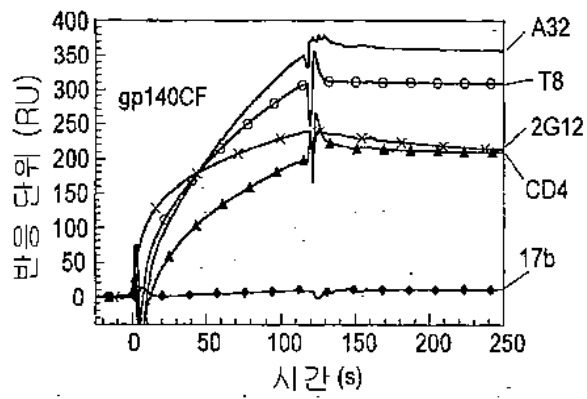
```

GCCACCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCATGATC
CTGGGCATGCTGATGATCTGCTCCGCGCGCGAGAACCTGTGGGTGACCGTGTAATACGGC
GTGCCCCGTGTGGAAGGAGGCCAACACACCCCTGTTCTGCGCCTCCGACGCCAAGGCCCTAC
GACACCGAGGTGCACAACCTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTG
GAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAG
CTGACCCCCCTGTGCGTGACCTGAACCTGCACCAACGTGCGCAACGTGTCTCCAACGGC
ACCGAGACCGACAACGAGGAGATCAAGAACTGCTCCTTCAACATCACACCGAGCTGCGC
GACAAGAAGCAGAAGGTGTACGCCCCTGTTCTACCGCTGGACGTGGTGGCCATCGACGAC
AAGAACTCCTCCGAGATCTCCGGCAAGAACTCCTCCGAGTACTACCGCCTGATCAACTGC
AACACCTCCGCTATCACCCAGGCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCAC
TACTGCGCCCCCGCGCGCTTCGCCATCTCGAAGTGCAACGACAAGAAGTCAACCGGCACC
GGCCCCCTGCAAGAACGTGTCCACCGTGCACTGCACCCACGGCATCAAGCCCGTGGTGTCC
ACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCGCTCCGAGAAC
ATCACCAACAACGCCAAGACCATCATCTGTCAGCTGAACGAGTCCGTGGAGATCAACTGC
ACCCGCCCCCAACAACAACACCGCAAGTCCATCCACATCGGCCCGCGCCAGGCTTCTAC
GCCACCGCGGAGATCATCGGCGACATCCGCCAGGCCCCACTGCAACATCTCCCGCACCAAG
TGGAACAAGACCCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACC
ATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACACCCACTCCTTCAACTGC
GGCGGCGAGTTCTTCTACTGCAACACCTCCGGCTGTTCAACTCCACCTGGATGTTCAAC
GGCAGCTACATGTTCAACGGCACCAAGGACAACCTCCGAGACCATCACCTGCGCTGCGGC
ATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATC
GAGGGCAAGATCACTGCAAGTCCAACATCACCGGCTGCTGCTGACCCGCGACGGCGGC
AACAACTCCAACAACAAGACCGAGACCTTCCGCCCGCGCGCGGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCC
CCCACCAAGGCCAAGCGCGCGTGTGGAGCGCGAGAAGCGCGCGTGGGCATCGCGCGC
GTGTTCTTGGGCTTCTTGGGCGCGCGCGCTCCACATGGGCGCGCGCTCCATCACCTTG
ACCGTGCAAGGCCCGCGAGCTGTCTGTCGGGCATCGTGCAGCAGCTCCAACTGCTGCGC
GCCATCGAGGCCCGAGCAGCTGCTGTCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAG
GCCCCGCTGCTGCGCGTGGAGCGCTACCTGAAGGACCGAGCAGCTGCTGGGCATCTGGGGC
TGCTCCGCAAGCTGATCTGCACCAACCAACGTGCCCTGGAACCTCCTCTGTTCCAACAAG
TCCAGGACGAGATC TGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCTCCAAC
TACACCGCATCATCTACCCCTGATCGAGGAGTCCCAAGAACGACGAGAGAAGAAGAG
CAGGAGCTGCTGGGCTTGGCAAGTGGGCTCCCTGTGGAACCTGTTTCGACATCACCAAC
TGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCTGATCGGCTGCGCATC
GTGTTGCGCGTGTGTCCATCGTGAACCGCTGCGGCCAGGGCTACTCCCCCTGTCTTC
CAGACCTGATCCCAACCCCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGC
GGCGAGCAGGCGCGCGACCGCTCCATCCGCTGGTGAACGGCTTCTGGGCTTGGCTGG
GACGACCTGCGCTCCTGTGCTGTTCTCCTACACCGCTGCGCGACTTCATCTCTGATC
GCCGCCCGCACCGTGGAGCTGCTGGGCGCGCTCCTTGGCGGCTGCGAAGGCTGG
GAGGCCCTGAAGTACCTGGGCAACCTGCTGCAGTACTGGGCGCAGGAGCTGAAGAATCC
GCCATCTCCCTGCTGGACACACCGCCATCGCCGTGGCGGAGGGCACCGCCGCTGATC
GAGATCGTGCAGCGCGCTGCCGCGCATCTGAACATCCCCCGCGCATCGCGCAGGGC
CTGGAGCCCGCCCTGCTGTAA
    
```

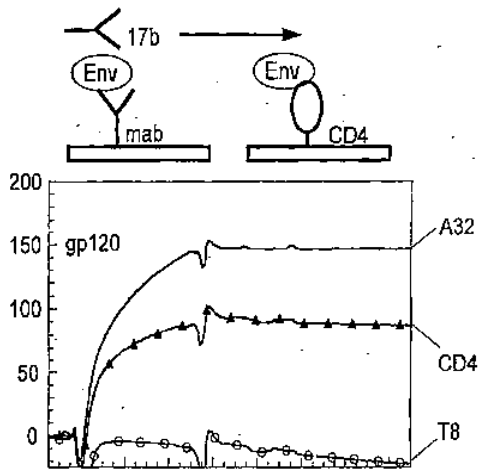
도면2a



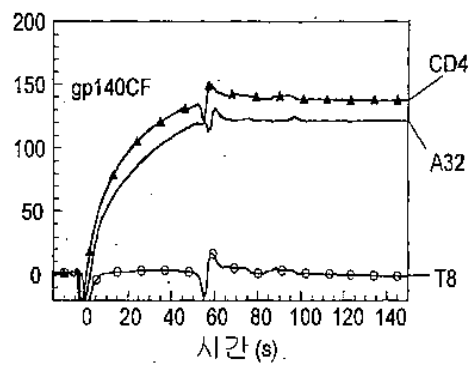
도면2b



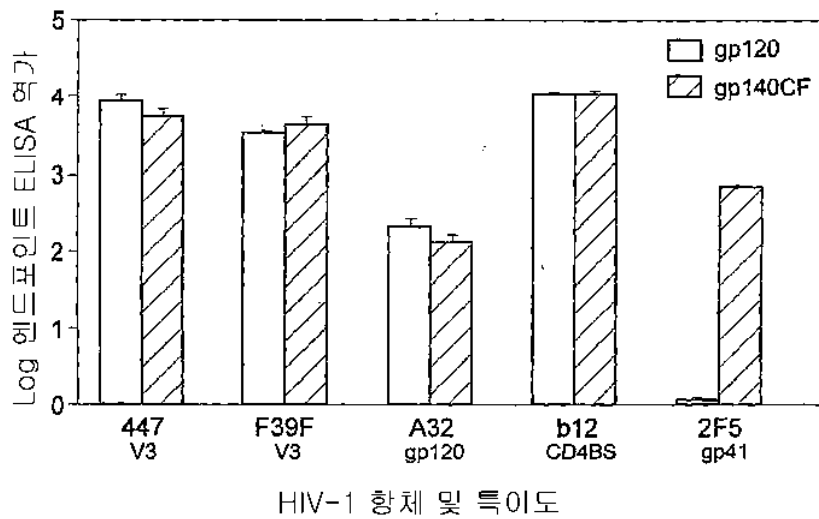
도면2c



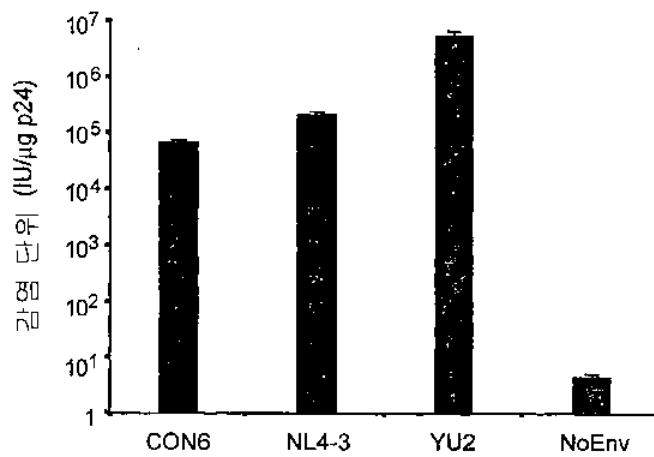
도면2d



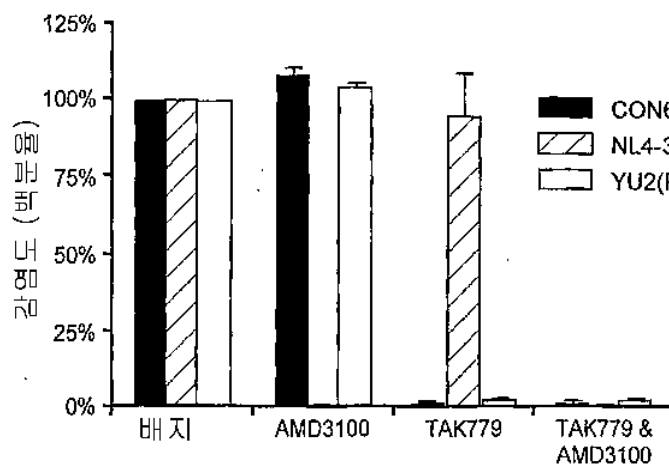
도면2e



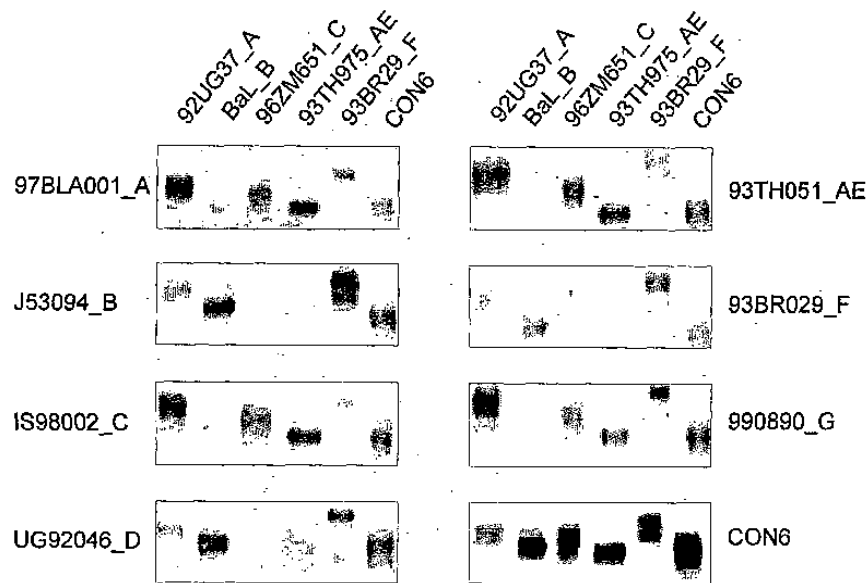
도면3a



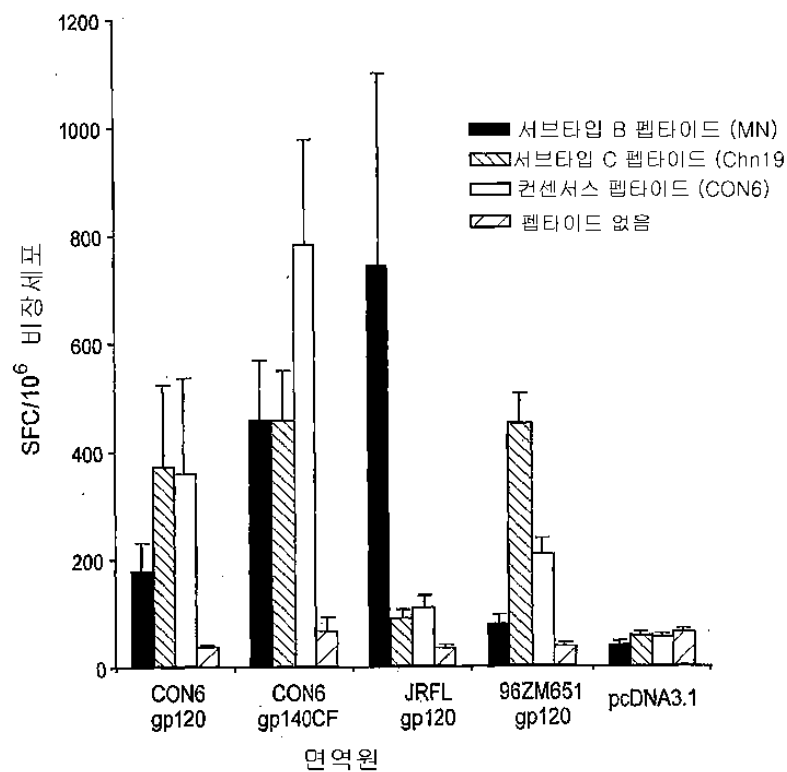
도면3b



도면4



도면5



도면6b

C.con.env (서브타입 C 컨센서스 env. 이 아미노산은
Los Alamos Database August 2002와 다름)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT
CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTGGTGGSCA
ACCTGTGGGTGACCGTGACTACGGCGTGCCCGTGGAAGGAGGCCAAG
ACCACCTGTCTGCGCTCCGACGCCAAGGCCTACGAGAAGG AGGTGCA
CAACGTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCCAGG
AGATGGTGTCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACCAC
ATGCTGACACGATGCAACGAGGACATCATCTCCCTGTGGGACCACTCCCT
GAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAACTGCCGCA
ACGTGACCAACGCCACCAACAACAACCTACAACGAGGAGATCAAG AACTGC
TCCTTCAACATCACCCAGAGCTGCGCGACAAGAAGAAGAGGTGTACGC
CCTGTTCTACCGCTGGACATCGTGCCCTGAAACGAGAACTCCTCCGAGT
ACCGCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCAAG
GTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCCGCTACGC
CATCTGAAGTGCAACAACAAGACCTTCAACGGCACCGGCCCTG CAACA
ACGTGTCCACCGTGGCTGACCCACGGCATCAAGCCCGTGGTGTCCACC
CAGTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCGCTC
CGACAACCTGACCAACAACGCCAAGACCATCATCGTGACCTGAACGAGT
CCGTGGAGATCGTGTGACCCGCCCAACAACAACCCCGCAAGTCCATC
CGCATCGGCCCGGCCGAGACCTTCTACGCCACCGCGCATCATCG GCGA
CATCGGCCAGGCCCATGCAACATCTCCGAGGACAAGTGAACAAGACCC
TGACGCGCGTGTCCAAGAAGCTGAAGGAGCACTTCCCAACAACACCATC
AAGTTCGAGCCCTCCTCCGGCGCGACCTGGAGATCACCAACCCACTCCTT
CAACTGCGCGCGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACT
CCACCTACAACAACAACCAACTCCAACCTCCACCATCACCTGCCCTGCG
CGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCGCGCCATGTA
CGCCCCCCCCCATCGCCGCAACATCACCTGCAAGTCCAACATCACCGGCC
TGCTGCTGACCCGCGACGGCGCGCAAGAAGAACACACCGAGATCTTCCGC
CCCGCGGGCGCGCATGCGCGACAACCTGCGCTCCGAGCTGTACAAGTA
CAAGGTGGTGGAGATCAAGCCCTGGGCGTGGCCCCACCAAGGCCAA GC
GCCGCGTGTGGAGCGCGAGAAGCGCGCGTGGGCATCGGCGCGCTGTTC
CTGGGCTTCTGGGCGCGCGCGCTCCACCATGGGCGCGCCCTCCATCAC
CCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGACGACGAGT
CCAACTGTGCGCGCATCGAGGCCAGCAGCATGTGCGAGCTGACC
GTGTGGGGCATCAAGCAGCTGCAGACCCGCGTGTGGCCATCGAGCGCTA
CCTGAAGGACCAAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGA
TCTGCACACCGCGCTGCCCTGGAACCTCCTCCTGCTCCAACAAGTCCAG
GAGGACATCTGGGACAACATGACCTGGATGCGAGTGGGACCGGAGATCTC
CAACTACACCGACACCATCTACCGCTGCTGGAGGACTCCAGAACCAGC
AGGAGAAGAACCAGAAGGACCTGCTGCCCTGGACTCCTGGAAGAACCTG
TGGAACTGGTTCGACATCACCAACTGGCTGTGTGTACATCAAGATCTTCAT
CATGATCGTGGGCGGCTGATCGGCTGCGCATCATCTTCGCGGTGCTGT
CCATCGTGAAACCGCGTGGCGCAGGGCTACTCCCCCTGTCTTCCAGACC
CTGACCCCCAACCCCGCGGCCCGACCGCTGGGCCGATCGAGGAGGA
GGGCGCGAGCAGGACCGCGACCGCTCCATCCGCTGTGTCCGGCTTCC
TGGCCCTGGCTGGGACGACCTGCGCTCCCTGTGCTGTCTTCTTACCAC
CGCTGCGCACTTCTCTCTGTGGCGGCCCGCGCTGGAGCTGTGCGG
CCGCTCCTCCTGCGCGGCTGCGAGCGCGCTGGGAGGCCCTGAAGTACC
TGGCTCCCTGTGTGAGTACTGGGCGTGGAGCTGAAGAAGTCCGCCATC
TCCCTGCTGGACACCATCGCCATCGCGTGGCGGAGGACCGACCGCAT
CATCGAGCTGATCCAGCGCATCTGCCGCGCCATCCGCAACATCCCCCGCC
GCATCCGCGAGGCTTGGAGCGGCTTCAAGTA

도면6c

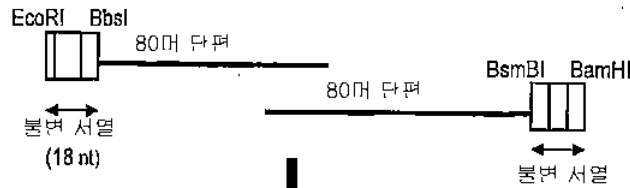
C.anc.env (서브타입 C 조상 env)
 MRVMGILRNCQQWIIWGIIGFWMLMICSVVGNLMVTYVYGVVVKAKTTILFCASDAKAYEREVHNWAT
 HACVPTDNPQEMVLENTENFMWKNMDVMQMHEDIISLWDQSLKPCVKLTPLCVTTINCTNVTINATNNT
 YNGEMKNCNFNITTELRDKKKRYALFYRLDIVPLNENSSEYRLINGNTSAITQACPKVSPDPIPIHYCA
 PAGYAILKCNKNTFNGTGPCNNVSTVQCTHGKIPVSTQLLNGLAEDEIIIRSENLTDAKTIIVQLN
 ESVEIVCTRPNNNTKSMRIGPGQTFYATGDIIGDIRQAHCNISEDKNWKTLOOVAEKLKGFHPNKTITF
 EPSSGGDLLEITTHSFNCRGEFFYCNSTKLENSTYNNNTNSNTITLPCRIKQILNMWQGVQAMYPAPLA
 GNITCKSNITGLLLRDGGKENTTETFRPGGDMRDNWRSELYKYKVEIKPLGVAPTEAKRRVVEREKR
 AVGLGAVFLGFLGAAGSTMGASITLTVAQRQLLSGIVQQQSNLLRAIEAQCHMLQLITYWGIKQLQARVL
 AMERYLKDQQLLGIVGCSGKLICTTAVPWNSSWSNKSLLDDIWDNMTWMEWDRFISNYTDTIYRLLEESQN
 QQEKNEQDLALDSWENLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSTVNRVROGYSPLSFQTLT
 ENPRGPRLERIIEEGGEQDRDRSIRLVSGFLALANDLRSCLFSYHRLRDFILIAARTVELLGRSLR
 GLQRGWEALKYLGSLVQYWQGQELKKSAISLIDTIAIAVAEGTDRIEVWQACRAILNIPRRIRQGFEAA
 LL

도면6d

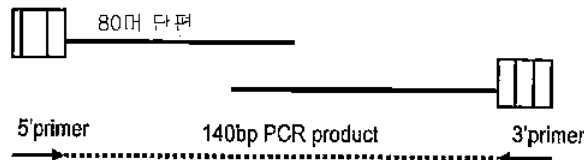
C.con.env (서버타임 C 컨센서스 env)
 MRVWGILRNCQQWIIWGIILGFWMIMICNVVGNLWVTYYGVVPVWKEA KTLFCASDAKAYEKEVHNWAT
 HACVPTDNPQEMVLENTENFNMWKNMDVQOMHEDIISLWDQSLKPCVKLTPLCVTLNCRNVTNATNT
 YNEEIKNCSFNITTELRDKKKVYALFYRLDIVPLNENSSSEYRLINCNTSALTQACPKVSPDPPIPIHYCA
 PAGYAILKCNKNTENGTPCANNVSTVQCTHGIKPVVSTQLLNGSLAEEEEIIRSENLTNNAKTIIIVHLN
 ESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIIROHCNISEDKWNTLQRVSKLKEHFPNKTIKP
 EPSSGGDLIITHSFNCRGEFFYCNTSKLFNSTYNNNTNSNTITLPCRIKQIINWQEVGRAMYAPPIA
 GNITCKSNITGLLLTRDGGKNTTEIFRPGGDMRDNWRSELYKYKVVEIKPLGVAPTAKARRVVEREKR
 AVGIGAVFLGFLGAAGSTMGAASITLTVOARQLLSGIVQOQSNLRAIEAQOQMLQLTVMGI KOLQTRVL
 AIERYLKDQQLLGIWCGSKLICTTAVPWNSSWSNKSQEDINDNMTWMQWDREISNYTDTIYRLLEDSON
 QQEKNEKDILALDSWKNLWNWFDTINWLWYIKIFIMIVGGLIGLRIIFAVLSTVNRVROGVSPLSFOTLT
 PNPRGPDRLGRIEEEGEGEQDRDRSIRLVSGFLALAWDDIRSLCLFSYHRLRDFILVAARAVELLGRSSIR
 GLQRGWEALKYGLSVQYWGLELKKSAISILDTI AIAVARGTDRIIELIQRICRAIRNIPRIIRQGFEEA
 LQ

도면6e

5' 말단에 있는 불변 서열로 3' 말단에 있는 20개 자기를 중첩시킴으로써
80-머 단편에 있는 전체 유전자 합성



불변 서열에 상보적인 프라이머를 이용해 5'에서 3'으로 단계별로
PCR을 수행함으로써 쌍으로된 80머 올리고를 연결

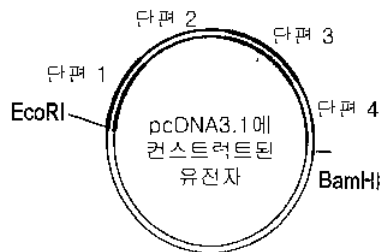


108bp PCR 단편을 pGEM-T에 클로닝한 뒤 서열분석.

적당한 서열을 가진 클론은 2개의 제한효소로 절단될 것이다.

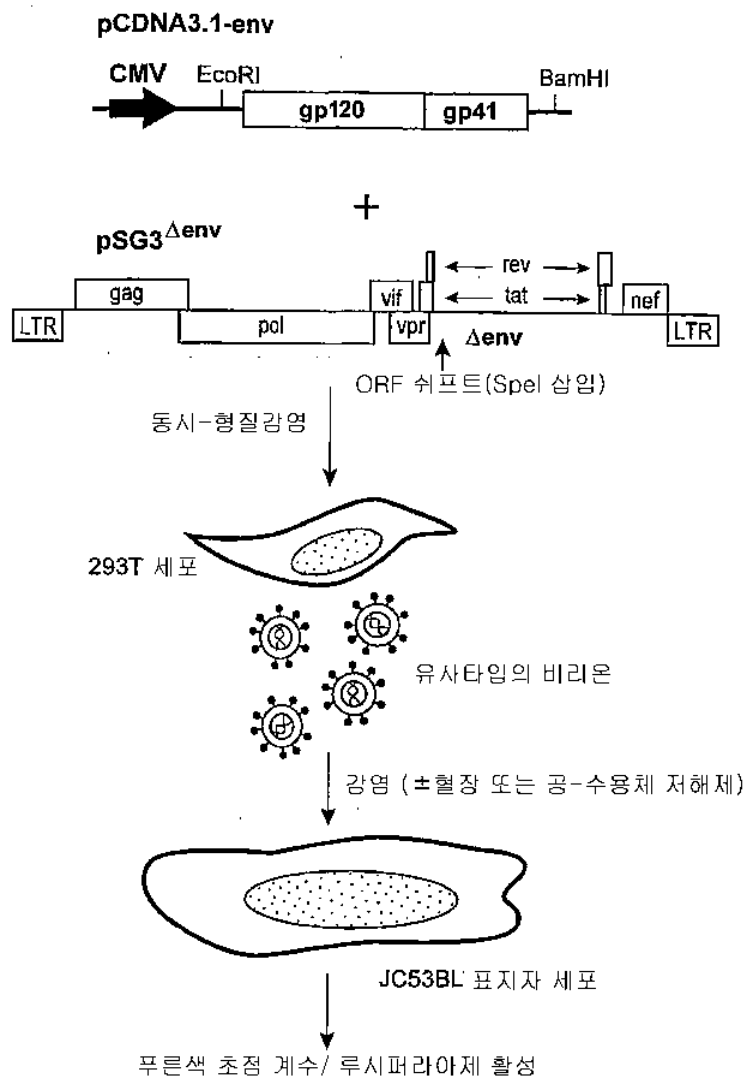
4개의 단편은 유전자의 5'에서 3' 말단으로 단계별로 pcDNA3.1과 함께 라이게이션될 것이다.

pcDNA3.1과 라이게이션되는 단편 (1-4는 5'에서 3'의 순이다)	단편을 절단하는데 사용되는 제한효소
Fragment 1	EcoRI/BsmBI
Fragment 2	BbsI/BsmBI
Fragment 3	BbsI/BsmBI
Fragment 4	BbsI/BamHI
pcDNA3.1	EcoRI/BamHI



전체 유전자가 pcDNA3.1에 클로닝될 때까지 5'에서 3' 방향으로
단계별로 라이게이션이 반복될 것이다.

도면7



도면8

4RVAGILRNQOQWIMGILGFNALMICNVVGNLWTVIYGVFVKEAKTTLFCAADAKAYEKEVHNWTHACHVPTDNPQEALENVTEFNFNKNKNDVDOHEDIIISLWDQSLKPCVKLTPLC
 +
 4RVAGILRNQOQWIMGILGFNALMICSVVGNLWTVIYGVFVKEAKTTLFCAADAKAYEKEVHNWTHACHVPTDNPQEALENVTEFNFNKNKNDVDOHEDIIISLWDQSLKPCVKLTPLC

V1
 VTLNCRNVTNATNNTYNEIEIKNCSENITTELRDKKKVVALFTRLDIVPELHENSSEYELNCHNTSAITQACPVSDPPIHYCAPAGYAILKCNKKTNGTGPCNNVSTVQCTHGIKPVVSTQL
 +
 ZILNCTNVTNATNNTYNGEKNCSNITTELRDKKKVVALFYRLDIVPELHENSSEYELNCHNTSAITQACPVSDPPIHYCAPAGYAILKCNKKTNGTGPCNNVSTVQCTHGIKPVVSTQL

V2
 4ANGSLAEZBIIIRSENLTNNAKTIIVHLNESVEIVCTRENNTKRSIRIGPGQTFYATGDIIGDROAHCHLSEDAKWKILQVSKLKEHFPNKTIFEPSGCELEITHSFNCRGEFFYCN
 +
 4ANGSLAEZBIIIRSENLTNNAKTIIVQLNESVELTCTRENNTKRSIRIGPGQTFYATGDIIGDROAHCHLSEDAKWKILQVSKLKEHFPNKTIFEPSGCELEITHSFNCRGEFFYCN

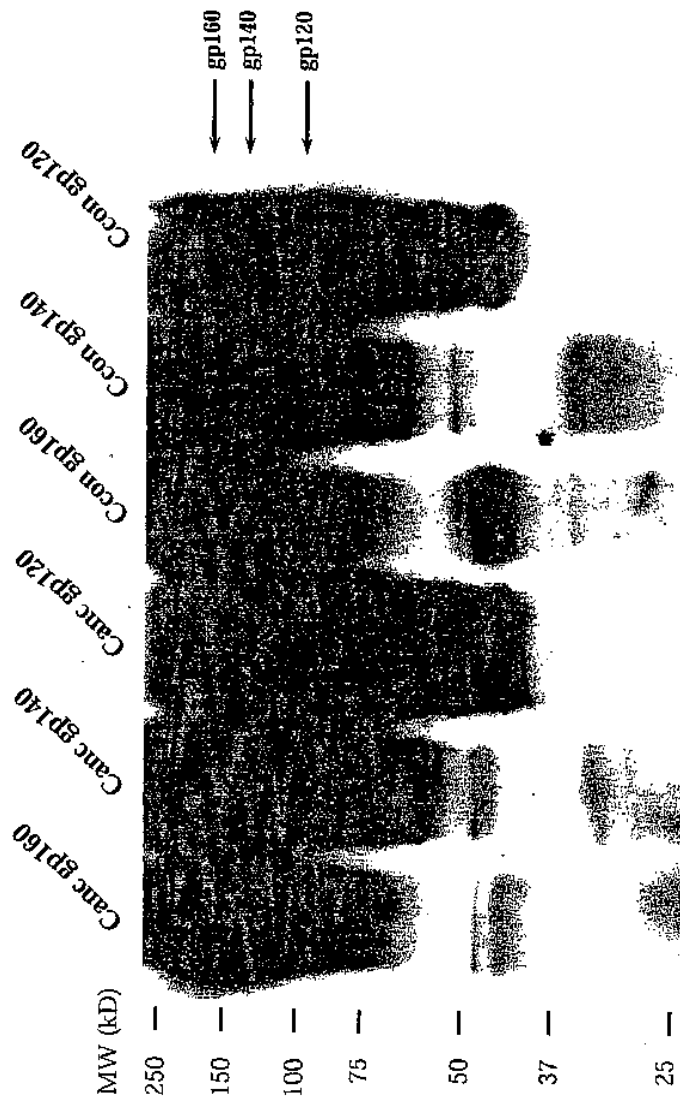
V3
V4
 4EKLFSNTNNNTNSNTITLFCRIKQITNNQGVGRAMYAPPIAGNITCKSNITGLLTRDGGKNTTEIFRPGGGDMDNDWASELYKYKVVEINPLGVAPTAKRRVYEREKRAVGLGAVFLG
 +
 4SKLFNSTYNGVNSNTITLFCRIKQITNNQGVQGMWAPPIAGNITCKSNITGLLTRDGGKNTTEIFRPGGGDMDNDWASELYKYKVVEINPLGVAPTAKRRVYEREKRAVGLGAVFLG
 gp120 ↑ gp41

 4LGAAGSTMGAASITLITVQARQLLSGIVQOQSNLLRAIEAQOHEMLQLTWGKIQLETRVLATERYLKDQQLGIWCCSKLICCTAVFNSNSWSKSCEDINDNWTWQDREISNHTDTITRLL
 +
 4LGAAGSTMGAASITLITVQARQLLSGIVQOQSNLLRAIEAQOHEMLQLTWGKIQLETRVLATERYLKDQQLGIWCCSKLICCTAVFNSNSWSKSCEDINDNWTWQDREISNHTDTITRLL

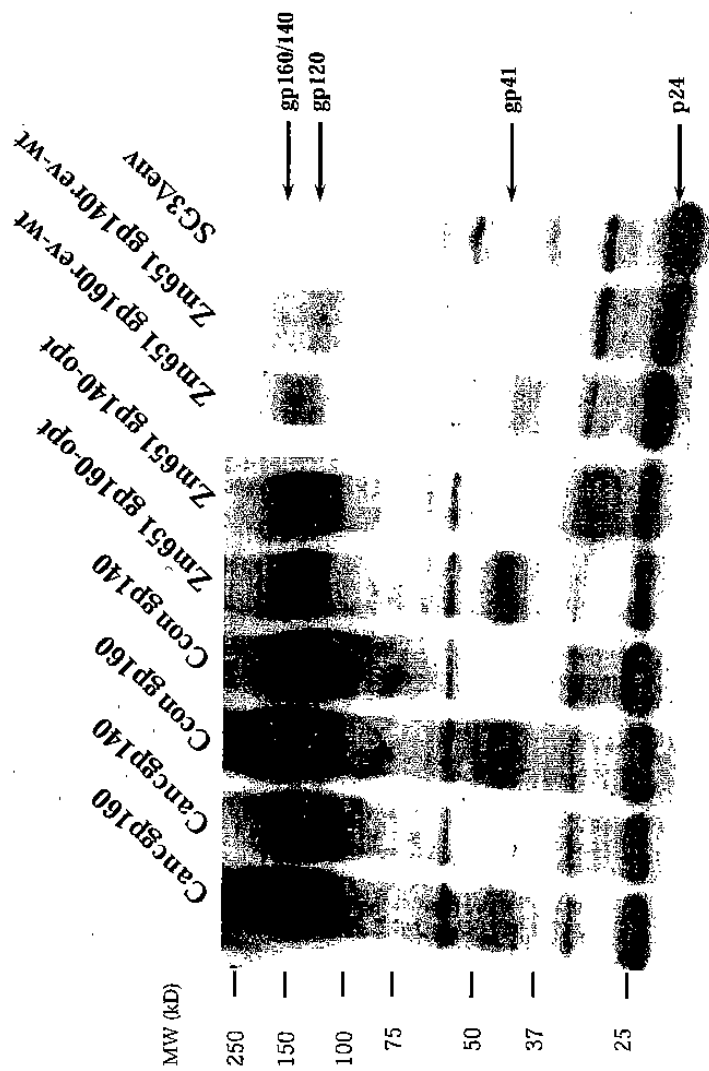
 4ESQOQOEKNEKNDILLALDSHKNLWNWEDIITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVQGYSELSFQTITPNRPGDRIERIEEGCEQDRDSIRLVSGFLALAWDDLRSICLFSYHRL
 +
 4ESQOQOEKNEQOLLALDSWENLWNWEDIITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVQGYSELSFQTITPNRPGDRIERIEEGCEQDRDSIRLVSGFLALAWDDLRSICLFSYHRL
 gp140 ↑

 4DFILVABRAVELIGRSSIRGLQRGWEALAYIGSLVQWGLEIKESAISLDDTIAIVAEGTDRILELIQRICRAINIPIRIRQGFZAAUQ 843
 +
 4DFILTAARTVELIGRSSIRGLQRGWEALAYIGSLVQWGLEIKESAISLDDTIAIVAEGTDRILEVVOACRAILINIPIRIRQGFZAAIL 843

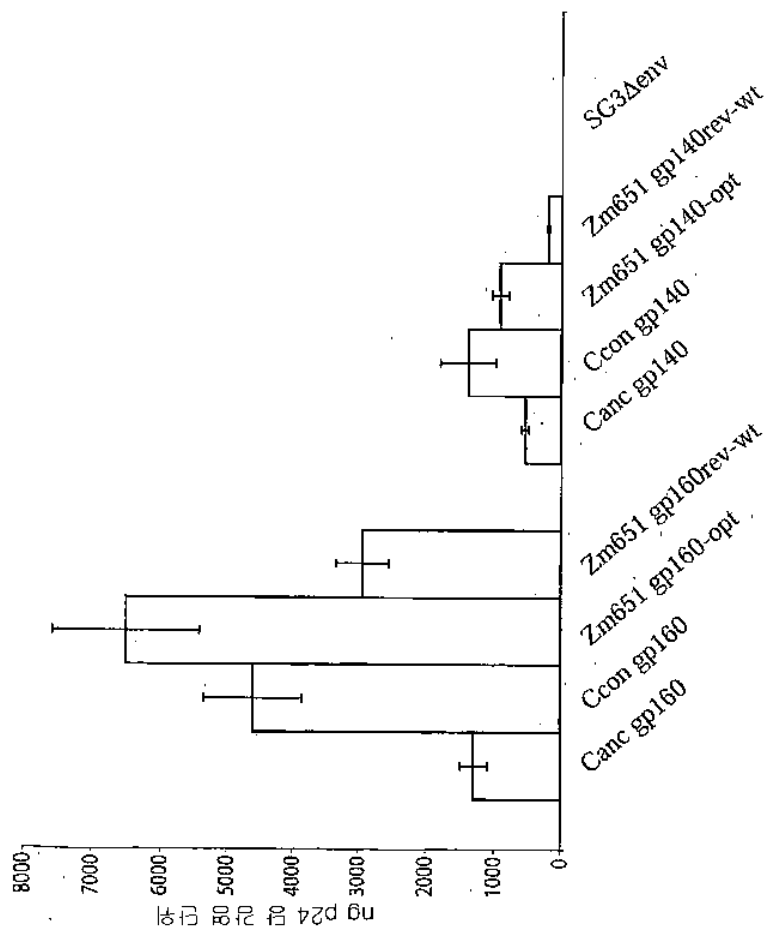
도면9



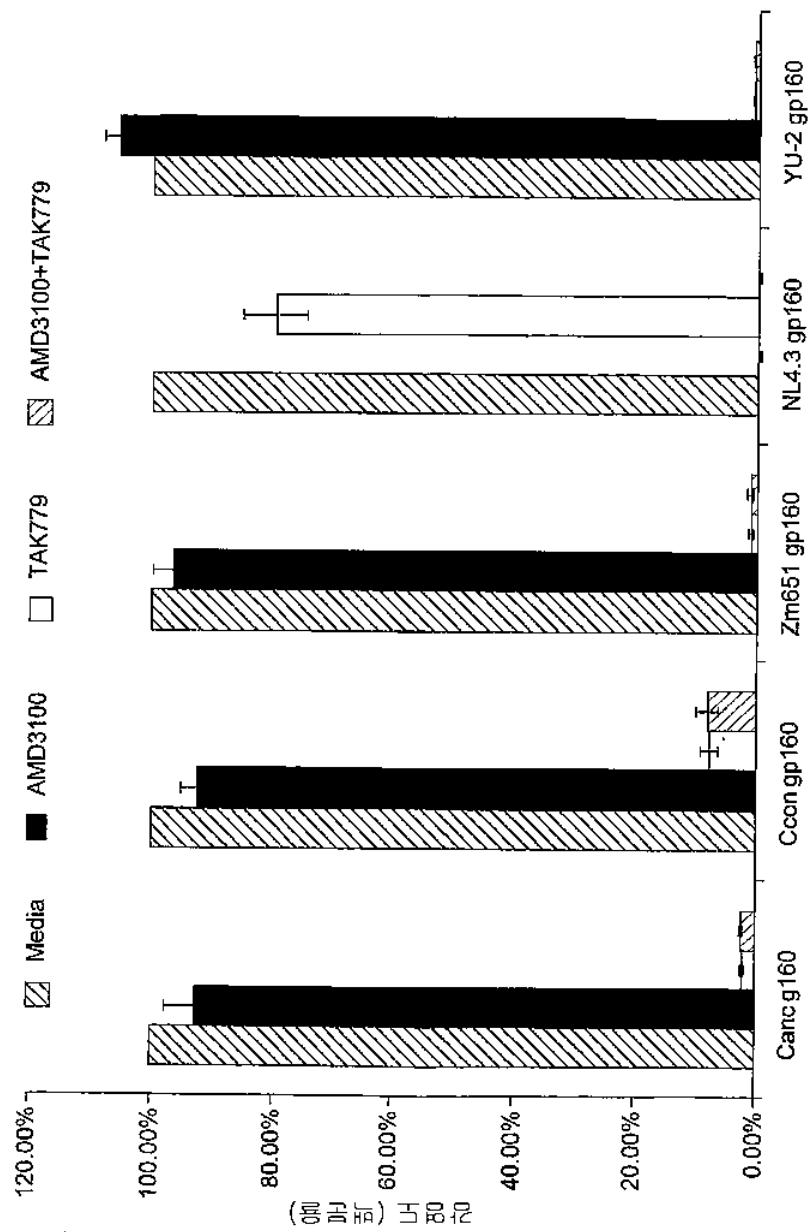
도면10a



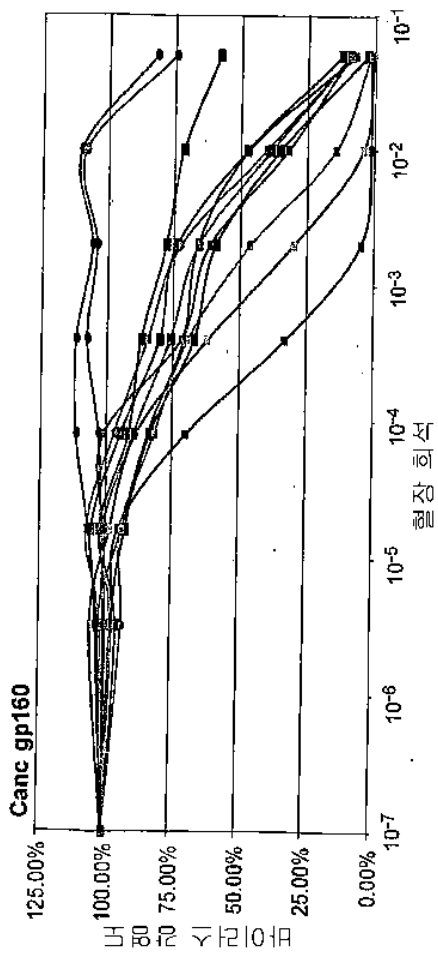
도면10b



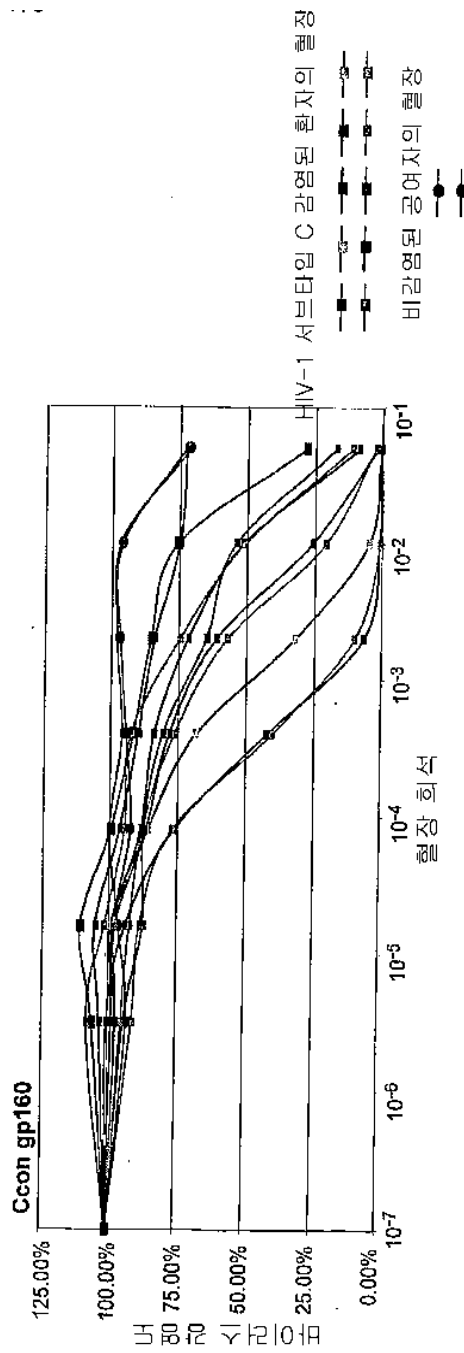
도면11



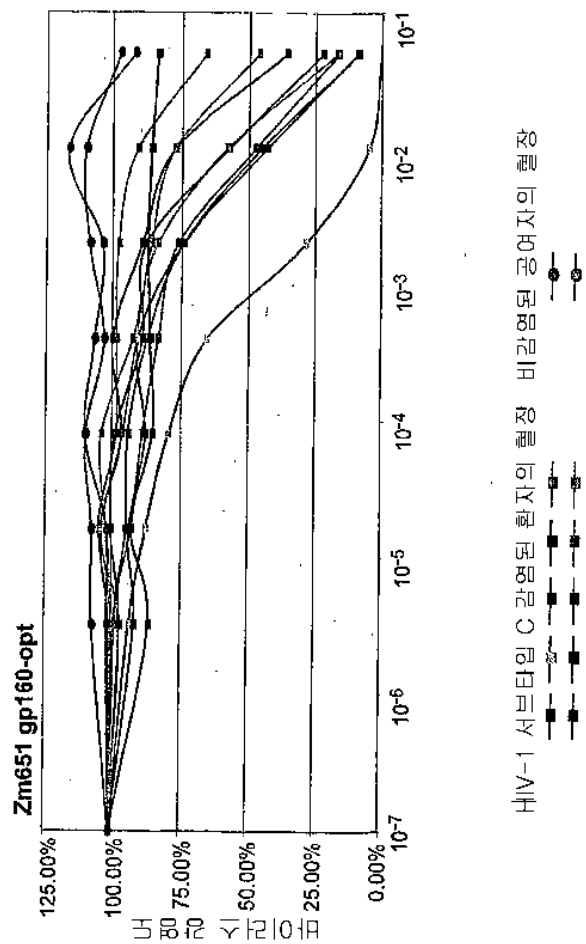
도면12a



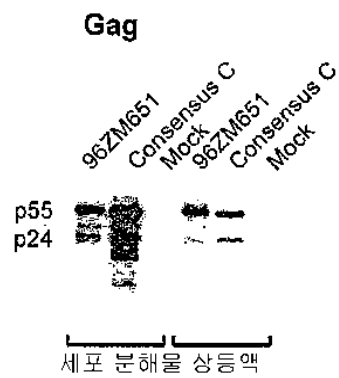
도면12b



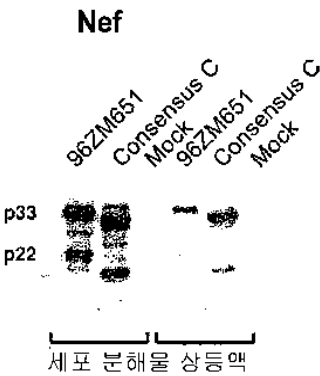
도면12c



도면13a



도면13b



도면13c

C.con.gag (서브타입 C 컨센서스 gag)

MGARASILRGGKLDITWEKIRLRPGGKKRYMIKHLVWASRELERPALNPGLLETSFGCKQIMKQLQPA
LOTGTEELRSLYNTVATLYCVHEKIHYEDTKEALDKIEEQNKSOQKTQQAFAAADGVSONYPI
VQNLGGQNVHQALSPRTLNAWVKVIEEKAFSPVIEPMFTALSEGATPDNLNTMLNTVGGHQAMQMLKDT
INEEAAEWDRLHPVHAGPTAPGQMRPRGSDIAGTTSTLQEQIAWMTSNPPVPVGDYIKRWIILGLNKIV
RMYSFVSILDIKQGPKEPFDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANPDCKTILRALGPGASLE
EMTACQGVGGFESHKARVLAEAMSQANNTNIMQRSNFKGPKRIVKCFNCGKEGHIARNCRAPRKKGCWK
CGKEGHQMKDCTERQANFLGKIWPESHKGRPCGNFLQSRPEPTAPPAESFREETTPA
PKQEPKDREPLTSLKSLFGSDPLSQ

도면13d

C.con.nef (서브타입 C 컨센서스 nef)
MGCKSKSSIVGWPVVRERIRRTEDPAEGVGAASQDLDKYGALTSNTATNNADCAWLEAQEEEEEV
GFEVRPQVPLRPMTYKAAFDSLFFLKEKGGLEGLIYSKKRQELDLWVYHTQGGFFDQWNYTFGPGVRYF
LTFGWCFKLVEVDPREVEEANEENCLLHPMSQHGMEDREVLKWKFDLSLARHWARELHPPEYYKDC

도면13e

C.con.gag (서브타입 C 컨센서스 gag. 공유 도메인에는 없음)

GCCGCGCCCATGGGCGCCCGCCAGCATCTGCGCGCGGCAAGCTGGACACCTGGGAGAGATCCGCC
 TGCGCCCGCGGCAAGAGCGCTACATGATCAAGCACCTGGTGTGGCCAGCCGGAGCTGGAGCGCTT
 CGCCCTGAACCCCGGCTGCTGGAGACGAGCGAGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCCGCC
 CTGCAGACCGGACCGAGGAGCTGCGCAGCTGTACAACACCGTGGCCACCTGTACTGCGTGACGAGA
 AGATCGAGGTGCGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAGAGCCAGCAGAA
 GACCCAGCAGGCGAGGCGCGCGGAGCGCAAGGTAGGCCAAGTAACCCATCGGTGCAGAACCTGCAG
 GGCCAGATGGTGCACAGGCCATCAGCCCGCCACCTGTAAGCCTGGGTGAAGTGATCGAGGAGAAGG
 CCTTCAGCCCGAGGTGATCCCATGTTCAACCGCTGAGCGGCGCCACCCCGAGGACCTTGAACAC
 CATGCTGAACACCGTGGGGGCCACCCAGGCCGCGCATGAGATGCTGAAGACACCCATCAACGAGAGGCC
 GCCGAGTGGACCGCCTGACCCCGTGCACGCCGCCCCATGCGCCCGCAGATGCGCGAGGCCCGCGT
 GCAGGACATCGCGGCAACCAAGCACCTGAGGAGCAGATCGCTGGATGACCCAGCAACCCCGCGT
 GCGGTGGGCGCATCTACAGCGCTGATCATCTGGGCGCTGAACAGATCGTGGCGATGTACAGCCQC
 GTAGCATCTGGACATCAGCAGGGGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTCAAGA
 CCTGCGCGCGGAGCAGGCCACCAAGAGCTGAAGAACTGGATGACCGACAACCTGCTGTGAGAACGC
 CAACCCGACTGCAAGACCATCTGCGGCCCTTGGGCCCGCGCGCGCTGGAGGATGATGACCGCC
 TGCCAGGGCGTGGCGGCCCGCAGCCACAGGCCCGCTGGCGGAGCCATGAGCCAGGCCACACAA
 CCAACATCATGATGAGCGGCAACTTCAAGGCCCGCAGCGCATCGTGAAGTGTCAACTGCGGCAA
 GGAGGCCACATCGCCCGCACTGCGCGGCCCGCGCAAGAGGGCTGCTGGAAGTGGGCAAGAGGGC
 CACCATGAGGAGCTGCACGAGCGCGAGGCCAATCTCTGGGCAAGATCTGGCCAGCCACAAGGGCC
 GCGCGGCAACTTCTCTGAGAGCGCGCCGAGGCCACCGCCCCCGCGGAGAGCTTCGGCTTCGAGGA
 GACCAACCCCGCCCAAGCAGGAGGCCCAAGGACCGCGAGGCCCTGACAGCCTGAGAGGCTGTTCGGC
 AGCGACCCCTGAGCCAGTAA

도면13f

C.con.nef (서브타입 C 컨센서스 nef. 공유 도메인에는 없음)
 GCCGCCCATGGCGGCAAGTGGAGCAAGAGCAGCATCGTGGGCTGGCCCCCGTGGCGCGAGCGCATCC
 GCCGACCCGAGCCCGCCGCGAGGGCGTGGGCGCGCCAGCCAGGACCTGGACAAGTACGGCGCCCTGAC
 CAGCAGCAACACCCGCCACCAACAAGCCGACTGGCCCTGGCTGGAGGCCCCAGGAGGAGAGAGGTG
 GGCTTCCCGGTGGCGCCCGCAGGTGCCCTGGCGCCCATGACCAAGGCCGCTTCGACCTGAGCTTCT
 TCCTGAAGGAGAGGGGCGCTGGAGGCGCTGATCTACAGCAAGAGCGCCAGGAGATCCTGGACCTGTG
 GGTGTACCAACCCAGGGCTTCTTCCCGACTGGCAGACCTACACCCCGGCCCCCGGCTGGCTACCCC
 CTGACCTTGGCTGGTGTCAAGCTGGTGGCCCTGGACCCCGGAGGTGGAGGAGCCAAAGAGGCG
 AGACAACCTGCTGTGCACCCCATGAGCCAGCAGCATGGAGGACGAGGACCGCGAGGTCTGAGTG
 GAAGTTGACAGCCACCTGGCCCGCCGCGCAATGGCCCGGAGCTGACCCCGAGTACTACAAGGACTGC
 TGA

도면14a

CONs.env (그룹 M 컨센서스 env 유전자. 이는 env 유전자에 있는
변이 부위에 대한 컨센서스 서열을 포함)

MRVRGIQRNCQHLWRWGTLLGLMLICSAENLWVTVYGVVPWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIIVLENTVENENWKNMVEQMHEDIISLNDQSLKPCVKLTPLCVTLNCTNVNVTN
TTNNTTEKGELKNCFSNITTEIRDKKQKVYALFYRLDVPIDDNNNSNYRLINCNTSAITQACPKVSF
EPIPIHYCAPAGFAILKCDKKFNGTGPCKNVSTVQCTHGLKPVVSTQLLNGSLABEEIIIRSENITNN
AKTIIVQLNESVEINCTRPNNNTRKSIIRIGPQAFYATGDIIGDIRQAHCHNISGTKNKTLQOVAKKLRE
HFNKTIIFKPSSGGDLBITTHSFNCRGEFFYCNISGLENSWIGNCTKNNNTNDTITLPCRKQIINM
WQGVGQAMYAPPIEGKITCKSNITGILLTRDGGNNNTNETIFRPGGDMRDNRSELYKYKVVKIEPLG
VAPTKAKRRVVEREKRAVIGCAVFLGLGAGSTNGAASITLTVQARQLLSGIVQQQSNLLRAIEAQOHL
LQLTWGIKQLQARVLAVERYLKDQQLLGIWGSCKLICITTVPNSSWSNKSQDEIWDNMTWMEWEREI
NNYTDIIYSLIEESQOQEKNEQELLALDKWASLNNWFDITNWLWYIKFIMIVGGLIGLRIVFAVLSIV
NRVRQGSPLSFQTLIPNRPGRPRPEGIEEGGEQDRDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDFI
LIAARTVELLGRKGLRRGWEALKYLNLLQYWQQLKNSAISLLDTTAIAVAEGTDRVIEVVQORACRAIL
NIPRRIRQGLERALL

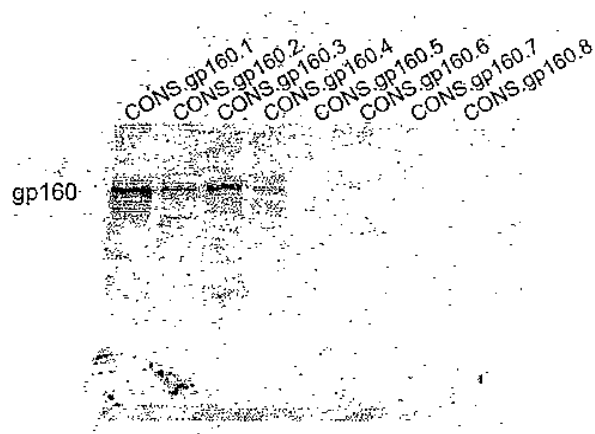
도면14b

CONS.env (그룹 M 컨센서스 env 유전자. 이는 env 유전자에 있는 변이 부위에 대한 컨센서스 서열을 포함.

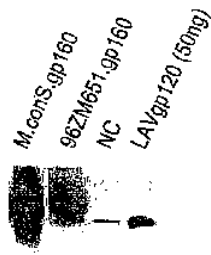
공유 도메인에 있는 것과 같은 동일 아미노산 서열)

GCCGCCGCCATGCGCGTGGCGCGGCATCCAGCGCAACTGCCAGCACCTGTG
GCGCTGGGGCACCCCTGATCCTGGGCATGCTGATGATCTGCTCCGCCGCCG
AGAACCCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC
AACACCACCCCTGTCTGCGCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCC
AGGAGATCGTGTCTGGAGAACGTGACCGAGAATTCAACATGTGGAAGAAC
AACATGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCACTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCCTGAATGCA
CCAACGTGAACGTGACCAACACCAACCAACACCGAGGAGAAGGGCGAG
ATCAAGAACTGCTCCTTCAACATCACCAACCGAGATCCGCGACAAGTAGCA
GAAGGTGTACGCCCTGTCTTACCGCCTGGACGTGGTGCCCATCGACGACA
ACAACAACAACCTCTCAACTACCGCCTGATCAACTGCAACACCTCCGCC
ATCACCCAGGCCCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCACTA
CTGCGCCCCCGCGGCTTCGCCATCTGAAGTGCAACGACAAGAAGTTCA
ACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCACTGCAACCCACGGC
ATCAAGCCCGTGGTGTCCACCCAGCTGTCTGTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCAACAACGCCCAAGACCA
TCATCGTGCGCTGAACGAGTCCGTGGAGATCAACTGCAACCCGCCCAAC
AACAACACCCGCAAGTCCATCCGCATCGGCCCGCGCCAGGCCCTTCTACGC
CACCGCGACATCATCGCGACATCCGCCAGGCCCACTGCAACATCTCCG
GCACCAAGTGGAAACAAGACCTGCAAGCAGGTGGCCAAAGAAGCTCGCGAG
CACTTCAACAACAAGACCATCATCTCAAGCCCTCTCCGCCGCGACCT
GGAGATCACCAACCACTCTTCAACTGCCCGCGGAGTTCTTCTACTGCA
ACACTTCCGCCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAAC
AACAACAACCAACGACACCATCACCCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGCGTGGGCCAGGCCATGTACGCCCGCCCATCG
AGGGCAAGATCACTGCAAGTCCAAACATCACCGGCCCTGTGTGACCCGC
GACGGCGGCAACAACAACCAACGAGACCGAGATCTTCCGCCCGCGCGG
CGCCGACATGCGCGACCAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGG
TGAAGATCGAGCCCTGGGCGTGGCCCCCACAAGGCCAAGCGCGCGTGT
GTGGAGCGCGAGAAGCGCGCGTGGGCATCGGCGCCGTGTCTTGGGCTT
CTGGGCGCGCGCGCTCCACCATGGGCGCGCCTCCATCACCTGACCG
TGCAGGCCCGCCAGCTGTGTCCGGCATCGTGAGCAGCAGTCCAACTTG
CTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG
CATCAAGCAGCTGCAGGCCCGCGTGTGGCGTGGAGCGCTACCTGAAGG
ACCAGCAGCTGCTGGGCATCTGGGCTGTCTCCGCCAAGCTGATCTGCACC
ACCACCGTGCCCTGGAACTCCTCCTGGTCCAAACAAGTCCAGGACGAGAT
CTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACATACA
CCGACATCATCTACTCCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAG
AACGAGCAGGAGCTGTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACCTG
GTTCCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCG
TGGGCGGCTGATCGGCTGCGCATCGTGTTCGCCGTGCTGTCCATCGTG
AACCCTGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACCCCTGATCCC
CAACCCCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGCGG
AGCAGGACCGCGACCGCTCCATCCGCTGTTGAACGGCTTCTGGCCCTG
GCCTGGGACGACCTGCGCTCCCTGTGCTGTCTCTACCACCGCCTGCG
CGACTTCATCTGATCGCGCCCGACCGTGGAGCTGCTGGGCGCAAGG
GCCTGCGCGCGCTGGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAG
TACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAC
CGCCATCGCGTGGCGGAGGGCACCGCGCTGATCGAGGTGGTGCAGC
GCGCTGCGCGCCATCTTGAACATCCCCCGCGCATCCGCCAGGGCCTG
GAGCGCGCTGCTGTTT

도면14c

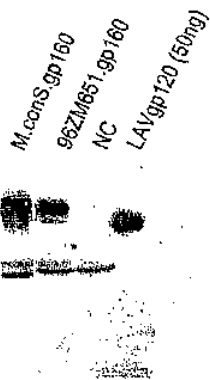


도면15a



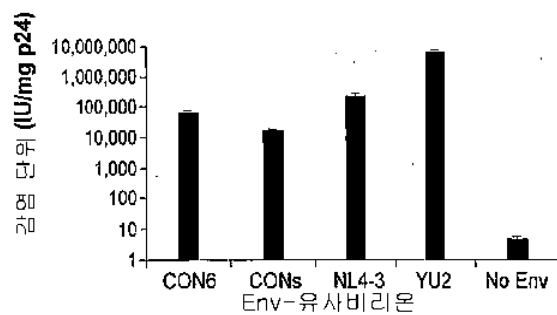
세포 분해물
포유동물에서의 A.con env 유전자의 발현

도면15b



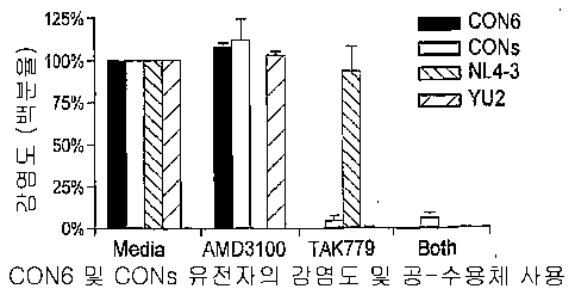
상등액
포유동물에서의 A.con env 유전자의 발현

도면16a

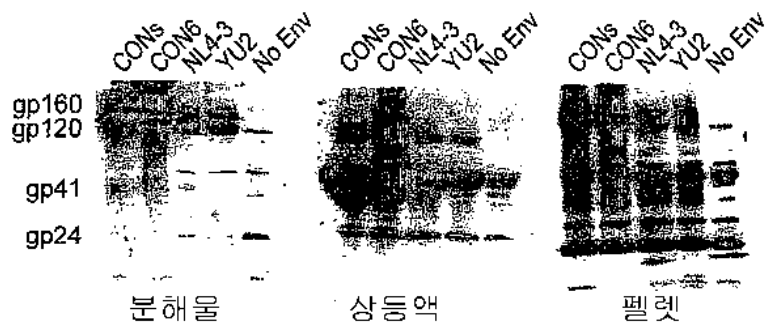


CON6 및 CONs 유전자의 감염도 및 공-수용체 사용

도면16b



도면17



CON6 및 CONs Env-유사비리온에서의 Env 단백질 결함

도 17A

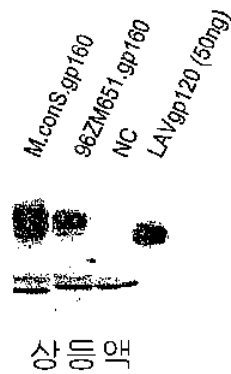
도 17B

도 17C

도면18a

A.con.env (subtype A consensus env)
 MRVNGIQRNQCQHLWRWGTMLGMILICSAAEENLWTVVYGVFVWKAETTLFCASDAKAYDTEVHNV
 WATHACVPTDNPQEIINLENVTEEPNWKNNMVEQMHTDIIISLWDQSLKPCVKLTPLCVTLNCSNVNVTI
 NITNITDNMKEIKNCSFNMTTELKDKKQKVSYLFYKLDVVQINKSNSSQYRLINCSAITQACPQVS
 FEPPIHYCAPAGFAILLCKDKKEFGTGPCKNVSTVQCTHGKIPVSTQLLLNGSIAEEEEVMIRSENITN
 NAKNLIIVQLTKPVKINCTEPNNTKRSIRIGPQAIFYATGDIIGDIRQAHCNVSRTEWNETTLOKVAKQLR
 KYFNKTIIFTNSSGGDEITTHSFNCGGEFFYCNSTGLFNSITWNGTKKKNSTESNDTITLPCRIKQI
 INMWQRVGQAMYAPPIQGVIRCESNITGLLLTRDGGDNNKNETFRPGGDMRDNRSELYKYKVVKIEP
 LGVAPTKARVWEREKRAVGIGAVFLGFLGAGSTMGASITLTIVQARQLLSGIVQQQSNLIRAI EAQQ
 HLLKLTWGIKQIQARVLAVERYLKDQQLLGIVGCSGKLICTTNVPNSSWSNKSQSEIWDNMTWLQWCK
 EISNYTDIIYNLIEESQOQERNEQDLIALDKWANLW NWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLS
 VLNVRQGYSPLSFQTHTPNPGGLDRPGRIEEEGGEQGRDRSIRLVSGFLALAWDDIRSLCLFSYHRLRD
 FILTAARTVELLGHSSSLKGLRLGWREGLKYLWNLLLYWGRELKSAINLLDTIAIAVAGWTDTRVIEIGQRI
 CRAILNIPRIRQGLERALL

도면18d



포유동물 세포에서의
A.con env 유전자의 발현

도면19a

M.con.gag (그룹 M 컨센서스 gag.

공유 도메인에 대한 동일 아미노산 서열)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGA
CGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCC
TGAAGCACCTGGTGTGGGCTCCCGCGAGCTGGAGCGCTTCGCCCTGAAC
CCCGGCTGTGGAGACCTCCGAGGGCTGCAAGCAGATCATCGGCCAGCT
GCAGCCCCCTGTCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
CCGTGGCCACCCTGFACTGCGTGCAACAGCGCATCGAGGTGAAGGACACC
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAA
GACCCAGCAGGCCCGCCGACAAAGGGCAACTCCTCCAAGGTGTCCCAGA
ACTACCCCATCGTGCAAGACCTGCAAGGGCCAGATGGTGCAACAGGCCATC
TCCCCCGCACCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTT
CTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC
ATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCG
CCTGCACCCCGTGCAACGCCCGGCCCATCCCCCGGCCAGATGCGCGAGC
CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC
GCCTGGATGACCTCCAACCCCCCATCCCCGTGGGCGAGATCTACAAGCG
CTGGATCATCCTGGGCTGAACAAGATCGTGCGCATGTACTCCCCGTGT
CCATCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG
GACCGCTTCTCAAGACCTGCGCGCCGAGCAGGCCACCAGGACGTGAA
GAACTGGATGACCGACACCTGCTGGTGCAAGCCCAACCCGACTGCA
AGACCATCCTGAAGGCCCTGGGCCCGCGGCCACCCTGGAGGAGATGATG
ACCGCTGCCAGGGCGTGGGCGGCCCGGCCCAAGGCCCGCGTGTGGC
CGAGGCCATGTCCCAGGTGACCAACGCCGCCATCATGATGCAGCGCGCA
ACTTCAAGGGCCAGCGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAG
GGCCACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAA
GTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCGGCCCGGCAACTTC
CTGCAGTCCCCCGGAGCCACCGCCCCCGCGGAGTCTTCGGCTT
CGGCGAGGAGATCACCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGC
CCCCCTGACCTCCTGAAGTCCCTGTTCCGCAACGACCCCTGTCCAG
TGA

도면19b

M.con.pol.nuc

GCCGCCGCCATGCCCCAGATCAACCTGTGGCAGCGCCCCCTGGTGACCAT
CAAGATCGGGCGGCCAGCTGAGGAGGCCCTGCTGGCCACCGCGCCGACG
ACACCGTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAGGCCCAAGATG
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
GATCGAGATCTGCGGCAAGAGGCCATCGGCAACCGTGCTGCTGGGCCCA
CCCCCGTGAACATCATCGGCGCAACATGCTGACCCAGATCGGCTGCACC
CTGAACCTCCCATCTCCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCC
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGA
TCAAGGCCCTGACCCAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAA
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAGCTGA
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCC
GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC
CTACTTCTCCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCA
CCATCCCTCCATCAACAAGAGAACCCCGGCATCGCTACAGTACAAC
GTGCTGCCCGCAGGGCTGGAAGGGCTCCCCCGCATCTTCCAGTCTCCAT
GACCAAGATCCTGGAGCCCTTCCGCACCCAGAACCCCGAGATCGTGATCT
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG
CACCGGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGCGCTGGGCTT
CACCACCCCGACAAGAAGCACCAAGGAGCCCCCTTCTGTGGATGG
GCTACGAGCTGCACCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCC
GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT
GAATGGGCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCA
AGCTGCTGCGCGCGCCAAGGCCCTGACCGACATCGTGCCCTGACCGAG
GAGGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGAAGGAGCCCGT
GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCGAGATCCAGA
AGCAGGGCCAGGACCAAGTGGACCTACAGATCTACCAGGAGCCCTTCAAG
AACCTCAAGACCGGCAAGTACGCCAGATGCGCTCCGCCACACCAACGA
CGTGAAGCAGCTGACCGAGGCCGCTGCAGAAGATCGCCACCGAGTCCATCG
TGATCTGGGGCAAGACCCCAAGTTCGGCTGCCCATCCAGAAGGAGACC

TGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATTCCCGAGTG
GGAGTTGCTGAACACCCCTCCCTGGTGAAGCTGTGGTACAGCTGGAGA
AGGAGCCCATCGCGCGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAAC
CGCGAGACCAAGCTGGGCAAGGCCGGCTACGTGACCGACCGCGCCGCCA
GAAGGTGGTGTCCCTGACCGAGACCAACCAAGAAAACCGAGTGCAGG
CCATCCACCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC
GACTCCCCAGTACGCCCTGGGCATCATCCAGGCCAGCCGACAAGTCCGA
GTCCGAGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG
TGTACCTGTCTGGGTGCCCGCCCAAGGGCATCGGCGGCAACGAGCAG
GTGGACAAGCTGGTGTCCACCGGCATCCGCAAGGTGCTGTTCTTGGACGG
CATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCACTCCAAGTGGCGCG
CCATGGCTCCGACTTCAACCTGCCCGCCATCGTGGCCAAAGGAGATCGTG
GCCTCCTGCGCAAGTGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGT
GGACTGCTCCCGCGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA
AGATCATCCTGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAG
GTGATCCCCCGGAGACCGGCCAGGAGACCGCTACTTCATCTGAAGCT
GGCCGGCGCTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAAC
TCACCTCCGCGCGCTGAAGGCCGCTGCTGGTGGCCGGCATCCAGCAG
GAGTTGGGCATCCCTACCAACCCAGTCCAGGGCGTGGTGGAGTCCAT
GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGCGCGACCAAGGCCGAGC
ACCTCAAGACCGCGTGCAGATGGCCGTGTTATCCACAACCTCAAGCGC
AAGGGCGGCATCGGCGGCTACTCCCGCGGCGAGCGCATCATCGACATCAT
CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCC
AGAAGTCCGCGTGTACTACCGGACTCCCGCGACCCCATCTGAAGGGC
CCCGCAAGCTGCTGTGAAGGGCGAGGGCGCGGTGGTGGATCCAGGACAA
CTCCGACATCAAGGTGGTGCCCGCGCAAGGCCAAGATCATCCGCGACT
ACGGCAAGCAGATGGCCCGGACGACTGCGTGGCCCGCGCCAGGACGAG
GACCTA

도면19c

M.con.nef (그룹 M 컨센서스 nef.

공유 도메인에 대한 동일 아미노산 서열)

GCCGCCGCCATGGGCGGCAAGTGGTCCAAGTCTTCATCGTGGGC TGCC
CGCCGTGCGCGAGCGCATCCGCGGCACCCCGCGCGCGAGGGCGTGG
GCGCCGTGTCCAGGACCTGGACAAAGCAAGGCGCCATCACCTCCTCCAAC
ACCGCGGCCAACACCCGAC TGCGCTGGCTGGAGGCCAGGAGGAGGA
GGAGGAGGTGGGCTTCCCGTGCGCCCGAGGTGCCCTTGCGCCCA TGA
CCTACAAGGCCCTGGACC TGTC CACTCTGGAAGGAAGAAGGCCCG C
CTGGAGGGCTGATCTACTCCAAGAGCGCAGGAGATCTGGACCTGTG
GGTGTAACCA CACCCAGGCTACTTCCCGACTGGCAGAACTACACCC CG
GCCCGGCACTCGCTA CCCCCGACCTTGGCTGGTGCTTCAAGCTGTG
CCCGTGGACCCGAGGAGGTGAGGAGGCCAACGAGGGCGAGAACAACTC
CCTGCTGCA CCGCATGTGCCAGCACGGCATGGAGGACGAGGAGCGCGAGG
TGCTGATGTGGAAGTTGACT CCGCTGGCCCTGCGCCACATCGCCCGC
GAGCTGCACCCGAGTACTACAAGGACTGCTAA

도면19d

C.con.pol.nuc
 GCCGCCGCCATGCCCCAGATCACCCCTGTGGCAGCGCCCCCTGGTGTCCAT
 CAAGGTGGGCGCGCCAGATCAAGGAGGCCCCCTGCTGGCCACCGCGCCGACG
 ACACCGTGTGGAGGAGATCAACCTGCCCGGCAAGTGAAGCCCAAGATG
 ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
 GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGTGGTGGGCCCCA
 CCCCCGTGAACATCATCGGCCCAACATGCTGACCCAGCTGGGCTGCACC
 CTGAACTTCCCCATCTCCCCATCGAGACCGTGCCTGGAAGCTGAAGCC
 CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGA
 TCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATC
 ACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAA
 GAAGAAGGACTCCACCAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA
 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCC
 GCCGCGCTGAAGAAGAAGAGTCCGTGACCGTGTGGACGTGGGCGACGC
 CTACTTCTCCGTGCCCTTGGACGAGGGCTTCCGCAAGTACACCGCCTTCA
 CCATCCCCCTCCATCAACAACGAGACCCCCCGCATCCGCTACCACTACAAC
 GTGCTGCCCGAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCCTCCAT
 GACCAAGATCCTGGAGCCCTTCCGCGCCCCAGAACCCCGAGATCGTGATCT
 ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG
 CACCGCGCCAGATCGAGGAGCTGCGCGAGCACCTGCTGAAGTGGGCGTT
 CACCACCCCCGACAAGAAGCACCAAGAAGGAGCCCCCTTCTGTGGATGG
 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCATCCAGCTGCC
 GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT
 GAAGTGGGCTCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA
 AGCTGCTGCGCGGCGCCAAAGGCCCTGACCGACATCGTGCCTGACCGAG
 GAGGCCGAGCTGGAGCTGGCCGAGAACCAGGAGATCCTGAAGGAGCCCT
 GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA
 AGCAGGGCCAGACCAAGTGGACCTACAGATCTACAGGAGCCCTTCAAG
 AACCTCAAGACCGGCAAGTACGCAAGATGCGCACCGCCCAACCAACGA
 CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGTCCATCG
 TGATCTGGGGCAAGACCCCCAAGTTCCGCTGCCCATCCAGAAGGAGACC
 TGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATTCCCGAGTG
 GGAGTTCTGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA
 AGGAGCCCTGACCTGTTGAGAGCTTCTGAGGAGTGGGCGCGCCCAAC
 CGCGAGACAAGATCGGCAAGGCCGCTACGTGACCGACCGCGCGGCCA
 GAAGATCGTGTCCCTGACCGAGACCAACCAAGCAAAACCGAGCTGCAGG
 CCATCCAGCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC
 GACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCGACAAGTCCGA
 GTCCGAGCTGGTGAACAGATCATCCAGCAGCTGATCAAGAAGGAGCGCG
 TGTACCTGTCTGGGTGCCCGCCACAAGGGCATCGGCGGCAACGAGCAG
 GTGGACAAGCTGGTGTCTCCGGCATCCGCAAGGTGCTGTTCTTGGACGG
 CATCGACAAGGCCAGGAGGAGCACGAGAAGTACCACTCCAAGTGGCGCG
 CCAATGGCCTCCGAGTTCAACCTGCCCGCATCGTGGCCAGGAGATCGTG
 GCCTCCTGCGACAAGTGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGT
 GGACTGCTCCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA
 AGATCATCTGGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAG
 GTGATCCCCCGCGAGACCGGCCAGGAGACCGCTACTTCATCTTGAAGCT
 GGCGGCGCGTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAACT
 TCACCTCCGCGCGCGTGAAGGCCGCTGCTGGTGGGCGCGCATCCAGCAG
 GAGTTGGGCATCCCCCTACAACCCCAAGTCCAGGGCGTGGTGGAGTCCAT
 GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGCGCGACCGAGGCCAGC
 ACCTCAAGACCGCCGTGCAGATGCCGTGTTTCATCCACAACCTTCAAGCGC
 AAGGGCGGCATCGGCGGCTACTCCGCGCGGAGCGCATCATCGACATCAT
 CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCC
 AGAAGTCCGCGTGTACTACCGGACTCCCGCGACCCCATCTGGAAGGGC
 CCGGCCAAGCTGCTGTGAAGGGCGAGGGCGCGTGGTGGTATCCAGGACAA
 CTCCGACATCAAGGTGGTGGCCCGCGCAAGGCCAAGATCATCAAGGACT
 ACGGCAAGCAGATGGCCGGGCGGACTGCGTGGCCGGCGCCAGGACGAG
 GACTAA

도면19e

M.con.gag (group M consensus gag)
 MGARASVLGGKLDANEKIRLRPGKKYRIKHLVWASRELERFALNPGLLETSEG CKQLIGQLQPA
 LQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEQNKSOQKTQQAADKGNSSKVSQNYPIVQN
 LQGQWVHQAI SPRTLNAWVKVIEEKAFSPFVIMFSALEGATPQDLNTMLNTVGGHQAAQMQLKDTINE
 EAAEWDRLHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIAWMTSNPPIPVGEIYKRWIIILGLNKKIVRM
 SPVSIIDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANPDCCKTILKALPGGATLEEM
 TACQGVGGPGHKARVLAEMSQVINAIMQRGNFQGRRIICKFNGCKEGHIANCRAPRKKGCWKCCK
 EGHQMKDCTERQANFLGKIWPSENKGRPGNELQSRPEPTAPPAESFGFGEIITPSKQEPKDKPEPPLTSLK
 SLFGNDPLSQ

도면19f

M.con.poi (group M consensus poi)
 MPQITLWQRPIVTKIGGQLKEALLaTGADDTVLEEINLPGRWKPKMIGGIGGFIKVRQYDQILIEIGK
 KAIGTVLVGTPVNIIGRNMLTGIGCTLNFFISPIETVPVKLPKGMDGPKVKQWPLTEKKKALTEICTE
 MEKESKISGIPENPYNTPIFAIKKXDSWKRLVDRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLD
 VGDYFVSPLDEDFRKYTAFTPSINNETHPGIRYQYNVLPQGWKSPALFQSSMTKILEPFTQNPPI
 YOYMDL YVGSDELIGQHRAKIEELREHLRWGFTTPDKKHQKEPFLWNGYELHPDKWTVQPIQLPEKD
 SWTVNDIQKLVGKLHWASQIYPGKVKQLCKLRGAKALTDIVPLTEBAELEAENREILKEPVHGVYD
 PSKQJAEIQKGGQDQWYTYQYQEPFKNLKTKYAKMRSATHTNDVKQLTEAVQKIATESVWNGKTPKFR
 LPIQKETWETWTEYMQATWPEWFEVNTPLVLWYQLEKEPIAGAEFTYVDGAANRETNLGAGVYTD
 RGRKWSLTETTNQTELOAHIALODSGSEVNIIVTDSQYALGIQADQPKSESELVNOIEQJIKKEK
 VYLSWVPAHKGGIGNEQYDKLVSTGIRKVPFDGIDKQEEHEKYHNNWRAMASDNLPPVAKIVASC
 DKCQLGEAMHGQVDCSPGIWQLDCTHLEGGKILVAVHVASGYEAPETGQETAYFLKAGRWPV
 KVIHTDNGSNFTSAVKAACWAGIQDEFIPYNPQSGVAVESMNKELKIGQVRDQAEHLKTAVQMAV
 FHNFRKKGIGGYSAGERIDIAIDITQTKELQKITKQINFRVYRDSRDPWKGFAKLLWKGEAVV
 IQDNSDIKVVPRIKAKIIRDYQKQMGDDICVAGRQDED

도면19g

M.con.nef (group M consensus nef)
MGGKWSKSSIVGWPAVRRERLRTRTHPAARGVAVSQDLDDKKHGAITSSNTAANNPDCAWLEAQEEEEVGFP
VRPQVPLRPMTYKAAIDLSHFLKEKGLEGLIYSKKRQEIIDLWVHTQGYFFPDWQNYTFPGGIRYPLTF
GWCEKLVPTDPEEVEEANEGENNSLIHFMCGHGMEDEREVILMNKFDSEIALRHIAHELHPEYKDC

도면19h

C.con.pd (subtype C consensus pd)
 MPQITLWDRPLVSKVGGQIKKALLaTGADDTVEEINLPKWKPKMIGGIGGFYKVRQYDQILIEICGK
 KAIGTVLGGTPVNIIGRNMLTQLCTLNFPSPIETVPVKLPGMDGPKVKQWPLTEEKIKALTAICEE
 MEKGGITKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTOODFWYQVGLGIPHPAGLKKKSVTVLD
 VGDAYSFPLDEGFRKYTAFTIP SINNETPGIRYQYNVLPQGWKSPAFQSSMTKILEFFRAQNPETM
 YQYMDLLVYGSDLGSGHRAKIEELREHLLKWGFTTPDKKHQKPEPFLWMGYELHPDKWTVQPIQLPEKD
 SWTVNDIQKLVGKLNWASQIYPGKVRQLCKLRGAKALTDIVPLTEAELELAENREILKEPVHGVYTD
 PSKDLIAEIQKQGHQWVYTIYQDEFFKRLNKTGKYAKMRTAHTNDVQQLTEAVQKIAMESIVWGTPTKFR
 LPIOKETWETWTDYNDQATWPEWEPATPPLVKLWYOLEKEPIAGAEYFDGAANRETKIGKAGYVTD
 RGRQKIVSLTETTNOKTELOALQDSSSEVNVTDSTYALGIQAQDPKSESELNQIEQLIKKER
 WYLSWPAHKHGGIGNEQYDKLVSSGIRKVLFLDGIDKAEHEHEKYHNSNRWMASEFNLPPIVAKEVASC
 DKCOLKGEAMHGQVDCSPGIWQLDCTHLEGKILVAVHVASGYEAEVPAETGQETAYFILKAGRWPV
 KVIHTDNGSNFTSAAVKAAACWMAQIDQEFQIPYNPQSGVYVESMKNELKKIIGQVRDQAEHLKTAVMAY
 FIHFKRKGIGGYSAGERIDIATDIQTKEQLKQIKIQRFRVYRDSDPIWKGPAKLLWKGEQAVV
 IQNDSIDIKVPRRKAKIKIDYQKIMAGADCVAGRDDED

도면20a

B.con.gag (서브타입 B 컨센서스 gag. 이 아미노산은 Los Alamos Database August 2002와 다름)

GGCGGCGCCATGGGCGCCCGCGCTCCGTGCTGTCGGGCGGCGAGCTGGA
 CCGCTGGGAGAAAGATCCGCGCTCGCGCCCGCGGCAAGAGAAGTACAAGC
 TGAAGCACATCGTGTGGGCGCTCCGCGAGCTGGAGCGCTTCGCGGTGAAC
 CCGGCGCTGCTGGGAGACCTCCGAGGGCTGCCGCCAGATCCTGGGCGAGCT
 GCAGCGCTCCCTGCGACACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
 CCGTGGGCGACCTGTACTGCGTGACACAGCGCATCGAGGTGAAGGACACC
 AAGGAGGCGCTGGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAA
 GGCCCGAGCAGGCGCGCGCGACACCGGCAACTCCTCCAGGTGTCCAGAA
 ACTACCCCATCGTGACAGAACCTGCAGGGCCAGATGGTGACACAGGCCATC
 TCCCCCGGCACCTTGAACGCGCTGGGTGAAGGTGGTGGAGGAGAAGGCGCTT
 CTCCCCCGAGGTGATCCCCCATGTTCTCCGCGCTGTCCGAGGGCGCCACCC
 CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCGCACAGGCGCGCC
 ATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCGCGCGAGTGGGACCG
 CCTGCACCCCGTGACGCGCGCGCCCATCGCCCCCGGCCAGATGCGCGAGC
 CCGCGGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGCAGATC
 GGCTGGATGACCAACAACCCCCCATCCCGTGGGCGGAGATCTACAAGCG
 CTGGATCATCCTGGGCGCTGAACAAGATCGTGCGCATGTACTCCCCACCT
 CCATCCTGGACATCCGCGCAGGGGCCCAAGGAGCCCTTCGCGGACTACGTG
 GACCGCTTCTACAAGACCTGCGCGCGGAGCAGGCGCTCCAGGAGGTGAA
 GAACCTGGATGACCGAGACCTGTGCTGGTGCAAGCGCAACCCCGACTGCA
 AGACCATCCTGAAGGCGCTGGGCGCGCGCGCCACCTGGAGGAGATGATG
 ACCGCGCTGCCAGGCGCTGGGCGCGCGCGGCCACAAGGCGCGCGTGTGGC
 CGAGGCCATGTCCAGGTGACCAACTCGGCCACCATCATGATGCAGCGCG
 GCAACTTCCGCAACCAGCGCAAGACCGTGAAGTGCTTCAACTGCGGCAAG
 GAGGCGCACATCGCCAAGA ACTGCGCGCGCCCCCGCAAGAAGGGCTGCTG
 GAAGTGGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCGCAGG
 CCAACTTCCCTGGGCAAGATCTGGCCCTCCCAACAAGGCGCGCCCCGGCAAC
 TTCTGCGAGTCCCGCGCGGAGCCCAACGCGCCCCCGGAGGAGTCCCTCCG
 CTTCGGCGAGGAGACCAACCAACCCCTCCAGAGAAGCAGGAGCCCATCGACA
 AGGAGCTGTACCCCTTGGCCTCCCTGCGCTCCCTGTTCGGCAACGACCC
 TCCTCCAGTAA

도면20b

B.con.env (서브타입 B 컨센서스 env. 이 아미노산은 Los Alamos Database August 2002와, 다음)

GCGCGCGCCATGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTG
GCGCTGGGGCACCATGCTGCTGGGCATGCTGATGATCTGCTCCGCGCGCG
AGAAGCTGTGGGTGACCGGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC
ACCAACCACCTGTTCTGCGCCTCCGACGCCAAGGCCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCTGCGTGCCTCAACGACCCAAACCC
AGGAGGTGGTGTCTGGAGAACGTGACCGAGAAGCTTCAACATGTGGAAGAAC
AACATGTTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACAGTCT
CCTGAAGCCCTGCGTGAAGCTGAACCCCTGTGCGTGACCTGAACTGCA
CCGACCTGAAGAACAACCTGCTGAACACCACTCCTCCTCGGCGAGAAG
ATGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAAATCAACACCTCAT
CCGCGACAAGGTGCGAGAAGGATACGCCCTGTTCTACAAGCTGCACTGG
TGCCCATCGACAACAAACAACACCTCTACCGCGCTGATCTCCTGCAAC
ACCTCGTGATCACCCAGGCTGCCCCAAGGTGCTCTCGAGCCCATCCC
CATCCTACTGCGCCCGCGCGGCTTCGCATCTGGAAGTGCAACGACA
AGAAGTTCAACGGCACCGGCCCTGCAACACGTGTCCACCGTGCAGTGC
ACCCACGGCATCCGCCCCGTGGTGTCCAACAGCTGCTGCTGAACGGCTC
CCTGCGCGAGGAGGAGTGGTGTCTGCTCCGAGAAGTTCACGACAACG
CCAAGACCATCATCTGTCAGCTGAACGAGTCCGTGGAGATCACTGCAC
CGCCCCAACAAACAACCCGCAAGTCCAACACATCGGCCCGGCGCGC
CTTCTACACCAACCGCGGAGATCATCGCGGACATCCGCGAGGCCAATGCA
ACATCTCCCGCGCAAGTGAACAAACACCTGAAGCAGATCGTGAAGAAG
CTGCGCGAGCAGTTCTGCGCAACAAGACCACTCGTGTCAACCAAGTCTC
CGGCGACCCCGAGATCTGTGATGCACTCTTCAACTGCGGCGGCGAGTTCT
TCTACTGCAACACCAACCGAGCTGTTCAAATCCACTGGAACGACAACGGC
ACCTGGAACAACACCAAGGACAAGAACAACATCACCTGCGCTGCGCGCAT
CAAGCAGATCATCAACATGTGGCAGGAGTGGGCAAGGCCATGTAACGCC
CCCCCATCCGCGGCGCAGATCCGCTGCTCTCCAACATCACCGGCTGCTG
CTGACCCGCGACGGCGGCAACAACAACAACGACACCGAGATCTTCCGCC
CGGCGGCGCGCGACATGCGCAACTGGCGCTCCGAGCTGTACAAGTACA
AGGTGGTGAAGATCGAGCCCCGGGCGTGGCCCCACCAAGGCCAAGCGC
CGCGTGTGCGCGCGAGAAGCGCGCGTGGGCGATCGGCGCCATGTTCT
GGGCTTCTGGGCGCGCGCGCTCCACCATGGGCGCGGCTCATGAACC
TGACCGTGCGGCGCGCGAGCTGCTGTCGGCATCGTGCAGCAGCAGAAC
AACCTGCTGCGGCCATCGAGGCCAGCAGCACCTGCTGCGAGCTGACCGT
GTGGGCGATCAAGCAGCTGCAAGCCCGCGCTGCTGGCGGTGAGCGCTACC
TGAAGGACCAGCAGCTGCTGGGCGATCTGGGCTGCTCCGCAAGCTGATC
TGCAACACACCGTGCCTTGGAAACGCTCTGCTGTCGCAACCAAGTCCCTGGA
CGAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACA
ACTACACCTCCCTGATCTACAACCTGATCGAGGAGTCCCAAGAACGAG
GAGAAGAACGAGCAGGAGCTGCTGGAGCTGACAAAGTGGCTCCCTGTG
GAAGTGGTTGACATCAACAACTGGCTGTGGTACATCAAGATCTTCACTA
TGATCGTGGGCGGCTGATCGGCTGCGCATCGTGTTCGCGCTGCTGTCC
ATCGTGAACCGGCTGCGCCAGGCTACTCCCTGCTCCTTCAGACCGG
CCTGCCCGCCCCCGCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGG
GCGGCGAGCGCGACCGCGACCGCTCCGCGCGCTGGTGGACGGCTTCTG
GCCCTGATCTGGGAGCAGCTGCGCTCCCTGTGCTCTCTCTACACCG
CCTGCGCGACCTGCTGCTGATCGTGACCGCATCGTGGAGCTGCTGGGCG
GCCGCGCTGGAGGTGCTGAAGTACTGGTGAACCTGCTGCAGTACTGG
TCCAGGAGCTGAAGAAGTCCGCGCTGTCCCTGCTGAACCGCACCGCAT
CGCGGTGGCGAGGGCACCGACCGCTGATCGAGGTGGTGCAAGCGCGCT
GCCGCGCATCTGCAATCCCGCGCGCATCCGCGAGGGCTGGAGCGC
GCCCTGCTGTAA

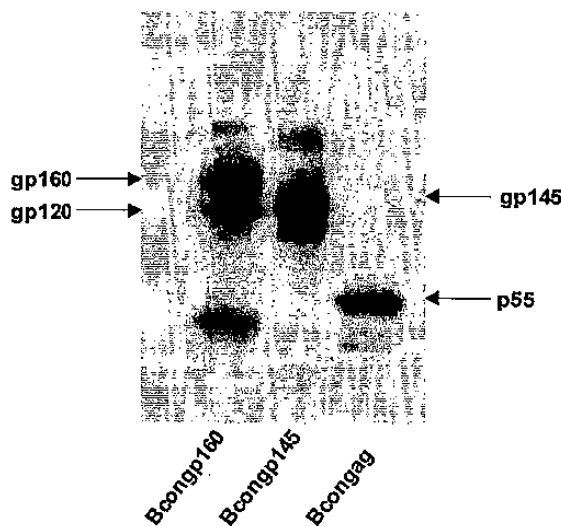
도면20c

B.con.gag (subtype B consensus gag)
 MGARASYLSGGELDRMEKIRLPFGKKYKILKHIVWASRELERFAVNPGLLETSEGCRLGLQLQPSLOT
 GSEELRSLYNTVATLYCVHORIEVKDTKEALEKIEEQNKSKKKAQAAADTGNSSQVSONYPIVQNLOG
 QNVHQALSPRTLNWVKVVEKAFSEVI PMESALSEGATPQDLNMTLVGGHQAAMQMLKETINEEAA
 EWDRLHPVHAGPIAPGQMRPEPGSD IAGTTSTLQEQIGWMTNNPPI PVGEIYKRWTILGLNKIV RMYSP
 STLDIROGPKPEPRDYVDRFYKTLRAFAQSQEVQNMWMTETILLVQNPDPCKTILKALGPAATLEEMMTAC
 QGVGGPGHKARVLAAMSQVINSATIMMORGNFRNORKTVKCNCKEGHIAKNCRAPRKKGCKGCKEG
 HQMKDCTERQANFLGKIWFESHKGRPGNFLQSRPEPTAPPESEFRGREETTPSQKQEPIDKELYPLASLR
 SLFGNDPSSQ

도면20d

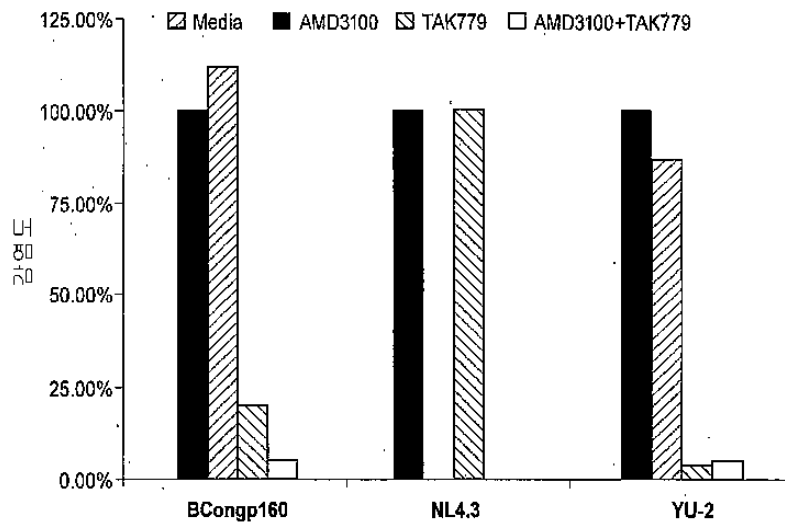
B.con.env (subtype B consensus env)
 MRVKGIRKHYQHLWRWCTMLLGLMLICSAAEKLVTV YYGVPYWKETITLFCASDAKAYDTEVHNWAT
 HACVPTDENPQEVVLENTENFMKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCLDLDKNNLLNT
 NSSGGEKMEKEIKNCSEFNITTSIRDKVQKEYALFYKLDVPEIDNNNTSYRLISONTSVITQACPVSF
 EPIPLHYCAPAGFAILKCNCKENGCTGVSTVQCTHGIPEVSTQLLNGSLAEFEVIRSEMETDN
 AKTIIVQLNESVEINCTRPNNNTKSIHIGPGRAFYTTGEIIGDIQAHCHNISRAKWNNTUKQIVKKLRE
 QFGNKTIVFNQSGGDPETVMHSFNCGEFFFYCNITQLFNSTWNDNGTWNNTKDKNTITLPCRITQIINN
 WQVCKAMYAPPPIRGQIRCSNITGLLLTRDGGNNNDTEIPRPGGDMRDNRSELYKVKVKTIEPLGV
 APTKAKRRVQREKRAVGIGAMFLGFLGAAGSTMGAASMTLTVPARQLLSGLVQQQNNLLRAIEAQOHL
 QLTVMGIGQLQARVLAVERYLKDQQLGIMGCSGLICTITVPWNASWSKSLDEIWDNNMTWMEWEREID
 NYTSLIYTLLEESQOQKNEQELLELDKXASLNNWFDITNWLWYIKIFIMIVGGILGLRVFAVLSIVN
 RVQGYSPLSFQTRLPAPRGDPDEGIEEGGERDRDRSGRLVDGFLALIWDRLSLCLFSYHRLDRLLL
 IVTRIVELLGRRGWEVLKYWNILLQYWSQELKNSAVSLINATAIAVARGTDRVIEVYQACRAILLIPRR
 IRQLERALL

도면21



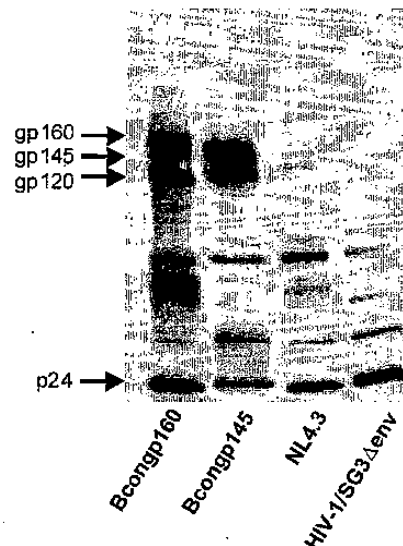
293T 세포에서 서브타입 B 컨센서스 env 및 gag 유전자의 발현, 코돈-최적화된 서브타입 B 컨센서스 gp160, gp140 및 gag 유전자를 293T 세포에 형질감염시키고, 웨스턴 블롯 분석으로 세포 분해물에 있는 단백질 발현을 확인하였다. 세포감염 48시간 후에, 세포분해물을 수득하고, BCA 단백질 분석법으로 단백질 함량을 확인하며, 2 μ g의 전체 단백질을 4-20% SDS-PAGE 겔의 각 레인에 로딩하였다. 단백질을 PVDF 막에 트랜스퍼하고 HIV-1 서브타입 B 감염된 환자의 혈청으로 탐침시켰다.

도면22



서브타입 B 컨센서스 외피의 공-수용체 사용
서브타입 B 컨센서스 gp160 Env를 포함하는 유사타입 임자를 AMD3100(CXCR4의 특이 저해제), TAK779(CCR5의 특이 저해제), AMD3000+TAK779의 존재하에서 DEAE-맥스트란 처리된 JC53-BL 세포에서 반응시켜 공-수용체 사용을 확인하였다. NL4.3, CXCR4 및 YU-2를 이용하는 것으로 알려진 분리물, CCR5-사용 분해물로 알려진 것을 대조군으로 포함시켰다.

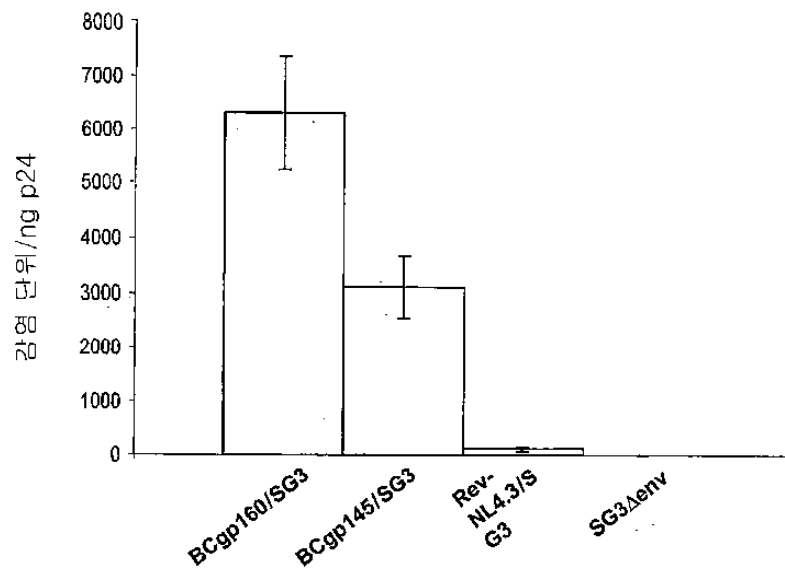
도면23a



코돈-최적화된 서브타입 B 컨센서스 gp160 및 gp140 유전자를 이용한 env-결핍 HIV-1의 트랜스 상보성

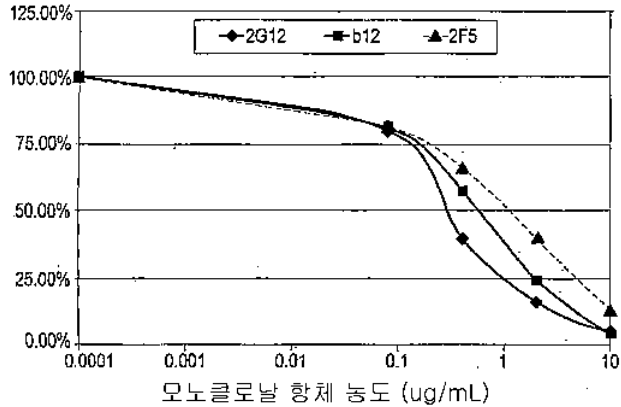
코돈-최적화된 서브타입 B 컨센서스 gp160 또는 gp140 유전자를 포함하는 플라스미드를 HIV-1/SG3Δenv 프로바이러스와 함께 293T 세포에 동시-형질감염시켰다. 형질감염 48시간 후에 유사타입 바이러스를 포함하는 세포 상등액을 수득하고, 테이בל탈 원심분리로 정화 시키며, 0.2uM 필터로 필터하고 20% 수크로즈 쿠션으로 펠렛화하였다. 각 바이러스 펠렛에 있는 p24의 정량은 쿨터 HIV-1 p24 항원 분석으로 확인하였으며; 25ng의 p24를 4-20% SDA-PAGE 젤의 각 레인에 로딩하였다. 단백질을 PVDF막으로 트랜스퍼하고 HIV-1 서브타입 B 감염된 환자 혈청의 항-HIV-1 항체로 탐침하였다. rev-의존 NL4.3env를 가진 트랜스 상보성을 대조군에 포함시켰다.

도면23b



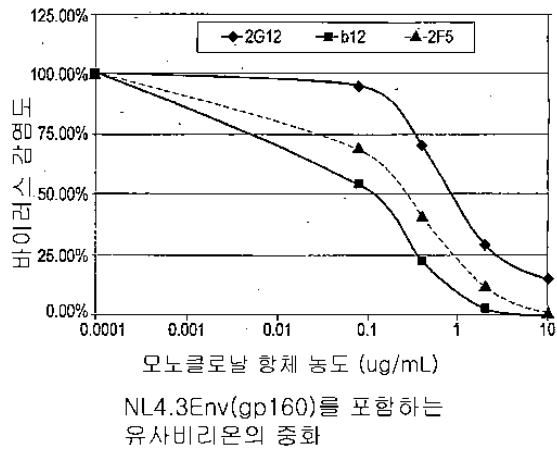
서브타입 B 컨센서스 외피를 포함하는 바이러스 입자의 감염도
컨센서스 B gp160 또는 gp140을 포함하는 유사타입 바이러스의 감염도는 JC53-BL 분석법으로 확인하였다. 수크로즈 쿠션 정제된 바이러스 입자를 쿨터 p24 항원 분석법으로 분석하고, 각 펠렛의 5배 연속 희석액을 DEAE-엑스트란 처리된 JC53-BL 세포와 반응시켰다. 48시간의 반응기간 후, 세포를 고정하고 염색하여 b-갈락토시다아제 발현 세포를 시각화하였다. 감염도는 p24 ng 당 감염 단위로 나타내었다.

도면24a

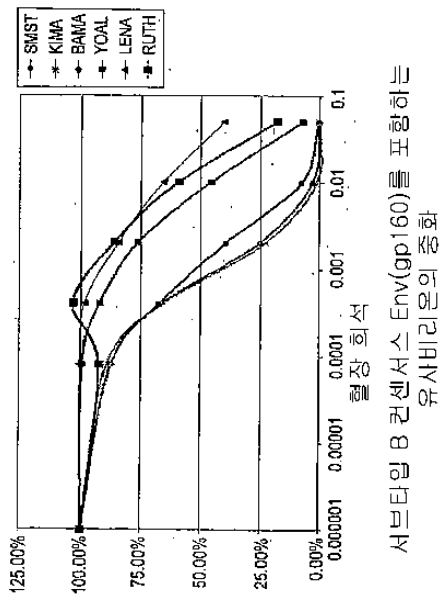


서브타입 B 컨센서스 Env(gp160)를 포함하는
유사비리온의 중화

도면24b



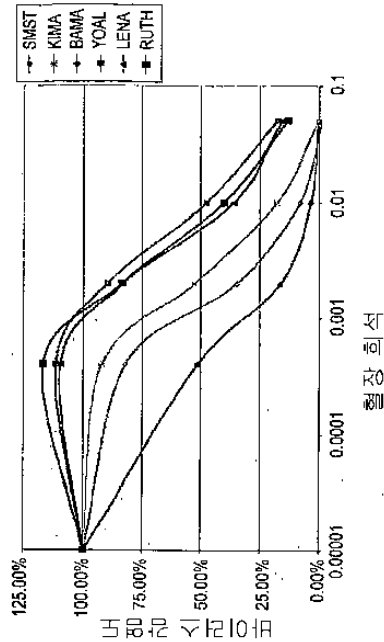
도면24c



서브타입 B 컨센서스 gp160 외피를 포함하는 비리온의 중화 민감도

서브타입 B 컨센서스 또는 NL4.3 Env(gp160) 을 포함하는 등가량의 유사비리온(1,500 감염단위)을 3개의 다른 모노클로날 중화 항체 및 HIV-1서브타입 B 감염된 환자의 혈장 샘플의 패널로 전반응시키고, 96웰 플레이트에 있는 JC53-BL 세포 단일층에 첨가하였다. 플레이트를 2일동안 반응시키고 바이러스 감염도의 표지자로 루시퍼라아제 활성을 측정하였다. 바이러스 감염도는 항체의 각 농도에서 나온 루시퍼라아제 단위(LU)를 대조군 감염에서 나온 LU로 나눔으로써 계산하였다. 평균 50% 지해 농도(IC50) 및 각 항체 희석액에서의 실제 % 중화를 각 바이러스에 대해 계산하였다. 루시퍼라아제 실험의 결과는 대응되는 감염에서의 푸른색 초점을 직접 계수하여 확인하였다.

도면24d

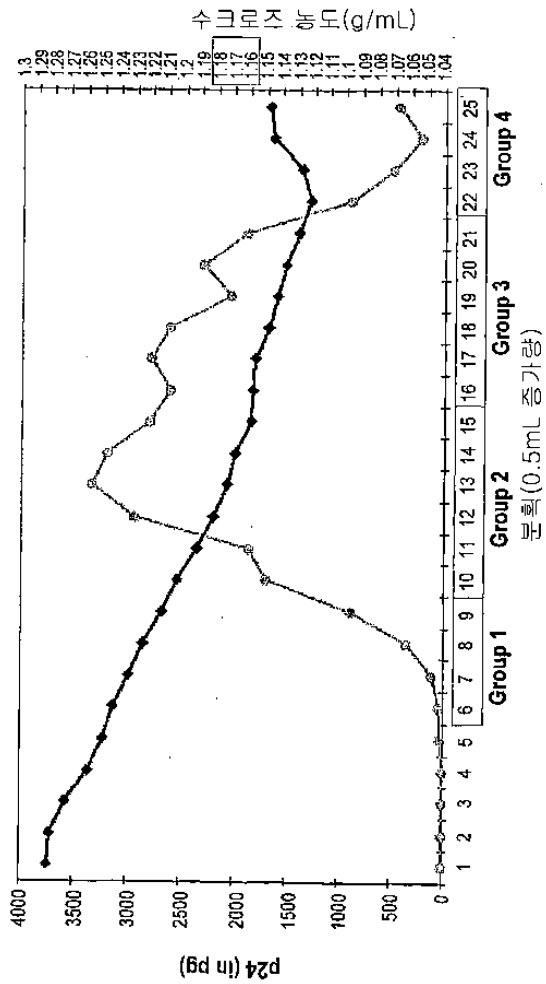


NL4.3Env(gp160)를 포함하는
유사비리온의 중화

서브타입 B 컨센서스 gp160 외피를 포함하는 비리온의 중화 민감도

서브타입 B 컨센서스 또는 NL4.3 Env(gp160)를 포함하는 등가량의 유사비리온(1,500 감염단위)을 3개의 다른 모노클로날 중화 항체 및 HIV-1서브타입 B 감염된 환자의 혈장 샘플의 패널로 전반응시키고, 96웰 플레이트에 있는 JC53-BL 세포 단일층에 첨가하였다. 플레이트를 2일동안 반응시키고 바이러스 감염도의 표지자로 루시퍼라아제 활성을 측정하였다. 바이러스 감염도는 항체의 각 농도에서 나온 루시퍼라아제 단위(LU)를 대조군 감염에서 나온 LU로 나눔으로써 계산하였다. 평균 50% 저해 농도(IC50) 및 각 항체 희석액에서의 실제 % 중화를 각 바이러스에 대해 계산하였다. 루시퍼라아제 실험의 결과는 대응되는 감염에서의 푸른색 초점을 직접 계수하여 확인하였다.

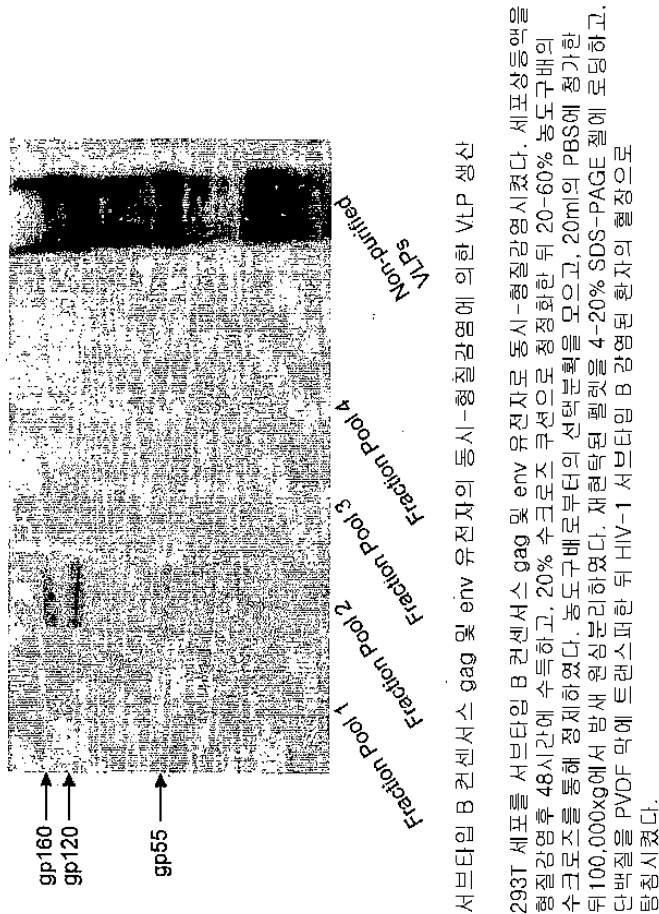
도면25a



수크로즈 농도구배 분획의 농도 및 p24 분석

0.5ml 분획을 20-60% 수크로즈 농도구배로부터 수득하였다. 분획번호 1은 농도구배 튜브의 바닥에서 얻은 가장 농도가 짙은 분획이다. 겔질계를 이용해 농도를 측정하고 각 분획에 있는 p24의 양을 롤더 p24 항원 분석법으로 확인하였다. 분획 6-9, 10-15, 16-21 및 22-25를 함께 모으고 웨스턴 블랏으로 분석하였다. 예측한 대로, 비리온은 1.16-1.18g/ml 농도에서 침전되었다.

도면25b



도면26a

Year 2000 Con-S 140CFI.Env

```

MRVRGIQRNCQHLWRWGTLLGLMLMICSAAENLWVTVYVGVVWKEANTLFCASDAKAYDTEVH
NVWATHACVPTDPNPQEIIVLENTNFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNC
TNVNVNTTNTNTEEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDNNNNSSNYRLINCNT
SAITQACPKVSFEPPIHYCAPAGFAILKNDKKFNGTGPKKNVSTVQCTHGIKPVVSTQLLNG
SLAEEELIIRSENIITNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQA
HCNISGTKWNKTLQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTW
IGNGTKNNNNNTNDTITLPCRILQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTIN
ETEIFRPGGGDMRDNRSELYKYVVKIEPLGVAPTAKLTVQARQLLSGIVQQQSNLLRAIEAQ
QHLLQLTVWGIKQLQARVLAVERYLKDQQLLEIWDNMWMEWEREINNYTDIISLIEESQNQQEK
NEQELLALDKWASLWNNWFDITNNLW
    
```

gp140CFI는 막형단 및 세포질 도메인의 결실에 추가로 절단-위치-결실(C), 융합-위치-결실(F) 및 gp41 면역우세 부위-결실(I)을 가진 HIV-1 외피 디자인을 지칭한다.

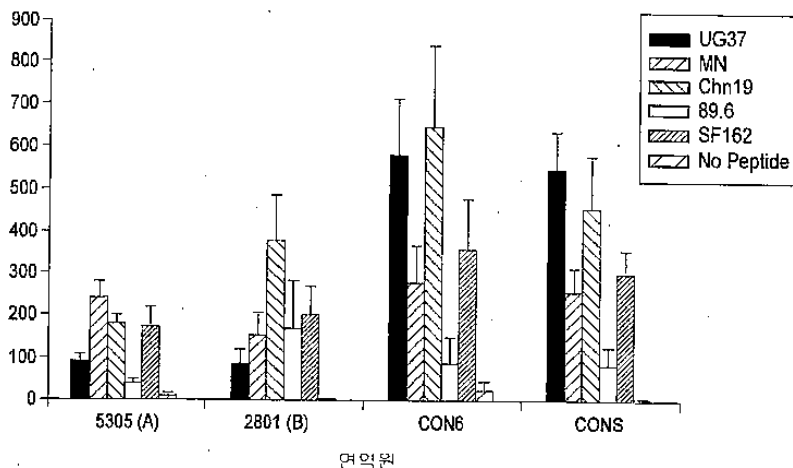
도면26b

Codon-optimized Year 2000 Con-S 140CFI. seq

ATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCCTGATCCTGGG
CATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGT
GGAAGGAGGCCAACACCACCTGTCTGCGCCTCCGACGCCAAGGCTTACGACACCGAGGTGCAC
AACGTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCGAGGAGATCGTGTGGAGAA
CGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCT
CCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCCCTGAAGTGC
ACCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTC
CTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTCTTACCCTGG
ACGTGGTGGCCATCGACGACAACAACAACAACTCCTCCAACCTACCGCTGATCAACTGCAACACC
TCCGCCATCACCAGGCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCACTACTGCGCCCC
CGCCGGCTTCGCCATCCTGAAGTGAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAGAAGC
TGTCCACCGTGCAAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGTCTGTAACGGC
TCCCTGGCCGAGGAGGATCATCTCCGTCGAGAACATCACCACAACGCAAGACCATCAT
CGTGCAGCTGAACGAGTCCGTTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCA
TCCGCATCGGCCCGGCGAGGCTTCTACGCCACCGCGACATCATCGCGGACATCCGCCAGGCG
CACTGCAACATCTCCGCCACCAAGTGGAACAAGACCTGCAGCAGGTGGCCAAGAAGCTGCGCGA
GCACCTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGCGACCTGGAGATCACCACCC
ACTCCTTCAACTGCCCGCGCGAGTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGG
ATCGGCAACGGCACCAAGAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAA
GCAGATCATCAACATGTGGCAGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGA
TCACCTGCAAGTCCAACATCACCAGGCTGTCTGCTGACCCGCGACGGCGGCAACAACAACCAAC
GAGACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAA
GTACAAGGTGGTGAAGATCGAGCCCTTGGCGGTGGCCCAAGGCCAAGCTTACCGTGCAGG
CCCCCAGCTGTCTGTCGGCATCGTGCAGCAGTCCAAACCTGCTGCGCGCCATCGAGGCCAG
CAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCTGCTGGCCGTGGA
GCGCTACCTGAAGGACGAGCTCGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCG
AGATCAACAACCTACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCGAGCAGGAGAAG
AACGAGCAGGAGCTGTGGCCCTGGACAAGTGGCCCTCCCTGTGGAAGTGGTTCGACATCACCAA
CTGGCTGTGGTGAGGATCC

도면27

HIV-1 외피 펩타이드에 대한 C56BL/6 마우스 T세포 반응



도면28a

Design of expression-optimized HIV-1 envelope gp140CF

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

MRVKGIRKNYQHILWRWGTMLLGLMLMICSAAEKILWVYVYGVVWKEATTTLCASDAKAYDTEVHNWTHACVPTDPNPQEVVL
 ENVTEFNMMKNNMVEQMHEDIISLWDQSLKFCVKLTPLCVTINCITDLMNATNTTTIYRWGEIKNCSENITTIRDKVQKEY
 ALFYKLDVPIDNNTSYRLISCSNTSVITQACPKVSEPIPIHYCAPAGFAILKCNDKFGTGPCTNVSTVQCTHGIRPVYSTQ
 LLNGSLAAEEVIVIRSENFIDNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGRAFYTTEIGDIRQAHCHISAKWNTLKQ
 IVKKLREQFGNKTIVFNQSSGGDPEIVMHSNCGGEFFYCNITQFNSTWNTTGNTEITPCRIKQIINWQEVGKAMYAPP
 IRGQIRCSNITGLILTRDGGNETEIFRPGGDMRDNRSELYKYVWKIEPLGVAPTAKRRVVOREKREHVGIGAMFLGFLGA
 AGSTMGAASMTLVQARQLLSGIVOOQNLLRAIEAQHLLQLTVNGIKQLQARVLAVEVYIKDQQLLGTWCSGKLCITFAPPW
 NASWSKSLDELWDNMTWMEWEREIDNYTSLYTLIESQCKNEQELLEDKWLWNNWFDITNWLVIKIFIMIVGGIVGL
 RIVFAVLSIVNRVRQGSPLSFQTRLPAPRGPDRPEIEEGGERDRSGRLVDGFLALIMDDLRSLCLFSYHRLDLLLIVTR
 IVELLGRRGWEVLKYWNWLLQYWSQELKNSAVSLLNATAIAVAEGTDRTVEVVRACRAILHIIPRRIRQGLERALL

*일출친 아미노산은 140CF 디자인에서 결실되어야 하는 용한 도메인을 지칭하고, 빨간색으로 밑줄친 "W"는 C 말단에 있는 마지막 아미노산이고, "W" 뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면28b

Con-B-140CF.pap (632 a.a.)
 Nick name: 002
 MRVKGIRKNIYQHILWRWGTMLLGMNICSAAEKLWVIVYGVFWKEATTLFCASDAKAYDTEVHNVWATHACVPTDPNPOEVVL
 ENVTEFNWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDIMNATNTTIIYRWGEIKNCSEFNITTSIRDKVOKEY
 ALFYKLDVVPIDNDNTSYRLISONTSVITQACPKVSEPIPIHYCAPAGFALKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQ
 ZLLNGSLAEVEVIRSENFTDNAKIIIVQLNESVEINCTRPNNTRKSIHIGPGRAPYTTGELIGDIROAHCHNISRAKWNNTLKQ
 IVKKLRQFGNKTIVFNQSSGGDPEIVMHSFNCGGGEFFYCNNTQLFNSTWNGTWNTEGNTLPCRIKQIINMWQEVGKAMYAPP
 IRGQIRCSNITGLLLTRDGGNNETEIPRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKATLTQVQARQLLSGIVQOQNNLLRA
 IEAQOHLQLFVWGIKQLQARVLAVERYLKQOQLLGIVGCSCKLICTTAVFWASWSNKSILDEIWNNTWNEWEREIDNTSLIY
 TLEESQOQEKNEQELLELDKWSLWFWFDITNLW*

* 푸른색으로 표시된 아미노산은 결실된 음화 절단 위치의 경계를 간단히 표시한 것이다.

도면28c

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

Nick name: 002

TTCCAGTCGACGGCCACCATCAGGGTGAGGGTATTCCGAAATAATCCAAACCTGTGGCGGCTGGGAACAACATGCTCTTGGTAT
CTTGTAATCTGGAGTGGCCCGGAGAAACTTGGTAACCTGTGTACTAGCGGCTCTGTCTGGAGGAAGCTACACCACTCTT
TTTGTGCACTCCGAGCTAAAGCTTACGACACAGAAGTCATATATGTTGGCCACCGATGTGGTCCCTACAGATCCCAAC
CCCAGGAAGTCGTCTTGAGATGTCACAGAGAAATTTAACATGTGGAAGATAATGGTAGAACAAATGACGAAGACATAT
TAGCTGTGGGACAGCTGTGAAGCCGTGGTGAACCTCACTCCATTTGGTCCACCTTACTGTACTGTATTCATGAAGCGCA
ACCAACAGGAATACCTATTATATATCCCTGAGGGGGGAATCAGAACTGCTTTACATCACTCACTCCATCCATAGGATAT
AGSTCCAGAAAGATATGGCTGTTTTATAAATCTGATGGTCCGATAGACATGACAACTACTTACGTATGATCTGCTTGGC
TAAACATCCGTGATACCCAAAGCTGCCAAAGGTCAGTTTGAAACCAATACCCATCTACTTGGCTCCGCTGGTGTCTTG
ATCTCAAGCTTGTGCAATTAATGAAAGCTGGAATTAAGTCCACAGACAGTCTGTATAGAGCGAAACATTTACCAATACGC
CTGTGTGATACCAAGCTCTTTGAGCGCTCTCGGAAGGAGGCTGTATAGAGCGAAACATTTACCAATACGC
TAAACAACTATGTCGCAATTAATGAAAGCTGGAATTAAGTCCACAGACCAACAAATAACAGAAATAATTTACCAATACGC
GGCCCGCGCGGCACTTTATACAACTGGCAATCAATTTGTGCAATCTACAGCAAGCTCATGCAATCTCCGCGGAATGGA
ACAAACCTGGAACAGATCGTAGAAGAACTCGAGAACATTCGTATTAATAAACCAATCATTTACCAACAACTCCGAGGGA
CTCTGAGATGATGTGCACTCATTCACTGCGCGGAGTTCTTTATTTAAACAACCTCAAGTTTATAAGCATCTGGAAT
GGAACTGGAACAACAGAGGAACAATCACTGCTGTCTGGATTAGCAATTAATATGGAAGCAAGTGGCAAAAG
CTATGTACCCCGCTATTGGCGGCAATAATAGATGCTAGTAAATTAATACCGGATTTGCTGACACCGGACGGAGAATAA
TGAACACAGAGATATTTCAGCTCGCGGAGCGCATAGAGATACTGGAGAAGTAGGCTGTAAATATAAGCTGTGAAGATA
ATAATCTTTCCGAGATAGACACAACTGTTGAGCTTCACAGTATGGGAATCAACAGCTTCAGGCAAGAGTGTCT
GGCCGTCGAGAGATACCTCAAGATCAACAATGCTGGGCATATGGGATCTCCGTAAACTATATGACATACCGCGCTGCC
TGAACCTGGAGCTGGTCTAATAATCTCTGGATGAATTTGGCAATAATCACTGATGGATGAATGGGAACGGGAATTTGCAACT
ATAATGTTGATTTATATCTGATCGGAAGATCTCGAACCAACAGGAGAACTGCTGGAACTGGACAAAGTG
GGCACTCAITGTGGAACCTGGTTTGACATATCAATCTGGCTGTGTAAAGCTTTACAA

(여기 및 아래에 표시된 모든 140CF 디자인에 대해, 140CF 유전자는 5'의 코작 서열(GCCACCATGG/A) 및 SalI 위치를 포함하는 "TTTCAGTCGACGGCCACC" 서열 및 3'의 송지 코돈을 포함하는 TAAAGATCTTACAA 및 BglII 위치 서열로 플렌킹될 수 있다)

도면29a

CON OF CON-S-2003 (829 a.a.)
 MRVGIQRNCQHLWRWGLIFGMLIICSAAENLWTVYGVVWKEANTTLFCASDAKAYDEVHNVWATHACVPTDPNPOEIVL
 ENVTENFMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTINCTDYNATNTNNEEIKNCSENIITEIRDKKKVYALFYKL
 DVVPIDNNSYRLINCNNTSAIQACPKVSEFPPIHYCAPAGFAILLKCNDDKKNFGPCKNVSTVQCTHGIKPVVSTQLLNGSL
 ABEELIIRSENITNNAKTIIVQLNESVEINCTRPNNTRKSIIGPGQAFYATGDIIGDROAHNCISRTKWNKTLOOVAKKLRE
 HFNKIIIFNPSSGGDLITTHSFNCGGEFFYCNITSEFNSTWNGTNTITLPCRIRQIINMWQGVGOAMYAPPIEGKIRCTSNIT
 GLLILTRDGGNNNTETPRPGGDMRDNRSELYKVKVIEPLGVAPTKARRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITL
 TVQARQLSGIVQQSNLLRAIEAQOHLQLTFVWGIKQLQARVLAVERYLKDQQLGIGWGGKLICTTNVPWNSSWSNKSQDEI
 WDMTWMEDWEDEINNYTDIIYSLIEESQOQEKNEOELLALDKWASLWNPDTINWLYIKLFIIMVGGGLIGLRIVFAVLSIVNR
 VRQGYSPLSFOTLI PNPRGPDREGEIEEGGEQDRDRSIRLUNGFLALAWDLRLSCLFSYHRIRDLILIAARTVELLGRGWEA
 LKYLWNLLOQWQZELKNSAISLLDITTAIAVAEGTDRVIEVQVRCRAILNIPRRIRQGFERALL

*밀출된 아미노산은 140CF 디자인에서 결실되어야 하는 용량 도메인을 지칭하고, 빨간색으로 밀출된 'W'는 C 말단에 있는 마지막 아미노산이고, 'W' 뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면29b

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

Nick name: 006

MRVGIQRNCQHWRWGILIFGMLIICSAENLWVTYVYGVPMKEANTILFCASDAKAYDTEVHNWATHACVPTDPNPQEIIVL
ENVTFENFMWKNMVEQMHEDIISLWDQSLKPCVKLTCLCVTLNCTDVNATNNTNNEIRKCSFNTTEIRDKKKVYALFYKL
DVVPLDDNNSYRLINCNTSAITQCPKVSFEPIHYCAPAGFAILKCNDDKNGTGCKNVSTVOCTHGLKPVVSTQLLNGSL
AEEETIIRSENITNNAKTIIVQLNESVEINCTRENNTKRSIRIGPGAFAATGDIIGDIRQAHCNISRTKWNKTLOQVAKKLRE
HFNKTIIFNPSSGGDEITTHSFNCGGEEFYCNTSELFNSTWNGTNNITILPCRKQIINMVGQVQAMYPPIEGKIRCTSNIT
GLLTRDGGNNNTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTAKATLVQARQLLSGIVQQSNILRAIEAQOHLIQLTV
WGKQLQARVLAVERYLKDQQLGNGSGKLICTTNTVFNSSWSNKSQDEIWDNMTWENDKEINNYTDIITYSLIEESQHQEK
NEQEILALDKWASLWNNWFDITNMLW*

*후분색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 간단히 표시한 것이다.

도면29c

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

Nick name : 006

TTTCAGTCCGACAGCCACCATCGGGTCATGGGGATACAGAGGAATTGCCAGCACTGTGAGGTGGGAATTTTGATATTGGGAT
GCTCATATCTGCTCTCCGCTGAGAACTCTGTGGTCACTGTGTATTACGGCGTCCCGCTGGAAGAGCTAATACCTACCCTG
TTTTGTGCAAGCGCGCCAAAGCATACGACACCGAAGTCCACATGTCTGGGTACCCACGGCTGTGTACTACTGTATCCAAATC
CCAGGAAATTTGTTCTTGAACGTAACGGAAACCTTTAAACATGTGGAAGATAATATGTTGGAGCAATGCACGAGATATAAT
CAGCCTGTGGACGAGTCCCTCAAACCATGCGTTAACTCACTCCACTCTGCTGACTCTGAACTGTACCGACGTGAACGCAAC
AATAATCAACAACATGAGGAGATAAGAAATGTTTCAATTAATAAACCACTGAGATACGGATAGAAAAAAGAGTTTATG
CACTCTTTACAGCTCGACGTGGTCCCCATAGACGCAATAATAGCTACCGACTCATTAATTGCAATACAGCGCTATAACCCA
GGCATGCCCCAAAGTTTCCTTCGAGCCCATACCGATTCACTACTGGCACCCGCGGATTCGCCATTCTTAAATGCAATGACAAG
AAGTTCAACGGCACCGGACCCCTAAGAACGTAAGCACTGTTCAATGTACACATGGAATTAAGCGGTAGTCAACGAGCTCC
TCCTCAACGGAGCCTTGCAAGAGAGAGATCATTAATCAGGTCAGAAAAATACATAACACCGGAACAATCATTTGTTCACT
GAATGAGTCTGTAGAAATCAATTTGTACCCGCCCTAATAATAACACAGAAAGTCAATTAGGATCGGACCGCGGCTTTCTAC
GCAACCGGAGATCATCGGGATATACGACAGCGCCACTGCAAGATTTCTAGAACTAAGTGAATAAACTTTGCGAGCAGGTAG
CCAAGAACTGCGGGAACATTTAATAAGACAATCATCTTCAATCCAAGTAGCGAGGGACCTTGGAAATCACTACATTCCTT
TAACGTGGGGCGAGTTTCTACTGTAACTCTGAACTGTCAACTGTTCAACTCAACATGGAATGGCACTAACAATACTATACTCTT
CCTTGCAGAAATRAAACAGATTATCAACATGTGGCAGGTGTGGGCAAGCAATGTATGCCACCAATCGAAGGCAAAATAAGAT
GCACCTCCAATATTACCGGATCCTCCTGACACGGGATGCGGAGGTGTTAAGATCGAGCCCTTGGAGTTGCCCAACCAAGCTATAAOC
TTGACCGTGCAAGCCAGCGAGTTGTTGTCAGGTATCGTACAGCAATCTAATCTTTGAGAGCCATTGAGGCTCAGCAGCACC
TCTTGCAGCTTACCGTCTGGGCGCATCAACAACCTTCAGGCACGCTCCTGGCGGTAGAGGCTATTGAAAGACCAACACTTCT
CGGATCTGGGGTGTCTTGGAAATTTGATCTGCACCAAAATGTGCTTGGAAACAGAGCTTGGTCAATAAAGCAAGACGAA
ATATGGGATAACATGACATGGATGGATGGATAAAGAAATTAATAATTACACTGACATATTACTACTTATOGAGGAATCAC
AAAATCAACAGGAAAAAATGAACAGGAACCTCTTGGCTCTGGACAAATGGGCTTCACTGTGGAACCTGCTCGACATCACAATTC
GCTCTGGTAAAGATCTTACAA

도면30a

CONSENSUS A1-2003 (845 a. a.)
 MRVNGIQRNQHLRWGTWILGIIICSAENLWTVVYGVFWKDAETTLFCASDAKAYETEMHNVNATHACVPTDPNPQEIHL
 ENVTEEFNNMKNMVEQMHTDIISLWQSLKPCVKLAPLCVTLNCSNVNVTNTTTHEEIKNCSFNMTELRDKKQKVYSLFY
 RLDVVQINENNSNXYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILCKDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLL
 LNSGLAEVEEVIIRSENITNNAKTIIVQITKPYKINCTRPNNNTRKSTRIGPQAFYATGDIIGDIRQAHCVSRSEWNTLQKVA
 KQLRKYPKNTIIFTNSSGGDLIITHSFNCGGEFFYCNSTSLFNSTWNNGTMKNTITLPCRKQIINMWQAGQAMYAPPIQGV
 IRCESNITGLLLTRDGGNNNTNETFREGGGDRDNWSELYKVKVIEPLGVAPTRAKRVVEREKRAVGIGAVELGFLGAAGS
 TMGAASITLTVOARQLLSGIVOQQSNLIRAJEAQQHLLKLTWGIKQLOARVLAVERYLKQQLLIGMGGCKLICITTNVFNWSS
 WSNKQNEIWDNMTWLQWDKEISNYTHIIYNLIEESONQOEKNEODLLALDKWANLWNEFISNWLWYIKIFIMIVGGGLIGLRIV
 FAVLSVINRVROGYSPLSFQHTBNPRGLDRPGRIEEGGQGRDKSIRLWSGFLALAWDDLSLCTFSYHRLRDFILIAARTVE
 LLGHSSLKGLRLGWEGLYLWNLNLLYWGRELKISAINLVDTIAIAVAGWTRVIEIGQRIQRAILLHI PRIRQGLERALL

* 밑줄친 아미노산은 140CF 디자인에서 결실되어야 하는 용할 도메인을 지칭하고, 빨간색으로 밑줄친 "W"는 C 말단에 있는
 마지막 아미노산이고, "W" 뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면30b

Con-A1-2003 140CF.pap (629 a.a.)

Nick name: 001

MRVNGIQRNCQHLLRWGTMILGMIIICSAENLWTVYGVYVWKAETTLFCASDAKAYETEMHNVWATHACVPTDPNPQEIHL
 ENVTBENMWKNNMVEQMHTDIISLWQSLKPCVKLTPLQVTLNCSNVNVTNNTTNTHEETIKNCSFNTTTELRLKQKQVYSLFY
 RLDVVQINENNSNSYRLINCSNTSAITQACPKVSFEPIPIHYCAPAGFALLCKCKDEFNGTGPCKNVSTVQCTHGKPKVYSTQLL
 LKSLAEEVITIRSENI'TNNAKTIIVQLTKPVKINCTRPNNNTRKSIRIGEPQAFYATGDIIGDIPQAHCVSRSEFNKTLQKVA
 KQLRKYFNKTIIFTNSSGGDEITTHSFNCGGEFFYCNTSGLFNSTWNNGTMKNTITLPCRIKQIINMQRAGQAMYPPIQGV
 IRCSNITGLILLTRDGGNNNTNETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTRAKTITVQAPQLSGIVQOOSNLLRAIEA
 QOHLKLTWGTIKQLOARVLAVERYLKDQQLIGINGCSGKLICTTVPNSSWSNKSQNEIWDNMFWLQWDKEISNYTHIYNLI
 EESQOQEKNEQDILLALDKWANLWTFDISNWLW*

*부분색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 간단히 표시한 것이다.

도면30c

CODON-OPTIMIZED Con-A1-2003.seq

Nick name: 001 (1918 nt)

TTCCAGTCGACGACCGCCATGAGGTTGATGGGAATCCACGGAACTGCCAGCATCTTCTCCGGTGGGAACCGATGATATCGGGAAT
GATTAATAATCTGCTTCCGCTGGAAACCTCTGGCTGACAGTCTGACTGCTGATGGAAGCAAGCTGAACCACTCTCT
TTTGTGTTCCGATCTAAAGCTTACGAAACCGAGATGSCAAATGTTTGGGCAACCCAGCTGGTGGCCAACTGATGCTTAATC
CACAGAATAATACATCTGGAGAAATGTTACTGAGGAATTTAAATGTTGGAAATAATATGCTGAGACAAATGACATGACATGATCAT
TTCACTCTGGGACCAATCACTCAAAACCTGCTTTAACTTACCCCTCTGCTGACCTCAATTTAGTACAACTCAACGCTACACA
AATAATACAAACCACTCAGCAAGGAAGAAATTAATAATGCTCTTTAATATGACCATGAACTCTGCGACAAAACAAATGCAATATG
TCTATTCACTGTTTTATAGCTGGAAGTCTCCCAATCAACGAGAACTATCAACCAATCCCTATCCACTACTGTGGCCCAAGTGGATTCCGAAATCTGT
AAGTCTATTACCAAGGTTCTCTTAAGTCTCTTTGAAACCAATCCCTTCCAAAGCTGTAGCATTCGCAATCAACCAAGTAG
TCAGCACTCAACTGCTCTGAAATGGCTCACTGCGAAGAGAGGATGATTCGGAAGACGAGCAATATCTTACTAGCAACAAATACTAACATCGGAGAC
AATAATTTCAATGACGAAACCACTGAGAAATCACTGTACTAGCAACAAATAACAACAAGAATACTATCAGAAATGGCCCTC
GGACAAGCCTCTTAGCCAAACAGGAAATCATAGGTGATCATGACAGGCCCAATTAAGACAACCAATATTTACTTACTCTCCGAGAGTGAATCTCGA
CACTCCGAAAGATGGCAAGCAGCTGAGAAATATTTAAAGACAACCACTTCAAGAACGCTTCAAGAACGAGTGGGAATPAAAA
AATAACCACTCATAGCTTTAATTTCTGGGGCGAATTTCTTACTGTAAACACATCTGGCTCTTAACTCTACCTCGGAGGTGATCTCGA
ACCATGAAATAATCATCACTTCCGACAAATAGCAATCATTAACATGTGGCAGGACGAGCAAGCCATGTATGATGCC
TCTCCATTAAGTGTGATTCGATGTGAAGCAATCACTAGCACTTCTCTGACCGGGATGGCGGAATAATATATCAATGA
GCAATTCAGACCCGCGCGGCGGATATCGCAGACAAATTTGGCAGATGCACTTTATAAATCAAGATGATTAAGATTGAGCCCTT
GGAGTTGCCCTTACTAGACAAACAACTTGCCTTCAGCCGAGCACTGCTCAGGAATCTGTCGAGCAAGCAAGTAACTCTC
TCCGAGCTACTGAGGACCAACAACTCTTGAAATTTGACCTTACCCCTATGGGAATCAAGCAATTCGACGCTAGGTTTGGCTGTGGGA
ACGCTATCTCAAGATCAGCAGCTTCTGGGAATCTGGGGATGCTCTGGGAAATGATATGTACTCAAAAGTACCTCGGAACATCA
AGCTGGGTAATAAAGCCGACAGCAATTTGGGATAAATACCACTGGCTCGAGTGGACAAAGAAATTTCTTAATATATCTCATTA
CTATPATACATCTGTCAAGAAATCAAGACCAAGCAAGAAAGATGACGAAGAGATCTTCTGGCCCTTGGACAAAGTGGGCTTAACTT
GTGGAACTGGTTTGACATTAGCACTGGCTGTGGTAAGATCTTACAA

도면31a

CONSENSUS C-2003 (835 a.a)
 MRVRLRNCCQWINGLGFWMICNVVGNLWTVYGVVWKEAKTTLFCASDAKAYKEVHNWATHACVTFDPNPOEIVL
 ENVTFENFMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTINCTNATNTMGKIKNSFNITELRDKKQKYALFYRLDI
 VELNENNSYRLINCNTSALTQACPKVSFDPIPIHYCAPAGYALLKCNKTNGTGPCNNVSVQCTHGIGIKFVSTQLLINGSLAE
 EELIRSENLTWNAKTIIVHLNESVELVCTRENNTKKSIRIGPGQTFYAIGDIIIGDIRQAHCNISBDKWNKTLOKVSKKLKEHF
 PNKTIKTFEPSSGGDLEITTHSPNCRGEFFYCNTSKLFNSTYNSTITLPCRKQIINWQEVGRAMYAPPIAGNITCKSNITG
 LLLTRDGGKNTTETFRPGGDMRDNRSELYKYKVEIKPLGIAPTAKRRVVEREKRAVGIGAVELGFIAGSTMGAASITLT
 VQARQLSGIVQOOSNLLRAIEAQHMLQLIVWGIKQIRVLAIERYLLKDQQLLGWCSGLICTTAVPMNSSWSNKSQEDIW
 DNMTWQWDREISNYTDTYRLLEDSONQCKEKNEKDLIALDSWKNLWNNWFDITNWLWYIKIFIMIVGGLIGLRITFAVLSIVNRV
 ROGYSPLSFQTLTPNPRGPDRLGRIEEEEGEQDRDRSIRLVSGFLAWDDLSLCLFSYHRLRDELLIAARAVELLGRSSLRGL
 QRGWEALKYLGSLVQVWGLELKKSAISLLDTAIAVAEGTDRIIELTIQICRAIRNIFERRIRQGFEEALQ

*밑줄친 아미노산은 140CF 디자인에서 결실되어야 하는 용합 도메인을 지칭하고, 빨간색으로 밑줄친 "W"는 C 말단에 있는
 마지막 아미노산이고, "W"뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면31b

Con-C 2003 140CF.pcp (619 a.a.)
 Nick name: 003
 MRVGRILRNCQOWIWIWGLGFWMLMNCNVVGNLWVIVYVYGVYVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPOEIVL
 ENVTFENNMKNMDVDMQHEDIISLWDQSIKPCVKLTCLCVTLNCTNATNATMTMGEIKNCSEFNITTELDRKKQKQVYALFYRLDI
 VPLNENNSIRLINCNTSALTQACPKVSEDFPIPIHYCAPAGYAILKCNNTFNCTGTCNNVSTVQCTHGIKPVYSTQILLNGSLAE
 EETIRSENLTNNAKTIIIVHLNESVEIVCTRPNNNTKRSIRIGPGOTFYATGDIIGDIRQAHCNISBDKWNKTLOKYSKKLKEHF
 ENKTIKFEPSGGDELEITHSENCRGEFFCYNTSKLENSTYNTNSTITLPCRIKQIINMWQEVGRAMYAPFIAGNITCKSNITG
 LLLTRDGGKNNTETFRPGGDMRONWRSELYKYKVEIKPLGIAPTAKATLVQARQLLSGIVQQSNLLRAIEAQOQMLQITVW
 GIKQLQTRVLAIERYLKQOQLLGWGCSSGKLICTTAVFWNSSWSNKSQEDINDNMTWQNDREISNTDTIYRLLEDSONQOEKN
 EKDIILALDSWKNLWFWFDTNMLW*

*부분색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 간단히 표시한 것이다.

도면31c

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

Nick name: 003

TTCACTCGACAGCCACCATCGGATGAGAGCAATCTCGGGAATGTGAGCAATGGTGGATCTGGGCATACTCGGATTTCTGGAT
GCTTTATAGCTAATGTTCTGGGAACCTGTGGTTACCGGTATACCTGATATCTATGGGTTCAGTCTGGGAAGGAGGTAAACAAGCTG
TTCTGTGCAAGTAGCGCCAAAGCCTTCAGAGAAAGATGACACACAGCTGTGGCTTACCCAGCTGTGTCTTCCAAACGATCAAAAC
CCAGAAATGCTCTCGAAGCTGACTGAATATTTAACTCTGGATGGAAGAATGATGTGATCAGATCGATCGACAGATATCAT
TTCAATGTGGCAACAATCTGAAACCACTGCTGTAACACTGACCCCTTCTGGTAAACACTTAATGACCACTTAATGACCACTTAATGCC
ACCAATACTATGGCGAATAAATAAATCTGACTTAATCATATACACAGGAATCCGGATAGAAACAAGAAGTCTACGGCTCT
TTTACCGACTCGAATATCGTCCCACTTAACCGAATAATGATTAACCGCTGATTAACCTTAACACATCAGCCACTAGCAAGCTTG
CCCCAAGTTCTTTTCGACCCATCCCAATTCATATTGTGCCCCGTGGATACGCTATACCTTAATGTCAACATAAACAACATTT
AATGCAACCGGACCATGTAACAGCTGACTACGTACAGTACAGCAGGAATAAATCTGTTCTCAACCCAGCTTCTCTCTTA
ACGGCTCATTTGGCGGAAGAATAATATATATCAGATCAGAAATCTGACACAAATGCAACAATGCAAAACACTCATGTGCACTCAATGA
ATCCGTGGAATCTGTGCACACAGCAAAATAACATACCCGGAATTAATCAGGAATTAATCAGGAATGGCGTGGCAGACATTTACGCTACA
GCTGATATAATTCGGCATATAGCAAGCCCATTAACATATCAGAAAGCAAGTGAATAAGACTCTCGAAGAAGTTTCTAAGA
AGCTTGAAAGCACTTCCCAATAAAGCAATTAAGTTCGAGCCTCTTCAGGAGGAGCCTTGAGATCAACAACACTCTTTTAA
TTGTAGAGGGGAGTCTCTTATTGTATATACATAAGACTCTTTTAAAGACTCTTTTAAAGACTCTTTTAAAGACTCTTTTAA
TGCAATAAAGCAANAATAATCATCTGTGCGAAGAGTGGCGAGCAATGTAGGCCCTCCCATCGCAGGACATATCATATGTA
ATCCAAATATCTGGCTCTTCTGTACACGGACGCGGGAAGATAACGTAGACCTCTCAGACTCGCGAGGCGGATGCG
CGATTAATTTGGCGAGCGAGCTCTACAGTATAAGTCTGTGAATCTGTGAACTAGCAACCAATCCAAATCTCTGCGAGCTATCGAAGCCCAACATCTC
ACTGTTAGCTAGACAGTGTCTCCGCAATGTGAACTGTGCAACAGCTGTGGCGTAGAGATATGAAAGATCAGCAACATC
TCCAGTTTACCGTTCTGGGAAATCAACAAATTCGAACACAGAGTCTGGCGTAGAGATATGAAAGATCAGCAACATC
GATTTGGGCTGTTCAAGTAAGCTCAATCTGTACAACCTCGGTCGGTGGAACTCAGCTCGAGTACAAAGCCAGAGATATA
TGGACACACATCAATTTGATGAGTGGATCGAGAAATAAGCACTATACAGATACCAATTTATCGCTCTCGGAGGACTCACAGA
ACGACAGAGAAATAATGAGAAAGATTGCTCGCGCTTGACAGTTGGAAGAATTTCTGGAAATTTGTTCCACATATACAACTGGCT
TGTGTAAGATCTTACAA

도면32a

CONSENSUS G-2003 (842 a.a.)

MRVAGIQRNQHLHWKQGLILGLGVLICSSANLWVTVYGVVWEDATTLFCASDRAKAYSTERHNWVWATHACVPTDPENPOBITL
 ENVTENENMWKNWVQMHEDTILSLWDSKPCVKVLTCLPCLVCLNDVNTNNNTNNNTKKEIFKCSNNITTEIRDRKKKKKEVALFY
 RLDPVPINDNGNSIYRLINCNSVTSIKQACPVTFDPIPIHVCAPEGAFALLKCRDKFNGTGPCKNVSTVQCTHGHKPVVSTQLL
 LQSLAAEEETIIRSENITNTKVIIVQLNETILNCTRPNNWTRKSIKIGPQAFVATGDIGIDRQAHCNVSRVNNENMLQKVK
 AOKKIKFNKSIITNSSSGDLEITTHSNCRGFVTCYNTSGFNFNLSTNSTITLPCPKTKQIVRMWQVQGMAYVAPPIAGNIT
 CRSNITGLLLTRDNGNNNTTEFGGQNDKRWSEYKYKIVKIKPLGVATRRVRVEREKRAVGLLGLFGLGAASGTMG
 AASTITLTVQVRQLLSGIVQOOSNLRATLAEAOQHLLQLTWGHLQOLARVLAVERYIKDQLLGLWGCSSGKICITTNVPMWNTWSN
 KSYNEIDWNTWTEWERISINYTQOISLIEESQNGEKNEQDILLADKWSLNNWFDITKLWLYIKIFIMVLGGILGLRIVEAV
 LSIWNRVQQGSYSLFSQTLTHQREPRPRETEGGGQKDKORSIRUVSGFIALAWDDIRSLCISYHYRDFDLIIAARVVELLIG
 RSSLKLRLGWEGLKYLWNLWLLYWGQELKNSAINLLDTIAIAVANWTRVTEVAQRACRAILNIPERRIQGLERAIL

* 밑줄친 아미노산은 140CF 디자인에서 결실되어야 하는 용합 도메인을 지칭하고, 빨간색으로 밑줄친 "W"는 C 말단에 있는 마지막 아미노산이고, "W"뒤의 진여 아미노산은 140CF 디자인에서 결실될 것이다.

도면32b

Con-G-2003 140CF (626 a.a.)
 Nick name: 007
 MRVKGIGQNNWQHLLWKWGTLLGLVLICSSASNNLWVYVYGVFVWEDADTTLFCASDAKAYSTERHNWVWATHACVPTIDENPQETTL
 ENVTEFENWKNWVEQMHEDLISLWDESLKPCVKLTPLCVILNCTDVNVNNTNNNTKKEIKNCSEFNITTEIRDKKKKEYALFY
 RLDVVPINDNGNSSIYRLINCNVSTIKQACPKVTFDPIPLHYCAPAGFAILKCRDKKNGTGCECKNVSTVQCTHGIKPVVSTQLL
 INGSIAEEELIIRSENITONTKVIIVOLNETIETCTRENNTKRSIRIGPGQAFYATGDIIGDIRQAHNCVSRTKWNEMLQKVK
 AQLKKIFNKSIITFNSSSGGDLIETHSEFNGRGEFFYCNSTGLFNNSLNNTSTITLPCIKQIVRWQVRGQAMYAPPIAGNIT
 CRSNIIGLLITRDGNNNTETFRPGGDMRDNRSELYKYKIVIKPLGVAPTRARLTIVQVQVQLLSGIVQOQSNLIRAEFAQOH
 LLQLTWGIKQLOARVLAVERYLKDQQLIGIWGCSGKLICTTNVPNTSNKSYSNEIWDNNWTIEWEREISNYTQOIYSLIEES
 QNQQEKNEQDLIALDKWASLWNNEDITXWLN*

*부분색으로 표시된 아미노산은 결실된 융합 클론 위치의 경계를 가단히 표시한 것이다.

도면32c

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

Nick name: 007

TTCAGTCCGACAGCCACCATGCGAGTGAAGGCAATCCAGAGAGAAATTGGCAGACACCTTTGGAAGTGGGGCCACTCATCTCGGCGCT
TGTGATCATATGCTCTGCGCTCAATTAACCTTTGGGTACAGTTTATACGGCTGCGCGTTTGGAGAGCGGCACACACACTCTT
TTTTGTGCGCAGCGACCGCTAGGCTTATTCACAGAGAGGAGATAAGTTTGGGCTACACATGCAATGCGTCCGACCGATCCTAATC
CCAGGAAATCAGCTCTTGAGAAATGTTACAGAGAAATTTAATATGTTGGAAGAACAACTGTTGAACAGATGCATGAAGACATAAAT
TTCTCTCTGGGATGATCTCTGAAACCTTGGTGAAGCTTACACCACTGTCGCTTACCTGAAATGCACTGACCTCAATGTCTACA
AATAATAACCAACAATACAPAAAAGAAATCAAAATTTCTTTCAACATACACCGGAGTACCGGATACGGAATAAAAAGAAAG
AATACGCCCTGTTCTACAGACTCGATGTGTCCCAATTAATGACACGGAAATTTTCATCCGACTTATCAATTTGTAAGCT
GTCACAAATCAACAGGCCCTGCTCAAGTCACATTTGACCTATCCCATTTCACTTACTGTGCCCCGCTGGCTTGGTATTCTTT
AAATGCCGAGCAAAAATTTAACGGAACAGGACCATGCAAGAAATGCTCAACAGTTCAATGCACTCAATGGAATTAACCGATCG
TTTCTACTCACTCCTTCTCAATGGAAGCCTGGCAGAGGAAATCATATCCGAGCGAAACATACCGACAAACACAAAGT
AATCATCGTACAGCTGACAGACCATTTGAATTAATTTACGAGACCTAATAACACAGAAAGCATACGCTCGGCCCC
GGACGGCTTTCTACGCCACAGGACATTTATCGGAGATATCCCGAGCTCACTGTATGTCTAGAACAAATGGAACGAA
TGCTTCAGAGGTCAAAGCTCAGCTCAAGAAATTTCAACAAATCTTACATCACTCAGCGCGGATCTGGAGAT
AACACTCATCTTCACTCTCCATGTAAGATCAACAAATCTCAGATGTTGGCAGCGAGTCGGTCAAGCTATGTACGCCCTCAA
TCGCGGTAATATCAGATAGAGCAATATCACAGGCTCTTGTTCACAGGACGCGGAAACACACACCCGAAACCTTCAG
ACAGGAGGAGGAGACATCGGAGCAATTTGGGAGCGAGCTGTATAATATAGATCGTAAATCAACCATTTGGGTGTAGCG
CCAACTAGAGCCCGAACACTGACCGTGCAGGTGAGGCACTGCTGAGCGGCTGTCCACACCAATCCAACTTTCTTAGAGCAA
TCGAGGCCCGAGCAGCATCTGCTCCAGCTTACTGTATGGGAATTAACAGACTCAAGCAAGAGTATGGCAGTGGAGGATCT
CAAGACCGCAGCAGCTCTGGGAATTTGGGTTGCGAGGAAAGCTCATATGTACAACCAATGTCCCTGGACACTAGTTGGAGT
AATAAGAGTTACATGAATCTGGGACAAATATGACATGGATCCAAATGGGAGCGGAAATATCCACTACTCAGCAATCTATT
CCCTCATTCGAAGAGTCAAGACAGGAGAAAGATGAGCAAGACCTCTCGCCCTGGATAAATGGGCATCTCTGTGGHACTG
GTTTGACATAACTAAATGGTTGTGAAGATCTTACAA

도면33a

CONSENSUS 01_AE-2003 (854 a.a.)
 MRVKETQNWPNLWKWGTLLGLVVICASDNLWVTVYGVVWRDADTTLFCASDAKAHEVEVHNWATHACVPTDPNPQRIHL
 ENVTEFNMWKNMVEQOEIVSLWQSLKPCVKLTPLCVNCTNANLTNNVITNVSNIIGNITNEVRNCSFNMTELRDKK
 QKVHALFYKLDIVQIEDNNSYRLINCNTSVIKQACPISFDPLPIHYCTPAGYAILKCNKNGTGPCKNVSSVQCTHGKPVV
 STQILLNGSLAEEIIIRSENLTNNAKTIIIVHINKSVEINCTRPSPNNTRTSITIGPQGVFYRTGDIIGDIRKAYCEINGTRWNEV
 LKQVTEKLKEHFNKTIIFQPPSGGDLITMHHFNCRGEFFYCNCTTKLFNNTCIGNETMEGNGTILPCIKIQIINMWOGAGOA
 MYAPPISGRINCVSNTIGILLTRDGGANNNTETFRPGGNIKDNWSELYKYKVOIETPLGIAPTRAKRRVVEREKRAVGIGAMI
 FGFLGAAGSTMGAASTILTVOARQLLSGIVQOOSNLRARIEAQOHLIQLTWGKIQLOARVLAVERYLKDQKFLGLWGCGRITIC
 TQAVPWNSTWSNRSFEETIWNNTWIEWEREISNYTNOIYEIITESQOQDRNEKDLLELDKWSLWNVFDITNWLWYIKITIMIV
 GGLIGLRLIFAVLSIVNRVROGYSPLSFQTPTHHQREFDRPERIEEGGEGGDRSVRLVSGFLALAWDDLRSLCLFSYHRLRDF
 ILTAARTVELLGHSSLRGLRGGWGLKYLGNLLLYWGQELKTSAISLLDATAIAGWTDVIEVAGQAWRAILHIFRRIRQGLE
 RALL

*밑줄친 아미노산은 140CF 디자인에서 결실되어야 하는 용합 도메인을 지칭하고, 빨간색으로 밑줄친 "W"는 C 말단에 있는
 마지막 아미노산이고, "W"뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면33b

Con-AE01-2003 140CF.pap (638 a.a.)

Nick name: 008

MRVKETQNNWPNLWKGTLILGLVICSASDNLWVTYVYGVVWRDADTTLFCASDAKAHEVEVHNWATHACVPTDPNPOETHL
ENVTFENFMWKNWVEQMQEDVTSIWDQSLKPCVKLTPLCVTLNCTNANLTNNITNVTNLTNITNEVRNCSFNMTELRDKK
OKVHALFYKLDIVQIEDNNSYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNKDNENGTGFCKNVSSVQCTHGKIPV
STOILLNGSLAEEHIIIRSENLTNNAKTIIIVHLNKSVEINCTRPSNNTRTSITIGPGQVYRTGDIIGDIRKAYCEINGTKWNEV
LKQVTEKLEHFNKTIIFQPPSGDLEITMHHFNCRGEFFYCNTTKLENNTCIGNETMEGCGTIIILPCKIKQIINMWQAGQA
MYAPPISGRINCVSNTITGILLTRDGGANTNETFRPGGNIKDNWRSELYKYKVQIEPLGIAPTRAKTLITVQARQLLSGIVQOQ
SNLIRAIEAQOHLITVWGKIQARVLAVERYLKQKFLGLWGCSCKIICTTAVFNNSTWNSRFEIWNMTWIEWEREISN
YTNQIYEILTESQOQDRNEKLELLELKWASLWNVFDTNWLW*

*부분색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 간단히 표시한 것이다.

도면33c

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

Nick name: 008

ttcagtcgacagccaccatgccagtcgaagaaacacaaatgaactggcctaatctgtggaactggggcaccctgatctcctgggttt
ggfcaattatttggcttcggagggacaaatctctgggttactgtctattacgggattcccggttggagagatgccgacactacactg
ttctggcctcagatgccaaagctcagaactgaagtgaatgttggcaacccacccctgttcttaccgaccccaacc
ccaaagaaatcacctggaaaggtagccggagaaactttaatgttggaagatgacatgggtgaacagatgaagagacgtta
cagcctgtgggtcaaaagcttgaagccttggctaaactgactccacttgggttaacacttaattgcaccaacgggaacccgaca
aaggttaacaacatcactaacgtctccaacatcatcgccacaataacgaactgagaattgagtttcaattgactacag
agctccggacaaagacagaggctccatgctctctttacaaactggacatcgctccagatcgagacaaataacagctacagact
tataaattgtaatacatccgtgattaaacagcatgcccaataagcttcgatcctattcctatccactactgactcctgcc
ggctatgctattcttgaatgcatgataagaaacttcaatggaccggacattgttagaagctgtctagtgtgcaatgcactcaag
gcattaacccagtgtgaagaccagctgctcctgaacggctctcggcagagagagattatttgcagtgagaaacactcac
caacacgcttagactatcatggtacatctcaataatcactgaagcttcaatgcaagagagagattatttgcagtgagaaacactcac
atcactatcgcccccagcacagtcttttatagacagagagatcataggaatgacagagagagattatttgcagtgagaaacactcac
caaatggaaacgaagtactcaaacagtcacagagagctttagggcaatttttttattgcaatagcacttaacttttcaacaat
tggcggagacctcgaaatcactatgacacctcaactggcggcgaaatttttttattgcaatagcacttaacttttcaacaat
acgtgcattcggaatgagaccatggaggctgcaatggacatcatcactcccatgcaagatgaacaaatcatttaacatgtggc
aaggtgctggacaagctatgtatgaccccccaatattccggtagaatttaattgcgtcagcaacatcactggcatatgctcactag
agacggaggagcaaatataacaaatgaacactccgaccagcgcgcgcaacatttaggacaactggcggtccgaactctataag
tacaagtctgacagatcgaaactcttggatagacacggactcgcggttagacatcacagtacggcccgacactcttcttg
gattcgtagcagcgaatccaaactctccgccaatcgagccccaaacacatctgcttcagctcacagtggggaatcaagca
gctccaggcacgctgctcgagtgaaagatccctgaaggatcagaatttcccttcttgggtatgttctggcacaataatc
tgcactaccgggttccctggatccaatgacatggagacccggagtttgaagagataggaacaaatgacatggatagatggg
aaagggaatttagatatacgaaacacagatatcgaatctcaccgaaaccaaatacagagatcgcaacgaaataagacct
cctcgagcttgatagtgggcatccctttgggaactggttcgacatcacaaattggctcttgstaagatcttataaa

도면34a

Wild-type subtype A Env
 00KE_MSA4076-A (Subtype A, 891 a.a),
 MGAMGIQNNWQNLWRWGTMLGLIICSVAEKSWTVYYGVVWRDAETTLFCASAKAHDKVEHNVWATHACVPTDPNPQEMIL
 ENVTEDENNWKNMVEQMHTDIISLWQSLKPCVKLTPELCVTLNCSDSNITSNSTKDSATLDMKSEIQNCSEFNMTELRDK
 KQKVISLFYRLDVQINENSSDYRLINCNTSAITQCPKVTPEPIPIHYCAPAGFALLKCNDKKFNGTGPCTNVSTVQCTHGKIP
 VVTITQLLLNGSLAEFEVMIRSENITENAKNIIVQKPEVQIICIRGNNTKSVHIGPGQAFYAIGDIIGDIRQAHCNVSRSLWN
 KTLQEVATQLRKHFRRNTKIIFTNSSGGDVELTTHSFNCGGEFFYCDTSGLEFSSWTASNDSMQEAHSTESNITLQCRIKQIINN
 WQRAGQAMYAPPPIPGIIRCESNITGLTITRDGEGNNTNETFPYGGNMRDNWRSELYKYKVVEPLGVAPTCSRVRVVEREK
 RAVGLGAVFIFELGAAGSTMGAASMTLVQARQLLSGIVQQSNLLRAIEAQQLHLKLVWGIKQLQARVLAVERYLRDQQLGI
 WGSGLICTINVPWNSWSNKSLEIENWTWQMDKEVSNYTMIIYNLLEESQOQEKNEQELIALDKWANTLWNTFNISNWLW
 YIKIFTNIVGGLIGLRIVFAVLSVINRVROGYSPLSFQTHUPNPRGLDRPCRIEEEGGEQDRDRSLRLVSGFLALAWDDLRSLCL
 PSYHRLRDFILIAARTLELLGHNSLKGLRWEGLYLWNLHAYWGRELKISALSIVDSIAIYAVAGWTDRIEIVQAIGRAILHI
 PRRIRQGLERALL

*밀줄친 아미노산은 140CF 디자인에서 결실되어야 하는 용할 도메인을 지칭하고, 빨간색으로 밀줄친 "W"는 C 말단에 있는
 마지막 아미노산이고, "W"뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면34b

00KE_MSA4076-A 140CF.pap (647 a.a)
 Nick name: 011
 MGAMGIQMNWONLWRGTMILGMLITCSVAEKSWTVVYGYVPVWRDAETTLFCASDAKAHDKEVHNVWATHACVPTDNPQEMIL
 ENVTEDEFNMWKNWSWVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTKDSATLDMKSEIQNGSFNMTTELROK
 KQKVYSLFYRLDVQINENSDDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILLKCNDDKKTNGTGCTNVTSTVQCTHGKIP
 VVTQILLNGSLAEERVMIRSENITENAKNIIVQKEPVOICIRPGNNTKRSVHIGPGQAFYATGDIIGDIRQAHNCNVSRELWN
 KTLQEVATQLRKHFRNNTKIIFTNSSGGDVEITTHSFNCGGEFFYODTSGLFNSWTASNDSMOEAHSTESNITLQCRITKQIINN
 WQRAGQAMYAPPIPGIIRCESNITGLILITRDGEGGNSTNETFRPVGGRNDRNWESELYKYKVVEPLGVAPTKSRITLTVOARQ
 LLSGIVQQQSNLIRATEAQOHLLKLTWVGKLOARVLAVERYLRDQQLLGWCCSGKLICTTNPVWNSWSNKSLSDEIWNMTW
 MOWDKVSNVTOMIYNLLESQOQKEKNEQELLALDKWANLWNNFNISNWLW*

*부분색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 간단히 표시한 것이다.

도면34c

CODON-OPTIMIZED O0KE_MSA4076-A 140CF.seq (1972 nt.)
 Nick name: 011
 ttccagtcgacagccaccatggggcgaatgggaatccacatgaaactggcagaaactctggcgatggggcgcacacatgatcctgggtat
 gctcatctctgctgtgtgcagaaaactcattgggtatgacgtctactacgggctaccagtgctggcgggacgcggcgaaccactctc
 ttctggcgctccgatggcgaagcacaccgataaagaagtcacaaatgtttgggtaccgatccgatccgatccgaatccctaaac
 cacaagaaatgatactccaaagcttactgaagacttcaacatgttgaaataattctatgttgaaagatgacacacgacataat
 atcactgtggatcagctgtcacaacccctgtgtcaaatgtgacccctctgggttgcactgaactgttccgactcaaatatcact
 tctaatccaacgagcaatagtagcaaaagactccgcaaccccttgatatcaaaagcgaatagacacgttctcatttaatatgacca
 cogaactgagagataaaagcagaaggttatctctgttctatcgattggacgttgacattacgaaatagcggcattt
 ccgactcatttaactgcataacacgaatcacaggttggccaaaggttaacatttgagccaatcccatttcactactcgcc
 cctgcaggatttgcctcctgtaatgcaacgataagagtttaattgggacaggaacctgacacacgtctccacccgtgcatgca
 cccacggcataaaacctgtgttacccacaatgtctgctcaatggatcacttgaagaggaactgcatgtccgtctgaa
 catcactgaaatgccaaaatattatagttcagttcaagaaacccgtccagatcatttgcatttccgctggtacaaactcgc
 aactcagtcacattggcccgccagcctttctatgcaacccggagaaattatagggcacatcagacaggacattgcaacgtca
 gccgggaattgggaacaaahctttgcaggaagtgtgctactcagctgcgaaacatttcagaaacatacaaaagattattttcac
 taattcatcagggctgacgtggagatcactaccatttcaatttaactgtggcggaatttcttatttggatagctctgggctc
 tttaattcctcatggactctagcaacgattcaatgcaagagcaattccacagaaagtattatcacactgcatgcccgaattt
 aagaaatcataatagtggcagcggccggtcagcaatgtagcaccctccatccccgaattatttgcattgtgagcttaatat
 cactggcctcattctgaccccgagacggtgggaagtpaataattctacaacgactttcagacccgtaggagcaatagcga
 gacaaattggcgatccgaactgtataatataaagtgtgaggttagacccttggagtgscctccaccataatcagaaacctga
 ctgtgacggcagccgaacttctgagcgaatagtcacaacagcaatccaatcttctgagagctatagaagccacagacacctgt
 taaacttacgggtggggaatacaaatgacggcaagagctgtggcagtggaagacttccctggaactcttctgtcaaacagagctctgacgaaat
 atctggggaatttccggttaagtgtgttgacgacaaacgttccctggaacttctgagagcttctgagagacacacacactctggga
 gggaataatagacatggatgcagtgggacaaagagtagcaactatcacagatcattcacaacctcctcgaaagatctcagaa
 tcaacaggaaaaaacgaacgaactgctcgccctcgatagtagggctaacctctgggaacttggtttaatttcaaacctggttg
 tggtaagaatctttacaa

도면35a

Wild-type subtype B
 QH0515.1g gp160 (861a.a)
 MRVKEIRRMNCQRLRRWGTMILGMLMICSAEQQLWVTVYGVVWKEATTTLFCASDAKAYVTEKHNWATHACVPTDPNPQEVVL
 ENVTENFNWKNMVEQNHEDIISLWESLKPCKVLTPLCVTLNCTDKLRNDTSGTNSSSWVKVQKEIKNCSFNITIGIRGRVQ
 EYSLFYKLDVIPIDSRNNSNNSSTEFSSYRLISNTSVITQACPISPEPIHYCAPAGFALIKNDKKFNGTGPCKNVSTVQCT
 HGIKPVSTQLLINGSLAEEVVIRSENFTNNVKSIIIVQLNKSVINCTRPNNTRKSIHIGAGKALYTGELIGDIRQAHNLSR
 AQWNTLKQIVIKLREQFNKTIIVFNQSSGGDVEIVMHSFNGSGEFFYCNSTQLFNSTWNGNDTWNTWKDTINDNITLPCRIRKQ
 IVNMQKVGKAMYAPPPIRGQIRCSKITGLILTRDGGTNGTNETETFRPGGNMKNDRSELYKYVVKIEPIGIAFTKAKRRV
 QREKRAVGTIGAMFLGETGAGSTMGASLFTLVQARLLSGIVQQNNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYRDQ
 QLLGWGSGRLICTTNVPWNTSWSNRSLNYTWNMTWMDREINNYTDYITLLEDAQOQKNEQELLELDKWSLWNVFDI
 TNWLWYIKIFIMIVGGLIRIVFAVLSIVNVRQGYSPISLQTHLPARRGPDREPIGEGGERDERSYRLVHGFALVWEDL
 RSICLFSYHRLRDLILIVARTVEILGQRGWELKYNWNLLEYWSLEIKNSAVSLVDTIAIAVAEGTDRIIIRRIARIRAFTHIPT
 RIRQGLERALL

*미표기 아미노산은 140CF 디자인에서 결실되어야 하는 융합 도메인을 지칭하고, 빨간색으로 미표기된 "W"는 C 말단에 있는 마지막 아미노산이고, "W" 뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면35b

QH0515.1g 140CF (651a.a)
 Nick name: 012
 MRVKEIRRNQORLRWGTMLLGMIMCSATEQLWVTVYGVPMWKEATTLFCASDAKAYVTEKHNVWATHACVPTDENPQEVVL
 ENVYTFENMWNMMVEQMHEDEIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSSTSSWEKVQKGEIKNCSFNITGIRGRVQ
 EYSLFYKLDVPIPIIDSRNNNNSTEFSSYRLISCVTSVITQACPKISTEPPIHYCAPAGFAILKCNDRKFNKGTGPKKNVSTVQCT
 HGIKPVVSTQILLNGSLAEVEVIRSENFTNNVKSIIVOLNKSVVINCTRPNNNTRKSIHIGAGKALYTGELIGDIRQAHCNLSR
 AQWNTLKQIVIKIREQFCNKTIIVFNQSSGGDVIVMHSFNCGGEFYCNSTQLENSTWNGNDTWKDTNDITLPCRIRKQ
 IVNMWQKVGKAMYAPPPIRGQIRCSSKITGLILTRDGGTNGTNETETFRPFGGNNKONRSELYKYKVVKIEPLGIAPTKAKTLIV
 QARLLSGIVQQONNLRRAEQOHLLOLTWVGIKQIQARVLAVERIARDQQLIGWCCSGRLICTTNVPMNTSWSNRSINYIWD
 NMTWQWDEIINNYTDIYITLLEDAQNOQEKNEQELLEIDKWSLWVFDIITWLIW*

*푸른색으로 표시된 아미노산은 결실된 응화 절단 위치의 경계를 간단히 표시한 것이다.

CODON-OPTIMIZED QH0515.1g 140CF.seq (1984 nt.)

Nick name:012

ttcagtcgacgagccaccatgagagttaaaagaaatcagacgcacaactgtcagaggttcagagagatggggaacgattctcttgggcgat
gcttgattattttcaggtccaccggaacagacttgggttaacccgtgtactatgggtgtaccttattggaaagaagccactacacccctg
tttttggscgtccgacscaraagctcagtgaaacagaaacacagactgtggcgccacacatgcatcgctgccracagatccraatc
ctcaggaagtcgttttggaatatttacacagaaaatttataatcttgrraaacaaatattggttagagagatgcatagagattcat
ctcactgtgggaacaactcttgaaacctttgtcacaacgacccacctttgcgttaacacacttactgtactgataagcttccccaat
gatagctccggaacaaattcuaagcactgggaataagtcgaaagggccaaatcaaaaattgcttcatttaacatcactactacoggtta
tcagaggggggtacaggaattattctttttctcaaaactcgagcttccoraatgcatccagaataactcaataaataagcac
agaatttagttagtatttcgcctttaagctgcacaacacgcgctgattacacagctggcctaaaaattctcttttgagcccatctct
attcactactggccaccagcggtgtcgcatctctcaaatgtatcagcaagaaattttaaaggaaacggcacctgttaagaattctct
ccacgcttcaatgcactcatgaaatcaacscctgtttctcccaacttttccatgcttcacgttttaataatcgctgttattgtacaaagtct
gattcgctccgaatttttacaacacactcaactcaatcttcctcagcttttaataatcgctgttattgtacaaagtct
aacanaarccacaghaaattcattcacaagggccgggaacactctgtataccgggaattttttggagacactcagacaagcac
actgttaacttgastgcggccacgagcaacacattgaaacagactctgtatgaactcagacagcagatttccgggaatttaagactat
cgtgttttaactcagagctccggggtgatttgcraattctgaattgcacttctttaaattgtgggggtgaaatttttttactgcaattct
acacaattgttttacacgacacttgaaagccaaatgacacatgnaatgacaccttggaatagatcagacaataatgatatcttactcttc
cgtgcgaagaacactcgaaagccaaatgacacatgnaatgacaccttggaatagatcagacaataatgatatcttactcttc
ttcttccaaagatatacacaactcgttaattgtggcaaaaagtggcgcaagccactgtacacacacatttaagaggacaaaattcggctg
ggcaacatcgaaggataactcggagaagtgaactttacagattaaagatcggcagcagcagacggacacccatttaagaggacaaaattcggctg
ctaaacactcaccgtgcgagctgagattgtcgttttccagggatagttcccaacaaacagacacacttcttagagccaatggaagcaca
acaaacacttctcgaattgacagctgtgggaatttaacagattgcagcccggttttcgctgtcgaaaggatcttagagatcag
cagcttttttgggtattctcggggtgttcaggccgctctatgtacacacaaatttcccttggaaatccttagagatcag
tttaattttattttgggaacataatgcattgattgcaatgggatagagaaatttaataactacacggactacattacacacttctgga
ggacccccaacattcagcaggagacagcaggaggaaactctctcgaattggataagtggggatcactgtgcaattgtgttcgataata
acnaattggctttggtaagaactcttaca

도면36a

Wild-type subtype C
DU123.6 gp160(854 a.a)
MRVKGIGORNWPQWINGILGFWMIIICRVVGNLWTVVYGVVWTEAKTTLFCASDAKAYEREVHVWATHACVPTDPNPQEIIVL
GNVTENFNWKNMDVDMQMHEDIISWDQSLKPCVKLPCLCVTINCTDVKNVATSNGTYYNSIDSMNGEIKNCSENIITTEIRDK
KOKVYALFYRPDVVPLNENSSSYILNONTSTTQACPKVSDPPIPIHYCAPAGYAILKCNKTFNGTGPCRNVSVQCTHCIRK
VYSTOLLNGSLAEEEEIIRSENLTNNAKTIIIVHLNESIEIVCTRPNNTRKSIIRIGPGQTVVATNDIIGDIRQAHCNISKTKN
TTLEKVKELKEHFPKAITFQPHSGDLEVTHTSFNCRGETFYCDTKLFNENLNTTNTTLTFCRIKQIVNMWQGVGRAMY
APPEVEGNITCNSSITGILLVRDGGNTSNSTPELFRPGGNNKDNWRSELYKYKVVEIKPLGVAPTAKRRVVERERAVGIGAVL
FGFLGAGSTMGCAASTILTVQARQLLSGIVQQQSNLIRAIERAOQHMLQLTVMGIKOLQARVLAIERYLKDQQLLGLWGCSKGLIC
PTTPWNSSWSNKSQTDIWDNMTMOWDREISNYTGTIYKLEESQNOQEKDLDLSDSKNLSWTFDTNWLWYIKIFIMIV
GGLIGLRITFGVLISIVKRVRCYSPLSFQTLTPNPRGLDRGLRIEEGGQDKRSIRLVNGFLALAWDDLRSLCTFSYHRLRDF
ILVAARAVELLGRSSLRGLQGWAEALKYLGNLVQYGGLELKRRAISLFDITIAIAVEGTDRLILEVILRIITKAIRNIPTRIQGFE
AALL

도면36b

DU123.6 140CF (638 a.a)
 Nick name: 013
 MRVKGQRNWPQWILGILGFWMIIICRVVGNLWVTVYGVFVWTEAKTLFCASDAKAYEREVHNWATHACVPTDPNPQELVL
 GNVTFENFNWKNDWVDQMHEDIISIWQSLKPCVKLITPLCVTLNCTDVKNATNSGTITNNNSIDSMNGEIKNCSENITTEIRDK
 KQVYALFYRPDVVPLNENSSYLLINCNTSTTQACPKVSFDPPIPHYCAPAGYAILKCNKIFNCTGPGCHNVSTVQCTHGIKE
 VVSTQILLNGSLAEELIIRSENLTNNAKIIVHLNESIEIVCTRPNNTRKSIIRIGQTVYATNDIIGDIRQAHGNISKTWN
 TTLEKVKELKEHFPKATTFQPHSGGDEVTTHSFNCRGEFFYCDITKLFNENLNTNTTLLPCRIKOIVNMWQGVGRAMY
 APPVEGNITCNSSITGLLLVRDGGNTSNSPEIFRPGGNMKNWRSELYKYKVEIKPLGVAPTRAKTLTVQARQILSGIVQQQ
 SNLRATFAQQHMLQITVWGIKQIQARVLAIERYLKDQQLGLNGCCSGKLICPTTVFNSSWSNKSQTDINDNMTWQMDREISN
 YTGFTYKLEESQOQERNEKDLALDSWKNLWSFDTITMLW*

*부분색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 가단히 표시한 것이다.

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

Nick name: 013

[illegible]

도면37a

Wild-type subtype CRF01_AE
97CNGX2F-AE (854 a.a.)
MRVKETQMNWPNLWKWGTLLGLVILCSASDNLWVTYYGYVPVWRDATTLCASDAKAHETEVNHWATHACVPTDPNPQEIHL
ENVTEFNMRNRMVEQMVEDVLSMDQSLKPCVKLLPCLCVTLNCTNANWTSNNFTNGPNKIGNITDEVKNCTFNMTTELKDKK
OKVHALFYKLDIVQINSSEYRLJNCNTSVIKOACPKISFDPIPIHYCTPAGYAILKCNDRKNFNGTGPKNVSSVQCTHGKPVVS
TQLLNGSLAEEIIIRSENLTNNAKTIIHLNKSVEINCRPSNNTRTSITMGPCQVFRYTGDIIGDIRKAYCEINGIKWNEVL
VQVTGKLEHENKTIIFQPPSGDDELIITHFSCRGEFFYCNITKLFNNTCIGNTSMEGCNNTIIPCKIKQIINMWQGVQAMV
APPISGRINCVSNIITGLLTRDGGADNNTNETFRPGGNIKDNWSELVKYKVVEIEPLGIAPTRAKRVVEREKRAVGIGAMI
FGELGAAGSTWGAASITLTVOARQLLSGIVQOQSNLLRAIEAQCHLLQLTVWGIQLQARVLAVERYLKQKFLGLWCSGKIIC
TTAVPWNSSWKNKSFEEINDMWIEWEREISNYTTSQIYEILTESQOQDRNEKDLLELDKWSLWNWFDITTNWLMYIKIPIIV
GSLIGLRITFAVLISIVNRVRQGYSPISFQPTHHQREPPDRFEIEIGEGGQSKDRSVRLVSGFTALAWDDLRLSLCLFSYHLRDF
ILIAARTVELLGHSSLSKGLRRGWGLKYLGNLLLYWQEIKISALSLLNATAIAYAGWTRVIEVAQRAWRALHLIPRRTRQGLE
RAIL

*밑줄친 아미노산은 140CF 디자인에서 결실되어야 하는 용할 도메인을 지칭하고, 빨간색으로 밑줄친 "W"는 C 말단에 있는 마지막 아미노산이고, "W"뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면37b

97CNGX2F-AE 140CF.pep (629 a.a.)
 Nick name: 018
 MRVKETQMNWPNLWKWGTILGLVILCSASDNLWVTVYGVFVWRDADPTLFCASDAKAHETEVENWYATHACVPTDPNPQEIHL
 ENVTFENFMRNRNWNVEQMEDVISLWDQSLKPCVKLTPLCVTLNCTNANWTNSNNTINGPNKIGNITDEVNCTFNMTTELKDKK
 QKVHALFYKLDIVQINSSEYRLINCNTSVIKOACPKISFDPIPIHYCTPAGYAILKCNDKNNGTGPCKNVSVQCTHGKIPVVS
 TQILINGSLAEEELIIRSENLNNAKTIIVHLKSVELNCTRPSNNTTSITMGPGQVEYRTGDIIGDIRKAYCEINGIKWNEVL
 VOYTGKLEHFNKTIIFQPPSGGDLLEIITHHFSRGEFFYCNCTKLFNNTCIGNTSMEGNCNTIILPCKIKQIINMWQGVGQAMY
 APPISGRINCVSNTITGILITRDGGADNNTNETFRPGGNIKNWRSELYKYKWEIEPLGIAPTRARTLTVOARQLLSGIVQQQ
 SNLLRAIEAQOHLLOLTWVGIKQLOARVIAVERYLKDQKFLGLWGCCKIICTTAVPWNSSWSNKSFEIWDNMTWIEWEREISN
 YTSQIYEILITSESQNDQDRNEKDLLELDKWSLWNV*

*부분색으로 표시된 아미노산은 결실된 응화 절대 위치의 경계를 간단히 표시한 것이다.

도면37c

CODON-OPTIMIZED 97CNGX2F-AE 140CF.seq (1921 nt.)

Nick name: 018

ttcagtcgacagccaccatgCGAGTAAAGAGACACAAATGGAATTGGCCCAATTTGTGGAAGTGGGAACATGTATCTCTGGCACT
 GGTGATAATCTGTAGTGCATCCGACAAATCTCTGGGTGACCGTTTACATGGTGTACCCAGTTTGGAGAGACGCTGTATACCCCTC
 TTCTGTGCAAGCAGCGCAAGGCCACGAAACTGAAGTCCATAATGTATGGCCACCCACGCTGCTATCCCAACCGGACCCCTAATC
 CCAAGAGATCCACCTTGAGATGTAAGTGAATTTTAACATGTGGAGAAATAACATGTTGGAACAAATGCAGGAGACGTTAT
 TTCTTTGGGACAGAGCCTTAAACCTTGTCAAATGACTCCCTGTGTGATCTCAATTTACAAAGCAATTTGGACC
 AACAGCAACACACTACCAACGGCCCTAACAAATTTGCAATATTAAGTGAAGTCAAGAACTGCATTTTAAACATGACACAG
 AACTGAAGGATAAGAAACAGAAAGTCCATGCTCTGTCTATAGCTCGACATAGTACAAATTAATAGCTCAGAAATAGACTGAT
 AACCTGCATACTTCCGTATCAACAGGCCCTGCCAAGATAGCTTCGATCCCTCATTTCACTACACACCGCGGT
 TACGCTATCTGAATGCAACGATAGAAATTTAACGACACAGTCCCTGCAAAAGCTTTCTCTCCAGCTGTACACACGGTA
 TCAGCCCTGTAGTATCAACAACTGCTCTGATGCTCTTGGCCGGAAGAGAGATCATCATTTAGAAAGTGAACCTTGAGAA
 CARGCCRAAGACTATATAGTGCACCTCAATAAATCTGTAGAAATCACTGACCCGACCCCTCAACACTCGAACTGAAGTGA
 ACAATGGCCCTGGCCCAAGTTTATTACGGGACCGCGACATATAGCGATATCAGAAAGGATATTCGAGATCAATGGCATCA
 AGTGGACGGAAGTACTGTTCTAGTAAGTGGAAACTCAAGACATTTTAATAGGCCATATATCCAGCCCGGAGTGGCGG
 CGACCTGAGATTATACCCATCACTTTTGTAGAGCGGAATTTTACTGTAACACCGCCAGCTCTTCAATAACAGTGC
 ATCGGGAACACTTTATGGAGAGGATTAATAATACCATTAATCTGCTGTAAGATCAAGCAGATTTATCAATGTGGCAGGAG
 TAGGTCAAGCAATGTACGACACCGGATTTACGACCGATCAATTCGGTATCAAAATCACCGGCAATCTGCTGACCCCGGACGG
 AGSCGACAGACAAATACCACTAACGAGACATTTAGACCTGGGCGGCAATTAAGGATTAATGGAGAAATGAGTGTATAAA
 TACAAGTCTGTAGAGATCGAACCCCTCGCATGCTCAACCGCGGCGGACTCTCACCGTACAGCTAGACAGCTGCTTTCTG
 GCATAGTCCAACAGACAGTCAACCTCCCTCGGCTATTGAAGCACACACACCTGCTCCAGCTGACTGTGTGGGAATCAACA
 ATTGAAGCAAGAGTCTCGCGGTGGACGCTATTTGAAGATCAGAAATTTCTTGGGCTTGGGCTGCAAGCGGCAAAATTAAT
 TGTACAACAGCGGTGCTTGGAACTCATCTGGAGTAAATAAGCTTTGAAGAACTGGGACAAATGACATGGATTGAGTGGG
 AGAGAGATTTCAAACTATACAGCCAAATTTACGAAATATGACAGAACTCAAAACCGACGACAGAAATGAGAAAGACCT
 GCTCGAAGTGGATAGTGGGCTCTTTGTGGAACTGgttaagatcttacaa

도면38a

Wild-type DRCEB-G (854a.a.)
 MRVGIQRNWOHLNWNWGIILGLVILICSAEKLWTVVYIGVPVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDFSPQEIINMR
 NVTENFNWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTEINNSTNITEYRNTNCSFNMTELRDKKAELYALFYR
 TDVYPIINEMNNENGTNSTWYRLTNCNVSTIKQACPVTTFEPIHYCAPAGFAILKCVDKKENGTCNNVSTVQCTHGIAPVV
 STOLLINGSIAEKDIIISSENI~~SDNAKVIIVHLNRSVEINCTRPNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVSWTKWNET~~
 LRDVOAKLQEFYFINKSIEFNSSGGDLEITTHSFNCGGEFFYCNSTGLFNNSILKSNISENNDTIILNCKIKQIVRMWQVRGQAM
 YAPPIAGNITCRSNITGLILTRGGDNNSTSEIFRPGGDMKNRSELYKYKTVKTSLSGIAPTRARRRVEREREKRAVGGAIF
 LGFLGTAGSTMGAASITLTQVRQLLSGIVQOQSNLIRAEAOQHLLQLTVWGIKQIPARVLADERYKDDQLLGIWGCSSGLIC
 TTNVPWNTSWNSKSYNEI~~WENMTWIEWEREIDNYTHIYSLTEQSQIQOEKNEQDLLALDQWASLWSFSISNWLWYIRIFWMIY~~
 GGLIGLRIVFAVLSIVNRVRCYSPLSFQTLJHHOREPDRPAGIEEGGCEODRDRSLVSGFLAIAWDDLRSLCLFSYHRLRDE
 ILIAARTVELLGRNSLKGRLGWEALKYIWNLLLYWARELKNSAINILDTIAI~~AVANWTDVIEVAQAGRAVLNIPRRIRQGLE~~
 RALL

*밑줄친 아미노산은 140CF 디자인에서 결실되어야 하는 융합 도메인을 지칭하고, 빨간색으로 밑줄친 'W'는 C 말단에 있는 마지막 아미노산이고, 'W' 뒤의 잔여 아미노산은 140CF 디자인에서 결실될 것이다.

도면38b

DRCBL-G 140CF.pep (630 a.a.)
 Nick name: 017
 MRVKGIORNWQHLLNNGILLGLVICSAEKLMVTYYGVVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPOEINMR
 NVTFENFMWKNMNVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFENMTTELDRKKKAEYALFYR
 TDVVPINEMNNENGTSTWYRLTNCNVSTIKQACPKVTFFPIPHYCAPAGFAILKCVDKKFGTGCNNVSTVQCTHGKIPVY
 STQLLNGSLAEKDIIISSENI SDNAKVIIVHLNRSEINCTRPNNTRRSVAIGPGQAFYTTGTVIGDIRKAHCNVSWTKWNET
 LRDVQAKLOEYFINKSIEFNSSGGDLEITTHSFNCGGEFFYCNSTGLFNNSILKSNISENNDTTLNCKIKQIVRMWQEVGOAM
 YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGDMKNWRSELYKYTKVIKISLGIAPTRARTLTVOVPQLLSGIVQOO
 SNLLRATEAQOHLQLTVWGIKQLRARVLALERYLKDQQLLGIWGCCKLICCTTNVPWNTSWSNKSYNEIWNMTWIEWEREIDN
 YTYHIYSLIEQSIOEKNEDLLALDQMASLWSW*

*푸른색으로 표시된 아미노산은 결실된 융합 절단 위치의 경계를 간단히 표시한 것이다.

도면38c

CODON-OPTIMIZED DRCL-G 140CF.seq (1921 nt.)

Nick name: 017

ttcagtcgacagccaccatgagagttaaaggatccaaaggcaattggcraacacacttggaaactggggcatattgattctttggact
 ggtgataattgtagcgctgaaanaactctgggttaactgtctatttacgggtgctgtgtggaggatgccaaagcccccctgtgtc
 tggcgaagtgtatgcaaaaggctcacagcaactgaatctcaacaacttggcccccacccacgctgtgtgccaaacccgacctagtcttc
 aggagatcaacatgagaaaggcttacggaaatttttaattatgtggaagaaataatggggagcaaatgcacgaacacataaatttc
 actctgggagagctcttgaaacuatgttgaaactttacccctgtggttcacccctgaactgtaccgaaatcaacraataaactca
 acggaataatcacagaaagataccgaatgactaactgttcccttaattatgacaacccgaactggagacaaaagaaggctgaat
 accgacttttctaccgaacagatgtgttaccaatcaacagatgaacaaatgaaacaaatggaaaggaaacttacctgtgtatgact
 gacaaactgtaacgttagacaatcaagcaggcttcccttaagtcaattcgaaactggaaacttgcataatgtatctacagtacatggacgcatg
 gaattacgctgtgtttcaaccagttctgctgaatggatcactcgcaaaaaggatatttattctcaagcgaaacacatac
 tgataatgcaaaaggctcatctgctccactcaacccgctcagttgaaataaactgcacactggccttaataaacacagacgctct
 gtcgcaatcgcccgagacaagctttttacactaccggggaagtattcgccgacatagcgaagcccactgccaaactgttagctgga
 ccaagtgaatgaaacactgcgcatgttcaagccaaacttcaagaaatcttcaataaacaattcaattgagttcaattctagctc
 tctatcctcaaaagtacatttctgaaataatgcacacatcacactgaattgcaagatcaagcagattgttaggatgtggcaac
 tggcggcgaactcgagattacactcactcctttactgcgcggaatttcttattgtataccctcggtctcttcaacaac
 gactcggacaagctatgtacgcccaccactcggcgaatataacgttcgaataatcactgcaattgcaagatcaagcagattgttaggatgtggcaac
 tggcggagacataatagcaccccgagatatttcagcagcgcgagcgatataaaaaaacactggaggtcagagctctacaag
 tacaaaacagatcaaaaataaaagcctggcattgtccactcgccgcccgcacactgacttccaaatccgacagctcctgtccg
 gratcgccaaacacagttccacttgcggcgctatagaggtcaaacacactccttcaactgactgtgtgggtgttccggcaaacactgac
 atgagacaaagatgtgggtggacgggtatcttaaggacacaaactcctggcatatgggggtgttccggcaaacactgac
 tgacacaaatgtaccctggacacacagctgctcaataaagtataatgagatgggaaacacatgacatggattgaatggg
 aaagggaattgacaaattatatacataccatatafactctctcagaaacattctcagatatacagagaaagaatgaacacagattt
 gtgggtcttgaccatggggtcttttgggagttgtaagatcttaca

도면39a

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

2003 Cons Env

MRVNGIORNCQHLMRWGILLFGMLIICSAENLWTVVYGVVWKEANTTLCASDAKAYDTEVHVNWATHACVPTDENPQEIIVLENTENF
 NAWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNAINTTNNNEEIKNCSEFNIITEIRDKKKVYALFYKLDVVPIDDDNNSYRLI
 NCNTSAITQACPVSFEPPIHYCAPAGFAILKCNDRKPNGTGPKNVSTVQCTHGKPKVSTQLLNGSLAEFIIRSENITNNAKTIIY
 QLNESVEINCTRPNNTRKSIRIGPQAFYATGDIIGDIQAHCHNISRTKWNKTLQOVAKKLREHFNKTIIFPSSGGDLEITTHSFNCGGE
 FFYCNTSELENSITWNGTNNITILPORIKQIINMWQGVQAWYAPPIEGKIRCTSNITGLLITRDGNNNTETIFPFGGDMRDNRSELYKYK
 VVKTEPLGVAFTRKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASTITVQARQLLSGIVQOOSNLRAIEAQOHLLQLTWYGIKQLQAR
 VLAVERYLKDQQLIGWCGSKLICTTNVFNSSWSNKSODEIWNMTWMEWDEINNYTDIIYSLIESQOQEKNEQELLALDKWASLWN
 WEDITNWLWYIKIFIMIVGSLIGRIVFAVLSTVNRVROGYSPISFQTLIPNRPDPDEGIEEGGQDRDRSIRLVNGFTLAWDDLRSI
 CLFSYHRLRLDILLIARTVELLIGRRGWEALKYLWNLLQYNGQELKNSAISLLDTALAVAEGTDRVIEVVQVCRAILNIPIRIROGFERAL
 L§

도면39b

2003 M. Group. and Env. seq. opt

ATGCGCCTGATGGGATACAGCGCAATCCCBAGCACTGTGCGGCTGGGGCAATCGATCTTGGCATCTTGGCATGCTGATGATCTGCTCGCGCGCGCG
 GAACTTGGTGGGTGATCGGGGTACTCGGGGTCCGGTGTGGAAGGAGGCAACCAACACCTGTTTGGGCTCTCGACGCCAAGGCTTACGACAT
 CCGAGTGAACAACTGTGGGCGACCGCGCTCGGCGCGGCAAGCGCCAAACGACGAGATCTGTGCTGGCAAGCTGCGGACGACGATGAC
 AACTCTGTGAGAGAAACAATGCGAGGAGATGACCGAGGACATCATCTCTGTGGGACGAGTCTCTGGAAGCCTCGCTGAAGCTGACCGC
 TGTGTGTGTGTGTAATCTGTAATCTGACGAGGTGAACCGACACATCCAACTGATGGGAGATCAAGAGACTGCTCTTCAACATCACTACCG
 CCGAGATCGCGGACAGACAGAGAGGTGTAGCCCTGTTTACCGCCCTGGACGTGTGCCTCAACGAGCTCTTACCGCTGTGATC
 AATCGAAGCATCTCGGCATCAACCAAGGCTCGCCGAGGCTCTCGAGCGCTCCATCACTCACTACCTGCGCCGCGGGGTGGCGGCAT
 CTTGAAGTCAACCAAGAGATCAAGGCAAGCGGCCCTTCGAAGAGTGTCCACGTGAGTGTACCTGCGCATCAAGGCATCAAGCGCTGTGT
 CACACCGACCTGCTGTGAAGGCTTCCCTTGGCCGAGGAGATCATCATCCGTCGAGACATCAACCGACCGAAGACATCATCGT
 CACCTGAACGCTGCTGTGGAGATCAATGCAATGCCACCGCCCAACAACAACCGGAATCCATCGCATCTGCGCCGCGCGGACGCTCTTACG
 CACCGGCGACATCATCTGGGCGACATCGGCGAGCGCACTGCAATCTCGGCGGAGTGGAGACAGACCTTCAGACAGGTGGCGCCACAGC
 TGGCGGAGCATCTTAACACACAGCAATCATCTCAAGCGCTCTCTCGGCGGAGCCGAGAGATCGAGAGATCCACCTCTCAATTCAGCGGAG
 GAGTCTTACTGAACACCTCGGCTGTTCATCTCACTGTGAAGGCGACCAACAGGACATCACTCCCTCTGCGCATCACTGACAGAT
 CGTGAATGTGGCAGCGGTGGCGAGGCAATGACGCCCGCCCATCGCGGCGCAATCACTCAAGATCCAAATCAACCGGCGGTGTC
 TGACCGCGACGCGGCGCATCAACACACCGCAATCTCGCCCGCGCGCGGACATGCGGCACTTGGGCTTCGAGCTTCAGATGT
 AAGGTGTGAATGAGCGCTGGGGGTGGCCCCCAACAGGCGCTGTGGAGTGGAGCGAGACAGCGCTGGGAGCTATGCGC
 CGTGTCTGGGCTTCTGGGCGCGCGGCTCCACTTGGGCGCGCTCATCTCACTGACCTGACCTGCGCGAGTGTGTCGCGCT
 TCTGCGACGAGATGCAACTCTCGCGGCTCAAGCGCGACGACCTTCAAGTCACTGAGCTGAGCGCCCAAGCGATGCTGACGAC
 CGGCTCTCGGCTGGGAGCGCTCACTGAAGACAGAGCTCTGGGCACTTGGGGTGTCTCGGCAAGCTGATCTGCACACCAAGTGC
 CTGAGACTCTCTGCTGTACACAGATCCGACAGACAGATCTGGAGACACATGACTGAAACAGATCTGGGCGGAGATCTTCCACTTACAGCG
 AACTCACTTCTCTGCTGGTGGAGGTCCGAGACAGATCTGGAGATCTAGATTTCAATCACTTGGCGGAGCTGAGCGGCTCTGTTGG
 AGTGTCTCACTCAACCGCTGGCGAGGATTAATCCCGCTGTCTTCAGACCTCTCACTCCCACTCGCGGCGCGACCGCTGTTGGTGG
 CACTCAGAGAGAGGGCGGAGGACCGGACCGTCCACTCGGCTGTGTCGGCTTCTGTGCCCCCTGTGGCTCTGGAGAGCACTCGCTGT
 CTGTGCGCTCTTCTCTCAACACCGCTGCGCGACCTCACTCTGATGCGCGCGCGAGCTGCTGGGCGCGCGGCTGGGAGGCGCT
 GAGATGACTGTGGAACCTGCTGATCTATGGGCGAGGATGAAGAACTCGGCACTCTGCTGGACACACCGCGAGCTGTGGCGT
 AGGGCAAGGACCGGTGATCGAGGTGTGTGACGCGGCTGGCGGCGCATCTCGCGCGCATTCGCGCAAGGCTTGGAGCGCGGC
 CTCTCTGTA

도면40a

2003 M. Group.Anc. Env
 MRVGIQRNCQHLLWRWGILIFGMLMICSAAENLWTVYGVVPVWKEANTLFCASDAKAYDTEVHNWATHACVPTDNPQEIIVLENTENE
 NMWKNMVEQMHEDIISLWDQSLAPCVKLTPLCVTLNCTDVNATNNTNNGE IKNCSEFNITTEIRDKKQKAYALFYRLDVVPINDNNSYRLJ
 NCNTSALTQACPKVSEFPIPIHYCAPAGFALLKCNDDKKNGTGCKNVSTVQCTHGIRPVVSTQLLNGLSLAEHEIIFKSENITDNAKTIIF
 QLNESVELNCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHNISGAENKTIQQVAAKLREHFNNKTIIFKPSGGDLLEITHSFNCGG
 EFFYCNTSGLFNSTWNGTNETITLPCRIKQIVNMWORGQAMYPPIAGNITCKSNITGILLTRDGGTNNTEFRPFGGDMRDNWRSELYKY
 KVKIEPLGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAGSTWGAASITLTVQARQLSGIVQQQSNLLRAEAQQHLLQLTVWGIKQLQA
 RVLAVERYLKDQQLGIWGCSGKLICTTNVPWSSWSNKSQDELDNMTWQWEREISNYTDIIVSLIEESQOQKNEQDIALDKWASLW
 NWFDITNMLWYIKIFIMIVGGGLIGLRIVFAVLSTVNRVQGYSPISFQULIPRPGDPDPGGIEEGGEGQDRORSIRLVSGFLALAWDDLRS
 LCLFSYHRLRDFILIAARTVELLGRRGWEALKYLWNLLQYWQQLKNSAISLLDITTAJAVAEGTDRVTEVVQACRAITHIPRRIRQCFERA
 LL\$

도면40b

2003 M. Group. and Env. seq. opt

[illegible]

도면41a

2003 CON A1 Env
 MRVMIQORNCQHLRGMGTMLIGSABENIMVTVYGVFVWKDETTLEFCASDAKAYETEMHNVWATHACVPTDINFQETHLENVTEEF
 NMMNNNVEQMHTDIISINQSLKPCVKLPPLCWLNCSSNWNVTNNTTHEEBIKKCSFNMTTELDKKQVYSLFYRLDVQINENNNS
 SYRLINCNTSALTQACEKVSFEPIPIHYCAPAGEFALLCKDKREFNGPCPNVSTVQCTHGKPKVSTQLLNGSLAEEDVIIRSENIITNNA
 KTLIVQLTKPVKINCTEPNNNTIRKSIRIGPQAFYATGDIIGIROAHCVSRSEMNKTLQKVAKOLRYKFNKTIITNNSGGDLIITHS
 FNGGFEFYCNTSGJFNSTWNGTMTITLPCRKOIINWDRAGOAMVAPPIQGVIRCESMTGLLPRDGGNNNTNTPFPGGCDMDN
 WRSELYKYVVKIEPLGVAPTRAKRVREREKRAVGIGAVELJELGAGSTMGRASTILIVQAPQLSGIVQQOSENLEATPAQOHLKLTIV
 WGLKQLQARVLAVERYLKDOQLLGWCSCKLICITINVPWNSSWKSQNEIMDNMTWLOMDKEISNYTHLIYNLIEESONQOEKNQDLLA
 LDKWANNNWEDISNMLWTIKIFIMVGGTIGURIVFAVLUSVINRVQGYSPUSQTHTNPRGLDPRGEIEEGGQGRDRSIRLVSGFLA
 LAWDDLRLCLCFSTHRLRDFLLIAARTVELLGHSSKGLRIGWEGLYLWNLIIYWGRELKISAINIIVDTIATAVAGMIDRVIEIGORIGRA
 ILHIPRRIRQGLERALLS

도면41b

[illegible]

도면42a

2003 Al. Anc Env
 MRVGIQRNCOHLRWGTMIFGMIICSAARENIMWTVVYGVVVKDAETTLKCSDAKAYDIEVHNWATHACVPTDPNPQELDENVTSEF
 NMKNNMVEOMHADIISLWDQSLKPCVKLTPELCVTINGSNVWVNNNTNTHHEEIKKCSFMMTELRDKKQKVYSLEFYRLDVPVPIFNFNNS
 SYRLINCNTSAITQACEKVSFEPLFIHYCAPAGAILKCKDKEENGTPCKRVSTVQCTHGKIKFVSTQLLNGSLAEFEVMIRSENITENA
 KTIILVQLTEPVKINCTRPNNTRKSIIRIGFSQAFATGDIIGTIRQAHCVSRTEWNTKQKVAQLRKHNNKTIIFNSSGGDLFIITHS
 FNCGEFFYCNISGLFNSITWNGTMDTITIPCRIRKQIINMQRVQAMVAPPITQGVIRCESNIGLLTRDGNNTNFTFAPGGGMDRDN
 WRSELYKYKVKIEPLGVAPTRAKRRVVEREKRAVGLGAVELGELGAGSTMGRASTILVQARQLLSGTVQQSNLIRAEQQHLLKITY
 WGIKOLQARVLAVERYLKDOQLLGLWGCSKGLICTTNVWNSNKSQDEIWDNNTWLOWDEKLSNYTDIILNLEESONQQKKEQDLIA
 LDKWANLWNNEDISNWLRYIKIFIMIVGGLIGLRIVFAYLSVINRVRCQTSPLSFQTLTENPEGPDPRGRIEEEGGQGRDSIRLVSGFLA
 LWDLRSGLCLFSYTHRLRDFILIAARTVEILGRSSLRGLFHWGLKYLWLLIYWGRELKISAINLLOTIAIYAVAGWTDRTVTEIGQRICRA
 ILNIPRIIROGLFSAALLS

도면42b

2003 A1.anc Env. seq. opt

[illegible]

도면43a

2003 CON A2 Env

MRVMTORNYCHLWENWGIILGMLIMCKATDLWTVYGVVWKDADDTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVNLENVTEDFN
 MWKNNVQMHEDIISLWDSLKPCKVLTPLCVTLNCSNANTTNNSTMEIKNCSYNITTELRDKTKVYSLFYKLDVVQLDESNKSEYYR
 LINCNTSAITQACPKVSEPIPIHYCAPAGFAILLCKKDBRPFNGTSCNNVSSVQCTHGIKPVASTQLLNGSLAEGKVMIRSENITNNAKNI
 IVQFNKPVPLTCIRPNNTNRKSIREFGQAFYTTNDIIGDIRQAHCNINKTKWNATLQVAEQLREHFPNKTIIFTNSSGGDLFTTTHSFNCG
 GEFFYCNTTGLFNSWNGTNTTEQMITLPCRKQIINMWQVGRAMYAPPIAGVIKCTSNITGIIILTRDGGNNETETTERPGGDMRDNR
 SELVKYKVKIEPLGVAPTRAKRRVVEREKRAVGMGAVELGFLGAGSTMGAASITLTVQAROLLSGIVQQSNLLIKATEAQOHLKLTVWG
 LKQLOARVLALERYLDQQLLGGWCSGKLICATVPMNSSWSNKTQETIWNNTWLOWDEISNYTNIYKLEESQOQKEQDLIALD
 KWANLWNNFNITNMLWYIRIFIMIVGGGLGLRIVIAIISVNNRVQGYSPLSFOIPTNPEGLDRPGRIEEGGEGQGRDSIRLVSGFIALA
 WDDLRSCLFSSYHRLRDCILIAARTVELLGHSSLSKGLRLGWEGCLKYLNLLLYWGREIKNSAISLLDTIAVAVAEWTRDRIETIGQACRAIL
 NIPRRIRQGFERALL\$

도면43b

2003 CON A2 Env.seq.opt

[illegible]

도면44a

2003 CON B Env

MRVKGIRKNYQHLWRGCTMLIGMLMICSAAEKLWVTYYGYVPVWKEATTTLFCASDAKAYDTEVHNWVATHACVPTDNPQEVVLENTENF
 NMWKNMVEQMHEDIIISLDQSLKPCVKLTPLCVTLNCTDLMNATNTTIIYRWGEIKNCSEFNTTISRDKVQKEYALFYKLDVVRIDND
 NTSYRLISCNSTSVITQACPVSFEPIPIHYCAPAGFAILLKNDKKNFTGTCTNVTSTVQCTHGIPIPVVSTQLLNGSLAEVEVIRSEFTD
 NAKTIIVQNLNESVEINCTRPNNNTRKSIHIGPGRFYTTGEIIGDIRQAHCNISRAKWNNTLKQIVKKLREQFGNKTIIVFNQSSGGDPPIVM
 HSFNCGGEFFYCNTTQLFNSTWNGTWNNTTEGNIITLPCRIKQIINMWQEVGKAMYPPIRGOIRGSSNITGLLTRDGGNNETEIRPGGGDM
 RDNWRSELYKYKVKTEPLGVAPTAKRRRVVOREKRAVGIGAMFLGAGSTMGASMTLTVQARQLLSGIVQQQNNLLRALEAQHLLQ
 LTVWGLQLOARVLAVERYLKDQQLGIGWCSGKLICTTAVFWNASWSNKSLELMDNMTWMEWEIIDNYTSLIYTLLEESQOQKEQE
 LLELDKASLWNVFDTINWLMYIKIFIMVGGSLVGLRIVEAVLSIVNRVQGYSLSFQRLPAPRGPDPKPEGTEEGGERDRDRSGRLVDG
 FLALIWDOLRLSLCLFSYHRLRDLIIIVTRIVELLGRGWEVLKYWNWLLQYWSQELKNSAVSLINATAIATAVAECTDRVIEVVQACRAILLHI
 PRIRQGLERALL\$

2003 CON B Env. seq. opt

ATGCGCGTGGAAGGGATCCGAAAGAACTACAGACACCTGTGGCGTGGGCGACCATGCTGTCTGGGCATGCTGATGATCTGCTCGCGCCGCGA
GAAGCTGTGGGTGACCTGTACTACGGGTCCGGTGTGGAAGAGGCCACCAACCCCTGTTCTCGCTCCGACGCCAAGGCTGACACCA
CCGAGGTGACCAACTGTGGCGACCAACCGCTCGGTGGCGACCGACCGACCCACGAGAGTGTGCTGGAGCAAGTGAACGAGAACTTC
AACAATGGAAGAAACAAGTGTGGAGCAGATGCACGAGGACAATCATCTCTCTGGGACAGTCTTGAAGCCCTGCTGTGAGCTGACCCC
TCTGTGGTGAACCTGAATCCAGCACTGATGAGCGACCAACCAACCAACCATCATCTACGCTGGCGGCGAGATCAAGAATCT
GCTCTCTCAACATCAACACTCCATCCGACGAGAGTGCAGAGAGTACGCCCTGTTCTACAGCTGGAAGCTGGTGCCTCCACACAGAC
AACAACCTCTACCGCTGATCTCTGACAACTCCGTGATCAACCAGGCTGCCACGAGCTCTTGGAGCCCTCCATCCATCCACACTGATG
CGCCCGCGCGGTTCCGATCTGTGAGTGCAGAGCAAGAAGTTCAACGGACACGGCCCTGCACCAACTGTCCAGCTGCAATTGCACCC
AGGCATCCGCGCGTGTGTCTCCACCTCCAGCTGCTGTGAGAGTCACTGGCCAGGAGGAGTGGTGTATCCGCTCGGAGCAACTCCAGGAC
AAGGCAAGACCATCATCTGTGCATGAGAGAGTCTGTGAGATTAATCTACCCGCCGCCCAACCAACCGCAAGTCCATCCATCGG
CCCGCGCGCTCTTACACACCGCGAGATCATCGCGACATCCGACGAGCCATGCAACATCTCCGCGCACTGGAAACAACAACC
TGAAGCAGATCTTGAAGAGGCTTGGCACAACAACCATCTGTCAACCATCTCTCGCGCGGCGACCCCGAGATCTGTGATG
CACTCTCTCAACTCGGCGCGAGTCTTCTACTGCAACCAACCCAGCTGTCAACTACCTGGAACGGCATCTGGAACTCGTAGG
TCAACATACCTCCCTCGCGCATCAAGCAGATCATCAACATGTGGAGGAGTGGGAAGGCCATGTACGCCCCCCCATCCGCGCCAGA
TCCGCTGCTCTCCATCAATCAACCGSCTTCTGTGAACCGCGGCGCAACAACAGAACCGAGATCTTCGCGCGCGCGGCGGACATG
CGGCAACTGGCCTCGAGCTGTACAGATCAAGTGGTGAAGTCAAGCTCAAGCTCGAGCCCTGGCGCGCCCAACAGCGCAAGCGCGGTGT
TGACGCGAGAGAGCGCGCTGGCATCTGGCGCATGTCTGGCTTCTTGGCGCGCGCGCTCCACTATGGCGCGCTTCATGACCC
TGACCGTGCAGGCCCGCHGCTGTCTGGGCATCTGTGACGACGACAACAACCTGTGCGGCCATCGAGGCCACGACCATCTGTCCAG
CTGACCGTGTGGGCATCAAGAGCTGACGCGCGCTGTGGCGGTACCTGTGGAGCGTACCTGTGGAGCGAGAGTCTGGAGCTTGGGCT
CTCGGGAAGCTGATCTGCACCAACCGCTGCCCTGGAAACGCTCTGTGTCCAAACATGCTTCGCAAGATCTGGGAACAACATGACCTGGA
TGAATGGAGCGCGAGATGCACAATACACTCCCTGATCTACACCTGTCAAGGATCCAGAACCCAGAGAGGAAGAACAGCGAGGAG
CTGTGGAGCTGAGCAATGTGGGCTCCCTGTGAAGTCTGATCTCAACATCAACTGGTGTGTACATAAGATCTTCACTGATGACTGGG
CGGCTGTGTGGCTCGGCATCGTGTGCTGTCCCTGTCTCATCTGAACCGGCTGGCGAGGCTACTCCCGCTCTCTCCAGACCGGCC
TTCGCGCCCGCGCGCGGACCGCGCCGAGGAGCATCGAGGAGGAGGCGGCGAGCGACCGCTCGGCCCTCTGGTGGACGCG
TTCCTGGCCTGATCTGGAGCAACTGCGCTCTCTGTGCTTCTACCAACCGCTCGGCACTGTCTGTGATCGTAGCGCGCTCGT
GGAGCTGTGCGCGCGCGCGCTGGAGGTGCTGAAGTACTGTGAAGTACTGTTGGAACCTGCTGCCAGGAGCTGAAGAACTCCGCGCTGT
CCCTGTGAACCGCACCGCATGCCGTGGCGGAGGAGGACCAACCGGTGATCGAGGTGTGCAAGCGCGCTTCGCGGCATCTGTGCACATC
CCCGCGCGCATCCGCAAGGCGCTTGGAGCGCGCTGCTTAA

도면45a

2003 B.anc Env

MRVKGIRKNCQHLLRWGTMILGMLMICSAAENLWVTYVYGVVWKEATTILFCASDAKAYETEVHNVWATHACVPTDENPQEVVLENTENF
 NEMKNNVVEQMHEDIISLWDQSLKPCVRLTFLCVTLNCTDLNATNTNSTMYRWGEIKNCSFNITTSIRDKMQKEYALFYKLDVVIDNN
 TSYRLINCNTSVITQACPKVSFEPIPIHYCTPAGFALLKCNCKKENGTCCKNVSTVQCTHGIRPVVSTQLLNGSLAEVEEVIRSENFTDN
 AKTILVQLNESVEINCTRENNTNRKSIHIGPGRAFYATGEIIGDIRQAHCNLSRAKWNNTLKQVVTKLREQFNKTIVENPSSGGDPPIVMH
 SFNCGGEFFYCNTTQLFNSTWNGTWNTEGNTILPCRIRKQIINMWQEVGKAMYAPPPIRGOIRCSNITGLLITRDGGNNETEIERPGGGDMR
 DNRSELYKYKVVKIEPLGVAPTKAKRVVQREKRAVGIGAMELGFELGAAGSTWGAASMTLIVQARQLLSGIVQQNNLLRAEAQQHLLQL
 TVWGIKQLOARVLAVERYDRDQLLGWTGCSGKLCITFTVPWNASWSKSLDEIWNNTWMEWEREIDNYTGLIYTLIEESONQOEKNEOEL
 LELDKWASLWNNFDTITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVROGYSPLSFOTRLPAPRGPDRPEGIEEGGERDRDRSGRLVNGF
 LALIWDRLRSICLESYHRLRLDLLIVARIVVELLGRRGWEALKYWNLLQYNSQELKNSAVSLNATAVAEGTDRVIEVVQACRAILHIP
 RRIRQGLERALL\$

2003 B.and Env. seq.opt

ATGCGCGTGAAGGGCATCCGCAAGAACTGCCAGGACCACTGTGGCGCTGGGCACCAATGCTGTCTGGGCATGCTGATGATGTGTCCTCGCGCGCGGA
GCAACTGTGGGTGACCGGTGTACTACGGGTGCGCGTGTGAAGAGGCGCACCACTACCACTGTTCTGGCGCTGGGCACCGCAAGCGCTACGAGA
CCAGTGTGCAACCTGTGGCGCTCCACCGCTGCCCGACCGACCCACCGAGAGTGGTGTCTGGAGAACTGAGTCAAGCGAGAACTTC
AACATGTGAAGAAACAACAATGTGTGAGAGATGACAGAGACATCATCTCCCTGTGGACAGTCCCTTGAAGCCCTTGGTGAAGCTGATCCCT
GCTGTGTGCTGAGCTGTGAACCTGTGAAGCTGTGAAGCGTACCAACCACTCCACCAACTATACCGCTGGCGTGGCGAGATCAAGAATCT
GCTCTTCACATCACCACCTCCATCCGCGACAGATGCGAAGAGTACGCCCTGTTCTACAAGCTGGAAGTGTGGCTGCCCATCGACACAAC
ACCTCTACCGCGCTGATCACTGCAACACTCCGTGTATCACCAAGGCCCTGCCCCAAGTGTCTTGAAGCCCACTCCCATCCACTACTGCAAC
CCCCCGCGGTTCGCGATCTGTGAATGACAGCAAGAAGATCTAACCGCAACCGGCCCTCGTGAAGAACTGTCCAACCTGTCAGTGCACCGACG
GCATCCGCGGTGTCCACACAGCTGTGCAAGCGCTCCTGGCCGAGGAGGTGTTGATCCGCTCGGAGAACTTCACCGACCAAC
GCCAAGACCATCATCTGTGCAGCTGAACGAGTCCGTGGAGATCACTGCACCCGCCCCCAACAACAACACCCGCAAGTTCATCCATCGCTGGCC
CGCGCGCGCTTACAGCAACCGCGGAGATCTGCGGCATCCGAGCCCACTGCAACTCTCCGCGCAAGTGAAGCAACAACCCCTGA
AGAGGTGTGAACAAGTTCGCGAGCGAGTTCGACACACAGACCATCGTGTTCACCCCTCTCCGCGGGCGACCCGAGATCTGTGATGCCA
TCCCTTCAACTGCGGGGGCGAGTTCCTTCTACTGCAACACCAACCACTGTTTCAACTCCACTGGAAGGCGCACTGGAAACAACACCGCGGACAA
CATCCCTCTCCGCGCATCAAGAGATCATCAACTATGTCAGAGGTGGACGCCATGTACGCCCCCTCCATCCGCGGCCAGATCC
GTCTCTCTCAACATCACCGGCTGCTACTGACCCCGCACGCTGGCAACAAGAGCGAGATCTTCGCGCGCGGGCGGCGACATCGCG
GADAACTGGCGCTCCGAGCTGACAAATGAAGTGTGAAGATCGAGCCCTTGGGCTGGCCCTCCACCAAGCGCAAGCGCTGGTGGA
CGCGAGAAACCGCGCTGGCGCATCTGCGGCCATGTCTCTGGGTCCTGGCGCGCGCGCTCCACCATGGCGCGGCTCCATGACCCCTGA
CCGTGACCGCCCGCAGTGTCTGTCTCCGCGATCTGTGCACGACGAACAACCTCTCTGGCGGCCATCGAGCGACGACCTGTCTGACGCTG
ACCTGTGGGGCATCAAGCAGCTCGAGCCCGGCTGTGGCGGTGGAGCGTAOCTGGCGCGACAGAGCTCTGGGCATCTGGGGCTGCTC
CGCAAGCTGTATCTACCAACCGCTCGGCTGGAAGCGCTCTGTGTCCAAAGATCCTCTGGACAGAGTCTGGAACAACAATCACTCATGATGG
AGTGGAGCGCGAGATACACAACTACACCGGCTGATCTACACCTGTAGTAGAGTCTCCAGAACCAACAGAGAAAGACGACAGAGCGTG
CTGGAGCTGGAAAGTGGGCTCCCTGTGGAAATGTTTCAGATACCAACTGGCTGTGTATCATCAGATCTTCATCATCATCTGTTGAGCGTG
CTGGTGGGCTCGGCATCTGTTCGGGTGTTCATGTGGAACCGGTGGCGAGCGCTACTCCCCCTGTCTCTCCAGACCCCGCTTCG
CGCGCCCGCGCGCCGACCGCCCGAGGGCATGAGAGAGGCGGCGAGCGGACCGCGACCGCTCCGCGCGCTGGTGAACCGGCTTC
CTGGCTGTATCTGGAAGCACTGCGTCTCTGTGCGCTGTCTCTACACCGCTGCGGAGCTGTGTCTCATCTGTCGTCGCGCATCGCTGA
TGTCTGGCGCGCGCGCTGGAGGCGCTGAATACCTGTTGGAACTTCTGTGCACTACTGGTCCCGAGAGCTGAAGAACTCCGCGCGTGTCCG
GTGTAAACGCAACCGCATCGCGTGGCGGAGGACGACCGCGGTGATCTGAGGTGTGACGCGCGCTGCCCGCGCCATCTCTGCACATCCCG
CGCGCATCTCGCCGCGCGCTGGAGCTGCTGTA

도면46a

2003 CON C Env
 MRVRGILLRNCQOWRLINGILGFWMMLCNVVGNLWTVYIGVPVWKEAKTTLFCASDAKAYEKEVHNWATHACVPTDPNPQEIIVLENTENF
 NMWKNMVDQMHEDIISLWDQSLAPCVKLTPLCVTLNCTNATNTMTGEIKNGSFNITTELRDKKOKVYALFYRLDIVPLNENNSYRLINC
 NTSALTQACPKVSFDPIPIHYCAPAGYAILKCNNKTENGTPCPCNNVSTVQCTHGIKPVVYSTOLLNGSLAEELIIRSENLTNNAKTIIVHL
 NESVEIVCTRENNTKRSIRIGPGOTFYATGDIIGDIRAHNCNISEDKWNKTLOKVSKKLKEHFNKTIKFEPSGGDIEITTHSFNCRGEF
 FYCNTSKLENSTYNTSTIITLPCRKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLITRDGCKNNTETFRPGGDMRDNRSELYKYKV
 VEIKPLGIAPTAKRRVVEREKRAVGIGAVFLAGLGAAGSTMGAASITLTVQARQLLSGLVQOOSNLLRAIEAQOHMLQLTVMGKIQLOTRV
 LAIBRYLKDOQLLGIWGCCKLICTTAVENWSSSNKSKQEDIMWMTWQMDREISNYTDTIYKLLLEDSONQOEKNEKDLIALPSWKNLWNW
 FDI TNMLWYIKIFTMIVGGLIGLRITFAVLISIVNRVROGYSPLSFOTLTNPRGPDRLGRIEEGEGEQDRDRSIRLVSGFLAWDDLRSLC
 LFSYHRLRDFILIAARAVELLGRSSLRGIGRGWEALKYLGSLVQYWGLEIKKSALSLDDTIALAVAEIGTDRIIELLIQICRAIRNIPRIRQ
 GFEAALQ\$

2003 CON_C Env. seq. opt

ATGCGCGTGTGGCGCATCTCTGCGCAACTGCCAGAGATGTGGATCTTGGGCAATCTTGGCTTCTTGGATGCTATCATCTGCAACGTGGTGGG
CAACCTGTGGTGAACCGTTTACGGGTGCCCTGTGAAGAGGCGGAAGCAACCTCTGTTTGGCTGGCGGCCCAAGGCGTACAGGA
AGAGAGTGTGCAACAGTGTGGCCACACGCTGCGTGGCCACCCACCCACAGGAGATCGTCTGGAGCACTGACCGAGAACTTC
AACAATGGGAAGAACGCAATGTTGGACCAAGATGACAGGAGCAATCATCTCTGTGGACCAAGTCCCTGAAGCCCTGGTGAAGCTGACCCG
TCTTGCTGTGACCTGTGAATGTCAACAGCCCAACCCCAACCAACATCCGCGGAGATCAAGAATGCTCTCTCAACATACACCGAGG
TCGCGCAAGAAGACGAGAGGTGACGCCCTGTCTACCGCTGGACAATCGTGCCCTGAACAGAGAACACTCTTACCGCTGATCAACTGC
AACACCTTCGCGCATCACCCAGGCTTGCCCAAGTGTCTTCGACCCCATCCCATCTCACTGTGGCCCGCGCGGTACGCCATCTGAA
GTGTGCACAAACAAAGCTTACAGCGACCGGCGCTGCATCAACCTGTGCAGTGCAGTGCACCAACCGCATCAAGCCCTGTGTGTCCAC
AGTGTGCTGTCAAGGCTTCCCTGCGCAGAGGAGATCATCATCGCTTCGAGAACCTGACCAACGCCAAGAACCATCATCTGTGACATG
AACAGAGTCCGTGGAGATGTGTGCACCCCGCCCAACAAACACCCGGAATCCATCCGCAATCGGCCCGCGGACAGACCTTCTACGCCACCGG
GACGACTATCTCGGCGACATCTCGGACGCGCATCGCAACTCTCGAGAGCAAGTGAACAGCACTCGAAGAGGTGTCCAAAGAGTGTGAAG
AGCAATCTCCCAACAAGACCATCAAGTTGAGCGCTCTCTCGCGCGCACTGTGAGATCAACACCCATCTCTCAACTCCGCGGCGAGTCT
TTTCTACTGAACACTCCAAGCTGTCAACTCCACTCAACTCCACAACTCCACCACTCACCTTGCCCTGCGCATCAAGAGATCATCAA
CATGTGGCAGAGTGGCGCGCATGTACGCGCCCGCCCATCGCGGCAACATCACTCGAATCCAACTCAACCTGCTGTCTGACCGC
GCGACGGCGGCAAGAAACACACAGAGACTTTCGCGCCGCGCGCGGCGACATCGCGGCACTTGGCGTCCGAGTGTACAATCAAGATG
GTGAGATCAAGCCCTCGGCAATCGCCCAACAAAGCAAGCGCGGTGTGGAGCGAGAAAGCGCGCTGGGCACTGGCGCGCTGT
CTCTGGGCTTCTGGCGCGCGGCTCCACATTTGGCGCGCTCTCATCACTGACCGTGTGAGGCCGCGCAGTGTCTCGGCACTGTGCG
ACGACGACTCCAACTGCTGTGCGGCCATCGAGGCCAGACAGATGTGCTGACGTGACCTGTGGGGAATCAAGCAGCTGCAGACCGGCTG
CTGGCAATGAGCGTACTCTGAAGAGCACTGTGGGCAATGCTGGCGTGTCCGCAAGTGTATGCAACCGTGTGACCGCTGGAA
CTCTCTCTGTGCCAAGAGTCCCAAGAGGACATCTGGACAACATGACTGGATCACTGTGGACCGCGAGATCTCAACTACACCGACACCA
TTTACCGCTGTTGGAGACTTCCAGACCAACAGAGAGAGGAGCTGCTTGCCCTGGACTCTTGGAAAGCACTTGTGGAACTGG
TTTGACATCAACAACTGGCTGTGATCAAGATCTTCACTGATGTGGCGGCTGATCGGCTGTGATCATCTCTGCGCGTGTGTC
CATCTGTGAACCGGTGCGCAGGCTACTTCCGCGCTGTCTTCAGACCTTGACCCCAACCCCGCGCGCCAGCGCTGAGCGGATCG
AGAGAGAGGCGCGGAGACGAGACCGGACCTCCATTCGCTGTGTCTCTGGCTTGGCTGGGAGACCTTGGCTCCCTGTGTC
CTTCTCTTACCAACCGCTGCGGAGTTCATCTGATTCGCGCGCGCTGTGGCTGTGGCTGTGCTCCCTGCGGCTGAGCG
CGGCTGGAGGCGCTGAACTACTCTGGCTGCGAGTACTGGGCTGTGAGCTGAAGAGTCCGCCATCTCTCCCTGTGACACACCATCG
CATTCGCGTGGCCGAGCACCGACCGCATCATTCGAGCTGATCCAGCGCATCTGCGCGCGCATCCGCAACATCTCCCGCGCATCGCCGAG
GCTTCCAGGCGCGCTTCAGTAA

도면47a

2003 C. anc Env

MRVGLLRNCQWNIWLGILGFWMMLMICNVVGNLWVTYYIGVPVWKEATTLFCASDAKAYEREVNVNATHACVPTDPNPFQEMVLENTENF
 NMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEMKNGSFNITTELRDKKOKYVALFYRLDIVPLNDNNSYELINC
 NTSATQACPKVSEFPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGKIKFVVSTQLLNGSLAEELIRSENITDNAKTIIVHL
 NESVEIVCTRFNNNTRKSIIRIGPGQTFYATGDIIGDIRQAHCNISEEKWNKTLORVGEKLEKHEFPNKTIKFASSGGDEITTHSFNGRGEF
 FYCNTSRLFNSTYNSKNSTJTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGKNNNTETFRGGGDMRDNRSELKYKYV
 VEIKPLGIAPTEAKRRVVEREKRAVGIGAVFLGFLGAGSTMGAASTITLVQARQLLSGIVQQSNLLRAIEAQOQMLQLTVMGICKQLQTRV
 LAIERYLKDOQLLGITWGCCKLICITAVPWNSSWNKSKQSEIWDNNTWQMDREISNYTDIIVRLLEDSONQOEKNEQOLLIADSWENLWNV
 FDI TNWLYIKIFIMIVGGGLGRLIFAVLSIVNRVROGYSPLSFQTLTENPRGPDRLGRIEEEGGEGQDRDRSIRLVSGFLAWDDLSLC
 LPSYHRLRDFILJARAFAVELLGRSSLRGQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTERIIELIQRICRAIRNIPRRIRQ
 GFEAAII\$

도면47b

2003 C. and Env. seq. opt

ATGCGCGTGATGGCATCTGGCAACTGCCACCACTGGTGATCTGGGCAATCTGGGATCTTGATGATCTGCAAGCTGGTGGG
CAACTCTGGGTGACCGTGTACTACGGCTGCCGTGTGGAGAGGCAAGCAACCTCTTCTGGCGCTGACCGCAAGCGCTACGAGC
CGGAGTGCACCAAGCTGTGGCGATCCACCGCTCGGTCCGACCGACCAACCCCGAGAGATGGTGTGGAAACCTGCAGACGAACTTC
AACATGTGAAAGACGACATGGTGGACCAAGATGCACGAGACATCATCTCCCTGGGACCATCTCTCTGTAAGGCGCTGGTGAAGCTGACCC
CTGTGGTGTACCTGAACTGCACACCGCACCAAGCCCAACCATGGCGAGATGAAGAACTCTCTCTCAACATCAACCCAGC
TGGCGCAAGAAGCAGAGGTGTACGCCCTGTCTACCGCTGGACATCTGGCCCTGAACGACAACACTCTCTACCGCTCATCACTGC
AAACCTCGCCATCAACGAGGCTGCCAGTGTCTTCGAGCCCATCCCATCACTGTGGCCCGCCGCGGACAGCAATCATCTGTGCACCTG
GTGCAACACACAGGCTCAACCGACCGGCTCGCAACAGCTGTCCACCTGCGAGTCGACCAAGCATCAAGCCGATCGGTGTGCCACCTG
AGCTGCTCTGAAGGCTCCCTGGCCGAGGAGGATCACTCCCTCGAGAACCTGACCCACAAGCCCAAGCAATCATCTGTGCACCTG
AAAGATCTGTGGAGATCTGTGCACCGCCCAACACACACCGCAAGTCACTCCGCTCGGCGGACAGCACTCTCTAGCACTCGG
CGACATCATCGGCAACATCCGCAAGCCCATCGCAACATCTCGAGGAAGATGGAAAGAACCTCGACGCGGTGGCGAGAGCTGAAG
AGCACTTCCCAACAAGACCATCAAGTTCCGCCCTCTTCGCGGGCGGACTGGAGATCAACACCATCTCTCAATCTCCGCGGCGAGTTC
TTTATCGCAACACTTCGCCCTGTCTCAACTCCACTCAATCACTCAACATACCTGCCCTGCCGCAACAGAGATCATCAA
CATGTGGCAGGGGTGGCGCGCATATGACCGCCCGCCATCGCCGCAACATCACTCTCAAGTCCAACTACACCGGCTGCTGCTGACCC
CGACCGCGCGCAAGACAACACCGAGACTTCGCGCGGCGGCGGACATGGCGCACTCTGGCGCTCCGAGTATACAGTACAGGTG
GTGGAGATCAAGCCCTGGGCATTCGCCACGAGGACAGGCGGCTGTGGAGCGGAGAGAGCGCTGGGATCGGACCGCGCTGT
CTGGGGTCTCTGGGCGCGCGGACTCCACATGGCGCGCTCTCATCCCTGACCTGACCGTGCAGGCCCGCCAGTGTCTCGGCATCGTGC
ACGAGCAGTCTCAACTGCTCGCGCCATCGAGCCCAAGACATGCTGTGAGTACGCTGTGGGCGCATCAGCAGCTGCACACCGCGTGA
CTGGCATCTGAGCGCTACTCGAGAGCAACAGCTGTGGGCTCTGGGCGTCTGGGCGTCTGTCACACCGCGCGCTCCCTGGAA
CTCTCTCTGTCCAAACACTCCAGGAGGAGATCTGGGACACATGACTGGATGGGACCGCGAGATCTCCAACTACACCGACACA
TTACCGCTCTGTGGAGACTCCAGAACACGAGAGAGAACAGCAGGAGACTGCTGGCGCTGGACTCTGGGAGAACTGTGGAATCG
TTGCACATCCCAACTGGTGTGTGTACTAAGATCTTATCATGATCTGGGCGGCTGATCGGCGCTGGCATCATCTCGCGCTGTCTG
CATGTGAAACCGCTGCGCAGGCTACTCCCGCTGTCTTCAGACCTGACCCCAACCCCGCGGCCCGACCCCTGGGCGCATCG
AGAGAGAGGCGCGGAGCAGGACCGGACCTCATCGCGCTGGTGTCCGGCTTCCTGGCCCTGGCTGGGACGACTCGGCTCCGTGTC
CTGTCTCTACACCGCTCGGCGACTCATCTGATCGGCTGGGCTGGAGCTCTGGCGGCTCTGGCGGCTCGGCGCTCGGCGG
CGGCTGGAGGCGCTGAAGTACCTGGGCTCTGGTGTGAGTACTGGGCGCTGAGCTGAAGAGTCCGCCATCTCCCTGTGTGACACCATCG
CAATCGCGCTGGCGAGGCGACCGGATCATCAGCTGTGAGTCTCAGCGACTTCCGCGGCACTCCGCAACATCTCCCGCGCGCATCCGCGAG
GGCTCTGAGGCGCGCTGTGA

도면48a

2003 CON D Env

MEVRG:QNRNYQHWRWGIMLIGMLMICSVAENLWTVVYGVVWKEATTTLFCASDAKSYKTEAHNIWATHACVPTDPNPQETELENTENE
 ANWKNNWVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVKRNTSNDTNEGEMKNCSENITTEIROKKQKQVHALFYKLDVVPIDDNNSNT
 SYRLINCNTSALTQACPVTPEPIPIHYCAPAGFALCKCKKNGTGPKNVSTVQCTHGIRPVVSTOLLNGSLAREEIIIRSENLTNNA
 KLIIVQINESVTINCTRPYNNTRQRTPIGPGQALYTRIKGDIRQAHCNISRAENKTLQQVAKKLGOLLNKTTIIFKSSGGDEITTHSF
 NCGGEFTYCNTRLENSTWNTKWNSTGKITLPCRKQIINWQGYGKAMYAPPIEGLIKCSNITGLITRDGANNSHNETFRPGGDMR
 DNWRSELYKYKVKIRPLGVAPTAKRRVVEREKRAIGLGAMFGLGAGSTMGAASWTLTVQARQLLSGIVQQNNLLRAIEAQOHLQL
 TVWGIQLOARILLAVERYLKDOQLGNGCSGKHICTTVFPMNSWSNKSLEIWNNTWNEWEREIDNYTGLIYSLIEESQOQKNEQEL
 LELDKWASLWNWFSITQWLWYIKIFTMIVGGGLIGRIVFAVLSLVNRVRQGYSPLSFQILLPAPRGPDRPEGIEEGGEGQGRGSRIRLVNGF
 SALIWDLRNLCLESYHRLRDILLIARIVELLGRGWELKYLWNLQYWIQELKNSAISLFDTTAIAVAEGTDRVIEIVQACRAINIP
 TRIRQGLERALL\$

도면48b

2003 CON D Env. seq. opt

ATGCGGTGGCGGCATCAGCGCAATACAGACCTGTGGCTGGGCATCATGCTGTGGCATGCTGATGATGCTCGTGGTGGCGCA
GCACTGTGGTGACCTGTACGGGTCCGCTGTGGAGAGGCCACCAACCTGTTCTGGCTCGACGCGCAAGTCTCTCAAGA
CGAGGCCCAACATCTGGCCATCCGCTCGGTCCGACCGACCAACCCACAGAGATCGACTGGAGAACCTGACCGAGACTTC
AACATGTGAGAGAACAACTGTGGAGAGATGACAGGACATCATCTCTGTGGACAGTCTCTGTGAAGCCCTGGTGAAGCTGACCC
CTGTGGTACCTGGAATCGACGAGCTGAGCGACACACCTCTCAAGCTGAGCTGTCCTCAAGCTGAGGAGATGAAGAATGCTCTCTCAACA
TCACACCGAGATCGGCGAAGAAGACGAGTGCACGCTTCTTCAAGCTGAGCTGTCCTCAAGCTGAGGAGATGTCACCGTGCAGTGCACCCACGGGA
TCCTACCGCTGATCACTGCAACACTCGGCCAICACCCAGGCTGCCCAAGTGACTTGGAGCCATCCCATCCACTACTTGGCCCT
TCGCGCGGTGTCCCATCTGAACTCGAAGCAAGAATTCAAGGCGACCGGCTCGAAGAACTGTCCACCGTCCGAGAACCTGCACCAACAGCC
AAGATCATATCTGTGAGCTGAAGAGTCGTGACCTCAACTGACCGCCCTACCAACAACCCCGCAGCGACCCCTTCGCGCCGCG
CGAGCCCTGTACACACCGCATPAGCGGCGACATCCGACGCCCATCTGCATCACTTCGAGCCCTCTCGCGCGGCGACCCCGACTCTCT
TGGCGAAGAGCTGGCGAGCTGCTGACCAAGACCACTCATCTTCAGCCCTCTCTCGCGCGGCGACCCCGAGATCAACACCCACTCTCT
AACTGGGCGGGAGTTCCTTACTGCBAACCTCCCGCTGTTCATCTCCCTGGAAACAACCAAGTGGAACTGCAACCGGCAAGTCACT
CTTCCCTCCGCATCAACAGATCATCAACTGGAGGCTGGGCAAGGCATCTAGCCGCCCTCTGAGGGCTGTATCAAGTGT
CTCTCAACATCAACCGCTGTCTGACCGGACGGCGGCGCAACACTCCACACGAGACTTCGCGCGGCGGCGGAGACATCGCG
GACACTGGCGCTCGAGCTGACAAATPACAAAGTGGTGAAGATCGAGCTCTGGGCTGGCCCTCCACGCGGCGGCGGCGGAGACTCGCG
GGCGAGAGACCGCATCTGGCTGGCGGCGCATCTCTCGGGCTCTGGGCGCGCGGCTACCATGGCGCGGCTCCATGACCTGA
CTCTGACCGCGCGAGCTGTGTTCGGGCTCTGTGCGGCACTGTGACGACGAAACACTCTGCTGGCGGCCATCGAGCCAGCACTGCTCGAGCTG
ACCTGTGGGCGATCAAGAGCTCGAGCGCCGATCTGGCCCTGGAGCTACTCTGTGACAGCTGTAAGAGACCGAGCTGCGGCATCTGGGGCTGCT
CGGTAAGACATGATCAACACACAGCTGCCCTCGGAACTCTCTGTGCTCCAAAGTCTCTGGAGAGACTTGAACAACAATCAACCTCGATGAG
ATGTGGAGCGGAGATGACACATPACACCGGCTGATCTACTCTGTGAGAGTCCGACACCGAGAGAGAGACGAGACGAGAGCTG
CTGAGCTGGACAAATGGGCTCTCTGTGGAATGGTCTCCATCACCACTGGCTGTGTATCATAGATCTCATATGATCTGTGGCG
CTGTGCGCTGATCTGTGTGCGCTGTCTCTGTGTAACCGCTGCGCGAGGGCTACTCCCTCCCTGCTCTTCCAGACCCGTGCTGC
CCGTCACCGCGCGGCCGACCGCCCCGAGGGCATCGAGAGGAGGGCGGCGACGAGCGCGCGGCTCCATCCGCTGGTGAAGCGGCTG
TCGCGCCATGATCTGGGAGCACTGCGCAACTGTGCGCTTCTTCTACACCGCTCGGACAGCTGATCTGATCGCGCGCGGATCGTGG
CTGTGCGCGCGCGGCTGGAGGCGCTGAAGTACTGTGAACCTGCTGCAGTACTGGATCCAGAGACTGAAGACTCCGCGCATCTCC
TGTTCGACACCAACCGCCATTCGCGGTGGCGGAGGCAACCGCTGATTCGAGATCTGTGCGAGCTGCGCGCGGCTGCGGACATCCCT
TCCGCACTTCGCGACGGGCTGGAGCGGCGCTCTGTAA

도면49a

2003 CON F1 End

MRVGMQRWHLGKMGLELGILLICNAENLWTVYICVPVWKEATTTLFCASDAKSYEKEVHNWATHACVPTDENPQEVVLENVTENF
 DMKNNMVEQMHTDIIISLWDQSLKPCVKLTFLCVTLNCTEDVNATNNDTNDNKTGAIONCSFNMTTEVRDKLYHALFYKLDIVPISNNNSK
 YRLJNCNTSTITQACPKVSWDPIPIHYCAPAGYALLKCNDRKRENGTGPCKNVSTVQCTHGIRKPVVSTQLLNGSLAEEDIIIRSQNISDNAK
 TII'VHLNESVQINCTRPNNTRKSIHLGPGQAFYAIGETIGDIRKAHCNLSGTQWNKTLQVRAKLKSHFPNKTIFKNSSSGGDEIITMHSF
 NCRGEFFYCNTSGLFNDTGSNGTITLPCRIKOIYNMWOEYGRWYAPAGNITCNSTNITGLLITRDGGONNTEFFRPGGNNKDNWSELY
 KYKVEIEPLGVAFTKAKQVVKERRRRAVGIGAVFLGLGAGSTMGAASITLTQAROLLSGIVQQONNLLRAIEAQOHLQLTVWGIKQL
 QARVLAVERYLKDOOLGLWGCGKLICTINVPWNSSWNKSDIWNNTWMEWEKEISNYNIIYRIEESQOQKEKEQELLALDKWAS
 LWNWFDISNMLWYKIFIMIVGGIGLRIVFAVLISVNRVKGYSPLSLQTLIPSPREPRPEGIEEGGEGQKDRSVRLVNGFLALVWDDL
 RNLCLFSYRHLRDFILIAARIVDRLRGWEALKYLGNIITQYNSQELKNSAISLNTTAVVZEGTDRVIEALQAGRAVINIPRRIRQGLE
 RALLS

2003 CON_F1 Env.seq.opt

[illegible]

도면50a

2003 CON F2 Env

MRVRENQNRNQHLCGWGLLELILICNARDNLWTVYGVVPVWKEATTTLFCASDAKAYEREVHVNVWATYACVPTDPSFQELVLGNVTENF
 NNNKNNMVDQMHEDIISLWDSLRKPCVKLTPELCVTLNCTDVNVTINTTNVTLGELKNCSEFNITTEIKDKKKKEYALFYRLDVVPINNNSIVYR
 LISCNTSTVTQACPKVSFEPIPIHYCAPAGFAILLKCNDRKFNGTGLCRNVSTVQCTHGIRPVVSTQLLINGSIAEEDIIRSENIISDNTKTI
 IVQENRSVEINCTRENNTNTRKSIIRIGPGRAFAYATGDIIGDIRKAYCNINRTILWNETLKKVAEEENKHNFTVTENPSSGGDLEITTHSFNCR
 GEFFYCNTSLLFNNTEVNNTKTIITLPCRIRKQFVNMWQVRGKRWAPPIAGQIQCNENITGLLLTRDGGKNGSETIRPGGGDMRDNRSELYK
 YKVKIEPLGVAPTAKRQVQREKRAVGIGAVLIGFLGAGSTWGAASITLAVQARQLISGIVQQQSNLLKATEAOOHLQOLTVMGKLOLQ
 ARILAVERYLKDOQLLGIWCGSGKLICTTNVPMNSWSNKSQDEYWDNMTWQWEKEISNYTDTIYRLLEDAQNQQEKNEQDLLALDKWDNL
 WSWFTITNWLWYIKLFIIMVGGGLIGLRIVEAVLSVNNRVQGYSPLSQTLIPNRPGRPPGGIEEGEQDRDRSIRLVSGFLALAWDDL
 SLCLFSYRHLRDFILIAARTVDMGLKRGWEALKYILMNLPOYWGQELKNSAISILDTTAIAVAEGTDRIIEVLQRAGRAVLHIPRRIRQGFER
 ALL\$

2003 CON_F2 Env.seq.opt

ATCGCGCTGCGGAGATGAGCGCAACTGGCAGCACTTGGCGAAGTGGGCGCTGCTGTTCTTGGGCATCTGTATCATCTGTCAAGCGCGCGCGCA
CAACTGTGGGTGACCGTGTACTACGGCTCGCGTGTGGAAAGAGGCGCACCAACCACTGTTCTGGCGCTCGGACGCCAAGGCGCTACGAGC
CGAGGTGGGACACAACGTGTGGCGCCACTACGCTTGGTGGCCACGACCCCTCCCGCCAGGAGTGGTGGCTGGGCAACGTGACCGAGAACTTC
AACAATGTGGAAGAACAACTGGTGGACAGAGTACGACGAGGACATATCTCTGTGGACAGTCTCTGTAAGCGCTGGGTGAAGCTGACGCC
CCTTGTGGTGCACCTGAATCGACCGACGTGAAGCTGACCATCAACACCAAGCTGACCTTGGCGGAGATCAAGAACTGCTCTTCAACA
TCACCCAGAGATCAAGACAGAGAAAGAAAGTACGCCCTTCTTACCGCGTGGAGCTGGTGGCCATCAACAACATCCATCGTGTACCGC
CTGATCTCTGAACACTCCACCGCTACCCGAGCTCGCCCAAGTGTCTTGGAGCCATCCCTCACTACTTGGCGCGCGCGCGGCTT
GGCATCTGTAAGTCAACGACAGAAAGTTCAACGACACCGGCTGTGGCGAAAGTGTCCACGTGGAGTGCACCCAGGCACTCCGCGCCG
TGGTGTCAACCACTGCTGCTGAAACGCTCTGGCGAGGAGCAATCATCTCCGTCGAGAAACATCTCCGAAACCAACCAAGACCAT
ATCGTGCAGTTCAACCGCTTCGTGGAGATCACTTCAACCGCGCCAAACAACACCGCAAGTCACTCCGATCGGCGCGCGCGGCGCT
TACGCCCAAGGACCATCATCGGACATCCGAAAGGCTACTGCAACATCAACCGACCTGTGGAACGAGACCCCTGAAGAAAGTGGCCG
AGGATTCAGGACCACTTCAACATCACGTGACTTCAACCGCTTCTCCGGCGGAGCTTGAAGTACCAACCTCCCTTCAACTGCTCGC
GGCGAGTCTTCTACTGAACACTCCGACTCTTCAACAACGAGTGAACATCAACAACATCACTCCCTGCCCTCGCATCGGACCGCA
TGTGTGAACATGTGGCAGCGGTGGCGGCGCATGACGCCGCCCATCGCCGCGCAGATCCAGTGCACATCCACATCACCGGCTGG
TGTGTACCCGCGAGCGCGCAAGACGGCTCGAGACCTCGCGCGCGCGGCGGACATGGCGCAACTGGCGTCCGAGTGTATCAAG
TACAAGTGTGAAGATCGAGCCCTGGCGTGGCCCTCACCAAGCGCAAGCGGAGTGGTCAGCGGAGAGAGCGCGTGGGACATCGG
CGCGTGTGTGGGTCTCTGGCGCGCGCGGCTCCACATGGCGCGCCCTCCATCACTGACCTGCAGGCCCGCAGTGTCTCG
GCATCGTGAAGACGAGTCAACTGTCTGAAGCGCATCGAGCGCCAGCAGCTTCTGTCAGTGCAGCTGGGGATCAAGGAGCTGCAG
CGCTGTGAACCTCTCTGGCGTGAAGCTTCTGAGACAGAGCTGTGGCGATCTGGGCTGTCCGCGAAGTGTGTGATGTGTCACCAACCAAGT
GCCCTGGAACCTCTCTGCTCCAAACATCCCGAGGACAGATCTGGGACAAATGACTGGATGAGTGGGAGAAAGAGATCTCCAACTACA
CGACACCATCTACCGCTGTACGAGAGCGCGCAACACGAGAGAGAAAGACGAGAGCACTGCTGGCTGGCAATGGGCAACCTCG
TGGTCTGGTTCACATCAACCAACTGGCTGTGTACATCAAGATTTCACTATGATCTGGCGGCTGTGATCGGCGTGCACATCGTGTCCG
CGTGTCTCGGTGTGAACCGGTGGCGCAGGGCTACTCCCGCTGTCTCGACAGACCTGATCCCAACCGCGCGCGCGCGAGCGCCCG
CGGCGATCGAGGAGGCGCGGAGAGGACCGGACCTGCTCCCTGCTGTGGTGTCTTGGCTGTGGCTGGCTGGAGCAGCACTGGC
TCCCTGCGCTGTCTCTACCGCACTCGGCACTCATCTGATCGCGCGCGACCGTGGAGATGGGCTGAAGCGCGGCTGGGAGG
CTGGAAGTACCTGCGCCACTACTGGGCGCAGGCTGAACATCGGCCATCTCCCTGCTGAAACACACCGCGCATCGCGGTG
CCGAGGCGACCGAGCGGATCATCGAGGTGCTGACGCGCGCGCGCGCTGCTGACATCCCGCGCGCATCGCGAGGCTTCGAGCGC
CGCGCTGTGTA

도면51a

2003 CON G Env

MRVKGIGQNRQHLWKKWGTLLGLVLICASNNLWTVYGVVWEDADTTLFCASDAKAYSTERHNWATHACVFTDPNPQETILENTENF
 NMMKNMVEQMHEDIISLWDESLKPCVKLTPLCVTINCTDVNTNNNTNNTKKEIKNCSEFNITTEIRDKKKKEYALFYRLDVVPINDMGNSS
 IYRLINCNVSTIKQACPKNVTFDPIPTHYCAPAGEAILKCRDKKENGTPCKNVSTVQCTHGKIKPVVSTQLLNGSLAEEIIIRSENITDNT
 KVIIVQINETIEINCTRENNTKRSIRIGGQAFYATGDIQDIOAHCVNVRTKNENMLQKVKRQKIFNKSITFNSSSGDLEITTHSF
 NCRGEFFYCNSTGLFNNSLNSTITLPCIKQIVRMQVRVQAMYPPIAGNITCRSNITGLITRDGGNNNTETFRPGGDMRDNWRS
 ELYKKIVKLPGLVAPTRARRVVEREKRAVGLGAVLLGFLGAGSTWGAASITLTVQVRQLSGIVQQQSNLRAEAQQHLLQLTVWGI
 KQLOARVLAVERYLKDQQLGIMWCGSKLICITTNVWNTSNKSYNEIWDNMTWIEWEREISNYTQOYLSIEESQOQEKNEQDLALDK
 WASLWNVFDTKWLWYIKIFIMIVGGLIGLRIVFAVLSIYNRVROGYSPLSFQTLTHOREPDRPERTEEGGGEQDKRSIRIVSGELALAW
 UDLRSICLESYHRLRDFILIAARTVELLGRSSSLKGLRLGHECLKYLWNLILLYWGOELKNSAINLLDTAIAVANWTDVIEVAQACRAILN
 IPRRINQGERALL\$

도면51b

2003 CON 6 Env. seq. opt

AACCTGTTGGAAGGCATCCAGCGCAACTGSCAGCAACCTTGGAAAGTGGGGCAACCTGTATCTCTGGGCGCTGGTGATCATCTGTCGGCTCGCTCCAA
CAATCGTTGGGTGACCGGTGTACTACGGCTGCCGTGTGGAGAGCGGACACACCTCTGTCTGGCGCTCCGACGCCAAGCGCTACTCCA
CCGAGCGCCACACAGCTGTGGGCCACCCAGCGCTGCTGCCACCGACCCCAACCCCGAGAGATCACCTCTGGAGAACGTACCGAGAACTTC
ACATGTGGGAAGAACACATGTTGGAGAGATGACGAGGACATCATCTCCCTTGGGACAGCTCTGGTGAAGCTGACGCC
CCCTGTGGTGTACCTGAATCCCGCTGAAGCTGACCAACAAACACCAACCAAGAGAGATCAAGCTGCTCTCTCAACA
TCACCACCGAGATCCGGACACGAGAGAGAGAGTACGCCCTGTTTACCGCTGGAGTGGTCCCATCAGCAACACGCAACTCTCTC
ATATACCGCTGTATCACTGTCAAGCTGTCCCATCAAGCAGGCTGTCCCAAGTGTACTTGACCCCATCCCATCACTTCTGGCGCTA
CGCGCGCTTCGCATCTGTGAAGTGTGGGACACAAAGATTAACAGCGCACGGCCCTGCAGAGAACTGTCCACCTGTGATCGACCCAGGCA
TCAAGCCCGTGTGTGTCACCGCTGTCTGTGAACGGCTCCCTGGCCGAGGAGAGATCATCATCCCTCCCGAGAACATCACCGACACAC
AAGTGATCATGTGACGCTGAACCGGACCATCGAGATCAACTGACCCCGCCCAACAACTCATCCGATCTCGCATCGGCCCGG
CCAGCGCTCTAGCGCATCATCGGACATCATCGGAGCGCATCCGAGCGGCATCAAGCTGTCCGCAAGTGTGAAGACAGATGCTCGAGA
AGGTGAAGCCCGCTGAGAGAGATCTTTCAACAAAGTCCATCACTTCAACTCTCTCTCGGCGCGCACTCGAGATCACCAACCACCTCTTC
AATGCGCGCGGAGTTCTTACTGTGATGAAACCTTCGGCTGTCTCAACAACTCCCTCTGTGATCTCCCAACTCCATCACCTTCGCCCTG
CAAGATCAAGACGCTGTGCGATGTGGACGCTGGCGCAGCGCAATATACCGCTCCATCTGCGCGGCACTACTCTGCTCCCAACA
TCACCGGCTGTGTGATGACCCGAGCGGCAACAAACACAGAGCTTCGCCCGCGCGGGGAGCATCGCGGACAACTGGCGCTCC
GAGCTGTACAAGTACAGATCGTTGAAGTCAAGCCCTTGGCGTGGCCCCACCCGCGCGCGCGCTGGTGGAGCGCGAGAGCGCG
GTGGGCTGGCGCGCTGTGCTGGGCTGTCTGGGCGCGCGCTCCATCATGGCGCGCTCCATCACCTGAGCGCTGAGGTGGCGC
AAGAGCTGTGCGGATCTGTGAGAGAGATCCAACTGCTGGCGCTCAGGCGCAGCAGCACTCTCGCTGAGCTGACCTGTGGGATCTG
AAGCAGCTGCAGCCCGGCTGTGCGCGTGGAGCGGTACTGTGAAGCAACAGCAGCTGCTGGGCACTCGGGCTCTCCGCGAAGCTGATCTG
CACACCAACGTGCTGGAAACACTCTCTGTGTCACAAAGTCTTACAAAGAGATCTGGGACAAATGACTGGATGGAGCGCGGAGA
TCTCTCAACTACAGCAGGATCTACTCTGTATCGAGAGTCCAGAGTCCAGAGACAGAGAGCTGTGGCTGCCCTGTGGCATG
TGGGCTCCCTGTGGAATGTGTGACATCAAGTGGCTGTGTGTCATCAAGATCTCATGATGATCTGGCGGCTGTGATCGGCTGG
CATGTGTTCGCGTCTGTCCATCTGTGAACCGGTGGCGCAGGCTACTCCCGCTCTCTTCCAGACCTTCAACCAACCGCGGAGC
CGGACCGCTCCGCGCGATCGAGAGAGGCGGAGAGACAGGACCGGTCCATCGCGCTGGTCTCGGCTCTTGGCGCTGGCGCTGG
GAGAGACTGTGCGTCTGTCTGTCTTCATACACCGCTGGGAGCTCATCTGATCGCGCGCGACCTCGGAGCTGCTGGCGGCTG
CTCCCTGAAGGCGCTGCGCTGGCTGGGAGGCGCTGAGTAACTCTGTGAAACCTGCTCTGCTACTGGGCGCAGGAGCTGAAGAACTCCGCCA
TCAACTGCTGAGCAACCATCGCATCGCGGTGGCCAACTGAGACCGGCTGATCGAGGTGGCCCGCTTCGCGCGCATCTGTAAC
TCCCGCCCGCATCCCGAGCGGCTGGAGGCGCGCTGCTGTAA

도면52a

2003 CON H Env

TRVNETQRNYPSLWHEWGTLLIGMLLICSZAGNLWVTYYGVPVKEAKTTLFCASDAKAYETEKHNWATHACVPTDPNPQEMVLENTENF
 NMWENDMEQMHTDILSLWDQSLKPCVKLTPLCVTLDCSNVNTTNATNSRFNMQELTNCFSFNTTVIRKQKQKXVHALFYRLDYVPIDMNS
 YQYRLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVSTVOCTHGIRPWSVSTQLLNGSLAEQVIRSKNISDN
 TKNLIVOLNKPVEITCTRNNNTRKS IHLGPGQAFYATGDIIGDIRQAHCNISCKKWNKTLHQVVTQLGKYFDNRTIIFKPHSGGDMETTH
 SFNCRGEFFYCNSTGLENSSWINSTNDTKNIITLPCRIKQIVNMQRVQAMYPPIKGNITCVSNITGLILTFDEGNTVTTFPGGDMRD
 NWSSELYKYKVVKIEPLGVAPTEARRRVVEREKRAVGMAFFLGFAGSTMGAAISITLTVQARQLLSGIVOOSNLIRAIQAQOHLQLT
 VMGKQLOARVLAYERLYKDQQLGIGWCSGKLICTTNVFNWSSWSNKSLEINWNTMWEWKQINNTEEYRLLVESQTQOEKNEQDLL
 ALDKWASLWNWFSITNWLWYIKIFIMVGGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLIPNRPDPPEGIEEGGQDRDRSVRLVNGFL
 PLVWDDLRSICLFSYRLLRDLILLIWRVTVELIGRRGREALKYLWNLIQYWGQELKNSAINLLNTTAIAVAEGTDRIIEIVQRAWRAILLHIPR
 RIRQGFERTIL\$

도면52b

2003 CON H Env. seq. opt.

ACACCGTGTATGAGACACGACGGAACATACCCTCTCTTGGCGCTGGGCGACCTGATCCTGGGCATCTGCTGATCTCTCCGCGCCGCGG
CAACCGTGTGGGTACTACTACGCGCTGCCGCTGTGGAGGAGGCCAAGACCAACCTGTCTGGCGCTCCGACGCGCAAGGCTCTACGAGA
CCGGAAGACCAACGTGTGGGCGACCCACGCTTGGTCCGCAACCCCAACCCCGAGGAGATGGTGTGAGAACTGTACCGAGAACTTC
ACAATGTGGAGAACGACATGGTGGAGAGATGCACACCGACATCATCTCTGTGGACCACTCTCTGTGAAGCCCTGGTGAAGCTGACCC
CCTTGTGGTGAACCTTCCAGTCTCCAGTGAAACACCAACGCCCAACAATCCCGCTTCAACATGCAGAGGAGCTGCACCAACTGCTCCT
TCAACTGTGACACGCTGATTCGGAACAGAGAGAGTGCACCCCTGTTTACCGCTTCAAGTGTGGCCATGACGACAACAACTTC
TACCAGTACCGCTGTCAACTCAACATCTCGTGTACCTCCAGGCTGCCCAAGTGTCTTTCAGCCATCCCATCTCACTTACTCTGGG
CCCGCCGCGCTTGGCCATCTGTGATGACCAAGAGCTTCAACGGACCGGCCCTGTCCACACTGTCCACCGTGCAGTGCACCCAC
GCATCTCGCCCGTGTGTCACCAAGTCTGTGTGAAGGCTCTCTGGCCGAGGAGTGTATCTCCGCTCCAGAAACATCTCCGGAAC
ACCAAGAAACATATCTGTCAGACTGCAACAGCCCTGGAGATCACTGCACCCGCCCAACAACAACCCGCAAGTTCATCTCACTCGCTGGCC
CGGCCAGGCTTTTACGCCACCGCGACATCAATCGCGGACATCGCCAGGCCCACTGTCAACATCTCCGGCAAGAGTGGAAACAGACCTGTG
ACCAAGTGTGACCGAGTGGCAAGTATCTGACAACCGACATCATCTCAAGCGCATCTCCGGCGAGAGATGAGTGTACCATCTCGGCC
TCCTTCAACTGCCGCGGAGTTCTTCTACTGCAACACTTCGCGCCTGTCTCACTCTCTTGGACCAACTCCACCAACGACACCAAGAACAT
CATCAACCTGTCCGCTCCGACATCAAGAGATGTGAACATGTGGCAGCGGTGGGCCAGCATGTAGCCGCCCTCCCAATCAAGGCAACATCA
CTCTGGTGTCCAACTACACCGGCTGATCTTGACTTCTCAGCAGGCGAACAACAACGTGACTTCCGCGCGGCGGCGAGCATGCGCGAC
AACTGGCCTCCAGCTGTACAAATACAGTGTGAATGAGCCCTTGGCGGTGSCCCCAACCGAGGCCGCCGCCGCTGTGGTGGAGCG
CGAAGAGCGCCCTGGGATGAGGCGCTTTCTTGGCTTCTTGGGCGGTCTCCACTTGGGCGGCGCTCACTTGGCGGCGCTTCACTCACTCGG
TGAAGCGCCGCTGTGTCGCGATCTGTGACGACGACCAACTGTCTGCGCGCATCTCAGGCCGACAGACATGCTTCAGTGTAAC
GTGTGGGGCATCAAGCAGTGCAGGCCGCTGTGGCGTGGAGCGCTACTCTGAAGACCAAGCAGCTGCTGGGCGATCTGGGCGTGTCTCGG
CAGCTGTGTGCAACAACAAGTSCCTGGTGAACCTCTCTGGTGTCAAGATCTCTGGAGATCTGGGAACAATGACCTGGATGGAGT
GGGAACAGCAGATCAACAACATACACGAGAGACTACCGCTCTGAGAGTGTCCAGAGCCGACGAGAGAGAACGACGAGCAGCTGTCTG
GCCCTGGAACAAGTGGCCCTCCCTGTGAACTGTGTTCTCCATCAACATGGCTGTGTGTATCAATCAAGATCTTCACTCATCTGTGGGCGGCGCT
ATCCGCTGTCCGATCATCTTCGCGTGTCTCCATCTGAAGAGCGCTCGCGAGCGTACTCCCGCTGTCTCCAGACCTTGATCTCCCA
AATCCCGCGGCCCGACCCCGCCAGGGCATTCGAGGAGAGGCGGCGCAGCAGCCGACCGCTCCGCTGCCTTGGTGAACCGCTTCCTG
CCCTGTGTGTGGAGACCTGTGGCTCCCTGTGCTGTCTTCTTACCGCTGTCTGCGCACTCTGTCTGTATCTGTGTGGCGACCGCTGGAGCT
GCTGGGCGCGCGCGCGAGCCCTGTGAATACCTGTGAACCTGTCTGATCTGGCGAGGAGCTGAGAACATCCGCGCATCACTGAGCTGC
TGACATCACACCGCATCGCCTGTGCGCTGAGGGCACTGACGCAATCATCGAGATGTGTGACGCGCGCTTGGGCGGCGCATCTGTGCATCTCCCGCG

GGATCTCCGCGAGGCTTCGAGCGCACCTCTGCTGTAA

도면53a

2003 CON 01 AE Env
 MRVRETQMWPNLAKWGTLLGLVICSZSDNLWTVVYCVPVWRDADTTLFCASDAKAHETEVEHNVWATHACVPTDPNPQELHENVTFENE
 NMWKNMVEQMDEVISLWDQSLKECVKLPPLCVTLNCTNANLTNVNNTITVNSNIIGNITNEVNCSTNMTELRDKKQKVHALFYKLDIVQ
 IEDNNSYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNKDNFNGTGPKNVSSVQCTHGIGKPVVSTQLLLNGSLAEELIIRSEN
 LTNNAKTIIVHLNKSVEINCTRPSNNTRTSITIGPGQVYRTGDIIGDIRKAYCEINGTKWNEVLKQTEKKEHFNKTIIFQPPSGGDLE
 ITMHFNCRGEEFFYCNTTKLENNTCIGNETMEGNGTIIIPCKTKQIINWOGAGQAMVAPPIISGRINCVSNTIGILLTRDGGANNNTETFR
 PGGGNIKONWRSLEYKYKVQIEPLGIAPTRAKRVVEREKRAVIGIGAMI EGTGAAGSTMGRASITLTVQARQLLSGIVQQQSNLLRAIEA
 QQHLLQITVWGIKQLQARVLAVERYLKQKFLGWCCSGKLICTTAVPWNSTWSNRSPFEEIWNMTWIEWERETSNTYQIYEILLTESQNOQ
 DRNEKDLLELDKWAASLWNWFDTIWNLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVROGYSPLSFQTPTHHQRPDRPERIEEGGEGQGRORS
 VRLVSGGTALAWDDIRSLCLFSYHRLRDFILIAARTVELLGHSSLKGLRRGWECIKYLGNNLLLYWGQELKISAILSDATAIAVAGWTDRTVI
 EVAOGAWRALIHI PRRIROGLERALL\$

도면53b

2003 CON 01 AE Env. seq. opt

ATGGCGCTGAAGAGACCCAGATACACCTGGCCCAACCTGTGGAAGTGGGAGCAACCACTGATCTCTGGCCCTGGTGATCTGCTCGCCCTCCGGA
CAACCTGTGGGTGACCTGTACTACGGCTCGCGTGTGGCGACGCGACACCACTGTTCTGGCCTCGACGCGCAAGCCGACACGAGA
CCGAGGTGGACAACCGTGTGGGCGACCAAGCCCTGGCGCCACCGCCACCGCCAGGAGATCCACCTGGAGAACTGACCGAGAACTTC
AACAATGTGGAAGAACAACTGTGTGAGCAGATCAGAGAGCGATGATCTCCCTTGTGGACAGTCCCTGAAGCCCTGGTGAAGCTGACCCG
CCTGTGGGTGACCTCGTGAACCTGACCAACGGCAACTGACCAACACTGAAACAACAATCAACAGCTGTCCAAACATCATCGGCAACATCAACGACG
AGGTGGCAAACTGCTCTTCTACATGACCAACGAGCTGGCGCAAGAGACGAGAGTGCAGCCCTGTTCTCAAGCTGGATCACTTCGACG
ATCAGGAGCAACAACACTCTTACCGCTGTATCAACCACTCCGTGATCAACGAGCCTGCCCAAGATCTCTCTCGACCCCATCCCCAT
CCACTACTGCAACCCCGCGCTAGCGCATCTGAAAGTGCACAGAACTTCAACGGCAACGGCCCTGCAAGAACCTGTCTCTCGGTGC
AGTGCCCAACCGGATCAAGCCCTGTCCACAGCTGCTGCTGAGAGTCAACTGCACCCGCCCTCCAAACAACCCGACCCCTTCAT
CTGACCAACAACGCCAAGCAATCATCTGCACCTGACAACTCCGTGAGATCAACTGCACCCGCCCTCCAAACAACCCGACCCCTTCAT
CACCATCGCCCGGCGAGTGTTCTACCGACCGGCGACATCATCGGCACTCCGAGAGCTTCTCGGAGATCAAGGAGATCAAGGAGATCAAGTGAG
ACAGGTTGCTGAAGCAGTGCAGAGAGCTGAAGAGCACTCAACAACAGCACTCATCTTCAGGCCCTCGCGCGGCACTCGATGAG
ATCAACATGCCACCACTTCACTGCGCGGCGAGTCTTCTTACTGCAACACCAACCAAGCTGTTCAACAACACCTGCATCGGCAAGAGACCAT
GGAGGGCTGCAACGGCACTCATCTGCGCTCGAAGATCAGCAATCATCAACTGTGGAGGGCGCCGGCGAGGCTGATGACGCCCTCCG
CCATCTCGCGCGCATCACTGCTGCAACATCAACGGCATCTGTGTGACCGCGAGGGCGCAACAACACCAAGCACTGATCCGCG
CCCGCGGGGGCAACATCAAGAHCACTGGCGCTCCGAGTGACAACTAAGTGTGACAGATCAGACCCCTTGGGATCGCCGCCACCGG
CGCAACCGCGCTGTGGAGCGAGAGCGCGCGTGGGATCGGCGCATATCTTCGCTTCGCTGGCGCGGCTGGCGGGTTCACCATGG
CGCGCGCTTCATCACTCGACCTGACGCGCGCAGCTGCTCGCGCATCTGTGACAGCAGTCCCAACTGCTGCGCGCATCGAGGCG
CAGACACACTGCTGACGTGACCGTGTGGGGATCAAGAGCTGACGCGCGCTGTGCGCGTGGAGCGCTACCTGAAGGACACCGAAGTT
CTGGGCGTGTGGGGTGTCTCGGCAAGTCACTGTCAACCGCTGCCCTGGAACTCACTGTGTCAACCGCTCTCTTCGAGGAGATCT
GGAAACAACATGACCTGGATCGAGTGGAGCGGAGATCTCAACTACACAACAAGATCTGACAGATCTTGACGAGTCCCGAGAACGACGAG
GACCGCAACGAGAAGCACTGCTGAGCTGGACAGTGGCCCTCCCTGTGGAACTGTTGCAATCAACAACCTGGCTGTGTGATCAAGAT
CTTATCATGATCTGGGGGCGCATATGCGCTCGGCATCATCTGCGCTGCTTCATCTGTGAACCGCGGAGCGGCTACTCCGCC
TGCTCTTCAGACCCCAACCAACGAGCGAGCCGACCGCCCGAGCGCATCGAGGAGGCGGGCGGCGAGCGGCGGACCGCTC
TGTGGCCTTGTTCTCGCGCTTCTGGCCCTGGCTGGAGACAGTTCGCTCCCTGCTGCTTCTTACCAACCGCTTCGCGGACTTCATCT
GATCCCGCGCGACCTGGAGTGTGGCACTCTCTCCCTGAAGGCGCTGCGCGCGGCTGGAGGCGGCTCAAGTACCTGGGCAACCTGC
TGTGTTATGCGGCGAGGATGAAGTATCCGCAATCTCCCTGTGACACCAACCGCATCGCGTGGCGGCTGGACGACCGGCTGATC
AGGTGGCCCGAGGCGAGTGCCTCTCGAATCTCCCGCGCGCATCCGACGAGCGCTGGAGCGCGCTGCTGTAA

도면54a

2003 CON 02 AG Env
 MRVNGLOKNYPILLWRGMLIFWIMLICHNAENLWTVVYGVVPWEDAETTLFCASDAKAYDTEVHNWATHACVPTDPNPQELHLENVTENEN
 MWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLDCNNITNSNTNNNAGEIKNCSTFMTTELDRKKQVYALFYRLDVQOINKNNSQYR
 LINCNTSALTQACPKVSEPIPIHYCAPAGEAILKCNCKEFGTGPCKNVSTVQCTHGKPKVYSTOLKINGSIAEEEIVIRSEMITNNAKTI
 IVQLVKVKINCITRPNNNTRKSVRIGPGOTFYATGDIIGDIRQAHCVNSTRKWNNTLOQVATQRLKYENKTIIFANPSGGDEITTHSFNCG
 GEFFYCNTSEFNSTWNNTNTEKCIITLCRIKQIVNMWQVQAMYPPIQGVIRCESNITGLLLTRDGGNNNSTNETFRFGGDMRDNW
 RSELYKYKVVKIEPLGVAPTRAKRVVEREKRAVGLGAVELGLGAGSTMGAASITLTVOARQLLSGIVQOQSNLLRAIEAQOHLKLTVM
 GIKQLOARVLALERYLKDOQLLGIWGCCKLICITTVFVNNSSWSNKTNDINDNMTWLQWDEISNYTDIILYNLIEESQOQEKNEQDLAL
 DKWASLWNNFDTNMLWYIKIFINIVGGLIGLRIVFAVLTIIINVRQGYSPLSFOTLTHHQREPDPRPERIEGGGEQDRDRSVRLVSGFLAL
 AWDDLRSGLFSYHRLRDFVLIARTVELLGHSSLKGLRLGWEALKYLGILLSYWGQELKNSAINLLDTIAIAVANWTDVRVIEIGQAGRAI
 INIPERRIROGLERALL\$

도면54b

2003 CON_02_AG Env.seq.opt

ATGGCGGTGATGGGCATCCAGAGAACTACCCCTGCTGTGGCGTGGGCATGATCATCTTCTGGATCAIGATCATCTGCAACGCCGCGAGAA
 CTTGTGGGACAGGTGTATACGGGCTGGCGTGTGGGCGAGCGGACACCCCTGATCTGTGGCTTCGGACCGTAGGCCCTACGACACCG
 AGGTGCACAACTGTGGGCACCAACCGCTCGTGGCCACCGACCCAGCCAGTCCACTGGAGAAAGCTGACGAGACTTCAAC
 ATGTGGAGAGAACACATGTTGGAGCAGATGCACAGGACATCATCTCCCTGTGGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCCTT
 GTGCGTGAACCTTGGATCGGACACAAACATCACCAACTCCACACACGACCGCGGAGATACAGAACTGTCTCTTCAACATGA
 CACCGAGCTGGCGACAGACAGAAAGTGTAGCCCTGTTTACCCCTCGACCTGGTGAGATCAACAGAACTCCAGTACCGG
 CTGATCAACTGCACACCTCCGCCATCACCCAGGCTGCCCAAAGTGTCTTGAGCCCATCCCATCTCACCTATCTGGCCGCCCGCGGCT
 CGCCATCTTGAAGTGCACGACAGAACTCAACGGCCAGCGCCCTCGAAGAACTGTGCACCTCGATGACGCCACACGACATCAAGSCCT
 TGGTGTCCACACAGTCTGTCTGAAGCGCTCCCTGGCGAGGAGATGTGATCCGCTCGGAGAACTACCAACACACCGACAGACCAT
 ATCTGTGCACTGTGAAGCCCGTGAAGATCAACTGACATCCGCCCAACAAACACCCGATCTGGCGTCTGGCCCGGCGAGACCTT
 CTACGCCACGGGCACATCATGGCGACATCCGCGAGGCCACTCAAGCTCTCCGCAACAGTGAACAAACACCTTCAGCAGGTGGCCA
 CCAAGTCGCGCACTACTTCAACAGAACCATCATCTTCGCAACCCCTCGCGCGGCGACCTGGAGATCACCACTCTTCAACTGGCG
 GCGAGTCTCTTACTGCAACACTCCGAGCTGTTCACCTCCACTGGAACTCCACCTTGGAAACACCGAGAAGTGCATCACTCTGACCT
 TCCATCAAGCAGATCTGAACATCTGGCAGAGTGGCGACCATATGCAACCCCTCCATCCAGGCGGTATCCGCTCGGAGTCCCAACA
 TCACCGGCTCGCTGACTGACCCGACGCGCGCGCAACAACTCCACACAGACCTCCCGCCCGCGGGCGGACATGCGCGACACTGG
 CGTCCGAGCTHACAAATGAAGTGTGAAGATCGACCTCGGGGTGGCCCAACCGCTCCGCGCGGCGGCGGACATGCGCGACACTGG
 CGCGCCGTGGGCTGGCGCGGTGTTCCTGGCTCTCTGGCGCGCGGCTCACCATCGGCGGCTCCATCACTCACTGACCTGCTCGAG
 CCSCCAGTCTCTTCCGACATCGTCSAGCAGGATCCAACTGCTCGCGCCATCGAGGCCACGACACTGCTGAAGCTCAACCGTGTGG
 GGATCAACAGACTGCAGGCCGCTGTGGCGTCTGGAGGCTACTGAAGAGATCAACAGACATCTGGCAACATCGGACCTGGCTGAGCTGGGACA
 GATCTGCACAACTGCTCCCTGGAATCTCTCTGTGTCCAAAGACTACAAACGACATCTGGCAACATCGACCTGGCTCGACTGGGACA
 AGGAGATCTCCAACTACACGACATCATCTACACTGATCGAGAGTCCCACACACGACGAGAGAAACAGACGAGACTGCTGTCGCCCTT
 GACAAATGGGCCCTCCCTGTGAACTGTTCGACATCACCAACTGGCTGTGTACATCAAGATCTTCATATATGTGTGGGCGGCTGATCGG
 CTTGCGCATCGTGTCCGCTGTGACCATCAACCGGTGGCGAGGCTACTCCGCCCTGTCTCCAGACCTTGACCCACCAACCG
 CGGAGCCGACAGCCCGCCGAGCGCATCGAGGAGGCGGCGGACGACCGACCGACCGCTCGCTGGCTGCTCGGCTGCTGCTG
 GCTTGGGACGACTCGCTTCCCTGCTGTCTTCTTACCGGCTCGGAGACTTGTGCTGATCGCGCCGCCACCGCTGGAGCTGCTGGG
 CCACTCTCTCCCTGAAGGCTTGCGCTCGGCTGGAGGCGCTCAAGTACTTGGGCAACTGCTGCTTACTTGGGCGCAGGAGCTGAAGA
 CTCCGACATCAACTGCTGGACATCGGCATCGCGTGGGCACTGGACCTCCGCGTGAATCGGACGCGCGCGCGCGGCCCTC
 TGGACATCTCCCGCGGCATCCGCAACGGGCTGGAGCGGCGCTTGTGTAA

도면55a

2003 CON 03 AB Env

MRVKEIRKHLRWGTLFLGMLMICSATENLWTVVYGVPUWKEATTTLPCASDAKAYSKEVHNWATYACVPTDPSQOEIPIENVTENENMG
 KNNVVEQMHEDEIISLWDQSLKPCVKLTPLCVTLNCTDLKKNVTSTNTSSIKMMEMKNCSEFNITTDLRKVKKEYALFYKLDVYQIDNDSYRL
 ISCNTSVVTQACPKISFEPIPIHYCAPAGFAILKCNDRKENGTFCTNVSTVQCTHGIKPVVSTQLLNGSLAEVEVIRSVNFTDNTKTII
 VQLEPVEINCTRPNNNTKGIHICPGRAFYATGDIIGDIRQAHONISITKWNNTIKQIVIKRKQFGNKTIVNQSSGGDPEIVMHSFNCG
 GEFYCNITTKLFNSTWNGTEELNNTGDIVTLPCRKQINMWDEVGKAWYAPPIAGQIRCSSNITGLLTRDGNQSNVTEIFRPGGDMR
 DNRSLEYKYKVVKIEPLGVAPTRAKRRVQREKRAVGIGAVFLGAGSTWGAASITLTQARQLLSGIVQQNNLLRAIEAQOHLQOL
 TWGIKQLOARVLAVERYLKDOQLGINGCSGKLICTTAVPWNTSWSNKSISDELWNNMTWMEWEIREINNYTGLIYNLIEESQOQEKNEQEI
 LALDKWASLWNWFEDISKWLWYIKIIFINIVGGLVGLRIIEAVLSIVNRVQGYSELSTFQKLFTQRPDRPEGIEEGGERDRDTSIRLVNGF
 LALLWDLRSICLFITYHHLRDULLIAARIVELLGRRGWEALKYWNLLQYWIQELKSSAINLIDTIAIYAVAGWTDREVIEIGQRTFCRAIRNIP
 RRIQGAEKALQ\$

도면55b

2003 CON 03 AB Env .seq. opt

ATGCGCGTGAAGAGATCCGCAAGACCTGTGGCGCTGGGCAACCTGTTCTTGGGCAATGCTGATGATCTCCGCCACCGAGAACCTGTGC
GGTAGCCTGTACTACCGCGTGCCTGTTGAAGAGAGCCACCAACCTGTTGTCGCTCCGACGCAAGGCTACTCCCAAGGAGTGTC
ACAAAGTGTGGCCACCTACCGCTGGTGCCTCCACCCAGAGATCCCCCTGGAGACGTGACCGAGACTTCAACATGGGC
AAGAACACATGGTGGAGCAGTGCAGGAGACATCTCCCTGTGGACCTGCTCCCTGAAGCCTGCTGAAGCTGACCTGCTGGT
GACCTGAACCTGACCGACCTGAAGAAGAACGTGACCTCCCAACACCTCTCCATCAAGATGATGGAGATGAAGAACTGCTCCCTTCACAA
TCACCAACGACCTGCGGACAGGTGAAGAGGAGTACGCGCTGTTCTAAGCTGGACGTGGAGATCGACACGACTCTTACCGCTG
ATCTCTGCAACACCTCCGTTGACCCAGGCTGCCCAAGATCTCTGAGGCCATCCCACTGACCTACTGCGCCCGCCGCTTCGC
CATCTGAAGTGCACGACACGAAGTTCACGGCACCGGCCCTGCAACCAAGTGTCCACGTTGCCAGTGCACCCAGGGCATCAAGCCCGTGG
TGTCCACCCAGCTGCTGACGGCTCCCTGGCCAGGAGGAGTGGTGAATCCGCTCCGTGAATTCACGACACCAAGACCATCATC
GTGAGCTGAAGAGGACCTGTGGAGATCAACTGCACCCGCCCAACCAACACCGCCAGGCAATCCACATCGGCCCCCGGCGCTTTCTA
CGCCACCGCGACATCATCGCGACATCCGCAAGGCCACATGCAATCTCGATCACCACTGAGAACACACCTGAGAGATCGTGAATCA
AGCTGGCAAGCAGTTCCGGCAACAGACCATGTTTCAACAGTCTCCGGCGGACCCCGAGATCGTGAATGCTCTCACTGGGCG
GGGAGTCTTTACTGCAACACCAACAGCTTTCAACTCCACCTGGAAGCGGACCCAGAGCTGTACGCCCTCCCATCGCCGCGAGATCCGCT
CCTGCCCTGCCCATCAAGCAGATCATCAATGTGGCAGAGTGGCAAGCCATGTACGCCCTCCCATCGCCGCGAGATCCGCTGCT
GACAACTGGCGCTCCGAGCTTCAAGTACAAAGTGTGAAGATCAAGCCCTGGCGTGGCCCTCCCAAGGCCAAGCGCGCTGTGCA
GCGGAGAAGCGCGCTGGCATCGCGCGTGTCTGGGCTTCTGGCGCGCCCGCTCCACCATGGGCGCGCTCCATCAACCTGA
CCGTCAGGCGCGCGCTGCTGTCGGGCAATCGTGAGCAGAGAACCTGCTCGCGCCATCGAGGCCAGCAGCAGCTGTGCTGCTGCTG
ACCGTGTGGGCATCAAGAGCTGCAAGCCCGCTGCTGGCGTGAAGATCAAGCCCTGGCGTGGCCCTCCCAAGGCCAAGCGCGCTGTGCA
CGCAAGCTGATCTGCACACCGCGCTGCCGTGGACACCTCTGCTGCCACCACTCCCTGGTCCACAGATCCCTGGACAGATCGGAGATGG
AGTGGAGCGGAGATCAACAACTACCGGCTGATCTACACCTGATCGAGAGTCCGAGAGTCCGAGAGTCCGAGAGTCCGAGAGTCCGAGAGT
CTGGCTTGACAAAGTGGGCTCCCTGTGGAATGGTTCGATCTCCAGTGGTGGTACATCAAGTCTTCAATCATCATGATGCTGGGCGG
CCTGTGGGCTCGGATCATCTCGCGCTGCTGCCATGTAACCGCTGCCAGGCTACTCCCGCTGCTTCCAGACCGCGCTGC
CCACAGCGGCGCGGACCGCGCGGAGGATCGAGAGGAGGCGGAGGCGGAGCGGAGCTCCATCGCGCTGGTGAACGCTTC
CTGGCTGATCTGGACGACTGGCTCCCTGTGCTGCTCATACCAACCTCGCGGAGCTGCTGATCGCGCGCGCATCGTGA
GTGTCGACACCATCGCCATCGCCGTGGCGCTGACCCGACCTGATCGAGATCGGCGCTTCTCGCGGCTTCTCGGACCATCGCC
CGCGCATCGCGCGCGGCGGAGAGGCGCTGCAGTAA

도면56a

2003 CON_04 CPX Env

MRVNGIQRTYPHLWNGTLLGLVVICASAKNLWTVYGVFVWRDAETTPFCASDAKAYDKEVHNINWATHACVPTDPNPQETALKNVTENF
 NMWKNMVEQMHEDIISLWDEGLKPCVKLTPLCVALNCNATINNSTKTNSTEEIKNCSEFNITTEIRDKKREYALFYRLDIVEPINDSANN
 SINSEYMLINCNASTIKOACPKVTFEPIPIHYCAPAGEAILKCNKNFTGLGPCITNVSSVOCTHGKPYVSTOLLNGSLATEGVIRSKNF
 TDNTKNIIVQLAKAVKINCITRPNNTRKSVHIGPCQWTWATGEIIGDIRQAHCNISGNDWNETLQKIVBELRKHFPNKTIIFAPSAGGDEI
 TTHSENCGGEFFYCNTSELFNSTYMNSTNTINKTITLFCRIKQIVSMQEVGQAMYPPIAGSINCSDITGILLTRDGGNNNTNNETFR
 PGGGDMRDNRSELYKVKVVGIEPVGAPTRARRRVQREKRAVGIGAVTIGLGAAGSTMGAASITLTVOARQLLSGIVQQSNLLRAIEA
 QQHLRLTVWGKIQIARVLALESYLDQQLLGWCGSGKLCITNVFVNSSWNKSYNDIWDNMTLQWDKEINNYTQIIYELLEESQNZQ
 EKNEQDLLALDKWANLNNWENISNMWYIKIFIMTVGGGLICRLIIFAVLSIVNRVQGYSPLSQTLIPTTORGPDRPEETEEEGEODRSR
 SIRLVNGFLPLIWDLLRNLCLEFSYRHLRNLLIJVARTVELIGIRGWEALKYLNWLLIYWGQELNSAINLLDTTAIAVAEGTDRIIEAVQRA
 CRAIRNIPRRITROGLERALLS

도면56b

2003 CON_04_CPX Env.seq.opt

ATGGCGTGATGTGGGATCTCAGCGCAATCTACCCCACTCTGGAGTGGGCACCTGATCTCTGGGCTGTGTAATCAATCTGCTCCGCTCCAA
 GAACCTGTGGTGAATGTGTATACGGCTGGCCGTGTGGCGAGCGCGAGACAACCTCTTGCCCTCCAGCCCAAGCCCTACGACA
 AGAGGTGCACCAATCTTGGCGACCGACCTCGTGGCCACCGACCCCAACCCAGAGATCGCCTTGAAGAAGCTGACGAGACTTC
 AACCTGTGGAAAGAACAACTGTGTGAGCAGATGCAOAGGACATCATCTCCCTGTGGACAGAGGCTTGAGCCCTGTGCTGAAGCTGACCC
 CATCTGGTGGCTGAACTGCTCCAAAGCCACATCAACTCTCCAAAGACAACACTCCACGAGGAGATCAAGAAGCTGCTCTTCAACA
 TCACACCGAGATCCGCGACAAAGAGAAGAGATCCGCTGTCTATCCGCTGGACATCTGCGCCATCAACGACTCCGCCAACAAAC
 TCCATCAACTCCGATCATGCTGATCAACTGAAGCAAGCTCCACATCAAGCAGCTGGCCCAAGTGAATCTTCCAGCTGCTCCGTCAG
 CTATCGCCCTCCGCGCTGCTCTCACTTGAATGAACGAAAGAACTCACCGGCTGGGCCCTGTCACCAAGTGTCTCTCCGTGTCAG
 GACACCGGATCAAGCCGTTGGTGTCCACCCAGCTGCTGTGAAGGCTCCCTGGCCACCGAGGCGTGGTGAATCTCCGCTCCAGAACTTC
 ACCGAACAACCAAGAACATCATCTGTGAGCTTGGCAGCTGGCCAGGCGGTGAAGATCAACTGCAACGCCCAACAACAACCGCAAGTCTGCA
 CATGGCCCGCGGCAAGCATGTGTATGGCAGCGAGATCTCCGCGACATCCGCGAGCCCACTGCAACATCTCCGGCAGCAGCTGGAAC
 AGACTCTGCAAGATGTGTGAGGAGCTGCGCAAGCAATCTCCCAACAGAACATCATCTTCCCTCCCTCGCCGCGGAGCTCGAGATC
 ACCAGCCACTCTTCAACTGGGCGGAGTTCTCTATCTGAACACCTCCGAGCTGTTCAACTCACTCATGAAGATCCCAACTCCAC
 TCGATCAACAAGACATCACTCTCGCTCGGCATCAAGCAATCTGTCTCATGTGGAGAGGTGGCGAGGCCATGTACGCGCCGCCCA
 CCGCGGCTCATCAACTGCTCTCCACATCAACCGGATCATCTGACCCCGACGCGGGCAACAACAACAGAACATGAGACCTTCGCG
 CCGCGCGCGGCGCAATGGCGCAACTTGGCTCCGAGCTGTACAAGATCAAGACTCGAGCCCTGGCGCTGAGCCCTGAGCCGACACCG
 CGCCCGCGCGGTGTGACGCGAGAACCGCGCTGGGATCCGCGCGTGTCTCTGGGCTTCTTGGGCGCGCGCGCTCCACCATGG
 GCGCGCCTCATACCTGATCCGTGAGCGCCCGACCTGCTGTCCGCGATCTGTGACAGACAGTCCAACTCTACTTGGGCCATGAGGCG
 CAGAGCACTGCTGGCCTGACGTTGGGGCATCAAGCACTGACGCGCGCTGTGGCCCTGAGTCTCACTTGAAGCAACGACGCT
 GTGGGCACTTGGGCTGCTCCGGAAGCTGATCTGCAACCAACGCTGCCCTGGAACTCTCTCTGGTCCAAACAGTCTTACAAACGACATCT
 GGGACAACATGCACTGGCTGCAGTGGGACAGAGATCAACAGATACATCAAGACTGCTGGAGAGTCCAGAGCTCCAGAACCAAGAG
 GAGAGAAGAGCAGCACTGCTGGCCCTTGACAGTGTGGCCAACTCTGTGAACTGTTTCAACATGTTTCAACTGAGCTCCAGAACCAAGAG
 TCTTCATGATCTGTGGGCGCTGAGCTCGGCTCGCATCATCTCGCGGTGTTCATCTGTCAACCGCTGGCCAGGCTACTCCCCC
 TGTCTCGAGACCTGATCCCCACACACCGCGGCCCCGACCGCCCGAGGCAACCGAGGAGGAGGAGGACGAGCACTCCCGC
 TCCATCCGCTGTGAACGGCTTCTTCCCTGTATGTGGAGCTTGGACCTGCGCAACTCTCTGTTCTTACCGGCACTGAGGCACTGCT
 ATGATGTGTGGCCGCACTGTGGAGCTGTGGCATCCGCGCTGGGAGGCCGTGAATCTTGTGACCTCCGCGCACTCGGCAAGCTGCT
 AGCTGCGAACTCCGCACTCAACCTCTGTGACACCAACCGCATCTCGCTGGCGAGGACCAACCGCATCACTGAGGCGCTGACAGCGGCC
 TCGCGGCCCATCCGCAACAATCCCCCGCCGATCCGCGAGGGCTGGAGCGCGCTGCTGTA

도면57a

2003 CON 06 CPX Env

MRVKGIOKRWQHLLWKWGTLLGLVICSASNMMVTVVYGVPAWEDADTILFCASDAKAYSAEKHNVWATHACVETDRNPQETALENVTENF
 NMWKNHVEQMHEDIISLWDESLKPCVKLTPLCVILNCINVTKNNTKINGREEIKNCSEFNVTETIRDKKKKEYALFYKLDVVPIDDNNNSY
 RLINCNASTLKQACPVSFEPPIPIHYCAPAGFAILKCRDNFNFGTGPCKNVSTVQCTHGKIPVYSTQLLINGSIAEEETIIKSENLTDNTKT
 IIVOLNKSVETRCTRPNNNTRKSISEFGQOAFYATGDIIGDIRQAHCVSRTDNNMLQNVTAKLKELKNITFNSAGGDLITTHSFNC
 GGEFFYCNTSOLFSTRPNETNTTILPCKIKQIVRMQRVGQAWYAPPIAGNITCTSNITGLLTROGNNNSSETFRPGGDMEDNWRSELY
 KYKVYKIKPLGIAFTFRARRRVGREGKRAVGLGAVELGFTGTAGSTNGAASITLTQVRQLLSGIVQQSNLLRAIEAOQHLLQLTVMGKQL
 QARVLAVERYLKDOQLLGHGWCSGKLICTNPVWNASMKNXYNEINDNMTWIEWDRENNYTQIYSLIEESQOQEKNEQDALLADKWS
 LWSFEDISNWLWYIKIFIMIVGGIIGLRIVFAVLSIVNRVROGYSPLSQTLIPNPTGADRPGEIEEGGEGQKTRISRLVNGFLALAWDDL
 RSLCLFSYHRLRDFVLIAARTVETLIGHRGWEILLKYIGNLVICYWGOELKNSAISLLDITTAIAVANWTDRAVIEVQVRFRAFLNIPRRIQGF
 RALL\$

도면58a

2003 CON 08 BC Env

MRVGRTRNYQOOWIAGVLGFWMMLMCNVENLWVTVYGVPMKEAKTTLFCASDAKAYETEYHNVWATHACVPTDRNPQEIYMNVTENF
 NMNNDMVNQMHEDVISLWDQSLKECVKLPCLCVTELECTNVSSNGNGTINETYNESVKEIKNCSFNATLLRDRKKTYVALFYKLDIVPLND
 ENSGRNSSEYRLINCNTSAITQACPVTDFPIPIHYCTPAGYALIKCNKKFNCTGQCHNVSTVQCTHGIRKPVVSTQILLNGSLAEREIII
 RSENTNNVXTIIVHINQSVELVCTRPNNTKRSIRIGQOTFYANGDIIGDIRQAHCHNISKDKWYETIQVSKKLAEHFPNKTIKEASSSG
 GDLEITHSENCRGEEFFYCNTSGLENGTYMNGTNNSSSIITIPCRIKQIJNMWQEVGRAMYAPPIEGNITCKSNITGLLLVROGGRTESNNT
 EIFRPGGDMRNMRNELYKVKVEIKPLGVAPTAAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQOOSNLLR
 AIEAQOHMLQLTVMGIKOLOPRVLAIERVLKDQQLGIWGCSSGKLICTTAVPWNSSWSNKSQQEINWDMWTMOWDKEISNYTNTIYRLLEDS
 QNQQERNEKDLLALDSWKNLWSWFDITNWLWYIKIIFMIVGGLIGLRIIFAVLSIVNVRQGSPLSFQIILTPNPGGPRLGRIEEEGEQD
 KTRSIRLVNGFIALAWDDLRLCLFSYHRLRDEIILLTARGVELLGRNSLRGLQRGWEALKYLGSLVQYWGLELKKSTISLVDITAIIVAEGT
 DRIINIVQIGICRAIHNIPIREIROQGEAAQ\$

2003 CON 08 BC Env seq. opt

ATATGCGCGGTGGCGGACACCGCGCAACTACCAAGCATGCTGGATCTGGGGCGTCTGGCTTCTTGGATGCTGATCATCTGCAAGCTCGAGGG
 CCAACTGTGTGGTGTACCGGTGTATACCGGTGTGTAAGAGAGGCAACAGACACCTGTGTTCTGGCTCTGACCGCAAGGCTTAGGAGA
 CGCGAGGTGGCAACAGCTGTGGCACTCCAGCGCTCGCTGCCAGCAACCCACCGAGGATCTGTGATGGAGAACTCCAGCGAGAACTTC
 AACATGTGGAACAAACGACATGTTGAAACCAATGACACGAGAACGTGATCTCCCTTGGGACCACTCCCTTGAAGCCCTGGTGAAGCTGACCC
 CCGTGTGTGGTGAACCTGGATGTGCAACAGTCTCTCCAAAGGCAACGGCACTTCCACAGAGACTTCCACAGGATCCGTGTGAAGAGATCAAGA
 ACTGTGCTCTTCAACCGCCACCTCTGCGCGACCGAAGAAAGACCGTGTAGCCCTGTCTTACCGCTGGATCTGTGCTGCTCCCTGAACAC
 GAGAACTCCGGCAAGAACTCTCGAGATACTCCGCTGATCACTGCAACACCTCCGCAATCCACAGGCGCTCCGCCAAGTGCACCTTCGAG
 CCCCATCCCATCTCACTACTGCAACCCCGCGCTACGCCATCTCTGAATGCAACCAAGAGATTCACCGGCAACGGCCAGTGTGCCACAAG
 TGTCTGCTGCTGATGTCAACCGGCACTAAGCGTGTCCACAGCTGTCTGTAAAGCTCTCCGTGGCGAGCGAGATCAATCATC
 CGCTCCGAGAACTCTACCAACAACTGTGAGAGACCATCATCTGTCACTTGAACCACTCCGTGTGGAATCTGTGACCCGCCCAACCAACAC
 CCGGAAAGTTCATCTGCACTGGCGCCGGCCAGACCTCTTACGGCAACCGGAGACATCTGGCGAGATTCGCGACGACCTCGACACATCTCCA
 AGCAAGAAGTGTGAGAGACTCTCGAGCGCGGTGTCCAAAGAACTGTGCGAGACATCTCCCAACAGACCATCAAGTTCGCTCTCTCTCGCGG
 GCGCACTCTGGAGATCAACACCACTCCTTCAACTCGCGCGGAGTCTTCTTACTGCACACCTCCGGCTGTTCACGCGACGTACATATGAA
 CGGCAACCAACAACCTCTCTCCATCATCAACATCCCTCGCGCATCAAGCAGTATCACACATGTGGCAGGAGGTGGCGCCGCTCATACG
 CCGCCCCCATCTGAGGGGCAACATCACTCGAAGTCCATCATCCCGCTGTGTGTGGCGAGCGGCGGACCGAGTCCAAACACACAC
 GAGATCTTCGCGCGCGGCGACATCGCAACAACTAGCGCAACAGCTTACAGTCAAGTGTGTGATCAAGCGTCAAGCGCTCGGCGCGG
 GCGCCCAACCGCCGACGCGCGCTGTGTGGAGCGGAGAACGCGCGCTGGCGCTCGGCGCGCTGTCTTCTGGCTTCTTGGCGCGCGCGCG
 GNTTCAACATGGGCGCGCTCCATCACCTTGACCTGCAAGGCCCGCCAGTCTGTCTCGGCACTGTGAGAGACGATCCAACTGTCTGGCG
 GGCATCTGAGCGCCAGCAATCTGTGACGTCTGTGGGCACTCAAGCACTGCAAGCAACCGCTGTCTGGCACTCGAGCTCACTGTGAA
 GAGCACAGCATGCTGGCATCTGGGCTGCTCGCGAAGTGTATCGCAACCGCGCTGGCTGTGATCTCTCTGTGTCGAACACCGCTCGG
 AGCAGGAGATCTGGGACAACTGACCTGGATGAGTGGACAGAGAGATCTCCAACTACACCAACCACTTACCCCGCTGTCTGGAGGACTCC
 CAGAACAGCAGGACAGCAACGAGAGGACCTGTCTGGCCTGGATCTCTGGAGAACTCTGGTCTGGTTCGATCACTCAACACTTGGCTGTG
 GTACATCAAGATCTTCATCATGTGTGGCGGCTGTATCGGCTGGCATATCTTCCGCTGTCTGTCTGATCGTGAACCGCTCGCGCAGG
 GCTACTCCGCCCTCTCTCCAGTCTTCAACCCCAACCCCGGCGCCCCCGCTGGGCGCGCATCGAGAGGAGGCGCGGAGAGGAGAC
 AAGACCTGCTCCATTCGCTGTGAACGCTTTCTGTGGCCTTGGCGTGGAGCACTGTGGCACTCTGTGCTTCTTACACCTCGCTGGG
 CGACTTTCATCTGCTGACCGCGCGGCTGTGGCGCACTCCTCTGCGCGCTCTGACGCGGCTGGAGGCGCTCTGAGCTAAC
 TGGGCTCCCTGGTGTGCACTACTGGGGCTCTGGAGTGAAGAGTCCACCATCTCCTCTGTGTGGACCACTATCGCCATCTGGCGAGGACAC
 GACCGCATCATACATCTGTGAGGGGATCTTGGCGCGCATCCCAACATCCCGCGCGCATCGCGCGAGGCTTCGAGGCGCGCGCTCGAGTA

도면59a

2003 CON_10 CD Env
 MRVMGIQRNCQWIIWGLGFWMMLICNATGNLWVTYYGVPMKETTTLFCASDAKAYKAEAHNIWATHACVPTDPNPQEIIVLENTENF
 NWWKNGMVDQMHEDIISLWQGLKPCVKLTPLCVTINCSDVNATNSATNTVVAGMNCSENIITEIRDKKKQYALFYKLDVVOIDGSSNTSY
 RLINCNTSATQACPKVTFEPIPIHYCAPAGFAILKCNDRKFNFTGFCRKNVSTVQCTHGIKPVVSTQLLNGSLAEETIIRSENLTIDNAKT
 IIVOLNESVTINCTRENNTNRKSIIRIGPGOTFYATGDIIGNIROAYCNIISGTEWKTLOQVAKKLGDLNKTTIIFKPSSGGDPETTTHTFN
 CGGEFYCNTSKLFNSWTSNNTGNTSTIITLPCRKQIINMWQGVGKAIYAPPIAGLINCSSNTGLLLTRDGGANNSETFRPGGDMRDNW
 RSELYKYKVVKIEPLGLAFTKAKRRVVEREKRAIGLGAAGSTMGAASTLTIVQARQLSGIVQQNNLLRAIEAQHLLQLTW
 GIKQLQARVLAVESYULKDOOLLGIWCSGKHICTTNVPWNSWSNKSLEIWDNMTWMEWEIEDNYTGLIYSLIEESQOQERNEQELIQL
 DKWASLWNNFESTNNLWYIKIFIMIVGGLLIGLRIVEFAVLSLVNVRQGYSPISQTLLPAPRGDPDRPEGIEEGEGEQGGRSIRLVNGFSAL
 IWDDLRLCLFESYHRURDLILIAIRIVEVLLGRRGWEAIKYIWMNLQYWIQELKNSAISLDDTTAFAVABGTDRAIEIVQRAVAVLINIFTRI
 ROGLERALLS

도면59b

2003 CON_10_CD Env.seq.opt

ATGCGCGTATGGGCATCCAGCGCAACTGCGCAGCAGTGGTGATCTGGGGCATCTTGGGCTTCTGGATGCTGATGATCTGCAAGCGCCACCGG
CAACTGTGGGTGACGGTGTACTAGCGGTCCGCTGGGAAGAGACCAACCAACCTGTTCCTGGCTCGACGCGCAAGGCTACAAAGT
CCGAGCGCCACAACATCTGGCGCATCCGACCGCTCGTGCGCCACCAACCCCAAGAGATCTGCTGGAGAACCTGACCGAGAACTTC
AACAATGTGAAGAAAGCGGATGGTGACACGATGACAGGAGCAATCTCCCTGTGGACAGGGCTGAAGCGTCTGCTGGAGCTGACCTC
CTCTGGTGACCTGAACTGCTCGACGTCAGCGCCAACTCCGCAACCAACCGTGTGGCGCGCATGAAGAAATGCGTCTTCAACA
TACCAACCGAGATCCGGCAAGGAGGAGGAGTACGCCCTGTTCTACAGCTGGACCTGGTGCAGATCAGGGGCTCCAAACCACTCTCTAC
CGCTGTGATCAACTGCAACACTCCGCCATACCCAGGCTGCCCAAGTGAOCTTCGACCATCCCATCACTGCGCGCCCGCGG
CTTGCGCATCTGGAAGTGCAGCAAGAGATCAACGCGCACCGGCTGTCCGAAGAAGTGTCCACCTGGCACTGCACCCAGCGCATCAAG
CCCTGTGTGTCACCGAGTGTCTGTGAACGGCTCCCTGGCCGAGGAGAGATCATATCCCTCGCGAACAATGACCGCAACCGCAAGAC
ATATATCTGACAGTGAACAGTCCGTGAATCAATCACTCAACCGCCCAACAACAACCGCAAGTCCATCCGATCGCCGCCGACAGAC
CTTCTACGCCACCGGACATCTGGCAACATCCCGCAGGCTACTGCAACATCTCCGCGACCGAGTGGACGAACCTCTGCACAGGTGG
CCAGAAGCTGGGCGACTCTGHAACAGACCACTCATCTTCAAGCCCTCTCCGGGCGGACACCGGAGTCAACACCCACCACTTCAAC
TGGGCGGGAGTCTTCTACTGCAACACTCCAAAGTCTTCACTCTCTCTGGACCTTCAACACACCGGACCGGACCGATTCACCCCT
GCCCTCGGCATCAAGCATCATCAACATCGAGGCGTGGGCAAGCCATACGCCCCCCCATCGCGGCCGTATCAACTGCTGCTCT
CCACATCAACCGGCTGCTGTCACCGGACGGCGGCGCCACAATCTCCGACAGCTTCCGCCCGCGCGCGCATGGCGACCAACTCG
CGTCCGAGCTGTACAGATPACAGTGTGAAGATCGAGCCCTGGGCTGGCCCAACCGCGCGTGGCGCGCGCGCATGGCGACAGAGAA
GGCGGCCATCGGCTGGGCGCGTCTCTGGGCTTCTCGGCGCGCGCGGCTCACCATGGGCGCGGCTCCCTGACCTGACCGTGCAGG
CCCGCAGCTGTCTGGCATCTGTGGACGACGAGACAACCTGCTGGCGCATCGAGGCCATCTGCGACCTTCTGCACTTCCGCTGTGG
GGATCAAGAGCTGTGAGGCGCGGTGGCGCTGGATCTTCACTGAAGAGCAGCAGCTGCTGGGACATCTGGGCTGCTCCCGCAAGG
CATCTGCAACCAACCGCTCCCTGGGATCTCTCTGTGCCAACAGTCCCTGGAGGAGATCTGGGAACAATGCACTGGATGGAGTGGGAGC
GCGAGATGGACACTACACCGGCTGATCTACTCTGATCGAGGAGTCTGCAACCAAGAGAGACGAGGAGCTGTGTCGACGTCG
GACAAATGGGCGCTCTCTGGAACTGGTGTTCATCAACCAAGCTGGCTGGATCTGGATCAAGATCTTCACTGATGTGGGCGGCGCTCATCGG
CTTTCGCACTGTGTTCCGCTGCTCTCTGTGTGAACCGCGCTGCGCAGGGCTACTCCCCCTGTCTTCCAGAACCTTGTGCGCGCGCGC
GGCGGCCGACCGCGCCCGCAACCTCGAGGAGAGGGCGCGGCGGCGGCGCTGATCCGCTGTGGTGAACGCTGTTCGCGCTG
ATCTGGGACATCTCGCAACCTCTGGCTGTCTTCAACACCGCTCGCGACCTGATCTGATCGCGCTCACCAGCTGTGGAGCTGTGCG
CCGCGCGGCTGGAGGCGCATCAAGTACATGTGGAACCTGTGCACTGACTGATCCAGGAGCTGAACAACTCCGCCATCTCCCTGCTGGACA
CCAGCGGCATCGCGTGGCGGAGGACCGACCGCGGCATCGAGATCTGTGCGAGCGCGCTGTGGAACTCTCGTGAACATCCGACCGCGCATC
CGCCAGGCGCTGGAGCGGCGCTGCTGTA

도면60a

2003 CON 11 CPX Env
 MRVKETORNWHNLWRGLMI FGLMLICNATENLWTVVYGVVPVKDADTTILFCASDAKAYSTERHNVWATHACYPTDPNPQOEI PLENVTENF
 NMWKNNVVEQMHEDIISLWDESLKPCVKILTFELCVTLNCTDVKNATNTTVEAAEIKNCSENIITTEIKDKKKKEYALEFYKLDVVPINDNNSIY
 RLINCNVSTVKQACPKVTFEIPFIHYCAPAGFALIKCNDKENGTPCNRVSTVQCTHGKFPVSTQLLNGSLAEGEVIRSENFTWNAKT
 ITIVQLNSSVRINCTRPNNTRKSIHIGPGQAFYATGDIIGDIROAHNCNISRAEWNTLQOVAKOLRENFNKTIIFNPNSSGGDLIITTHSFNC
 GGEFFYCNTSRLNSTNNNDTNDTKQMHITLPCRISOIYNNMORVQAMVAPPIQGIKCNSTITGLLITROGNNNTNETERTGGDMRD
 NWRSELYIKYVVEIKELGVAETRAKRVREREKRAVGIGAVLLGELGAAGSTMGASITITVQAKQLLSGIVQOQSNNLLKAIETAOQHLLKLT
 VMGIKQOARVIAVERYLKDQQLLGIWGCCKLICTNVNPNWFSWSNKSYSDEIWNMTWIEWEREINNYTQTIYTLLESQNOQKEQQLL
 ALDKWASLWNVFDISNWLWYIKIFIMIVGSLGLRIIFAVLSIVNRCROGYSPLSFOTLTPNHKEADRPCGIEGGGEQDRTRIRLVSQFL
 ALAWDDLRLNCLFSYHRLRDEILIAARIIVETLGRGWEILKYLGNLAQYWGQELKNSAISLLNATAIAVAEGTDRIIEVVHRVLRAILHIPR
 RIRQGFERALLS

도면60b

2003 CON 11 CPX Env. seq. opt.

ATGCCGTGAAGAGACCCAGCGCAACTGGCAACAACCTGTGGCGTGGCGCTGATGATTTGGCATCTGATCTGCAACGCCACCGA
 GAACCTGTGGTGACCTGTACTAGCGGTGCCGTGTGGAAGGACGCCGACACCAACCTGTTTGTGGCTCCGACGCCAAGGCTTACTCCA
 CCAGAAGCACAACCTGTGGGCCACCCACGCTGGTGCACACCCAGGAGATCCCTGTGGAGATCCCTGTGGAGATCCCTGTGGAGATCC
 AACTGTGGAAGAACAACTGTGGGAGATGACGAGGAGATCATCTCTGTGGGAGAGTCCCTGAAGCCCTGCTGAAGCTGACCC
 CCTGTGCTGACCTGAACGACCGACCTGAAGACGCGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCG
 CCACCGAGATCAAGGACAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 CGCTGATCAACTGCAACGCTGTCCACCGTGAAGGAGCGCTGCCCAAGGTGACCTTCGAGCCCATCCCATCTACTGCGGCCCGCGCG
 CTTGCGCATCTGAGTGAAGTGAACGACAGAGATTCAACGGGACCGGCCCTGCAAGAGCTGTCCAGGAGCTTCCAGGAGCTCAAGC
 ATCATCTGTCAGCTGAACTCTCTGCTGCGATCAACTGACCGCCCGACCAACAAACACCCCGAAGTCCATCCACATCGGCCCGCGCGG
 CTTCTACGGCACCGCGACATCATCGGCGACATCGGCGAGCTGCGCGAGCGCTGCAACATCTCCGCGGAGTGGAAACACACCTGAGCAGGTG
 CCAAGCAGCTGCGCGAGACTTCAACAGACCATCATCTCAACACCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT
 GGCGCGAGTTCTTACTGCAACACCTCCCGCTGTTCAACTCCACTGGAACACACACACACACACACACACACACACACACACACACACAC
 CCTGCCCTGCCGATCAAGCAGATCGTGAACATGTGGAGCGGTGGCGAGCGCTGACCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 ACTCCAATCACCGGCTGTGTGACCGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
 AACTGGGCTCCGAGCTGTAAAGTACAAGGTGAGTCAAGCTTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG
 CGAAGGCGCGCTGGCATCGCGCGCTGCTGGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 TGAGCGCGCGAGCTGCTGTCGCGCATGTGAGCAGCATGCTGAGCAGCATGCTGAGCAGCATGCTGAGCAGCATGCTGAGCAGCATGCTGAGC
 GTGTGGGCATCAAGCAGCTGAGCGCGCGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 CAAGCTGATCTGCACCAACACGTCGCGCTGGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 GGGAGCGGAGATCAACAACTACCCAGACCTGCTGCACTCTCCAGTGGCTGTGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT
 GCCTGGACAAGTGGGCTCCCTGTGGAACTGTTCCGATCTCCAGTGGCTGTGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT
 GATCGGCTGGCATCATCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 ACCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 GCGCTGGCTGGGACGACTGCGCAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 CCTGGGCGCGCGCGCTGGGAGATCTGAGTACCTGGGCAACCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 TGAAGCGCACCGCATGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
 CGCATCGCGCGCGCGCTTGAAGCGCGCGCGCTGCTGTA

도면61a

2003 CON 12 BF Env

MRVRGMORRWOHLGKWLLEFLGILLICNATENLWTVVYGVVWKEATTTLFCASDAKSYEREVHNWATHACVPTDPNPQEVLENVVTENF
 DMWKNMVEQMHTDILSLWDSLRKPCVKLTPLCVTLNCTDANATANATKEHPEGRAGAIQNCSEFNMTTEYRDKQMKVQALFYRLDIVFTSON
 NSNEYRLINCNTSTITQACPKVSWDEPIPIHYCAPAGYAILKCNDRKFNGTGPCKNVSTVQCTHGJKPKVYSTQLLNGSLAEELIIRSQNIS
 DNAKTIIVHLNESVQINCTRNNNTKSIHIGPGRAFYATCDIIGDIRKAHCNVSGTQWNKTLQVKKLRSEYFNTTIKFNSSGGDPFITM
 HSFNCRGEFFYCNTSKLENDIVSNDTIILPCRIQIWNMQEVEGRANYAAP IAGNITCTSNITGLLTRDGGHNETNKTETFRPGGNNMKN
 WRSELYKYKVVEIEPLGVAPTRAKQVVKREKRAVGIGALFLGFLGAGSTMGAASITLTQARQLISGIVQQOSNLLRAIEAQHLLQLTY
 WGIKQIQARVLAVERYLKDOQLGLWCCSKLICITNVPWNSSWSNKSQEEIWNMTWMEWEKEINNYSEIYRLIEESQOQKEKNEQELLA
 LDKWASLWNWFDISNWLWYIRIFIMIVGGLIGLRIVFAVLSIVNRVKGYSPLSQTHIPSPREPDRPEGIEGGEGQKDRSVRLVNGFLA
 LIWDDLRSLCLFSYHRLRDLILLIVTRIVELIGRRGWEVLKYWNLLQYWSQELKNSAISLINTTAIVAEGTDRVIEALQRVGRAILNIPRR
 IRQGLERALLS

도면61b

2003 CON_12_BF Env.seq.opt

[illegible]

도면62a

2003 CON 14 BG Env

MKAKGTORNQSLWKWGTLILGLVLICASNDLWVTYYGYVPVWKEATTILFCASDAKAYDAEVHNVWATHACVPTDPNPQOEVALENVTENF
 NWENNMYDQMCEDILSLWDQSLKPCVELTLCVTILNCTOFNNTNNTNNTNNDGGEIKNCSENIITSLRDKTKKEYALFYNLDPVQMDND
 NSSYRLTSCNTSIIITQACPKVSFTPIPIHYCAPAGFVILKCNKKTNGTGPCTNVTQCTHGRPVVSTQLLNGSLAFEEIIVIRSKNFTD
 NAKTIIVOLKDPRIEINCTRNPNNTNRKRITWGPGRVLYTTGOIIGDIRKAHCNI SKTKWNNTLQIIVKKLREQFNKTIIVQRRSSGGDPPIVM
 HSFNCGGEFFYCNTTQLENSTWRSNSTWNTTETNNTDILTLPCKRIKQIYVMWQKVGKAMYAPPISSQIRCSSNITGLLLIRDGGSNNTETTF
 RFGGNNKDNWRSELYKYVVKIEPLGVAPTRAKRRVQREKRAYGIGALIFGELGAGSTMGAASMTLTVOARQLLSGIVQQQNNLLRAIE
 AQQHMLQLTVMGIKOLOARYLAVERYIKDQQLLGIWGCSEKLICTTTPVWNASNKSLLDIWNNTWMEWEREDNYTGLIYTLIEQSQNO
 QERNCELELDKWAASLWNWFNITNLWYIKIFIMLIGGLIGLRIVFAVLSIINRVKGYSPLSFQTLTHHOREPDRPGRIEEGEGEQDKDR
 SIRLVSGFLALAWDDLRSLCLFSYHRLRDFTLIAARTVELLGRSSILKGLRLGWGELKYLWNLLLYWGREIKNSAINLLDPTVAIAVANWTDRA
 IEVVQVRGRAVINIPVRIRQGLERALL\$

2003 CON_14_BG Env.seq.opt

[illegible]

도면63a

1. 2003 CON S gag.pgp
 MGAASVLSGGKLDWEKIELRPGGKKKYRLKHLVWASRELERFALNPGELLETSGCCQIIEQLQALQTSSEELRSLYNTVATLYCVHQRI
 EVKDTKEALDKIEEQNKSKQKTQQAADTGNSSKVSQNYPIVQNLQGMVHQALSPRTINAWKVVEKAFSEFVIMFSAISEGATPQDL
 NTMLNTVGGHQAAMQMLKDTINEEAABWDLHPVHAGPIPPGOMREPRGSDIAGTISTLQEQIGWMTSNPPPIPVGEIYKRWIILGLNKIVRM
 YSPVSLDILRQGPKEPFRRDYVDRFEKTLRAEQATQDVKNWMTDTLLVQMANPDCKTILKALGPGATLEEMNTACQGVGQFSSHKARVLAAPMS
 QVTNTTIMQRGNFKGQKRIIKCFNCGKEGHIARNCRAPRKGCWKCGKEGHQWKDCTERQANFLGKIWPSNKRPGNFLQSRPEPTAPPAE
 SFGFGEIITPSPKQEPKDKELYPLASLKSIFGNDPLSOS

도면63b

2003 CON S gag.OPT
 ATGGGGCCCGCGCCTCCGTGCTGTCCGGGGCAAGCTGACGCTGGAGAACGATCCGCTGCGCCCCCGGGCAAGAGAGTACCGCCT
 GAAGCAGCTGTGTGTGGCCCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGGCTGCGAGAGATCATCG
 AGCAGCTGCAGCCGCCCTGCAGACCGGCTCCGAGAGCTGCGTCCCTTACAACACCGTGGCCACCCCTGTACTGCGTGCAACAGCGCATC
 GAGGTGAAGGACACCAAGAGGCCCTTGGACAAATCGAGAGGACAGACAGTCCAGCAGAGACCCAGCAGGCCCGCGCGACACCGG
 CAATCCTCCAGGTGTCCCGAATACCCCATCGTGCAAGACCTGCAAGGCCAGATGGTGACCAAGGCCATCTCCCGCGCACCCCTGAACG
 CCTGGGTGAAGTGTGGAGAGAAGCCCTTCTCCCGAGGTGATCCCATGTTCTCCGCCCTGTCGAGGGCGCCACCCCGCAGGACCTG
 AACACATGCTGAACACCGTGGGGGCCAACAGGCCGCGCATGCTGAAGACACCATCAACGAGGAGCGCCGAGTGGGACCGCCT
 GCACCGGTGCACGCCGCGCCCATCCCGCGCGCAGATGCGCGAGCCCGCGCTCCGACATCGCCGACACCATCTCCACCTGCGAGG
 AGATCGGCTGGATGACTCCAAACCCCGCATCCCGTGGCGAGATCTACAAGCGCTGGATCATCTGGGCTTGAACAGATCGTGCGCATG
 TACTCCCGGTGTCCATCTTGACATCCGCGAGGCCCCAGGAGCCCTTCGCGACTAGTGGACCGCTTCTTCAAGACCTTCGCGCGCGA
 GCAGCCACCCAGGACGTGAAGAACTGGATGACCGACACCTGTGTGTGAGAACCGCAACCCGACTCGAAGACCATCTGGAAGCGCGTGG
 GCCCGCGCGCACCCGTGGAGGAGATCATGACCGCTGCCAGGGCGTGGCGGCCCTCCACAGCCCGCGCTGTGGCCGAGGCCCATGTCC
 CAGGTGACCAACACCACTCATGATGCAGCGCGGCACTTCAAGGCCAGAGCGCATCATAGTGTCTTAACTGCGGCAAGGAGGCCA
 CATCGCCGCACTGCGCGCCCCCGCAAGAGGGCTGTGGAATGCGGCAAGGAGGCCCAAGATGAGGACTGACCCGAGCGCCAGG
 CCAACTTCCTGGGCAAGATCTGGCCCTCCAAACAGGGCGCCCCCGGCACTTCTGCGCTCCCGCCCCGAGCCACCGCCCCCGCGGAG
 TCCTTCGGCTTCGGCGAGGAGATCACCCCTCCCGCAAGCAGGAGCCCAAGGACAAGGAGCTGTACCCCGCTGGCTCCCTGAAGTCCCTGTT
 CGGCAACGACCCCTGTCCAGTAA

도면64a

2. 2003 M.GROUP.anc gag.pgp

MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLEETAGCCQIMGQLQPALQTGTEELRSLYNTVATLYCVHORI
 EVKDTREALDKIEEQNKSSQQTQAAADKGDSSQVSPYIVQNLQGMVHQAIAPRTLNAWKVEKAFSPYIPMFSALSEGATPQDL
 NTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIGHWTSNPPPIPVGEIYKFWIILGLNKIVRM
 YSPVSLIDIRQGPKEPRDYVDRFFTKILRAEQATQDVKNWNTDTLLVQANPBCKTILKALGPGATLEENMTACQGVGGPGHKARVLAEAMS
 QVTNANIMQKGNFKGPRRIYKCFNCGKEGHIARNCRAPRKKGWKCGKEGHQMKDCTERQANFLGKIWFSNKGPRGNFLQSRPEPTAPPAE
 SFGFGHEITPSPKQEPKDKELYPLASLKSLSFGSDPLSOS

도면64b

2003_M.GROUP.anc gag.OPT

ATGGCGCGCGCGCTCCGCTGCTGCGCGGCGAGCTGGACGCGCTGGAGAAAGATCCGCTTGGCCCCCGGCGCAAGAAGATACCGCCT
GGAACACTGTGTGGCGCTCCGGAGGTGGAGGCTCCGCTTGAACCCCGCGCTGTGGAGACCGCGAGGCTGCGGACGASATCATGG
GCACACTGTGACCGCGCTTCAGACGGGACGAGGTTCCGCTCTGTCAACACCGTGGACCACTGTACTGCTGGTCCACGACGATCG
GGCTTGAGGACACCGAAGGCGCTTGAACAAGATCGAGGAGGACAGACAGTCCACAGAGAACCACGACGCGCGCGCGACAGGG
CGACTCTCCCAAGTGTCCAGACTACCCATGTGTCAGAACTGTGAGGGCCAGATGTGCACGAGGCACTTCCCCCGACCTCTGAACG
CTGGGTGAAGGTGTGGAGGAGAAAGCTTCTCCCGAGAGTATCCCATGTTCGCGCTGTTCGAGGGCGCCACCCCGCCAGGACCTG
GAACCCGCTGAAACACTGTGGCGGCCACGAGCGCATGAGATGCTGAAGACACGATCAACAGGAGGCGCGCGAGTGGGACCGCCCT
GACCCGCTGACGCGCGGCCATTCGCCCGGACGATGCGGAGCCCGCGGCTTCGACATTCGCGGCGCACCTTCACTCCCTGAGGAGC
AGATGGCTGGATGACCTCCACCCCCCTCCCGTGGGCGGACATCAAGGCTGCGATCATCTGGCGCTGAACAGATCGTGCGCAT
TACTTCCCGCTGCTCATCTGCTGATCATCCGACGCGCCCAAGGACGCTTCGCGACTACCTGGACCGCTTTCACAGACCTTCGCGGCCGA
GAGCGACCCGACGACTCAAGAACTGGATGACGACACCACTGCTGTGTCGAGACGCCACCCGACATCTCTGAAGACCACTCTGAGAGCCCTTCG
GCCCGGCGCCACCTGGAGGAGATGATACCGCTGCCAGGGCGTGGCGGCGCCCGGCGCCACAGGCCCGCTGTGCGCGAGGCTATTC
CAGTGACCAACGCTACATCATGATGACGGCGCAACTTCAGGGCCCGCGCGATCTCAAGTGTCTCAACTCGCGCAAGGAGGCGCA
CATCGCGCAACTCGCGCGCCCGCGAAGAGGGCTCTGGAAGTCGGGCAAGGCGGCCGACCAAGATGAGGACTGCACGAGCGCCGAG
CGAACTTCCCTGGGCAAGATCTGGGCTCTCAACAGGCGGCCCGGCAACTTCGTGAGTCCGCGCGAGCCACCGCCCCCGCGCGAG
TCCCTGGCTTCGGCGAGGAGATCACCCCTCCCGCCACGAGGACCCAGGACGAGGAGCTTACCCCTTGGCTTCCGCTGAAGTCCGCTGT
CGGTCCGACCCCTGTCCGAGTA

도면65a

3. 2003 CON A1 gag.PEP
 MGARASVLSGCKLDAREKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETTEGCOQINEQLQPALKTGTEIRSLYNTVATLYCVHORI
 DVKDTREALDKIEEQNKSKQKTQAAAADIGNSSKVSQNYPIVQNAQGMVHQSLSPRTLNANWKVIEEKAFSEVIFPMFSALSEGATFQDL
 NMMLNTVGGHOAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMEPRGSDIAGTSTTPOEQIGWMTGNPPIPVGDIYKRWIILGLNKIVRM
 YSPVSLDIKQGPKEFRDYVDRFFKILRAEQATQEVKNWMTETLLVQANPDCKSILRALGPGATLEEMMTAQQGVGGPGHKARVLAEAMS
 QVQHTNIMQQRGNFRGQKRKICFNCGKEGHLARNCRAPRRKKGCWKCGKEGHQMKDCTERQANFLGKIWESSKGRFGNFPQSRPEPTAPPAEI
 FGMGEITSPFKQEQKDRQDPPFIVSLKSLFGNDPLSQ\$

도면65b

3. 2003 CON A1 gag.OPT

ATGGGCGCGGGCTCCGTGCTGTCCGGCGGCAAGCTGGACCGCTGGGAGAAATCCGCTGGCGCCCGCGGCAAGAGAAATACCGCT
 GAAGACCTGGTGTGGGCTCCGCGAGCTGGAGCGCTTCGCGCTGAACCCCTCCCTGCTGGAGACCCACCGAGGGCTGCCAGCAGATCATGG
 AGCAGCTGCAGCCCGCCCTGAAGACCGGCAACCGAGGAGCTGCGCTCCCTGTACACACACCGTGGCCACCTGTACTGCGTGCAACGAGCGCATC
 GACGTGAAGACACCAAGAGGCGCTGGACAGATCCAGAGATCCAGAACAAATCCAGCAGAGACCCAGAGCGCCCGCGGACACCGG
 CAATCTCCAGGTGCCAGACTACCCCATGCTGCAGAACGCCAGGCGCAGATGGTGACCAAGTCCCTGTCCCGCCGACCCCTGAACG
 CCTGGGTGAAGGTGATCGAGGAGAGGCTTCTCCCGAGGTGATCCCATGTTCTCCGCGCTGTCCGAGGCGCCACCCCGAGGACCTT
 AACATGATGCTGAACATCGTGGGCGGCAACAGGCGCATGATGAGGACACCAATCAACGAGGAGCGCCGAGTGGGACCGCTT
 GCACCCGCTGACCGCGGCCCATCCCGCGCAGATGCGGAGCCCGCGCTCCGACATCGCCGACACCACTCCACCCCGCAGAGC
 AGATCGGTGGATGACCGCAACCCCGCAICCCCGTGGCGACATCTACAGCGTGGATCATCTGGGCTGAACAGATCGTGGCATG
 TACTCCCGCTGTCCATCGTGGATCATGACGAGGCGCCAGAGGCGCCAGAGGCGCTTCCGCGACTACGTGACCGCTTCTTCAAGACCTGCGGCGCA
 GCAGGCCACCCAGAGGTGAAGACTGGATGACCGAGACCTGCTGTGTCAGAACGCCAACCCCGACTGCAATCTCGCGCGCTGG
 CAGGTGACACACCAATCATGATGATGAGCGGCAACTTCCGCGGCGAAGCGCATCAAGTGTCACTGCGGCAAGGAGGCGCACCT
 GGCCCGCAACTGCGCGGCGCCCGCAAGAGGCTGTGGAATGCGGCAAGAGGCGCACAGATGAGGACTGCACCGAGGCGCGCA
 ACTTCTGGGCAAGATCTGGCCCTCTCCAGGGCGCGCCCGGCAACTCCCGAGTCCCGCGCGACCCCGCCCGCGCGAGATC
 TTCGGCATGGGCGAGGAGATCACTCCCGCCCGCAAGCAGAGAGAGACCGCGAGCAGGACCCCGCTGGTGTCCCTGAGTCCCTGTT
 CGGCAACGACCCCTGTCCAGTAA

도면65c

4. 2003_A1.anc gag.PEP

MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETAEGCQQTMGQLQPALKTGTEELRSLYNTVATLYCVHQRI
 EVKDTKEALDKIEETQNKSKOKTQQAADTGNSSKVSQNYPIVQNAQGMVHQSLSPRTINAWVKVIEKAFSEVIFPFSALSEGATQODL
 NMMLNTVGGHQAAMQLKDTINEEAAEWDRLHPVHAGPIPPQMRPRGSDIAGTISTLQEQIGMTGNPPIPVGDIYKRWIILGLNKIVRM
 YSPVSLDIRQCPKEPFRDYYDREFKTLRAEQATQEVKNWMTETLLVQNAHPDCSILRALGPGATILEMNTACQGVGGPGHKARVLAEAMS
 QVQNTDIMMRGNFRGPKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKINPSSKGRPGNFTQSRPEPTAPPAEN
 FGMGEWISSPKQEQKDRQYPLVSLKSLFGNDPLSQ\$

도면65d

2003_A1.anc gag.OPT

ATGGGCGCCCGGCGCTCCGTGCTCCGGCGGCAAGCTGGACGCCCTGGGAGAGATCCGCCTCGGCCCGCGGCAAGAGTAACCGCCT
 GAAGCACCTGCTGTGGCGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTCTGGAGACCCCGAGGCTGCGAGCGAGATCATGG
 GCCAGTGCAGCCCGCCTGAAGACCGGCAACGAGAGCTGCGCTCCCTGTACACACCTGGCACCCTGTACTGCTGCACCCAGCGGATC
 GAGGTGAAGGACACCAAGGAGGCCCTGGACAGATCGAGGAGATCCAGACAAGTCCAGAGAACCCAGAGGCCCGCGGACACCGG
 CAACTCCTCCAGGTGTCCAGAACTACCCATCTGTCAGAACGCCAGGCCAGATGGTGCACCAATGCTCCCTGTCCTCCCGGCACTGACG
 CCTGGGTAAAGTGATCGAGGAGAGGCTTCTCCCGGAGGTGATCCCGATGTTCTCCGCCCTGTCCGAGGGGCCACCCCGAGGACCTG
 AACATGATGCTGAACATCTGTGGCGGCCACCGAGCGCCATGTCAGATGCTGAAGGACACCATCAAGAGAGGCCCGAGTGGGACCGCCT
 GCACCCGTTGCACGCCGCCCATCCCCCGGCCAGATGCCGAGCCCGGCTCCGATCCCGGACCATCAAGCGCTGGATCATCCTGGGCTGAACAAGATCGTGGCATG
 AGATCGGCTGGATGACCGCAACCCCGCATCCCGTGGCGACATCTACAAGCGTGGATCATCCTGGGCTGAACAAGATCGTGGCATG
 TACTCCCGGTCTCCATCTGGACATCCGCCAGGGCCCAAGGAGCCCTTCCGCCACTAGTGGACCGCTTCTCAAGACCCCTGGCGCCGA
 GCAGGCCACCGAGGTGAAGACTGGATGACCGAGACCTGTGGTGAAGACGCCCAACCCCGCTGCAAGTCCATCGTGGCATG
 GCCCGCGCCACCCCTGGAGGAGATGATGACCGCTGCCAGGCGTGGGCGGCCGCCACAGCCCGCTGCTGGCGGAGCCATGTCC
 CAGGTGAGAAACACCGACATCATGATGACGCGCGCACTTCCGGGCGCCCAAGGCAATCAAGTCTTCACTGGGCAAGAGGGCCACCT
 GCCCGCAACTGCGGCGCCCGGCAAGAGGCTCTGGAAGTGGGCAAGGAGGCCACCAAGTGAAGGACTGCACGAGCGCCAGGCCA
 ACTTCCTGGGCAAGATCTGGCCCTCCTCAAGGGCGCCCGGCACTTCCCGAGTCCCGCGAGCCACCGCCCCCGCCCGGAGAAC
 TTCGGCATGGCGAGGAGATGATCTCCTCCCCCAAGCAGGAGCAGAGGCCGAGAGTACCCCCCTGGTGTCCCTGAAGTCCCTGT
 CGGCAGGACCCCTGTCTCCAGTAA

도면66a

3. 2003 CON A1 gag.PEP
 MGARASVLSGCKLDAREKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETTEGCOQINEQLQPALKTGTEIRSLYNTVATLYCVHORI
 DVKDTREALDKIEEQNKSKQKTQAAAADIGNSSKVSQNYPIVQNAQGMVHQSLSPRTLNANWKVIEEKAFSEVIFPMFSALSEGATFQDL
 NMMLNTVGGHOAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMEPRGSDIAGTSTTPOEQIGWMTGNPPIPVGDIYKRWIIILGNKIIVRM
 YSPVSLDIKQGPKEFRDYVDRFFKILRAEQATQEVKNWMTETLLVQANPDCKSILRALGPGATLEMMTACQGVGGPGHKARVLAEMS
 QVQHTNIMQQRGNFRGQKRKICFNCGKEGHLARNCRAPRRKKGCWKCGKEGHQMKDCTERQANFLGKIWESSKGRFGNFPQSRPEPTAPPAEI
 FGMGEITSPFKQEQKDRQDPPFIVSLKSLFGNDPLSQ\$

도면 66b

2003_CON_A2 gag.OPT

ATGGCGCGCGCGCTCCATCTCTCCGGCGGCAAGCTGCAAGCCTGCGAAGAATCCGCTGGGCCCGCGCGCAAGAGAAGTACCGCCCT
GAACACCTTGGTGTGGGCTCCCGGAGAGTGGGAAGTTCTCAATCAACCCCTCCCTGCGAGGCTGCGCGAGATCATCC
GACCTTCGACCGCGCGCTCGACACGGCAACCGAGAGTGAATCTGTTACCAACCGTGGCGCTGCTGTACTCGTGCACACGCGCATC
GACCTTGAAGCAACCAAGGAGCGCTTGACAAAGATTCGAGGAGGACAGACAATGTCGACAGAGAAGACCCAGCACGCGCGCGACACGG
CACTCTCTCTCTCTCCAGATACCCCATCTGTGAGAACGCGCGGCGAGTGTGTGCACCGGCGCATCTCCCGCGGACCTCGCTGAACG
CTTGGTGAAGTGGTGGAGGAGAGCGTTCTCCCGAGGTACCCCATCTCAACCGCTGTGCGAGGCGCCACCCCGCGAGCTCTGAAAC
GAACCTGTGACACCGCTGGCGCGCACCAACGCGCGCATGAGATGTTGAAGCAACCATCAACAGGAGGCGCGCGAGTGGGACCGCCCT
GACCCCGTGCACCGCGCGCCCATCCCCCGCGAGTGGCGAGCCCGCGCTTCGACATGCGCGGCAACCATCTCCACCTTCGAGGATC
AGATCGGCTGGATGAATTCACACCCCGCATCCCCGCTGGGCGAGATCTACAGGCTCGCATCTCTGGGCTGTGAACAGATGTGGCGCAT
TACTCCCGCTGCTCATCTTGACATCCGACGCGCCACAGAGCCTCTCGGACATCTGGACCGCTCTTCAAGACCTGTGGGCGCGCA
GAGCCGACCCAGAGGTGAAGATCTGATGACGACACCTCTGTGTCGACAGCCCAACCGCGACTTCATCTCGCGCGCTTGG
GCCCGGCGCCACCTCGAGGAGATGTAACCGCTTCGAGGGCTGGGCGGCCCTCCCAAGCGCGGCTGCTGCGCGCGCTTGC
CAGTGGAAACACCAACACCAATCTGATGACGAGCGCTGCTGAGCGCGCAAGAGGCGCACCAAGTGTCAACTGTGGCAAGGAGGG
CCACTGGCGCGCAATCGCGCGCCCGCGCAAGAGGCTGTGAAGTGGGCAAGGAGGCGCACCAAGTGTCAACTGTGGCAAGGAGGG
GAGGCAATTCCTGGGGAAGATGTGGGCTCCACACAAGGGCGCCCGCGCACTTCGCGCAGTTCGCGACCGAGCCACCGCGCGCGCC
CGATCTTGGCATGGGCGAGGATACCTCTCTCTTGAAGAGGAGCTGAAGACCGCGAGCCCTACAAACCGCGCGCATCTCTCGAAGTC
CTGTTGCGGACGAGCGCGCTGTCCCCAGTAA

도면67a

6. 2003 CON B gag-PEP
 MGRASVLSGGLDRWEKIRLRPGGKKKYLKHIYVASRLELRFVAVNPGELLETSBGCRLQLQPSLQTSGLRSLNTVATLYCVHQRI
 EVKDTKEALEKIEEQNKSKKQAADTGNSSQVSNYPVQNLQGMVHQALSPRTINAWKVVEKAFSEVVPNFSALSEGATQDL
 NTMLNTVGGHOAAMQMLKETINEEAAEWDRLLHPVHAGPIAPGQMRPRGSDIAGTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
 YSPTSILDIRQGPKEPFRDYVDFYKTLRAEQASQEVKNWMTETILIVQNPDPCKTILKALGPATLEEMMTACQGVGGPGCHKARVLALAMS
 QVTNSATIMQRGNEFNQRKIVKCNCGKEGHIARNCRAPRRKGOMKCGEKGQNKDCTERQANFLGKIWP SHKGRPGNFLQSRPEPTAPPE
 RSRFRGFETTTTPSOKORPTDKET.YPIASS

도면67b

2003 CON B gag, OPT

ATGGGCGCCCGGGCCTCCGTGCTGTCGCGCGGCGAGCTGACCCGCTGGGAGAGATCCGCTGCGCCCCCGCGGCGGCAAGAGAGATACAGCT
GAAGACATCGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCGTGAACCCCGGCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCCTGG
GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTCGCTCCCTGTACACACCCGTGGCCACCTGTATGCGTGCACCCAGCGCATC
GAGTGAAGGACACCAAGAGGCCCTGGAGAGATCGAGGAGAGAGACAAGTCCAAGAAGAGGCCAGAGCGCGCGCCGACACCGG
CAACTCCTCCAGGTGTCAGAACTACCCATCGTGCAGAACCTGCAGGCCAGATGGTGCACAGGCCATCTCCCCCGCACCCCTGACG
CCTGGGTGAAGTGGTGGAGAGAGAGCCCTTCTCCCGGAGTGTATCCCATGTTCTCCGCTGTCGAGGGGCGCCACCCCGAGGACCTG
AACACATGCTGAACACCGTGGCGGCCACAGGCCGCTATGCAATGCTGAAGAGACCATCAAGAGAGGCCCGCGAGTGGGACCGCCT
GCACCCGTGCACGCGGCCCATCCGCCCGGCCAGATGGCGAGCCCGCGGCTCCGACATCCCGGACCCACCTCCACCTCCAGGAGC
AGATCGCTGGATGCCAACACCCGCCATCCCGTGGGAGATCTACAAGCGCTGGATCACTCTGGGCTGATCAAGATCGTGGCATG
TACTCCCCACCTCCATCTGGACATCCGCGAGGCCCGAGAGCCCTTCGCGACTACGTGACCGCTTCTACAGACCTTGGCGCGGA
GCAGGCTCCAGAGGTGAGAACTGGATGACCGAGACCTGCTGGTGAAGAGCCAGCCCGGACTGAGAGACCATCTGAAGGCCCTGG
GCCCGCGCCACCTGGAGAGATGATGACCGCTGCCAGGCCCTGGGCGGCCCGGCCACAGSCCGCTGCTGGCGGAGGCCATGTCC
CAGTGACCCAACTCCGCCACCATGATGAGCGCGGCACTTCGCAACAGCGAGACCGGTGAAGTCAACTGCGGCAAGAGGG
CCACATCGCCAAAGACTGCGCGCCCGCCGCAAGAGGGCTGCTGAAGTCCGCAAGGAGGCCACCAATGAGGATGCAAGGAGGGCC
AGGCCAACTTCCTGGGCAAGATCTGGCCCTCCACAGGGCGGCCCGCGCACTTCCTGCACTCCCGCGCGGAGGCCACCGCCCCCGGAG
GAGTCCTTCCGCTTCGGCGAGGAGACCAACCCCTCCAGAGAGAGAGGCCCATCGAAGAGAGCTGTACCCCTGGCCTCCTAA

도면67c

7. 2003 B. and gag. pep

MGARASVLSGGKLDKWEKIRLPGCKKKYKIKHIVWASRELERFAVNPGLLETSEGCRQLGQLQPALQTGSEELRSLYNTVATLYCVHQRI
 EVKDYKDALDKIEEQNKSKKKAQQAADTGNSSQVSNYPYIVQNLQGMVHQAI SPRINAWYKVEKAFSEVYIPMSALSEGATPODL
 NTMLNTVGGHQAAOMLKETINEEAAEWDRLHPVHAGPIAPGOMREPRGSDIAGTSTLQEQIGWNTNPPPIPVGEIYKRWIILGLNKIVRM
 YSPISLIDIFQGPKEPFRDVIYDRFYKTLRAEQASQDVKNWMTETLLVQANPDCKTILKALGPAATLEENMTACQGYGGFGHKARVLAEAMS
 QVTNSTTIMQRGNFRDQRKIVKCFNCGKEGHIARNCRAPRKKGCCKGEGHQMCKCTERQANFLGKIWPSPSHKGRPGNFLQSRPEPTAPPE
 ESFRFGEETTPSQKQEPIDKELYPLASLSLFGNDPSSOS

도면67d

2003 B.anc gag.OPT

ATGGGCGCGCGGCTCCGTCTCTCCGGCGGCGAGCTGGACAGTGGAGAGAGATCCGCTGGCGCCCGGGGCAAGAGAGTACAGCT
GAAGCACATCGTGTGGGCTCCGGGAGCTGGAGGCTTCGGCTGAACCCGGCTGTGGAGACCTCCGAGGCTGCCGCGAGATCCTGG
GCCAGCTGCAGCCCGCTGCAGACCGCTCCGAGGAGTGGCTCCCTGTACACACCGTGGCCACCTGTACTGTGCTGCACAGCCGATC
GAGTGAAGGACACCAAGGAGCCCTGGACAGATCGAGAGGAGCAGACAGTCCAGAGAGGCCAGAGGCCGCGCGACACCGG
CAACTCCTCCAGTGTCCAGACTACCCCTCTGTGCAAGACTCGAGGCGCAGATGTCAGGCGCAGATGTCAGGCGCATCTCCCGCGACCTGAACG
CCTGGGTGAAGTGGTGGAGGAGAGGCTTCTCCCGAGGTGATCCCATGTTCTCCGCTGTCCGAGGCGCCACCCCTAGSACTG
AACACCATGCTGAACACCGTGGGCGGCCACAGGCGCATGGAGATGCTGAAGAGACCATCAAGGAGGCGCGGAGTGGGACCGCT
GCACCCGTCACGCGCGCCCATCGCCCCCGCCAGATCGCGAGCCCGGCTCCGACATCGCGGACCCACCTCCACCTGCAGGAGC
AGATCGGCTGGATGACCAACACCCCGCATCCCGTGGCGGAGATCTACAAGGCTGATCTCTGGGCTGAACAGATGCTGGGATG
TACTCCCATCTCCATCCTGGACATCCCGCAGGCGCCAGAGGCTTCCGGACTACGTGGACCGCTTACAAGACCTGCGCGCGA
GCAGGCTCCAGGACGTGAAGACTGGATGACCGAGACCTCTGTGTGTCAGACGCAACCCGACTCAACCATCTCTGAAGSCCTGG
GCCCGCGCCACCTGGAGGAGATGATCACCGCTGCAGGCGCTGGGCGGCCCGCCACAGGCCCGCTGTGCTGAGTGTCTCACTGCGGCAAGGATCC
CAGTGAACCACTCCACCATCATGATGACGCGCGCACTTCGCGACACGCGAGATGCTGAGTGTCTCACTGCGGCAAGGATCC
CCATCGCGCGCACTGCGCGCCCGCCCAAGAGGCTGCTGGAGTGGCGCAAGGAGGCCACAGATGAAGGACTGCACCGAGCGCC
AGGCCAATCTCTGGGCAAGATCTGGCCCTCCACAGGCGCCCGCGCACTTCTGAGTCCCGCGCGAGGCCACCGCCCTCCCTGAGT
GATCTCTCCGCTTCGGCGAGGAGACCAACCCCTCCAGAGGAGGAGCCCATCCACAGSAGCTGTACCCCTGGGCTCCCTGAAGTC
CTGTTCGGCACGACCCCTCTCCAGTAA

도면68a

8. 2003 CON C gag.pgp
 MGARASILRGOKL DKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLETSSEGGKQIKQLPALQTGTEELRSLYNTVATLYCVHEKI
 EVRDTKEALDKIEEQNKSQKTQQAADGKVSONYPIVQNLQGMVHQAISPRILNANWVKVIEEKAESPEVTFMFTALSEGATPQDINTM
 INTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGFIAPGQMRPEFGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWTILGLNKLIVRMYS
 VSILDIKQGPKEPERDYVDRFEKTLRAEQATQDVKNWMTDILLVQANPDCKTILRALGPGATLEENMTACQYGGGPGSHKARVLAEMSQAN
 NTNIMQRSNEKPGFKRIVKFCNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKINWPSHKGRFCNQLQNRPEPTAPPAESFR
 FEETTPAPKQFPKDRPLETSLKSLFGSDPLSQ\$

도면68b

2003 con c gag.oft
 ATGGGGCGCGGCTCCATCTCGCGCGGCAAGCTGGACAAAGTGGAGAGATCCGCTGCGCCCGCGGCAAGAAAGCACTACATGCT
 GAAGCACTGGTGTGGGCTCCGCGAGCTGGAGCGCTCGCCCTGAACCCGCGCTGTGGACACTCCGAGGCTGCAAGCAGATCATCA
 AGCAGCTGCAGCGCGCTGAGACCGGACCGAGGAGGCTGGCTCCCTGTACACACCGTGGCCACCTGTACTGCGTGCAGAGAGATC
 GAGTGGCGCACACCAAGAGGCGCTGGACAAGATCGAGAGGAGCAGAACAGTCCAGCAGAGACCCAGAGCGCAAGCGCGGACGG
 CAAGGTGTCCAGAACTACCCCATCTGTCCAGAACTGCAGGGCCAGTGGTGCACAGGCCATCTCCCGCGCACCTCGCTGGGTGA
 AGGTGATCGAGGAGAGGCTTCTCCCGAGGTGATCCCATCTTACCGCTGTCCGAGGCGCCACCCCGAGGACCTGAACACCATG
 CTGAACACCGTGGCGGCAACAGGCGCCATGAGATGCTGAAGGACACCATCAACGAGGAGCGCGGAGTGGGACCGCTGCACCCCT
 GCAGCGGCGCCATCGCCCGGCGAGTGCAGCGAGCCCGCGGCTCGACATCGCGGCAACCATCCCTGCAGGAGGATCGCT
 GGATGACCTCAACCCCGATCCCGTGGCGGACATCAAGGCTGGATCATCTGGGCTGAAAGATGTCGCGCATCTACTCCCC
 GTGTCCATCTTGACATCAAGCAGGCGCCCAAGAGCCCTCCCGACTACGTGACCGCTTCTCAAGACCTCTGCGCGCTGGCGCGGCG
 CCAGGACGTGAAGACTGATGACCGCTGCTGTGCAGACGCCAACCCGACTGCAAGACCTCTGCGCGCTGGCGCGGAGCGGAC
 AACACCAATCATGATGACGCGCTCCACTTCAAGGCGCCCAAGCGCATCTGTAAGTCTTCACTGGCGCAAGGAGGCGCATCGCCG
 CAACTGCCGCGCCCGCAAGAGGCTGTGGAAGTGGGCAAGGAGGCGCCACAGATGAAGACTGCAAGGCGCAGGCGCAACTTC
 TGGGCAAGATCTGGCCCTCCACAGGCGCGCCGCGAATCTCTGCAAGACCGCCGAGGCCACCGCCCGCGGAGTCTCTTCGG
 TTCGAGGAGACCAACCCCGCGCCCAAGCAGGAGGCCAAGGACCGGAGCCCTGACCTCCCTGAAGTCCCTGTTCGGCTCCGACCCCGCTGTC
 CCACTAA

도면 68c

9. 2003_C.anc.gag.pEP

MGARASILRGCKLDTWEKILRLPGCKKHHVYKHLVWASRELERFALNPGLELTSEGCKQIMKOLOPALQTGTEELRSLYNTVATLYCVHERI
 EVRDKEALKEIEGQNSQKTOQAAEADGNDGKVQSYPIVQIQMGQHQVQASPTLNANWVVEKAFSEPIEMFTALSEGAFTPODL
 NTLNLTUVGHQAAMQMLCTINERAAEALHLPVHAGVPAGQWREPGSDIAGTSTLQEQELAWMTSNPPIEVDGIYKRWITLGLNGLIVRM
 YSPYSILDTKQGPPEFRVDVDFEKTILRAEQAAQTDVKNWMTDILLVONANPOCKTILRALGPGATILEMTACQGVGGPCHKARVLAEAMS
 QANTNIMQWRSNFKGPKLIVKCNCEGHIARAPFKKGWCKGKEGHQMKDCTERQANFLGKIWPESHKGRPNFLQSRPEFTAPPAE
 SFRENTTITAPKQEGDEPLTSLKSLFGSDPLTSQ\$

도면68d

2003 C.and.gag.OPT

ATGGCGCCCGCGGCTCCATCCTGGCGCGGCGCAGCTGGACACACTGGGAGAAAGATCCGCTGGCGCCCGCGGCAAGAAGCACTACTATGAT
CAAGACCTGGTGTGGGCTCCCGGAGGTGGAGCGCTTCGCGCTGAACCCGGGCTGTGGAGACCTCCGAGGCTGCAAGCAGATCATGA
AGCAGCTGAGCGCCGCTTGAGACCGGACCGAGGAGTGGCTCCCTGTACACACCGTGGCCACCTGTACTGGTGCAAGCGCATC
GAGGTGGGAGACACCAAGAGGCTTGACAAAGTCCGAGAGAGCAGACAAATCCAGCAGAGAGCCAGAGGCGGCGGCGGACGG
CGACACGGCAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCGGGCCAGATGCCATGTTCACGCGCTGTCCGAGGCGCCACCCGAGGACCTG
CCTGGTGAAGTGTGGAGGAGAGGCTTCCTCCCGAGTGTATCCCATGTTCACGCGCTGTCCGAGGCGCCACCCGAGGACCTG
AACACCATGCTGAACACCGTGGCGGCGCACCCAGGCGCATGCRAGATGCTGAAGACACCATCAACGAGGAGGCGCGGAGTGGACCGCCT
GCACCCGTGCAAGCCGCCCCGTGGCCCCCGGCGAGATCGCGAGCCCCGCGCTCCGACATCGCGGCGCACCTCCACCTGCGAGG
AGATCGCTGGATGACCTCAACCCCCCATCCCGTGGCGGACATACAAAGGCTGGATCATCTGGGCTGAACAAAGATCGTGCGCATG
TACTCCCCGTGTCATCTGTGACATCAGCAGGCCCCAAGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAGACCTTCGCGGCGGA
GCAGGCCACCCAGACGTGAAGACTGGATGACCGACACCTGTGTGTCAGAACCGCAACCCGACTGCAAGCCATCTCTGGCGGCTG
GCCCGGCGCCACCTGGAGGAGATGATGACCGCTGCAGGCGTGGCGGCCCGGCGCACAGGCGCGCTGTGGCGGAGGCCATGTCC
CAGGCCAACACCAACATCATGATGACGCGTCCAACTCAAGGGCCCCAAGCGCATCGTAACTGCTCAACTGCGGCAAGGAGGCCA
CATCGCCGCAACTGCGGCGCCCCCGCAAGAGGGCTGTCTGAGAGTGGCGCAAGGAGGCCACCAAGTGAAGGACTCACCGAGCGCAGG
CCAACTTCCTGGGCAAGATCTGGCCCTCCACAGGCGCGCCCGGCACTTCCTGCACTCCGCGGAGGCCACCGCCCCCGGCGGAG
TCCTTCGCTTCGAGGAGACACCCCCCGCCCCCAAGCAGGAGGCCACCGAGGCTCCCTGAAGTCCCTGTTCGGCTCCGA
CCCCCTGTCCCCAGTAA

도면69a

10. 2003 CON D gag .pep

MGARASVLGGKLDWEKIRLRPGGKKKYLKHIVWASRELERFALNPGILLETSEGCKQIIIGQLQPAIQMGSEELRSYNTVATLYCVHERI
 EVKDTKEALEKIEEFQNKSKKKAQAAADTGNSSOVSONYPIVONLQGMVHQALSPRTINAWKVIIEKAFSPYIIMFSALSEGATPODL
 NTMLNTVGGHQAAMQMLKETINEERAEWDRLHPVHAGVPAGQMPREPRGSDIAGTTSTLQEQICWMTSNPPPIPVGEIYKRWIILGLNKIVRM
 YSPVSLIDIRQGPKEFFRDYVDREYKTLRAEQASQDVKNWMTETLLVQANPDCKTILKALGPEATLEEMMTACQGVGGFSPHKARVLAEMS
 QATNSAAVMQRGNEFKGPRKIIKCNCGKEGHIKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSPHKGKRPNGNFLQSRPEPTAPPA
 ESFGEGEEITPSQXQEQKDKELYPLTSLKSLFGNDPLSQ\$

도면 69b

2003 CON D *gag*. OPT
ATGGCCGCGCGGCTCCGTGCTCCGGCGGCAAGCTGGGAGCAATCCGCTGGCCCGCGGCAAGAAAGATACCGCCT
GAGGCACATCTGTGGGCTCCCGGAGGTGGAGGCTTCGCCCTGACCCCGGCTGCTGGACCTCCGAGGCTCAAGACGATCATCG
GCCACTGGAGCCCGCCATTCAGACGGCTCCGAGAGTGGCTCCGTATCAACACCTGGCCACCTCTGCTGACGACGACATCG
TGCTGAAGACACCAAGGAGCTCTGGAGAAATCGAGGAGGACGACAACCTCAAGAAGCGCAGCGGCCGCCGACACACCG
CAACTCTCCAGGTGCCAAGAACTCCCATCTGTCAAACTCGAGGCGAGATGTCGACAGGCCATCTCCGCCGACCCCTGAAC
CTTGGTGAAGTATCGAGGAGAGGCTTCTCCCGAGGTGATCCCATGTTCTCGCCCTGTCTGAGGGGCCACCCCGAGGACTG
GACACCTGTGAACACCTGGGCGCCACAGCGCGCATGATGATCTGAAGAGGCTCATCAAGGAGGCGCTCCGAGTGGGACGCGCT
CAACCCGTGCACCGCGCCGCGCGCGAGATGTCGAGCGCCGCGCTCGACATCGCGGCAACCTCCACCTCCAGAGG
AGATCGGTGGATCACTCAACCCCGCATCCCGTGGGCGAGATCAACAAGCGTGGATCATCTGGCCCTGAACAAGATCGTGGCATG
TACTCCCGCTCCATCTGTGACATCCCGAGGCGCCAGAGGACCTTCGCGACTCTCGGACCGCTTCAACAACCTCTCGCGCGG
GAGGCTCTCCAGAGCTGAAGATGATGACAGCAGACCTGCTGTGACAGACCAACCCGACATGCAAGCATCTCTGAAAGCCCTCG
GCCCGAGGCGCACCTGGAGAGATGATACCGCTTCGAGGGGTGGCGGCCCTCCCAACAGGCGCCGTGCTGCGCGAGGCCATGTCC
CAGGCAACCACTCCGCGCGGTGATGATGAGCGGGCACTTAAGGGCCCCGGAAGATCTCAATGTGTCGCGCAAGAGGG
CCATCGCCCAAGAACTCGCGGCCCGCGCAAGAGGCTGTGGAAGTGGCGAGGCGCCACAGATGAAGGACTGCACCGAGCGGC
AGGCCAATCTCTGGGACAGATTGGCCCTCCACAGGCGCCGCCGCAACTTCCTGCAGTCCCGCCCCGAGGCCACCCGCCCGCCG
GATGCTTCGGCTTCGGCGAGGATACCCCTCCCAAGAGGAGGACGAAGGAGTGTACCCCTGACCTCCCTGAGTCCCTGAGTCCCT
TCTGGCACAGACCCCTTCTCCAGTAA

도면70a

11. 2003 CON F gag. PEP
 MGARASVLSGGKLDWEKIRLRPGGKKYRMKHLVWASRELERFALDPGLLETSEGCQKIIGQLQPSLQTGSEELRSLYNTVAVLYCVHQKV
 EVKDTKEALEKLEEEQNKSOQKTOQAAADKGVSONYPIVONLQGMVHQAI SPRTLNAWKVIEEKAFSFEVIEFMSALSEGATPQDLNTML
 NTVGGHQAANQMLKDTINEEAAEWDRLHPVHAGTIPPGQWREPRGSDIAGTTSTIQEQIOWMTSNPPVPVGDIVKRWI IIGLNKIVRMYSVP
 SILDIRQGPKEPFRDYVDRPFKTLRAEQATQEVKGWMTDILLVQVANPDCKTILKALGPGATLEEMWTACQGVGGPGHKARVLAEMSQATN
 TAIMQKSNFKGQRRIVKCNCGKEGHIARNCRAPRKKGCWKCGREGHQNKDCTERQANFLGKIWPSNKG RFGNFLQSRPEPTAPPAESFGF
 REEITPSPKOEQKDEGLYPPLASLKSIFGNDP\$

도면70b

2003 CON F gag .OPT

ATGGGCGCCGCGGCTCCGTCGTCCGCGGCAAGCTGGACGCCCTGGAGAGATCCGCTGCGCCCCGCGGCAAGAGAGTACCGCAT
 GAAGACCTGGTGTGGGCTCCCGAGCTGGAGCGCTTCGCTTGGACCCCGGCTGTGGAGACCTCGAGGCTGCCAGAGATCATCG
 GCCAGCTGACGCTCCCTGAGACCGGCTCCGAGAGCTGCGCTCCCTGTACACACCGTGGCGTGTCTACTGCGTGCACCCAGAGAGTG
 GAGTGAAGGACACCAAGAGGCCCTGGAGAGCTGGAGGAGAGAGACAAGTCCCAGAGAGACCCAGAGGCCGCCGCCACAGGG
 CGTGTCCAGAACACCCATCGTGCAGAACCTGCAGGCCAGATGGTGCACCGCATCTCCGCCCGCACCCCTGACCCCTGGGTGAAGG
 TGATCAGGAGAGGCTTCTCCCGAGGTGATCCCATGTTCTCGGCTGTCCGAGGGCGCACCCCGCAGGACCTGAACACCATGCTG
 AACACGTTGGCGGCCACCAAGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGAGTGGGACCGCTGCACCCCGTGCA
 CGCGGCCCATCCCGCCGCGGAGATGCGGAGCCCGGCTCCGACATCGCGGCAACCATCCACCTGCAGGAGCATCCAGTGA
 TGACCTCAACCCCGCTGCGGTGGCGACATCTACAAGCGCTGGATCATCTGGGCTTGAACAAGATCGTGGCATGTACTCCCGCTG
 TCCATCTGGACATCCCGAGGCGCCCAAGAGCCCTTCGCGACTACGTGGACCGCTTCTCAGACCTTGGCGCCGAGCAGGCCACCA
 GGAGTGAAGGCTGGATGACCGACCCCTGCTGTGCAGACGCCAACCCTGAGACCATCTGAAGGCTTGGGCCCCGGGCCCCA
 CCTGGAGGAGATGATGACCGCTGCCAGGCGTGGCGGCGCCCGGCCACAAGCCGCTGTGCGGAGGCCATGTCCAGGCCACCAAC
 ACCGCATCATGATGCAGAGTCCAACTTCAAGGCCAGCCCGCATCTGAAGTCTTCACTGCGGAGAGGACCATCGCCAGAGAA
 CTGCGCGCCCCCGCAAGAGGGCTGCTGAAGTGCAGGCGCGAGGCCACCCAGATGAGGACTGCACGAGGCCAGGCCAACTTCTGTG
 GCAAGATCTGGCCCTCCACAAGGCGCGCCCGCAACTCTGTGAGTCCGCGCGAGCCACGCGCCCCCGCGAGTCTCTTGGCTTC
 CGCAGGAGATCACCCCTCCCGCAGCAGAGCAGAGGCGCTGTACCCCGCCCTGGCTCCCTGAAGTCCCTGTTGGGCAACGA
 CCCCTAA

도면71a

12. 2003 CON G gag.pgp

MGARASVLSGCKLD AWEKIRLRPGGKKKYEMKHLVWASRELERFALNPDLLETAEGCQIMGQLQPALQTGTEELRSLNTVATLYCVHQRI
 EVKDTKEALEEVEKIQQKSOOKTQQAANDEGNSQVSYFIVQNAQGMVHQALSPRTINAWKVVEKAFSPVPIPMFSALSEGATPQDL
 NTMLNTVGGHQAAQMQLKDTINEEAAEWDHRMHPQOAGPIPPGOIREPRGSDIAGTSTLQEQIRWMTSNPPIPVGEIYKRWIILGLNKIVRM
 YSPVSLDIFQGPKEPFRDYVDRFETLRRAEQATQEVKGMWTDLLVQANANPDCFTILRALGPGATLEENMTACQGVGGPSHKARVLAEAMS
 QASGAAAAMWQSNFKGPRRTIKCFNCGEGHLLARNCRAPRKKGCKGKGHOMKDCETERQANFLGKLTWPSNKGKRPGNFLQNRPEPTAPP
 AESFGFGEETAPSPKOEKEKELYPLASLKSLEFGSDP\$

도면71b

2003 CON E gag-DPI

ATGGGCGGCGCGCTCCGTGCTGTCGGGGCAAGCTGGACGCTGGAGACAGATCCGGCTGGCCCCCGGCCAAGAAGAGTACCGCAT
 GAGGACCTGGTGGGGCTCCCGGAGCTGGAGCGCTTGGCCCTGAACCCGACCTGTGGAGACCCGGAGGGCTCCAGCAGATCATGG
 GCCAGCTGACGCCGCCCTGACACCGGACCGAGAGTGGCTCCCTGTTCAACACCTGGCACCCCTGTACTGCTGCACCCAGCGCATC
 GAGGTGAAGGACACCAAGAGGCGCTGGAGAGTGGAGAATCCAGAGAAGTCCAGAGAGACCCAGCAGGCCGCCCATGGACGAGGG
 CAATCCTCCAGGTGTCAGAACTACCCCATGTCAGAACGCCAGGCCAGATGTGACCAAGGCCATCTCCCCCGCAACCTGAACG
 CCTGGTGAAGTGGTGGAGAGAGGCCCTTCTCCCCGAGGTGATCCCATGTTCTCCGCTTCTCCAGGGGCCACCCCCAGGAOCTG
 AACACCATGCTGAACDCTGGGCGGCCACAGGCCGCCATGCAGATGCTGAAGSACCAATCACAGAGAGGCCGCCGAGTGGGACCCGAT
 GACCCCGAGAGCGGCCCATCCCCCGGCGAGATCCCGGAGCCCGGGCTCCGATCCCGGACCACTCCACCTTCAGGAGC
 AGATCCGCTGGATGACCTCAACACCCCCCATCCCGTGGCGAGATCTAAGAGCTGGATCATCTGGGCTGAACAGATCGTGGCATG
 TACTCCCCCTGTGCACTCCTGGACATCCGCAAGGCCCCCAAGGAGCCCTTCGCGACTAGTGGACCGCTTCTCAAGACCTCGCGCCGA
 GAGGCCACCCAGAGGTGAAGGCTGGATGACCGACACCTGCTGGTGCAGAACGCCAACCCGACTCAAGACCATCTGCGGGCCCTGG
 GCCCGGGCCACCCCTGGAGAGATGATGACCGCTGCCAGGGCTGGGGGCCCTCCACAGGGCCGCGTGGCCGAGGCCATGTCC
 CAGGCTCCGGCGCGCGCCGCAATCATGATGAGAAATCACTCAAGGGCCCCCGCCACCATCAAGTGTTCAACTGCGGCAAGGA
 GGGCCACTGGCCCGCAACTGCGCGCCCCCGCAAGAGGGCTGCTGAAAGTGGGGCAAGGGGCCACAGATGAGGACTGCACCGAGC
 GCCAGGCCAACTTCTGGGCAAGATCTGGCCCTCAACAGGGGCCCGCGGCACTTCTGCAAGAACCGCCGAGGCCACCGCCCCC
 GCCGAGTCCCTCGGCTTCGGCGAGGAGATCGCCCCCTCCCCCAAGCAGGAGCAGAGGAGAGGCTGTACCCCTGGCCCTCCCTGAGTC
 CCTGTTCGGCTCCGACCCCTAA

도면72a

13. 2003_CON_H_gag.pep
 MGRASVLSGGKIDAEKIRLRPGGKKYRLKHLVWASRELERFALNPGLETAEGCLQIEQLQPAIKTGTETELQSLFNTVAVLYCVHQRI
 DVKDTKEALGKIEETQNKSOQKTQAAADREKDNKVSQNPVQNAQGMVHQAIAPRTINAWKVVEKAFSPVPIPMFSALSEGATPDDL
 NAMLNTVGGHOAAMQMLKDTINEEAAEWDRLHPVHAGPIPPQOMREPRGSDIAGTSTLQEQIAWMTGNPPIPVGDIIYKRWIILGLNKIVRM
 YSPVSLIDIKQGPKEPFRDYDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGQGASIEENMTACQGVGGESHKARVLAAMS
 QVTNANAATMQKGNFKGPRKIVKCFNCGKEGHIARNCRAPRKKGCKGREGHQMKDCTERQANFLGKIWPSSKGRPGNFIQSRPEPTAPP
 AESFGFGEEMTPSPKQELKDKEPPLASLRSILFGNDPLSQ\$

2003_CON_H_gag.OPT

ATGGGCGCGCGGCGCTCCGTGCTCTCCGGCGGAAGCTGGACGCGCTGGAGAAACATCCGCTTGGCCCCGGGCAAGAAGATACCGCCT
GAACACCTCGGTGTGGGCTTCCGCGAGCTGGAGGCTTCCGCTGAACTCCGGGCTGTGGAGACCGCGAGGAGTGGCTGCAGATACATCG
AGACCTGGAGCGTGGCGCTCAAGACCGGACGAGGAGTGCAGTCTCTTGGACACCGTGGCGGTGTGCTGACTGGTGCACACGCGCATC
GACCTGAAGGACACCAAGGAGGCGCTGGCAAGATCGAGGATCCAGAACATCCAGACAGAAGACCCAGCAGCGCCGCGCGACAGGA
GAGGACACAAAGTGTCCCAAGATCTGCCAATCTGCAGAACGCCAGGGCGAGTGTGCACAGGCGCATCCCTCCCGCCACTCTGAACG
CTTGGGTGAAGTGTGTGAGGAGAAGGCTTTCTCCCGAGGTATCCCATGTTCCGCTGTCCGCTGTCCGAGGGCGCCACCCCGCAGGACTG
AAGCCATCTGTGAACACCGCTGGGCGCGACCAAGCGGCATCAGATGTTGAAGACACCATCAACAGGAGGCGCCGCGAGTGGGACCTGCCT
GAACCCCTGACGCGCGGCGCCCATCCCGCCGACGCGGAGCTCGACATCGCCGCGGACCACTCCACCTCGACGAGT
AGATCGCTGGATGACGGCAACCCCGCTCCGTTGGCGGACATCTACAGGCTGATCATCTGGCGCTGAACAGATCGTGCAGATG
TACTCCCGCTTCACTCGATCAAGCAGGCGCCCAAGAGCGCTTCCGAGATACCTGGACGCTTCTCAAGACCTCTCGGCGCGA
CGAGGCCACCGAGCGCTGAAGACTGATGACCGACACTCTGCTGTGCAGACGCCAACCGGATTCGACACCATCTCGCGCCCTGG
GCCAGGGCGCTCCATCAGAGGATGATGACCGCTTGCAGGGCTGGGCGGCGCCCTCCACAAAGCGCTGGTGTGGCGAGCGCATTC
CAGGTGAACAAGCCACCGCGCATCATGATCAGAGAGGCACTTCAAGAGGCGCCGACAGGCTTCAAGTCTCACTCGGCGAAGGA
GGGACATCTCGCGCGCACTCGCGCGCCCGCAAGAGGCTGTCAAGTGCGGCGCGAGGGGCGACAGATGAGGAGCTGCACGAG
GCCAGGCAACTTCTTGGGCAATCTGGCCCTCTCAAGGGCGCGCGGCAACTTCTGAGTCCCGCCCGGCGACCGCGCCCGC
CGGAGTCTTGGCTTGGCGGAGGATGACCCCTCCCGCAAGCAGGAGTGAAGGACCAAGAGCGCGCCCTGGGCTCCTCGGCTCCCT
TTTCGGCAACAGCCCGCTTCCAGATAA

도면73a

14. 2003 CON K gag . pep
 MGRASVLSGGKLDWEKTRLRPGGKKYRLKHLVWASRELERFALNPSSILETTEGCRQIIROLQPSLOTGSEEEKSLFNTVATLYCVHORI
 EVRTKRALDKLEEEQNKSQKTOQTADKGVSONYPIVONLQGMVHQALSPTILNAWVKVIEEKAESPVIEMFSALSEGATPQDLNMTL
 NTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMPREPRGSDIAGTTSTLQEQITWMTSNPPVPVGEIYKRWILLGLNKIVRMKSPV
 SILDIROGPKPEPRDYVDRFFKTLRAEQATQEVKNWMTDILLVQANPDCKTIILKALGPASLEEMMTACQGVGGPGHKARILAEAMSOVTN
 TAYMQRGNFKGQRKIIKCFNCGKEGHIANCRAPRKKGCWKCKEGHQKDCITERQANFLGKIWPSNKGCRPGNFIQSRPEPTAPPAESFGF
 GEETPSRQETKDKKEQGPPLTSLKSLFGNDPLSQ\$

도면73b

2003 CON K gag.OPT
 ATGGGGCCTGGGCTCCGCTGTCCGGGGCAAGCTGACACCTGGAGAGATCCGCTGGCCCGGGCAAGAGAACTACCCCT
 GAAGCACCTGTGTGGGCTCCCGGAGCTGGAGGCTTGGCCCTGAACCCCTCCCTGTGGAGACCAACGAGGGCTGCGCCAGATCATCC
 GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGAAGTCCCTCTTCAACACCGCTGGCCACCCCTGTATCTGGTGCACCAAGCATC
 GAGTCCGGACACCAAGGAGCCCTGGACAAGCTGGAGGAGGACAGAACAGTCCACAGAGAACCCACGAGGAGACCCGACCAAGGG
 CGTGTCCCAAGACTACCCCTCCGTGAGAACCTGCAGGGCCAGATGGTGCACAGGCCCTGTCCCGCCGACCCCTGAAGCCCTGGGTGAAG
 TGATCGAGGAGAAGGCTTCTCCCGAGGTGATCCCATGTTCTCCGCTGTCCGAGGGCCACCCCGAGACCTGAACCCATGCTG
 AACACCGTGGGGGCAACCAAGCCGCAATGCAGATGCTGAAGGACACCATCAACGAGGAGCCGCGACTGGGACCGCTGCACCCCTGCA
 CGCGGCCCATCCCGGCCAGATGGCGAGCCCGGCTCGGACATCGCGGACCCACTCCACCTGCAGGAGCAGATCAACCTGGA
 TGACCTCAACCCCGCTGCTGGGGAGATCTACAGCGCTGGATCATCTGGGCTGAACAGATCTGCGCATGTACTCCCGGTG
 TCCATCTGGACATCGCCAGGGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACCTGCGCGCAGGAGCCAGCCCA
 GGAGTGAGAACTGGATGACCGACCCCTGCTGTGTCAGAACGGCAACCCGACTGCAGACCATCTGAAGGCCCTGGGCCCGGCGCCT
 CCTTGAGGAGATGATACCGCTGCAGGCGTGGGGGCGCCGCGCAAGTCAAGTGCTTCAACTGCGCAAGGAGGCCACATCGCCGCAA
 ACCGCGTGATGATGCAGCGGGCACTTCAAGGCGCAGCGCAAGATCATCAAGTGCTTCAACTGCGCAAGGAGGCCACATCGCCGCAA
 CTGCGGGGCGCCCGCAAGAGGGCTGCTGGAAGTCCGGAAGGAGGGCCACCATGAAGGACTGCACGAGGCGCCAGCTTCTG
 GCAAGATCTGGCCCTCAACAAGGCGCGCCGCAACTTCTGAGTCCCGCCGAGCCACCGCCCGCGAGTCTCTCGGCTTC
 GGCGAGGAGATACCCCTTCCCGCCCGCAGGAGACCAAGGACAGGAGAGGGCCCCCTGACCTCCCTGAAGTCCCTGTCGGCAACGA
 CCCCCGTCTCCAGTAA

도면74a

15. 2003 CON 01 AE gag. PEP
 MGAASVLSGGKLDANEKIRLRPGKKKYRMKHLVWASRELERFALNPGILLETAEGCCQIIIEQLQSTLKTGSEELKSLFNTVATLWCVHQRI
 EVKDTKEALDKIEVONKSOQKTQOAAAGTSSSKVSNYPPIVONAQOMVHQPLSPRTLNAMVKVVEKGFNPEVIMFSALSEGATPQDL
 NMNLNIVGGHQAAQMLKETINEEAAEWDVRVHPVHAGPIPPGOMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGDIYKRWIILGLNKLVRM
 YSPVSLDITROGPKPEPFDYVDRFYKTLRAEQATQEVKNWMTETLLVQANPDCKSILKALGTGATLEENMTACQGVGGPSHKARVLAAMS
 QAQHANIMQRGNEKGQKRIKCFNGCKEGHLCARNCRAPRKKGCWKCKGDKCTERQANFLGKIWPSNKGEPGNFPQSRPEPTAPPAEN
 WGMGFEITSLPKOEOKDKEHPPPLVSLKSLFGNDPLSQ\$

도면74b

2003 CON 01 AE gag.OPT
 ATGGGGCCCGGCGCTCGCTGCTCGGCGGCAAGCTGGAGCGCTGGGAGAAGATCGGCTCGGCCCGGCGGCAAGAAGAAGTACGGCAT
 GAAGACCTGTGTGGGCTCCGCGAGCTGGAGCGCTTCGCCCTGAGCCCGGCCCTGCTGGAGACCGCCGAGGGCTGCCAGAGATCATCG
 AGCAGCTGAGTCCACCTGAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCGTGGCCACCCCTGTGGTGGTGACACGACGATC
 GAGTGAAGGACACCAAGGAGGCTGGACAAAGATCGAGGAGTGCAGAACAGTCCAGCAGAAAGACCCAGCAGGCGCGCGGCGCACCGG
 CTCTCTCCAAAGTGTCCAGAACTACCCCAATCGTGCAGAACGCCAGGGCCAGATGGTGCACAGCCCTCTCCCCCGCACCCCTGAACG
 CCTGGGTGAAGTGGTGGAGGAGAAGGCTTCAACCCGAGTGCATCCCATGTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGTGGGACCGCGT
 AACATGATCTGACATCGTGGGCGGCCACAGGCGGCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCGAGTGGGACCGCGT
 GACCCCGTGCACGCGGCCATFCCCCCGGCGAGATGGCGAGCCCGGCTCCGACATGCCGACATGCCGACACCTCCACCCCTGCAGGAGC
 AGATCGGTGGATGACCAACACCCCGCATCCCGTGGCGACATCTAACGCTGGATCATCTGGCCCTGACAAAGATGATGGCATG
 TACTCCCCGTGTCCATCTTGACATCCGCCAGGCCCCAAGAGCCCTTCGCGACTACGTGGACCGCTTCTACAGACCTGCGGCCGA
 GCACGGGCCACCTGGAGGAGATGATCACCGCTGCCAGGGCGGCGCCCTCCACAGGCCCGCGCTGGCCGAGCCCATCTCTGAAGSCCTTG
 CAGGCCAGCACGCCCAACATCATGATGCAGCGCGGCACTCAAGGGCCAGAGCGCATCAAGTCTTCACTGCGGCAAGGAGGGGCAACCT
 GGCCCGCACTGCCGCGCCCCCGCAAGAGGGTGTCTGGAAGTCCGGCAAGGAGGGCACCAAGATGAGGACTGCACCGAGGCGGAGCCCA
 ACTCTCTGGGCAAGATCTGGCCCTCCACAAGGGCGCGCCGCACTCCCGAGTCCCGCCGAGCCCAACCGCCCGCGAGAAC
 TGGGGCATGGGCGGAGATCACTCCCTGCCCGAGCAGGAGCAGAGACCAAGGAGCACCCCCCCCCCTGGTGTCCCTGAAGTCCCTGTT
 CGCAACGACCCCTGTCCAGTAA

도면75a

16. 2003 CON 02 AG gag.pgp

MGAPASVLSGGKLDWEEKIRLRPGGKKKYRLKHLVWASRELERFALNPGELLETAEGCCQIMEQLQSALFTGSEELKSLYNTVATLWCVHQRI
 DIKDTKEALDKIEEVQNKSKQKTQOAAATGSSSSONYPVQNAQSQMTHQSMSPRTLNWVKVIEKAFSPVIFMFSALSEGATPQDLNMM
 LNIYGGHQAAMQMLKDTINEEAAEDWDRVHPVHAGPIPPQGMREPRGSDIAGTTTLOEQIGWMTSNPPIVGEIYKRWIVLGLNKIVRMYSP
 VSILDIRQGPKEPFRDYVDREFFKTLRAEQATQEVKNWMTETLLVQNAEDCKSLRALGPGATLEEMMTACQGVGGPGHKARVIAEAMSQVQ
 QSNIMMQNGNFRGQRTIKCFNCGKEGHLARNCKAFPRKKGGCWKCGKEGHQMKDCTERQANFLGLWPSSKGRPGNEPQSRPEPTAPPAESFGM
 GEETSSPKQEPDRDKGLYPPLTSLKSLFGNDP\$

도면 75b

2003 CON 02 AG gag.OPT

ATGGCGCGGGCGCTCCCTGCTCTCCGCGGCGAGCTGGACGCGCTGGGAGAGATCCGCTCGCGCCGCGGCGCAGAGAGACTACCGCCT
GAAGACACCTGGTGGGCGTCCCGCGAGTGGAGCTTCCCTGAACTCCGCGCTGTGGACCGCGAGGCTGCCAGAGCATATGG
AAGAGCTCAGATCGCGCTCGCGACGGCTCGAGGAGTGAACCTGTGTAACAAACCGTGGCCACCTGTGTGCGTCCAGCAGCGCATC
GACATCAAGACACCAAGAGGCGCTTGACAAATTCAGGAGGTGCAGACAAATCCAGACAGAGACCCAGCAGGCGCGCGCGCGACCGG
CTCTCTCCCGAATACCTCCATCTGTGGAGAACCGGAGGTCAGCCACAGTTCATCCCGCGCATCTGAACTGCTGAGCTTGGGTGA
AGGTGTCCAGAGAGAGGCTTCTCCCGAGGTGATCCCATCTCTCCGCTCTCGAGGCGGCACCCCGAGGACTGCAACCTGGATG
CTGAACATCTGGCGGCCACAGGCGGCATGCAGATCTGAGAGACACCATCAACAGAGAGGCGCGCGAGTGGGACCGCTGTGCACCCCGT
GCACGCCGCGCCATCCCGCGGAGTGGGAGCCCGCGCTCGCATATCGCGGCACCATCCACTCGAGGAGAGAGATCGGCT
GGATGACTCCACCCCGCATCCCGCTGGGCGAGATATCAACGCTGGATCTGTGGCGCTGAAACAGATGTGAGCATCTACTCTCCCT
CTGTCTCATCTGGACATCGAGGCCCGCAAGACCGCTCTCCGAGATGTCGAGCGCTTCTTCAAGCCTCTCGCGCGCGAGCAGGCCAC
CGAGGTGTGAACATGTGATGACCGAGACCTCTCGTGTGCAGAACCCCAACCGACTCATCTGGCGGCGCTTGGCGCGCGGCG
CCACCTGAGAGAGATGATACCGCTCGCAGGGGTGGGCGGCCCGCGCAGGCGCGGCTGTGCGCGAGGCGCATGTCCCGAGTGG
CAGTCCAACTCATGTGAGGCGAACTTCCTCGCGCGACGACCATCAATGCTTTCAGTCCAGGAGGCGGCGCATCTGGCGCGAA
TCTGAAGGCCCGCGCAAGAGGCTGCTGGAATCGCGCAAGAGGCGCACAGATGAAGGACTGCACCGAGCGCGAGCCAACTTCTCTG
GCAGATCTGGCCCTCTCCAAAGGCGCGCCCGGCAATCTCCCACTAGTCCCGCGCGCGCCCGCGCGAGTCTCTTGGCGATG
CGCGGAGAGATCACTCTCTCCCAAGAGGAGCGCGCGAGGGCTGTATCCCGCCCTGACCTCGCTGAGTCCGCTGTTCGGCAACGA
CCCCGTAAG

도면76a

17. 2003 CON 03 ABG gag . pep
 MGARASVLSGGKIDAEKIRLRPGKKKRYIKHLVWASRELERFALNPISLETSEGCCQILEQLQPTLTGTSEELKSLYNTVATLYCVHORI
 EIKDKREALDKIEIQNKSKQTQQAATGTGSSSKVSNYPVQNAQGMTHQSMSPRTLNWVKVIEKAFSEPIPMFESALSEGATPQDL
 NMLNIVGGHQAAMQMLKDTINERAEWDRIHPAQAGPPEPGQMEPRGSDIAGTTSTLQEQIGWMTSNPPIPVGDIYRRWIIIGLNKIVRM
 YSPVSLIDIROGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTETILLVQNANPDCKTILRALSGATLEEMNTACQGVGGPGHKARVLAPEMS
 QVQANIMMQKSNFRGPKRIKCFNCGKDGHLARNCRAPRKKGCWCKGEGHQMKDOCTERQANFLGRIFWSSKGRFGNFPQSRPEPSAPPAEN
 FGMGERITPSLKQEQKDREQHPPSISLSLFGNDEL\$Q\$

도면76b

2003 CON 03 ABG gag. OPI

ATGGGCGCCGCGCCCTCGTCTGTCGGGGCGCAAGCTGGACGCCCTGGGAGAGATCCGGCTGCGCCCGCGCGGCAAGAGAGTACCGCAT
 CAGACCTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCTCGAGGGCTCCAGCAGATCCTTGG
 AGCAGCTGACGCCACCTGAAGACCGGCTCCGAGGAGCTGAGTCCCTGTACACACCCGTGGGACACCTGTACTGGTGCACCGGCACTTC
 GHGATCAAGGACACCAAGGAGCCCTGGACAGATCGAGGAGATCCAGAACAGTCCAAAGCAGAGACCCAGCAGCGCCACCGGCAACCGG
 CTCTCTCCCAAGGTGTCCAGAACTACCCCATCGTCAGAACGCCAGGSCCAGATGACCCACAGTCCATGTCTCCCGGCGCCACCCCGAGGACCTG
 CTTGGTGAAGGTGATCGAGGAGAAGSCCTTCTCCCGGAGTGATCCCATGTCTCCCGGCTGTCCGAGGGCGCCACCCCGAGGACCTG
 AACATGATGTAACATCGTGGGGGCGCCACAGGCGCCATGCGAGTCTGAAGGACACCATCAAGAGGCGCGCGGAGTGGGACCGCTT
 GCACCCGCGCCAGGCGCGCCCTTCCCGCGCGCAGATGCGCGAGCCCGCGCTCCGACATGCGCGGACACCTCCACCTGCGAGGAGC
 AGATCGGCTGGATGACCTCCACCCCGCCATCCCGTGGGCGACATCTAACAGCTGGATCTCTTCAAGACCTGTCAGAGATGTCGCGATG
 TACTCCCGGTCCATCTGGACATCCCGAGGCGCCCAAGGAGCCCTTCGCGACTAGTGGACCGCTTCTTCAAGACCTGTCGCGCGCA
 GCAGGCCACCGAGGACGTGAAGACTGTGATGACCGAGACCTGCTGGTGCAGAACGCCACCCGACTGGAAGACCATCTGCGCGCGCTGG
 GCTCCGCGCGCACCTGGAGAGATGATGACCGGCTGCCAGGGCGGGGGCGCCCGGCGCAAGCGCGCTGCTGGCGAGGCCATGTCC
 CAGGTGAGAACGCCAACATCATGATGACAGATCCAACTTCGCGGCGCCCAAGCGCATCAAGTGTTCAACTGCGGCAAGGACGCGGACCT
 GCGCGCAACTGCGCGCGCCCGCAAGAGGGTGTCTGGAAGTCCGGGCAAGGAGGCCACCATGAGGACTGCAAGGAGCGCGGAGGCA
 ACTTCTGTGGCGGATCTGGGCTCTCTCCAGGGGCGCCCGGGGAACTTCCCGAGTCCCGCGGAGCCCTCGCGCGCGCGGAGAAC
 TTCGGCATGGGCGGAGAGATCACCCCTCCCTGAAGCAGGAGCAAGAGGACCGGAGCAGCACCCCCCTCCATCTCCCTGAGTCCCTGT
 CGGCAACGACCCCGCTGTCCGATGA

도면77a

18. 2003 CON 04 CFX gag' pep

MGARASVLSGGKLDAMERIRLRPGGKKYKHLVWASRELERFALNPGLLETAEGCQQLMEQLQSTIKTGSEELKSLFNTIATLWCVHQRI
 DVKDTKEALDKVEEMQNKSKQKTQQAADITGSSNVSQNYPIVQNAQGMVHQSIPTINAWYKVIIEKAFPEVIMFSALESGATPQDL
 NMMLNIVGGHQAAQMQLKDTINEFAAEWDRAHPVHAGPIPPGQREPRGSDIAGTTSTIQEIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
 YSPVSLIDIRQGPKEPFROYVDREPKCLRAEQATQEVKNWMTETILLVQNNANPDCCKSILKALGIGATLEEMMTACQGVGGPSHKARVLAAMS
 QASNAAAANMQKSNFKGQRIIKCFNCGKEGHLARNCRAPRKGCWKCGKEGHQMKDCTERQANFLGRMWPSKGRPCGNFLQSRPEPTAPP
 AESLEWKEETTSPPKQEPKDELYPLTSLKSLFGSDPLSQ\$

도면77b

2003 CON 04 CFX gag.OPT
 ATGGGCGCGGCGCTCGGTGCTGTCCGGCGGCAAGCTGGACGCTGGAGCGCATCCGCTGCGCCCCGGCGGCAAGAAGAGTACCGCCT
 GAAGACACCTGGTGTGGGCTCCCGGAGGTGGAGGCTTCGCCCTGACCCCGGCTGCTGGAGACCGCGAGGCTGCCAGAGCTGATGG
 AGAGCTGCAGTCCACCTGAAGCCGCTCCGAGGAGTGAAGTCCCTGTTCAACACCATCCACCTGTGTGCTGCACAGCGCATC
 GACGTGAAGGACACCAAGGAGGCTTGACAAAGTGGAGGAGATGCAACAAGTCCRAGCAGAAGACCCAGCAGCCCGCGACACCGG
 CGGCTCCTCCAAAGTGTCCAGACTACCCCATCGTGCAGAACCCCGGCGAGATGGTGACAGTCCATCTCCCGCCGACCCCTGAAG
 CTTGGGTGAAGTGTATCGAGGAGAAGGCTTCTCCCGAGGTGATCCCATGTTCTCCGCTGTCCGAGGCGCCACCCCGAGGACCTG
 AACATGATGCTGAACATCGTGGCGGCGCACCGCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCGCCGAGTGGACCGGC
 CCACCCCGTGACCGCGGCCCATCCCGCGGCGAGTCGCGAGCCCGCGCTCCGATCGCGGACCCACCTCCACCTGCGAGGC
 AGATCGGTGGATGACCTCAACCCCATCCCGTGGCGGAGATCTACAAGGCTGGATCATCTGGGCTGAACAAGATCGTGCGATG
 TACTCCCGGTGTCATCTGTGACATCCCGAGGCCCCAAGAGCCCTTCCCGACTACGTGACCGCTTCTCAAGTGCCTGGCGCGGA
 GCAGGCCACCCAGAGGTGAAGACTGGATGACCGAGCCCTGCTGGTGCAGAACCCGACCTCCATCTGGAAGGCTTGG
 GCACCGGCGCCACCTGGAGGAGATGATACCGCTGCGGCGTGGCGGCGCTCCCAAGGCGCGCTGCTGGCGGCGCATGTCC
 CAGGCTCCAAACGCGCGCGCATATGATGACAGAGTCCAACTTCAAGGSCCAGCGCCGATCATCAAGTCTTCACTGCGGCAAGGA
 GGGCCACTGCGCGCACTGCGCGCGCCCGCAAGAGGGTGTGTAAGTGCAGAGGCGCACCAAGATGAGGACTGCAAGGAGC
 GCCAGGCCAATTCCTGGGCGGATGTGGCCCTCTCCAAGGCGCGCGGCAACTCTGAGTCCCGCGGAGCCACCGCGCCCCC
 GCGAGTCCCTGGAGATGAAGGAGGAGACCACTCTCTCCCGCAAGCAGGAGCCCGCGCACAGGAGTGTACCCCTGACCTCCCTGAAGTC
 CCTGTTCCGCTCCGACCCCTGTCTCCAGTAA

도면78b

2003_CON_06_CPX_gag.OPT

ATGGGGCCCGGCGCTCCGTGCTGTCCGGCGGCAAGCTGGACGAGTGGGAGAGATCCGCTGCGCCCCCGGGCGCAAGAGAAAGTACCGCTT
 GAAGACCTGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCGCGAGGCTGCCAGCAGATCATCG
 AGCAGTGCAGTCCGCCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTACAAACACCGTGGCCACCTGTACTGCTGACACGCGCATC
 AAGTGACCGACACCAAGAGGCGCTGGACAAGATCGAGGAGATCCAGAACAAATCCAAAGAGGCCAGCAGGCCCGCGCCACCGG
 CAATCCTCCACCTGTCCAGAACTACCCATCTGTGAGACGCCAGGCCAGATGGTGCCACAGGCCATCTCCCCCGCACTGAAACG
 CCTGGTGAAGGTGATCGAGGAGAGGCTTCTCCCGGAGTGTATCCCATGTTCTCCGCTGTCCGAGGGCCACGCCCCAGGACCTG
 AACATGATGCTGAACATCGTGGGCGGCCACCGAGCGCCCATGCGATGCTGAAGACACCATCACGAGGAGCGCCGAGTGGGACCCGGT
 GCACCCGTGCACGCCGCCCATCCCCCGGCCAGATGCGCGAGSCCCCGGCTCCGACATCGCGGACCACTCCACCTTCAGGAGC
 AGATCGGCTGGAAGCTCCAAACCCCATCCCGTGGCGAGATCTAAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGCGCATG
 TACTCCCCGTGTCCATCTGGACATCCGCGAGGCCCCCAAGGAGCCTTCCGCGACTAGTGGACCGCTTCTCAAGACCTCGCGCGCGA
 GCAGGCCACCCAGAGGTGAAGACTGGATGACCGACACCTGTGTGTGGAGACCGCCACCCGACTGCAAGACCATCTGAGGCGCTGG
 GCCCGCGCCACCTGGAGGAGATGATGACCGGCTCCAGGGCGTGGCGGCCCGCCGACCAAGGCCCGCTGTGGCGGAGCGCATGTCC
 CAGGCTCCGGCAACGAGGCGCCCATCATGATGCGAAGTCCAACTTCAAGGGCCCAAGCGCTCCATCAAGTCTTCACTGCGGCAAGGA
 GGGCCACCTGSCCGCAACTGCGCGGCCCGCCGAGAGAGGCTGTGGAGTGGGCAAGGAGGCGCCAGATGAGACTGCAACCGGAGC
 GCCAGGCCAACTTCTGGGCAAGATCTGGCTCCAAACAAGGGCGCCCGGCACTTCTGCAACCGCCCGAGCGCCACCGCCCCC
 GCGAGTCCCTTCGGCTCGGCGAGGAGACCGCCCCCTCCCCCAAGCAGGAGCCCCAAGGAGAGAGCTACCCCCCTGCGCTGAGTCTC
 CCTGTTCGGCAACGACCCCTAA

도면79a

20. 2003 CON 07 BC gag .PEP

MGARASILRGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGILLETSBGCKQIKQLQPALQTGTTELRSLFNTVATLYCVHTEI
 DVRTKEALDKIEEQNKIQOKTQQAKEADGKVSQNYPIVQNLQGMVHQPISPRTLNAWVKVVEKAFSPVLPMSALSSEGATPQDLNTM
 LNTVGGHQAAQIILKDTINEEAAEWDRLEHPVHAGEIAPGOMREPRGSDIAGTTSNLQECIAWMTSNPPVPVGDLYKKRWILLGLNKIVRMYS
 TSILDIKQGEKPEPRDYVDREFFKTLRAEQATQDVKNWMTDTLLVQONANPDCKTILRALGPGASLEEMTACQGVGGPSHKARVLAEMSQT
 STILMQRSNFKGSKRIVKFCNCGKEGHIARNCRAPRKGCWKCGKEGHQMKDCTERQANFLGKIWPSPHKGKRGPNFLQSRPEPTAPPEESFRF
 GEETTPSQQEPIDKELYPLTSLKSLFGNDPSSQ\$

도면79b

2003 CON 07 BC gag.OPT

ATGCGCGCGGCTCCATCCTGCGCGCGGCAAGCTGGACAACTGGAGAAAGATCCGCTGCGCCCCGCGGCAAGACCACTACATGCT
 GAGACACCTGGTGTGGGCTCCCGGAGCTGGAGGCTTGCCCTGAAACCCGCTGTGGACCTCGAGGCTGCAAGCAGATCATCA
 AGAGCTGCAAGCCGCTGAGACCCGACCGAGGAGTGGCTCCCTGTTCAACACCTGGGCCACCTGTACTGCTGCAACCCGAGATC
 GACGTGGCGACACCAAGGAGCCCTGGACAAGATCGAGAGGAGCAGACAAATCCAGCAGAACCCAGAGGCCAAGGAGCCGACGG
 CAAGGTGTCCTCAGAACTACCCATGCTGCAGAACTGCAAGGCCAGATGGTGACCACTCTCCCCGACCTGAACGCTGGGTGA
 AGTGGTGGAGGAGAAAGGCTTCTCCCGAGGTGATCCCATGTTCTCGCCCTGTCCGAGGGGCCACCCCCAGGACCTGAACACCATG
 CTGACACCGTGGGGGCCACCAAGCCGCCATGAGGACACCATCAAGAGGAGGCCCGGAGTGGACCCCTGCACCCCGT
 GCACCGGCCCCCATCGCCCCCGGCGAGATCGGAGACCCCGGCTCGACATCGCGGACCATCTGGCCCTGAAACAGATCGTGCCATGTACTCCCC
 GGATGACCTCCAACCCCGCTGCCGCTGGGCGACATCAACAGCTGATCATCTGGCCCTGAAACAGATCGTGCCATGTACTCCCC
 ACCTCCATCTTGACATCAAGCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGGACCGCTTCTCAACCCCTGCGGCCGAGCAGGCCAC
 CCAGGAGTGAAGACTGGATGACCGACCCCTGCTGTGTCAGAACCCGACCTGCAAGACCATCTGCGGCCCTGGGCCCTGGGCCCGGCG
 CCTCCATCGAGGAGATGATGACCGCTGCGAGGGCTGGCGGCCCTCCCAAGGCCGCTGTGCGCCGAGGCCAATGTCAGACCAAC
 TCCACCATCTGTGAGGGCTCCAACTCAAGGGCTCCAAAGCATCTGAAGTCTCAACTGCGGCAAGGAGGCCACATCGCCGCAA
 CTGCCGCGCCCCCGCAAGAGGCTGTGGAATGCGGCAAGAGGGCCACAGATGAAGGCTGACCGAGGCCAGGCCACTTCTCTGG
 GCAAGATCTGGCCCTCCACAGGGCCCCCGGCAACTTCTGAGTCCCGCCCCGAGCCACCGCCCCCGGAGAGTCTTCTCGCTTC
 GCGAGGAGACCAACACCCCTCCAGAGAGAGAGCCATCGACAGGAGCTGTACCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA
 CCCCTCTCCCAAGTAA

도면80a

21. 2003 CON 08 BC gag .pep

NGRASILGGKLDKWEKTLRPGGKKHYNLKHILWASRELERFALNPGLLETSEGCKQIIKQLPALQTGTTEELRSLFNFTYATLYCVHAEI
 EVRDTKEALDKIEEFQNKIQOKTOOAKEADEKVSONYPIVONLQOMVHQPISPTLNWVKVVEEKAFSPVTPMFTALSEGATPQDINTM
 LNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGFVAPGQMRPRGSDIAGTSTLQEOIGWNTNNPIPVGEIYKRWIILGLNKIVRMYS
 TSILDIKQGPKEPERDYVDRFFETLRAEQATQDVKNWMTDTLLYONANPDCKTILRALGFGASIEEMTACOGVGGPSHKARVLAEMSQT
 NTILMORSNFKGSKRIVKCFNCGKEGHTAKNCRAPRKKCKGCKGEGHQMCKDCTERQANFLGKIWPSHKGRFGNFIQSRPEPTAPPAESFR
 EETTPAKQEPKDRPLETSLRSLFGSDPLSQ\$

도면80b

2003 CON 08 BC gag.OFT
 ATGGGGCCGCGGCTCCATCTCTGGGGGGGCAAGCTGGACAACTGGGGAAGATCCGGCTGCCCTCCGGGGGGAAGAGCACTACATGCT
 GAAGACCTGTGTGGGCTCCCGAGCTGGAGCGCTTCGCCCTGAACCCGGCTGTGGACCTCGAGGCTGCAAGAGATCATCA
 AGCAGCTGAGCCGCTCTGAGACGGCCTGAGAGCTGGCTCCCTCTTCAACCTGGCCACCTGTACTGCTGCAAGCCGAGATC
 GAGGTGCGGACACCAAGAGGCGCTGGACAAGATCGAGGAGAGACAGATCCAGAGAGACCCAGAGGCGCAAGAGGCGGACGA
 GAAGTGTCCAGAACTACCCATCGTGAGAACCTGAGGGCCAGATGGTGACACCCCTGTCCCGGAGGCTGACCCAGGCTGGGTGA
 AGTGGTGGAGGAGAGGCGCTTCTCCCGAGGTGATCCCATGTTACCGCCCTGTCCGAGGGCCACCCCGAGGCTGACACCATG
 CTGAACACCGTGGGGGCGCCAGGCGCATGCGATGCTGAAGACACCATCAAGAGGAGGCGCGGAGTGGGACCGCTGCACCCGCT
 GCAGCGGCGCCCGTGGCGCCCGGCGAGATCGCGAGCCCGCGGCTCCGACATGCGCGGACCATCCACCTGCAAGAGAGATCGTGGCATGTACTCCCC
 GGATGACCAAGAACCCCGCATCCCGTGGCGGAGATCTAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCATGTACTCCCC
 ACCTCATCTGGACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACCTTGGCGCGGAGGAGGAC
 CCAGAGCTGAAGAACTGGATGACCGACACCTGCTGGTGAGAACGCCAACCCCGACTGCAAGACCTTCTGGCGCTGGCGCGGCGG
 CCTCCTGGAGGAGATGATACCGCTGCCAGGCGCTGGCGGCGCTCCACAGGCGCGGTGCTGGCGGCGCATGTCCAGACCAAC
 AACACATCTGATGCGAGCGCTCCACTTCAAGGCTCCAGCGCATCGTGAAGTCTTCACTGGGGAAGAGGCGCAATCGGCGAAGAA
 CTGCGCGCGCCCGCAAGAGGCGTGTGAAGTGGGCAAGAGGCGCACCGAGATGAAGGATGACCGAGGCGCGCAACTTCTCTGG
 GCAAGATCTGGCCCTCCCAAGGCGCGCCCGCAACTTCTGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC
 GAGGAGACCAACCCCGCGCCCAAGCAGGAGCGCGGAGCCCTGACCTCCCTGGCGCTCCGAGCTCGAGCCCGCGCGCGCTTC
 GTAA

도면81a

22. 2003 CON 10 CD gag pep

MGARASVLSGGKLDWEKIRLRPGGKKKYLKHLVWASRELERFALNPGLLETSEGGKQIIGQLQPAIQTGSSEIKSYNTVATLYCVHERI
KVTDIKEALDKIEEEQTKSKKKQAQATADTGNSSQVSNYPVQNQQMVHQPLSPRTINAWKVIIEKAFSPVIMFSAISEGATPODL
NTMLNTVGGHQAAMQMLKETINEFAAEWDRLHPVQAGPVAPGQIREPRGSDIAGTSTLQEQIRWMTSNPPIPVGEIYKRWIILGINKIVRM
YSPVSIIDIRQGPKEPFRDYVDREYKTLRAEQASQDVKNWMTETLLVQANANPDCKTILKALGPAATLEEMMTACQGVGSPSHKARVLAEMS
QATSGNALMMQGNFKGPKRIIKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSNKGKRGPNFLQSRPEPTAPPA
ESTFGEEITPSQKQEQDKELHPLASIKSLFGNDPLSQ\$

도면81b

2003_CON_10_CD_gag.OPT
 ATGGCGCCGGCGCTCCGTGCTGTCGGGGCAAGCTGACAGTGGGAGAAGATCCCCCTGGCCCCGGCGGCAAGAAGAAGTACCGCCT
 GAAGACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGACACTCCGAGGGCTGCAAGAGATCATCG
 GCCAGTGCAGCCCGCATCCAGACCGGCTCCGAGGAGATCAAGTCCCTGTACAAACCGTGGCCACCTGTACTGGTGACAGCGCATC
 AAGGTGACCGACACCAAGGAGCCCTGGACAAGATCGAGGAGCAGACCAATCCAGAGAGGCCCCAGAGGCCACCGCGACACCGG
 CAATCCTCCAGGTGCCAGAACTACCCATCTGCAGAACCTCGAGGCCAGATGGTGCCACAGCCCTGTCCCGCCGCAACCCCTGAACG
 CCTGGGTGAAGGTGATCGAGGAGAGGCCCTTCTCCCGAGGTGATCCCATGTCTCCGCTGTCCGAGGGCCACCCCGAGGACCTG
 AACACCATGCTGAACACCGTGGGGGCCACCCAGCCCGCATGAGATGCTGAAGAGACCATCAACGAGGAGGCCCGCGAGTGGACCGCCT
 GCACCCGTGACGGCCGCCCGTGGCCCCCGGCCAGATCCCGAGCCCGGCTCCGATCGCCGACCTCAACGAGGAGGCCCGCGAGTGGACCGCCT
 AGATCCGCTGGATGACCTCAACCCCGCATCCCGTGGCGGAGATCTACAGGCTGGATCATCTGGGCTGAACAAGATCGTGCGCATG
 TACTCCCGGTGTCATCTGTGGACATCCCGCAGGCCCCCAAGGAGCCCTTCCGCGACIAGTGGACCGGTTCTACAAACCTTCGCGCCGA
 GCAGSCCTCCAGACGTGAAGAACTGGATGACCGAGACCTGTGTGGTGAAGAGCCCGGACTGCAACCATCTCTGAAGGCCCTGG
 GCCGCCCGCCACCTGGAGGAGATGATACCGCTGCCAGGGGTGGGGCCCTCCACAGGCCCGCGCTGCTGCCGAGGCCCATGTCC
 CAGGCCACTCCGGCAACGCCATCATGTGACGCGCGCACTTCAAGGCCCCCAAGAGATCATCAAGTGTCTCACTGGGCAAGAGGG
 CCACATCGCAAGAACTGCGCGGCCCGCCGCAAGAGGCTGTGGAGTGGCGCGGAGGCCACCCAGATGAGGACTGCAACCGAGCGCC
 AGGCCAACTTCTGGGCAAGATCTGGCCCTCCAAAGGCCCGCCCGCAACTTCTGAGTGGCGCCGAGGCCACCGCCCGCCCGCCG
 GAGTCCCTTCGGCTTCGGCGAGGAGATCACCCCTCCCAAGAGGAGCAGAGGAGTGCACCCCTGGCCTCCCTGAAGTCCCT
 GTTCGGCAACGACCCCTGTCCAGTAA

도면82a

23. 2003_CON_11_CPX_gag.gag.pep

gag.pepMGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETAECCQIMCQLQALGCTEELRSLYNTVATL
YCVHRIEYKDTKEALDKIEEIQNKSKQKQQAADTGNSSKVSONYPIVQNAQGMVHQAI SPTLNAWVKVVEKAPSPVETIPMFSALSE
GATPQDLNMLNIYGGHQAMQMLKDTINEEAAEWDVHPVHAGPIPPGQMRERGS DIAGTTSTLQEQIGWMTGNPPVFGELIYRRWILG
LNKIVRMYSPVSIILDIRQGPKEPFDYVDRFFKTLRAEQATQEVKSWMTETLLIQNANPDCKSLRALGPGATLEEMTACQGVGGPGHKAR
VLAEMSQVQQTINIMMQRSNFKGQKRIRKFCNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSSSKGRPGNFLQSRPEP
TAPPAESFGFGEETAPSPKQEPKEKELYPLTSLKSLFGSDPLSQ\$

도면82b

2003_CON 11_CPX_gsg_OPT
ATGGCGCCGCGCTTCGGTGTCTCCGGCGCAAGCTTGSAGCGCTTGGAGAAGATCCGCTCGCGCCCGCGCGCAAGAAGTACACGCCCT
AGGAGCACTGTTGTGGCTTCGCGAGCTGGAGCGTTTCGCCCTGAACCTCCCTTGGAGACCCCGAGGCTCGACAGAGATCATGTG
GCCAGCTGGAGCGCCCTTGGGCAACCGGACCGAGGAGCTCGCTCCCTGTACAACACCGTGGCCACCCCTGTACTCGTGTCCACACCGGATC
GAGGTGAAGCAACTCAAGGAGCTCGTAGAAGATCGAGAGATTCAGAACAAATTCARACAGAAGACAGCAGCGCGCGCCGACACACGG
CAACTCTCCAAAGTCTCCAGAACTCCCACTCGTCAGAACCCGACAGGCGAGATGGTGCAACAGCCATCTCCCCCGCACCCCTCGAACG
CCTTGGGTCAAGTGTGGAGGAAGGCCCTTCTCCCGAGGTGATCCCAATGTTCTCGCCCTGTCCGAGGGGCCACACCCCAAGGACTG
AACATATGTCTGAACATCTGGGCGGCCACCGCGCGCATCGATGATGCTGAAGAGCATCATCAACAGGAGCGCGCGAGTGGACGCGCT
GACACCCCTGTAGCGCGGCCCTCCATCCCGCGCGAGCTCGAGCCCGCGGCTCCGACATCGCGGCACCCATCCAGCTCGACGAG
AGATCCGGCTGGATGACCGGCACACCCCGGTGCGGTGGCGAGATCTACCGCGCTGGATCATCTCGGCCCTGAACAGAATGCTGGCATG
TATATCCCGCTGTCCATCTCTGGACATCCCGAGGCCCAAGAGCCTTCGCGAGCTATGGAGACCCCTTCTTACAGCACTCTCGCGCCGA
GAGGACCACTCAAGAGTGAATCTGTGATGACCGAGCCCTGTGATTCAGAACCAACCCGATCGAAGTCTCTTCGCGCGCCCTGG
GCCCGCGGCCACCTCGAGGAGATGATGACCGCTGCCAGGCGTGGCGGCCCGCCCAAGGCCCGCTGCTGGCCGAGGCCCATGTCC
CAGTGGACAGACCAATCATGTATGACAGGCTCCAACTCTAAGGCGCCAGAGCGCTGGCGGCCCGCCCAAGGCCGAGGGCCACCT
GGCCCGCAATGCGCGGCCCGCGAAGAGGCTGTGGAAGTGTGGCAAGAGGCGCACCATGTAAGGACTGCACCGAGGCGCGAGGCCA
ACTTCTTGGGCAAGATCTGGCCCTCTCCAAAGGCGCCCGCCGGCACTTCTCGATGTCGCCCGCGAGCCCAACCGCCCGCGAGTCC
TTCGCTTGGCGGAGGAGATCGCCCCCTCCCCCAAGAGGAGGCCAAGAGAGGAGCTGTACCCCCCTGTAACCTCCCTGGAAGTCCCTGTCCG
TCCGACCCCTGTCCCACTAA

도면83a

24. 2003 CON 12 BF.gag.pfp

MGARASVLSGGELDRWEKIRLRPGGKKKYLKHIVWASRELERFAVNPGLLETSEGCRKIIGQLQPSLQTSSEILRSLYNTIYVLYFYHOKV
 EVKUTKEALDKLEEQNKSOQKTQQAADKGVSNYPFIVQLQGQWVHQALSPRTLNAWVKVVEEKAFSPETPFMSALSEGATPQDLNTML
 NTVGGHQAAQMQLKDTINEEAAEWDLHPVHAGPTPPGQWREPRGSDIAGTISTLQEQIQWMTSNPPVPVGEIYKRWIILGLNKIVRMYSVP
 SILDIRQGPKEPFRDYVDREFTLRAEQATQEVKGMWTDLLVQANANPDCKTILKALGEGATLEEMNTACQGVGPGHKARVLAEAMSOVTN
 TTVMMQKSNFKGQRRIVKCFNCGKEGHIARNCRAPRKKGCMKCGEGHQMKDCTERQANFLGKIWPSNKGKRPGNFLQNRPEPTAPPAESEGF
 GEEITSPKOEOKDEGLYPPLASLSLFGNDP\$

도면83b

2003 CON 12 BF.gag.OPT
 ATGGGCGCCGCGCTCCGTCTGTCCGGCGCGAGCTGGACCGCTGGAGAGATCCGCTGCGCCCGCGGCAAGAGAGATACCGCT
 GAAGCAATCGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCGGTGAACCCGCGCTGCTGGAGACCTCGAGGGCTGCCGCAAGATCAICG
 GCCAGTGCAGCCCTCCCTGAGACCGGCTCCGAGAGCTGCGCTGCTTACACACCATGCGCGTGTCTACTCTGTGACACAGAGAGTG
 GAGTGAAGGACACCAAGGAGGCCCTGGACAAGCTGGAGGAGGAGAGACAAGTCCACAGAGACCCAGCAGCGCGCGGACCAAGGG
 CGTGTCCAGAACTACCCCATCGTGCAGAACTGCAGGCCAGATGGTGCACAGGCCCTGTCCCGCCACCCCTAACGCCCTGGGTGAAGG
 TGGTGGAGGAGAAGGCTTCTCCCGAGGTATCCCATGTTCTCCGCTGTCCGAGGGCGCACCCCGCAGGACCTGAACACCATGCTG
 AACCGTGGGCGGCGCACAGCCGCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGCGCGGAGTGGACCGCTGCACCCCGTGA
 CGCGGCGCCATCCCGCCGAGATGGCGAGCCCGGCTCCGACATCCCGGACACCTCCACCTGCAGGAGCAGATCCAGTGA
 TGACCTCAACCCCGCTGCGCGGAGATCAACAGCGCTGGATCCTGGGCTGAACAAGATCGTGGCATGTACTTCCCGCTG
 TCCATCTGGACATCCGCGAGGCCCGCAAGAGCCCTCCGCGACTACGTGGACCGCTTCTTCAAGACCTGGCGGCCGAGCAGGCCACCCA
 GGAGGTGAAGGCTGGATGACCGACACCTGCTGGTGCAGAACCGAACCCGACATGCAGACCATCTTGAAGGCCCTGGGCCCGGCGCCA
 CCTGGAGGAGATGATGACCCCTGCGAGGCGTGGCGGCGCCCGGCGCCAGGCCCGGCTGGCGAGCCATGTCCAGGTGACCAAC
 ACCACCGTGTATGAGAGTCCAACTTCAAGGGCCAGCGCCGCTCGTGAAGTCTTCAACTGGGCAAGGAGGCCACATCGCCAGAA
 CTGCCGCGCCCGGCAAGAGGCTGCTGGAACTGCGCGCGGAGGGCCACCAATGAAGGACTGCACGAGCGCCAGGCCCACTTCTCTGG
 GCAGATCTGGCCCTCAACAGGCGCGCCCGGCAACTTCTCTGAGAACCGCCGAGCCACCGCCCGGAGTCCCTTCGGCTTC
 GCGGAGGAGATCACCCCTCCCGCAAGCAGGAGCAGAAGGAGGAGGCCCTGTACCCCGCTGGCTCCCTGAAGTCCCTTCCGCAACGA
 CCCCTAA

도면84a

25. 2003 CON 14 BG gag. pep
 MGRASVLSGGKLDWEKIRLRPGGKKYRMKHLVWASRELERFALNPDLLETAEGCQIMGQLQPALQTGTETIRSLFNTVATLYCVHQKI
 EVKDTKEALEEVEKXQKKSQKKQQAAMDEGNNSQASQNYPIVQNAQGMVHQALSPRTINAWKVVEEKAESPVPIPMFSALSREGATPQDLN
 TMLNTVGGHQAAMQWLKDTINEEAAEWDRMHPQQAQPIPPGQIREPRGSDIAGTSTLQEQIRWMTSNPPIFVGEIYKRWIILGLNKIVRMY
 SPVSLDIDRQGPKEPFRDYVDRFEFTLRAEQATQEVKGWMTDTLLIVONANPDCKTILRALGPGATLEEMMTACQGVGGFESHKARVLAEMSQ
 ASGATIMQKSNFKGPRRNKCFNCGKEGHLARNCRAPRKKGCWKCKEGHQMKDCTESKANFLGKIWPSNKGREGNFIQNRPEPTAPPAES
 FGFGEIAPSPKOEPEKEIYPLASLKSIFGSDPSSQ\$

도면84b

2003 CON 14 BG gag.OPT.
 ATGGGCGCCTGGGCTCCGTGCTGTCCCGGGCAAGCTGACGCGCTGGAGAGATCCGCTGCGCCCGGGCGCAAGAGAGTACCCCAT
 GAAGCACCTGCTGTGGGCTCCCGGAGCTGGAGGCTTCGCCCTGAACCCCGACTGCTGGAGACCGCGGAGGCTGCGAGCAGATCATGG
 GCCAGCTCAGCCCGCTGAGACCGGACCGAGAGATCCGCTCCCTGTCAACACCGTGGCCACCCCTGTACTGCTGCACCCAGAGATC
 GAGGTGAAGGACACCAAGAGGCTTGGAGAGTGGAGAGGAGGAGAGTCCCAAGAGAGCAGAGGCGCCATGGACGAGGGCAA
 CAATCCCGAGGCTCCAGAGACTACCCCATGCTGCAGAACGCCAGGCGGAGATGCTGCACACGAGGCTTCGCCCTGACGAGGCTTGAAC
 GGGTGAAGTGTGGAGAGAGGCTTCTCCCGAGTGTATCCCATGTTCTCGCCCTGTCCGAGGCGCCACCCCGAGGACCTGAAC
 ACCATGCTGAACACCGTGGGCGCACCGCGCATGAGATGCTGAAGGACACCATCAAGGAGGAGCGCGGACACCCCTCCACCTGCGAGGAGAG
 CCCCCAGAGCGGCCCCATCCCCCGGCGAGATCCGCGAGCCCCGCGCTCCGACATCGCGGACACCCCTCCACCTGCGAGGAGAG
 TCCGCTGGATGACCTCAACCCCCCATCCCGTGGCGAGATCTACAAGGCTGGATCATCTGGGCTGAACAGATGCTGCGCATGATC
 TCCCCGTGTCATCTTGACATCCGACAGGCCCCAAGGACCCCTCCGCGACTACGTCGACCTTTTCAGACCTGCGGCGCGAGCA
 GGCCACCCAGAGGTGAAGGCTGGATGACCGACACCTGCTGTGTCAGACGCAACCGGACTGCAAGACCATCTCGGCGGCTGGCC
 CCGGCGCACCTGGAGGAGATGATGACCGCTGCGAGGCTGGGCGGCCCCCTCCACAGGCGCGGCTGCTGGCCGAGCCATGTCCAG
 GCCTCGGCGCGCACCATCATGATGACAGAGTCCAACTTCAAGGCGCGCGCGGACCATCAAGTGTCTTCACTGGGCGAGGAGGCGCC
 GGCCCGCAACTGCCCGCCCCCGGAGAGAGGCTGCTGGAGTGGGCGAGGAGGCGCCACAGATGAAGGACTCACCGAGTCCAGGCTCA
 ACTTCTGGGCAAGATCTGGCCCTCAACAGGCGCGCGCGCAACTTCTGAGAACGCCCCGAGCCACCCCGCCCCCGGAGTCC
 TTCGCTTCGGGAGGAGATCGCCCTCCCCCAAGCAGAGCCCCAAGAGAGATCTACCCCTTCCCTGAAGTCCCTGTTCGG
 CTCCGACCCCTAATCCAGTAA

도면85a

31. 2003 CONS nef.PEP
MGGKWSKSSIVGWPAVRERIRRTPRAAEGVGAVSQDLDKHGAITSSNTAATNADCAWLEAQAQEEVGFVVRPOVELRPMYKGAFDLSHFYK
EKGLDGLIYSKKRQELIDLWVYHTQGYFFPDWQNYTFGPGIRYFPLTFGWCFKLVPVDPPEVEEANEENNCCLGHPMCQHGMEDEOREVLMWK
FDSRLAARHTARELHPEFYKDC\$

도면85b

203 CONS nef .OPT
 ATGGGCGGCAAGTGGTCCCHAGTCCATCGTGGGCTGGCCCGCGCGTGGCGGCGCATCCGCGGCACCCCCCGCCCGAGGGCGTGGG
 CGCCGTGTCCAGGACCTGGACAGCACGGCGCCATCACTCTCCACACCGCGCCGCCACAGCCGACTGGCCTGGCTGGAGGCCGAGG
 AGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCGAGGTGGCCCGCATGACCTACAGGGCGCTTGGACCTGTCCCACTTCCCTGAAG
 GAGGAGGGCGGCTGGACGGCCCTGATCTACTCCAGAGAGCGCCAGGAGATCCTGGACCTGTGGTGTACACACCCAGGGCTACTTCCCCGA
 CTGGAGAACTACACCCCGGCGCGGCACTCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGGCCCTGGACCCCGAGGAGGTGG
 AGGAGGCCAACGAGGCGAGAACAACTGCTGTGCACTGTCAGCCAGCATGGAGCAGGACCCGCGAGGTGCTGATGTGGAG
 TTCGACTCCCGCCCTGGCCCTGGGCCACATCGCCCGGAGCTGCACCCCGAGTTCTACAGGACTGCTAA

도면86a

32. 2003 M. GROUP.anc nef.pap
MGGKWSKSIIVGWPAVYRERMRTAPAAEGVAVSODLDKHGAIITSSNTAATNADCAWLEAQEEEEVGFVRPOVELRPMYKAAFDLSHFLK
EKGGLDGLIYSKKRQELLDLWVYHTQGYFFDWQNYTPGPGIRYPLTFGWCFKLVFVDPPEVEEANEENNCCLLHFMCGMEDEEREVINMK
FOSRLALRHIARELHFEFFYKDC\$

도면86b

2003_M_GROUP.anc nef.OPT
 ATGGGCGGCAAGTGGTCCCAAGTCCTCCATCGTGGCTGGCCCGCCGCTGCGGAGCGCATGCGCCGACCGCCCCCGCCCGCCCGAGGGCGGTGGG
 CGCCGTGTCCAGGACCTGGACAGCACGGGCGCCATCACCTCTCCAAACCGCCGCCACCAAGCCCGACTGCGCCTGGCTGGAGGCCCGAGG
 AGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCGCCAGGTGCCCGCCATGACCTACAGGGCGCCTTCGACCTGTCCACTTCCCTGAAAG
 GAGAGGGCGGCTGGACGGCTGATCTACTCCAGAAGGCCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCGGA
 CTGGCAGACTACACCCCGGCGCATCCGCTACCCCGTGACCTTCGGCTGGTCTTCAAGCTGGTGGCCCGTGGAGCCCGGAGGAGTGG
 AGGAGGCCAACGAGGGCGAGAACAACTGCCGTGCTGCACCCCATGTGCCAGCACGCAATGGAGGAGGAGCGGAGGTGCTGATGTGGAG
 TTCGACTCCCGCTGGCCCTGGGCCACATCGCCCGGAGCTGCACCCCGAGTTCTACAGGACTGCTAA

도면87a

33. 2003 CON A nef. PEP
MGCKWSKSIYGVWPDIREIRIRTPPAKGVGAVSQDLKYGAVTINNATQASCAMLEAQEEEEVEGFPVPPQVPLRPMTFKGAFDLSFFL
KEKGLDGLIYSOKRQEIIDLWVYNTQGYFPDQNYTEPGGTNPPLTEGWCFKLVVDDEVEATEGENNCLLHFICQHGMDDEKEVLMW
KFSRLARRHIALEMHPFYKDCS

도면87b

2003 CON A nef .OPT

ATGGGGGCAAGTGTCTCCAGTCTCTCCATCGTGGGTGGCCCGACATCCCGAGCGCATCCGCGCACCCCGCGCCCAAGGGCGTGGG
CGCCGTGTCCAGGACCTGGACAACTAGTACGGCGCCCTGACCATCAACACACCGCGCCACCCAGGCCCTCTCTGCCCTGGAGGCCCCAGG
AGGAGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCGCCAGTCCCGCTGGCGCCCATGACCTCAAGGGCGCTTCGACCTGTCTCTTCTCTG
AAGGAGAAGGCGCGCTGGAGGCGCTGATCTATCTCCAGAAAGCGCCAGGAGATCTGGACCTGTGGGTGTACACACCCAGGCGTACTTCCC
CGACTGGCAGAACTACACCCCGCGCCAGCCGCTTCCCGCTGACCTTCGGCTGTGCTTCAAGCTGGTGGCCCGTGGACCCCGACGAGG
TGGAGGAGGCCACCGAGGGGAGAACTGCTGCTGACCCCATCTGGCAGCACGGCATGGAGGAGAGGAGGTGCTGATGTGG
AAGTTGCACTCCCGCTGGCGCCGCGCCACATCGCCCTGGAGATGACCCCGAGTCTTCAAGGACTGCTAA

도면88a

34. 2003 CON A1 nef.PEP
MGKWKSSIVGMPEYVERMRTPPAATGVGAVSQDLDKHGAVTSSNINHFSCVWLEAQEEEEVGFPVRQVPLRPMYKCALDLSHFLKEK
GGLDGLIYSRKNQEIILLWVYHTQGYFEDWQNTPGGIRYPLTFGWCFKLVVPVDEVEKATEGENNSSLHPICQHGMDDEEREVLKWKFD
SRLAKHRAQELHPEFYKCS

도면88b

2003 CON A1 nef. Opt
 ATGGCGGCAAGTGGTCCAGTCTCTCCATCGTGGGCTGGCCCGAGGTGCGCGGCAATGCGCGCAACCCCGCGCGCCACCGGCGTGGG
 CGCCGTGTCCAGGACCTGGACAAGCACGGCGCCGTGACCTCTCCACATCAACACCCCTCCTGCGTGTGGCTGGAGGCCAAGGAGGAGG
 AGGAGTGGGCTTCCCGTGGCGCCCGCCAGGTGCCCTGGCCCGCATGACCTACAGGGGCCCTGGACCTGTCCCCTCTGAAAGGAGAAG
 GGCGGCTGGACGGCTGATCTACTCCCGAAGCGCCAGGAGATCCTGGACCTGTGGGTACACACCCAGGCTACTTCCCCTGACTGGCA
 GAATACACCCCGGCGCCGCTACCGCTACCCCTGACCTTCGGCTGGGTTCAGCTGGTGGCCCTGGACCCCGAGGAGGTGGAGAAG
 CCACGAGGCGGAGAACAACCTCCCTGCTGACCCCATCTGCCAGCACGGGATGGACGACGAGGAGCGCGAGGTGCTGAAGTGAAGTTCGAC
 TCCCGCCTGGGCGCTGAAGCACCGCGCCCGAGGAGCTGCACCCCGAGTTCTACAGGACTGCTAA

도면88c

35. 2003 A1.anc nef.pgp
MGKWSKSSIVGWPEVREMRRTPEAAKGVGAVSODLDKKGAVTSSNTAANNPCAWLEAQEEEEVGFPVRPQVPLRPMTYKGAFDLSHFLK
EKGLDGLIYSKKRQEILDWVYHTQGYFPDWQNYTPGPGIRYPLITFGWCFKLVVDPAEVEERATEGENSSLHPICQHGMDDEEREVIMWK
FDSRLALKHFARELHPFEFYKDC\$

도면88d

2003 A1.anc nef.OPT

ATGGCGGCAAGTGGTCCCAAGTCCATCTGGCTGGCCCGGAGGTGGCGGACGCGATGGCGGCAACCCCGCCCAAGGGCGTGGG
CGCGTGTCCAGGACCTGGACACGACGCGCGCTGACCTCCTCCAAACCGCGCGCAACACCCCGGCTGGCCCTGGCTGGAGGCCAGG
AGGAGGAGGAGGTGGCTTCCCGTGGCGCCCGGAGTGGCCCTCGCGCCCATGACCTACAGGCGCCCTCGACCTGTCCCACTTCCCTGAAG
GAGAAGGGCGGCTGGACGGCTGATCTACTCCAGAGACGCGCCAGGAGATCCYGGACCTGTGGTGTACCAACCCAGGGCTACTTCCCGA
CTGGCAGACTACACCCCGGCGCCGGCATCCGCTACCCCTGACCTTGGCTGGCTGGCTCAAGTGTGCGGTGGACCCCGCCGAGGTGG
AGGAGGCCACCGAGGGCGAGAACAACTCCCTGTGACCCCATCTGCCAGCAGGCATGGACGAGGAGCGGAGGTGCTGATGTGGAAG
TTGGACTCCCGCCTGGCCCTGAGCACCGCGCGCGAGCTGCACCCGAGTTCTACAAGGACTGCTAA

도면89a

36. 2003 CON A2 nef .pep
MGKTKSKSIVGWPAIRERNKRTEPAAEGVAVSODLATRGAVTSNTAATNPDCAWLEAQEEEEVGFVRPOVPLRPMTFKGAFDLSHFL
KEKGLDGLLYSQKQDILDLWVYHTQGYFPDQNTPGPCTRYPLTFGWCFLVPVDPSEVEERATEGENNSLLHPICQHGIEDPPEREVLRW
KFDSRLALRRARELHPEFYKDC\$

도면89b

2003 CON A2 nef.OPT

ATGGCGGCAAGTGTCCAAAGTCCTCCATGCTGGCTGGCCCGCCATCCGGAGCGCATGCGCAAGCGAACCCCGCCCGCGAGGCGGT
GGGCGCCGTGTCCAGGACCTGGCCACCCGGGGCCGTGACCTCTCCAAACACCGCCGCCACCAACCCGACTGCGCTGGCTGGAGGCCCC
AGGAGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCGCCAGGTGCCCTGCCCTGACCTTCAAGGGGGCTTGACCTGTCCACTTCTTG
AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCCAGAGCGCCAGACATCCTGGACCTGTGGGTGTACCAACCCAGGGCTACTTCCC
CGACTGGCAGAACATACACCCCGGCCCGCCACCGCTACCCCTGACCTTCGGCTGGTCTCAAGCTGGTCCCTGGACCCCTCCGAGG
TGGAGGAGGCCACCGAGGGCGGAGAACATCCCTGCTGCAACCCCATCTGCCAGCACGGCATCGAGGACCCCGAGCGGAGGTGTGCGCTGG
AAGTTCGACTCCCGCTGGCCCTGGCGCCACCGGGCCCGAGCTGCAACCCCGAGTTCACAAAGGACTGCTAA

도면90a

37. 2003 CON B nef.pgp
MGKKSKRSVVGWETVREHMRRAEPAADGVCVSRDLEKHGAISSNTAANNADCAWLEAQEEEEVGFEVRPQVPLRENTYKCALDLSHFLK
EKGGLGLTYSQKQDILLDWVHTQGYFFDWQNYTPGPGIRYELTFGWCFKLVVEPEKVEEANEENNSLLHPMSLHGMDPPERELVVK
FDSRLAEHHWARELHPEYYKDC\$

도면90b

203 CON-B nef.OPT
 ATGGGCGGCAAGTGGTCCGAAGGCTCCGTGGTGGCTGGCCACCCGTGGCGGAGCGCATGCCGCGCGGAGCCCGCCGCCGACGGCGTGGG
 CGCGGTGTCCCGGACCTGGAGAGCAGCGGCCATCACTCTCTCCAACACCGCGCCCAACAGCCGACTGGGCTGGGCTGGAGGCCCCAGG
 AGGAGGAGGAGTGGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTAAGGGCGCCTGGACCTGTCCCACTTCTCTGAAG
 GAGAAGGGCGGCTGGAGGGCTGATCTACTCCAGAGCGCCGAGGACATCTGGACCTGTGGTGTACCCACCCAGGGCTACTTCCCCGA
 CTGGCAGAACTACACCCCGGCGGCTCCGCTACCCCTGACCTTGGCTGGTGTCTCAAGCTGGTCCGCGGAGCCCGAGAGGTGG
 AGGAGGCCCAACGAGGGCGAGAACCTCCCTGCTGCACCCATGTCCCTGACGGCATGGACGCCGCCGAGGCGGAGGTGCTGGTGTGGAG
 TTTCGACTCCCGCCTGGGCTTCCACACATGGCCCGGAGCTGACACCCGAGTACTACAGGACTGCTTA

도면90c

38. 2003_B_and_nef.pep
MGGKWSKSSMGGWPAYVERMKRAEPHADGAVSRDLEKHGAI TSSNTAATNADCAWLEAQQEEEVGFVRPQVPLRPMYKAAIDLSHFLK
EKGGLEGLIYSQKRQDILDWVYHTQGYFPDWQNTFPGGIRYPLTFGWCFKLVEPEPEKVEEATEGENNSLLHFMQCQHGMDDPEKEVLWVK
FDSRIAFAHHMARELHPEYYKDC\$

도면90d

2003 B.anc nef.Opt
 ATGGGCGGCAAGTGGTCCAAAGTCCTCCATGGGCGGCTGGCCGCGCCCTGCGGAGCGCATGAAGCGCGCCGAGCCCGCCGACGGCGGTGGG
 CGCCGTGTCCCGGACCTGGAGAGCACGGGCCCATACCTCTCTCAACACCGCCGCCACCAAGCCCGACTGCGCCTGGCTGGAGGCCCGAGG
 AGGAGGAGGAGGTGGGTTCOCGTGGCCGCCAGGTGCCGCCCATGACCTACAGGCGCCCTGGACCTGCCACTTCCCTGAAG
 GAGAAGGCGGCTGGAGGCGCTGATCTACTCCCAAGAGCGCCAGACATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCGGA
 CTGGCAGAACTACACCCCGCCCGGCATCCGCTACCCCTGACCTTGGCTGGTCTTAAGTGGTCCCGTGGAGCCCGAGAAGGTGG
 AGGAGGCCACCGAGGGGAGAACACTCCCTGCTGCAACCCCATGTGCCAGCAGGCTGGACGACCCCGAGAGAGGAGGTCTTGGTGTGGAG
 TTCGACTCCCGCCTTCCACCACATGGGCGGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

도면91a

39. 2003 CON 02 AG nef pep
MGKKWSKSSIVGWPKVREIRIQTEPPAATCGGAASQDLDRHGAISSNTAATNADCAWLEAOEEEEVGFPVRPOVPLRMTYKZAAVDLSHFLK
EKGGLGLIYSKKRQEIILDLWVYHTQFFPDWQNTTPGGTRFPLTFGWCFKLVPMDFAEVELEANEGENNSLLHFPICQHGMEDEDEVLVWR
FDSSIAFKHRARELHPEFYKCS

도면91b

2003 CON 02 Ag nef. OPT
 ATGGGCGCAAGTGGTCCAAAGTCCCTCCATCGTGGGCTGGCCCAAGGTGGCGAGCGCATCCGCCAGACCCGCCGCCACCGGGGTGGG
 CGCGGCTCCCAAGGACCTGGACCGCCACGGCGCATCACCTCCCTCCACACCCGCCCAACAGGCCGACTGGGCCCTGCTGAGGCCCCAGG
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCCCAGGTGCCCCCATGACCTACAGGCCGCGTGGACCTGTCCCACTGTCCCACTTCTCTGAAG
 GAGAAGGGGGCTGGAGGCTTATCTACTCCAAGAAAGCCCGAGGATCCCTGGACCTGTGGGTGTACCAACCCAGGCTTCTTCCCCGA
 CTGGCAGAACTACACCCCGGCCCGGACCCGCTTCCCCCTGACCTTCGGCTGGTGTCAAGCTGGTGGCCATGGACCCCGCCGAGGTGG
 AGGAGGCCAACGAGGGCGAGAACACTCCCTGTGTGACCCCATCTGCCAGCACGGCATGGAGACGAGACCGCGAGGTGCTGGTGTGGCGC
 TTGGACTCCTCCCTGGGCTTCAAGCACCGCGCGCGAGCTGCACCCCGAGTTCTACAAAGGACTGCTAA

도면92a

40. 2003 CON C nef . PEP
MGKWSKSSIVGMPAVRERIRRTTEPAAGYCAASODLDKHGALTSSNTATNNADCAWLEAQEEEEVGFVRPOVPLRPMTYKAAFDSLFFL
KEKGLEGLIYSKKRQEILDWVYHTQGYEPFDWQNTTFCGCVRYELTFGWCCKLVDPREVEEANEENNCILHPMSQHGMEDEDDREVLKW
KFDSHLARRHMARELHPFYKDC\$

도면92b

2003 CON C nef .OPT
 ATGGGCGGCAAGTGGTCCAACTCCTCCATCGTGGGCTGGCCGCGCGGTGGGAGCGCATCCGCGCACGAGCCCGCCGAGGGCGGTGGG
 CGCCGCTCCAGGAGCTGGACAGACGGGCGCTGACCTCTCCAAACACCGCCACCAACGCCGACTGCCCTGGAGGCCCGAGG
 AGGAGGAGGAGGAGTGGGCTTCCCGTGGCCCGCCAGGTGCCCTGGCCCGCATGACCTACAGCCGCTTCGACCTGCTCTTCTCTG
 AAGGAGAAGGGGGGCTGGAGGCGCTGATCTACTCCAAGAGCGCCAGGATCCTGGACTGTGGGTGTACCAACCCAGGGCTACTTCCC
 CGACTGGCAGAACTACACCCCGGCGCTGGCTACCCCTGACCTTCGGTGGTGTCTCAAGCTGGTGGCCCGGACCCCGCGAGG
 TGGAGGAGGCCAACGAGGGCGGAGAACACTGCTGCTGCAACCCCGATGTCCAGCACGGCATGGAGGACCGCGAGGTGCTGAAGTGG
 AAGTTCGACTCCCACTGGCGCGCGCCACATGGCCCGGAGCTGCAACCCGAGTACTACAAGGACTGCTAA

도면92c

41. 2003 C.anc nef.pap
MGKWSKSSIVGWPAVRERMRTEPAAGVGAASQDLDKHGALTSSNTAANNACAWLEAQEEEEVGFPVRPQVPLRPMTYKAAFDLSEFL
KEKGLDGLIYSKKRQELLDIMVYHTQGYFFDWQNYTPGFGVRYPLTFGRCFKLVPVDPREVEEANEENNCLLHPMSQHGMEDDREVIKW
KFDSHLARRHVARLHPYYKDC\$

도면92d

2003 C.anc nef.Opt

ATGGCGGCAAGTGGTCCAGTCTCTCCATCGTGGGTGGCCCGCCGTGGCGAGGCAATGGCGCCGACCCGAGCCCGCGAGGGCGTGGG
CGCCGCTCCGAGGACCTGGACAAGCACGGCGCCCTGACTCTCTCAACACCGCGCCCAACACGCCGACTGGCTGGAGGCCCCAGG
AGGAGGAGGAGGAGGTGGGCTTCCCGCTGCGCCCCCAGGTGCCCCCTGGGCCCCATGACCTACAGGCGGCTTCGACCTGCTCTTCTG
AAGGAGAAGGCGGCGCTGGACGGCTGATCTACTCCAGAAGCGCAGGATCCTGGACTGTGGGTGTACCAACCGAGGGTACTTCCC
CGACTGGCAGAACTACACCCCGGCGGTGGCTACCCCTGACCTTCGGTGGTCTCAAGCTGGTGGCCCTGGACCCCGCGAGG
TGGAGGAGGCCAACGAGGGGAGAACACTGCTGTGACCCCAATGTCCCAGCHGGAAGGACCGCGAGGTGCTGAAGTGG
AAGTTCGACTCCCACTGGCCCGCGCCACATGGCCCGGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

도면93a

42. 2003 CON D nef.pgp
MGKWSKSSIVGWPALRERIRRTTEPAADGCVASRDLEKHGAI TSSNTAATNADCANLEAQEEDVEVGFVVRPOVLRPMYKAAALDLSHFL
KEKGGLEGLVWSQKHQEI LLDLWVYNTQGFFPDQNYTPGPGIRYPLITFGWCFELVPDPEEVEEATEGENNCLLHPMCQHGMEDPEREVLW
RNSRLAFEHKARVILHPEFYKDC\$

도면93b

2003 CON D nef .OPT
 ATGGGCGGCAAGTGGTCCAGTCTCTCCATCGTGGGCTGGCCGCCCATCCGGGAGGCGCATCCGCGCACCGAGCCCGCGCCGACGGCGTGGG
 CGCCGTGTCCCGGACCTGGAGAGACGGCGCCCATCACTCTCCAAACCGCGCCGACCAAGCCCGACTGCGCTGGCTGGAGGCCCCAGG
 AGGAGACGAGGAGGTGGCTTCCCGTGGCGCCCGCAGGTGCCCTGGGCCCCCATGACCTACAGGCCGCCCTGGACCTGTCCACTTCCTG
 AAGGAGAAGGCGGCTGGAGGCTGTGTGTGCCAGAGCGCCAGGATCTGGACCTGTGGGTACACACCCAGGCTTCTTCCC
 CGACTGGCAGAACTACACCCCGGCCCTACCGCATCCGCTACCCCCCTGACCTTCGGCTGGTCTTCGAGTGGTCCCGGACCCCGAGGAGG
 TGGAGGAGGCCACCGAGGGCGAGAACACTGCTGTGCACCCCATGTGCCAGCAGGATGGAGGACCCCGAGGCTCTGATGTGG
 CGCTTCAACTCCCGGCTGGGCTTCGAGCACAGGCGCGGTGCTGCACCCCGAGTTCTACAGGACTGCTAA

도면94a

43. 2003 CON F1 nef .PEP
MGKWSKSSIVGWPVVRERMRPTPEAAGVCAVSODLERCAITSSNTGATNPDIAWLEAQEEVGEFVRPQVPLRPMYKGAVDLSHFLK
EKGGLEGLIYSKKRQBIIDLWVIHTQGYFPDWQNYTPGCGIRYPLTFGWCFLVPEDEEVERKANEGENNCLLHPMSQHMEDEDEREVLIWK
FDSRLALRHIARERHPEFYQD\$

도면94b

2003 CON F1 nef.OPT
 ATGGGCGGCAAGTGGTCCCAAGTCCATCGTGGCTGGCCCCGCGGTGGCGGCGCATGCGCCCCACCCCCCGCGCGAGGGCGTGGG
 CGCGGTGTCCCAAGACCTGGAGCGCCGGCGCCATCACTCCTCCACACCGGGGCCACCAACCCCGACCTGGCTGGAGGCCGAGG
 AGGAGGAGGAGGTGGGCTTCGCCGTGGCCCCCAGGTGCCCTGGCGCCATGACCTACAAGGGCGCGGTGGACCTGTCCCACTTCCCTGAAG
 GAGAAGGGCGGCTGGAGGGCCTGATCTACTCCAAGAGAGGGCCAGGAGATCCTGGACCTGTGGGTGTACCAACACCCAGGGCTACTTCCCCGA
 CTGGCAGACTACACCCCGGGCCCCGGCATCCGCTACCCCTGACCTTGGGCTGGGTTCAGCTGGTGGCCCCCGGAGGAGGTGG
 AGAGGGCCAAACGAGGGCGAGAACACTGCTGCTGCACCCCATGTCCAGCACGGCATGGAGGACGAGAACCCGGAGGTGCTGATCTGGAAG
 TTCCACTCCCGCCTGGCCCTGGGCCACATCGCCCGCGAGGGCCACCCCGAGTTCTACAGGACTAA

도면95a

44. 2003_CON_F2_nef.pep
MGGKWSKSSIVGMPPTIRERIRRTPTVAAEGVAVSQDLIDKHGAIITSSNTRATNADLAWLEAQDEDEVGFFVVRPQVPLRPMYKAAFDLSHFLK
EKGGLGLIYSKKRQELIDLVYHTOGYFPDQNYTFPGGTRYPLTFGWCFKLVPDPEEVEKANEKENNCLLHPMSLHGMEDEDEVLKWK
FDSRLALRHTARERHPEYYKD\$

도면95b

2003 CON F2 nef.OPT
 ATGGGGGCAAGTGGTCCAAATCCTCCATCGTGGGCTGGCCACCATCCGGAGCGCATCCGCCGACCCCGTGGCCCGAGGGCGTGGG
 CGCCGTGTCCAGGACCTGGACAAGCAGGGGCCATCACCTCTCCAACACCGCGCCACACGCCGACCTGGCTGGCTGGAGGCCGAGG
 AGGACGAGGAGGTGGGCTTCCCGTGGCGCCCGCAGGTGCCCTGGCGCCCATGACCTACAGGCGCCCTTCGACCTGTCCACTTCCCTGAAG
 GAGAAGGGCGGCTGGAGGGCTGATCTACTCCAAGAAGCGCCAGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCGGA
 CTGGCAGAACTACACCCCGGCCCGGCACCGCTACCCCGTACCTTGGCTGGTTCAGTGGTGGCCCGTGGACCCCGAGGAGTGG
 AGAAGGCCAACGAGGCGAGAACACTGCCTGTGACCCCATGTCCCTGCACGGCATGGAGACGAGGACCCGCGAGGTGCTGAAGTGGAG
 TTCGACTCCCGGCTGGGCCCTGGGCCACATCGCCCGGAGGCCACCCCGAGTACTACAGGACTAA

도면96a

45. 2003 CON G nef . pep
MGKWSKSEIVGWEVREIRIOTPPAAEGYGAVSODLARHGAITSSNTAANNPCAWLEAQEEDSEVGFPVRPOVPLRPMTYKGAFDLSFFL
KEKGLDGLIYSKKRQDILDWVYNTQGFEDWQNYTPGPGTRFELTFGWCEKLVPMDEAEVEEANKGENNSLLHPICQHGMEDEDREVLVW
RFDSSLARRHIAREIHPEYKDC\$

도면96b

2003 CON S nef.OPT
 ATGGCGGCAAGTGGTCCAAAGTCCATCGTGGGCTGGCCCGAGGTGCGCGAGCATCCGCCAGACCCCCCGCCGAGGGCGGTGG
 CGCGTGTCCAGGACCTGGCCCCGACGGGCCATCAGCTCCTCCACACGCGCGCCAAACACCCGACTGGCCTGGCTGGAGGCCCAGG
 AGGAGGACTCCGAGGTGGCTTCCCGTGGCGCCCGAGTGGCCCTGCGCCCAAGACTACAGGGCGCCTTCGACCTGTCTTCTCCTG
 AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCCAGAAGCGCCAGGACATCCTGGACCTGTGGGTACACACCCAGGGCTTCTTCCC
 CGACTGGCAGAACTACACCCCGGGCCCCGGCACCCTTCCCCGTGACCTTCGGCTGGTGTCAAGCTGGTGCCCATGGACCCCGCGAGG
 TGGAGGAGGCCCAACAAGGCGGAGAACACTCCCTGTGACCCCATCTGCCAGACGGCATGGAGGACCGCGAGGTGCTGTGTGG
 CGCTTCGACTCCTCCCTGGCCCGCGCCACATCGCCCGGAGTGTGACCCGAGTACTACAGGACTGCTAA

도면97a

46. 2003 CON H nef.PEP
MGGKWSKSSIGGWPATREIRRAEPAAEGVAVSRDIDRRGAVTINNTASTNPDSAWLEAQEEEEVGFEVRPQVPLRPMYKGAFDLSHFL
KEKGLEGLIYSKKRQELDLWVYNTQGYFEDWQNTPGGERYPITFGWCFKLVDPDQVEKANEKENNSLLHPICQHGMEDEEREVLMW
KFD SRLAFRHIARELHPEFYKDC\$

도면97b

2003_CON_H_nef.OPT
 ATGGGCGGCAAGTGGTCCCAAGTCCTCCATCGGCGGCTGGCCGCCATCCGCCGCCGAGCCCGCCGAGGGCGGTGGG
 CGCCGTGTCCCGGACCTGGACCGCGCGGCGCCGTCACCATCAACAACCCGCTCCACCAACCCGACTCCGCTGGCTGGAGGCCCAGG
 AGGAGGAGGAGGAGGTGGGCTTCCCGTGGCCCGCCAGGTGCCCCCTGGCGCCCATGACCTACAAGGCGCCTTCGACCTGCCACTTCTTG
 AAGGAGAAGGCGCGCTGGAGGCTGATCTACTCAAGAAAGCGCCAGGATCCGACTGTGGGTGTACACACCCAGGGCTACTTCCC
 CCACTGGCAGAACTACACCCCGCGCGGAGCGCTACCCCTGACCTTCGGCTGGTGTCTCAAGCTGGTGGCCCGTGGACCCCGCCAGGAGG
 TGGAGAAGGCCAACGAGGCGGAGAACCTCCCTGTGCAACCCATCTGCAGCAGGGCATGGAGACGAGGAGCGGAGGTGCTGATGTGG
 AAGTTCGACTCCCGCCTGGCCTTCCGCCACATCGCCGCGGAGCTGCACCCCGAGTCTACAAGGACTGCTAA

도면98a

47. 2003 CON 01 AE nef.psp
MGGKWSKSSIVGWPQVRERIKQTEPPATEGVGAVSQDLDKHGAVTSSNMNADCVWLRAQEEEEVGFPVPQVPLRPMTYKGAFLDLSFFLKEK
GGLDGLIYSKKRQELLDLMVYNTQGFFPDWQNYTPGPGIRYPLCFGWCFKLVPDPREVEEDNKGNNCLLHPMSQHGTEDEEREVLWVKED
SALARKHIARELHPEYYKDCŞ

도면98b

2003 CON 01 AE nef.OPT
 ATGGGGCAAGTGGTCCAAAGTCTCCATCGTGGGTGGCCCAAGGTGCGGAGCGCATCAAGCAGACCCCGCCGCCACCGAGGGCGGTGGG
 CCGCGTGTCCAGGACCTGGACAGCACGGCCCGTGAACCTCTCCAAACATGAACAACGCCACTGCGTGTGGTGTGGCGCCCGCAGGAGGAGG
 AGGAGTGGGCTTCCCGTGGCCCGCCAGGTGCCCTGGCCCATGACCTACAAAGGGCGCTTCGACCTGTCTTCTTCCTGAAGGAGAAG
 GCGGCTGGAGGGCTGATCTACTCCAAGAGCGCCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTTCTTCCCGACTGGCA
 GAACTACACCCCGGCCCGGCATCCGCTACCCCTGTGCTTCGGTGGTCTTCAAGCTGGTGGCCCGCTGGACCCCGCGAGGTGGAGGAGG
 ACAACAGGGCGAGAACAACTGCCTGTGACCCCATGTCCAGCAGGATCGAGGACGAGGACCGCGAGGTGCTGATGTGGAGTTCCGAC
 TCCGCCCTGGCCCGCAAGCACATCGCCCGCGAGCTGCACCCCGAGTACTACAGGACTGCTAA

도면99a

48. 2003_CON 03 AE nef pep
MGGKWSKSSIVGMPQVRERIRAPAPARGVPSODLDKYGAVTSSNTAANNADCAWLEAQKEEVGFVRPOVPLRPMWTKGAFDLSHFL
KEKGGDLGLIYSKKRQEIILDLWVIHTQGIFFPDWQNYTPGPGIRFPLTFGWCYKLVDPDEVEEATEGENNSLIHPICQHGMDDEKEVIMW
KFDSRLATHRRARELHPEFYKDC\$

도면100a

49. 2003 CON 04 CFx nef .pep
MGKWSKSIVGWPATRRMRQRGPAQAEFAAGVCAVSQDLDKHGAIITSSNTAATNPDKAWLEAQEEEEVGFPVRQVPIRPMTFKAALD
LSHFLKEKGGLDGLIYSKKRQEIILDLWVYNTQGYFPDWQNTPGGERPFLCFGWCFKLVPVDFEQVEEATEGENNCLLHPISQHGMEDEER
EVLKWKFD SRLAYKHIARELHPEFYKDC\$

도면100b

2003 CON 04 CFX nef. OPT
 ATGGGCGGCAAGTGGTCCCAAGTCCCTCCATCGTGGGCTGGCCCGCATCCGCGAGCGCATGCGCCAGCGGCCCCGCCAGGCGGAGGCCGC
 CGCCGCCGCGGTGGGCGCCGTGTCCAGGACCTGGACACAGCACGGCGGCATCAGCTCTCCACACCGCGGCCACCAACCCCGACACAGGCCT
 GGCTGGAGGCCCCAGGAGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCCAGGTGCCCTGGCGCCCATGACCTTCAAGGCGCCCTGGAC
 CTGTCCCACTTCTGAAGGAGAAGGCGGCTGGACGGCTGATCTACTCCAGAAGCGCCAGGATCTGGACCTGTGGGTGTACACAC
 CCAGGCTACTTCCCGACTGGCAGAACTACACCCCGGCCCGCGAGGCTTCCCGCTGTGCTTCGCTTCAAGCTGGTGCCCG
 TGGACCCCGAGGAGTGGAGGAGGCCACCGAGGGGAGAACACTGCTGTGTCACCCCATCTCCAGCACGGCATGGAGGACGAGGACCG
 GAGGTGCTGAAGTGAAGTTCGACTCCCGCTGGCTACAAGCACATCGCCGAGCTGCACCCCGAGTTCTACAGGACTGCTAA

도면101a

50. 2003 CON 06 CFx nef . pep
MGKWSKSSIVGWPOVRERMRNPTEGAAGVGAVSODLDKHGAI TSSNTATTNNAACAWLEAOTEDVGFVVRPQVPLRPMTYKGAFDLSFF
LKEKGLDGLIYSKKRQEILDLDLWVYHTQGEFFPDWQNYTPGFGIRYPLITFGWCYKLVDPKVEEDTKGENNCLLHPMCQHGVEDEEREVLN
WKFDSSLARRHIAREMHPEFYKDC\$

도면101b

2003 CON 06 CFx nef. OPT
 ATGGCGGCAAGTGGTCCCAAGTCTCCATCGTGGGCTGGCCCCAGGTGGCGGATGCGCAACCCGCCACCGAGGGCGCCGCGAGGG
 CGTGGGGCGCGGTGTCCCAAGGACCTGGACAGCGGGGGCCATACCTCCTCCACACCGCCACCAACGCGCCTGGCGCTGGCTGGAGG
 CCCAGACCCGAGGACGAGGTGGGCTTCCCCGTGCGCCCCAGGTGCCCTGCGCCCCATGACCTACAAGGGGCGCTTCGACCTGTCTCTTC
 CTGAAGGAGAAGGGCGGCTGGACGGCTGATCTCAAGAAGCGCGGAGATCCTGGACCTGTGGGTGTACCAACCCAGGGCTCTT
 CCCCAGTGGCAGAACTACACCCCGGGCCCCGGCTACCCCTGACCTTCGGCTGGTGTACAGCTGGTGCCTGGACCCCCAAGG
 AGGTGGAGGAGGACCAAGGGCGAGACAACCTGCTGTGTCACCCCTATGTGCCAGCAGGCTGGAGACGAGGCGCGAGGTGCTGATG
 TGGAAATTGCACTCCTCCCTGGCGCGGCCACATCGCCCGGAGATGCACCCCGAGTCTTACAAGGACTGCTAA

도면102a

51. 2003 CON 08 BC nef pep
MGKWSKSSIVGWPAIRERIRRTTEPAADGVGAVSRDLEKKGATSSNTADTNADCAWLETQEEEEVGFPVRPQVPLRPMTFKGALDLSFFLK
EKGLEGLIYSKKQEIILDIWVYHTQGYFPDWHNITPGPGVRFPLTFGWCFKLYPVDPREVEEANEEDNCLLHPVCQHGMEDEHREVLKWK
FDSQLAHRHARELHPEFYKDC\$

도면102b

2003 CON 08 BC nef .OPT

ATGGCGGGAAGTGGTCCAGTCCCTCCATGCTGGGCTGGGCGGCCATCCGCGAGGGCATCCGCGCACCGAGCCGCGCGCGCGCGGTGGG
CGCCGTGTCCCGGACCTGGAGAGCACGGGCGCCATCACCCTCCCAACACCGCGGACCAAGCCGACTGCGCTGGCTGGAGACCCAGG
AGGAGGAGGAGTGGGCTTCCCGGTGCGCCCGCCAGGTGGCCCTGGGCGCCATGACCTTCAAGGGCGCCCTGGACCTGTCTTCTTCTGAA
GAGAGGGCGGCTGGAGGCGCTGATCTACTCCAGAAGCGCCAGGAGATCCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCGA
CTGGCACAACTACACCCCGCGCGCGGTCCGCTTCCCGCTGACCTTCGGCTGGTGTCTCAAGTGGTSCCGTGGACCCCGCGGAGTGG
AGGAGGCCAACGAGGCGAGGACAACTGCTGCTGACCCGCTGTGCCAGCACGGCATGAGGACGAGCACCGGCGAGGTGCTGAAAGTGGAG
TTCCGACTCCAGCTGGGCCACCGCCACCGGCGCCCGAGTGCACCCCGAGTTCTACAGGACTGCTAA

도면103a

52. 2003 CON 10 CD nef pep
MGKWSKSEIVGWPAVBRIRRTDPAAGVGAASRDLEKYGAISSNTAQTNPDCAWLEAQEEEEVGFPVRPQVPLREMTYKGAFDLSFFL
KEKGLEGLIYSKRQDILDWVYNTQGFPPDWONYPGPGIRYPLTFGWCYKLVVPDPREVEEANEGENNSLIHPMSLHGMEDPHGEYIMW
KFDNIAHKHWARELHPEYKDC\$

도면103b

2003 CON 10 CD nef.OPT
 ATGGGCGGCAAGTGGTCCCAAGTCCTCCATCGTGGGCTGGCCCGCCGTGCGGAGCGCATCCGCCCGACCCGACCCCGCCGAGGGCGTGGG
 CGCCGCTCCCGCGACCTGGAGAGTACGGGCGCATACCTCTCTCCACACCCGCCAGACCAACCCCGACTGCGCTGGCTGGAGGCCGAGG
 AGGAGGAGGAGGAGGTGGCTTCCCGTGGCCCGCAGGTGCCCCGATGACCTACAGGGCGCTTGACCTGTCTTCTTCTG
 AAGGAGAAGGGCGGCTGGAGGGCTGATCTACTCCAGCGCCGCGAGACATCCTGGACTGTGGGTGTACACACCCAGGGCTTCTTCCC
 CGACTGGCAGAACTACACCCCGGCTCCGATCCGTAACCCCTGACCTTCGGGTGGTGTACAGCTGGTGGCCCGTGGACCCCGGAGG
 TGGAGGAGGCCAACGAGGGCGAGACACTCCCTGTGACCCCTATGTCACGGATGGAGACCCCCACGGCGAGGTGGTGTATGTGG
 AAGTTCGACTCCAACTGGCCACACAGACATGGCCCGCGAGCTGACCCCGAGTACTACTACAGGACTGCTAA

도면104a

53. 2003 CON 11 CFX nef .pep
MGKWSKSEIVGWPETIRELRRTPTAAEGVGAVSKDLEKHGAVTSSNTAQTNACANLEAQDEEEVGFPVREPQVPLHPMTYKGAFDLGEF
LKEKGLDGLLYSKRQELLDLWVYHTQGYFPDWQNYTPGPGIRYPLCFGWCFLVPVEPREVEANEGENNCILHPSQHGMDEREVLN
WKFDSSLARRHIARELHPDFYKDC\$

도면104b

2003 CON 11 CFX nef.OPT
 ATGGGGCAGTGGTCCAGTCCATCGTGGGTGGCCCGAGATCCGCGAGGCGCTGCGCGCACCCCCCCCCACCGCCGCGCCGAGGG
 CGTGGGCGCCGTGTCCAAGGACCTGGAGAGCAAGGGGCGGTGACCTCCTCCAAACCGCCCGACCCAGCCGCTGCGCTGGCTGGAGG
 CCCAGGAGGAGGAGGTGGGCTTCCCGGTGGGCCCCCAGGTGGCCCTGCGCCCAAGCTACAGGGCGCTTCGACCTGGGCTTCTTC
 CTCAGGAGAGGGCGGCTGGAAGGCTGATCTACTCCAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCCACCCAGGGCTACTT
 CCCCAGTGGCAGAACTACCCCGGCGCGGATCCGCTACCCCTGCTTGGGTGGGTCAAGCTGGTGGCTGGAGCCCGCG
 AGGTGGAGGAGGCCCAAGGAGGCGGAGAACAACTGCTGTGCAACCCATGTCCAGACAGGCACTGACGAGGAGCGGAGGTGCTGATG
 TGGAASTTCGACTCCTCCCTGGCCCGCGCCACATCGCCCGGAGCTGCACCCCGACTTCTACAGGACTGCTAA

도면105a

54. 2003 CON 12 BF nef . pep
MGKWSKSSIVGWPDIIRERMRA2PAAEGVGAVSODLENRGATSSNFRANNPDIANLEAQEEVGFVVRPOVPLRPNTYKGAIDLSHFLK
EKGLEGLIYSKKRQEIILDIWVYHTQGYFDWQNYTPGPGIRYPLITIGWCFKLVVDPEEVEKANEENCLLHPMSQHGMEDDEDREVLWVK
FDSRLALRHIAREKHFYQDC\$.

도면105b

2003_CON_12_BF_ref.OPT

ATGGGGGCAAGTGGTCCAAATCCTCCATCGTGGGCTGGCCGAGATCCGCGAGGCAATGCGCGCGCCGCCGCCCGCGCGCGCGGTGGG
CGCCGTGTCCAGGACCTGGAGAACCGCGGGCCATCACCTCTCCAACACCCGCGCCACACCCCGACCTGGCTGGAGGCCCCAGG
AGGAGGAGGAGGTGGGTTCGCCGTGGCCGCCAGGTGCCCTGGCGCCATGACCTACAGGGCGCCCTGACCTGTCCCTGACCTTCCTGAAG
GAGAGGCGGCTGGAGGCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGGGTGTACCAACCCAGGCTACTTCCCCGA
CTGGCAGAACTACACCCCGCGCATCCGCTACCCCTGACCTTCGCTGTGCTTCAAGCTGGTCCCGTGACCCCGAGGAGGTGG
AGAGGCCAACGAGGGGAGAACACTGCTGTGACCCCATGTCCAGACCGCATGGAGGACCGGAGGTGCTGTGTGGGAG
TTCGACTCCCGCTGGCCCTGGCGCCACATCGCCCGGAGAGCACCCCGAGTTCTACCGGACTGCTAA

도면106a

55. 2003_CON_14_BG_ref.PEP
MGGKWSKCSIVGWPETREIRRTPEAAVGVGAVSQDLAKHGAIITSSNTAANNPDCAWLEAQEEDSEVGTFVRPOVPLRPMTYKGAFDLSFFL
KEKGLDGLLYSKQODILLDWVINTQGFPPDWQNTFGPGTRYPLTFGWCEKLEPVDPAEVELEATKGENNSSLHPICOHGMEADADNEVLW
RFDSSIARRHIARELHPDFYKDC\$

도면106b

2003_CON_14_BG_nef.Opt
 ATGGGCGGCAAGTGGTCCAAAGTGCCTCCATCGTGGGCTGGCCCGAGGTGCGCGAGCGCATCCGCCGCAACGCCCGCGCGCGCGGTGGG
 CGCCGTGTCCAGGACCTGGCCAGACGGCGCCATCACCTCTCCAAACACCGCGCGCCACACGCCGACTGCGCTGGCTGGAGGCCGAGG
 AGGAGACTCCGAGGTGGCTTCGCCGTGGCGCCCGAGGTGCCCTGCGCCCATGACCTACAGGGCGCTTCGACCTGTCCCTTCTCCTG
 AAGAGAAAGGCGGCTGGACGGCTGATCTACTCCAGCAGCGCCAGGACATCCTGGACCTGTGGGTGTACAAACCCAGGGCTTCTTCCC
 CGACTGGCAGAACTACACCCCGGCCCGGCTACCCCTGACCTTCGGTGGTGTCTCAAGCTGGAGCCCGTGGACCCCGCCGAGG
 TGGAGGAGGCCACCAAGGGCGAGAACACTCCCTGCTGACCCCATCTGCCCAGCACGGATGGAGGACGCCCAACAGAGGTGCTGATCTGG
 CGCTTCGACTCTCCCTGGCCCGCGGCCACATCGCCCGGAGCTGACCCCGACTTCTTACAAGGACTGCTAA

도면107a

61. 2003 2003 2003 con s pol pep

PERENLAFQGEAREFSSSEQTRANSPTSRELRVRGDNPLSEAGAEQGTIVSLSPFQITLWQRPVTVTKIGGQKKEALLDTGADDTVLEIN
 LPKWKPKWIGGIGGFIKVRQYDQILIEICGKKALCTVLVGPTPVNIIGRNMLTQIGCTINFPISPIETVPVKLKEGMDPKVKQWPLTEEK
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIAKKDSTKWKYVDFRELNKRTQDFWEVQLGIFHEPAGLKKKSVTVLVDGDYFVSVELDE
 DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFTQNPFIYIYQIMDDLYVGSDEIGQHRTKIEELREHLLRWGF
 TTPDKKHQKEPFFLWNGYELHPDKWTVQPIQLPEKDSWTYNDIQKLVGKLWASOILYPGIKVKOLCKLLRGAKALTDIVPLTEAELELAEN
 REILKEPVHGVYDPSKDLIAEIOKQGDQWTYOIYQEPFNKLTGKYAKMRSANTNDVKQLTAEVQKIATESIIVGKTPFRFLPIQKETW
 ETWWTYMOATWIPMEFEVNTPPLVKLWYLEKEPIVGAETFYVDGAANRETKLKGAGYVTDGRQKVVSLSLTETTNQKTELQATHLALQDSG
 SEVNLVTDQYALGLIOAPDKSESELVNOIEQLIKKEKVLVSWVPAAHKGIGGNEQVDKLVSTGIRKVLFLDGLDKAQEEHEKHHSWRAM
 ASDFNLPPIVAKELIVASCDCQKGEAMHQVDCSPGIWQLDCTHLEGKILLVAVHVASGYIEAEVIPAETGOETAYFILKLAGRWPVKVIH
 TDNGSNFSAAVKACWAGIQQEFGIPYNPQSGVVESENKELKKIIGQVRDQAEHLKTAVOQMAVFIHNFRRKGGIGGYSAGERIIDIAT
 DIQTKELQKQITKIQNTFRVYIRDSDRDPITWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKITROYGKQWAGDDDCVAGRQDEDS

2003 CON S pol.OPT

TTCTTCCGCGAGAACCTTGGCTTTCGAGCAGGGGACCGCGGACATCTCCCACTCCCGCGAGCTGGCGTGGC
CGCGGCGCAACCCCGCTTCGAGCGCGCGCGAGCGCGGACCTGTCTCTCTTCCCCAGATCAACCTTGGTGTGACGG
TGAAGTTCGCGGGCGCGACTTCGAGAGAGCCCTGCTGACACCGGCGCGACACACCTGCTGGAGGATCAACCTGCGCGAGATCGAAGACCCCAAGT
ATCGGGGCGCATCGCGCGCTTCATCAAGTGGCGCACTGACACAGATCTCGGACAGAGATCTCGGACAGAGCCATCGCACCTGCTGCTGGCGCCACG
CCCGTGTAAATATATCGCGCGCCAACTGCTGACCAAGTCGGCTGCACCTTGAATTCCTCCATCTCCCACTCGAGACCGTGTCCCTGTAAGCTGAAGCCCG
AGATGAGCGCGCCAAAGCTGAAGACATGGCCCTGACCGAGAGAGAGTAAGGCGCTTGAACGAGATCTGCACAGATGAGAGAGGAGCGACAGATCTCC
AAGATCGCCCGCGAGAACCCCTACAACACCCCGATCTTGGCAATCAGAGAGAGATCTCAACAAGTGGCGCGAAGTGTGTGACTTCGCGAGCTGAACA
GGGACACACAGGACTCTTGGAGGTGACATCGCCATGGGATCTCCACACCGCGCGCTTGAAGAGAGAGATCTGCTACCGCTGAGCTGGGCGAGCGCTTACT
TCTTCGTGCCCTTGGACAGAGATCTCCGAGATCACACCGCTTCACCATCCCTCTCATCAACAGAGAGACCCCGGATCTCGGTACCAGTACAAGAGTGTGT
CCCGAGGCTTGAAGAGGCTCCCGGCCATCTTTCAGTCTCCATGACCAAGATCTGAGCGCTTTCGCAACCGAGAACCCCGAGATCGTGATCTATCCAGTA
CATGAGAGACTGTACAGTGGGCTCGAGCTTGAGATCTGGCACGACGACCAACATCGAGAGTGGCGGACGACTGTGCTGCTGGGCTTCACCAACC
CTCTGGACCGTGAGACATATCAGAGAGTGTGTGGCAGATGAATCTGGGCTCCAGATCTACCCCGGACTGAGGAGTGGAGAGTGTGCAAGTGTGTGG
CGGCGCACAGGCGCTGACCGACATCTGTGCCCTGACCGAGGAGCGGAGCTGGAAGTGGGAGCTCGCGGAGAACCCGAGATCTCTGAGGAGCGCTGTGCAAGCGGCTGT
ACTACAGACCTCTCCAGAGACCTGTATGCCCGAGATCCAGAGAGGCGGACGACAGTGCAGCTGAGGAGTCAACAGATCTGCCACGAGTCCATGTATCTGGGCGCA
GGCAATATCCCTCGAGAGCTCGCGCACACACAGACGACGTGAGAGCTGACCGAGCGCTGCAGAGATCGCCACGAGCGCGCCCAACCGCGAGACCAAGCTG
GCAACCCCAAGTCTCGCTGCCATCTCAGAGAGAGCTGGGAGCTGTGTGAGCGAGTACTGTGGAGGCGACCTGAGTCCCGAGTGGGAGTCTGCTGAGCA
CCCCCCCCCTGTGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCTGTGGGCGCGAGACCTTCTACTGGAGCGCGCGCCCAACCGCGAGACCAAGCTG
GGCAGCGCGGCTACTGTCAGCAGCGCGGCGCAGAGGTGTGTCCTTACGAGACCAACACAGAGACCCAGTGCAGGCGCATCTACCTTGGCGCTT
GAGATATCCGAGTCCGGCTCGAGGTGAACATGTGACGACTCCCACTGCTGGGCTCAATCAGGCGCCGACCGGACAGTCCGAGTCCGAGTGGTGAACC
ACCGGATCOGCAAGTGTCTTCTTGACGCGCATCGCAAGGCCAGAGAGAGACAGAGATACCACTCCACTGGCGCGCATGGCCCTCCGACTTCAA
CCGTGCCCATCGTGGCGAGAGAGATGTGGCTCTCTGGAGATGTCGAGCTGAGAGGCGAGCCATCGAGGCGAGGAGTGGTGGTGGTCTCCCGCGCATCT
GGGACGCTGCACTTCTTCTCGAGGCGCAAGTATCTCTGGAGGTGGCGCTGGCCGTGAGTGTGAGAGTGTATCCACCGAGAGCTTCACTCCCGCGCGAGCGGC
CCGCTGTGTGGCGCGAGATCTCAGCAGAGTTCGAGATCCGCTTACACCCCGATCCAGGGCTGTGTGAGTCCATGAACAGGAGCTGAAGAGATCA
TCTGGCCAGGTGGCGACAGCGCCAGAGACCTGAGAGACCGCGTGCAGATGGCTGTTCTTCAACATTCAGGCGCAGGCGGCGCATCGCGGCTACTCC
GCGCGGAGCGCATCTATCGACATCATCGCACAGATCGCACCGGCGCGCGCGAGAGTGTGTGATCCAGGACAACTCCGAGTCAAGTGTGTCCCGCG
CTTCGCGAGCCAAAGTATATCCGCGACATTCAGGCGAGAGATGCGCGCGAGACTGTGGCGCGCGCGGAGAGTGTGTGATCCGAGTCAAGTGTGTCCCGCGC

도면108a

62 2003_M GROUP anc pol.PEP
 FFRENLAFOQGEAREFSSEQTRANSPTSRELRVRGGDNPLSEAGAEQCTVSSESFPQITLWQRPVLTIKIGGQREALLDTGADDTVLEEIN
 LPGKWKPMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPIETVPVKLPKGMGPKVKQWPLTEEK
 IKALTEICTEMEKDGKISKIGPENPINTPVFAIKKKDSTKWKLVDFRELNKRTQDEWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVELDE
 DFRKYTAFTIPSINNETPGIRYOYNVLPOGWNKSPAI FQSSWTKILLEPFRKNPEIYIYQMDLTVGSQDLEIGQHRAKIEELREHLLRWGF
 TTPDKKHQKEPPFLWMGYELHPDKWTVQFIQLPEKDSWTVNDIQKLVGKLNWASQIYPGKVKQCLCKLLRGAKALTDIVPLTEAELELAEN
 REILKEPVHGVIYDFPSKDLIAETIQKQGDQWTYQIYOEPFNKLTGKYAKMRSHTNDVKQLEAVQKIATESIVIMNGKTPKFRPLIQKETW
 ETWWTYWOATWIPWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLKAGYVIDEGRQKVVSLETTNOKTELOAIHIALQDSG
 SEVNIIVTDSOYALGIIQAOPDKSESELVNOIIEQLIKKEKVIYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHNSWRAM
 ASDENLPPVVAKEIVASCDKQCKGEAMHGQVDCSPGIWLOCTHLEKGVILVAVHVASGYLEAEVIPAETGOETAYFILKLGRWPVKVIH
 TDGNSFTSAAVRAACWAGIOOEFGI PYNPQSGVVESMKNELKKIILGOVRQAEHLKTAVMQAVFIHNFRRKGGIGGYSAGERIIDIILAT
 DIQTKELQKQITKIQNFRVYYRDSRDLWKGPKALLWKGEAVVIQDNSEIKVVPRRKAKIIRDYGKQMGAGDDCVAGRQDEd\$

2003_M.GROUP and pol.OPT

TTCTTCCGCGAAGACCTGGCCCTTCGACGAGGCGAGCCCGCGAGTTCTCCTCGAGCAGACCCCGGCGCAATCTCCCGCAGACTGCGCGCTGGCCGCGCGCAGAACCCCTTCTCGAGGCGCGCGCGAGCGCCGAGGCGACCGCTGTCTCTCTCCCGCAGATCAACCTGTGGCAGCGCCCGCTGTGTGACCA
TACAGGTCGCGCGCAGCTGTCGCGAGCCGCTGCTGGACACCGCGCAGACACCTGCTGGAGAGATCAACTTCCCGCGCAAGTGAGACCCCAAGATG
ATCGCGGCGATTCGCGCGCTTCATCAAGGTTCGCGAGTACGACAGATCTCATCGAGATCTTCGCGAAGAGGCCATCGCCACCGTCTGCTGTGGCGCCCGC
CCCGTGAACTCATTCGCGCGCAATGCTGACCCAGATCGGCTGCACCTTGAACTCCCGCATCTCCCGCATTCGACGCTGCCCTTGAAGCTTGAAGCTCG
AGATGAGCGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGGATCAAGGCCCTGACCGAGATCTGCACCGAGATCTGCACCGAGATGSGAAGAGGSGCAAGATCTCC
AGATCGGCCCGCGAATCCCTTACACACCCCGCTGTGCBCTTCAAGAGAGGAGACTCACCAAGTGGCGAGCTGTGTGTGATCTTCGGCAGCTGACACAA
GGCGACCCAGAGACTTCTGGAGTTCAGCTGGGCATCCCGCACCCCGCGGCTTACCATCCCTCATCAAGACGAGACCCCGGCTACCGTGTGACGATACACGCTGCT
TCTGCTGCGCCCTCGACGAGGACTTTCGCGAAGTACACCGCTTACCGCTTCCATGACCAAGATCTTGAAGCCCTTCCGCAACCAAGACCCCGAGATCTGATCTACAGTA
CCCCAGGCTTGAAGGGTCCCGCGCATTTCCAGTCTTCATGACCAAGATCTTGAAGCCCTTCCGCAACCAAGACCCCGAGATCTGATCTACAGTA
CATGACGAGCTGTACAGGAGTCCCGATCGAGATTCGGCCAGCAGCGCAGATCGAGAGGCTGCGCAGCTGCTGTGCGCTGGGGCTTCAACCAACC
CGGACAGAGACACCAAGAGGCGCCCTTCTCTGTGTGAAGTGGCTGCGACCCGACAGTGCAGCTGACGCCATTCGAGCTGCCGAGAGGAC
TCTTGACCGTGAACATCAAGAAAGTGTGGCGAGCTGAATGGGCTCCAGATCTACCGGCTCATCGGCGATCAAGGTGAAGAGCTGTGCAAGTGTGTGG
CGGCGCCACGSCCTGTACCGACATGTCGCCCTGACCGAGGAGCGAGCTGGAGCTGGCGAGAACCGCAGATCTTGAAGGACCCGCTGCACGGCGTGT
ACTAGACGCCCTCAAGAGCTGTATCGCGGAGTCCAGAGACAGCCAGCTGAAGAGCTGCGCAGAGATCGCCACCGAGTCCATGTGATCTTGGGCGCA
GACCCCGCAAGTTCGGCTGCCATTCGAGAGAGACTTGGGAGACTGTGTGACCGAGTATGTGGCAGGCTACCTGTGATCCCGGAGGAGCTTCAAGAACCTTCAAGAC
GCCCCCGCTGTGAAGCTGTGGTACCTGCGGAGACACCAAGCACTGAAGAGCTGACGAGGCTGCGCAGAGATCGCCACCGAGTCCATGTGATCTTGGGCGCA
GGCAGCGCGGCTACTGTACCGACCGCGCGCGCAGAGTGTGCTCTTCGCGCAGACCAACAGAGAGAGCTGAGCGGCTACCGCTTGGCGCT
CGAGACTCCGGCTCGAGGTGAATCGTGACCGACTGGAAGAGNCCCATGCTGGGCGCGAGACCTTCTAGTGAGCGCGCGCGCAACCGCGAGACCAAGCTG
AGATCATCGAGAGCTCATAGAGAGAGAGTGTACTCTGCTGAGTGGCGCATCAAGGGCATCAAGGCCATCAAGGCGCAGAGCTCGAGTCCGAGTGGTGAAC
TCCGGCATCCGCAAGTGTCTTCTTGACGGCATCGACNAGGCCAGAGAGACACAGAAATACACTCCAACTGSGCGCGCATGGCTTCCGACTTCAA
CTTGGCCCCGTGTGGCGCAGGAGATCTGGCTCTCTGCAAGTGTGGCAGCTGAAGCGCGAGCGCATGACGGCGCAGCTGAGCTGCTCCCGCGGATCT
GGCAGTGTGACTTCACTTCACTCGAGGCAAGGTGATCTGTGTGGCGCTGTCAGCTTGGCTCCCGCTACATCGAGCGCGAGTGTGATCCCGCGCGAGACGGC
CAGGAGCGCGCTACTTCACTTCTGAGTCTGGCGGCTGTGGCGCTGAGTGTGTCACACAGCTGAAGGCGAGCGCATGACGGCGCAGCTGAGCTGCTCCCGCGGATCT
CGCTGCTGTGGCGCGCATTCAGCAGAGATTCCGATTCCTCATCAACCCCGCTGCCAGTCCCGAGGCGTGTGTGATCCATGAACAGAGAGCTGAAGAGATCA
TCGGCCAGAGTGGCAGACCGCGCGCATCGAAGCGCGCTGACATGSCCTGTTTCACCAACTTCAAGCGCAGAGGGCGGCTACGCGGCTACTCC
GCGCGAGCGCATTCAGCAATGATCGCCACCGATCATCAAGAGCTGCGAGAGCTGCGAGTGAATCAAGATCAAGAGATCTTCGCGTGTACTACCGCGA
CTCCGCGACCCCATCTGGAAGTCCCGCGCAAGCTGTGGAGAGGCGCGAGCGCGTGTGTGATTCAGGACAACCTCGAGATCAAGTGTGTGCCCGCG
CAAGGCGCAAGATCATCCGCGACTACGGCAACAGATGTCGCGCGAGCTGTGGCGCGCGCGAGGAGACTTAA

도면109a

63. 2003 CON A1 pol. pep

FFRENLAFOQGEARKTSSEOTGANSPTSRLMDGGRDLSFSEAGAEROGTGPTFSFPOITLWORPLVTVRIGGQIKKALLDTGADDTVLEDI
 NLPCKWKPMIGGIGGFIKVKQYDQILIEICGKKAIGTVIVGTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKGMDGPKVKQWPLTEE
 KIKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLDVDFRELNRKHTQDFEWVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD
 ESFRKYTAFTIPSTNNETEGIRYQYNVLFQGWKSPAIPOSSMTKILEPFRSKNPEIILYQYMDLVYSSDLGIGQHRTKIELRAHLLSWG
 FTTPDKKHQKEPPFLMWGYELHPDKWTQPIELPEKESWTNDIKLVCKLNWASQIYAGIKVKQCKLLRGAKALTIDIVTLTEAELELAE
 NRELLKDPVHGTVYDPSKDLIAELQKQGDQWYQIYQBPENLTKGKYARKRSAHTNDVKQAEVVQVVMESIVIWGKTPKFKLPIQKET
 WETWMDYQATWIPWEFEVNTPLVLMWYQLEKDPYVGAETFYVDGAANRETKLGAGYVTDGRQKVVSLETNNQKTELHAIHLALQDS
 GSEVNI VTDSQYALGIIQAQPDSESELVNIIEKLGKDKVLSWPAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAQEEHERYHSNWR
 MASDFNLPPIVAKREIVASCDKCOLKGEAMHGQVDCSPGIWQLDCTHLEGVILIVAVHVASGYIEAEVIPAETGQETAYFELLKLAGRWPKVV
 HTDNGSNFTSAAVKAAACWANIIOEEFGIPYNQSGVVESSMKNELKKTIGQVEQAEHLKTAVQMAVETIHNKRKGGIGGYSAGERIDIIA
 TDIQTKELQKQITKIQNFRVYRDSRDPITWKGPAKLLMKGEGAVVIQDNSDIKVVPRKAKIINDYKQMGAGDDCVAGRQDEDS

2003 CON A1 pol.OPT

[illegible]

도면109c

64. 2003 Al.anc pol.pap

FFRENIAEQGEARKEFSSEQTRANSPTSRELWDGGRDLSLSEAGHERQGTVPSPFPQITLWQRPVTVVKIGSQKKEALLDTGADDTVLEDI
 NLPCKWKPWIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTINFISPIETVPVKLKPGMDGPKVKQWPLTEE
 KIKALTEICTEMEKEGKISKIGPENPNYNTVEFAIKKDSIKWRKLVDFRELNKRKTQDFWEVQLGIPHAGLKKKSVTVLDVGDAYFSVPLD
 ESFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGSPIFQSSMTKILEPFRSKNPEIIVIQYMDLYVGSDEIGQHRAKIEELRAHLLSWG
 ETPDKKHQKEPFFLWNGYELHPDKWTVPKLPKEDSWTVNDIQLVGLKNWASQIYAGIKVKQLCKLIRGAKALTDIVLTTEEALELAE
 NREILKDPVHGYYDESKDIAVEIQOGQDQWTYQIYQEPFNLTGKYAKKRSANTNDVKQTEVQKVATESIIVWGKTPKFRPLQKET
 WETWWEYXQATWIPWEFVNTPPLVKLWYQLEKEPIAGAEFTYVDGAANRETKLGAGYVTDGRQKVYSLTETTNOKTELHAIHLALQDS
 GSEVNIVTDSQYALGIIQAPDRSESELVNOIIEKLEIEKVKVLSWVPAHKGIGNEQVDKLYSSGIRKVLFDGIDKRAQEEHEKYHSNWRA
 MASDFNLPPPIVAKETIVASCQKQKGEAMHGQVDCSPGIWQLDCTHLEGVILVAVHVASGYIEAEVIPAETQETAYELLKLAGRWPKVY
 HTDNGSNFTSAAVKAACWAWANIQQEFGIPYNPQSGVVESMNKELKTIQVREQAELHTAVQMAVFIHNEKKGIGGYSAGERIIDIIA
 TDIQTKEHQIITKIQNFRVYIRDSRDPWKGPAKLLWRGEGAVVIQDNSDIKVVPRKAKIIRDYGMAGDDCVAGRQDEUS

2003_A1.anc pol.OPT

TTCCTTCGCGAGAACCTCGGCGTTCAGAGAGGGGAGCGCCGCGAAGTTCCTCTCGAGCAGACCCGCGCCAAATCCCCCACTCCGCGAGCTGTGGGAGG
CGGCGCGAGTCCCTGTCTCCGAGCGCGCGCGAGCGCAGCGCACCGTCCCTCTCTCTTCCCGAGATCACCTGTGGCAGCGCCCTCGTGTGTGA
CCTGTGAAGATCGGCGCGGCGAGTGAAGAGGAGCCCTGCTGCGAGCAGCGCGACGACACCTGTCTGGAGAGACATCAACCTCGCGAGTGGAGAGCCCAAG
ATGATCGGCGGATCGGCGGCTTATCAAGTGTCCGAGTACGACAGATCTGTATCGAGTATCGGCGACGAAGGCGCATCGGCACCTGTGCTGTGGCGCC
CACCCCTGTGAACATATPCCGCGCGCAATGCTAGCCACAGATCGGTGTGCACCTGTAACTCCCATCTCCCATCGAGACCGTGTGAGTGTGAGTGAAGC
CGCGATGAGCGGCGCCCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAAATCAAGGCCCTGACCGAGATCTGACCGGATCGGAGATGGAGAGGAGGCGCAAGATC
TCCAAAGTCGGCCCGAGAACCCCTTACGACACCCCGCGTGTTCGCCATPACAGAAAGAGCATCAACCAAGTGGCGAGTCTGTGAGATCTCGCGAGTGA
CAAGCGCACCCAGAGTATCTGGAGGATCGAGCTGGCATCCGCCACCCCGCGGCTGAAGAGAGAGATCCGTGTGACCTGTCTGGAGCTGGCGGACGCT
ACTTCTTCGTCGCCCTGACAGATCTCTCCGCAAGTACAGCGCTTCACCATCCCTCCATPACAGAGAGACCCCGCATCCGTACCAAGTACAGGCT
GTATCATGGACCTGTACAGGCTCCCGCGCATCTCCAGTCTCTCATGACCAAGATCTGTGAGCCCTTCCGCTCAGAGACCCCGAGATCGTGTATCTACCA
CCCCCGAGAGATGTACAGTGGCTCCGAGTCCGAGATCGGCGACCGCGCCAGATCAGAGAGTGGCGCCACCTGTGTCTGTGGGCTTACCA
GACTCTGTGACCTGTGAACAGCATCAGAAAGCTGTGGGAGATGTAACTGAGCTCGAGCTCGAGCTCCCGACAGTGGCGACATCAAGTGTGAAGCACTGTGCGAGAG
ATCCGCGGACCTCCAGAGACCTGTGGCGAGCTGTGGCGAGATCGTGGACACCAACCAAGATCTGAGCACTCACTGAGCACTGACCTGGAGTCTGGTGA
CAGAGACCCCAAGTTCGCGCTCCCATCAGAGAGACCTGGAGACCTGTTGATGAGTACTTGGAGGCCACCTGTGATCCCGAGTGGGAGTCTGTGA
ACAGACCCCGCCCTGTTGAACTGTGGTACACTGGAGAGACCTGGCGCGCGAGACCTTACGTGTGAGCGGCGCGCCGCCAACCGCGAGACCAAG
TGTGGCAGCGGCTAGTGCAGCGCGCAGAGTGTGTCCCTGACGAGACACCAACAGACAGAGCTGCAGCCATCCACTGCC
CCCTGGAGGACTCGGCTCGGAGTGAACATCGTACCGACTCCGATCGCCTGGGATCATCAGCGCCAGCGCGACCGTCTCGAGTGGTGA
TACCAAGATATCAGAGATGTACAGAGAGACAGTGTACTCTGTGGTCCCGCCACAGGCGATCGCGCGCAGAGCAGAGCTGGACAGCTGGT
TCTCTCGGCACTCGCAGGTGCTTCTCTGGACGGATCGACAGGCCAGAGGACAGAGAAGTACTCATCAACTGTGGCGGCCAATGGCCCTTCGAGT
CAACTCTGCCCACTGTGGCCAGAGACATCTGGTGTGCTGTGCAAGTGTGCTGACGTTGGCTTACATCGAGGCGGAGTGTGATCCCGCGCGAGC
TCTGTGGAGTGGACTGCACCCACAGAGGCAAGTGTCTGTGGTGGCGCGCGCTGGCGTGAAGTGTGTACACCGACCAAGCGGCTCAACTTCACTCCCGCGCGAGCC
GGCGCGGCTGCTGTGGGCGACATCCAGAGAGTTCGCGATCCCTTACACCCCACTCCAGGGCGTGTGGAGTCCATGAAACAGAGCTTGAAGAGA
TATCATGGCAGTGTCCGAGAGCGGACACTGAGACCGCGCTGCATGCGCTGTTCATCCACACTTACAGCGCACAGGGCGCATCGCGGCTAC
TCCGCGCGAGGCGCATATCGACATPATGGCCACCGCATATCGACATCCAGAGACTGCAGAGAGAGATTCACAGATCAGAGTTCGCGTGTACTACCG
CGACTCCGCGCGACCTCATGTGTGGAGAGGCGCGCGCAAGCTGTCTGGAGAGGCGCGGCTGTGTATCGAGCAACTCCGACATCAAGTGTGGTCCCG
CGCGAGGCGCAAGATCATTCGCGACTACGGGAGAGATGGCGCGGCGAGCTGTGGTGGCGCGCGCGAGGAGGAGTAT

도면110a

65. 2003 CON A2 pol.pap

FERENLAFQREARKFSSEQNRANSTSRRLNRNGGRDNLLSEAGAEEOQTVHSCNFPQITLWQRLVTVVKIEGQIREALLDTGADDTVLEDI
 NLPGRWKPRMIGGIGGFIVRQYDQIAIEICGKRAIGTVLGGPTPVNIIIGRNMVLGCTLNFPIETVPVVKLPKPGMDGPKVKQWPLTEE
 KIKALTEICKEMEKEGKISKIGPENFYNTVFIAIKKDSFKWRKLVDFRELNRKTPQDEFWEVQLGIPHPAGLKKKSVTVLDVGDAFYFVPLH
 EDPRKYTAFTIPSTNNETFGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFRSKNPEMVIYQYMDLYVGSDELIGOHRAKIEELRAHLLRWG
 FTTDPKKHQKEPPFLWNGYELHPDKWTVQPIKLPEKDSWTVNDIOKLVGKLNWASQIYAGIKVQOLCKLIRGTAKLTDTVTLKEAELELEE
 NREILLKNPVHGVYIDPFSKDLIAEIQKQGDQWTYQIYQEPFNLTGKYAKRKSTHNDVKQLTEAVQKIAESIVWGTTPKFRPIQKET
 WETWTEYWOATWIPWEFVNTPLVKLWYQLETEPTAGAEFTYVDGAANRETKLGAGVTDGRQKTIVSLTETTNOKTELHAIYLAQDS
 GLEVNIVTDSQYALGIIQAQPDSESELVNOIIEKILIEKERVYLSWPBAHKIGGNEQVDKLYSSGIRKXVLFLDGIDKAQEEHERRYHNSWRA
 MAHDENLEPIVAKETIVASCDCOLKGEAMHGQVDCSPGICWQLDCTHLEGKVIILVAHVHVASGYIEAEVIFAETGGQETAYFILKLGRWPVKVI
 HTDNGENFTSATVKAACWAGVQQEPGIPYNPQSGVVSMMNKELKIIIGQVRDQAEHLKTAVQMAVFLHNFRRKGGIGGYSAGERIIDIIA
 TDIQPRELQKQIIKIQNFRVYVYRDSRDPIMKGPALKLMKGEAVVIQDNSDIKVVPRRKAKIIIRDYKQMGAGDDCVAGRQDEDS

도면110b

2003_con_22 pol.OPT

TTCTTCGGGAGAACCTGGCTTCAGAGAGGGAGGCCCGCAGTTCTCTCCAGAGAGAACCGGCGCAATCTCCCACTCTCCCGGAGATGGCGCAAGGG
 CGGCGGAGCAACTGTGTCCGAGGCGGCGCCGAGAGCAGGGCACCTGTGCACTCTCTCAACTTCCCTCAGATCACCTGTGGCAGCGCCCTCTGTGTGA
 CGGTGAAGATCGAGGCTGCGGAGGCGCTGTGTGACACCGGGCCCGACACACCTGGTGTGAGGACATCAACTCTCCCGGCAAGTGGAGAGCCCAAG
 ATGATCGGCGCATCGGCGGCTTCTATCAAGTTCGCGCAGTACGACACGATCGCCATCGAGATCTGCGGCAAGCTGCGACCTGCTGTGTGGCGCC
 CACCCCGTGAATATCATCGGCGCAATGCTGTGTGACAGTGGGCTGCACCTGAACTTCCCACTCTCCCACTCTCCCACTCTGAGACCGTGTGAGCTGAAGC
 CGGCAITGAGCGCGCCAGGTGAAGGTGAGGCTGCGCTGACCGAGAGAGATCAGGCGCTGACCGAGATCTCAAGAGATGGAGAGGAGGGCAAGATC
 TCCAAGTCCGCGCCGAGAACCCCTATACACCCCGGTGTTCGCATCAAGAGAGATCCACAACTGGGCGAAGCTGTGTGACTTCCCGGAGCTGAA
 CAAAGCGCACAGGACTTCTGGGAGGTGAGCTGGGCTCCCGCCAGCGGCTTCACTATCCCTTCACTCAACAGAGACCCCGGATCCGCTACAGTACAGGTG
 ACTTCTCCCTGCGCTGCACGAGGACTTCCGCAAGTACACCGCTTCACTATCCCTTCACTCAACAGAGACCCCGGATCCGCTACAGTACAGGTG
 CTGCGCAGGCTGGAAGGCTCCCGGCTATCTCCAGTCTCTCAAGACCAAGATCTGAGCGCTTCCGCTCCAGAGACCCCGAGATGTGTATCTACCA
 GTACATGAGACCTGTACGTGGGCTCCGACCTGGAGATCGGCGACCGCGCAAGATCGAGAGCTGGCGGCCACTGTGTGCTGGCTGGGCTTACCA
 CCCCACAAAGAACACAGAGAGAGCCCGCTTCTGTGGATGGGCTACAGCTGACCCGACAGTGGACCTGTGACCGCATCAAGTGTCCCGAGAG
 GACTCTGGACCTGAGACATCCAGAGCTGCTGGGAGCTGACTGGGCTCCCGACATCTACCGGSCATCAAGTGAAGCACTGTGTGCAAGCTGCT
 GCGGCAACCAAGGCTTACCTGACCTTGTGACCTTCCCAAGAGAGCGGAGCTGGAGCTGGAGAGACCGGAGATCTGAGAGACCCCGGCTGACGCGG
 TGTACTAGACCCCTCCAGGACCTGATCGCGAGATCCAGAGCAGGCGCAGGACCTGAGACCTACAGATCTACCAAGAGCCCTTCAAGACCTGAAG
 ACGGCAAGTACGCGAGGCGAGTCCACCCACACAGAGCTGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG
 CAGAGCCCAAGTCCGCTGCTGATCCAGAGGAGACCTGGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG
 CTGCGCAAGCTGCTGACCGAGCGCGCGCAGAGATCTGCTGCTGACCGAGACCCACACAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG
 CTTGCAAGACTCCGCGCTGGAGGTGAACATCTGACACTCCAGTACGCGCTGGGCTATCCAGGCGCGAGCGCTCCGAGTCCGAGTCCGAGTGGTG
 ACCAGATCATCGAGAAGCTGATCGAGAAGAGCGGCTGACTGCTCTGCTGCTGCGCGCCACAGGSCATCGGCGCAGGAGCTGAGAGAGCTGAGAG
 TCTCCGGCATTCGCAAGGCTGTCTGAGAGCGGCTGAGAGAGCGGCTGAGAGAGCGGCTGAGAGAGCGGCTGAGAGAGCGGCTGAGAGAGCGGCT
 CAACCTGCGCGCATCGTGCGCAAGGAGATGCTGGCTCTCTGAGACAGTSCAGAGTGAAGGCGGAGCGCATTCACCTCCAGCTGAGAGCTGAGAG
 TCTGGAGCTGAGTGCACCTGGAGGCAAGGTGATCTCTGCTGGCTGGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCT
 GGCTAGGAGACCGCTACTTCTGAGCTGGCGCGCGCTGGCGCTGAGAGTGCACACGCAACGCGCGCTGAGAGTGCATGAGAGAGCTGAGAGAGCT
 GCGCGCTGCTGTGTGGCGCGGCTGAGAGAGTTCGGGCTTCCCTACACCGCGCTGAGAGTGCATGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG
 TCATCGGCGAGTGGCGAGCGGAGCACTGAGAGCGGCTGAGAGTGCATGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCT
 TCGCGCGGAGGCGATCATGACATCATTCGCGACCGACATCCAGAGCAGGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG
 CGACTCCGCGAGCCCATCTGAGAGGCGCGCGCAGCTGCTGAGAGGCGGAGGCGCGCTGAGTCCAGAGCACTCCGACATCAAGTGTGTGCGCC
 GCGCAGGCGCAAGATCATCCGCGACTAGGCGCAAGCAGATGCGCGCGCGGAGAGCTGCTGCTGGCGCGCGCGAGGAGGAGCTAA

도면111a

66. 2003 CON B pol. pep

FERDLAFQCKAREFFSSEQTRANSPTRRELQWGRDNNSLSEACADROQTVSFSPQITLWQRPLVTIKIGGOLKEALLDTGADDTVLEEM
 NLPGRWKPMIGGIGGFIVKVRQYDQILIEICGHRKAGIVLVGPTPVNIIGRNLLTQIGCTINFPISPIETVPVKIKPGMDGPKYQWPLTEE
 KIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKSTKWRKLVDFRELNKRITQDFWEVQLGIPHPAGLKKKKSVTVDVGDAYFSVFLD
 KDFRKYTATTIPSNNETGIRYQYNVLPQGWKGS PAIFOSSMTKILEFFRKQNPDIYIYQMDLQYVSDLEITGQHRTKIEELRQHLLRWG
 FTTDPKKHOKEPFFLMMGYELHPDKWTVQPIVLPEKDSWTVDIQLVGLKNWASQIYAGIKVKQLCKLLRGTKALTEVILPTEEAELLEAE
 NRELLKEPVHGVVYDPPSKOLIAETQKOGOGQWTVQIYQEPFKNLTKGYARMRGAHTNDVKQLTEAVQKIATESIIVWGKTPFKLPIQKET
 WEAWTXYWQATWIPWEFEVNTPELVKLWYQLEKEPIVGAETFYVDGAANRETILGKAGIYVDRGRKVVSLTDTTQKTELQAIHLAQDS
 GLEVNIVTDSQYALGIIQAPDKSESELYSQIIEQLIKREKVYLAWVPAHKGIGGNEQVDKLVASAGIRKVLFLDDGIDKAEQEEHEKYHNSWRA
 MASDFNLPEWVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLECKIILVAHVHVASGYTEAEVIPAETQOETAYFLLKLAGRWPKTI
 HTDNGSNFTSTTVKAACHWAGIKQEFGPYNPQSGVVESENMKELKIIGQVRDQAEHLKTAVQMAVFIIHFRKKGIGGCYSAGERIVDIIA
 TDIQTKELOKQITAKIQNFRVYIRDSRDFLWKGEPAKLLWKGEAVVIQDNNDIKVVPRRKAKIIRDYGKQMGAGDCVASRQDEDS

도면111b

2003_CON_B_pol.opt

TTGTTGGGAGGACCTGGGCTTCCCCAGGGAGAGGCGCGGAGTTCTCTCCGAGAGAGCCCGGCGCAATCCCCCAACCGCGCGAGAGTGGAGGTGTG
GGCCCGGACACAACTCCCTGTCCGAGCGCGCGCGGACCGGACAGGCGACCGGTCTCTCTCCCTCCCGAGATCACCTCTGGCAGCGCGCCCTGTGTGA
CATTCAAGATCGGCGGCGAGCTGAGGAGGCGCTGTGACACCGGCGCGCGAGAACCTGCTGGAGAGATGAACTGCTGGCGCGCGCTGGAGCGCAAG
ATGATCGGCGGCTTCATCAAGTGCAGATACGACAGATCTGATCAGATCTCGGCGCAGAGCCCTGCGCACCCTGCTGTGTGGCGCC
CACCCTGTGAATCATCGCGCGACCTGTGACCAATCGGCTGACCTGAGTTCCCTATCCCTCCATCGAGACCGCTGGCGGTGAAGTGAAGC
CGGCTGAGAGCGCCCAAGGTGAAGTGCCTGTCGAGAGAGATCAAGGCTGAGATCTGACAGATCGAGATGAGAGAGGAGGCAAGTATC
TCCAAGTCGGCGCGAGACCCCTACAAACCCCGCTGTTCGCTCATAGAGAGACTTCCAGATGGCGGCAAGCTGTGGAGCTTCCGCGAGCTGAA
CRAGGACCCAGGACTTCTGGAGGTGAGCTGGGCTCCCGCGCGCTCAAGAGAGAGTCCGTGACCTGTGGAGCTGGCGGAGCTGGCGGAGCTG
ACTTCTCGGTGCGCTGGACAGGACTTCCGCAAGTACACCGCTTACCATCCCTCCATCAACAGAGAGACCCCGGCAATCCGATCAAGTAAAGTG
CTGCGCAGGCTGGAAGGCTCCCGCGCTTCCAGTCTCCATGACCAATCTGGAGCCCTCCGCAAGAGAACCCCGATCTGTGATCTACCA
GTACATGGACGACTGTACGTGGCTCCGACTGGAGATGGCGACACCGACCAAGATGAGAGCTGGCGCGACACCTGCTGTGGCTGGGCTTACCA
CCCCGACAGAGACACAGAGAGGAGCGCCCTTCTGTGGATGGGTAGAGTGGAGTGCACCCGCAAGTGGACCTGCTGTGGCTGGCGCTGCAAG
GACTCTGGACCGTGAGGACATCCAGAGCTGTGGCAAGTGAATCTGGGCTCCAGATCTAGCGCGGATCAAGGTGAAGCTGTGCAAGCTGT
GGCGGACAGGCGCTGACCGAGTATCCCTGACCGAGGACCGAGTGGAGTGGCGCGGAGAACCGGAGATCTGAGAGGAGCGCTGTGACGCGG
TGACTACGACCCCTCCAGGACCTGATCGCGAGATCCAGAGCAGGCGCAGGCGAGTGGACCTACCAAGATCTACAGGAGCGCTTCAAGACCTGAAG
ACCGCAAGTACCGCCCGATCGCGCGCGCCACACCAAGACCTGAAGCTGAGAGCTGAGCGGCGGAGAGTCCCTGAGAGAGTCCGATCTGTGGG
CAAGACCCCAAGTCAAGCTGCCATCCAGAGAGAGACTGGAGGCTGTGGACCGAGTACTGGCAGGCGACCTGGATCCCGAGTGGAGTTCGTGA
ACACCCCGCGCTGTGAGTGTGTACCGAGTGGAGAGGAGCGCTCCAGTACCGCTGGGCTGGTGGCGCGCGAGACCTTACCTGGAGGCGCGGAGAG
CTGGCAGCGGCTAGTCAAGAGGAGTGTGTCTCCCTGACCGACACCAACCGAGAGTCCAGGCGCGCGCGAGAGTCCGAGTCCGAGTCCGAGTGT
CCGAGATCTCGAGCAGCTGATCAAGAGGAGAGTGTACCTGGCTGGTGGCGCGCGCGAGAGTCCGCGCGCGAGAGTCCGAGTCCGAGTGT
TCCGCGCGCATCCGCAAGGTGCTGTCTGGAGCGGATCGACAGGCGCGAGGAGCGGAGAGTACCACTCCAGTCCGCGCGCGAGTCCGCGCGAG
CAACTGCGCGCGGTGGTGGCGCAAGGAGTCTGTGGCTCTCTGGAGAGTCCAGCTCAAGGCGGAGGCGGAGTCCGCGCGAGTCCGCGCGAG
TGTGCGAGCTGGACTGCAACCTGGAGGCGAGATCATCTGTGGCTCTCTGGAGAGTCCAGCTCAAGGCGGAGGCGGAGTCCGCGCGAGTCCGCGCGAG
GGCGCGCGCTGTGTGGCGCGCGCATCAAGCAGAGTTCGGCATCCCTTACACCGCGCTGGAGTCCGCGCGAGTCCGCGCGAGTCCGCGCGAG
TCATCGCGCGAGTTCGGCGCGAGCAGCTGAAGACCGCTGACAGTGGCTGTCTATCCCACTTCAGCGCAAGGCGGCGCATCGCGCGGCTAC
TCCGCGCGAGCGCTCTGGAGATCTTCGCCACCGAGATCCAGAGCTGAGAGGAGTCCAGAGTCCAGAGTCCAGAGTCCGCGCGAGTCCGCGCG
CGACTCCCGCGAGCGCTGTGGAAGGCGCGCGAGCTGTGGAAGGCGGCGGCTGTGTGATCCAGGAGCACTCCGAGATCAAGTGGTGGCGCG
GGCGAGGCGCAAGATCATTCGGCGACTACGGCAAGAGATGGCGCGCGAGTGTGGCTTCCCGCGCGAGGAGCTAA

도면111c

67. 2003 p.anc pol.pap

FFRENIAFTQKAREFSSEQTRANSPTRRELQVWGRDNNPLSEACADROGTVSEFPQITLWQRPVLTIKIGGQKKEALLDTGADDTVILEEM
 NLPQKWKPMIGGIGGFIKVRQYDQILIEICGHRKAIGTVLVGPTFVNIIGRNLLITQIGCTLNFPISPIETVPVKLPGNDGPKVKQWPIITEE
 KIKALVEICTEMEKEGKISKIGPENPNYNTPVFAIKKQSTKWRKLVDFRELNKTODFEVQVQIGIPHPAGLKKKSVTVLVDVGDAYFSVPLD
 KDFRYTATIP SINNETPGIRYQYNVLPGWKGSFAITQSSMTKILEFPERKQNPETVIYQYMDLIVGSDLELGQHRKIEELREHLRWG
 FTFPKKHQKEPPFELWNGYELHPDKWTVOPIVLPEKDSWTVNDIOKLVGKLNWASQIVAGIKVKQKCLLRGTAKLTAEVYVPLTTEEALELAE
 NREIIEKPEVHGYYDPPSKDLIAEIQKQGGQWTVQIYQEPFKNLKTGYARMRGHTNDVKOLTEAVQKIATESIIVGKTPKFKLPQKET
 WEAMTEYQATWIPWEFVNTPELVKLVQLEKEPIVGAETFYVDGAANRETGLKGAGVTDGRQKVVSLLTDTTNQTELQAIHLALQDS
 GLEVNIVTDSQYALGIIQAPDKSELSVSQIIEQLIRKEKVILAWVPFAHKGIGGNEQVDKLVSAGIRKVLFDGIDKQAEHEHEKYHSNWRA
 MASDFNLPPVWAKETIVASCDKQKGEAMHGQVDCSPGIWQLDCTHLEKILIVAVHVASGYTAEVTPAETGQETAYFILLKLAGRWPVKVI
 HTDNGSNFTSTTVKAACWAGIKQEFGIYPNPQSGVSVSMNKKKIIIGQVRQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIVDIIA
 TDIQTKELQKQITKIQNFRVYVYRDSRDPINWGPAPKLLMRGEGAVVIQDNSDIKVPVPRKAKIIRDYGRQMGAGDDCQVSRQDEDS

2003 B. and pol. OPT

[illegible]

도면112a

68. 2003 COM C pol. per

FFRENLAFTQGEAREFPFSQTRANSPTSRELOVRGDNPRSEAGAERQGTINFPOITLMQRPLVSIKVGQIKKALLDTGADDTVLEENLPG
 KWPKMIGGIGGFIKVRQYDQILLIEICGKKAIGTVLVGPTPVNLIGRNMLTQCGTTLNFPISPTETVPVKLPGMDGPKVKQWPLTEEKIKA
 LTAICEEMEKEGKITKIGPENPYNTPVFAIKKRDSTKWRKLVDFRELNRKTQDFWEVQIGIPHPAGLKKKSVTVLDVGDAYFSVPLDEGFR
 KYTATTPSINNETPGIRYQYNVLPQGWKSPAIQSSNTKIIEFPFRAQNPETIYQYMDDLVVGSDLEIGQHRAKIEELREHLLKWGFTTP
 DKKHQKEPFLMGGYELHPDKWTVQPIQLPEKDSWTVNDIQLVGKLNWASQIYEGIKVRQCKLLRGAKALTDIVPLTEEALELEAENREI
 LKEPVHGVYDPSKDLIABLQKQCHDQWTYQIYOEPFNKLTGKYAKMRTAHNDVQOLTEAVQKIAMESIVIWGKTPKFRPLPIQKETWETW
 WTDYQATWIPPEWETVNTPLVLWYQLEKEPIAGAEITYVDGAANRETIGKAGYVTDGRQKIVSLTETTNQKTELQAIQIALQDSGSEV
 NIVTDSQYALGIIQAQPKSESELVNOITQOLIKKERVYLSWVPAHKGITGNGEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRWASE
 ENLPPIVAKETIVASCDCQOLKGEALHGVDCSPGIWQLDCTHLEGGKIILVAVHVASGVIEAEVIPAETQETAYIILKLAGRWPVKVTHTDN
 GSNFTSAAVKACWAWAGIQOEFGIPYNPQSGVVESMNKELKIGQVRDQAEHLKTAVOMAVFIHNFRRKGGIGGYSAGERIIDIIATDIQ
 TKELQKQIKIKIQNFRVYRDSRDFIWKGPAKLIWKGEGAVVIQDNSDLKVVPRRKAKIKDYCKQMACDCVAGRQDEDS

2003 CON C pol.OPT

[illegible]

도면112c

69. 2003 C. and pol. pep

FFRENLAFFQGEAREFPSEQTRANSPTSRELQVGRDNPRSEAGAEQGTLTINFPQITLLWQRPVLVSIKYGGQIKEALLDTGADDTVLEENL
 PGKWKPKMIGGIGGFIKVRQYDQILJEICGKKAIGTVLVGTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPKPGMDGPKVKQWPLTEKI
 KALTAICEEMKEGKITIKIGENPNTPVFAIKKDDSTKWRKLVDFRELNKRTOQDWEVOLGIPHPAGLKKKSVTVLDVGDAYESVPDEG
 PRKYTAFTIFSINNETPGIRYOYNVLPQGWKSPAI FQSSWTKILEPFRQNPEIYIYQWMDLIVGSDLEIGQHRAKTEELREHLLKMGFT
 TPDKKHQKEPPFLMGGYELHPDKWTVPQIQLPEKDSWTVNDIQKLVGKLNWASQIYPGKVRQCKLLRGAKALTDIVPLTERAELELAENR
 EILKEPVHGVYDPSKDLIAEIQKQGHQWTYQIYQEPFKNLKTGKYAKRMTAHTNDVKOLTEAVQKIAMESIYVWGKTPKFRLPIQKETWE
 TWWTYQATWIPPEWFFVNTFPLVKLWYQLEKEPIAGAEYFYVDGAANRETKIGKAGYVTDRCQKIYSLTETNQKTELQALQALQDSGS
 EVNIVTDSQYALGIIQAOQPKSESELVNOIIEQLIKKEKVVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEHKEHYHNSNRAMA
 SEFNLPIIVAKEIVASCDCQKQKGEAMHGQVDCSPGIWQDCTHLEKILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHT
 DNGSNFTSAAVKAAQWAGIQQEFQIPYNPQSQGWESMKNELKILIGQVRDQAEHLKTAVQNAVFHNFRKGGIGGYSAGERIIDIIATD
 IQTKRELQKLIKIQNFRVYVYRDSRDP IWKGPALKLNKSEGAVVIQDNDIKVVFRRKAKIIRDYQKQWAGADCVAGRQDEDS

2003_c.and pol.opt

TTCTTCCGCGAGAACTGGCCTTCCCCGAGGCGAGSCCGCGAGTTCCCTCCGAGAGACCCCGGCCAATCCCCCACTCCCGGAGAGTGCAGGTGGG
CCGCGACAAACCCCGCTCCGAGGCGGCGCCAGAGCCAGGCAACCTGACCTTGAAGTATCCCCAGATCAACCTGTGGCAGCGCCCCCTGCTGTCTCAATCA
AGTGGGCGGCCCAATCAAGGAGGCCCTGCTGGACACCGGCGCCGACACAGACTGTGGAGGAGTCAACCTGCCGCGACGTGGAGGCCCAAGATGATC
GGGGGCAATCGCGCGCTTCACTAAGTGGCGAGTAGACAGATCTGATCGAGATCTGGCAAGGCAAGGCCATCGGCATCTGCTGGCGCCCCCAAC
CTGTGAACATCATCGGCGCAACATGCTGAACCAAGTGGGCTGCACCTGAAATTCCTCCATTCGAGACCTGAGACCTGCGCTGAAGCTGAAGCGGCA
TGAAGCGCCCAACAGGTGAAGAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCGGCATCTCGAGAGGAGATCTGGAGGAGGCGAGATCACCAAG
ATGCGCCCGAGAACCCCTCAACAACCCCGCTGTCCGCAATGAAGAGAGGATCAACCAAGTGGCGGCACTGCTGGAGCTTCCGCGAGTGAACACG
CACCCAGGACTTCTGGAGTGCAGCTGGGATCCCAACCCCGCGGCTTGAAGAAAGAGTCCGCTTCCATCAACAGAGACCCCGGATCTGAGTGGGCTTCA
CCGTGCCCTTGGACAGAGGCTTCCGCAAGTACACCGCTTCCACATCCCTCCATCAACAGAGACCCCGGATCTGAGTGGGCTTCAACCAAGTACAT
CAGGCTGGAAGGCTCCCGGCCATCTTCAGTCTCCATGACCAAGATCTCGAGGCTTCCGCGCCAGAACCCCGAGATCTGATCTACAGTACAT
GAGACCTGTACTGTGGCTCCGACTTGGAGATCGGCGAGTACGAGTGGCTAGCAGTCCGCGGAGTTCGAGAGTGGAGTGGGCTTCAACCAAGTACAT
ACAAGAGCACCACAAGAGAGCCCCCCTTCTCTGGGATGGGCTAGCAGTCCAGCCCGAGATTCGAGAGTGGAGCTGAGTGGGCTTCAACCAAGTACAT
TGGACCGTGAACGACATCCAGAGAGTGTGGGCAAGCTCCCAATCTACCGGATCTACCGGATCAAGCTGGCGGATCTGCGAGTGGCTGGAGTGGCT
GSCCAAGGCCCTGACCGACATCTGSCCCTGACCGAGAGGCGGAGCTGAGCTGGGCGAGAACCGGAGATCTTACAGAGGACCTTCAAGAGCTGAAGACCGG
AATACCCCTCAAGAGCTGACACCCACACAGCAAGTGAAGCAGTGAAGCAGTGAAGCGGCGTGACAGATCGCATGAGTCTCATCTGGGAGTCTGGGCGAAGAC
CCCCAGTTCGGCTGGCCATCCAGAGAGAGACTGGAGAGCTGTGTGACCGACTTGGCAGGCCACTGGATCCCGAGTGGGAGTTCGTGGAACAGCC
CCOCCCTGTGTGAAGCTGTGTGTTACAGCTGGAGAGAGGCCCATCTGCGGCGCCAGACTTTCAGTGAAGGCGCGCCCAACCGGAGACCAAGATCGGC
AAGGCGCTGACTGTGACCGCGGCGCGAGAACTGTGCTTACCGAGACCAACACAGAGACCAAGCTGAGGCGATCGAGGCGATCCAGCTGGCCCTGCA
GAGCTCGGCTCCGAGTGAACATGTAAGCTCCAGTACCGCTTGGGATCATACGGCCAGACCGGCAAGTCCGAGTCCGAGCTGGTGTGAACAGGA
TCTATCAGAGCAGCTGTGAGCTGAACATGTAAGCTCTGCTGGTTCGCGCCACAGGCGATCGGCGGAGGCGATCGAGGAGTGGAGTGGTGTGCTTC
GGCATCCGCAAGGTGTGTTCTTGACCGGCATCGACAAGGCCAGGAGGAGCAGAGATCACTCAACTGGCGCGGCATGGCTCCGAGTTCAACTT
GGCCCCCATCTGTGGCGAAGGATGTGGTGCTCTGCAAGTGGCGATGAGGCGGCGCATGACGGCAGGTCGACTGTCGCCGCGATCTGGC
AGTGTGACTGCACCACTGAGGCGCAGATCATCTGTGGCTGTACAGTCCCGCTCATCAACCCCAAGTCCGAGGCGTGGTGGAGTCCATGACAGGAGATCGGCGAG
GAGAGCGCTTACTTATCTGAGCTGGCGCGCTGGCTGGCTGAGGTGATCCACACACAGGCTCCAACTCCGCGCGCTGAGGCGGAG
CTGCTGGTGGCGGCGATCAGCAGGATTCGGCATCCCTCAACCCCAAGTCCGAGGCGTGTGATCCACAACCTCAAGCGCAAGGCGGCATCGCGGCTACTTCCGC
GGCAGGTGGCGACAGGCGAGCAGCAGTGAAGCCCGCTGCAATGGCGGTGAGAGGAGTTCGACAGCAGATCATCAAGATCCAGGATGGAAGAGATCATCG
CGCGACCCCATCTGGAAGGCCCGCGCAAGTGTGTGAAGAGGCGAGGCGCGCTGGTGGTATCCAGAGCAACTCCGACATCAAGTGTGCTCCCGCGGCA
AGGCGAAGATCATCCGCGCATCGGCAAGCAGATGGCCGGGCGGACTGCTGCTGGCGGCGCGGACAGGAGACTTAA

도면113a

70. 2003 COM D pol.PEP

FFRENLAFFQKAGELSSSEQTRANSTSRRLRVWCGDNPLSETGAERQQTVSFNFPQITLWQRPLVTTIKIGGQKEALLDTGADDTVLEEIN
 LFGKWKPKMIGGIGGTIKVQYDQILIEICGHKAICTVLVGPPTPVNIIGRNLLTQIGCTINFPISPIETVPVKLPGMDGPKVKQWPLTEEK
 IKALTEICTEMEKEGKISRIGPENPNYNTPIFAIKKDDSTKWRKLVDFRELNRKRTQDFWEVQLGIPHYPAGLKKKSVTVLDVGDAYFSVLDE
 DFRKYTAFTIPSINNETFGIRYQYNVLFQGWKGSFAIFQSWTKILEPFRKQNPETVIYQYMDLTVGSDLEIGQHRRTIEELREHLLRWGF
 TTPDKKHQKEPPFLWGYELHPDKWTVPFIKLPEKESWTVNDIOKLVGKENWASQIYPGKVRQCLKLRGTKALTEVPLTERAELELAEN
 REILKEPVHGYYDPKSKDLIAEIOKQOGQWTYQIYQEPKNLKTCKYARMGAHTNDVKQKTALESIVWGKTPKFLPIQKETW
 ETWWTYQATWIPWEFEVNTFFLVKLWYLEKEPIIGAETFFYDGAANRETTLKAGYVTDGRQKVPLTDTTNQKTELQAINIALQDSG
 LEVNAVTDQSQYALGIIQAQPKSESELVSQIEOLIKKEKVILAWPFAHKGIGGNEQVDKLVSGIRKVLFDGIDKAQEBEHKYNHNNWRAM
 ASDENLPPVVAKEIVASCDKQOLKGEAMHQVDCSPGIWOLDCTHLEGVILVAVHVASGYIEAEVIPAETGQETAYFLKLKLAGWPVKVWH
 TDNGSNFTSAVRKACWAGIKQEFGIYPNPQSQGVVESNMKELKKIIGQVRDQHEHLTAVOQMAVFIHNFRRKGGIGGYSAGERIIDIIAT
 DIQTKELQKQIIKIQNFRVYRDSRDPIMKGPALKLWKGEQAVVIQDNSDIKVPFRKVKIIRDYGMAGDDDCVASRQDEDS

2003_CON_D pol:OPT

TTCTTCGGGAGAACCTTGGGCTTCCCCGAGGGGAAGGCCGGGAGCTGTCTCTCGAGGACAGACCCGGGBCAACTCCCCCACTCCCGGAGCTGGCGCTGT
GGGCGGAGCAAGCCCTGTTCGAGACCGGCGCGAGCCACGCTGTCTTCACTTCCCCAGATCACTCTGCTGGAGGAGATCAACTGCTCCCGGCAAGTGGAGSCCAAGATG
TCAGATTCGGGCGAGCTTGAAGGAGGCCCTTGTCTGGACAACGGCGCCAGACACCTGTCTGGAGGAGATCAACTGCTCCCGGCAAGTGGAGSCCAAGATG
ATCGGCGGCTATTCGCGGCTTTCATCAAGTGTGGCCAGTACGACAGATCTCTCATCGAGATCTGCGCCCAAGSCATCGCCCTGCTGTGGGCCCCAC
CCCGTGAACTCATTCGCGCGCACTCTGTACCCAGATCGGCTGACACCTTGAACTCCCATCTCCCATTCGAGACCTGGCCCTGAAGCTGAAGCTG
GCATGGACGGCCCAAGGTGAAGAGAGTGGCTCTACCTGAGGAGAGATCAAGGCGCTGACCGAGATCTGACCCGAGATGGAGAGAGGCAAGATCTCC
CCATCGGCCCGGCAATCTCAACAACCCCATCTTGGCATCAAGAACGAGATCTCAACATGGCAGAGCTGTGTGACTCTCCGCGAGTGAACAA
GGCACAACCCAGCATCTGGGAGGTGAGCTGGCATCTCCACCOCGCGCTCGAHGAGAAGAACTCCCTGACCTGTTCGAGCTTGGCGAGACGCTACT
TCTCTGTGCCCTTGAGAGAGACTTCGCGAAGTACACCGCTTACCATCCCTCCATCAACAAGAGACCTCCGCTACCGATACAGACAGCTGTG
CCGACGGCTGGAGGGTCCCGGCATCTTCCAGTCTTCATGACCAAGATCTCTGGAGCCCTCCCGAAGCAGAACCCGAGATCTGATCTACAGTAA
CATGGACACTGTACGTGGGTCCGACTGGAGATGGGCGACCGACACAGATCAGAGAGCTGGCGAGACCTGCTGGCTGGGCTTCAACACCT
CCGAGAGAGCAACAGAGAGGCGCCCTTCTCTGTGGATGGCTACGAGTCCGACCCCGACCCGCAAGCTGGACCTCTCAAGTGCCTCAGACAC
TCTCGGACCGTGAACACATCCAGAACTGTGTGGGAGCTGATCCCCCTGACCGAGGCGGAGTCCGAGATCCGAGAGAGAGAGAG
CGSCACTACGCCCTCCAGGCTGATCGCTGGATCCAGAACGACGCTGAGCACTGACGAGCTGACGACCCCGCAAGCTGGACCTCTCAAGTGCCTCAGAC
GACCCCAAGTTCGCTCTGCCATCAAGACCATCTGGAGACTGTGGAGCTGTGAGCGAGTACTGCGAGCGACCTGGATCCCCGAGTGGAGTCTGGACAA
CCCCCTCTGGTGAAGCTGTGGTACGAGCTGGAAGAGCCCATTCATCGCGCCGAGACTTCTAATGAGGCGCCGCCCAACCGGAGACCAAGCTG
GGCAGCGCGCTAGTACGCAACCGGCGCGCAGAGCTGTGCCCTTACCGACACCAACAGAGACAGGCTGAGGCTGACGCTCTCAACTGAGCCT
CGAGACTCGGCTGAGSCTGAACATCTGTGAGGACTCCCATGCTGCCCTGGGCTATCCAGGCCATCCAGGCCCAAGCCGACAGCTCGAGCTGTGCTCC
AGATCATCGAGACTCATAGAGAGAGAGAGCTGTACCTGGCTGGGCTCCGCGCCCAAGGGCATCGGCGGCAAGCTGAGGCTGAGCTGTGCTCC
AAGGCACTCGCAAGTCTCTTCTGGAGGCTATCGAACGGCCAGGAGGACAGAGAGTACCAACAATGCGCGGCCCATGGCTTCGAGCTTCAA
CTTCCGCCCTGGCGCGCAGAGAGATGTGGCTCTCGCAAGTGTGAGCTGAGGCGGCGGCTGACCGGCGAGGCTGGAGTGTCTCCCGGAGATCT
GGCAGTGGACTGCACCACTGAGAGGCAAGSTGATCTGTGTGGCGGTGAGCTGGGCTTCCGCTATCAGGCGCGAGTGTATCCCGCGCGAGACCGG
CAGGAGAGCGCTACTTCTCTGAAGCTGGCGGCGCTGGCCCTGGGCTGGGCTGGGCTGAGGCTGGGCTGACCGGCGAGTGGAGTGTCTCCCGGAGATCT
CGGCTGCTGGTGGCGCGGCACTACAGCAGAGTTTCGCACTCCCTCAACCCCACTCCAGGCGGCTGGTGGATCTCATGAACAGGAGTGAAGAGATCA
TCGGCCAGGTGGCGGACACCGCGAGGACCTGAAGACCGCGGTGAGATGGCTGTTCATCCAACTTCAAGCCCAAGGGCGGCTACGGCGGCTACTCC
GCTCCGCGGACCCCATCTGGAGAGGCGCCGCCAAGCTGTGGAAGGGCGGCGGCTGGTGGAACTCCGAGATCCAGAACTTCCGCGTGTACTATCCGCGA
DUAAGTGAAGATCATCCGCGACTACCGGCGAGGATGGCGCGGAGGATGGTGGCTCTCCGCGAGGACGAGACTTAA

도면114a

71. 2003 CON F1 pol. PEP

FFRENIAFQGEARKEFPSEOTRANSPASRELRVQGRDNELSEAGAERRCTVPSSLFPPQLTWORPLVTIKIGGOLKEALLDTGADDTVLEDI
 NLPKWKPMIGGIGGFIAVKQYDHLIELICGHRKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPIPIETVPVKLPGMDGPKVKQWPLTEE
 KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDDSTKWRKLVDFRELNRKTQDFWEVQLEGIHPFAGLKKKSVTVLDVGDAYFSVPLD
 KDERKYTATLPSYNNETPGIRYOYNVLPQGWKGSPIITQCSMTKILIEPERTKNPDIVIOYMDLIVGSDLEIGQHRTKIEELREHLLKWG
 FTFDDKKHKEPPELMMGYELHPDKWTVOPIQLPDKDSWTVNOIQLVKLNWASQIXPGIKVKQOLCKLLRGAKALTDIVPIATAEAELELAE
 NREILKEPVHGYYDFPSKOLIAELIQOGQWTFYIOYEPFKNLKTGYAKMRSAHTNDVKOLTEAVQKIALESIIVWGKTPKFRLPILKET
 WDTWNTDYQATWIPWEWTFVNTPELVKLVQLETEPIVGAETFYVDGASNRETKKGACGYTDEGRQKVVSLETTNQAELOAIHLALQDS
 GSEVNIVTDSQYALGIIQAPDKSESELVNQIIEQLIOKEKVIYLSWVPAAHKGIGGNEQVDKLYSAGIRKILFLDGDKAQEEHEKXHNWRA
 MASDENLPPVWAKEIVASCDKCOLKCEAMHGQVDCSPGIWQLDCTHLEGIILVAVHVASGYIEAEVIPAETGQETAVFILKLGRWVPVKII
 HTDNGSNFTSAAVKAACWAGIQQEFGLPYNPQSGVVSNNKELKKIIGQVRDQAEHLKTAVQMAVFHNFKRKGGIGGYSAGERIIDIIA
 TDIQTRELOKQITKIQNFRVYYRDSRDPVWKGPAKLLMKGEGAVVIQDENSEIKVVPRRKAKIIRDYQKQMGAGDDCVAGRQDEDS.

도면114b

2003 CON F1 pol.OPT

[illegible]

도면115a

72. 2003 CON F2 pol.PEP

FFRENLAFCQGEARKESSEQTRANSASPARELVRGRDNLPEAGAERQGTGSSLDFFPQITLWQRPVITIKVGGQIREALLDTGADDTVLEDI
 NLPKWKPKMIGGIGGFIVKVRQYDQIPETCGQKAIGTVIVGPTPVNIIGRNMLTQIGCTLNFPISPETVPYKLRPGMDGPKVKQWPLTEE
 KIKALTEICTEMEREKISKIGPENPNYTPVFAIKKOSTKWRKLVDFEKLNRKTQDFWEVOLGIPHPAGLKKKSVTVLDVGDAYFSVPLD
 KEFRKYTAFTIPSNNETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEFFRAKNPEIYIYQXMDLIVGSDLEIGQHRKXIEELREHLLRWG
 FTTDPKKHOKEPPEFLMMGYELHPDKWTVQAIQLDPKSSWTVNDIQLVGLKNWASQIYPGIRVRKHLCKLRLGAKALTDVVPLTAEAELELAE
 NREILKEPVHGVYDPSKDLIAETQKQGHDOWTYQIYQEPHKNLKTGYARRKSAHTNDVKQLTEVVQKIATEGLVINGKVPKRLPIQKET
 WEIWMTEYQATWIPWEFEVNTPELVKLWYQLETEPIVGAETFYVDGAANRETKLKGAGYVTDGRQKVVPLETNTNQTELQAIHLALQDS
 GSEVNIIVTDSQYALGIIQAHPPDKSESELVNOIIRQLIQKERVYLSWVPAHKGIGNEQVDKLYSTGIRKVLFLDGDIDKAEHEKHYHGNRA
 MASDENLPPVVAKELVASCDKCOLKGEAMHGQVDCSPGIWQLDCTHLECKIILVAVHVASGYIEAEVIPAETQETAYFILKLAGRWPKII
 HTDNGSNFTSTVVKACWAGIQQEFPIYFNPQSGVVESMKNELKKIIGVRODAEHLKTAVMAYFIHNFRRKGGIGGYSAGERIIDI
 TDIOTKELQKQITKIQNERVYFRDSRDPVWKGPAPKLLMRGEGAVVIQDNEIKVYPRRKAIIIRDYGMAGDDCVAGRQDEDS

2003_CON_F2_POL.OPT

TTTCTTCGGGAGAACCTTGGCCCTTCCAGCABGGCGAGGCCCGCAASHTTCTCTCTCGACGACAGACCCCGGCGCAACTCTCCCGGCTCCCGCGGAGACTGCGCGGTGGCC
 GCGCGCGAGAACTCTCTCGCCGAGGCGCGCGCGAGCGAGGCAAGCGGCTCTCTCTGGACTTCCCGGAGATCACCTCTGTGGACGCGGCGGCGGCGGCGGCTGTGA
 CCATCAAGGTTGGGCGGAGCTGGCGAGGCGCTGTCTGGACACCGGCGCGCCACGACACCTGCTGGAGGACATCAACTGCTCCGGGCAAGTGGGAAGCCGCGCAAG
 ATGATCGGGGGGATTCATCAAGGTTGGCGAGTACGACAGATCTCCCATCGAGATCTCGCGCGAGAGCGCATCGGACCGCGCTGTGTGTGGGCC
 CACCCCGGTGAACATCATCTGGCGCGCAATGCTGACCCAGATCGGCTGCACTTCTCCCATCTCCCATCGAGCCGCGGCGGCGGCGGAGCTGAAGC
 CGCGGATGAGCGGCGCCAAAGTGAAGTGGCTGGCCCTGACCGGAGATCAAGAGGAGCTTCCACCGAGTCTGCACCGAGATCGAGGAAGGAGGCGCAAGATC
 TCCAGATCGGCTCGGACACCTCTAGACACACCCCGTTTGGCATAGAGAGAGACTTCCACGATGCGCGAAGCTGTGTGGATCTTCGCGGAGCTGAA
 CAAAGGATCCCAAGBACTCTTGGGAGTGGAGTGGGCATCTCCACACCCCGCGGCTTAAGAGAAGAGAGTCCGTGACCCGTGTGTGAGCGTGGGCGACGCGT
 ACTTCTCTCGTCCGCTTGGAGAGAGTTCGCAAGTACACCGCTTCCACATCCCTCCATCAAACAGAGACCCCGGCACTCCGCTACCAGTACCAAGCTG
 CTGCGCCGAGGCTGAAGAGCTTCCCGCGCATCTTCCAGTCTCCATGACCAAGATCTCTGGAGCCCTTCCGCGCGCAAGAACCCCGAGACTGTGATCTTACCA
 TCCCGGACAGAGACACAGAGAGAGCCCTCTCTGTGTGATGGGTACAGCTGCACCCGACAAAGTGGACGCTGTGAGGCCATCAAGCTGCCCGCAAG
 TCTCTCTGGACCTGAGACATACAGAGACTGTGTGGCAAGCTGAATGGGCTCCAGAGTACCCCGGATCTACCCCGGATCTGGCTGAGACATCTGTCCRANGTCT
 GTGTGGGCGCGCAAGGCTGACCTGACCTGAGCTGGTGGCTTACCGCGGAGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCGCTGACACGGG
 GTGTATTCAGACCTCCAGAGACTGTATGTGGGATATGGGAGAGAGGCGCACAGCTGAGCATCCAGATCTACGAGGCGCCACAGAGATCTGAG
 ACCGGCAAGTACGCCCGCGGACATCCGCCACACCAACAGACTGAGAGCTGACGAGGTTGGTTCAGAGATCCGCAACCGGAGGCGATCTGTATCTGGGG
 CAAAGTGTCCCACTTCGCTCTCCCATCCAGAGAGACTCTGGAGATGTGTGGCCGATCTGTGGCCGATCTGAGGAGCCGACCTGGCATCCCGGAGGAGTCTGTGA
 ACACCCCGCCCTGTGTGAGCTGTGTGACCTGAGAGACCGAGCCCATCTGTGGCGCTGAGACCTCTAGCTGGAGCGCGCCGACCTCCGAGCGAGACCTGGA
 CTGTGGGAGCGCGCTGAGTACCGACCGCGCGCCAGAGTGTGCCCTCAGCGAGACACCAAGAGAGACCGAGCTTCAGAGGCCGCGCGCCGACCTCCAGCTGGA
 CCGTCAGGACTCCGCTCCGAGTGAACATCTGACGACTCCAGTACGCGCTGGCATCATCAGGCGCCACCCCGCAAGTCCGAGTCCGAGCTGGTGTGA
 CCGAGATATCGAGCGCTGATCCGAGAGGCGGTGACTCTCTGGGTGGCGCCCGCCAGAGGCTATGGCGCGCATCGGCGCAAGAGGAGCTGGT
 TCCACCGCATCCGCAAGTGTCTGTGTGAGCGCATCGAGAGGCGCCAGGAGGACAGAGATACCTTCCACTGAGCGGCGCATGGCGCTCCGAGT
 CAACCTGCCCCCTGTGTGGCGCAAGAGATCTGTGGCTCTCTGGACAAAGTGTCCAGCTGAAGGCGGAGGCCATTCAGCGCCAGGTGACTGCTCCCGCGGCA
 TGTGCGAGCTGTGACTCACCACTGTGAGGCGCAAGATCTCTGTGGCTGAGTGTGGCTCCGCTTACATCGAGGCGAGGTATCCCGCGCGAGCT
 GCGCAGGAGACCGGCTACTTACTCTGTGAAGTGTGGCGGCTGTGCCCTGAAGATCATCACACGACAGACGGCTCCAACTTCACTCCACTGTGTGA
 GGGCGCGCTGTGTGTGGCGGCTATCCAGAGAGGTTTGGCATCTACACCCCACTCCAGGCGCTGTGTGACTCATGAACAGAGGCTGAAGAGA
 TATCGGCGAGGTGCGGACCAAGCGCGAGCACTTGAAGACCGCGGTGCATGGCCGTGTATCCCAACTTCAAGCGCAAGGGCGGCTACCGGCGCTAC
 TCGCGCGGAGGCGCATATCGACATCTCGCCACCGACATCCAGACAGGCGAGGCTACCAAGATCCAGAGATCTCGGCTGTACTTCCG
 CGGCGAGGCGAAGATCTCGGCATCTAGCGGCGAGGCTGGCGGCGAGCTTCTGTGGCGCGCGCGAGCAACACAGATCAAGGTGTGTGGCTGCGC
 TTTCTTCGGGAGAACCTTGGCCCTTCCAGCABGGCGAGGCCCGCAASHTTCTCTCTCGACGACAGACCCCGGCGCAACTCTCCCGGCTCCCGCGGAGACTGCGCGGTGGCC

도면116a

73. 2003 CON G pol. pep

FERENLAFQGEAREFSSQARANSPTRELVRREDSPLEAGAEKGGAISLSFPQITLWQRPVTVKIGGQLEALLDTGADDTVLEEIN
 LPCKWKPMIGGIGGFVKVRQYDQILLIELSGKKAIGTVLVGPTPINIIGRNMLTQIGCTLNFISPIETVPVKLPKPGMDGPKVKQWELTEEK
 IKALTEICTEMEGKISKIGPENPYNTPIFAIKKDDSTKWRKLVDFRELNRRTQDFEVQLGIPHPAGLKKKRSVTVLDVGDAYFSVPLDE
 NFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPIFOSSMTKILEPRTKNPEIIVYQYMDLYVGSDLLETGHRAKIEELREHLRWGF
 TTPDKKHQKEPPFLWMGYELHPDKWTVQIQLPKESWTVNDIQKLVGKLNWASQIYPGIKVKQKQCKLRGAKALTDIVPLTAEAELELAEN
 REILKEPVHGVIYDFPSKELIAEVQKGLDQWTYQIYQEPYKNLTKGYAKRGSHTNDVKQLETVVQKIATESIIVGKTPKFKLPTRKETW
 EVWTEYWOATWIPWEFVYNTPLVKLWYRLETERIPGAETIYVDGAANRETAKGAGVTDKGQKIITLTETTNOKAELQAIHLALQDSG
 SEVNI VTD SQYALGLIQADPDRSESELVNOIIEQLIKKVKVYLSWVPAHKGIGGNEQVQDKLVSSGIRKVLFLDGDIAQAEHEHRYHSNWRAM
 ASDFNLPPIVAKETIVASCDKQKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH
 TDNGSNFTSAAVKACWMANITOREFGIPYNPOSQSVESMNKELKILIGQVRQAEHLKTAVQMAVFTIHNFRKKGIGGYSAGERIIDIIAS
 DIQKELQKQITKIQNFRVYIRDSDRPIWKGPAKLILWKGECAVVIQDNNEIKVYPRRRAKIIIRDYKQWAGDDCVAGQDEDS

도면116b

2003_CON_G pol.OPT

[illegible]

도면117a

74. 2003 CON H pol.pap

FFRENLAFOQREARKFSPQARANSPTSRELVRRGDDPILSEAGAEGOGTSLSPPOITLWQRPVTVTKIEGOLREALDGTGADDTVLEINL
 PGKWKPMIGGIGGIFIKVRQYEQVAIEICGKKAIGTVLVGPTPPNIIIGNILTOIGCTLNFPSPIETVPVKLPGMDGPKVKQWPLTEEKI
 KALTEICIEMEKEGKISKICPENPYNTPIFAIKKDDSTWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKRSVSLDVGDAYFSVPLDKD
 FRKYTAETIPISINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPENIIYQMDLIVGSDLEIGOHRAKIEELRAHLIRWGFT
 TPDKKHQEPFFLWNGYELHPDKWTVQPVKLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKOLCKLLRGAKALTDIVPLTKAELELEAENR
 EILREPVGYYDPSKDLIAEIOKQPPDQWTYQIYQEPFNLTGKYAKRTAHTNDVKLTAEVQKIATSIWNGKIPKFRLP IQRETWE
 TWTEHWQATWIPENEFVNTPHLVKLNQLETEPIAGAETIYVDGAANRETKIGKAGYVTDRGKQKVSLTETTNQKTELQAIYLAIQDSGL
 EVNIVTDSQYALGIIQAQPKSESELVNOIEELIKKEVYLSWVFAHKGIGGNEQVDKLVSSGIRKVLFLQGLDKAQEEHERYHNNWRAMA
 SDFNLPPIVAKEIVASCDCQQLKGEAMHGQVDCSPGIWQDCTHLEKVTLLVAVHVASGYIEAEVIPAETGQETAYTILKLAGRWPVRMIHT
 DNGSNFTSAVKAACWADIQOETGIPYNEPQSQGVESNMKELKXIGQVRDQAEHLRTAVQMAVFIHNFKRKGIGGYSAGERIIDITATD
 IQTKRELQKQISKIQFRVYIRDSRDFIWKGPALKLWKGEAVVIQDNSEIKVVPFRKAKIIRDYGKQMGDDDCVAGRQDEDS

2003 CON H POL.OPT

[illegible]

도면118a

75. 2003 CON 01 AE pol. pep

FFRENIAFQOGKAGEFSSEOTRANSPSTSRKLGGRDNLITAGAEROGTSSTSFPOITLWORPLVTVKIGGOLKEALLDYGADDTVLEDI
 NLPKWKPNIGGIGGFIKVRQYDQILLIELCGKKAIGTVLVGPTPVNIIIGRNMLTQIGCTINFPISPIDTVPVTLKPGMDGPKYQWPLTEE
 KIKALTEIOCKEMEEEGKISKIGPENFYNTFVFAIKKDSKWKRLVDRELNKKTDQFWEVQLGIPHPAGLKKKKSVTVDVGDAYFSVPLD
 ESFRKYTARTIPSIINNETPGIRYOYNVLPQGWKSPAIFOSSWTKILEPFRJKNPEMVIYOYMDLYVGSDDLFIGQHRTKIEELRAHLISWG
 FTTPDKKHQREPPFLNMGYELHPDRTVQDIELPEKDSWTVNDIQKLVKLNWASQIYAGIKVKQOLCKLLRGAKALTDIVPLTEEALELAE
 NREILKTPVHGYYXDPSKDLVAEVQKQGDWTYQIYQEPFNKLTGKYARKRSANTINDVRLTEVVQKIATESISIVMGKTPKFRLPQRET
 WETWMEYWCATWLPWEFEVNTPTLVKLWYQLEKDPVIGAEFVYVDGAASNETKLGAGYVTDGROKVVSLTETTNOKTELHAIHLAQDS
 GSEVNI VDSQYALGIIQAOPDRSESEVNVNQTIELIKKBYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAOEHEERYHNSWRT
 MASDENLPPTVAKETIVANCCKOLKGEAMHGOVDCSPGIWOLDCTHLEGKVLVAVHVASGYIEAEVI PAETGGQETAYELLKLAGRWPPYKVI
 HTDNGSNFTSAAVKACWVANVRQFFGIPYNPOSGVVESEMNKELKKIIGQVRQAEHLKTAVOQMAVFTHNFKKGGGIGGYSAGERIIDI
 TDIQTKELQKQITTKIQNFRVYIYRDSRDPIMKGPALLWKGGGAVVIQDNSDIKVPVPRKAKIIRDYGKQMGAGDCVAGRQDEDS

도면119a

76. 2003 CON 02 AG pol.pap

FFRENLAFOQGEARKFSSEQGTNSPTSRFLWDGGRDNLISEAGTEGQGTISSFNFPQITLWQRPVTVVRIGGQIEALLDTGADDTVLEEI
 NLPGRKPKWIGGIGGFIKVRQYDQALLIEICGKKALGTVLVGPTEFVNIIIGRNMLTQIGCTLNFPIETVPVKLPKPGMDGPKVKQWPLTEE
 KIKALTDICTEMKEGKISKIGPENFINTPVFAIKKOSTKWRKLVDFRELNKRQDEWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD
 KDFRYTAFITPSVNNETPGIRYQYNVLPQWKGSPAIPOASMTKILEPRTKNPEIIVYQMDLYVGSDELIGQHRAKIEELREHLLRWG
 FTTDDKKHQREPPFLWNGYELHPDKWTVQPIQLPEKDSWTVNDIQLVGLNWSAQIYAGIKVKQOLCKLRGAKALTDIVTLTEEALELAE
 NREILKEPVGYYDPTKDIAEIQKGGQDWTYQIYQEPFNLTGKYAKMRSHTNDVKQLTEVVQKVATESIVWKTPKFPLPIQRET
 WEAWMEYWQATWIEPEWEFYNTPIVAKLWQLEKDPVGAETFYVDGAANRETCLGKAGYVTDGRQKVSLTETNQKTELHAIHLALQDS
 GSEVNIIVTDSQYALGIIQAQDRSESELVNOIIEKLEKDVYLSWVPAAHKGIGGNEQVDKLVSNKIRKVLFLDGDIDKAQEEHERYHSNRA
 MASDFNLPPIVAKEIVASCDCQLKGEAMHGQVDCSPGIWQDCTHLEGKIIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI
 HTDNGSNFTSAVKACWNAVTOREGI BYNPQSGVVEWMKELKKIIIGQVRDOAEHLKTAQVMAVFIHNFKRKGIGGYSAGERIIDIILIA
 SDIOTKELOKOITKTIONFRVYIRDSRDPIMWGPALKLWKGEAVVIQDNDSDIKVYPRRKAIIIRDYKQKQAGDDCVAGRQDEDS

도면119b

2003 CON 02 AG pol.OPT

TTTTCTTCGGCAGAACTTGGCCCTTCAGCAGGCGGAGCCCGCAGTTCTCTCGAGCAGACCGGCAACAATCCGCCACTCTCCGCAGCTGTGGCAGG
 CGCGCGCGCATCCGCGCTGTCCGAGCGCGACCCAGGCGACCATCTCTCTTCAACTTCCCCAGATCACTCTGTGGCAGCGCCCTCTGGTGA
 CCGTGGCGCATCCGCGCGCTGATCGAGGCTGCTGGACACGGCGCGACGACACACCTCTCTGGAGAGATCAACTCTCCGGCAAGTCCGACG
 ATGATCGCGCGCATTCGCGGCTTATCAGAGTTCGCGAGTACGACAGATCTGATCAGAGTTCGCGGACGAAGCGCATCGGACCGTGTGTGGCGCG
 CACCCCGGTGAACATCATCGSCGCAATGCTGACCATCGCTGACTTCCCATCTCCCATCGACCGCTGCGCTGAAGCTGAAGC
 CCGGCGATGAGAGGCCCCAAGAGTGAAGCTGGCCCTCGCTGGAGAGAAATCAAGGCGCTTACCGACTCTCCACAGATGAGAGGAGCGACATC
 TCACAGATCGCGCGCGGAGAACCTCTACAAACCCCGTGTGCCATTAAGAACGAGCATCCACCACTGGCGAAGCTGTGGATCTCCGCGAGTGA
 CAAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCATCCGCCACCCCGCGGCTGAAGAAGAAAGATTCGTGACCTGTCTGGACGTGGCGACGCGCT
 AGCTCTTCGTTCGCTCGCCTGGCAAGAGACTTCGCAAGTACACCGCTTCACCATCCCTCCGTGAACAACGAGACCCCGGCATCCGTTACAGTACAAGCTG
 CTGACCCAGGCTGGACAGTGTGCGCTCGCATCTCCAGCGCTCTCAACAGAGATCTCGAGGCTTCGCGACACCAAGAACCCCGAGATCTGATCTACCA
 CTGACATGAGCACTGTAGCTTGGCTCGCGCTGGAGATCGCGCGACGCGCAAGATCAGAGGCTTCGCGACCTGCTGCGCTGGGCTTCTCACCA
 CCCCAGCAAGAGCACAGAGAGAGCCCCCTTCTCTGGATGGGTACAGCTGCACCCAGCAAGTGGACCGTGTGGAGCCCATCAAGCTGCCGAGAG
 TCTCTGGACCTGTAGACCATCAGAAAGCTGTGGCAAGCTGMACTGGGCTTCCAGATCTAGCGCGCATCAAGTGTGAAGCAGCTGTGCAAGCTGCT
 GCGCGGCGCAAGCGCTTGACCATCTGTCAGCTTCAGAGAGCGCAGCTGAGCTGCGCGAGAACCGGAGATCTCTGAGAGGAGCCGTGCACTGGAG
 TGTGTACTTACCAACCCCAAGCACTGATCGCGAGCTGACAGAGCGGCGACAGCACTGACCATCAGATCTACAGAGAGCCCTTCAAGAACCTGAG
 ACGGCGAAGTACGCCAAGATGGCTTCGCGCACCAACCAAGCATGAAGCTGACCGAGGTGTGCAGAAAGTGGCCACCAGTCCATCTGTATCTGGGG
 CACAGACCCCGCAAGTTCGCGTGGCATCCAGCGCGAGCACTGGGAGGCTTGGATGGAGTACTTGGCAGGCCACCTGATCGTCCCGAGTGGAGTCTGTGA
 CTGGGCARGCGCGCTAGCTTACCGAGCGCGCGCCAGAGAGTGTGTCTCTGACGCGCGACCCACAGCAGCAGGCTGCAGCCCATCAACTGGC
 CCTCTGAGGACTCTCGGCTTCGAGTGAACATCTGTGACCGACTCCCACTACGCCCTTGGCATCAICGAGGCCAGCGCGTCCAGTCCGAGCTGGTGA
 ACCAGATCATCGAGAACTCATCGAAGAGCAAGTGTACTGTCTGGTGGCGGCCACAGGGCATCGCGCGACAGCGAGTGGACAGCTGTGTG
 TCACAGGCACTCCGAGAGTGTCTCTGGACGCACTGCAGACGCCCGAGGAGCAGCAGAGGCTTACCATCTCAACTCGCGCCATGSCCTCCGCTT
 CAACCTTGCCTCCCATCTCTGGCCCAAGGAGATCTGGCCCTCTTCGCAAAAGTTCGAGCTGAAGGGCGAGGCCATTCAGCGGCTAGTGGACTCTCCCGCGGCA
 TCTTCTGCGAGCTGGACTCACCCTCATCTGGAGGCAAGATCATCTGTGTGGCGTGCAGCTGGGCTCTCCGGTTCATCTAGGCGAGCTGATCCCGCGCGAGAC
 GGGCGAGGAGCGCTTACTTCACTGAAAGCTTGGCCGCGCTGGCCCTGAAAGGTGATCCACACCGACCAAGCGTTCACACTCACTCCCGCGCGGTGA
 TCTCATCGGCAGTGTGGCGACACAGCGCHAGCACTGAGACCGCGCTGCAGATGGCCGTGTTTCAPCCAAACTTCAAGCCCAAGCGCGGATTCGGCGGTAC
 CTCTCGCGCGGAGCGCATCTCGACATCATCGCTCCGATCATCAGACCAAGAGACTGCAGAGCAGATCACAGATCAGAGACTTTCGGGTGTACTACCG
 CGACTCCCGCGACCCATCTTGGAGAGGCCCGCCAGCTTCTGTGHAAGGCGAGGGCGCGGTGGTGTATCGAGCAAACTCCAGACTCATAGGTGGTGGCCCG
 CGCTCGCGCGCAGATCATCTCGGAGTACGGGCAAGAGATGGCCGGCGACCATCTGCTGGCGCGCGCGCGAGGAGACTTAA

도면120a

77. 2003 CON 03 AB pol. PEP
 FFERNLAFQOREARKEFSSEQTRAISPTSRKIMWGGGRDNPLPETGTERTQGTASSFNFPQITLWQREPLVTVRI GGQKKEALLDTGADDTVLEDI
 NLP GKWKPMIGGIGGFIKVRQYDQILIEICGKKAIGTVLAVGPTFVNIIIGRNMLTQLGCTLNFPISEPIETVPVTLKPGWDGPKVKQWPITEE
 KIKALTDICKEMEREGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHAGLKKKSVTVL.DVGDAYFSVPLD
 QDERKYTAFTTIPSTNNETPGIRYOYNVLPOGWKGSFAIFQSSMTKILEPPRKQNEEIVTYQYMDLIVGSDLEIGQHRTKIEELREHLLRWG
 FTTPKKHQKEPPFLWNGYELHPDKWTVQPIVLPEKDSWTVNDIOKLVGKLNWASQIYAGIKVRQLCKLIRGAKALTEVILPTAAEAELAE
 NREILKEPVHGVYIDPSKDLVAEIQOGQONTYQIYQEPFNKLTGKYARLRGAHTNDVKQLTEAVQKIATESIVIMGKTPEFKLPIQKET
 WETWTEYMOATWIPMEFEFVNTPLVKLWYQLEKEPIVGAETFYVDGAANRETKSGKAGVITDRGRQKVSVLTDITNQKTELOALHLALQDS
 GLEVINVTDSQVALGILQAPDKSESELVSQIIEQLIKKVKYLAWVPAHKGIGNEQVDKIVSAGIRKVFELDGIDKKAQEAHEKHYHSNWRA
 MASDFNLPPVVAKEIVASCDCQLGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYILEBEVIPAETGQETAYFVLKLAGRWPVKII
 HTDNGSNFISTAVKACWAGIKQERGIYPNPQSGWESMNKQKQIIGQVRDQAEHLKTAVQMAVFTHNEKRKGGIGGYSAGERIIDLIA
 TDIOITRELOKQILIKIQNFRVYRDSRDPIMWGPALKLWKGGGAVI.QDNNDIKVYPRRKAKIIRDYGMQWAGDDCVASRQDED\$

2003 CON 03 AB POL.OPT

[illegible]

도면121a

78. 2003 CON 04 CPX pol. pep

FERENVAFOQREARKFSSEQARANSFARRRELROERGDNLLSEAGTEGQGTISFNFPQITILWQRELVTIKIGGQIREALLDTGADDTVLEEIN
 LFGKWKPMIGGIGGFIVKQYDQIPFIEICGKKAIGTVLVGPTFYNIIIGNMLTQLGCTINFPISPIETVPVKLPKPGMDGPKVKQWPLTEEK
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIRKKNSTRWKLVDFRINKRQTQDEWVOLGIPHPAGLKKKSVTVLDVGDAYFSVPLDP
 EERKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSFALFQCSMTKILEPPTKKNPEIIVYQYMDDLVYQSDLEICQHRAKIEELREHLLRWGF
 STPDKKHQKEPPFLMWGYELHPDKWTVQPIQLAEKDSWTVNDIQKLVGKLNWASQIYPGKVKQOLCKLLRGAKALTDIVPLTTAELELAEN
 REILKEPVHGAYYDPSKDLIAEIOKQOGOWTYOYQEPYKNLKTGYAKTRSAHTNDVROLTAZAVOKIAMECTIVMGKTPKFELPIQKETW
 DTWTEYWCATWIPWEFVNTFPLVLKLYOLETDPIAGAEFFYVDAASRETQKQAGYVTDGRQKVVSLSSETNQKTELQAIYLAQDSG
 SEVNLVDSQYAGIIGIQAQEDRSDDLNVQIIIEQLIQKDKVYLSWVPARKGIGGNEQVDKLVNSGIRKIVFLDGDIDKAQEEHEKYHNNWRAM
 ASDENLPPVVAKEIVASCNKCOLKGEAMHGQVDCSPGIWOLDCTHLEGKIILVAVHVASGYIEAEVIPAETGOETAYFTILKLAGRWPKLIH
 TDNGENFTSAVKAACWWADIQEEGIPYNPQSQGVVESNNKELKKIIGQVRDQAEHLKTAVMQMAVFIHNFKRGGGIGYSAGERIIDIAS
 DIQTKELOQKITIKIQNFRVYIRDSRDPIMKGPAILMKGEGAVVTQDNSDIKVVPRRKAKIIRDYGKQWAGDDCVAGRQDEDS

도면121b

[illegible]

도면122a

79. 2003 CON 06 CPX pol .PEP

FFRENLAQGEAREESESSEQARANSPTREELRVBRGDSPLPEAGAGQGAISLSPQITLWQRELVTVRIGGQLEALLDTGADDTVLEDIN
 LPGKWKPMIGGIGGFIKVRQYDQILLIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTINFFLSPFETVPVKLPGMGDPKVKQWPLTEEK
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIRKKKSTKWRKLVDFRINKRRTQDFWEVOLGPHPHAGLKKKSVTVLVDGDVAFSVFLDE
 DERKYTAFTIPSINNETPGIRYOYNVLPQGWKSPAIPOSSMIKLEPFRINKNEIIVIQYMDLIVGSDLEITGQHRAKIEELREHLLKMGF
 TTPDKKHQKEPFFLMWGYELHPDKWTVQPIQLPDKDSWTVNDIQKLVGKINWASQIYPGKVKQCLCKLRGAKALTDIVPLTAAELEIAEN
 REILKEPVHGVIYDESKDLIAEIQKQGGQWTVQIYQEPHKNLKTGYARIKSAHTNDVKOLTEAVOKIALESTIVWGKTPKFRPLPIOKETW
 ETWWTEYWCATWIPWEFEVNTPELVKLWYQLETEPIVGRRETFFYVDGAANRETKKKGAGYVTDGRQKVVSLETETNQKTELQAINLALQDSG
 SEVNLVDSQYALGIIQAFDKSESELVNQIIEQLIKREKVYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGDIDKAQEDHERYHSNWRAM
 ASDENLPPIVAKETVASCCKQLKGEAMHGQVDCSPGIWQLDCTHLEKIIIVAVHVASGYIEAEVIPAETGQETAYFTILKLAGRWPVKVIH
 TDGCSNFTSAAVKAACWNANITQETGIPYNPQSGQVWESMNKELKTIIGQVRDQAEHLKTAVQMAVFIHFKRKGIGGYSAGERIIDIILAS
 DIQTKELQKQITKIQNFRVYRDSRDPIMKGPALKLLWKGEGAVVIQDNSEIKVYVRRRAKIIIRDYCKQKQAGDDCVAGRQDEDS

2003 CON 06 CPX pol.OPT

[illegible]

도면123a

80. 2003 CON 08 BC pol pep
 FERILAFQGEAREFFPEQTRANSPTSRELOVRGDNPSSEAGTEROGTINFPOITLWQRPVLSIKVGGQIKREALLDTGADDTVLEEVNLP
 KWKPKMIGGIGGFTKVRQYEQIPIEICGKKAIGTVLVGFTPVNLI GRNMLTQLGCTINEPISLETVEVKLKFCMDGPKVKQWPLTEBKKA
 LTAICDEMEKEGKTKIGEDNPNYNTPIFAIRKKDSKWKRLVDFRELNRKTQDFEWQLGIPHPAGLKKKSVTVLDVGDAYSVPDLKDFR
 KYTAFITPSVNNETPGIRYQYNNVLPQGWKGSPIATQCSMTKILEPFRKQNPDIYIYOYMDLVVGSdleIGQHRTKTEELREHLLKWGFTTP
 DKKHQKEPFLMWGYELHDPKWTVPQIQLPEKDSWTVNDIQKLVGKLNWASQIYPGKVRQKLLRGAKALTDIVPTEEALELEAENREI
 LKEPVHGAYYDPSKELIARIQOGQDOWTYOIYQEPFKNLKTGYAKMRTAHTNDVKOLTEAVQKIAMESIVINGKIPKFRLPQKETWETW
 WTDYQATWIPENEFVNTPEPLVKLWYQLEKDPFAGVETPYVDGAANRETKIGRAGYVTDGRKKIVSLTDTTNQKTELQAIYIALQDSGSEY
 NIVTDSQYALGIIQAQPPKSESELVNOILEQLIRKERVYISWPAHKGIGGNEQVDKLVNSGIRKVLFDGIDKAQEEHEKHVSNWRAMASD
 ENLPPIVAKELIVASCDQOLKGEAMHGOVDCSPGIWOLDCTHLEGKIILVAVHVASGVTEAEVLPATGQETAVFILKLAGRPVKVLIHTDN
 GSNETSAAVKAAACWWAGIQCFGIPYNPQSGVVESEMNKELKLIQVDRDQAEHLKTAVOMAVFIHNNKRRKSGSIGGYSAGERIVDIIATDIQ
 TRELQKQILIKIQNFRVYVRDSRDPWKGPAKLWKEGAVVIQDNSDIKVVERKAKTIKDYGKQNAQADCVAGRQDEDS

2003_con_08_BC_pol.1_of2

TTTCTTCGCGAGATCTCTGGCCTTCCCCAGGCGGAGGTCCTCCCGCGAGATTCCTCCCGCAACTCCCGCACTCTCCCGGAGCTGCAGGTGGC
CGCGGACACCCCTCTCCGAGGCGGACCGAGCGCCGAGACCTGTAACTTCCCGCAGATACCTGTGGCAGCGCCCTTGTGTCCATCAAGGTGG
GGCGGAGATTCGAGAGGCGCCTGTGTGACACCGGCGCGAGCACCGTGTGAGGAGGTGAACCTGCCCGGCAAGTGGAGAGCATGATGCGCGG
ATCGCGCGCTTCATCAAGGTGCGCAGTACGAGCAGATCCCATCGAGATCTCGGCAAGAGGCCATCTGGCAGTCTGTGGCGCCACCCGCTGGAA
CATATCGCGCGGACATGTGTACCCAGCTGGGCTGTACCTGAACTTCCCATCTCCCACTCGAGACCTGCCGTGAAAGCTGAAAGCTCGGATGAGG
GCCCAAGGTGAGCAGTGGCCCTGTACCGGAGGAGAACTCTCGCATGTGGCGCAAGTCTCGCGAGCTTCCGCGAGCTGAAACAAGAGTACCGG
CGCAACCGCTTACACACCCCATCTTCGCAATCCGGAAGAGACTCTCCAGTGGCGCAAGTCTCGAGCTGGCGGAGCGCTACTTCTCTCGGTG
GCACTTCTGGAGGTGCAGCTGGGCATCCCGATCCCGAAGAGACTCTCCATGTGGCGCAAGTCTCGAGCTGGCGGAGCGCTACTTCTCTCGGTG
CTCGTGAACAAGACTTCCGCAAGTACCGCTTACCATCCCTCCGTGAACAACAGAGATCCGTGTGAGCTACCAAGCTGTGCTGCCCGCGAGG
TGGAAGGCTCCCCGGCATCTCCAGTACCGCTTACCATCCCTCCGTGAACAACAGAGATCCGTGTGAGCTACCAAGCTGTGCTGCCCGCGAGG
CCTGTACTGTGGCTTCGACGCTGGAGATCGGCGAGACCGCAACAATCGAGAGCTGTGTAAGTGGGGCTTCACCAACCGCCGCAAGA
AGCAACAGAGAGGCCCTCTGTGTGATGGGCTACAGCTGCACCCGACAGTCCCGGCACTGAGCTGCGGAGCTGCGCGAGCATCTGTAAGTCTG
GTGACAGATCCGAAGTGGTGGCGAAGCTGAATGGGCTCCAGACTTACCCCGGCACTGAGTGGCGGCGTGTGAGGAGCGCTGTGAGGAGCGCT
GGCTTGAACGATGTCGTCGACAGGAGCGAGCTGAGAGCTGACCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
GCCCTTGAACGATGTCGTCGACAGGAGCGAGCTGAGAGCTGACCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
GCTTCAAGGAGCTGATCCGCGAGATTCGAAGACGGCGAGGACAGTGGACCTACCAATCTACAGAGCGCTTCAAGAGCTGAGAGCTGAGAGCT
GCACAAGTGGCAACCGGCGAGTGAAGAGTGGTGGCGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT
CTTCCGCTGCCCATCAAGAGGAGCTGGGAGA CTTGGTGGCGGCTTACGTGGAAGGAGCGCTGTGAGGAGCGCTGTGAGGAGCGCTGTGAGGAGCGCT
TGTGAGCTGTGTGATCCAGTGGAGAGCGGCTGCGCGGCTGGAGACTTTTACGTGGAAGGAGCGGCTGTGAGGAGCGCTGTGAGGAGCGCT
GGTATGTHGACCGCGCGCGGCGAGAGATCGTGTCTGACGACCAACCAAGGAGCGGCTGTGAGGAGCGGCTGTGAGGAGCGCTGTGAGGAGCGCT
CGGCTCCGAGGTGACATGTGACCGACTCCAGTACGCTGGGCTGCGCGGCTGGAGACTTTTACGTGGAAGGAGCGGCTGTGAGGAGCGCTGTGAGGAGCGCT
AGCATGTGATCAAGAGAGCGGCTGCTTCTGCTGGTCCCGCGGCAAGGAGCGGCTGTGAGGAGCGGCTGTGAGGAGCGGCTGTGAGGAGCGCT
CGAAGGTGTGTTCTTGAAGCGGCTGACGAGCGCGGAGCGGCTGATGACCGGCTGGAGCTTCTCCCGGCTGTGAGGAGCGGCTGTGAGGAGCGCT
CATCGTGGCGAAGAGATCGTGTCTTCGACCATGCGAGCTGAGGCGGAGCGGCTGATGACCGGCTGGAGCTTCTCCCGGCTGTGAGGAGCGCT
ACTGCAACCGCTGAGGSCAGATCATCTGTGTGGCGCTGACGCTGGGCTGCGGCTGGAGCTTCTCCCGGCTGTGAGGAGCGGCTGTGAGGAGCGCT
GCCATCTTCACTGAGATGGCGCGGCTGGCGCTGAGGCTGATGCAACCGGCTGAGGCTGATGCAACCGGCTGAGGCTGATGCAACCGGCTGAGGCT
TGGCGGACCGGCGGACCATGAAGACCGCGGTGAGATGGCGCTGTATCCCAACTTCAAGCGAGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCT
GCGATGTGTGACATCATGCGCAACCGCTGAGATGGCGCTGTATCCCAACTTCAAGCGAGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCT
CCCATCTGGAGAGGCGCGGACAGTGTGTGAGAGGCGGCGGCTGGTGTGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT
AGATCTCAAGAGCTACGCGCAAGCTATGCGCGGCGGCTGAGTGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT

도면124a

81. 2003 CON 10 CD pol. PEP
 FERENLAEQQRKARELPSEOTRANSPTSRELVRWGGNTLSETGAERQGVLSLSPQITLWQRPVTVKIGGOLKEALLDTGADDTVLEEMN
 LFGKWKPMIGGIGGFIKVRQYDQILIEICGYKAIGTVLVGPTPNYNIIGRNLITQICCNLFPISPIETVPVKLKEGMDGPKVKQWPLTEEK
 IKALTEICTEMEKEGKISRIGPENPYNTPIFAIKKDDSTKWKLVDFRELNKRTQDFWEVLGIGIPHPAGLKKKSVTVLDVGDYFVSVELYE
 DFRKYTAFTLPSINNETPGIRYQYNVLPQGWKGSIPAIFQSSMTKILEPFRQNPEWIIYOYMDOLYVGSDEIGOHRIKIEELRGHLLKMGF
 TTPDKKHQKEPFFLMWGYELHPDKWTVQPIQLPEKDSWTVNDIOKLVGKLNWASQIYPGIKVRQOLCKLLRGAKALTDIVPLTEEALELAEN
 REILKEPVHGVYDPSKDLIAEIQKQGDQWTYQIYOEPHNKLTGKYAKERTKLGAGVYTDGROKVISITDTNOKTELOAINLALQDSG
 ETWTDYQATWIPEWEFVNTPPVLKLYQLEKEPIVGAETFYVDGAANRETKLGAGVYTDGROKVISITDTNOKTELOAINLALQDSG
 SEVNIIVTDSQYALGIIQAQPDKSESELVNTIEQLIKKEKVYLSWVPAHKGIGGNEQVVKLVSSGIRKVLFDGLDKAQEEHEKYHNNWRAM
 ASDFNLPVVAKEIVASCQKQKGEALHGOVDCSPGIWOLDCTHLEGVLLVAVHVASGYIEAEVIPAETGQETAYFLKLKLAGRWPVKVWH
 TDNGSNETSAVKAACWAGIKQEFGIPYNEFSQGVVESNMKELKKIIGQVRDQAEHLTAQVQMAVFTHNFKRKGIGGYSAGERIIDIIAT
 DIQTRELQKQIITIKIQNFRVYTRDSROPIWKGPAKLLWKGEAVVIQDNSDIKVPVRKVKIIKDYKQKMGADCVASRQDEQ

도면124b

2003 CON 10 CD pol.OPT

[illegible]

도면125a

82. 2003 CON 11 CPX pol. PEP

FFRENLAFOQGEAREESPEQARANSPTSREILVRGGDSPLPETGAEGEGAISFNFPQITLWQRPVLTIKVAGOLKEALLDTGADDTVLEEID
 LFGWRKPMIGGIGGFIVKRYEELIIEIEGKKAIGTVLVGPTFVNIIGRNMLQIGCINLFPISPIDTVPVKLPGMDGPKVKQWPLTEEK
 IKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDDSTXWRKLVDFRELNKRTQDFWEVQLGIEHPAGLKKKSVTVLGVGDYFSPVPLDE
 SFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGSZALFQSSMTKILEPFTQNPETVVIYQYMDLVGSDLEIGOHREKVEELRKHLKWF
 TTPDKKHQKEPFFLWNGYELHPDKWTVQPIQLPDKECWTVNDIQKLVGKLWASQIYPGKVKQKCKLLRGTKALTDIVPLTAAEAELELAEN
 REILKEPVHGYYDPSKDLIAEVQKQGLDQWTYQIYQEPKRLTKYAKRRTAHTNDVRQLAEVVQKISMESIVIWGKIPKRFPIQRETW
 ETWTDYQATWIPEWEFVNTPPLVKLWYQLEKEPIGAEFFYVDGAANRETGLKAGVYTDKGRQKVTVLTETNOKTELEATHALQDSG
 LEVNIIVTDSQYALGIIQAQPKSESELVSQIEQLIKKEKRVLSWVPAHKGIGGNEQVDKLVSSGIRKVFLLDGDIDKAOEEHERYHSNWRAM
 ASDFNLPPIVAKELIVASCDKQLKGEAMHGOVDCSFGINQIDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFIKLKAGRWPVKVIH
 TDNGSNETSAAVKAACWMANIQOEFQIPYNPQSQGVVESNNKELKKIIGOVREQAHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIAT
 DLQTRQLQKQITTKIQNFRVYTRDSRDPFWKGPALKLWKGEGAVVTIQDNSDIKVPFRKAKIIRDYGMAGDDCVAGRODED\$

도면126a

83. 2003 CON 12 BF pol. PEP

FFRENLAFOQGEARKFPSPQARANSPPASRELWVRGDNPLSEAGAERRGTVPSLSFPQITLWQREPLVTIKVGQKKEALLDTGADDTVLEDI
 NLPCKWKPMIGGIGGFIVKVKQYDNILIELCGHKAISTVLVGPFPVNIIGRNLLTQLGCTLNEFISPIETVPVKIKPGMDGPKVKQWPLTEE
 KIKALTEICTENEREKGIKSKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNNKRTIQDFEWQLGIPHAGLKKKSVTVLDVGDAYFSVPLD
 KDFRKYTAFTIPSYNNETPGIRYOYNVLPQGWKGSFAIFOSSMTKILEPFRKQNPDIIVYQMDLIVGSDLEIGQHRKIEELRQHLLRWG
 ETTDDKKHOKEPPELWMMGYELHPDKWTVQPIVLPBKDSWTVNDIQKLVGKLINWASQIYPSIKVKQLCRLLRGTKALTTEVPLTKAELELAE
 NREILKEFPVHGYYDPDSKDLIAEIQKQGGQWTYQIYOEPFNKIKIGKYARMRGHTNDVKQLEAVQKITTESIIVWGKTPKFRLPIKET
 WDTWTEYWOATWLPWEWTFVNTPLVLKLYQLETEPIAGAEFTFYDGSANRETKKKGAGYVTDGRQKAVSLTETTNQKAEHLHAIQALQDS
 GSEVNIVTDQYALGLIQAPDKSESEIVNQIIEQLIKKXVYLSWVPAHKGIGGNEQVDKLVASAGIRKILELDGIDKAEHEHEKYHNNRA
 MASDFNLPPVVAKEIVASCDKQKGEAMHGQVDCSPGIWQDCTHLEGIILVAVHVASGYLAEVIEAETGQETAYFILKLGRWPVKTI
 HTDNGENFSSAAVKAACWAGIQOEFGI PYNPQSQGVFESMKNELKKIIRQVROQAEHLKTAVMAVEFHNFRKGGGICGYASGERIIDIS
 TDIQTRELQKQIKIQNFRVYRDSRDPVWKGPAKLLWKGECAVVIQDNSEIKVVPERRAKIIRDYGMAGDDCQVAGRQDED\$

도면127a

84. 2003 CON 14 BG pol. pep
 FFRENIAFQQGEAREFSEQARANSPTREELWVRGDSPLPEARAEGKGDIPLSLPQITLWQRPVTVRIGGQLEALLDTCAGDVTVEIDIN
 LPGKWKPMIGGIGGFIKVRQYDQILTEICGKAIGTVILVGPPIINILGRNMLTQIGCTLNPEISPIETVVKLKPGMDGPKVKQWFLTEEK
 IKALTDICTEMEREGKISKIGPENPYNTPIFAIKKKDDSTKWKLVDFELNRRTQDFEYVQLGIPHPGSLKKKSVTVLDVGDAYTSVPLDE
 SFRKYTAFTIPSTNNETGIRYQYNVLPQGWKGPATQSSMTKILEPFRIKNPEIVYQYMDLYVGSDDLEIGQHRAKIEELRKHLISWGF
 TTPDKKHQKEPPFELWGYELHPDKWTVOPIQLPDKESWTVDIOKLVGKLNWASQIYPGKVKQOLCKLLRGAKALTIDIVPTAAELELAEN
 REILKEPVHGYYEPPSKELIAEVOKQGLDQWTYQIYORPYKNLTGKYAKRGSHTNDVKQLEFVVQKIATESIVTWGKTPFKLPKPKETW
 EVWTEYQWATWLPDWEFVNTPLVKKWYRLETEPIAGAETYVVDGAANRETILGKACVYTDGKOKIITLTETTNOKAELQAIHIALQDSG
 SEVNIIVTDSQYALGIIQAOQDRSESEVNVQIIIEQLIKKEKVYLSVVPARKHGIGNEQVDKLVSSGICKVLFLDGDIDKAQEEHEKHYHSNWRAM
 ASDENLPPYVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEKILLVAVHVASGYIEAEVPAETGQETAFILKLAGRPVKLIH
 TDNGSNFTSAAVKACWVANITQEFGIPYNPQSQGVESMNKELKKIIGQYRQDAEHLKTAVQMAVEIHNFRKGGIGGYSAGERIIDIIAS
 DIQTKELQKQITIKTQNFVRVYFRDSRDPWKGPAKLLWRGEGAVVIQDNNEIKVVPRRKAKIINDYKQKQAGDDCVAGRQDED\$

도면127b

2003 CON 14 BG pol.OPT

TTCTTTCGGCAGAACCTTGGCCCTTCCAGCAGGGGAGGCCCGCGAGTTCTCCCGCAGCAGGCCCGCGCCAACTCCGCCAACCCGCGCGAGCTGTGGGTGCG
 CCGCGCGGCACTCCCCCTGGCCGAGCGCGCGAGGCGCAATCCCTCTGCTGCCAGATCACCCTTGTGGAGCGCGCCCTGCTGGTGGACCG
 TGGGCATCGCGCGGCTTATCGAGGCGCTGCTGGACATCCAGCCGCGCGAGCTGTGCGGAGCACTAACTTGC CGGCGAAGTGGAAACCCCAAGCAG
 ATCGCGGCGCATCGCGGCTTATCAAGGTGGCGACTAGCAGCAGATCTCTGGGCAACAAGAGCACTGCGACCGTGCTGCTGGGCCCCGAC
 CCCCATTACATATCAGCCCGGCAAAATGTCAGCCAGATGGGTGCACCTGAACTTCCCATCTCCCCATCGAGACCTGCCGCTGAAGTCGAGCCG
 GGATGAGCGGCCCAAGGTGAGCAGTGGCCCTGACGAGGAGAAGATCAAGGCCCTGACCGACATCTGCACCGAGATGGAGCGGCGCAAGATCTCC
 AAGATCGGCCCGCGAGACCTCTACACACCCCATCTCTCCCATCAGAGAGAGGACTCTACAAATGTCGCAAGTGTGTGAACTTCGCGAGGTGGACAA
 GCGCACCCACGAGACTCTGGAGTGGCATCCCCACCCCTCGCGCTTGAAAGAGAAGTCTGCTGACGTGGGCGACGCTTACT
 TTCTCGTCCCTTTCGCGAGTCTTCGCGAGTATACCGGCTTACATCCCTCCACCAAGAGACGCCCGGCATCCGTACCAAGTACACGCTCTG
 CCCCAGGCTTGGAGGGCTCCCGCCCATCTCCAGTCTCATGACCAAGATCTGAGGCCCTTCGCAATCAAGAACCCGAGATGCTGATCTACCAAT
 CATGACGACACTCTAGCTGGGCTCCGACTGAGATGGCCAGACCGGCGAGATGAGAGGCTGGCAGCACTGCTGTCTGGGCTTACACACCC
 CCGGAGAGGACCAAGAACCTTCTCTGTGGATGGCTACGACTGCACTCCCGACAGATGAGAGCTGGCAGCACTGCTGTCTGGGCTTACACACCC
 TCTTGACCTGTGACGACATCCAGAGGCTGTGGCGAGTGAATCGGCTCCAGATCTACCCCGCATCAAGGTGAAGCACTGCTGAGGTGCTGCG
 CGGCGCCAGGCCCTGACCGACATGCTGCCCTGACCCGAGGCGCAGCTGGAGCTGGCCAGAAACCCGAGATCTTGAAGAGCCCTGTGCACGGCGCTGT
 AGCTACAGGCTCTCCAGAGGCTCTCCCGAGGTGCAGAGAGGCGCTGACAGCTGGAATCTACCATCTACAGAGAGCCCTACAAAGAACCTTCAAGAC
 GGGCAAGTACGCCCAAGCGGGCTCGCCACACAGCACTGTAAGTGAAGCTGACCGAGTGTGAGAGATGCCACCGATCTGATCTTGGGCGCA
 GGGACCCCAAGTTCAGAGTGGCCATCGCGAAGAGACATCACTCCCTGACGAGCACCAACAGAGCCGAGCTGCAAGGCCATCCCATATCGGCCCT
 CGAGACTCGGCTCCGAGTGAACATGTAGCGACTCCAGTACGCCCTGGGCACTATCAAGGCCAGAGCCGAGCTGCAAGGCCATCCCATATCGGCCCT
 AGATCTCGAGCTCCAGTCAAGAGAGAGAGTGTACTCTGTCTGAGTCCGCCACCAACAGGCGATCGGCGCAACGCGATCCGAGAGTGGTGAAC
 TCGGCATTCGAGAGTGTCTTCTTGACGGCATGCAAGGCCOAGAGAGAGCAAGAGAATACCACTCAACTGGCGCGCCATGGCCCTTCGACTTCAA
 CCGTCCCGCTGTGGCGCAGAGAGATGCTGGCTCTCTCGACAAATGCCAGTGAAGGGCGAGGCCATGACGGCATGAGTGACCTCTCCCGCGCATCT
 GCGACCTGCACTGCACCACTGAGGGCAAGCATATCTGTGGCGCTGAGCATCTGGGCTCCGGCTATCATCGAGGCGAGGTGATCCCGCGCGAGCGGC
 CAGAGACCTCGACTTCTATCTGAAGTCCGCGCTGTGGCTGGAAGATCTCACACAGCGACATCGGCTCCAACTCTCACTCGGCGCGCTGAAGCG
 CGCCTGCTGTGGGCGCAACATCACCCAGAGTTTCGCATCCCTTACAACCCCGAGTCCAGGGCTGTGGAGTCTATCAACATTCAGCGCCAGGGCGGATCGCGGCTACTCC
 TCTGGCCAGGTGGCGGACAGCGCGAGCACTCAAGAGCCCGTGCAGATGGCGTACTGATCCAGAGAGAGATCTCAAGCAAGATCA
 GCTCCCGCGAGCGCATATCGACATCATCGCTCCGACATCCAGACATCCAGAGATGCTGATCCAGGACACACAGCATCAAGTGTGGTGGCT
 TCTCCGCGGCGACCCCATCTGGAGAGGCCCGCGCGAGTGTGTGAAAGGGCGGCGGCTGCTGATCCAGGACACACAGCATCAAGTGTGGTGGCT
 CAAGGCCAAGATCATCCGGACACTACGGCAGACAGATGCGCGGCGACGATGCGTGGTGGCGGCGGCGAGCGAGCGAGGAGCACTAA

서 열 목 록

<110>	DUKE UNIVERSITY
<120>	CONSENSUS/ANCESTRAL IMMUNOGENS
<130>	IP-2012-0060
<150>	PCT/US2004/03097
<151>	2004-09-17
<150>	60/604,722
<151>	2004-08-27
<150>	60/503,460
<151>	2003-09-17
<160>	321

<170> PatentIn Ver. 3.3

<210> 1

<211> 866

<212> PRT

<213> Human immunodeficiency virus

<400> 1

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Met Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asn Val Arg Asn Val Ser Ser Asn Gly Thr Glu Thr Asp

130 135 140

Asn Glu Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg

145 150 155 160

Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val

165 170 175

Pro Ile Asp Asp Lys Asn Ser Ser Glu Ile Ser Gly Lys Asn Ser Ser

180 185 190

Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala

195 200 205

Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro
 210 215 220
 Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly Thr
 225 230 235 240

 Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys
 245 250 255
 Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
 260 265 270
 Glu Ile Ile Ile Arg Ser Glu Asn Ile Thr Asn Asn Ala Lys Thr Ile
 275 280 285
 Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn
 290 295 300
 Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Gln Ala Phe Tyr

 305 310 315 320
 Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile
 325 330 335
 Ser Arg Thr Lys Trp Asn Lys Thr Leu Gln Gln Val Ala Lys Lys Leu
 340 345 350
 Arg Glu His Phe Asn Asn Lys Thr Ile Ile Phe Lys Pro Ser Ser Gly
 355 360 365
 Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe
 370 375 380

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

450 455 460
 Asn Lys Asn Lys Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
 465 470 475 480
 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu
 485 490 495
 Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Glu Arg
 500 505 510
 Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly
 515 520 525

 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 530 535 540
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
 545 550 555 560
 Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly
 565 570 575
 Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys
 580 585 590
 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys

 595 600 605
 Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Asp
 610 615 620
 Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Ser
 625 630 635 640
 Asn Tyr Thr Asp Ile Ile Tyr Arg Leu Ile Glu Glu Ser Gln Asn Gln
 645 650 655
 Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp Lys Ala Ser Leu
 660 665 670

 Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe
 675 680 685
 Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val
 690 695 700
 Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe

705 710 715 720
 Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile
 725 730 735
 Glu Glu Glu Gly Gly Glu Gln Gly Arg Asp Arg Ser Ile Arg Leu Val

 740 745 750
 Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys Leu
 755 760 765
 Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg Thr
 770 775 780
 Val Glu Leu Leu Gly Arg Arg Ser Leu Arg Gly Leu Gln Lys Gly Trp
 785 790 795 800
 Glu Ala Leu Lys Tyr Leu Gly Asn Leu Leu Gln Tyr Trp Gly Gln Glu
 805 810 815

Leu Lys Asn Ser Ala Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val
 820 825 830
 Ala Glu Gly Thr Asp Arg Val Ile Glu Ile Val Gln Arg Ala Cys Arg
 835 840 845
 Ala Ile Leu Asn Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala
 850 855 860

Leu Leu

865

<210> 2

<211> 2607

<212> DNA

<213> Human immunodeficiency virus

<400> 2

gccaccatgc gcgtgatggg catccagcgc aactgccagc acctgtggcg ctggggcacc 60

atgatcctgg gcatgctgat gatctgctcc gccgccgaga acctgtgggt gaccgtgtac 120

tacggcgtgc ccgtgtggaa ggaggccaac accaccctgt tctgcgcctc cgacccaag 180

gcctacgaca ccgaggtgca caacgtgtgg gccacccacg cctgcgtgcc caccgacccc 240

aacccccagg agatcgtgct ggagaacgtg accgagaact tcaacatgtg gaagaacaac 300

atggtggagc agatgcacga ggacatcatc tcctgtggg accagtcct gaagccctgc 360

gtgaagctga cccccctgtg cgtgaccttg aactgcacca acgtgcgcaa cgtgtcctcc	420
aacggcaccg agaccgacaa cgaggagatc aagaactgct ccttcaacat caccaccgag	480
ctgcgcgaca agaagcagaa ggtgtacgcc ctgttctacc gcctggacgt ggtgcccac	540
gacgacaaga actcctccga gatctccggc aagaactcct ccgagtacta ccgcctgac	600
aactgcaaca cctccgcat caccaggcc tgcccaagg tgccttcga gcccatcccc	660
atccactact gcgccccgc cggtctcgcc atcctgaagt gcaacgacaa gaagttcaac	720
ggcaccggcc cctgcaagaa cgtgtccacc gtgcagtga cccacggcat caagcccg	780
gtgtccacc agctgctgt gaacggctcc ctggccgagg aggagatcat catccgctcc	840
gagaacatca ccaacaacgc caagaccatc atcgtgcagc tgaacgagtc cgtggagatc	900
aactgcacc gccccaacaa caacaccgc aagtcaccc acatcgccc cggccaggcc	960
ttctacgcca cggcgagat catcgccgac atccgcccagg cccactgcaa catctccgc	1020
accaagtga acaagacct gcagcagggt gcccaagaagc tgcgcgagca cttcaacaac	1080
aagaccatca tcttcaagcc ctctccggc ggcgacctgg agatcaccac ccactcctc	1140
aactgcggcg gcgagttctt ctactgcaac acctccggc tgttcaact cacctggatg	1200
ttcaacggca cctacatgtt caacggcacc aaggacaact ccgagaccat caccctgcc	1260
tgccgatca agcagatcat caacatgtgg cagggcgtgg gccaggccat gtacgcccc	1320
cccatcgagg gcaagatcac ctgcaagtcc aacatcacg gcctgctgt gacccgcgac	1380
ggcggcaaca actccaacaa gaacaagacc gagaccttc gcccggcg cggcgacatg	1440
cgcgacaact ggcgtccga gctgtacaag tacaaggtgg tgaagatga gccctgggc	1500
gtggccccc ccaaggcaa gcgcccggt gtggagcgc agaagcgcgc cgtgggcatc	1560
ggcgccgtgt tctgggctt cctgggcgc gccggctcca ccatgggcgc cgcctccatc	1620
acctgaccg tgcaggcccg ccagctgtg tccggcatcg tgcagcagca gtccaacctg	1680
ctgcgcgcca tcgaggccca gcagcacctg ctgcagctga ccgtgtggg catcaagcag	1740
ctgcaggccc gcgtgctggc cgtggagcgc tacctgaagg accagcagct gctgggcatc	1800
tgggctgtc ccgcaagct gatctgcacc accaacgtgc cctggaactc ctctgtgtcc	1860
aacaagtccc aggacgagat ctgggacaac atgacctgga tggagtggga gcgcgagatc	1920
tccaactaca ccgacatcat ctaccgctg atcgaggagt cccagaacca gcaggagaag	1980
aacgagcagg agctgtggc cctggacaag tgggcctccc tgtggaactg gtctgacatc	2040
accaactggc tgtgttacat caagatctt atcatgatcg tgggcggcct gatcggcctg	2100

cgcatcgtgt tcgccgtgct gtccatcgtg aaccgcgtgc gccagggcta ctccccctg 2160

tccttccaga cctgatccc caacccccgc ggccccgacc gccccgagg catcgaggag 2220

gagggcggcg agcagggccg cgaccgctcc atccgcctgg tgaacggctt cctggccctg 2280

gcctgggacg acctgcgctc cctgtgccctg ttctcctacc accgcctgcg cgacttcac 2340

ctgatcgccg cccgcaccgt ggagctgctg ggccgcccgt cctgcgcgg cctgcagaag 2400

ggctgggagg cctgaagta cctgggcaac ctgctgcagt actggggcca ggagctgaag 2460

aactccgcca tctccctgct ggacaccacc gccatcgccg tggccgaggg caccgaccgc 2520

gtgatcgaga tcgtgcagcg cgctgccgc gccatcctga acatccccg ccgatccgc 2580

cagggcctgg agcgcgcctt gctgtaa 2607

<210> 3

<211> 2535

<212> DNA

<213> Human immunodeficiency virus

<400> 3

gccgccatgc gcgtgatggg catcctgcgc aactgccagc agtggaggat ctggggcatc 60

ctgggcttct ggatgctgat gatctgctcc gtgggtggga acctgtgggt gaccgtgtac 120

tacggcgtgc ccgtgtggaa ggaggccaag accacctgt tctgcgcctc cgacgccaag 180

gcctacgagc gcgaggtgca caactgtgtg gccaccacg cctgcgtgcc caccgacccc 240

aacccccagg agatggtgct ggagaacgtg accgagaact tcaacatgtg gaagaacgac 300

atggtggacc agatgcacga ggacatcatc tcctgtggg accagtcct gaagccctgc 360

gtgaagctga cccccctgtg cgtgacctg aactgcacca acgtgaccaa cgccaccaac 420

aacacctaca acggcgagat gaagaactgc tccttcaaca tcaccaccga gctgcgcgac 480

aagaagaaga aggagtacg cctgttctac cgctggaca tcgtgccct gaacgagaac 540

tcctccgagt accgctgat caactgcaac acctccgcca tcaccaggc ctgccccaa 600

gtgtccttcg accccatccc catecactac tgcgccccg ccggtacgc catcctgaag 660

tgaacaaca agaccttcaa cggcaccggc ccctgcaaca acgtgtccac cgtgcagtgc 720

accacggca tcaagccgt ggtgtccacc cagctgtgc tgaacggctc cctggccgag 780

gaggagatca tcatccgctc cgagaacctg accgacaacg ccaagacat catcgtgcag 840

ctgaacgagt ccgtggagat cgtgtgcacc cgcaccaaca acaacaccg caagtccatg 900

cgcatcgcc cggccagac ctctacgcc accggcgaca tcatggcga catccgccag 960

gcccactgca acatctcca ggacaagtgg aacaagacc tgcagcaggt ggccgagaag 1020

ctgggcaagc acttcccaa caagaccatc accttcgagc cctcctccgg cggcgacctg	1080
gagatcacca cccactcctt caactgccgc ggcgagttct tctactgcaa cacctccaag	1140
ctgttcaact ccacctaaa caacaacacc aactccaact ccaccatcac cctgccctgc	1200
cgcataaagc agatcatcaa catgtggcag ggcgtgggcc aggccatgta ccccccccc	1260
atcgccggca acatcacctg caagtccaac atcaccggcc tgcctgtgac ccgcgacggc	1320
ggcaaggaga acaccacga gaccttcgc cccggcggcg gcgacatgac cgacaactgg	1380
cgtctcgagc tgiacaagta caagtggtg gagatcaagc ccctgggcgt ggcacccacc	1440
gaggccaagc gccgcgtggt ggagcgcgag aagcgcgccg tgggcctggg cgcctgttc	1500
ctgggcttcc tgggcgcgcg cggtccacc atgggcgccg cctccatcac cctgaccgtg	1560
caggccccgc agctgtgtc cggcatcgtg cagcagcagt ccaacctgct gcgcgccatc	1620
gaggccccgc agcatatgct gcagctgacc gtgtggggca tcaagcagct gcaggccccgc	1680
gtgtggcca tggagcgcta cctgaaggac cagcagctgc tgggcatctg ggcctgtctc	1740
ggcaagctga tctgcaccac cgccgtgcc tggaaactct cctgttcaa caagtcctg	1800
gacgacatct gggacaacat gacctggatg gactgggacc gcgagatctc caactacacc	1860
gacacatct accgcctgct ggaggagtcc cagaaccagc aggagaagaa cgagcaggac	1920
ctgttgccc tggactcctg ggagaacctg tggaaactgt tcgacatcac caactggctg	1980
tggtacatca agatcttcat catgatcgtg ggcggcctga tcggcctgac catcatcttc	2040
gccgtgctgt ccatcgtgaa ccgcgtgcgc cagggtact cccccctgtc cttccagacc	2100
ctgacccca acccccgccg ccccgaccgc ctggagcgca tcgaggagga gggcgccgag	2160
caggaccgcg accgctccat ccgcctggtg tccggttcc tggccctggc ctgggacgac	2220
ctgcgtccc tgtgctgtt ctctaccac cgctgcgcg acttcactct gatcgccgc	2280
cgcacgtgg agctgtggg ccgtcctcc ctgcgcggcc tgcagcgcgg ctgggaggcc	2340
ctgaagtacc tgggctcct ggtgcagtac tggggccagg agctgaagaa gtccgccatc	2400
tcctgctgg acaccatgc catcgccgtg gccgaggga ccgaccgat catcgaggtg	2460
gtgcagcgcg cctgccgcgc catcctgaac atccccgcc gcatccgca ggccttcgag	2520
gccgcctgc tgtaa	2535
<210> 4	
<211> 2535	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 4	

gccgccatgc gcgtgatggg catcctgcgc aactgccagc agtggaggat ctggggcatc	60
ctgggcttct ggatgctgat gatctgcaac gtggaggga acctgtgggt gaccgtgtac	120
tacggcgtgc ccgtgtggaa ggaggccaag accacctgt tctgcgcctc cgacgccaag	180
gcctacgaga aggaggtgca caacgtgtgg gccacccacg cctgcgtgcc caccgacccc	240
aacccccagg agatgggtgt ggagaacgtg accgagaact tcaacatgtg gaagaacgac	300
atggaggacc agatgcacga ggacatcatc tccctgtggg accagtccct gaagccctgc	360
gtgaagctga cccccctgtg cgtgacctg aactgccgca acgtgaccaa cgccaccaac	420
aacacctaca acgaggagat caagaactgc tccttcaaca tcaccaccga gctgcgcgac	480
aagaagaaga aggtgtacgc cctgttctac cgcctggaca tcgtgccctt gaacgagaac	540
tcctccgagt accgcctgat caactgcaac acctccgcca tcaccaggc ctgccccaa	600
gtgtccttcg accccatccc catccactac tgcgccccg ccggtacgc catcctgaag	660
tgcaacaaca agaccttcaa cggcaccggc cctgcaaca acgtgtccac cgtgcagtgc	720
accacggca tcaagccgt ggtgtccacc cagctgtgc tgaacggctc cctggccgag	780
gaggagatca tcatcgcctc cgagaacctg accaacaacg ccaagacat catcgtgac	840
ctgaacgagt ccgtggagat cgtgtgcacc cgcccaaca acaacaccg caagtccatc	900
cgcacggcc ccggccagac cttctacgc accggcgaca tcacggcgga catccgccag	960
gcccactgca acatctccga ggacaagtgg aacaagacc tgcagcgct gtccaagaag	1020
ctgaaggagc acttcccaa caagaccatc aagtgcgag cctcctccg cggcgacctg	1080
gagatcacca cccactcctt caactgccg ggcgagtct tctactgaa cacctccaag	1140
ctgttcaact ccacctaca caacaacacc aactccaact ccaccatcac cctgccctgc	1200
cgcacgaagc agatcatcaa catgtggcag gaggtgggccc gcgccatgta ccccccccc	1260
atgcgcggca acatcacctg caagtccaac atcaccggcc tgctgtgtgac ccgcgacggc	1320
ggcaagaaga acaccaccga gatcttccg cccggcgggc gcgacatgcg cgacaactgg	1380
cgtccgagc tgtacaagta caaggtgtg gagatcaagc cctgggctg ggccccacc	1440
aaggccaagc gccgcgtggt ggagcgcgag aagcgcgccg tgggcatcgg cgccgtgttc	1500
ctgggcttcc tgggcgccc cggctccacc atgggcgcc cctccatcac cctgaccgtg	1560
caggccccgc agctgtgtc cggcatcgtg cagcagcagt ccaacctgt gcgcgccatc	1620
gaggccagc agcatgtct gcagctgacc gtgtggggca tcaagcagct gcagaccgc	1680
gtgtggcca tcgagcgta cctgaaggac cagcagctgc tgggcatctg gggtgtctc	1740
ggcaagctga tctgaccac cgccgtgcc tggaaactc cctggtccaa caagtcccag	1800

gaggacatct gggacaacat gacctggatg cagtgggacc gcgagatctc caactacacc 1860

gacaccatct accgcctgct ggaggactcc cagaaccagc aggagaagaa cgagaaggac 1920

ctgctggccc tggactcctg gaagaacctg tggaaactggt tcgacatcac caactggctg 1980

tggtagatca agatcttcat catgatcgtg ggcggcctga tcggcctgcg catcatcttc 2040

gccgtgctgt ccatcgtgaa ccgcgtgcgc cagggctact cccccctgtc cttccagacc 2100

ctgaccccca acccccgagg ccccgaccgc ctgggcccga tcgaggagga gggcggcgag 2160

caggaccgag accgctccat ccgcctgggtg tccggcttcc tggccctggc ctgggacgac 2220

ctgcgtctcc tgtgctgtt ctctaccac cgcctgcgcg acttcatcct ggtggccgcc 2280

cgcgccgtgg agctgctggg ccgctcctcc ctgcgcggcc tgcagcgagg ctgggaggcc 2340

ctgaagtacc tgggctccct ggtgcagtac tggggcctgg agctgaagaa gtccgccatc 2400

tccctgctgg acaccatcgc catcgccgtg gccgagggca ccgaccgcat catcgagctg 2460

atccagcgca tctgccgcgc catcgcaac atccccgcc gcatccgcca gggttcgag 2520

gccgccctgc agtaa 2535

<210> 5

<211> 842

<212> PRT

<213> Human immunodeficiency virus

<400> 5

Met Arg Val Met Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Ser Val Val Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Arg Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asn Val Thr Asn Ala Thr Asn Asn Thr Tyr Asn Gly Glu
 130 135 140
 Met Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg Asp Lys Lys

 145 150 155 160
 Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asn
 165 170 175
 Glu Asn Ser Ser Glu Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile
 180 185 190
 Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr
 195 200 205
 Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe
 210 215 220

 Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His
 225 230 235 240
 Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu
 245 250 255
 Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Ala
 260 265 270
 Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Val Cys Thr
 275 280 285
 Arg Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln

 290 295 300
 Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His
 305 310 315 320
 Cys Asn Ile Ser Glu Asp Lys Trp Asn Lys Thr Leu Gln Gln Val Ala
 325 330 335
 Glu Lys Leu Gly Lys His Phe Pro Asn Lys Thr Ile Thr Phe Glu Pro
 340 345 350
 Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg

355 Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr
 370 375 380
 Asn Asn Asn Thr Asn Ser Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile
 385 390 395 400
 Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Gln Ala Met Tyr Ala
 405 410 415
 Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu
 420 425 430
 Leu Leu Thr Arg Asp Gly Gly Lys Glu Asn Thr Thr Glu Thr Phe Arg
 435 440 445
 Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys
 450 455 460
 Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala
 465 470 475 480
 Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Leu Gly Ala
 485 490 495
 Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 500 505 510
 Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val
 515 520 525
 Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met
 530 535 540
 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu
 545 550 555 560
 Ala Met Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 565 570 575
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser
 580 585 590
 Trp Ser Asn Lys Ser Leu Asp Asp Ile Trp Asp Asn Met Thr Trp Met
 595 600 605

Glu Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu
 610 615 620
 Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu
 625 630 635 640
 Ala Leu Asp Ser Trp Glu Asn Leu Trp Asn Trp Phe Asp Ile Thr Asn
 645 650 655
 Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile
 660 665 670
 Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg
 675 680 685
 Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Pro Arg
 690 695 700
 Gly Pro Asp Arg Leu Glu Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp
 705 710 715 720
 Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp
 725 730 735
 Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
 740 745 750
 Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Ser Ser
 755 760 765
 Leu Arg Gly Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser
 770 775 780
 Leu Val Gln Tyr Trp Gly Gln Glu Leu Lys Lys Ser Ala Ile Ser Leu
 785 790 795 800
 Leu Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile
 805 810 815
 Glu Val Val Gln Arg Ala Cys Arg Ala Ile Leu Asn Ile Pro Arg Arg
 820 825 830
 Ile Arg Gln Gly Phe Glu Ala Ala Leu Leu
 835 840
 <210> 6
 <211> 842

<212> PRT

<213> Human immunodeficiency virus

<400> 6

Met Arg Val Met Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Val Val Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Arg Asn Val Thr Asn Ala Thr Asn Asn Thr Tyr Asn Glu Glu

130 135 140

Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg Asp Lys Lys

145 150 155 160

Lys Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asn

165 170 175

Glu Asn Ser Ser Glu Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile

180 185 190

Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr

195 200 205

Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe

210 215 220

Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His

225 230 235 240

Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu

245 250 255

Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ala

260 265 270

Lys Thr Ile Ile Val His Leu Asn Glu Ser Val Glu Ile Val Cys Thr

275 280 285

Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln

290 295 300

Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His

305 310 315 320

Cys Asn Ile Ser Glu Asp Lys Trp Asn Lys Thr Leu Gln Arg Val Ser

325 330 335

Lys Lys Leu Lys Glu His Phe Pro Asn Lys Thr Ile Lys Phe Glu Pro

340 345 350

Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg

355 360 365

Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr

370 375 380

Asn Asn Asn Thr Asn Ser Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile

385 390 395 400

Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala

405 410 415

Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu

420 425 430

Leu Leu Thr Arg Asp Gly Gly Lys Lys Asn Thr Thr Glu Ile Phe Arg

435 440 445

Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys

450 455 460

Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Ala

465 470 475 480
 Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala
 485 490 495
 Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 500 505 510
 Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val

 515 520 525
 Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met
 530 535 540
 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu
 545 550 555 560
 Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 565 570 575
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser
 580 585 590

 Trp Ser Asn Lys Ser Gln Glu Asp Ile Trp Asp Asn Met Thr Trp Met
 595 600 605
 Gln Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu
 610 615 620
 Leu Glu Asp Ser Gln Asn Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu
 625 630 635 640
 Ala Leu Asp Ser Trp Lys Asn Leu Trp Asn Trp Phe Asp Ile Thr Asn
 645 650 655
 Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile

 660 665 670
 Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg
 675 680 685
 Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Pro Arg
 690 695 700
 Gly Pro Asp Arg Leu Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp
 705 710 715 720
 Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp

725 730 735
 Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
 740 745 750
 Phe Ile Leu Val Ala Ala Arg Ala Val Glu Leu Leu Gly Arg Ser Ser
 755 760 765
 Leu Arg Gly Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser
 770 775 780
 Leu Val Gln Tyr Trp Gly Leu Glu Leu Lys Lys Ser Ala Ile Ser Leu
 785 790 795 800
 Leu Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile

 805 810 815
 Glu Leu Ile Gln Arg Ile Cys Arg Ala Ile Arg Asn Ile Pro Arg Arg
 820 825 830
 Ile Arg Gln Gly Phe Glu Ala Ala Leu Gln
 835 840
 <210> 7
 <211> 842
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 7
 Met Arg Val Met Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp
 1 5 10 15
 Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Val Val Gly Asn

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

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

Lys Lys Leu Lys Glu His Phe Pro Asn Lys Thr Ile Lys Phe Glu Pro
 340 345 350
 Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg
 355 360 365
 Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr
 370 375 380

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

 450 455 460
 Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Ala
 465 470 475 480
 Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala
 485 490 495
 Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 500 505 510
 Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val
 515 520 525

 Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met
 530 535 540
 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu
 545 550 555 560
 Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 565 570 575
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser
 580 585 590

Trp Ser Asn Lys Ser Gln Glu Asp Ile Trp Asp Asn Met Thr Trp Met

595 600 605

Gln Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu

610 615 620

Leu Glu Asp Ser Gln Asn Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu

625 630 635 640

Ala Leu Asp Ser Trp Lys Asn Leu Trp Asn Trp Phe Asp Ile Thr Asn

645 650 655

Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile

660 665 670

Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg

675 680 685

Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Thr Arg

690 695 700

Gly Pro Asp Arg Leu Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp

705 710 715 720

Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp

725 730 735

Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp

740 745 750

Phe Ile Leu Val Ala Ala Arg Ala Val Glu Leu Leu Gly Arg Ser Ser

755 760 765

Leu Arg Gly Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser

770 775 780

Leu Val Gln Tyr Trp Gly Leu Glu Leu Lys Lys Ser Ala Ile Ser Leu

785 790 795 800

Leu Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile

805 810 815

Glu Leu Ile Gln Arg Ile Cys Arg Ala Ile Arg Asn Ile Pro Arg Arg

820 825 830

Ile Arg Gln Gly Phe Glu Ala Ala Leu Gln

835 840

<210> 8

<211> 842

<212> PRT

<213> Human immunodeficiency virus

<400> 8

Met Arg Val Met Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Ser Val Val Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Arg Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asn Val Thr Asn Ala Thr Asn Asn Thr Tyr Asn Gly Glu

130 135 140

Met Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg Asp Lys Lys

145 150 155 160

Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asn

165 170 175

Glu Asn Ser Ser Glu Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile

180 185 190

Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr

195 200 205

Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe
 210 215 220
 Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His
 225 230 235 240
 Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu
 245 250 255
 Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Ala
 260 265 270
 Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Val Cys Thr
 275 280 285
 Arg Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln
 290 295 300
 Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His
 305 310 315 320
 Cys Asn Ile Ser Glu Asp Lys Trp Asn Lys Thr Leu Gln Gln Val Ala
 325 330 335
 Glu Lys Leu Gly Lys His Phe Pro Asn Lys Thr Ile Thr Phe Glu Pro
 340 345 350
 Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg
 355 360 365
 Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr
 370 375 380
 Asn Asn Asn Thr Asn Ser Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile
 385 390 395 400
 Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Gln Ala Met Tyr Ala
 405 410 415
 Pro Pro Ile Ala Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu
 420 425 430
 Leu Leu Thr Arg Asp Gly Gly Lys Glu Asn Thr Thr Glu Thr Phe Arg
 435 440 445
 Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys

450 Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala
 465 470 475 480
 Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Leu Gly Ala
 485 490 495
 Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 500 505 510
 Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val
 515 520 525
 Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met

 530 535 540
 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu
 545 550 555 560
 Ala Met Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 565 570 575
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser
 580 585 590
 Trp Ser Asn Lys Ser Leu Glu Asp Ile Trp Asp Asn Met Thr Trp Met
 595 600 605

 Glu Trp Asp Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu
 610 615 620
 Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu
 625 630 635 640
 Ala Leu Asp Ser Trp Glu Asn Leu Trp Asn Trp Phe Asp Ile Thr Asn
 645 650 655
 Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile
 660 665 670
 Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg

 675 680 685
 Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Thr Arg
 690 695 700

Gly Pro Asp Arg Leu Glu Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp
705 710 715 720
Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp
725 730 735
Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
740 745 750

Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Ser Ser
755 760 765
Leu Arg Gly Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser
770 775 780
Leu Val Gln Tyr Trp Gly Gln Glu Leu Lys Lys Ser Ala Ile Ser Leu
785 790 795 800
Leu Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile
805 810 815
Glu Val Val Gln Arg Ala Cys Arg Ala Ile Leu Asn Ile Pro Arg Arg
820 825 830
Ile Arg Gln Gly Phe Glu Ala Ala Leu Leu
835 840
<210> 9
<211> 493
<212> PRT
<213> Human immunodeficiency virus
<400> 9
Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Thr Trp
1 5 10 15
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Arg Tyr Met Ile Lys
20 25 30
His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
35 40 45
Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Met Lys Gln Leu
50 55 60
Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn

65	70						75						80						
Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Glu	Lys	Ile	Glu	Val	Arg	Asp				
				85							90							95	
Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Gln				
				100							105							110	
Gln	Lys	Thr	Gln	Gln	Ala	Glu	Ala	Ala	Ala	Asp	Gly	Lys	Val	Ser	Gln				
				115							120							125	
Asn	Tyr	Pro	Ile	Val	Gln	Asn	Leu	Gln	Gly	Gln	Met	Val	His	Gln	Ala				
				130							135							140	
Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Ile	Glu	Glu	Lys				
				145							150							155	
											160								
Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Thr	Ala	Leu	Ser	Glu	Gly				
				165							170							175	
Ala	Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His				
				180							185							190	
Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Asp	Thr	Ile	Asn	Glu	Glu	Ala	Ala				
				195							200							205	
Glu	Trp	Asp	Arg	Leu	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly				
				210							215							220	
Gln	Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr				
				225							230							235	
											240								
Leu	Gln	Glu	Gln	Ile	Ala	Trp	Met	Thr	Ser	Asn	Pro	Pro	Val	Pro	Val				
				245							250							255	
Gly	Asp	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val				
				260							265							270	
Arg	Met	Tyr	Ser	Pro	Val	Ser	Ile	Leu	Asp	Ile	Lys	Gln	Gly	Pro	Lys				
				275							280							285	
Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Phe	Lys	Thr	Leu	Arg	Ala				
				290							295							300	
Glu	Gln	Ala	Thr	Gln	Asp	Val	Lys	Asn	Trp	Met	Thr	Asp	Thr	Leu	Leu				
				305							310							315	
											320								

Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly

325 330 335

Pro Gly Ala Ser Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly

340 345 350

Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala

355 360 365

Asn Asn Thr Asn Ile Met Met Gln Arg Ser Asn Phe Lys Gly Pro Lys

370 375 380

Arg Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg

385 390 395 400

Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu

405 410 415

Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly

420 425 430

Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser

435 440 445

Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu

450 455 460

Thr Thr Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr

465 470 475 480

Ser Leu Lys Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln

485 490

<210> 10

<211> 207

<212> PRT

<213> Human immunodeficiency virus

<400> 10

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Val

1 5 10 15

Arg Glu Arg Ile Arg Arg Thr Glu Pro Ala Ala Glu Gly Val Gly Ala

20 25 30

Ala Ser Gln Asp Leu Asp Lys Tyr Gly Ala Leu Thr Ser Ser Asn Thr
 35 40 45
 Ala Thr Asn Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 55 60
 Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
 65 70 75 80
 Thr Tyr Lys Ala Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly
 85 90 95
 Gly Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp

100 105 110
 Leu Trp Val Tyr His Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr
 115 120 125
 Thr Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe
 130 135 140
 Lys Leu Val Pro Val Asp Pro Arg Glu Val Glu Glu Ala Asn Glu Gly
 145 150 155 160
 Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Met Glu Asp
 165 170 175

Glu Asp Arg Glu Val Leu Lys Trp Lys Phe Asp Ser His Leu Ala Arg
 180 185 190
 Arg His Met Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys
 195 200 205

<210> 11
 <211> 1491
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 11

gccgcgcgcca tgggcgccccg cgccagcatc ctgcgcggcg gcaagctgga cacctgggag 60
 aagatccgcc tgcgccccgg cggaagaag cgctacatga tcaagcacct ggtgtgggcc 120
 agccgcgagc tggagcgctt cgccctgaac cccggcctgc tggagaccag cgagggtgc 180
 aagcagatca tgaagcagct gcagcccgcc ctgcagaccg gcaccgagga gctgcgcagc 240
 ctgtacaaca ccgtggccac cctgtactgc gtgcacgaga agatcgaggt gcgcgacacc 300

aaggaggccc tggacaagat cgaggaggag cagaacaaga gccagcagaa gacccagcag 360
 gccgaggccg ccgccgacgg caaggtgagc cagaactacc ccatcgtgca gaacctgcag 420
 ggccagatgg tgcaccaggc catcagcccc cgcacctga acgcctgggt gaaggtgatc 480
 gaggagaagg ccttcagccc cgaggtgatc cccatgttca ccgcctgag cgagggcgcc 540
 acccccagg acctgaacac catgctgaac accgtgggcg gccaccaggc cgccatgcag 600

atgctgaagg acaccatcaa cgaggaggcc gccgagtggg accgcctgca ccccgtagc 660
 gccggcccca tcgccccgg ccagatgcgc gagccccgcg gcagcgacat cgccggcacc 720
 accagcacc tgcaggagca gatcgctgg atgaccagca accccccgt gcccgtaggc 780
 gacatctaca agcgctggat catcctgggc ctgaacaaga tcgtgcgcat gtacagcccc 840
 gtgagcatcc tggacatcaa gcagggcccc aaggagccct tccgcgacta cgtggaccgc 900
 ttcttcaaga ccttcgcgcg cgagcaggcc acccaggacg tgaagaactg gatgaccgac 960
 accctgctgg tgcagaacgc caaccccgac tgcaagacca tcctgcgcgc cctgggcccc 1020

ggcgccagcc tggaggagat gatgaccgcc tgccagggcg tgggcggccc cagccacaag 1080
 gcccgctgc tggccgagc catgagccag gccaacaaca ccaacatcat gatgcagcgc 1140
 agcaacttca agggcccca gcgcacgtg aagtgttca actgcggcaa ggagggccac 1200
 atgccccga actgccgcgc cccccgaag aagggtgct ggaagtgcgg caaggagggc 1260
 caccagatga aggactgcac cgagcgccag gccaacttcc tgggcaagat ctggcccagc 1320
 cacaaggcc gccccggcaa ttctctgcag agccgccccg agcccaccgc cccccgcgc 1380
 gagagcttcc gcttcgagga gaccacccc gccccaagc aggagcccaa ggaccgcgag 1440

cccctgacca gcctgaagag cctgttcggc agcgacccc tgagccagta a 1491

<210> 12
 <211> 633
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 12

gccgccgcca tgggcggcaa gtggagcaag agcagcatcg tgggctggcc cgccgtgcgc 60
 gagcgcatcc gccgcaccga gcccgccgc gagggcgtgg gcgccccag ccaggacctg 120
 gacaagtacg gcgcccagc cagcagcaac accgccacca acaacgccga ctgcgctgg 180
 ctggaggccc aggaggagga ggaggagggt ggcttccccg tgcgccccca ggtgccctg 240
 cgccccatga cctacaaggc gccttcgac ctgagcttct tctgaagga gaaggcggc 300

ctggagggcc tgatctacag caagaagcgc caggagatcc tggacctgtg ggtgtaccac 360

accagggct tcttccccga ctggcagaac tacacccccg gccccggcgt gcgctacccc 420
ctgaccttcg gctgggtgctt caagctgggtg cccgtggacc cccgcgaggt ggaggaggcc 480
aacgagggcg agaacaactg cctgtctgcac cccatgagcc agcacggcat ggaggacgag 540
gaccgcgagg tgctgaagtg gaagttcgac agccacctgg cccgccgcca catggcccgc 600
gagctgcacc ccgagtacta caaggactgc tga 633

<210> 13
<211> 852
<212> PRT

<213> Human immunodeficiency virus

<400> 13

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asn Val Asn Val Thr Asn Thr Thr Asn Asn Thr Glu Glu

130 135 140

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg

145 150 155 160

Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val

165 170 175

Pro Ile Asp Asp Asn Asn Asn Asn Ser Ser Asn Tyr Arg Leu Ile Asn
180 185 190

Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu
195 200 205

Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys
210 215 220

Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser
225 230 235 240

Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu
245 250 255

Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu
260 265 270

Asn Ile Thr Asn Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser
275 280 285

Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
290 295 300

Arg Ile Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly
305 310 315 320

Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Thr Lys Trp Asn Lys
325 330 335

Thr Leu Gln Gln Val Ala Lys Lys Leu Arg Glu His Phe Asn Asn Lys
340 345 350

Thr Ile Ile Phe Lys Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr
355 360 365

His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly
370 375 380

Leu Phe Asn Ser Thr Trp Ile Gly Asn Gly Thr Lys Asn Asn Asn Asn
385 390 395 400

Thr Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn
405 410 415

Met Trp Gln Gly Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly

420 425 430
 Lys Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp
 435 440 445
 Gly Gly Asn Asn Asn Thr Asn Glu Thr Glu Ile Phe Arg Pro Gly Gly
 450 455 460
 Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val
 465 470 475 480
 Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg
 485 490 495

 Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu
 500 505 510
 Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr
 515 520 525
 Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln
 530 535 540
 Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu
 545 550 555 560
 Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu

 565 570 575
 Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly
 580 585 590
 Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn Ser Ser Trp Ser Asn
 595 600 605
 Lys Ser Gln Asp Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Glu
 610 615 620
 Arg Glu Ile Asn Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu
 625 630 635 640

 Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp
 645 650 655
 Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp
 660 665 670
 Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg

675 680 685
Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr

690 695 700
Ser Pro Leu Ser Phe Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Asp

705 710 715 720
Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg

725 730 735
Ser Ile Arg Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu

740 745 750
Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu

755 760 765
Ile Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Lys Gly Leu Arg Arg

770 775 780

Gly Trp Glu Ala Leu Lys Tyr Leu Trp Asn Leu Leu Gln Tyr Trp Gly
785 790 795 800

Gln Glu Leu Lys Asn Ser Ala Ile Ser Leu Leu Asp Thr Thr Ala Ile
805 810 815

Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val Val Gln Arg Ala
820 825 830

Cys Arg Ala Ile Leu Asn Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu
835 840 845

Arg Ala Leu Leu

850
<210> 14
<211> 2568
<212> DNA
<213> Human immunodeficiency virus
<400> 14

gccgccgcca tgcgcgtgcg cggcatccag cgcaactgcc agcacctgtg gcgctggggc 60
acctgatcc tgggcatgct gatgatctgc tccgccgccg agaacctgtg ggtgaccgtg 120
tactacggcg tgcccgtgtg gaaggaggcc aacaccaccc tgttctgcgc ctccgacgcc 180
aaggcctacg acaccgaggt gcacaacgtg tgggccaccc acgcctgcgt gcccaccgac 240

cccaaccccc aggagatcgt gctggagaac gtgaccgaga acttcaacat gtggaagaac	300
aacatggtgg agcagatgca cgaggacatc atctccctgt gggaccagtc cctgaagccc	360
tgctgaagc tgacccccct gtgcgtgacc ctgaactgca ccaacgtgaa cgtgaccaac	420
accaccaaca acaccgagga gaaggcgag atcaagaact gctccttcaa catcaccacc	480
gagatccgcg acaagaagca gaaggtgtac gccctgttct accgcctgga cgtggtgccc	540
atcgacgaca acaacaacaa ctctccaac taccgcctga tcaactgaa cacctccgcc	600
atcaccagg cctgccccaa ggtgtccttc gagcccatcc ccatccacta ctgcgcccc	660
gccggcttcg ccatcctgaa gtgcaacgac aagaagtca acggcaccgg cccctgcaag	720
aacgtgtcca ccgtgcagtg caccacggc atcaagccc tggtgtccac ccagctgctg	780
ctgaacggct ccttgccga ggaggagatc atcatccgt ccgagaacat caccaacaac	840
gccagacca tcatcgtgca gctgaacgag tccgtggaga tcaactgca cgccecaac	900
aacaacccc gcaagtccat ccgcatcggc ccggccagg ccttctacgc caccggcgac	960
atcatcggcg acatccgcca ggcccactgc aacatctccg gcaccaagt gaacaagacc	1020
ctgcagcagg tggccaagaa gctgcgcgag cacttcaaca acaagaccat catcttcaag	1080
ccctctccg gggcgacct ggagatcacc acccactcct tcaactgccg cggcgagttc	1140
ttctactgca acacctccg cctgttcaac tccacctgga tcggcaacgg caccaagaac	1200
aacaacaaca ccaacgacac catcacctg cctgccgca tcaagcagat catcaacatg	1260
tggcaggcg tgggccaggc catgtacgcc ccccccacg agggcaagat cacctgcaag	1320
tccaacatca ccggcctgct gctgaccgc gacggcggca acaacaacac caacgagacc	1380
gagatcttc gcccggcgg cgcgacatg cgcgacaact ggcgctccga gctgtacaag	1440
tacaagtg tgaagatgca gccctgggc gtggcccca ccaaggcca gcgccgctg	1500
gtggagcgcg agaagcgcg cgtgggcac ggcgccgtgt tctgggctt cctgggcgc	1560
gccggctcca ccatgggcgc cgcctccat accctgaccg tgcaggccc ccagctgctg	1620
tccggcatcg tgcagcagca gtccaacctg ctgcgcgcca tcgaggcca gcagacctg	1680
ctgcagctga ccgtgtggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc	1740
tacctgaagg accagcagct gctgggcac tgggctgct ccggcaagct gatctgcacc	1800
accacctgc cctggaactc ctctggtcc aacaagtccc aggacgagat ctgggacaac	1860
atgacctgga tggagtggga gcgcgagatc aacaactaca ccgacatcat ctactccctg	1920
atcgaggagt cccagaacca gcaggagaag aacgagcagg agctgctggc cctggacaag	1980

tgggcctccc tgtggaactg gttcgacatc accaactggc tgtggtacat caagatcttc 2040

atcatgatcg tgggcggcct gatcggcctg cgcacgtgtg tcgccgtgct gtccatcgctg 2100

aaccgcgtgc gccagggcta ctccccctg tccttcaga cctgatccc caacccccgc 2160

ggccccgacc gccccgaggg catcgaggag gagggcggcg agcaggaccg cgaccgctcc 2220

atccgcctgg tgaacggctt cctggccctg gcctgggacg acctgcgtc cctgtgcctg 2280

ttctcctacc accgcctgcg cgacttcac ctgatcgccg cccgcaccgt ggagctgctg 2340

ggccgcaagg gcctgcgccg cggctgggag gcctgaagt acctgtgaa cctgctgcag 2400

tactggggcc aggagctgaa gaactccgcc atctccctgc tggacaccac cgccatcgcc 2460

gtggccgagg gcaccgaccg cgtgatcgag gtggtgcagc gcgcctgccg cgccatcctg 2520

aacatcccc gccgcatccg ccagggcctg gagcgcgcc tgctgtaa 2568

<210> 15

<211> 857

<212> PRT

<213> Human immunodeficiency virus

<400> 15

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Met Ile Leu Gly Met Ile Ile Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Asn Leu Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

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

Gly Leu Phe Asn Ser Thr Trp Asn Gly Asn Gly Thr Lys Lys Lys Asn
385 390 395 400

Ser Thr Glu Ser Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln
405 410 415

Ile Ile Asn Met Trp Gln Arg Val Gly Gln Ala Met Tyr Ala Pro Pro
420 425 430

Ile Gln Gly Val Ile Arg Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu
435 440 445

Thr Arg Asp Gly Gly Asp Asn Asn Ser Lys Asn Glu Thr Phe Arg Pro
450 455 460

Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr

465 470 475 480

Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys
485 490 495

Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val
500 505 510

Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser
515 520 525

Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln
530 535 540

Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu
545 550 555 560

Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala
565 570 575

Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys
580 585 590

Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp
595 600 605

Ser Asn Lys Ser Gln Ser Glu Ile Trp Asp Asn Met Thr Trp Leu Gln

610 615 620

Trp Asp Lys Glu Ile Ser Asn Tyr Thr Asp Ile Ile Tyr Asn Leu Ile

625 630 635 640
 Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala
 645 650 655
 Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp
 660 665 670
 Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly
 675 680 685

 Leu Arg Ile Val Phe Ala Val Leu Ser Val Ile Asn Arg Val Arg Gln
 690 695 700
 Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Thr Pro Asn Pro Gly Gly
 705 710 715 720
 Leu Asp Arg Pro Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Gly Arg
 725 730 735
 Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp
 740 745 750
 Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe

 755 760 765
 Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly His Ser Ser Leu
 770 775 780
 Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr Leu Trp Asn Leu
 785 790 795 800
 Leu Leu Tyr Trp Gly Arg Glu Leu Lys Ile Ser Ala Ile Asn Leu Leu
 805 810 815
 Asp Thr Ile Ala Ile Ala Val Ala Gly Trp Thr Asp Arg Val Ile Glu
 820 825 830

 Ile Gly Gln Arg Ile Cys Arg Ala Ile Leu Asn Ile Pro Arg Arg Ile
 835 840 845
 Arg Gln Gly Leu Glu Arg Ala Leu Leu
 850 855
 <210> 16
 <211> 2583
 <212> DNA

<213> Human immunodeficiency virus

<400> 16

gccgccgcca tgcgcgtgat gggcatccag cgcaactgcc agcacctgtg gcgctggggc	60
accatgatcc tgggcatgat catcatctgc tccgccgccc agaacctgtg ggtgaccgtg	120
tactacggcg tgcccgtgtg gaaggacgcc gagaccaccc tgttctgcgc ctccgacgcc	180
aaggcctacg acaccgaggt gcacaacgtg tgggccaccc acgcctgcgt gcccaccgac	240
cccaaccccc aggagatcaa cctggagaac gtgaccgagg agttcaacat gtggaagaac	300
aacatggtgg agcagatgca caccgacatc atctccctgt gggaccagtc cctgaagccc	360
tgcgtgaagc tgacccccct gtgcgtgacc ctgaactgct ccaacgtgaa cgtgaccacc	420
aacatcacca acatcacga caacatgaag ggcgagatca agaactgctc cttcaacatg	480
accaccgagc tgcgcgacaa gaagcagaag gtgtactccc tgttctacaa gctggacgtg	540
gtgcagatca acaagtccaa ctctctctcc cagtaccgcc tgatcaactg caacacctcc	600
gccatcacc accgcctgccc caaggtgtcc ttcgagccca tccccatcca ctactgcgc	660
cccgcgggtc tgcctacct gaagtgcgaag gacaaggagt tcaacggcac cggcccctgc	720
aagaacgtgt ccaccgtgca gtgcacccac ggcatcaagc ccgtggtgtc caccagctg	780
ctgtgaacg gctccctggc cgaggaggag gtgatgatcc gtcgagaa catcaccaac	840
aacgccaaga acatcatcgt gcagctgacc aagcccgtga agatcaactg caccgcccc	900
aacaacaaca cccgaagtc catccgcatc ggccccggcc aggccttcta cgccaccggc	960
gacatcatcg gcgacatccg ccaggccac tgcaacgtgt cccgcaccga gtggaacgag	1020
accctgcaga aggtggccaa gcagctgcgc aagtacttca acaacaagac catcatcttc	1080
accaactcct cggcgcgga cctggagatc accaccact cttcaactg cggcgcgag	1140
ttcttctact gcaacacctc cggcctgttc aatccacct ggaacggcaa cggcaccaag	1200
aagaagaact ccaccgagtc caacgacacc atcacctgc cctgccgat caagcagatc	1260
atcaacatgt ggcagcgcgt gggccaggcc atgtacgccc ccccatcca gggcgtgatc	1320
cgctgcgagt ccaacatcac cggcctgctg ctgaccgcg acggcggcga caacaactcc	1380
aagaacgaga ccttccgccc cggcggcggc gacatgcgcg acaactggcg ctccgagctg	1440
tacaagtaca agtggtgaa gatcgagccc ctggcggtgg cccccacaa ggccaagcgc	1500
cgcgtggtgg agcgcgagaa gcgcgccgtg ggcatcggcg ccgtgttcct gggttctctg	1560
ggcgccgccc gctccaccat gggcgccgcc tccataccc tgaccgtgca ggcccgcag	1620
ctgtgtccg gcatcgtgca gcagcagtc aacctgtgc gcgccatcga ggcccagcag	1680
cacctgtga agctgaccgt gtggggcatc aagcagctgc aggcccgct gctggccgtg	1740

gagcgtacc tgaaggacca gcagctgctg ggcatctggg gctgctccgg caagctgac	1800
tgcaccacca acgtgccctg gaactcctcc tggccaaca agtcccagtc cgagatctgg	1860
gacaacatga cctggctgca gtgggacaag gagatctcca actacaccga catcatctac	1920
aacctgatcg aggagtccca gaaccagcag gagaagaacg agcaggacct gctggccctg	1980
gacaagtggg ccaacctgtg gaactgggtc gacatctcca actggctgtg gtacatcaag	2040
atcttcatca tgatcgtggg cggcctgac ggcctgcgca tcgtgttcgc cgtgctgtcc	2100
gtgatcaacc gcgtgcgcca gggctactcc ccctgtcct tccagacca ccccccaac	2160
cccgccggcc tggaccgcc cggccgcac gaggaggagg gcggcgagca gggccgcgac	2220
cgctccatcc gcctgggtgc cggcttcctg gccctggcct gggacgacct gcgctccctg	2280
tgctgttct cctaccaccg cctgcgcgac ttcatcctga tcgccgccg caccgtggag	2340
ctgtgggcc actcctccct gaaggccctg cgctgggct gggaggccct gaagtacctg	2400
tggaacctgc tgctgtactg gggccgcgag ctgaagatct ccgccatcaa cctgctggac	2460
accatcgcca tcgccgtggc cggctggacc gaccgcgtga tcgagatcgg ccagcgcac	2520
tgccgcgcca tctgaacat ccccgccgc atccgccagg gcctggagcg cgccctgctg	2580
taa	2583
<210> 17	
<211> 1503	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 17	
gccgccgcca tgggcgccc cgccctcctg ctgtccggcg gcaagctgga cgccctgggag	60
aagatccgcc tgcgccccgg cggcaagaag aagtaccgcc tgaagcacct ggtgtgggcc	120
tcccgcgagc tggagcgctt cgccctgaac cccggcctgc tggagacctc cgagggtgc	180
aagcagatca tcggccagct gcagccgcc ctgcagaccg gctccgagga gctgcgctcc	240
ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgaggt gaaggacacc	300
aaggaggccc tggagaagat cgaggaggag cagaacaagt cccagcagaa gaccagcag	360
gccgccgccc acaagggcaa ctctccaag gtgtcccaga actaccccat cgtgcagaac	420
ctgcagggcc agatggtgca ccaggccatc tcccccgca cctgaacgc ctgggtgaag	480
gtgatcgagg agaaggcctt ctccccgag gtgatccca tgttctccgc cctgtccgag	540
ggcgccacc cccaggacct gaacacatg ctgaacaccg tgggcggcca ccaggccgcc	600
atgcagatgc tgaaggacac catcaacgag gaggccgcc agtgggaccg cctgcacccc	660

gtgcacgccg gcccacatccc ccccgccag atgcgcgagc cccgcggctc cgacatcgcc	720
ggcaccacct ccacctgca ggagcagatc gcctggatga cctccaaccc ccccatcccc	780
gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt gcgcatgtac	840
tccccctgtt ccatcctgga catccgccag ggccccaagg agcccttccg cgactacgtg	900
gaccgcttct tcaagacct gcgcgccgag caggccaccc aggacgtgaa gaactggatg	960
accgacaccc tgctggtgca gaacgccaac cccgactgca agaccatcct gaaggccctg	1020
ggccccggcg ccaccttgga ggagatgatg accgcctgcc agggcgtggg cggccccggc	1080
cacaaggccc gcgtgctggc cgaggccatg tcccaggtga ccaacgccgc catcatgatg	1140
cagcgccgca acttcaaggg ccagcgccgc atcatcaagt gcttcaactg cggcaaggag	1200
ggccacatcg cccgcaactg ccgcgcccc cgcaagaagg gctgctggaa gtgcggcaag	1260
gagggccacc agatgaagga ctgcaccgag cgccaggcca acttcctggg caagatctgg	1320
ccctcaaca agggccgccc cggcaacttc ctgcagtcgc gcccagacc caccgcccc	1380
cccgcgagt ccttcggctt cggcgaggag atcacccct ccccaagca ggagcccaag	1440
gacaaggagc cccccgtac ctccctgaag tcctgttcg gcaacgacc cctgtcccag	1500
taa	1503
<210> 18	
<211> 2856	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 18	
gccgcgccca tgccccagat caccctgtgg cagcgcccc ttggtgacct caagatcggc	60
ggccagctga aggaggccct gctggccacc ggcgccgacg acaccgtgct ggaggagatc	120
aacctgcccg gcaagtggaa gccaagatg atcgcgccga tcggcggctt catcaagggtg	180
cgccagtacg accagatcct gatcgagatc tgcggcaaga aggccatcgg caccgtgctg	240
gtgggcccc cccccgtgaa catcatcggc cgcaacatgc tgaccagat cggctgcacc	300
ctgaacttcc ccatctcccc catcgagacc gtgccgtga agctgaagcc cggcatggac	360
ggccccaagg tgaagcagtg gcccctgacc gaggagaaga tcaaggccct gaccgagatc	420
tgcaccgaga tggagaagga gggcaagatc tccaagatcg gccccgagaa cccctacaac	480
acccccatct tcgccatcaa gaagaaggac tccaccaagt ggcgcaagct ggtggacttc	540
cgcgagctga acaagcgcac ccaggacttc tgggaggtgc agctgggcat cccccaccc	600
gccggcctga agaagaagaa gtccgtgacc gtgctggacg tggcgcacgc ctacttctcc	660

gtgcccttgg acgaggactt ccgcaagtac accgccttca ccatccctc catcaacaac	720
gagacccccg gcatccgcta ccagtacaac gtgctgcccc agggctggaa gggtccccc	780
gccatcttcc agtcctccat gaccaagatc ctggagccct tccgcacca gaaccccag	840
atcgtgatct accagtacat ggacgacctg tacgtgggct ccgacctgga gatcgccag	900
caccgcgcca agatcgagga gctgcgcgag cacctgctgc gctggggctt caccacccc	960
gacaagaagc accagaagga gcccccttc ctgtggatgg gctacgagct gcaccccag	1020
aagtggaccg tgcagcccat ccagctgccc gagaaggact cctggaccgt gaacgacatc	1080
cagaagctgg tgggcaagct gaactgggcc tcccagatct accccggcat caaggtgaag	1140
cagctgtgca agctgctgcg cggcgccaag gccctgaccg acatcgtgcc cctgaccgag	1200
gaggccgagc tggagctggc cgagaaccgc gagatcctga aggagcccg gcacggcgtg	1260
tactacgacc cctccaagga cctgatcgcc gagatccaga agcagggcca ggaccagtgg	1320
acctaccaga tctaccagga gcccttcaag aacctcaaga ccggcaagta cgccaagatg	1380
cgctccgcc acaccaacga cgtgaagcag ctgaccgagg ccgtgcagaa gatcgccacc	1440
gagtccatcg tgatctgggg caagaccccc aagtccgccc tgeccatcca gaaggagacc	1500
tgggagacct ggtggaccga gtactggcag gccacctgga ttcccagtg ggagtctgtg	1560
aacaccccc ccttggtgaa gctgtggtac cagctggaga aggagcccat cgccggcgcc	1620
gagaccttct acgtggacgg cgccgccaac cgcgagacca agctgggcaa ggccggctac	1680
gtgaccgacc gggccgcca gaagtggtg tccttgaccg agaccaccaa ccagaaaacc	1740
gagctgcagg ccattccact ggccctgcag gactccggt ccgaggtgaa catcgtgacc	1800
gactcccagt acgcccctgg catcatccag gccagcccg acaagtccga gtccgagctg	1860
gtgaaccaga tcatcgagca gctgatcaag aaggagaagg tgtacctgtc ctgggtgcc	1920
gccacaagg gcatcgccgg caacgagcag gtggacaagc tgggtgccac cggcatccgc	1980
aaggtgctgt tcctggacgg catcgacaag gccagaggagg agcacgagaa gtaccactcc	2040
aactggcgcg ccatggcctc cgacttcaac ctgccccca tegtggccaa ggagatcgtg	2100
gcctcctgcg acaagtgcc gctgaagggc gaggccatgc acggccaggt ggactgctcc	2160
cccggcatct ggcagctgga ctgacccac ctggagggca agatcatcct ggtggccgtg	2220
cacgtggcct ccggctacat cgaggccgag gtgatecccg ccgagaccgg ccaggagacc	2280
gcctacttca tcctgaagct ggccggccgc tggccctga aggtgatcca caccgacaac	2340
ggctccaact tcacctccgc cgccgtgaag gccgcctgct ggtgggcccg catccagcag	2400

gagttcggca tccctacaa cccccagtcc cagggcgtgg tggagtccat gaacaaggag	2460
ctgaagaaga tcatcgcca ggtgcgcgac caggccgagc acctcaagac cgccgtgcag	2520
atggccgtgt tcatccaaa cttcaagcgc aagggcggca tcggcggcta ctccgccggc	2580
gagcgcatca tcgacatcat cgccaccgac atccagacca aggagctgca gaagcagatc	2640
accaagatcc agaacttccg cgtgtactac cgcgactccc gcgaccccat ctggaagggc	2700
cccgccaaagc tgctgtggaa gggcgagggc gccgtggtga tccaggacaa ctccgacatc	2760
aaggtggtgc cccgccgcaa ggccaagatc atccgcgact acggcaagca gatggccggc	2820
gacgactgcg tggccggccg ccaggacgag gactaa	2856
<210> 19	
<211> 633	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 19	
gccgccgcca tgggcccga gttgtccaag tctccatcg tgggtggcc cgccgtgcgc	60
gagcgcatcc gccgcacca ccccgccgc gaggcgtgg gcgccgtgtc ccaggacctg	120
gacaagcacg gcgccatcac ctctccaac accgccgcca acaaccccga ctgcgcctgg	180
ctggaggccc aggaggagga ggaggaggtg ggcttccccg tgcgcccga ggtgccctg	240
cgcccatga cctacaaggc cgccctggac ctgtccact tctgaagga gaaggcggc	300
ctggagggcc tgatctactc caagaagcgc caggagatcc tggacctgtg ggtgtaccac	360
accagggtct acttccccga ctggcagaac tacaccccg gcccggcat ccgtacccc	420
ctgaccttcg gctggtgtt caagctggtg cccgtggacc ccgaggaggt ggaggaggcc	480
aacgaggcg agaacaactc cctgtgcac ccatgtgcc agcacggcat ggaggacgag	540
gagcgcgagg tgctgatgtg gaagttcgac tccgcctgg ccctgcgcca catcgccgc	600
gagctgcacc ccgagtacta caaggactgc taa	633
<210> 20	
<211> 2856	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 20	
gccgccgcca tgccccagat caccctgtgg cagcgcccc tggtgtccat caaggtgggc	60
ggccagatca aggaggccct gctggccacc ggcgccgacg acaccgtgct ggaggagatc	120

aacctgcccc gcaagtggaa gccaagatg atcggcggca tcggcggctt catcaaggtg	180
cgccagtacg accagatcct gatcgagatc tgcggcaaga aggccatcgg caccgtgctg	240
gtgggcccc ccccggtgaa catcatcggc cgcaacatgc tgaccagct gggtgcacc	300
ctgaacttcc ccatctcccc catcgagacc gtgcccgtga agctgaagcc cggcatggac	360
ggccccaagg tgaagcagtg gccctgacc gaggagaaga tcaaggcct gaccgccatc	420
tgcgaggaga tggagaagga gggcaagatc accaagatcg gccccgagaa cccctacaac	480
acccccgtgt tcgcatcaa gaagaaggac tccaccaagt ggcgcaagct ggtggacttc	540
cgcgagtga acaagcgac ccaggacttc tgggaggtgc agctgggcat ccccccccc	600
gcccgcctga agaagaagaa gtccgtgacc gtgctggacg tggcgacgc ctacttctcc	660
gtgcccctgg acgagggtt ccgcaagtac accgccttca ccatcccctc catcaacaac	720
gagaccccc gcatccgcta ccagtacaac gtgctgcccc agggctggaa gggtcccc	780
gccatcttc agtctccat gaccaagatc ctggagccct tccgcgcca gaaccccgag	840
atcgtgatct accagtacat ggacgacctg tacgtgggt cgcacctgga gatcgccag	900
caccgcgcca agatcgagga gctgcgag cacctgtga agtggggctt caccacccc	960
gacaagaagc accagaagga gcccccttc ctgtggatgg gctacgagct gcacccgac	1020
aagtggaccg tgcagccat ccagctgcc gagaaggact cctggaccgt gaacgacatc	1080
cagaagctgg tgggcaagct gaactgggc tccagatct acccggcat caaggtgcgc	1140
cagctgtgca agctgtcgc cggcgccaag gcctgaccg acatcgtgcc cctgaccgag	1200
gaggccgagc tggagctggc cgagaaccgc gagatcctga aggagccgt gcacggcgtg	1260
tactacgacc cctccaagga cctgatgcc gagatccaga agcagggcc cgaccagtgg	1320
acctaccaga tctaccagga gcccttcaag aacctcaaga ccgcaagta cgccaagatg	1380
cgcaccgccc acaccaacga cgtgaagcag ctgaccgagg ccgtgcagaa gatcgccatg	1440
gagtccatcg tgatctgggg caagaccccc aagtccgcc tgcccatcca gaaggagacc	1500
tgggagacct ggtggaccga ctactggcag gccacctgga ttcccagtg ggagtctgtg	1560
aacaccccc cctggtgaa gctgtggtac cagctggaga aggagccat cgccggcgcc	1620
gagaccttct acgtggacgg cgccgccaac cgcgagacca agatcggcaa ggccggtac	1680
gtgaccgacc gggcgccga gaagatcgtg tccctgaccg agaccacaa ccagaaaacc	1740
gagctgcagg ccatccagct ggccctgcag gactccggt ccgaggtgaa catcgtgacc	1800
gactcccagt acgccctggg catcatccag gccagcccg acaagtccga gtccgagctg	1860
gtgaaccaga tcatcgagca gctgatcaag aaggagcgcg tgtacctgtc ctgggtgccc	1920
gcccacaagg gcatcggcgg caacgagcag gtggacaagc tgggtgtctc cgcatccgc	1980

aaggtgctgt tcttgacg catcgacaag gcccaggagg agcacagaaa gtaccactcc 2040
aactggcgcg ccatggcctc cgagttcaac ctgcccccca tctgggcaa ggagatcgtg 2100
gcctcctgcg acaagtgcc gctgaagggc gaggccatgc acggccaggt ggactgctcc 2160

cccggcatct ggcagctgga ctgcaccac ctggagggca agatcatcct ggtggccgtg 2220
cacgtggcct cgggctacat cgaggccgag gtgatccccg ccgagaccgg ccaggagacc 2280
gcctacttca tctgaagct ggccggcgcg tggccgtga aggtgatcca caccgacaac 2340
ggctccaact tcacctccgc cgccgtgaag gccgcctgct ggtgggcccgg catccagcag 2400
gagttcgga tccctacaa ccccgagtc caggcggtgg tggagtccat gaacaaggag 2460
ctgaagaaga tcacggcca ggtgcgcgac caggccgagc acctcaagac cgccgtgcag 2520
atggccgtgt tcaccacaa cttcaagcg aaggcgggca tcggcggtta ctccgcccgc 2580

gagcgcatca tcgacatcat cgccaccgac atccagacca aggagctgca gaagcagatc 2640
atcaagatcc agaacttccg cgtgtactac cgcgactccc gcgaccccat ctggaagggc 2700
cccgccaagc tgctgtggaa gggcgagggc gccgtggtga tccaggacaa ctccgacatc 2760
aagtggtgc cccgcgcgaa ggccaagatc atcaaggact acggcaagca gatggccggc 2820
gccgactgcg tggccggccg ccaggacgag gactaa 2856

<210> 21
<211> 497
<212> PRT
<213> Human immunodeficiency virus
<400> 21

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

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

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp

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

340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365

 Gln Val Thr Asn Ala Ala Ile Met Met Gln Arg Gly Asn Phe Lys Gly
 370 375 380
 Gln Arg Arg Ile Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile
 385 390 395 400
 Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly
 405 410 415
 Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe
 420 425 430
 Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe Leu

435 440 445
 Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Phe
 450 455 460
 Gly Glu Glu Ile Thr Pro Ser Pro Lys Gln Glu Pro Lys Asp Lys Glu
 465 470 475 480
 Pro Pro Leu Thr Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu Ser
 485 490 495
 Gln

<210> 22
 <211> 948
 <212> PRT
 <213> Human immunodeficiency virus

<400> 22
 Met Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Thr Ile Lys Ile
 1 5 10 15
 Gly Gly Gln Leu Lys Glu Ala Leu Leu Ala Thr Gly Ala Asp Asp Thr
 20 25 30
 Val Leu Glu Glu Ile Asn Leu Pro Gly Lys Trp Lys Pro Lys Met Ile
 35 40 45

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

290 295 300
 Leu Arg Glu His Leu Leu Arg Trp Gly Phe Thr Thr Pro Asp Lys Lys
 305 310 315 320
 His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro
 325 330 335
 Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp Ser Trp
 340 345 350
 Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser

 355 360 365
 Gln Ile Tyr Pro Gly Ile Lys Val Lys Gln Leu Cys Lys Leu Leu Arg
 370 375 380
 Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu Ala Glu
 385 390 395 400
 Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly
 405 410 415
 Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln
 420 425 430

 Gly Gln Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn
 435 440 445
 Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Ser Ala His Thr Asn Asp
 450 455 460
 Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Thr Glu Ser Ile
 465 470 475 480
 Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln Lys Glu
 485 490 495
 Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro

 500 505 510
 Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln
 515 520 525
 Leu Glu Lys Glu Pro Ile Ala Gly Ala Glu Thr Phe Tyr Val Asp Gly
 530 535 540
 Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asp

545 550 555 560
 Arg Gly Arg Gln Lys Val Val Ser Leu Thr Glu Thr Thr Asn Gln Lys
 565 570 575

 Thr Glu Leu Gln Ala Ile His Leu Ala Leu Gln Asp Ser Gly Ser Glu
 580 585 590
 Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala
 595 600 605
 Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln
 610 615 620
 Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser Trp Val Pro Ala His Lys
 625 630 635 640
 Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Thr Gly Ile

 645 650 655
 Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His
 660 665 670
 Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu
 675 680 685
 Pro Pro Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln
 690 695 700
 Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile
 705 710 715 720

 Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Ile Ile Leu Val Ala
 725 730 735
 Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu
 740 745 750
 Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu Lys Leu Ala Gly Arg Trp
 755 760 765
 Pro Val Lys Val Ile His Thr Asp Asn Gly Ser Asn Phe Thr Ser Ala
 770 775 780
 Ala Val Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln Glu Phe Gly

 785 790 795 800

Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Met Asn Lys
805 810 815
Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu
820 825 830
Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys
835 840 845
Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile Asp Ile Ile
850 855 860

Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile
865 870 875 880
Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Ile Trp Lys
885 890 895
Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln
900 905 910
Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile
915 920 925
Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Gly Arg

930 935 940
Gln Asp Glu Asp
945
<210> 23
<211> 207
<212> PRT
<213> Human immunodeficiency virus
<400> 23

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

Ala Ala Asn Asn Pro Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu

50 55 60
 Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
 65 70 75 80
 Thr Tyr Lys Ala Ala Leu Asp Leu Ser His Phe Leu Lys Glu Lys Gly
 85 90 95
 Gly Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp
 100 105 110
 Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr

 115 120 125
 Thr Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe
 130 135 140
 Lys Leu Val Pro Val Asp Pro Glu Glu Val Glu Glu Ala Asn Glu Gly
 145 150 155 160
 Glu Asn Asn Ser Leu Leu His Pro Met Cys Gln His Gly Met Glu Asp
 165 170 175
 Glu Glu Arg Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Leu
 180 185 190

 Arg His Ile Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys
 195 200 205
 <210> 24
 <211> 948
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 24
 Met Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Ser Ile Lys Val
 1 5 10 15
 Gly Gly Gln Ile Lys Glu Ala Leu Leu Ala Thr Gly Ala Asp Asp Thr
 20 25 30
 Val Leu Glu Glu Ile Asn Leu Pro Gly Lys Trp Lys Pro Lys Met Ile

 35 40 45
 Gly Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Asp Gln Ile Leu
 50 55 60

Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly Thr Val Leu Val Gly Pro
65 70 75 80
Thr Pro Val Asn Ile Ile Gly Arg Asn Met Leu Thr Gln Leu Gly Cys
85 90 95
Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu
100 105 110
Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu
115 120 125
Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu Lys Glu
130 135 140
Gly Lys Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val
145 150 155 160
Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp
165 170 175
Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu
180 185 190
Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val
195 200 205
Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Gly Phe
210 215 220
Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro
225 230 235 240
Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser
245 250 255
Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg
260 265 270
Ala Gln Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr
275 280 285
Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile Glu Glu
290 295 300
Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp Lys Lys
305 310 315 320

His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro

325 330 335

Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp Ser Trp

340 345 350

Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser

355 360 365

Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg

370 375 380

Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu Ala Glu

385 390 395 400

Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly

405 410 415

Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln

420 425 430

Gly His Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn

435 440 445

Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr Asn Asp

450 455 460

Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met Glu Ser Ile

465 470 475 480

Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln Lys Glu

485 490 495

Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp Ile Pro

500 505 510

Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln

515 520 525

Leu Glu Lys Glu Pro Ile Ala Gly Ala Glu Thr Phe Tyr Val Asp Gly

530 535 540

Ala Ala Asn Arg Glu Thr Lys Ile Gly Lys Ala Gly Tyr Val Thr Asp

545 550 555 560

Arg Gly Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn Gln Lys

565 570 575
 Thr Glu Leu Gln Ala Ile Gln Leu Ala Leu Gln Asp Ser Gly Ser Glu
 580 585 590
 Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala
 595 600 605
 Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln

 610 615 620
 Leu Ile Lys Lys Glu Arg Val Tyr Leu Ser Trp Val Pro Ala His Lys
 625 630 635 640
 Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ser Gly Ile
 645 650 655
 Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His
 660 665 670
 Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Glu Phe Asn Leu
 675 680 685

 Pro Pro Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln
 690 695 700
 Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile
 705 710 715 720
 Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Ile Ile Leu Val Ala
 725 730 735
 Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu
 740 745 750
 Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu Lys Leu Ala Gly Arg Trp

 755 760 765
 Pro Val Lys Val Ile His Thr Asp Asn Gly Ser Asn Phe Thr Ser Ala
 770 775 780
 Ala Val Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln Glu Phe Gly
 785 790 795 800
 Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Met Asn Lys
 805 810 815
 Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu

820	825	830
Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys		
835	840	845
Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile Asp Ile Ile		
850	855	860
Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Ile Lys Ile		
865	870	875
Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Ile Trp Lys		
885	890	895
Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln		
900	905	910
Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile		
915	920	925
Lys Asp Tyr Gly Lys Gln Met Ala Gly Ala Asp Cys Val Ala Gly Arg		
930	935	940
Gln Asp Glu Asp		
945		
<210>	25	
<211>	1512	
<212>	DNA	
<213>	Human immunodeficiency virus	
<400>	25	
gccgccgcca tgggcgcccc cgcctccgtg ctgtccggcg gcgagctgga cgcctgggag	60	
aagatccgcc tgcgccccgg cggcaagaag aagtacaagc tgaagcacat cgtgtgggcc	120	
tcccgcgagc tggagcgctt cgccgtgaac cccggcctgc tggagacctc cgagggtgc	180	
cgccagatcc tgggccagct gcagccctcc ctgcagaccg gctccgagga gctgcgctcc	240	
ctgtacaaca ccgtggccac cctgtactgc gtgcaccagc gcatcgaggt gaaggacacc	300	
aaggaggccc tggagaagat cgaggaggag cagaacaagt ccaagaagaa ggcccagcag	360	
gccgccgccc acaccggcaa ctctcccag gtgtccaga actaccccat cgtgcagaac	420	
ctgcagggcc agatggtgca ccaggccatc tcccccgca cctgaacgc ctgggtgaag	480	
gtggtggagg agaaggcctt ctccccgag gtgatccca tgttctccgc cctgtccgag	540	

ggcgccaccc cccaggacct gaacaccatg ctgaacaccg tgggcggcca ccaggccgcc 600
atgcagatgc tgaaggagac catcaacgag gaggccgccg agtgggaccg cctgcacccc 660
gtgcacgccg gccccatcgc ccccgccag atgcgcgagc cccgcggctc cgacatcgcc 720
ggcaccacct ccacctgca ggagcagatc ggctggatga ccaacaaccc ccccatcccc 780
gtgggcgaga tctacaagcg ctggatcatc ctgggcctga acaagatcgt gcgcatttac 840
tccccacct ccatcctgga catcgccag ggccccaagg agcccttcg cgactacgtg 900
gaccgttctt acaagacct gcgcgccgag caggcctccc aggaggtgaa gaactggatg 960

accgagaccc tgctggtgca gaacccaac cccgactgca agaccatcct gaaggccctg 1020
ggccccgccg ccacctgga ggagatgatg accgcctgcc agggcgtggg cggccccggc 1080
cacaaggccc gcgtgctggc cgaggccatg tcccaggtga ccaactccgc caccatcatg 1140
atgcagcgcg gcaacttcg caaccagcg aagaccgtga agtgcttcaa ctgcggcaag 1200
gagggccaca tcgccaagaa ctgccgcgcc ccccgcaaga agggctgctg gaagtgcggc 1260
aaggagggcc accagatgaa ggactgcacc gagcgccagg ccaacttcct gggcaagatc 1320
tggccctccc acaagggccg ccccgccaac ttctgcagt cccgccccga gccaccgcc 1380

cccccgagg agtccttcg cttcggcgag gagaccacca cccctccca gaagcaggag 1440
cccatcgaca aggagctgta cccctggcc tccctgcgt cctgttcgg caacgacccc 1500
tcctccagt aa 1512

<210> 26
<211> 2562
<212> DNA
<213> Human immunodeficiency virus
<400> 26

gccgccgcca tgcgcgtgaa gggcatccgc aagaactacc agcacctgtg gcgctggggc 60
accatgctgc tgggcatgct gatgatctgc tccgccgccg agaagctgtg ggtgaccgtg 120
tactacggcg tgcccgtgtg gaaggaggcc accaccacc tgttctgcgc ctccgacgcc 180

aaggcctacg acaccgaggt gcacaacgtg tgggccaccc acgcctgcgt gccaccgac 240
cccaaccccc aggaggtggt gctggagaac gtgaccgaga acttcaacat gtggaagaac 300
aacatggtgg agcagatgca cgaggacatc atctccctgt gggaccagtc cctgaagccc 360
tgcgtgaagc tgacccccct gtgcgtgacc ctgaactgca ccgacctgaa gaacaacctg 420
ctgaacacca actcctctc cggcgagaag atggagaagg gcgagatcaa gaactgctcc 480
ttcaacatca ccacctccat ccgcgacaag gtgcagaagg agtacgcct gtcttacaag 540

ctggacgtgg tgcccatcga caacaacaac aacacctcct accgcctgat ctctgcaac	600
acctccgtga tcaccaggc ctgccccaaag gtgtccttcg agcccatccc catccactac	660
tgcccccccg cgggcttcgc catcctgaag tgcaacgaca agaagttcaa cggcacccggc	720
ccctgcacca acgtgtccac cgtgcagtgc acccacggca tccgccccgt ggtgtccacc	780
cagctgctgc tgaacggctc cctggccgag gaggaggtgg tgatccgctc cgagaacttc	840
accgacaacg ccaagaccat catcgtgcag ctgaacgagt ccgtggagat caactgcacc	900
cgccccaaaca acaacacccg caagtccatc cacatcggcc cggcccgcg cttctacacc	960
accggcgaga tcacggcgga catccgccag gccactgca acatctcccg cgccaagtgg	1020
aacaacaccc tgaagcagat cgtgaagaag ctgcgcgagc agttcggcaa caagaccatc	1080
gigttaacc agtcctccgg cggcgacccc gagatcgtga tgactcctt caactgcggc	1140
ggcgagtctt tctactgaa caccacccag ctgttcaact ccacctggaa cgacaacggc	1200
acctggaaca acaccaagga caagaacacc atcacctgc cctgccgcat caagcagatc	1260
atcaacatgt ggcaggaggt gggcaaggcc atgtacgcc ccccatccg cggccagatc	1320
cgtgctcct ccaacatcac cggcctgctg ctgaccgcg acggcggcaa caacaacaac	1380
gacaccgaga tcttccgcc cggcggcggc gacatgcgcg acaactggcg ctccgagctg	1440
tacaagtaca aggtggtgaa gatcgagccc ctgggcgtgg cccccacaa ggccaagcgc	1500
cgcgtggtgc agcgcgagaa gcgcgccgtg ggcatcggcg ccatgttctt gggttctctg	1560
ggcgccgccc gctccaccat gggcgccgcc tccatgacc tgaccgtgca ggcccgcag	1620
ctgctgtccg gcacgtgca gcagcagaac aacctgtgc gcgccatga ggcccagcag	1680
cacctgtgc agctgaccgt gtggggcatc aagcagctgc agggccgct gctggccgtg	1740
gagcgtacc tgaaggacca gcagctgtg ggcatctggg gctgtccgg caagctgac	1800
tgcaccacca ccgtgcctg gaacgcctcc tggccaaca agtcctgga cgagatctgg	1860
gacaacatga cctggatgga gtgggagcgc gagatcgaca actacacctc cctgatctac	1920
acctgatcg aggagtccca gaaccagcag gagaagaac agcaggagct gctggagctg	1980
gacaagtggg cctccctgtg gaactgggtt gacatcacca actggctgtg gtacatcaag	2040
atcttcatca tgatcgtggg cggcctgatc ggctgcgca tcgtgttcgc cgtgctgtcc	2100
atcgtgaacc gcgtgcgcca gggctactcc ccctgtcct tccagaccg cctgcccgc	2160
cccccgggc ccgaccgcc cgagggcac gaggaggagg gcgcgagcg cgaccgcgac	2220
cgtccggcc gcctggtgga cggttctctg gcctgatct gggacgacct gcgtccctg	2280

tgcctgttct cctaccaccg cctgcgcgac ctgctgctga tcgtgaccgc catcgtggag 2340
 ctgctggggc gccgcggctg ggaggtgctg aagtactggt ggaacctgct gcagtactgg 2400
 tcccaggagc tgaagaactc cgccgtgtcc ctgctgaacg ccaccgccat cgccgtggcc 2460
 gagggcaccg accgcgtgat cgaggtggtg cagcgcgcct gccgcgccat cctgcacatc 2520
 cccgcgcga tccgccaggg cctggagcgc gcctgctgt aa 2562

<210> 27

<400> 27

000

<210> 28

<211> 500

<212> PRT

<213> Human immunodeficiency virus

<400> 28

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys

20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys

100 105 110

Lys Lys Ala Gln Gln Ala Ala Asp Thr Gly Asn Ser Ser Gln Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His

130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala

 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285

 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser

 355 360 365
 Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg
 370 375 380
 Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
 385 390 395 400
 Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys

405 410 415
 Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
 420 425 430

 Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe
 435 440 445
 Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
 450 455 460
 Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
 465 470 475 480
 Lys Glu Leu Tyr Pro Leu Ala Ser Leu Arg Ser Leu Phe Gly Asn Asp
 485 490 495
 Pro Ser Ser Gln

500
 <210> 29
 <211> 850
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 29
 Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
 1 5 10 15
 Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Lys
 20 25 30
 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
 35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

 50 55 60
 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80
 Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
 85 90 95
 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asp Leu Lys Asn Asn Leu Leu Asn Thr Asn Ser Ser Ser
130 135 140

Gly Glu Lys Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile
145 150 155 160

Thr Thr Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Leu Phe Tyr
165 170 175

Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Asn Thr Ser Tyr Arg
180 185 190

Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val
195 200 205

Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala
210 215 220

Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Thr
225 230 235 240

Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser
245 250 255

Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile
260 265 270

Arg Ser Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu
275 280 285

Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg
290 295 300

Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly Glu
305 310 315 320

Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys
325 330 335

Trp Asn Asn Thr Leu Lys Gln Ile Val Lys Lys Leu Arg Glu Gln Phe
340 345 350

Gly Asn Lys Thr Ile Val Phe Asn Gln Ser Ser Gly Gly Asp Pro Glu

355 360 365
 Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn
 370 375 380
 Thr Thr Gln Leu Phe Asn Ser Thr Trp Asn Asp Asn Gly Thr Trp Asn
 385 390 395 400
 Asn Thr Lys Asp Lys Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln
 405 410 415

 Ile Ile Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro
 420 425 430
 Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu
 435 440 445
 Thr Arg Asp Gly Gly Asn Asn Asn Asn Asp Thr Glu Ile Phe Arg Pro
 450 455 460
 Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
 465 470 475 480
 Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys

 485 490 495
 Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Met
 500 505 510
 Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser
 515 520 525
 Met Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln
 530 535 540
 Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu
 545 550 555 560

 Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala
 565 570 575
 Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys
 580 585 590
 Ser Gly Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp
 595 600 605
 Ser Asn Lys Ser Leu Asp Glu Ile Trp Asp Asn Met Thr Trp Met Glu

610 615 620
 Trp Glu Arg Glu Ile Asp Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile

 625 630 635 640
 Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu

 645 650 655
 Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp

 660 665 670
 Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly

 675 680 685
 Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln

 690 695 700

 Gly Tyr Ser Pro Leu Ser Phe Gln Thr Arg Leu Pro Ala Pro Arg Gly
 705 710 715 720
 Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg

 725 730 735
 Asp Arg Ser Gly Arg Leu Val Asp Gly Phe Leu Ala Leu Ile Trp Asp

 740 745 750
 Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu

 755 760 765
 Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp

 770 775 780
 Glu Val Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr Trp Ser Gln Glu
 785 790 795 800
 Leu Lys Asn Ser Ala Val Ser Leu Leu Asn Ala Thr Ala Ile Ala Val

 805 810 815
 Ala Glu Gly Thr Asp Arg Val Ile Glu Val Val Gln Arg Ala Cys Arg

 820 825 830
 Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala

 835 840 845

 Leu Leu
 850

<210> 30
 <211> 610
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 30
 Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
 1 5 10 15
 Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn
 20 25 30
 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn
 35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50 55 60
 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80
 Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
 85 90 95
 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asn Val Asn Val Thr Asn Thr Thr Asn Asn Thr Glu Glu
 130 135 140
 Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg
 145 150 155 160
 Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val
 165 170 175
 Pro Ile Asp Asp Asn Asn Asn Asn Ser Ser Asn Tyr Arg Leu Ile Asn
 180 185 190
 Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu
 195 200 205
 Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys

210 215 220
 Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser
 225 230 235 240
 Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu
 245 250 255
 Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu
 260 265 270

 Asn Ile Thr Asn Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser
 275 280 285
 Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
 290 295 300
 Arg Ile Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly
 305 310 315 320
 Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Thr Lys Trp Asn Lys
 325 330 335
 Thr Leu Gln Gln Val Ala Lys Lys Leu Arg Glu His Phe Asn Asn Lys

 340 345 350
 Thr Ile Ile Phe Lys Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr
 355 360 365
 His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly
 370 375 380
 Leu Phe Asn Ser Thr Trp Ile Gly Asn Gly Thr Lys Asn Asn Asn Asn
 385 390 395 400
 Thr Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn
 405 410 415

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

465 470 475 480

Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Leu Thr

Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn

500 505 510

Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val

515 520 525

Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr

530 535 540
Leu Lys Asp Gln Gln Leu Glu Ile Trp Asp Asn Met Thr Trp Met Glu
545 550 555 560

Trp Glu Arg Glu Ile Asn Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile
565 570 575

Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala
580 585 590

Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp
595 600 605

Leu Trp

610

<210> 31

<211> 1837

<212> DNA

<213> Human immunodeficiency virus

<400> 31

atgcgcgtgc gcggcattcca gcgcaactgc cagcacctgt ggcgctggggg cacctgatcc 60

tgggcatgct gatgatctgc tccgccgccg agaacctgtg ggtgaccgtg tactacggcg 120

tgcccggtgtg gaaggaggcc aacaccacc tgtttctgcgc ctccgacgcc aaggcctacg 180

acaccgaggt gcacaacgtg tgggccaccc acgcctgcgt gccaccgac cccaaccccc 240

aggagatcgt gctggagaac gtgaccgaga acttcaacat gtggaagaac aacatggtgg 300

agcagatgca cgaggacatc atctccctgt gggaccagtc cctgaagccc tgcgtgaagc 360

tgacccccct gtgcgtgacc ctgaactgca ccaacgtgaa cgtgaccaac accaccaaca 420

acaccgagga gaagggcgag atcaagaact gtccttcaa catcaccacc gagatccgcg 480

acaagaagca gaaggtgtac gccctgttct accgcctgga cgtggtgccc atcgacgaca 540

acaacaacaa ctctccaac taccgcctga tcaactgcaa cacctccgcc atcaccagg 600

cctgccccaa gggttccttc gagcccatcc ccatccacta ctgcgcccc gccggcttcg 660

ccatcctgaa gtgaacgac aagaagttca acggcaccgg cccctgcaag aacgtgtcca 720

ccgtgcagtg caccacggc atcaagcccg tgggtgtccac ccagtgtctg ctgaacggct 780

ccctggccga ggaggagatc tcatccgtc cgagaacatc accaacaacg ccaagaccat 840

catcgtgcag ctgaacgagt ccgtggagat caactgcacc cgccccaaca acaacaccg 900

caagtccatc cgcacggcc ccggccaggc ctctacgcc accggcgaca tcatcggcga 960

catccgccag gccactgca acatctccgg caccaagtgg aacaagacc tgcagcaggt 1020

ggccaagaag ctgcgcgagc atttcaacaa caagaccatc atcttcaagc ctctctccg 1080

cggcgacctg gagatcacca cccactcctt caactgccgc ggcgagttct tctactgcaa 1140

cacctccggc ctgttcaact ccacctggat cggcaacggc accaagaaca acaacaacac 1200

caacgacacc atcacctgc cctgccgat caagcagatc atcaacatgt ggcagggcgt 1260

gggccaggcc atgtacgcc ccccatcga gggcaagatc acctgcaagt ccaacatcac 1320

cggcctgctg ctgaccgcg acggcggcaa caacaacacc aacgagacc agatcttccg 1380

ccccggcggc ggcgacatgc gcgacaactg gcgctccgag ctgtacaagt acaagtggt 1440

gaagatcgag cccctgggag tggccccac caaggccaag cttaccgtgc aggcccga 1500

gtgtgtgtcc ggcatcgtgc agcagcagtc caacctgctg cgcgccatcg aggcccagca 1560

gcacctgctg cagctgaccg tgtggggcat caagcagctg caggcccgcg tgctggccgt 1620

ggagcgtac ctgaaggacc agcagctcga gatctgggac aacatgacct ggatggagt 1680

ggagcgcgag atcaacaact acaccgacat catctactcc ctgatcgagg agtcccagaa 1740

ccagcaggag aagaacgagc aggagctgct ggccctggac aagtgggcct ccctgtggaa 1800

ctggttcgac atcaccaact ggctgtggtg aggatec 1837

<210> 32

<211> 841

<212> PRT

<213> Human immunodeficiency virus

<400> 32

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp

1 5 10 15

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

260 265 270
 Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser
 275 280 285
 Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
 290 295 300
 His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly Glu Ile Ile Gly
 305 310 315 320
 Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Asn
 325 330 335
 Thr Leu Lys Gln Ile Val Lys Lys Leu Arg Glu Gln Phe Gly Asn Lys
 340 345 350
 Thr Ile Val Phe Asn Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met
 355 360 365
 His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln
 370 375 380
 Leu Phe Asn Ser Thr Trp Asn Gly Thr Trp Asn Asn Thr Glu Gly Asn
 385 390 395 400
 Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu
 405 410 415
 Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys
 420 425 430
 Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn
 435 440 445
 Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
 450 455 460
 Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly
 465 470 475 480
 Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg
 485 490 495
 Ala Val Gly Ile Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly
 500 505 510
 Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln

515 520 525
 Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile
 530 535 540

 Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
 545 550 555 560
 Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln
 565 570 575
 Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala
 580 585 590
 Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp
 595 600 605
 Asp Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr

 610 615 620
 Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
 625 630 635 640
 Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
 645 650 655
 Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met
 660 665 670
 Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser
 675 680 685

 Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr
 690 695 700
 Arg Leu Pro Ala Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu
 705 710 715 720
 Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser Gly Arg Leu Val Asp Gly
 725 730 735
 Phe Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser
 740 745 750
 Tyr His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu

 755 760 765

Leu Leu Gly Arg Arg Gly Trp Glu Val Leu Lys Tyr Trp Trp Asn Leu
770 775 780
Leu Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu
785 790 795 800
Asn Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu
805 810 815
Val Val Gln Arg Ala Cys Arg Ala Ile Leu His Ile Pro Arg Arg Ile
820 825 830

Arg Gln Gly Leu Glu Arg Ala Leu Leu
835 840

<210> 33
<211> 632
<212> PRT
<213> Human immunodeficiency virus
<400> 33

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

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
50 55 60
His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80
Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95
Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110
Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asp Leu Met Asn Ala Thr Asn Thr Asn Thr Thr Ile Ile

130 135 140
 Tyr Arg Trp Arg Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr
 145 150 155 160
 Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu
 165 170 175
 Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Arg Leu Ile Ser
 180 185 190

 Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu
 195 200 205
 Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys
 210 215 220
 Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser
 225 230 235 240
 Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu
 245 250 255
 Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Glu

 260 265 270
 Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser
 275 280 285
 Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
 290 295 300
 His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly Glu Ile Ile Gly
 305 310 315 320
 Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Asn
 325 330 335

 Thr Leu Lys Gln Ile Val Lys Lys Leu Arg Glu Gln Phe Gly Asn Lys
 340 345 350
 Thr Ile Val Phe Asn Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met
 355 360 365
 His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln
 370 375 380
 Leu Phe Asn Ser Thr Trp Asn Gly Thr Trp Asn Asn Thr Glu Gly Asn

385 390 395 400
 Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu

 405 410 415
 Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys
 420 425 430
 Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn
 435 440 445
 Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
 450 455 460
 Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly
 465 470 475 480

 Val Ala Pro Thr Lys Ala Lys Thr Leu Thr Val Gln Ala Arg Gln Leu
 485 490 495
 Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu
 500 505 510
 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
 515 520 525
 Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu
 530 535 540
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val

 545 550 555 560
 Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asp
 565 570 575
 Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr Ser
 580 585 590
 Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
 595 600 605
 Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
 610 615 620

 Phe Asp Ile Thr Asn Trp Leu Trp
 625 630

<210> 34
 <211> 1927
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 34

ttcagtcgac ggccaccatg aggggtgaagg gtattcggaa aaattaccaa cacctgtggc 60
 gctggggaac catgctcctt ggtatgctga tgatttgag tgccgccgag aaactttggg 120
 taactgtgta ctacggcggtt cctgtctgga aggaagctac aaccactctt ttttgtgcat 180
 ccgacgctaa agcttacgac acagaagtgc ataatgtttg ggccacccat gcttgcgtcc 240
 ctacagatcc caacccccag gaagtcgtcc ttgagaatgt cacagagaat tttaacatgt 300

ggaagaataa tatggtagaa caaatgcacg aagacattat tagcctgtgg gaccagtctt 360
 tgaagccctg cgigaaactc actccacttt gcgtcacact taactgtact gatttgatga 420
 acgcaaccaa cacaatact actattatat atcgctggag gggggaaatc aagaactgct 480
 ctttcaacat caccacttcc ataaggata aggtccagaa agaatagcc ctgttttata 540
 aacttgatgt ggicccgata gacaatgaca aactagcta tcgactgac tcttgaaca 600
 catccgtgat tacccaagct tgcccaaagg tcagctttga accaataccc attcactact 660
 gcgctccgc tggttttgcc atcctcaagt gtaacgacaa aaaattcaat gggaccggac 720

cttgcacaaa cgtgtccacc gtgcaatgta ctcacggaat cagacctgtt gtcagtaccc 780
 aactcctctt gaacgggtct ctgcggaag aggaggtcgt gattagaagc gaaaacttta 840
 ccgataacgc taaaacaatc attgtgcaac ttaatgaaag cgtcgaaatt aactgcacca 900
 gaccaacaa taataccaga aaatctattc acatagggcc cggccgcgca tttatacaa 960
 ctggcgaaat cattggtgac atcagacaag ctattgcaa tatctccgc gcgaaatgga 1020
 acaacaccct gaaacagatc gtgaagaaac ttcgagaaca attcggtaat aaaacaatcg 1080
 tattcaacca aagctccgga ggcgaccctg agatagtat gcactcattc aactgtggcg 1140

gcgagttctt ctattgtaac acaactcaac tttttaatag cacttggaat ggaacatgga 1200
 acaacacaga agggaacatc actctgcctt gtcggattaa gcagatcatt aatatgtggc 1260
 aagaagtggg aaaagctatg tacgccccgc ctattcgcgg acaataaga tgctctagta 1320
 atattaccgg attgttgctg acacgcgacg gaggaataa tgaaacagag atatttagac 1380
 ctggcggagg cgacatgaga gataactgga gaagttagct ttacaaatat aaagtcgtaa 1440
 agatagaacc attgggggta gcaccaacca aagcaaaaac cttgacagta caggctaggc 1500
 agctgctgag cggaatcgtg caacaacaaa ataatcttct ccgagccata gaagcacaac 1560

aacatctgtt gcagctgaca gtatggggaa tcaaacagct tcaggcaaga gtgctggccg 1620
 tcgagagata cctcaaagat caacaactgc tgggcatatg gggatgttcc ggtaaactca 1680
 tatgcactac cgccgtgccc tggaacgcga gctggtctaa taaatccctg gatgaaattt 1740
 gggacaacat gacttggatg gaatgggaac gggaaattga caactatact agtttgattt 1800
 atactctgat cgaagaatct cagaaccaac aggagaaaaa cgaacaggaa ctgctggaac 1860
 tggacaagtg ggcatcattg tggaactggt ttgacattac taactggctg tggtaaagat 1920
 cttacaa 1927

<210> 35

<211> 829

<212> PRT

<213> Human immunodeficiency virus

<400> 35

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Ile Leu Ile Phe Gly Met Leu Ile Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Val Asn Ala Thr Asn Asn Thr Thr Asn Asn Glu Glu

130 135 140

Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys

145 150 155 160

Lys Lys Val Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp
 165 170 175
 Asp Asn Asn Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr
 180 185 190
 Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys
 195 200 205
 Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn
 210 215 220
 Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly
 225 230 235 240
 Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala
 245 250 255
 Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile Thr Asn Asn Ala Lys
 260 265 270
 Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg
 275 280 285
 Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Ala
 290 295 300
 Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys
 305 310 315 320
 Asn Ile Ser Arg Thr Lys Trp Asn Lys Thr Leu Gln Gln Val Ala Lys
 325 330 335
 Lys Leu Arg Glu His Phe Asn Lys Thr Ile Ile Phe Asn Pro Ser Ser
 340 345 350
 Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu
 355 360 365
 Phe Phe Tyr Cys Asn Thr Ser Glu Leu Phe Asn Ser Thr Trp Asn Gly
 370 375 380
 Thr Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn
 385 390 395 400
 Met Trp Gln Gly Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly

405										410					415									
Lys	Ile	Arg	Cys	Thr	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp									
420										425					430									
Gly	Gly	Asn	Asn	Asn	Thr	Glu	Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met									
435										440					445									
Arg	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile									
450										455					460									
Glu	Pro	Leu	Gly	Val	Ala	Pro	Thr	Lys	Ala	Lys	Arg	Arg	Val	Val	Glu									
465										470					475					480				
Arg	Glu	Lys	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu									
485										490					495									
Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val									
500										505					510									
Gln	Ala	Arg	Gln	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu									
515										520					525									
Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp									
530										535					540									
Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	Val	Glu	Arg	Tyr	Leu									
545										550					555					560				
Lys	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile									
565										570					575									
Cys	Thr	Thr	Asn	Val	Pro	Trp	Asn	Ser	Ser	Trp	Ser	Asn	Lys	Ser	Gln									
580										585					590									
Asp	Glu	Ile	Trp	Asp	Asn	Met	Thr	Trp	Met	Glu	Trp	Asp	Lys	Glu	Ile									
595										600					605									
Asn	Asn	Tyr	Thr	Asp	Ile	Ile	Tyr	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn									
610										615					620									
Gln	Gln	Glu	Lys	Asn	Glu	Gln	Glu	Leu	Leu	Ala	Leu	Asp	Lys	Trp	Ala									
625										630					635					640				
Ser	Leu	Trp	Asn	Trp	Phe	Asp	Ile	Thr	Asn	Trp	Leu	Trp	Tyr	Ile	Lys									
645										650					655									

Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe
660 665 670

Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu
675 680 685

Ser Phe Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Asp Arg Pro Glu
690 695 700

Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile Arg
705 710 715 720

Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu
725 730 735

Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu Ile Ala Ala
740 745 750

Arg Thr Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr
755 760 765

Leu Trp Asn Leu Leu Gln Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala
770 775 780

Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val Ala Glu Gly Thr Asp
785 790 795 800

Arg Val Ile Glu Val Val Gln Arg Val Cys Arg Ala Ile Leu Asn Ile
805 810 815

Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu
820 825

<210> 36
<211> 620
<212> PRT
<213> Human immunodeficiency virus
<400> 36

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
1 5 10 15

Gly Ile Leu Ile Phe Gly Met Leu Ile Ile Cys Ser Ala Ala Glu Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn

35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50 55 60
 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

 65 70 75 80
 Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
 85 90 95
 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asp Val Asn Ala Thr Asn Asn Thr Thr Asn Asn Glu Glu
 130 135 140

 Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys
 145 150 155 160
 Lys Lys Val Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp
 165 170 175
 Asp Asn Asn Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr
 180 185 190
 Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys
 195 200 205
 Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn

 210 215 220
 Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly
 225 230 235 240
 Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala
 245 250 255
 Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile Thr Asn Asn Ala Lys
 260 265 270
 Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg
 275 280 285

Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Ala
290 295 300

Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys
305 310 315 320

Asn Ile Ser Arg Thr Lys Trp Asn Lys Thr Leu Gln Gln Val Ala Lys
325 330 335

Lys Leu Arg Glu His Phe Asn Lys Thr Ile Ile Phe Asn Pro Ser Ser
340 345 350

Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu
355 360 365

Phe Phe Tyr Cys Asn Thr Ser Glu Leu Phe Asn Ser Thr Trp Asn Gly
370 375 380

Thr Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn
385 390 395 400

Met Trp Gln Gly Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly
405 410 415

Lys Ile Arg Cys Thr Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp
420 425 430

Gly Gly Asn Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met
435 440 445

Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
450 455 460

Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Thr Leu Thr Val Gln
465 470 475 480

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
485 490 495

Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly
500 505 510

Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys
515 520 525

Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys
530 535 540

Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Asp
 545 550 555 560
 Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Asp Lys Glu Ile Asn
 565 570 575

Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln
 580 585 590
 Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp Lys Trp Ala Ser
 595 600 605
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp
 610 615 620

<210> 37
 <211> 1891
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 37

ttcagtcgac agccaccatg cgggtcatgg ggatacagag gaattgccag cacttgigga 60

ggtaggggaat ttgatattc gggatgctca taatctgctc tgccgctgag aacctgtggg 120
 tcactgtgta ttacggcggt cccgtctgga aagaagctaa tactaccctg ttttgtgcaa 180
 gcgacgcaa agcatacgac accgaagtcc acaatgtctg ggctaccac gcctgtgtac 240
 ctactgatcc aaatccccag gaaattgttc ttgaaaacgt aacggaaaac tttaatatgt 300
 ggaagaataa tatgggtggag caaatgcacg aggatataat cagcctgtgg gaccagtccc 360
 tcaaaccatg cgttaaactc actccactct gcgtgactct gaactgtacc gacgtgaacg 420
 caaccaataa tacaacaac aatgaggaga taaagaattg ttcatttaat ataaccactg 480

agatacggga taagaaaaaa aaggtttatg cactctttta caagctcgac gtggtgcca 540
 tagacgaaa taatagctac cgactcatta attgcaatac tagcgctata acccaggcat 600
 gccccaaagt ttccttcgag cccataccga ttactactg cgcacccgcc ggattcgcca 660
 ttcttaaatg caatgacaag aagttcaacg gcaccggacc ctgtaagaac gtaagcactg 720
 ttcaatgtac acatggaatt aagccggtag tgtcaacgca gctcctctc aacggaagcc 780
 ttgcagaaga agagatcatt atcaggtcag aaaatatcac taacaacgcg aaaacaatca 840
 ttgttcagct gaatgagtct gtagaaatca attgtaccgc ccctaataat aacacaagaa 900

agtcaattag gatcggaacc ggccaggcct tctacgcaac cggagatatc atcggggata 960

tacgacaggc cactgcaac atttctagaa ctaagtggaa taaaactttg cagcaggtag 1020
ccaagaaact gcgggaacat ttttaataaga caatcatctt caatccaagt agcggagggg 1080
acctggaaat cactacacat tcctttaact gtggggcgga gtttttctac tgtaatacct 1140
ctgaactgtt caactcaaca tggaatggca ctaacaatac tataactctt ccttgagaa 1200
taaaacagat tatcaacatg tggcagggtg tggggcaagc aatgtatgca ccaccaatcg 1260
aaggcaaaat aagatgcacc tccaatatta cgggactcct cctgacacgg gatggcgga 1320

acaataacac ggagaccttt aggccaggcg gcggcgatat gagagataac tggcgctccg 1380
agctctataa atacaaagtc gttaagatcg agccccttgg agttgcgcca accaaagcta 1440
aaaccttgac cgtgcaagcc aggcagttgt tgcaggtat cgtacagcag caatctaate 1500
ttttgagagc cattgaggct cagcagcacc tcttcagct taccgtctgg ggcataaac 1560
aacttcaggc acgcgtcctg gccgtagagc gctatttgaa agaccaacaa ctctcggga 1620
tctgggggtg ttctggaaaa ttgatctgca cgacaaatgt gccttggaac agcagctggt 1680
caataaaaag ccaagacgaa atatgggata acatgacatg gatggaatgg gataaagaaa 1740

ttaataatta cactgacatt atttactcac ttatcgagga atcacaaaat caacaggaaa 1800
aaaatgaaca ggaactcttg gctctggaca aatgggcttc actgtggaac tggttcgaca 1860
tcacaaattg gctctggtaa agatcttaca a 1891

<210> 38

<211> 845

<212> PRT

<213> Human immunodeficiency virus

<400> 38

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Leu Arg Trp

1 5 10 15

Gly Thr Met Ile Leu Gly Met Ile Ile Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Thr Glu Met

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile His Leu Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys

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

Gln Lys Val Ala Lys Gln Leu Arg Lys Tyr Phe Lys Asn Lys Thr Ile
340 345 350

Ile Phe Thr Asn Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser
355 360 365

Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe
370 375 380

Asn Ser Thr Trp Asn Asn Gly Thr Met Lys Asn Thr Ile Thr Leu Pro
385 390 395 400

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Ala Gly Gln Ala
405 410 415

Met Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Glu Ser Asn Ile
420 425 430

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Asn Glu
435 440 445

Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu
450 455 460

Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro
465 470 475 480

Thr Arg Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly
485 490 495

Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
500 505 510

Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
515 520 525

Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln
530 535 540

Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
545 550 555 560

Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
565 570 575

Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp
580 585 590

Asn Ser Ser Trp Ser Asn Lys Ser Gln Asn Glu Ile Trp Asp Asn Met

595 600 605

Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr His Ile Ile

610 615 620

Tyr Asn Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln

625 630 635 640

Asp Leu Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp

645 650 655

Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly

660 665 670

Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Val Ile Asn

675 680 685

Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Thr Pro

690 695 700

Asn Pro Arg Gly Leu Asp Arg Pro Gly Arg Ile Glu Glu Glu Gly Gly

705 710 715 720

Glu Gln Gly Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala

725 730 735

Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg

740 745 750

Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly

755 760 765

His Ser Ser Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr

770 775 780

Leu Trp Asn Leu Leu Leu Tyr Trp Gly Arg Glu Leu Lys Ile Ser Ala

785 790 795 800

Ile Asn Leu Val Asp Thr Ile Ala Ile Ala Val Ala Gly Trp Thr Asp

805 810 815

Arg Val Ile Glu Ile Gly Gln Arg Ile Gly Arg Ala Ile Leu His Ile

820 825 830

Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu

	835				840				845						
<210>	39														
<211>	629														
<212>	PRT														
<213>	Human immunodeficiency virus														
<400>	39														
Met	Arg	Val	Met	Gly	Ile	Gln	Arg	Asn	Cys	Gln	His	Leu	Leu	Arg	Trp
1				5					10					15	
Gly	Thr	Met	Ile	Leu	Gly	Met	Ile	Ile	Ile	Cys	Ser	Ala	Ala	Glu	Asn
				20				25					30		
Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys	Asp	Ala	Glu
				35				40					45		
Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Glu	Thr	Glu	Met
	50						55					60			
His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro
65						70					75				80
Gln	Glu	Ile	His	Leu	Glu	Asn	Val	Thr	Glu	Glu	Phe	Asn	Met	Trp	Lys
						85					90				95
Asn	Asn	Met	Val	Glu	Gln	Met	His	Thr	Asp	Ile	Ile	Ser	Leu	Trp	Asp
				100				105					110		
Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu
				115				120					125		
Asn	Cys	Ser	Asn	Val	Asn	Val	Thr	Asn	Asn	Thr	Thr	Asn	Thr	His	Glu
	130						135					140			
Glu	Glu	Ile	Lys	Asn	Cys	Ser	Phe	Asn	Met	Thr	Thr	Glu	Leu	Arg	Asp
145						150					155				160
Lys	Lys	Gln	Lys	Val	Tyr	Ser	Leu	Phe	Tyr	Arg	Leu	Asp	Val	Val	Gln
					165					170					175
Ile	Asn	Glu	Asn	Asn	Ser	Asn	Ser	Ser	Tyr	Arg	Leu	Ile	Asn	Cys	Asn
				180					185					190	
Thr	Ser	Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Glu	Pro	Ile
	195							200						205	

Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys
210 215 220

Asp Lys Glu Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val
225 230 235 240

Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu
245 250 255

Asn Gly Ser Leu Ala Glu Glu Glu Val Ile Ile Arg Ser Glu Asn Ile
260 265 270

Thr Asn Asn Ala Lys Thr Ile Ile Val Gln Leu Thr Lys Pro Val Lys
275 280 285

Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile
290 295 300

Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile
305 310 315 320

Arg Gln Ala His Cys Asn Val Ser Arg Ser Glu Trp Asn Lys Thr Leu
325 330 335

Gln Lys Val Ala Lys Gln Leu Arg Lys Tyr Phe Lys Asn Lys Thr Ile
340 345 350

Ile Phe Thr Asn Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser
355 360 365

Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe
370 375 380

Asn Ser Thr Trp Asn Asn Gly Thr Met Lys Asn Thr Ile Thr Leu Pro
385 390 395 400

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Ala Gly Gln Ala
405 410 415

Met Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Glu Ser Asn Ile
420 425 430

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Asn Glu
435 440 445

Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu

450 455 460
 Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro
 465 470 475 480
 Thr Arg Ala Lys Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly
 485 490 495
 Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln
 500 505 510
 His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg
 515 520 525

 Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile
 530 535 540
 Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn
 545 550 555 560
 Ser Ser Trp Ser Asn Lys Ser Gln Asn Glu Ile Trp Asp Asn Met Thr
 565 570 575
 Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr His Ile Ile Tyr
 580 585 590
 Asn Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp

 595 600 605
 Leu Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp Ile
 610 615 620
 Ser Asn Trp Leu Trp
 625
 <210> 40
 <211> 1918
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 40

 ttcagtcgac agccaccatg agggatgatgg gaatccaacg gaactgccag catcttctcc 60
 ggtggggaac gatgatactg ggaatgataa taatctgctc tgccgtgaa aacctctggg 120
 tcacagtgtg ctacggagtg cctgtatgga aggacgtga aaccactctc ttttgtgctt 180

 ccgatgctaa agcctacgaa accgagatgc acaatgtttg ggccaccac gcctgcgtgc 240

caactgatcc taatccacaa gaaatacatc tggagaatgt tactgaggaa tttacatgt 300
ggaaaaataa tatggtagag caaatgcaca ctgacatcat ttcactctgg gaccaatcac 360
tcaaaccctg cgttaaactt accccctctt gcgtgaccct caattgtagc aacgtcaacg 420
tcacaaataa tacaaccaac actcacgagg aagaaattaa aaattgctcc tttaatatga 480
ccactgaact tcgcgacaaa aaacaaaaag tctattcact gttttatagg ctggacgtcg 540
tccaaatcaa cgagaacaat tctaacagta gctatcgact tatcaattgc aatacctctg 600

ctattaccca agcttgtcct aaagtctctt ttgaaccaat ccctatccac tactgtgccc 660
cagctggatt cgcaattctg aagtgaagg ataaggaatt caacggaact ggcccttgca 720
agaacgttag cactgtccaa tgcactcacg gaatcaaacc agtagtcagc actcaactgc 780
tcctgaatgg ctactcgcc gaagaaggagg tgattatccg aagcgagaac ataactaaca 840
atgcgaagac aataattgtt caattgacga aaccagtga gatcaactgt actagaccaa 900
ataacaacac aagaaaatct atcagaattg gccccggaca agccttctac gcaacaggag 960
atatcatagg tgacatcaga caggccatt gcaacgtttc aagaagcgag tggaataaaa 1020

cactccagaa agtggcaaag cagctgagaa aatacttta gaacaagaca atcatattha 1080
ctaactctc cggagggtgat ctgaaataa ccactcatag ctttaattgt gggggcgaat 1140
tcttctactg taacacatct ggccctttta attctacctg gaataacggc accatgaaaa 1200
atactatcac cctcccttgc agaattaagc aaatcattaa catgtggcag agagcaggac 1260
aggccatgta tgccctccc attcaagggtg tgattcgatg tgaaagcaac attactggac 1320
ttcttctgac ccgggatggc ggaaataata ataccaatga gacattcaga cccggcggcg 1380
gcgatatcg agacaattgg cgaagtgaac ttataaata caaagtagtt aagattgagc 1440

cccttggagt tgccctact agagcaaaaa cattgaccgt tcaggccagg cagctgctct 1500
caggaatcgt gcagcagcaa agtaacctcc tccgagctat cgaggcaca caacatctct 1560
tgaaattgac cgtatgggga atcaagcaat tgcaggctag ggttttggct gtggaacgct 1620
atctcaagga tcagcagctt ctgggaatct ggggatgctc tgggaaattg atatgtacta 1680
caaacgtacc ctggaactca agctggagta ataaaagcca gaacgaaatt tgggataata 1740
tgacctggct gcagtgggac aaagaaattt ctaattatac tcatatcata tacaatctga 1800
tcgaagaatc acagaaccag caggaaaaga atgagcaaga ccttctggcc ttggacaagt 1860

gggctaactt gtggaactgg tttagacatta gcaactggct gtggtaaaga tcttaca 1918

<210> 41

<211> 835

<212> PRT

<213> Human immunodeficiency virus

<400> 41

Met Arg Val Arg Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Val Val Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asn Ala Thr Asn Ala Thr Asn Thr Met Gly Glu Ile Lys

130 135 140

Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys

145 150 155 160

Val Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asn Glu Asn

165 170 175

Asn Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala

180 185 190

Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro

195 200 205

Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr

210 215 220

Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys

225 230 235 240

Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
245 250 255

Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile
260 265 270

Ile Val His Leu Asn Glu Ser Val Glu Ile Val Cys Thr Arg Pro Asn
275 280 285

Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr
290 295 300

Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile
305 310 315 320

Ser Glu Asp Lys Trp Asn Lys Thr Leu Gln Lys Val Ser Lys Lys Leu
325 330 335

Lys Glu His Phe Pro Asn Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly
340 345 350

Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe
355 360 365

Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr Asn Ser Thr
370 375 380

Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
385 390 395 400

Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn
405 410 415

Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
420 425 430

Gly Lys Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
435 440 445

Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys
450 455 460

Pro Leu Gly Ile Ala Pro Thr Lys Ala Lys Arg Arg Val Val Glu Arg
465 470 475 480

Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly

485 490 495
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 500 505 510
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
 515 520 525
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
 530 535 540
 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys

 545 550 555 560
 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys
 565 570 575
 Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Glu
 580 585 590
 Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile Ser
 595 600 605
 Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Asn Gln
 610 615 620

 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
 625 630 635 640
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
 645 650 655
 Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala
 660 665 670
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser
 675 680 685
 Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Pro Asp Arg Leu Gly Arg

 690 695 700
 Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile Arg Leu
 705 710 715 720
 Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys
 725 730 735
 Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg

740 745 750
Ala Val Glu Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly
755 760 765
Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp Gly Leu
770 775 780
Glu Leu Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala Ile Ala
785 790 795 800
Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys
805 810 815
Arg Ala Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala
820 825 830
Ala Leu Gln

835
<210> 42
<211> 619
<212> PRT
<213> Human immunodeficiency virus
<400> 42
Met Arg Val Arg Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp
1 5 10 15
Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Val Val Gly Asn
20 25 30
Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
35 40 45
Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val
50 55 60
His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80
Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95
Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asn Ala Thr Asn Ala Thr Asn Thr Met Gly Glu Ile Lys
130 135 140

Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys
145 150 155 160

Val Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asn Glu Asn
165 170 175

Asn Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala
180 185 190

Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro
195 200 205

Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr
210 215 220

Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys
225 230 235 240

Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
245 250 255

Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile
260 265 270

Ile Val His Leu Asn Glu Ser Val Glu Ile Val Cys Thr Arg Pro Asn
275 280 285

Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr
290 295 300

Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile
305 310 315 320

Ser Glu Asp Lys Trp Asn Lys Thr Leu Gln Lys Val Ser Lys Lys Leu
325 330 335

Lys Glu His Phe Pro Asn Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly
340 345 350

Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe

355 360 365
 Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr Asn Ser Thr
 370 375 380
 Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
 385 390 395 400
 Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn
 405 410 415

 Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
 420 425 430
 Gly Lys Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
 435 440 445
 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys
 450 455 460
 Pro Leu Gly Ile Ala Pro Thr Lys Ala Lys Thr Leu Thr Val Gln Ala
 465 470 475 480
 Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg

 485 490 495
 Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile
 500 505 510
 Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys Asp
 515 520 525
 Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr
 530 535 540
 Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Glu Asp
 545 550 555 560

 Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile Ser Asn
 565 570 575
 Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Asn Gln Gln
 580 585 590
 Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn Leu
 595 600 605
 Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp

610	615	
<210>	43	
<211>	1888	
<212>	DNA	
<213>		
>	Human immunodeficiency virus	
<400>	43	
ttcagtcgac agccaccatg cgagtgagag gcattctgcg gaattgtcag caatggtgga		60
tctggggcat actcggattc tggatgctta tgatatgcaa tgttgtgggg aacctgtggg		120
ttaccgtata ctatgggggtt ccagtctgga aggaggctaa aacaacgctg ttctgtgcaa		180
gtgacgcaa agcctacgag aaagaagtgc acaacgtctg ggctaccac gcttgtgttc		240
caaccgatcc aaacccccag gaaatcgtcc tcgagaacgt gactgaaaaa tttaacatgt		300
ggaagaatga tatggtagat cagatgcacg aagatatcat ttcattgtgg gaccaatcat		360
tgaaacatg cgtaaaactg acccccctct gcgtaaacct taactgcacc aatgcaacta		420
atgccaccaa tactatgggc gaaataaaaa actgtagctt taacattaca acggaactcc		480
gggataagaa acaaaaggtc tacgcgctct ttaccgact cgatatcgtc ccacttaacg		540
agaataatag ttaccgctg attaaactgta acacatcagc cattacgcaa gcttgcccca		600
aagtttcttt cgaccccatc ccaattcact atttgtcccc cgctggatac gctatactta		660
aatgcaacaa taaaacattt aatggaaccg gaccatgtaa caacgtcagt accgtacaat		720
gtacgcacgg aattaaacct gttgtctcaa ccagcttct ccttaacggc tcatitggcgg		780
aggaagaaat tattatcaga tcagaaaact tgaccaacaa tgccaaaacc atcatcgtgc		840
acctcaatga atccgtggaa atcgtgtgca ccagaccaa taacaatacc cggaatcaa		900
tcaggattgg gcctggccag acattttacg ctacaggtga tataattggc gatattagac		960
aagccattg caacatatca gaagacaagt ggaataagac tetgcagaag gtttctaaga		1020
agctgaagga acactttccc aataaaacga ttaagttcga gccctcttca ggaggagacc		1080
ttgagatcac aacacactct tttaattgta gaggggagtt cttctattgt aatacatcaa		1140
agctctttaa cagtacctac aactccacta atagtacat cacactcccc tgcagaataa		1200
agcaaataat caacatgtgg caagaagttg gccgagcaat gtacgccct cccatcgcag		1260
gcaacattac atgtaaacc aatattactg gccttttgct gacacgggac ggcggaagaa		1320
ataaactga gaccttcaga cctggcggag gcgatatgcg cgataattgg cggagcgagc		1380
tctacaagta taaagtcgtt gaaatcaagc cactgggcat agctcctacg aaagcaaaga		1440
cactcactgt tcaggctaga cagctgctct ccggcatagt gcaacagcaa tccaatctcc		1500

tcgagctat cgaagcccaa caacatatgc tccagcttac cgtctgggga atcaaacaat 1560
tgcaaacacg agtgctggcg atagagagat atttgaaga tcagcaactc ctggggattt 1620

ggggctgttc aggtaagctc atctgtacaa ctgcggtgcc gtggaactca agctggagta 1680
acaaaagcca agaggatata tgggacaaca tgacttggat gcagtgggat cgagaaataa 1740
gcaactatac agataccatt tatcggtccc tggaggactc acagaaccag caggagaaaa 1800
atgagaaaga ttgctcgcg cttgacagtt ggaagaattt gtggaattgg ttcgacatta 1860
caactggct ctggtaaaga tcttacia 1888

<210> 44
<211> 842
<212> PRT
<213> Human immunodeficiency virus
<400> 44

Met Arg Val Lys Gly Ile Gln Arg Asn Trp Gln His Leu Trp Lys Trp

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

Gln Glu Ile Thr Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys		
85	90	95
Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp		
100	105	110
Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu		
115	120	125
Asn Cys Thr Asp Val Asn Val Thr Asn Asn Asn Thr Asn Asn Thr Lys		
130	135	140
Lys Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp		

145 150 155 160
 Lys Lys Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro
 165 170 175
 Ile Asn Asp Asn Gly Asn Ser Ser Ile Tyr Arg Leu Ile Asn Cys Asn
 180 185 190
 Val Ser Thr Ile Lys Gln Ala Cys Pro Lys Val Thr Phe Asp Pro Ile
 195 200 205
 Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Arg
 210 215 220

 Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val
 225 230 235 240
 Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu
 245 250 255
 Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile
 260 265 270
 Thr Asp Asn Thr Lys Val Ile Ile Val Gln Leu Asn Glu Thr Ile Glu
 275 280 285
 Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile

 290 295 300
 Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile
 305 310 315 320
 Arg Gln Ala His Cys Asn Val Ser Arg Thr Lys Trp Asn Glu Met Leu
 325 330 335
 Gln Lys Val Lys Ala Gln Leu Lys Lys Ile Phe Asn Lys Ser Ile Thr
 340 345 350
 Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe
 355 360 365

 Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn
 370 375 380
 Asn Ser Leu Leu Asn Ser Thr Asn Ser Thr Ile Thr Leu Pro Cys Lys
 385 390 395 400
 Ile Lys Gln Ile Val Arg Met Trp Gln Arg Val Gly Gln Ala Met Tyr

405 410 415
 Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly
 420 425 430
 Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Glu Thr Phe Arg

 435 440 445
 Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys
 450 455 460
 Tyr Lys Ile Val Lys Ile Lys Pro Leu Gly Val Ala Pro Thr Arg Ala
 465 470 475 480
 Arg Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Leu Gly Ala
 485 490 495
 Val Leu Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 500 505 510

 Ser Ile Thr Leu Thr Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val
 515 520 525
 Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu
 530 535 540
 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu
 545 550 555 560
 Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 565 570 575
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Thr Ser

 580 585 590
 Trp Ser Asn Lys Ser Tyr Asn Glu Ile Trp Asp Asn Met Thr Trp Ile
 595 600 605
 Glu Trp Glu Arg Glu Ile Ser Asn Tyr Thr Gln Gln Ile Tyr Ser Leu
 610 615 620
 Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu
 625 630 635 640
 Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Lys
 645 650 655

Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile
660 665 670
Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg
675 680 685
Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr His His Gln Arg
690 695 700
Glu Pro Asp Arg Pro Glu Arg Ile Glu Glu Gly Gly Gly Glu Gln Asp
705 710 715 720
Lys Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp

725 730 735
Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
740 745 750
Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Ser Ser
755 760 765
Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr Leu Trp Asn
770 775 780
Leu Leu Leu Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala Ile Asn Leu
785 790 795 800

Leu Asp Thr Ile Ala Ile Ala Val Ala Asn Trp Thr Asp Arg Val Ile
805 810 815
Glu Val Ala Gln Arg Ala Cys Arg Ala Ile Leu Asn Ile Pro Arg Arg
820 825 830
Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu
835 840

<210> 45
<211> 626
<212> PRT
<213> Human immunodeficiency virus
<400> 45

Met Arg Val Lys Gly Ile Gln Arg Asn Trp Gln His Leu Trp Lys Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asn Asn

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

275 280 285
 Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile
 290 295 300

 Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile
 305 310 315 320
 Arg Gln Ala His Cys Asn Val Ser Arg Thr Lys Trp Asn Glu Met Leu
 325 330 335
 Gln Lys Val Lys Ala Gln Leu Lys Lys Ile Phe Asn Lys Ser Ile Thr
 340 345 350
 Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe
 355 360 365
 Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn

 370 375 380
 Asn Ser Leu Leu Asn Ser Thr Asn Ser Thr Ile Thr Leu Pro Cys Lys
 385 390 395 400
 Ile Lys Gln Ile Val Arg Met Trp Gln Arg Val Gly Gln Ala Met Tyr
 405 410 415
 Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly
 420 425 430
 Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Glu Thr Phe Arg
 435 440 445

 Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys
 450 455 460
 Tyr Lys Ile Val Lys Ile Lys Pro Leu Gly Val Ala Pro Thr Arg Ala
 465 470 475 480
 Arg Thr Leu Thr Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val Gln
 485 490 495
 Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu
 500 505 510
 Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala

 515 520 525

Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys
 530 535 540
 Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Thr Ser Trp
 545 550 555 560
 Ser Asn Lys Ser Tyr Asn Glu Ile Trp Asp Asn Met Thr Trp Ile Glu
 565 570 575
 Trp Glu Arg Glu Ile Ser Asn Tyr Thr Gln Gln Ile Tyr Ser Leu Ile
 580 585 590

Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala
 595 600 605
 Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Lys Trp
 610 615 620

Leu Trp

625

<210> 46

<211> 1909

<212> DNA

<213> Human immunodeficiency virus

<400> 46

ttcagtcgac agccaccatg cgagtgaagg gaatccagag aaattggcag cacctttgga 60

agtggggcac atcatcctc ggccttgtga tcatatgctc tgcctcaaat aacctttggg 120

tcacagttaa ttacggcgtg cccgtttggg aggacgcaga cacaactctt ttttgtcca 180

gcgacgctaa ggcttattca acagagaggc ataacgtttg ggctacacat gcatgctgac 240

cgaccgatcc taatccccag gaaatcactc ttgagaatgt tacagagaat tttaatatgt 300

ggaagaacaa catggttgaa cagatgcatg aagacataat ttctctctgg gatgaatctc 360

tgaacacttg cgtgaagctt acaccactgt gcgttacctt gaattgcact gacgtcaatg 420

tcacaataa taataccaac aatacaaaaa aagaaatcaa aaattgttct ttcaacataa 480

ccaccgagat acgcgataaa aaaaagaaag aatagccctt gttctacaga ctgatgtgg 540

tccaattaa tgacaacgga aattcttcca tctaccgact tatcaattgt aacgtgtcta 600

caatcaaaca ggctgtcctt aaagtcacat ttgacctat tccattcat tactgtgcc 660

ccgtggctt cgtattctt aaatgccgag acaaaaaatt taacggaaca ggaccatgca 720

agaatgtctc aacagttcaa tgcactcatg gaattaaacc agtcgtttct actcaactcc 780

ttctcaatgg aagcctggca gaagaggaaa tcataatccg cagcgaaaac ataaccgaca 840
acacaaaagt aatcatcgta cagctgaacg agaccattga aataaattgt acgagaccta 900
ataataacac aagaaaaagc atacgcatcg gccccggaca ggctttctac gccacaggag 960

acattatcgg agatatccgc caggctcact gtaatgtgtc tagaacaaaa tggaacgaaa 1020
tgcttcagaa ggtcaaagct cagctcaaga aaatattcaa caaatctatt acattcaact 1080
catcatcagg cggcgatctg gagataacaa ctatttcctt caactgtcgg ggagaatfff 1140
tttactgtaa cagctccggc ctgttcaaca attcactcct gaatagcact aactccacca 1200
tcactctccc atgtaagatc aaacaaatcg tcagaatgtg gcagcgagtc ggtcaagcta 1260
tgtaccccc tccaatcgcc ggtaatatca catgtagaag caatatacaca gggctcttgc 1320
ttacaaggga cggcgggaac aacaacaccg aaaccttcag accaggagga ggagacatgc 1380

gagacaattg gcggagcgag ctgtataaat ataagatcgt aaaaatcaaa ccattgggtg 1440
tagcgccaac tagagccga acactgaccg tgcagggtgag gcaactgctg agcggcattg 1500
tccaacaaca atccaatctt cttagagcaa tcgaggccca gcagcatctg ctccagctta 1560
ctgtatgggg aatcaaacia ctgcaagcaa gagtattggc agtggagagg tatctcaagg 1620
accagcagct tctgggaatt tggggttgca gcggaaagct catatgtaca accaatgtgc 1680
cctggaacac tagttggagt aataagagtt acaatgaaat ctgggacaat atgacatgga 1740
tcgaatggga gcgcgaaata tccaactata ctacagcaat ctattccctc attgaagaga 1800

gtcagaacca gcaggaaaag aatgagcaag acctcctcgc cctggataaa tgggcatctc 1860
tgtggaactg gtttgacata actaaatggt tgtggtaaag atcttacia 1909

<210> 47

<211> 854

<212> PRT

<213> Human immunodeficiency virus

<400> 47

Met Arg Val Lys Glu Thr Gln Met Asn Trp Pro Asn Leu Trp Lys Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asp Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Asp

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala His Glu Thr Glu Val

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

305 310 315 320
 Ile Ile Gly Asp Ile Arg Lys Ala Tyr Cys Glu Ile Asn Gly Thr Lys

 325 330 335
 Trp Asn Glu Val Leu Lys Gln Val Thr Glu Lys Leu Lys Glu His Phe
 340 345 350
 Asn Asn Lys Thr Ile Ile Phe Gln Pro Pro Ser Gly Gly Asp Leu Glu
 355 360 365
 Ile Thr Met His His Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn
 370 375 380
 Thr Thr Lys Leu Phe Asn Asn Thr Cys Ile Gly Asn Glu Thr Met Glu
 385 390 395 400

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

 465 470 475 480
 Gln Ile Glu Pro Leu Gly Ile Ala Pro Thr Arg Ala Lys Arg Arg Val
 485 490 495
 Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Met Ile Phe Gly
 500 505 510
 Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu
 515 520 525
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 530 535 540

 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr
 545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg
565 570 575

Tyr Leu Lys Asp Gln Lys Phe Leu Gly Leu Trp Gly Cys Ser Gly Lys
580 585 590

Ile Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Thr Trp Ser Asn Arg
595 600 605

Ser Phe Glu Glu Ile Trp Asn Asn Met Thr Trp Ile Glu Trp Glu Arg
610 615 620

Glu Ile Ser Asn Tyr Thr Asn Gln Ile Tyr Glu Ile Leu Thr Glu Ser
625 630 635 640

Gln Asn Gln Gln Asp Arg Asn Glu Lys Asp Leu Leu Glu Leu Asp Lys
645 650 655

Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr
660 665 670

Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile
675 680 685

Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser
690 695 700

Pro Leu Ser Phe Gln Thr Pro Thr His His Gln Arg Glu Pro Asp Arg
705 710 715 720

Pro Glu Arg Ile Glu Glu Gly Gly Gly Glu Gln Gly Arg Asp Arg Ser
725 730 735

Val Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg
740 745 750

Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile
755 760 765

Ala Ala Arg Thr Val Glu Leu Leu Gly His Ser Ser Leu Lys Gly Leu
770 775 780

Arg Arg Gly Trp Glu Gly Leu Lys Tyr Leu Gly Asn Leu Leu Leu Tyr
785 790 795 800

Trp Gly Gln Glu Leu Lys Ile Ser Ala Ile Ser Leu Leu Asp Ala Thr
805 810 815

Ala Ile Ala Val Ala Gly Trp Thr Asp Arg Val Ile Glu Val Ala Gln
820 825 830

Gly Ala Trp Arg Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln Gly
835 840 845

Leu Glu Arg Ala Leu Leu
850

<210> 48
<211> 638
<212> PRT
<213> Human immunodeficiency virus
<400> 48

Met Arg Val Lys Glu Thr Gln Met Asn Trp Pro Asn Leu Trp Lys Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asp Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Asp
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala His Glu Thr Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile His Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asn Met Val Glu Gln Met Gln Glu Asp Val Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asn Ala Asn Leu Thr Asn Val Asn Asn Ile Thr Asn Val
130 135 140

Ser Asn Ile Ile Gly Asn Ile Thr Asn Glu Val Arg Asn Cys Ser Phe
145 150 155 160

Asn Met Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys Val His Ala Leu

165 170 175
 Phe Tyr Lys Leu Asp Ile Val Gln Ile Glu Asp Asn Asn Ser Tyr Arg
 180 185 190
 Leu Ile Asn Cys Asn Thr Ser Val Ile Lys Gln Ala Cys Pro Lys Ile
 195 200 205
 Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Thr Pro Ala Gly Tyr Ala
 210 215 220
 Ile Leu Lys Cys Asn Asp Lys Asn Phe Asn Gly Thr Gly Pro Cys Lys
 225 230 235 240
 Asn Val Ser Ser Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser

 245 250 255
 Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile
 260 265 270
 Arg Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu
 275 280 285
 Asn Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Ser Asn Asn Thr Arg
 290 295 300
 Thr Ser Ile Thr Ile Gly Pro Gly Gln Val Phe Tyr Arg Thr Gly Asp
 305 310 315 320

 Ile Ile Gly Asp Ile Arg Lys Ala Tyr Cys Glu Ile Asn Gly Thr Lys
 325 330 335
 Trp Asn Glu Val Leu Lys Gln Val Thr Glu Lys Leu Lys Glu His Phe
 340 345 350
 Asn Asn Lys Thr Ile Ile Phe Gln Pro Pro Ser Gly Gly Asp Leu Glu
 355 360 365
 Ile Thr Met His His Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn
 370 375 380
 Thr Thr Lys Leu Phe Asn Asn Thr Cys Ile Gly Asn Glu Thr Met Glu

 385 390 395 400
 Gly Cys Asn Gly Thr Ile Ile Leu Pro Cys Lys Ile Lys Gln Ile Ile
 405 410 415

Asn Met Trp Gln Gly Ala Gly Gln Ala Met Tyr Ala Pro Pro Ile Ser
 420 425 430
 Gly Arg Ile Asn Cys Val Ser Asn Ile Thr Gly Ile Leu Leu Thr Arg
 435 440 445
 Asp Gly Gly Ala Asn Asn Thr Asn Glu Thr Phe Arg Pro Gly Gly Gly
 450 455 460

 Asn Ile Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
 465 470 475 480
 Gln Ile Glu Pro Leu Gly Ile Ala Pro Thr Arg Ala Lys Thr Leu Thr
 485 490 495
 Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn
 500 505 510
 Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val
 515 520 525
 Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr

 530 535 540
 Leu Lys Asp Gln Lys Phe Leu Gly Leu Trp Gly Cys Ser Gly Lys Ile
 545 550 555 560
 Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Thr Trp Ser Asn Arg Ser
 565 570 575
 Phe Glu Glu Ile Trp Asn Asn Met Thr Trp Ile Glu Trp Glu Arg Glu
 580 585 590
 Ile Ser Asn Tyr Thr Asn Gln Ile Tyr Glu Ile Leu Thr Glu Ser Gln
 595 600 605

 Asn Gln Gln Asp Arg Asn Glu Lys Asp Leu Leu Glu Leu Asp Lys Trp
 610 615 620
 Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp
 625 630 635
 <210> 49
 <211> 1945
 <212> DNA
 <213> Human immunodeficiency virus

<400> 49

ttcagtcgac agccaccatg cgagtcagg aacacaaat gaactggcct aatctgtgga	60
agtggggcac cctgatcctg ggtttgggtca ttatttgctc tgcgagcgac aatctctggg	120
ttactgtcta ttacggagtc cccgtttgga gagatgccga cactacactg ttctgcgcct	180
cagatgccaa agctcatgaa actgaagtgc ataatgtttg ggcaaccac gcctgtgttc	240
ctaccgacce aaacccccaa gaaatacacc tggaaaacgt gaccgagaac ttaatatgt	300
ggaagaataa catgggtgaa cagatgcaag aagacgtaat cagcctgtgg gatcaaagtc	360
tgaaaccttg cgtaaaactg actccacttt gcgtaacact taattgcacc aacgcgaacc	420
tgacaaacgt taacaacatc actaacgtct ccaacatcat cggcaacata acgaacgaag	480
tgagaaattg cagtttcaat atgactacag agctccggga caagaaacag aaggtccatg	540
ctctctttta caaacctgac atcgtccaga tcgaagacaa taacagctac agacttataa	600
attgtaatac atccgtgatt aaacaagcat gccccaaat aagcttcgat cctattccta	660
tccactactg tactcctgcc ggctatgcta tcttgaaatg caatgataag aacttcaatg	720
ggaccggacc ttgtaagaac gtgtctagtg tgcaatgcac tcacggcatt aaaccagtgg	780
taagcaccca gctgctcctg aacggctctc tggcagagga agagattatt attcgaagtg	840
agaacctcac caacaacgt aagactatca tcgtacatct caataaatca gtcgaaatta	900
attgcaccag accctccaat aatactagaa cttaatcac tatcggccca ggacaagtct	960
tttatagaac aggagatatc ataggagata tcagaaaggc atattgcgag ataaacggga	1020
caaatggaa cgaagtactc aaacaagtca cagagaagct taaggaacat ttcaacaata	1080
aaaccattat tttcaacc ccaagtggcg gagacctga aatcactatg caccacttca	1140
actgccgagg cgaatTTTTT tattgcaata ccactaaact tttcaacaat acgtgcatcg	1200
gaaatgagac catggagggc tgcaatggaa caatcatact cccatgcaag ataaaacaaa	1260
tcattaacat gtggcaaggt gctggacaag ctatgtatgc accccaata tccggtagaa	1320
ttaattgcgt cagcaacatc actggcatac tgctcactag agacggagga gcaaataata	1380
caaatgaaac attccgacca ggcggcggca acattaagga caactggcgg tccgaactct	1440
ataagtacaa agtcgtacag atcgaacctc ttggaatagc accgactcgc gctaagacac	1500
tcacagtaca ggccccgaaa cttctttctg gaatcgtaca gcagcaatcc aacctcctcc	1560
gcgcaatcga ggccaacaa catctgcttc agctcacagt ttggggaatc aagcagctcc	1620
aggcacgcgt gctcgcagtg gaaagatacc tgaaggatca gaaattcctt ggtctctggg	1680
gatgtttctg caaaataatc tgcaactacc cgtttccctg gaattcaaca tggagcaacc	1740
ggagttttga agagatatgg aacaatatga catggataga gtgggaaagg gaaattagta	1800

actatacgaa ccagatatac gaaatcctca ccgaaagcca aaatcagcag gatcgcaacg 1860

aaaaagacct cctcgagctt gataagtggg catccctttg gaactgggtc gacatcacia 1920

attggtcttg gtaaagatct tacaa 1945

<210> 50

<211> 863

<212> PRT

<213> Human immunodeficiency virus

<400> 50

Met Gly Ala Met Gly Ile Gln Met Asn Trp Gln Asn Leu Trp Arg Trp

1 5 10 15

Gly Thr Met Ile Leu Gly Met Leu Ile Ile Cys Ser Val Ala Glu Lys

20 25 30

Ser Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Glu

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala His Asp Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Ile Leu Glu Asn Val Thr Glu Asp Phe Asn Met Trp Lys

85 90 95

Asn Ser Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Ser Asp Ser Asn Ile Thr Ser Asn Ser Thr Ser Asn Ser Thr

130 135 140

Lys Asp Ser Ala Thr Leu Asp Met Lys Ser Glu Ile Gln Asn Cys Ser

145 150 155 160

Phe Asn Met Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys Val Tyr Ser

165 170 175

Leu Phe Tyr Arg Leu Asp Val Val Gln Ile Asn Glu Asn Ser Ser Asp

180 185 190
 Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro
 195 200 205
 Lys Val Thr Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly
 210 215 220
 Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro
 225 230 235 240
 Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val
 245 250 255

 Val Thr Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val
 260 265 270
 Met Ile Arg Ser Glu Asn Ile Thr Glu Asn Ala Lys Asn Ile Ile Val
 275 280 285
 Gln Phe Lys Glu Pro Val Gln Ile Ile Cys Ile Arg Pro Gly Asn Asn
 290 295 300
 Thr Arg Lys Ser Val His Ile Gly Pro Gly Gln Ala Phe Tyr Ala Thr
 305 310 315 320
 Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Val Ser Arg

 325 330 335
 Glu Leu Trp Asn Lys Thr Leu Gln Glu Val Ala Thr Gln Leu Arg Lys
 340 345 350
 His Phe Arg Asn Asn Thr Lys Ile Ile Phe Thr Asn Ser Ser Gly Gly
 355 360 365
 Asp Val Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe
 370 375 380
 Tyr Cys Asp Thr Ser Gly Leu Phe Asn Ser Ser Trp Thr Ala Ser Asn
 385 390 395 400

 Asp Ser Met Gln Glu Ala His Ser Thr Glu Ser Asn Ile Thr Leu Gln
 405 410 415
 Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Ala Gly Gln Ala
 420 425 430
 Met Tyr Ala Pro Pro Ile Pro Gly Ile Ile Arg Cys Glu Ser Asn Ile

435 440 445
 Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Glu Gly Asn Asn Ser Thr
 450 455 460
 Asn Glu Thr Phe Arg Pro Val Gly Gly Asn Met Arg Asp Asn Trp Arg

 465 470 475 480
 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Val Glu Pro Leu Gly Val
 485 490 495
 Ala Pro Thr Lys Ser Arg Arg Arg Val Val Glu Arg Glu Lys Arg Ala
 500 505 510
 Val Gly Leu Gly Ala Val Phe Ile Gly Phe Leu Gly Ala Ala Gly Ser
 515 520 525
 Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu
 530 535 540

 Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu
 545 550 555 560
 Ala Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu
 565 570 575
 Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Asp Gln Gln Leu
 580 585 590
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val
 595 600 605
 Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Glu

 610 615 620
 Asn Met Thr Trp Met Gln Trp Asp Lys Glu Val Ser Asn Tyr Thr Gln
 625 630 635 640
 Met Ile Tyr Asn Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
 645 650 655
 Glu Gln Glu Leu Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp
 660 665 670
 Phe Asn Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile
 675 680 685

Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Val
690 695 700
Ile Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His
705 710 715 720
Thr Pro Asn Pro Arg Gly Leu Asp Arg Pro Gly Arg Ile Glu Glu Glu
725 730 735
Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe
740 745 750
Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr
755 760 765
His Arg Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg Thr Leu Glu Leu
770 775 780
Leu Gly His Asn Ser Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu
785 790 795 800
Lys Tyr Leu Trp Asn Leu Leu Ala Tyr Trp Gly Arg Glu Leu Lys Ile
805 810 815
Ser Ala Ile Ser Leu Val Asp Ser Ile Ala Ile Ala Val Ala Gly Trp
820 825 830
Thr Asp Arg Ile Ile Glu Ile Val Gln Ala Ile Gly Arg Ala Ile Leu
835 840 845
His Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Ile
850 855 860
<210> 51
<211> 647
<212> PRT
<213> Human immunodeficiency virus
<400> 51
Met Gly Ala Met Gly Ile Gln Met Asn Trp Gln Asn Leu Trp Arg Trp
1 5 10 15
Gly Thr Met Ile Leu Gly Met Leu Ile Ile Cys Ser Val Ala Glu Lys
20 25 30
Ser Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Glu

35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala His Asp Lys Glu Val
 50 55 60
 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80
 Gln Glu Met Ile Leu Glu Asn Val Thr Glu Asp Phe Asn Met Trp Lys
 85 90 95

 Asn Ser Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Ser Asp Ser Asn Ile Thr Ser Asn Ser Thr Ser Asn Ser Thr
 130 135 140
 Lys Asp Ser Ala Thr Leu Asp Met Lys Ser Glu Ile Gln Asn Cys Ser
 145 150 155 160
 Phe Asn Met Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys Val Tyr Ser

 165 170 175
 Leu Phe Tyr Arg Leu Asp Val Val Gln Ile Asn Glu Asn Ser Ser Asp
 180 185 190
 Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro
 195 200 205
 Lys Val Thr Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly
 210 215 220
 Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro
 225 230 235 240

 Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val
 245 250 255
 Val Thr Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val
 260 265 270
 Met Ile Arg Ser Glu Asn Ile Thr Glu Asn Ala Lys Asn Ile Ile Val
 275 280 285
 Gln Phe Lys Glu Pro Val Gln Ile Ile Cys Ile Arg Pro Gly Asn Asn

290 295 300
 Thr Arg Lys Ser Val His Ile Gly Pro Gly Gln Ala Phe Tyr Ala Thr

 305 310 315 320
 Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Val Ser Arg

 325 330 335
 Glu Leu Trp Asn Lys Thr Leu Gln Glu Val Ala Thr Gln Leu Arg Lys

 340 345 350
 His Phe Arg Asn Asn Thr Lys Ile Ile Phe Thr Asn Ser Ser Gly Gly

 355 360 365
 Asp Val Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe

 370 375 380

 Tyr Cys Asp Thr Ser Gly Leu Phe Asn Ser Ser Trp Thr Ala Ser Asn
 385 390 395 400
 Asp Ser Met Gln Glu Ala His Ser Thr Glu Ser Asn Ile Thr Leu Gln

 405 410 415
 Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Ala Gly Gln Ala

 420 425 430
 Met Tyr Ala Pro Pro Ile Pro Gly Ile Ile Arg Cys Glu Ser Asn Ile

 435 440 445
 Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Glu Gly Asn Asn Ser Thr

 450 455 460
 Asn Glu Thr Phe Arg Pro Val Gly Gly Asn Met Arg Asp Asn Trp Arg
 465 470 475 480
 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Val Glu Pro Leu Gly Val

 485 490 495
 Ala Pro Thr Lys Ser Arg Thr Leu Thr Val Gln Ala Arg Gln Leu Leu

 500 505 510
 Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala

 515 520 525

 Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
 530 535 540

Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Asp Gln Gln Leu Leu
 545 550 555 560
 Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro
 565 570 575
 Trp Asn Ser Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Glu Asn
 580 585 590
 Met Thr Trp Met Gln Trp Asp Lys Glu Val Ser Asn Tyr Thr Gln Met

595 600 605
 Ile Tyr Asn Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu
 610 615 620
 Gln Glu Leu Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe
 625 630 635 640
 Asn Ile Ser Asn Trp Leu Trp
 645

<210> 52
 <211> 1972
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 52

ttcagtcgac agccaccatg ggggcaatgg gaatccagat gaactggcag aacctctggc	60
gatggggcac aatgatcctg ggtatgctca tcatctgctc tgttcagaa aagtcatggg	120
taacagtcta ctacggcgta ccagtgtggc gggacgccga aaccactctc ttctgcgcct	180
ccgatgccaa agcacacgat aaagaagtcc acaatgtttg ggctacccat gcgtgcgtgc	240
caaccgatcc taaccacaa gaaatgatac tcgaaaacgt tactgaagac ttcaacatgt	300
ggaaaaattc tatggttgaa cagatgcaca ccgacataat atcactgtgg gatcagtcctc	360
tcaaaccctg tgicaaattg acccccctct gcgttacact gaactgttcc gactcaaata	420
tcacttctaa ttcaacgagc aatagtacga aagactccgc aacccttgat atgaaaagcg	480
aaatacagaa ctgttcattt aatatgacca ccgaactgag agataaaaag cagaaggttt	540
attctctgtt ctatcgattg gacgtgggtc agattaacga aaatagcagc gattaccgac	600
tcattaactg caatacatca gcaatcacac aggcttgccc aaaggtaaca tttagccaa	660
tccctattca ctactgcgcc cctgcaggat ttgccatcct gaaatgaac gataagaagt	720
ttaatgggac aggaccctgc accaacgtct ccaccgtgca atgcaccac ggcataaaac	780

ctgttggttac cacacaattg ctgctcaatg gatcacttgc tgaagaggaa gtcattgattc 840
ggctctgaaaa catcactgaa aatgccaaaa atattatagt tcagttcaaa gaacccgtcc 900

agatcatttg cattcgccct ggtaacaaca ctgcgaagtc agtgcacatt gggcccggcc 960
aggctttcta tgcaaccgga gatattatag gcgacatcag acaggcacat tgcaacgtca 1020
gccgggaatt gtggaacaaa actttgcagg aagttgctac tcagctgcga aaacatttca 1080
gaaacaatac aaagattatt ttcactaatt catcaggcgg tgacgtggag atcactaccc 1140
attcatttaa ctgtggcgga gaattcttct attgcgatac ctctgggctc ttttaattcct 1200
catggactgc tagcaacgat tcaatgcaag aagcacattc cacagaaaagt aatatcacac 1260
tgcagtgccg aattaaacaa atcatcaata tgtggcagcg ggccgggtcaa gcaatgtacg 1320

cacctcccat ccccgaatt attcgatgtg agtctaatat cactggcctc attctgaccc 1380
gagacggtgg cgaaggtaat aattctacaa acgagacttt cagaccgta ggaggcaata 1440
tgcgagacaa ttggcgatcc gaactgtata aatataaagt ggtgaaggta gaacctcttg 1500
gagtggcacc caccaaatca cgaaccctga ctgtgcaggc acgccaactt ctgagcggaa 1560
tagtccaaca gcaatccaat cttctgagag ctatagaagc ccagcaacac ctgcttaaac 1620
ttacggtgtg gggaatcaaa caattgcagg caagagtgtc ggagtgga cgatacttga 1680
gagaccaaca actcctggga atctggggat gttccggtaa gttgatttgc acgacaaacg 1740

ttccctggaa ctcttctgg tcaacaaga gtcaggacga aatatgggaa aatatgacat 1800
ggatgcagtg ggacaaggaa gttagcaact atacacagat gatctacaac ctctcgaag 1860
aatctcagaa tcaacaggaa aaaaacgaac aagaactgct cgcctcgat aagtgggcta 1920
acctctggaa ctggtttaat atttcaaact ggttgtggta aagatcttac aa 1972

<210> 53

<211> 861

<212> PRT

<213> Human immunodeficiency virus

<400> 53

Met Arg Val Lys Glu Ile Arg Arg Asn Cys Gln Arg Leu Arg Arg Trp

1 5 10 15

Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu Gln

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

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

290 295 300
 Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Ala Gly Lys
 305 310 315 320
 Ala Leu Tyr Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys
 325 330 335
 Asn Leu Ser Arg Ala Gln Trp Asn Asn Thr Leu Lys Gln Ile Val Ile
 340 345 350
 Lys Leu Arg Glu Gln Phe Gly Asn Lys Thr Ile Val Phe Asn Gln Ser
 355 360 365
 Ser Gly Gly Asp Val Glu Ile Val Met His Ser Phe Asn Cys Gly Gly

 370 375 380
 Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn
 385 390 395 400
 Gly Asn Asp Thr Trp Asn Asp Thr Trp Lys Asp Thr Thr Asn Asp Asn
 405 410 415
 Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Val Asn Met Trp Gln Lys
 420 425 430
 Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys
 435 440 445

 Ser Ser Lys Ile Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Thr Asn
 450 455 460
 Gly Thr Asn Glu Thr Glu Thr Phe Arg Pro Gly Gly Gly Asn Met Lys
 465 470 475 480
 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu
 485 490 495
 Pro Leu Gly Ile Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg
 500 505 510
 Glu Lys Arg Ala Val Gly Thr Ile Gly Ala Met Phe Leu Gly Phe Leu

 515 520 525
 Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Leu Thr Leu Thr Val
 530 535 540

Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu
545 550 555 560
Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
565 570 575
Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu
580 585 590
Arg Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Arg Leu Ile
595 600 605
Cys Thr Thr Asn Val Pro Trp Asn Thr Ser Trp Ser Asn Arg Ser Leu
610 615 620
Asn Tyr Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile
625 630 635 640
Asn Asn Tyr Thr Asp Tyr Ile Tyr Thr Leu Leu Glu Asp Ala Gln Asn
645 650 655
Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala
660 665 670
Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys
675 680 685
Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe
690 695 700
Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu
705 710 715 720
Ser Leu Gln Thr His Leu Pro Ala Arg Arg Gly Pro Asp Arg Pro Glu
725 730 735
Gly Ile Gly Glu Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser Val Arg
740 745 750
Leu Val His Gly Phe Leu Ala Leu Val Trp Glu Asp Leu Arg Ser Leu
755 760 765
Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Leu Leu Ile Val Ala
770 775 780
Arg Thr Val Glu Ile Leu Gly Gln Arg Gly Trp Glu Ala Leu Lys Tyr
785 790 795 800

Trp Trp Asn Leu Leu Leu Tyr Trp Ser Leu Glu Leu Lys Asn Ser Ala

805 810 815

Val Ser Leu Val Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr Asp

820 825 830

Arg Ile Ile Glu Ile Ala Arg Arg Ile Phe Arg Ala Phe Leu His Ile

835 840 845

Pro Thr Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu

850 855 860

<210> 54

<211> 651

<212> PRT

<213> Human immunodeficiency virus

<400> 54

Met Arg Val Lys Glu Ile Arg Arg Asn Cys Gln Arg Leu Arg Arg Trp

1 5 10 15

Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu Gln

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Val Thr Glu Lys

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Glu

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Lys Leu Arg Asn Asp Thr Ser Gly Thr Asn Ser Ser

130 135 140

Ser Trp Glu Lys Val Gln Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn

145 150 155 160

Ile Thr Thr Gly Ile Arg Gly Arg Val Gln Glu Tyr Ser Leu Phe Tyr

165 170 175

Lys Leu Asp Val Ile Pro Ile Asp Ser Arg Asn Asn Ser Asn Asn Ser

180 185 190

Thr Glu Phe Ser Ser Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile

195 200 205

Thr Gln Ala Cys Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr

210 215 220

Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe

225 230 235 240

Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His

245 250 255

Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu

260 265 270

Ala Glu Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn Val

275 280 285

Lys Ser Ile Ile Val Gln Leu Asn Lys Ser Val Val Ile Asn Cys Thr

290 295 300

Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Ala Gly Lys

305 310 315 320

Ala Leu Tyr Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys

325 330 335

Asn Leu Ser Arg Ala Gln Trp Asn Asn Thr Leu Lys Gln Ile Val Ile

340 345 350

Lys Leu Arg Glu Gln Phe Gly Asn Lys Thr Ile Val Phe Asn Gln Ser

355 360 365

Ser Gly Gly Asp Val Glu Ile Val Met His Ser Phe Asn Cys Gly Gly

370 375 380

Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn

385 390 395 400
 Gly Asn Asp Thr Trp Asn Asp Thr Trp Lys Asp Thr Thr Asn Asp Asn
 405 410 415
 Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Val Asn Met Trp Gln Lys
 420 425 430
 Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys

 435 440 445
 Ser Ser Lys Ile Thr Gly Leu Ile Leu Thr Arg Asp Gly Gly Thr Asn
 450 455 460
 Gly Thr Asn Glu Thr Glu Thr Phe Arg Pro Gly Gly Gly Asn Met Lys
 465 470 475 480
 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu
 485 490 495
 Pro Leu Gly Ile Ala Pro Thr Lys Ala Lys Thr Leu Thr Val Gln Ala
 500 505 510

 Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg
 515 520 525
 Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile
 530 535 540
 Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Asp
 545 550 555 560
 Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Arg Leu Ile Cys Thr
 565 570 575
 Thr Asn Val Pro Trp Asn Thr Ser Trp Ser Asn Arg Ser Leu Asn Tyr

 580 585 590
 Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile Asn Asn
 595 600 605
 Tyr Thr Asp Tyr Ile Tyr Thr Leu Leu Glu Asp Ala Gln Asn Gln Gln
 610 615 620
 Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu
 625 630 635 640
 Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp

645

650

<210> 55

<211> 1984

<212> DNA

<213> Human immunodeficiency virus

<400> 55

ttcagtcgac agccaccatg agagtaaaag aatcagacg caactgtcag aggttgagga 60
gatggggaac gatgctcctg ggcatgctga tgatttcag tgccaccgaa cagctttggg 120
taaccgtgta ctatggtgta cctgtatgga aagaagccac tacaaccctg ttttgcgcgt 180
ccgacgcaaa agcctacgta acagaaaagc acaacgtgtg ggccacacat gcatgcgtgc 240
caacagatcc aaatcctcag gaagtcgttc tggaaaatgt aacagaaaat tttaatatgt 300
ggaaaaacaa tatggtagag cagatgcatg aagatatcat ctactgtgg gaacaatcct 360

tgaaaccttg tgtcaaactg accccacttt gcgtaacact taactgtact gataagcttc 420
gcaatgatac gtccggaaca aattcaagca gctgggaaaa agtgcaaaag ggcgaaatca 480
aaaattgttc atttaacatc actaccgta tcagagggcg ggtacaggaa tattctcttt 540
tctacaaact cgacgtcacc ccaatcgact ccagaaataa ctcaataat agcacagaat 600
ttagtagtta tcgccttata agctgcaaca ccagcgtgat tacacaagcg tgccctaaaa 660
tctcttttga gccatttctt attcactact gcgcaccagc cggcttcgcc atcctcaaat 720
gtaacgacaa gaaatttaac ggaaccggac cctgtaagaa tgtgtccacc gttcaatgca 780

ctcatggaat caagcccgtc gtttctaccc aacttcttct caatggtagc cttgcggagg 840
aggaagtgtg gattcgctcc gaaaatttta caaacaacgt caagtcaatc atcgtccagc 900
ttaataaatc cgtcgttatt aattgtacaa gaccaacaa taacaccaga aaatccattc 960
acataggggc cgggaaagct ctgtataccg gggaaattat tggagacatc agacaagcac 1020
actgtaactt gagtcgcgcc cagtggaaca acacattgaa acagatcgtg atcaagctca 1080
gagagcagtt cgggaataag actatcgtgt ttaatcagag ctccggcggt gatgtcgaaa 1140
tcgtaatgca ctcttttaat tgtgggggtg aattttttta ctgcaattct acacaattgt 1200

ttaacagcac ctggaacggc aatgacacat ggaatgacac ctggaaagat acgacaaatg 1260
ataatattac tcttccgtgc agaataaagc aaatcgtaaa tatgtggcaa aaagtgggca 1320
aggccatgta cgcaccacct ataagaggac aaattcgtgt tctttccaag atcacaggtc 1380
tgatactcac acgggacgga ggcacgaacg ggacaaacga gaccgagacc ttccgaccag 1440
gaggcggcaa catgaaggat aactggagaa gtgaacttta caagtataaa gtggtcaaga 1500

ttgagcctct gggatcgcc cctactaagg ctaaaacact caccgtgcag gctagattgc 1560
 tgctttcagg gatagtccaa caacagaaca accttcttag agccattgaa gcacaacaac 1620

acttgcgtga gttgacagtg tggggaatta aacagttgca ggcccgggtt ctgcgtgtcg 1680
 aacggatatct tagagatcag cagcttttgg gtatctgggg gtgttcaggc cgcctcatat 1740
 gcaccacaaa tgtcccttgg aatacctcat ggagtaacag gtctcttaat tatatttggg 1800
 acaatatgac atggatgcaa tgggatagag aaattaataa ctacaccgac tacatctaca 1860
 cacttctgga ggacgccag aatcagcagg agaagaacga gcaggaactc ctgcaattgg 1920
 ataagtgggc atcactgtgg aattgggttcg atataactaa ttggctttgg taaagatctt 1980
 acaa 1984

<210> 56
 <211> 854
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 56

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

130 135 140
 Tyr Asn Asn Ser Ile Asp Ser Met Asn Gly Glu Ile Lys Asn Cys Ser
 145 150 155 160
 Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys Gln Lys Val Tyr Ala
 165 170 175
 Leu Phe Tyr Arg Pro Asp Val Val Pro Leu Asn Glu Asn Ser Ser Ser
 180 185 190
 Tyr Ile Leu Ile Asn Cys Asn Thr Ser Thr Thr Thr Gln Ala Cys Pro

 195 200 205
 Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly
 210 215 220
 Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro
 225 230 235 240
 Cys His Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val
 245 250 255
 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile
 260 265 270

 Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile Val
 275 280 285
 His Leu Asn Glu Ser Ile Glu Ile Val Cys Thr Arg Pro Asn Asn Asn
 290 295 300
 Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Thr Val Tyr Ala Thr
 305 310 315 320
 Asn Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Lys
 325 330 335
 Thr Lys Trp Asn Thr Thr Leu Glu Lys Val Lys Glu Lys Leu Lys Glu

 340 345 350
 His Phe Pro Ser Lys Ala Ile Thr Phe Gln Pro His Ser Gly Gly Asp
 355 360 365
 Leu Glu Val Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr
 370 375 380
 Cys Asp Thr Thr Lys Leu Phe Asn Glu Ser Asn Leu Asn Thr Thr Asn

385 390 395 400
 Thr Thr Thr Leu Thr Leu Pro Cys Arg Ile Lys Gln Ile Val Asn Met
 405 410 415

 Trp Gln Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Val Glu Gly Asn
 420 425 430
 Ile Thr Cys Asn Ser Ser Ile Thr Gly Leu Leu Leu Val Arg Asp Gly
 435 440 445
 Gly Asn Thr Ser Asn Ser Thr Pro Glu Ile Phe Arg Pro Gly Gly Gly
 450 455 460
 Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
 465 470 475 480
 Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val

 485 490 495
 Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Leu Phe Gly
 500 505 510
 Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu
 515 520 525
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 530 535 540
 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr
 545 550 555 560

 Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu Arg
 565 570 575
 Tyr Leu Lys Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys
 580 585 590
 Leu Ile Cys Pro Thr Thr Val Pro Trp Asn Ser Ser Trp Ser Asn Lys
 595 600 605
 Ser Gln Thr Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg
 610 615 620
 Glu Ile Ser Asn Tyr Thr Gly Thr Ile Tyr Lys Leu Leu Glu Glu Ser

 625 630 635 640

Gln Asn Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser
645 650 655

Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr
660 665 670

Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile
675 680 685

Ile Phe Gly Val Leu Ser Ile Val Lys Arg Val Arg Gln Gly Tyr Ser
690 695 700

Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Leu Asp Arg
705 710 715 720

Leu Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp Lys Asp Arg Ser
725 730 735

Ile Arg Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg
740 745 750

Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Val
755 760 765

Ala Ala Arg Ala Val Glu Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu
770 775 780

Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Asn Leu Val Gln Tyr
785 790 795 800

Gly Gly Leu Glu Leu Lys Arg Arg Ala Ile Ser Leu Phe Asp Thr Ile
805 810 815

Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Leu Glu Val Ile Leu
820 825 830

Arg Ile Ile Arg Ala Ile Arg Asn Ile Pro Thr Arg Ile Arg Gln Gly
835 840 845

Phe Glu Ala Ala Leu Leu
850

<210> 57
<211> 638
<212> PRT
<213> Human immunodeficiency virus

<400> 57

Met Arg Val Lys Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Val Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Thr Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Arg Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Ile Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Val Lys Val Asn Ala Thr Ser Asn Gly Thr Thr Thr

130 135 140

Tyr Asn Asn Ser Ile Asp Ser Met Asn Gly Glu Ile Lys Asn Cys Ser

145 150 155 160

Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys Gln Lys Val Tyr Ala

165 170 175

Leu Phe Tyr Arg Pro Asp Val Val Pro Leu Asn Glu Asn Ser Ser Ser

180 185 190

Tyr Ile Leu Ile Asn Cys Asn Thr Ser Thr Thr Thr Gln Ala Cys Pro

195 200 205

Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly

210 215 220

Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro

225 230 235 240

Cys His Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val

245 250 255
 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile

 260 265 270
 Ile Ile Arg Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile Val
 275 280 285
 His Leu Asn Glu Ser Ile Glu Ile Val Cys Thr Arg Pro Asn Asn Asn
 290 295 300
 Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Thr Val Tyr Ala Thr
 305 310 315 320
 Asn Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Lys
 325 330 335

 Thr Lys Trp Asn Thr Thr Leu Glu Lys Val Lys Glu Lys Leu Lys Glu
 340 345 350
 His Phe Pro Ser Lys Ala Ile Thr Phe Gln Pro His Ser Gly Gly Asp
 355 360 365
 Leu Glu Val Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr
 370 375 380
 Cys Asp Thr Thr Lys Leu Phe Asn Glu Ser Asn Leu Asn Thr Thr Asn
 385 390 395 400
 Thr Thr Thr Leu Thr Leu Pro Cys Arg Ile Lys Gln Ile Val Asn Met

 405 410 415
 Trp Gln Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Val Glu Gly Asn
 420 425 430
 Ile Thr Cys Asn Ser Ser Ile Thr Gly Leu Leu Leu Val Arg Asp Gly
 435 440 445
 Gly Asn Thr Ser Asn Ser Thr Pro Glu Ile Phe Arg Pro Gly Gly Gly
 450 455 460
 Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
 465 470 475 480

 Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Thr Leu Thr
 485 490 495

Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn
500 505 510
Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val
515 520 525
Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu Arg Tyr
530 535 540
Leu Lys Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu

545 550 555 560
Ile Cys Pro Thr Thr Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser
565 570 575
Gln Thr Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu
580 585 590
Ile Ser Asn Tyr Thr Gly Thr Ile Tyr Lys Leu Leu Glu Glu Ser Gln
595 600 605
Asn Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp
610 615 620

Lys Asn Leu Trp Ser Trp Phe Asp Ile Thr Asn Trp Leu Trp
625 630 635

<210> 58
<211> 1945
<212> DNA
<213> Human immunodeficiency virus
<400> 58

ttcagtcgac agccaccatg cgcgtaaagg ggattcaaag aaattggccg caatggtgga	60
tttggggaat tctgggcttt tggatgataa ttatatgccg cggtgtcgga aatttgtggg	120
tgactgtgta ctacgggggtg cccgtgtgga ctgaggcaaa gaccaccctg ttctgtgcta	180
gcgatgccaa agcctatgaa gcggaagtgc acaatgtttg ggctactcat gcctgtgtcc	240
ctaccgacce aaaccctcag gaaatagtgc tcggcaatgt aacggaaaac ttcaacatgt	300
ggaaaaatga tatggtggat cagatgcacg aagacattat ctcaatctgg gaccaaagcc	360
tgaaaccctg cgttaaactg actcctctct gcgtcactct caattgcaca gatgtcaaag	420
tgaatgccac ctcaaacggt acgacaactt acaacaattc tattgactct atgaacggcg	480
aaatcaaaaa ttgttccttt aacatcacca ccgagatacg cgacaaaaag cagaaggtct	540

atgccctttt ttaccgcccc gacgtagtcc cactcaacga gaattccagc tcatacatcc 600
tcatcaactg caatacatca actaccacac aagcatgccc gaaagttagc ttgatccaa 660

ttcctataca ttactgcgcc cccgccggt acgtataact gaaatgcaat aataagactt 720
ttaacgggac cggcccatgt cacaacgtgt caaccgtgca atgcactcat ggcatcaagc 780
ccgtggtgtc aaccagctg ctgctcaatg gctcacttgc agaagaagaa attattatcc 840
gctctgagaa tcttactaac aatgcaaaaa cgattatcgt gcaccttaat gaatcaatag 900
aaatcgtgtg tactcggccc aacaataata ctgaaaaag cattcgcac ggacctggcc 960
agacagttta cgcaactaat gacatcatcg gggacatccg acaggcccat tgcaacattt 1020
ctaaaaccaa gtggaataca accctggaaa aagtaaagga aaaacttaaa gaacattttc 1080

cctctaaggc gatcacgttt caacctcaca gtggcggaga cttggaagtc acaacacatt 1140
cttttaactg ccgcggagaa tttttttatt gtgatacaac aaaacttttt aatgaatcaa 1200
atctcaacac cacaatata accacactga cctcccttg tagaatcaaa caaatcgtaa 1260
acatgtggca aggggttga agggctatgt acgtccccc cgtcgaagga aatataacgt 1320
gtaacagcag catcactggg ctgcttcttg ttcgagacgg aggcaatact tetaattcaa 1380
ctcctgaaat ttttaggcct ggcgttgga atatgaaaga taactggcgc tcagaactgt 1440
acaaatacaa agttgttgaa attaagcccc tgggagtcgc tccaacaaa gctaaaacac 1500

tcacagtga agcaagacag ctctttcag gcatcgtcca gcaacagtca aatctcctta 1560
gagcaatcga agcccaacag catatgctcc aactcacagt ctgggggatt aaacagcttc 1620
aagcccgct gcttgcctac gaacgtatc ttaaagacca acagcttctt ggcctctggg 1680
gtttagtggtg aaaactcatc tgccccacca ccgtgccttg gaatagttct tggagtaata 1740
aatcacagac cgatatttgg gacaacatga cctggatgca atgggatagg gaaatttcta 1800
attatactgg cacaatctac aaactcttgg aagaaagtca aaatcagcaa gaaaaaacg 1860
aaaaggacct cctgccttg gactcctgga agaactttg gagctggttc gacataacta 1920

attggctgtg gtaaagatct taaa 1945

- <210> 59
- <211> 854
- <212> PRT
- <213> Human immunodeficiency virus
- <400> 59

Met Arg Val Lys Glu Thr Gln Met Asn Trp Pro Asn Leu Trp Lys Trp

1 5 10 15

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

260 265 270
 Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn
 275 280 285
 Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Ser Asn Asn Thr Arg Thr
 290 295 300
 Ser Ile Thr Met Gly Pro Gly Gln Val Phe Tyr Arg Thr Gly Asp Ile
 305 310 315 320
 Ile Gly Asp Ile Arg Lys Ala Tyr Cys Glu Ile Asn Gly Ile Lys Trp
 325 330 335

 Asn Glu Val Leu Val Gln Val Thr Gly Lys Leu Lys Glu His Phe Asn
 340 345 350
 Lys Thr Ile Ile Phe Gln Pro Pro Ser Gly Gly Asp Leu Glu Ile Ile
 355 360 365
 Thr His His Phe Ser Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Thr
 370 375 380
 Lys Leu Phe Asn Asn Thr Cys Ile Gly Asn Thr Ser Met Glu Gly Cys
 385 390 395 400
 Asn Asn Thr Ile Ile Leu Pro Cys Lys Ile Lys Gln Ile Ile Asn Met

 405 410 415
 Trp Gln Gly Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Ser Gly Arg
 420 425 430
 Ile Asn Cys Val Ser Asn Ile Thr Gly Ile Leu Leu Thr Arg Asp Gly
 435 440 445
 Gly Ala Asp Asn Asn Thr Thr Asn Glu Thr Phe Arg Pro Gly Gly Gly
 450 455 460
 Asn Ile Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
 465 470 475 480

 Glu Ile Glu Pro Leu Gly Ile Ala Pro Thr Arg Ala Lys Arg Arg Val
 485 490 495
 Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Met Ile Phe Gly
 500 505 510
 Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu

515 520 525
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 530 535 540
 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr

 545 550 555 560
 Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg
 565 570 575
 Tyr Leu Lys Asp Gln Lys Phe Leu Gly Leu Trp Gly Cys Ser Gly Lys
 580 585 590
 Ile Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys
 595 600 605
 Ser Phe Glu Glu Ile Trp Asp Asn Met Thr Trp Ile Glu Trp Glu Arg
 610 615 620

 Glu Ile Ser Asn Tyr Thr Ser Gln Ile Tyr Glu Ile Leu Thr Glu Ser
 625 630 635 640
 Gln Asn Gln Gln Asp Arg Asn Glu Lys Asp Leu Leu Glu Leu Asp Lys
 645 650 655
 Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr
 660 665 670
 Ile Lys Ile Phe Ile Ile Ile Val Gly Ser Leu Ile Gly Leu Arg Ile
 675 680 685
 Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser

 690 695 700
 Pro Leu Ser Phe Gln Thr Pro Thr His His Gln Arg Glu Pro Asp Arg
 705 710 715 720
 Pro Glu Glu Ile Gly Glu Gly Gly Gly Glu Gln Ser Lys Asp Arg Ser
 725 730 735
 Val Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg
 740 745 750
 Ser Leu Cys Leu Phe Ser Tyr His Leu Leu Arg Asp Phe Ile Leu Ile
 755 760 765

Ala Ala Arg Thr Val Glu Leu Leu Gly His Ser Ser Leu Lys Gly Leu
770 775 780

Arg Arg Gly Trp Glu Gly Leu Lys Tyr Leu Gly Asn Leu Leu Leu Tyr
785 790 795 800

Trp Gly Gln Glu Ile Lys Ile Ser Ala Ile Ser Leu Leu Asn Ala Thr
805 810 815

Ala Ile Ala Val Ala Gly Trp Thr Asp Arg Val Ile Glu Val Ala Gln
820 825 830

Arg Ala Trp Arg Ala Leu Leu His Ile Pro Arg Arg Ile Arg Gln Gly
835 840 845

Leu Glu Arg Ala Leu Leu
850

<210> 60
<211> 630
<212> PRT
<213> Human immunodeficiency virus
<400> 60

Met Arg Val Lys Glu Thr Gln Met Asn Trp Pro Asn Leu Trp Lys Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asp Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Asp
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala His Glu Thr Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile His Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Arg
85 90 95

Asn Asn Met Val Glu Gln Met Gln Glu Asp Val Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125
 Asn Cys Thr Asn Ala Asn Trp Thr Asn Ser Asn Asn Thr Thr Asn Gly
 130 135 140
 Pro Asn Lys Ile Gly Asn Ile Thr Asp Glu Val Lys Asn Cys Thr Phe
 145 150 155 160
 Asn Met Thr Thr Glu Leu Lys Asp Lys Lys Gln Lys Val His Ala Leu
 165 170 175
 Phe Tyr Lys Leu Asp Ile Val Gln Ile Asn Ser Ser Glu Tyr Arg Leu

 180 185 190
 Ile Asn Cys Asn Thr Ser Val Ile Lys Gln Ala Cys Pro Lys Ile Ser
 195 200 205
 Phe Asp Pro Ile Pro Ile His Tyr Cys Thr Pro Ala Gly Tyr Ala Ile
 210 215 220
 Leu Lys Cys Asn Asp Lys Asn Phe Asn Gly Thr Gly Pro Cys Lys Asn
 225 230 235 240
 Val Ser Ser Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr
 245 250 255

 Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg
 260 265 270
 Ser Glu Asn Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn
 275 280 285
 Lys Ser Val Glu Ile Asn Cys Thr Arg Pro Ser Asn Asn Thr Arg Thr
 290 295 300
 Ser Ile Thr Met Gly Pro Gly Gln Val Phe Tyr Arg Thr Gly Asp Ile
 305 310 315 320
 Ile Gly Asp Ile Arg Lys Ala Tyr Cys Glu Ile Asn Gly Ile Lys Trp

 325 330 335
 Asn Glu Val Leu Val Gln Val Thr Gly Lys Leu Lys Glu His Phe Asn
 340 345 350
 Lys Thr Ile Ile Phe Gln Pro Pro Ser Gly Gly Asp Leu Glu Ile Ile
 355 360 365
 Thr His His Phe Ser Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Thr

370 375 380
 Lys Leu Phe Asn Asn Thr Cys Ile Gly Asn Thr Ser Met Glu Gly Cys
 385 390 395 400

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

 465 470 475 480
 Glu Ile Glu Pro Leu Gly Ile Ala Pro Thr Arg Ala Arg Thr Leu Thr
 485 490 495
 Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn
 500 505 510
 Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val
 515 520 525
 Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr
 530 535 540

 Leu Lys Asp Gln Lys Phe Leu Gly Leu Trp Gly Cys Ser Gly Lys Ile
 545 550 555 560
 Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser
 565 570 575
 Phe Glu Glu Ile Trp Asp Asn Met Thr Trp Ile Glu Trp Glu Arg Glu
 580 585 590
 Ile Ser Asn Tyr Thr Ser Gln Ile Tyr Glu Ile Leu Thr Glu Ser Gln
 595 600 605
 Asn Gln Gln Asp Arg Asn Glu Lys Asp Leu Leu Glu Leu Asp Lys Trp

 610 615 620

Ala Ser Leu Trp Asn Trp

625 630

<210> 61

<211> 1921

<212> DNA

<213> Human immunodeficiency virus

<400> 61

ttcagtcgac agccaccatg cgagtaaaag agacacaaat gaattggccc aattttgtgga 60

agtggggaac attgatcctg ggactggtga taatctgtag tgcacccgac aatctctggg 120

tgaccgttta ctatggtgta ccagtttggg gagacgctga taccaccctc ttctgtgcaa 180

gcgacgccaa agcccacgaa actgaagtcc ataatgtatg ggccaccac gcgtgcgtac 240

caaccgaccc taatcccca gagatccacc ttgagaatgt aactgagaat tttaacatgt 300

ggagaaataa catggtggaa caaatgcagg aagacgttat ttccttgtgg gaccagagcc 360

ttaaaccctt tgtaaatgt actccctgt gtgtgactct caattgtaca aacgcaaatt 420

ggaccaacag caacaacact accaacggcc ctaacaaaat tggcaatatt actgatgaag 480

tcaagaactg cacttttaac atgacaacag aactgaagga taagaaacag aaagtccatg 540

ctctgttcta taagctcgac atagtacaaa ttaatagctc agaatataga ctgataaact 600

gcaatacttc cgttatcaaa caggcctgtc caaagataag cttcgatccc atccctatc 660

actactgcac accagccggt tacgtatcc tgaatgcaa cgataagaat tttaacggca 720

caggtccctg caaaaacgtt tcctctgtcc agtgtacaca cggatatcaag cctgtagtat 780

caacacaact gtcctgaat ggctccttgg ccgaagaaga gatcatcatt agaagtgaga 840

acctgacgaa caacgccaag actataatag tgcacctcaa taaatctgta gaaatcaact 900

gtaccgacc ctcaaacaac actcgaacaa gtataacaat gggccctggc caagtttttt 960

accggaccgg cgacataata ggcgatatca gaaaggcata ttgcgagatc aatggcatca 1020

agtggaacga agtactggtt caagtaactg gaaaactcaa agaacatttt aataagacca 1080

taatatcca gccccgagt ggcgcgacc tcgagattat caccatcac ttttcttgta 1140

gaggcgaatt ttttactgt aacacgacca agctcttcaa taacacgtgc atcggaaca 1200

cttctatgga aggatgtaat aatacattta tactgcctg taagatcaag cagattatca 1260

acatgtggca gggagtaggt caggcaatgt acgcaccacc gatttcagga cggatcaatt 1320

gcgtatcaaa tatcaccggc attctgctga cccgggacgg aggcgcagac aacaatacca 1380

ctaacgagac atttagacct ggaggcggca atataaagga taattggaga agtgagctgt 1440

ataaatacaa agtcgtagag atcgaacccc tcggcattgc tccaacccgg gcccgactc 1500

tcaccgtaca agctagacag ctgctttctg gcatagtcca acagcagtca aacctcctcc 1560

gcgctattga agcacaacaa cacctgctcc agctgactgt gtggggaatc aaacaattgc 1620

aagcaagagt gctcgccgtg gaacgctatt tgaaagatca gaaatttctt ggactttggg 1680

gctgcagcgg caaaattatt tgtacaacag cggcgccttg gaactcatcc tggagtaata 1740

aaagctttga agaaatctgg gacaatatga catggattga gtgggagaga gagatttcaa 1800

actatacaag ccaaatttac gaaatactga cagaaagtca aaaccagcag gacagaaatg 1860

agaaagacct gctcgaactg gataagtggg cctctttgtg gaactggtaa agatcttaca 1920

a 1921

<210> 62

<211> 854

<212> PRT

<213> Human immunodeficiency virus

<400> 62

Met Arg Val Lys Gly Ile Gln Arg Asn Trp Gln His Leu Trp Asn Trp

1 5 10 15

Gly Ile Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Glu Lys Leu

20 25 30

Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Glu Asp Ala Asn Ala

35 40 45

Pro Leu Phe Cys Ala Ser Asp Ala Lys Ala His Ser Thr Glu Ser His

50 55 60

Asn Ile Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Ser Pro Gln

65 70 75 80

Glu Ile Asn Met Arg Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn

85 90 95

Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Glu

100 105 110

Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn

115 120 125

Cys Thr Glu Ile Asn Asn Asn Ser Thr Arg Asn Ile Thr Glu Glu Tyr

130 135 140
 Arg Met Thr Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys
 145 150 155 160
 Lys Lys Ala Glu Tyr Ala Leu Phe Tyr Arg Thr Asp Val Val Pro Ile
 165 170 175
 Asn Glu Met Asn Asn Glu Asn Asn Gly Thr Asn Ser Thr Trp Tyr Arg
 180 185 190

 Leu Thr Asn Cys Asn Val Ser Thr Ile Lys Gln Ala Cys Pro Lys Val
 195 200 205
 Thr Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala
 210 215 220
 Ile Leu Lys Cys Val Asp Lys Lys Phe Asn Gly Thr Gly Thr Cys Asn
 225 230 235 240
 Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser
 245 250 255
 Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Asp Ile Ile Ile

 260 265 270
 Ser Ser Glu Asn Ile Ser Asp Asn Ala Lys Val Ile Ile Val His Leu
 275 280 285
 Asn Arg Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg
 290 295 300
 Arg Ser Val Ala Ile Gly Pro Gly Gln Ala Phe Tyr Thr Thr Gly Glu
 305 310 315 320
 Val Ile Gly Asp Ile Arg Lys Ala His Cys Asn Val Ser Trp Thr Lys
 325 330 335

 Trp Asn Glu Thr Leu Arg Asp Val Gln Ala Lys Leu Gln Glu Tyr Phe
 340 345 350
 Ile Asn Lys Ser Ile Glu Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu
 355 360 365
 Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn
 370 375 380
 Thr Ser Gly Leu Phe Asn Asn Ser Ile Leu Lys Ser Asn Ile Ser Glu

385 390 395 400
 Asn Asn Asp Thr Ile Thr Leu Asn Cys Lys Ile Lys Gln Ile Val Arg

 405 410 415
 Met Trp Gln Arg Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Ala Gly
 420 425 430
 Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu Ile Leu Thr Arg Asp
 435 440 445
 Gly Gly Asp Asn Asn Ser Thr Ser Glu Ile Phe Arg Pro Gly Gly Gly
 450 455 460
 Asp Met Lys Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Thr Val
 465 470 475 480

 Lys Ile Lys Ser Leu Gly Ile Ala Pro Thr Arg Ala Arg Arg Arg Val
 485 490 495
 Val Glu Arg Glu Lys Arg Ala Val Gly Val Gly Ala Ile Phe Leu Gly
 500 505 510
 Phe Leu Gly Thr Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu
 515 520 525
 Thr Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 530 535 540
 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr

 545 550 555 560
 Val Trp Gly Ile Lys Gln Leu Arg Ala Arg Val Leu Ala Leu Glu Arg
 565 570 575
 Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys
 580 585 590
 Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Thr Ser Trp Ser Asn Lys
 595 600 605
 Ser Tyr Asn Glu Ile Trp Glu Asn Met Thr Trp Ile Glu Trp Glu Arg
 610 615 620

 Glu Ile Asp Asn Tyr Thr Tyr His Ile Tyr Ser Leu Ile Glu Gln Ser
 625 630 635 640

Gln Ile Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp Gln
645 650 655

Trp Ala Ser Leu Trp Ser Trp Phe Ser Ile Ser Asn Trp Leu Trp Tyr
660 665 670

Ile Arg Ile Phe Val Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile
675 680 685

Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser
690 695 700

Pro Leu Ser Phe Gln Thr Leu Leu His His Gln Arg Glu Pro Asp Arg
705 710 715 720

Pro Ala Gly Ile Glu Glu Gly Gly Gly Glu Gln Asp Arg Asp Arg Ser
725 730 735

Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg
740 745 750

Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile
755 760 765

Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Asn Ser Leu Lys Gly Leu
770 775 780

Arg Leu Gly Trp Glu Ala Leu Lys Tyr Leu Trp Asn Leu Leu Leu Tyr
785 790 795 800

Trp Ala Arg Glu Leu Lys Asn Ser Ala Ile Asn Leu Leu Asp Thr Ile
805 810 815

Ala Ile Ala Val Ala Asn Trp Thr Asp Arg Val Ile Glu Val Ala Gln
820 825 830

Arg Ala Gly Arg Ala Val Leu Asn Ile Pro Arg Arg Ile Arg Gln Gly
835 840 845

Leu Glu Arg Ala Leu Leu
850

<210> 63
<211> 630
<212> PRT
<213> Human immunodeficiency virus

<400> 63

Met Arg Val Lys Gly Ile Gln Arg Asn Trp Gln His Leu Trp Asn Trp

1 5 10 15

Gly Ile Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Glu Lys Leu

20 25 30

Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Glu Asp Ala Asn Ala

35 40 45

Pro Leu Phe Cys Ala Ser Asp Ala Lys Ala His Ser Thr Glu Ser His

50 55 60

Asn Ile Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Ser Pro Gln

65 70 75 80

Glu Ile Asn Met Arg Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn

85 90 95

Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Glu

100 105 110

Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn

115 120 125

Cys Thr Glu Ile Asn Asn Asn Ser Thr Arg Asn Ile Thr Glu Glu Tyr

130 135 140

Arg Met Thr Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys

145 150 155 160

Lys Lys Ala Glu Tyr Ala Leu Phe Tyr Arg Thr Asp Val Val Pro Ile

165 170 175

Asn Glu Met Asn Asn Glu Asn Asn Gly Thr Asn Ser Thr Trp Tyr Arg

180 185 190

Leu Thr Asn Cys Asn Val Ser Thr Ile Lys Gln Ala Cys Pro Lys Val

195 200 205

Thr Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala

210 215 220

Ile Leu Lys Cys Val Asp Lys Lys Phe Asn Gly Thr Gly Thr Cys Asn

225 230 235 240

Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser

245 250 255
 Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Asp Ile Ile Ile
 260 265 270
 Ser Ser Glu Asn Ile Ser Asp Asn Ala Lys Val Ile Ile Val His Leu
 275 280 285
 Asn Arg Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg
 290 295 300
 Arg Ser Val Ala Ile Gly Pro Gly Gln Ala Phe Tyr Thr Thr Gly Glu
 305 310 315 320
 Val Ile Gly Asp Ile Arg Lys Ala His Cys Asn Val Ser Trp Thr Lys
 325 330 335
 Trp Asn Glu Thr Leu Arg Asp Val Gln Ala Lys Leu Gln Glu Tyr Phe
 340 345 350
 Ile Asn Lys Ser Ile Glu Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu
 355 360 365
 Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn
 370 375 380
 Thr Ser Gly Leu Phe Asn Asn Ser Ile Leu Lys Ser Asn Ile Ser Glu
 385 390 395 400
 Asn Asn Asp Thr Ile Thr Leu Asn Cys Lys Ile Lys Gln Ile Val Arg
 405 410 415
 Met Trp Gln Arg Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Ala Gly
 420 425 430
 Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu Ile Leu Thr Arg Asp
 435 440 445
 Gly Gly Asp Asn Asn Ser Thr Ser Glu Ile Phe Arg Pro Gly Gly Gly
 450 455 460
 Asp Met Lys Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Thr Val
 465 470 475 480
 Lys Ile Lys Ser Leu Gly Ile Ala Pro Thr Arg Ala Arg Thr Leu Thr
 485 490 495

Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn
500 505 510
Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val
515 520 525
Trp Gly Ile Lys Gln Leu Arg Ala Arg Val Leu Ala Leu Glu Arg Tyr
530 535 540

Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu
545 550 555 560
Ile Cys Thr Thr Asn Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser
565 570 575
Tyr Asn Glu Ile Trp Glu Asn Met Thr Trp Ile Glu Trp Glu Arg Glu
580 585 590
Ile Asp Asn Tyr Thr Tyr His Ile Tyr Ser Leu Ile Glu Gln Ser Gln
595 600 605
Ile Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp Gln Trp

610 615 620

Ala Ser Leu Trp Ser Trp

625 630

<210> 64

<211> 1921

<212> DNA

<213> Human immunodeficiency virus

<400> 64

ttcagtcgac agccaccatg agagttaaag gaatccaacg caattggcaa cacctttgga 60

actggggcat attgattctt ggactggtga taattttag cgctgaaaaa ctctgggtaa 120

ctgtctatta cggcgtgcct gtctgggagg atgccaacgc cccctgttc tgcgcaagtg 180

atgcaaaggc tcacagcact gaatctcaca acatttgggc caccacgcc tgtgtgccaa 240

ccgaccctag tcctcaggag atcaacatga gaaacgttac cgaaaatttt aatatgtgga 300

agaataatat ggtggagcaa atgcacgaag acataatttc actctgggac gagtctctga 360

aacctgtgt gaaacttacc cccctgtgcg tcacctgaa ctgtaccgaa atcaacaata 420

actcaacgag aaatatcaca gaagaatacc gaatgactaa ctgttccttt aatatgacaa 480

ccgaactgcg agacaaaaag aaggctgaat acgcactttt ctaccgaaca gatgtgttac 540

caatcaacga gatgaacaat gaaaacaatg gaacgaactc tacctggtat agactgacaa 600
actgtaacgt tagcacaatc aagcaggcct gccctaaagt cacattcgaa ccaataccaa 660

ttcactactg cgcacccgcc ggattcgcta ttcttaagtg cgtggataag aagtttaacg 720
gaactggaac ctgcaataat gtatctacag tacaatgcac gcatggaatt aagcctgtcg 780
tttcaaccca gttgctgctg aatggatcac tcgcagaaaa ggatattatt atctcaagcg 840
aaaacatata tgataatgca aaggatcatca tcgtccacct caaccgctca gttgaaataa 900
actgcactcg gcctaataat aacacaagac gctctgtcgc aatcggccca ggacaagctt 960
tttactactac cggggaagtt atcggcgaca tacggaaagc ccactgcaac gttagctgga 1020
ccaagtggaa tgaaacactg cgcgatgttc aagccaaact tcaagaatac ttcataaaca 1080

aatcaattga gttcaattct agctctggcg gcgacctga gattacaact cactccttta 1140
actgcggcgg cgaattcttt tattgtaata cctccgtct cttcaacaac tctatcctca 1200
aaagtaacat ttctgaaaat aatgacacaa tcacactgaa ttgcaagatc aagcagattg 1260
ttaggatgtg gcaacgagtc ggacaagcta tgtacgcccc acccatcgcc ggaaatataa 1320
cgtgtcgatc aaatatcact ggcctcatcc ttactagaga tggcggagac aataatagca 1380
ccagcgagat attcagacca ggcggaggcg atatgaaaaa caactggagg tcagagctct 1440
acaagtacaa aacagtcaaa attaaaagcc tgggcattgc tcccactcgg gcccgcacac 1500

tgactgtcca agtccgacag ctctgtccg gaatcgcca acaacagtcc aacttgctgc 1560
gcgctataga ggctcaacaa catctcttc aactgactgt gtggggatc aaacaattga 1620
gagcaagagt gctggcgtg gaacggatc ttaaggacca acaactcctg ggcataatggg 1680
ggtgttccgg caaactgac tgcacaacaa atgtacctg gaacaccagc tgggtcaaata 1740
aaagtataa tgagatatgg gaaaacatga catggattga atgggaaagg gaaattgaca 1800
attatacata ccatatatac tctctcatcg aacaatctca gatacaacag gaaaagaatg 1860
aacaagattt gttggctctt gaccaatggg ctctttgtg gagttggtaa agatcttaca 1920

a 1921

- <210> 65
- <211> 829
- <212> PRT
- <213> Human immunodeficiency virus
- <400> 65

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

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

260 265 270
 Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg
 275 280 285
 Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Ala
 290 295 300
 Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys
 305 310 315 320
 Asn Ile Ser Arg Thr Lys Trp Asn Lys Thr Leu Gln Gln Val Ala Lys
 325 330 335

 Lys Leu Arg Glu His Phe Asn Lys Thr Ile Ile Phe Asn Pro Ser Ser
 340 345 350
 Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly Glu
 355 360 365
 Phe Phe Tyr Cys Asn Thr Ser Glu Leu Phe Asn Ser Thr Trp Asn Gly
 370 375 380
 Thr Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn
 385 390 395 400
 Met Trp Gln Gly Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly

 405 410 415
 Lys Ile Arg Cys Thr Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp
 420 425 430
 Gly Gly Asn Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met
 435 440 445
 Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
 450 455 460
 Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Glu
 465 470 475 480

 Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
 485 490 495
 Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val
 500 505 510
 Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu

515 520 525
 Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
 530 535 540
 Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu

 545 550 555 560
 Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
 565 570 575
 Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln
 580 585 590
 Asp Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Asp Lys Glu Ile
 595 600 605
 Asn Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln Asn
 610 615 620

 Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp Lys Trp Ala
 625 630 635 640
 Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys
 645 650 655
 Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe
 660 665 670
 Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu
 675 680 685
 Ser Phe Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Asp Arg Pro Glu

 690 695 700
 Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile Arg
 705 710 715 720
 Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu
 725 730 735
 Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu Ile Ala Ala
 740 745 750
 Arg Thr Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr
 755 760 765

Leu Trp Asn Leu Leu Gln Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala
 770 775 780
 Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val Ala Glu Gly Thr Asp
 785 790 795 800
 Arg Val Ile Glu Val Val Gln Arg Val Cys Arg Ala Ile Leu Asn Ile
 805 810 815
 Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu
 820 825
 <210> 66
 <211> 830

 <212> PRT
 <213> Human immunodeficiency virus
 <400> 66
 Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
 1 5 10 15
 Gly Ile Leu Ile Phe Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn
 20 25 30
 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn
 35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50 55 60

 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80
 Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
 85 90 95
 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asp Val Asn Ala Thr Asn Asn Ser Thr Asn Met Gly Glu
 130 135 140
 Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys

145 150 155 160
 Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro Ile Asn
 165 170 175
 Asp Asn Asn Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr
 180 185 190
 Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys
 195 200 205

 Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn
 210 215 220
 Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly
 225 230 235 240
 Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala
 245 250 255
 Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile Thr Asp Asn Ala Lys
 260 265 270
 Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg

 275 280 285
 Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Ala
 290 295 300
 Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys
 305 310 315 320
 Asn Ile Ser Gly Ala Glu Trp Asn Lys Thr Leu Gln Gln Val Ala Ala
 325 330 335
 Lys Leu Arg Glu His Phe Asn Asn Lys Thr Ile Ile Phe Lys Pro Ser
 340 345 350

 Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly Gly
 355 360 365
 Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Thr Trp Asn
 370 375 380
 Gly Thr Asn Glu Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Val
 385 390 395 400
 Asn Met Trp Gln Arg Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Ala

405 410 415
 Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg

 420 425 430
 Asp Gly Gly Thr Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp
 435 440 445
 Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys
 450 455 460
 Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val
 465 470 475 480
 Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe
 485 490 495

 Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr
 500 505 510
 Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn
 515 520 525
 Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val
 530 535 540
 Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr
 545 550 555 560
 Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu

 565 570 575
 Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser
 580 585 590
 Gln Asp Glu Ile Trp Asp Asn Met Thr Trp Met Gln Trp Glu Arg Glu
 595 600 605
 Ile Ser Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln
 610 615 620
 Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp Lys Trp
 625 630 635 640

 Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile
 645 650 655

Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Val
660 665 670
Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro
675 680 685
Leu Ser Phe Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Asp Arg Pro
690 695 700
Gly Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile

705 710 715 720
Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser
725 730 735
Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala
740 745 750
Ala Arg Thr Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys
755 760 765
Tyr Leu Trp Asn Leu Leu Gln Tyr Trp Gly Gln Glu Leu Lys Asn Ser
770 775 780

Ala Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val Ala Glu Gly Thr
785 790 795 800
Asp Arg Val Ile Glu Val Val Gln Arg Ala Cys Arg Ala Ile Leu His
805 810 815
Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu
820 825 830

<210> 67
<211> 2493
<212> DNA
<213> Human immunodeficiency virus
<400> 67

atgcgcgtga tgggcatcca gcgcaactgc cagcacctgt ggcgtgggg catcctgatc 60

ttcggcatgc tgatgatctg ctccgccgcc gagaacctgt gggtgaccgt gtactacggc 120
gtgccctgtt ggaaggagge caacaccacc ctgttctgcg cctccgacgc caaggcctac 180
gacaccgagg tgcacaactg gtgggccacc cagcctgcg tgcccaccga ccccaacccc 240
caggagatcg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300

gagcagatgc acgaggacat catctccctg tgggaccagt ccctgaagcc ctgctgaag	360
ctgaccccc tgtgctgac cctgaactgc accgacgtga acgccacaa caactccacc	420
aacatgggcg agatcaagaa ctgctccttc aacatcacca ccgagatccg cgacaagaag	480
cagaaggtgt acgccctgtt ctaccgctg gacgtggtgc ccatcaacga caacaactcc	540
taccgctga tcaactgcaa cacctccgcc atcaccagg cctgccccaa ggtgtccttc	600
gagcccatcc ccatcacta ctgcccccc gccggcttcg ccatcctgaa gtgcaacgac	660
aagaagtcca acggcaccgg cccctgcaag aacgtgtcca ccgtgcagtg caccacggc	720
atcaagcccc tgggtgtccac ccagctgctg ctgaacggct ccctggccga ggaggagatc	780
atcatccgt ccgagaacat caccgacaac gccaagacca tcatcgtgca gctgaacgag	840
tccgtggaga tcaactgcac ccgccccaac aacaacacc gcaagtccat ccgcatcggc	900
cccgccagg ccttctacgc caccggcgac atcatcggcg acatccgcca ggcccactgc	960
aacatctccg gcgccgagt gaacaagacc ctgcagcagg tggccgcaa gctgcgcgag	1020
cattcaaca acaagaccat catcttcaag cctcctccg gcggcgacct ggagatcacc	1080
accactcct tcaactgcgg cggcgagtgc ttctactgca acacctccg cctgttcaac	1140
tccacctgga acggcaccaa cgagaccatc accctgcct gccgcatcaa gcagatcgtg	1200
aacatgtggc agcgctggg ccaggccatg tacgcccc ccatcgccg caacatcacc	1260
tgcaagtcca acatcacgg cctgtgtgtg acccgcgacg gcggcaccaa caacaccgag	1320
accttccgc ccggcgccgg cgacatgcgc gacaactggc gtcaggct gtacaagtac	1380
aaggtggtga agatcgacc cctgggctg gccccacca agccaagcg ccgctggtg	1440
gagcgcgaga agcgcgccgt gggcatcggc gccgtgttc tgggcttct gggcgcgcc	1500
ggctccacca tgggcgccg ctccatcacc ctgacctgc agggccgcca gctgtgttc	1560
ggcatcgtgc agcagcagtc caacctgctg cgcgccatcg agggccagca gcacctgctg	1620
cagctgaccg tgtggggcat caagcagctg caggcccgcg tgcggccgt ggagcgctac	1680
ctgaaggacc agcagctgct gggcatctgg ggctgtccg gcaagctgat ctgcaccacc	1740
aacgtgccct ggaactctc ctggccaac aagtcccagg acgagatctg ggacaacatg	1800
acctggatgc agtgggagcg cgagatctc aactacacc acatcatcta ctccctgatc	1860
gaggagtccc agaaccagca ggagaagaac gagcaggacc tgctggccct ggacaagtgg	1920
gcctccctgt ggaactggtt cgacatcacc aactggctgt ggtacatcaa gatcttcac	1980
atgatcgtgg gcggcctgat cggcctgcgc atcgtgttc ccgtgtgtc catcgtgaac	2040
cgctgcgc agggctact cccctgtcc ttccagacc tgatcccaa cccccgggc	2100

cccgaccgcc ccggcgcat cgaggaggag ggccggcgagc aggaccgcga ccgctccatc 2160

cgcttggtgt ccggttctt ggccctggcc tgggacgacc tgcgtccct gtgcctgttc 2220

tcctaccacc gcctgcgga cttcatcctg atcgccgcc gcaccgtgga gctgctgggc 2280

cgccgcggct gggaggccct gaagtacctg tggacctgc tgcagtactg gggccaggag 2340

ctgaagaact ccgccatctc cctgctggac accaccgcca tcgccgtggc cgagggcacc 2400

gaccgcgtga tcgaggtggt gcagcgcgcc tgcgcgcca tctgcacat ccccgccgc 2460

atccgccagg gcttcgagcg cgccctgctg taa 2493

<210> 68

<211> 845

<212> PRT

<213> Human immunodeficiency virus

<400> 68

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Leu Arg Trp

1 5 10 15

Gly Thr Met Ile Leu Gly Met Ile Ile Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Thr Glu Met

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile His Leu Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Ser Asn Val Asn Val Thr Asn Asn Thr Thr Asn Thr His Glu

130 135 140

Glu Glu Ile Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp

145 150 155 160
 Lys Lys Gln Lys Val Tyr Ser Leu Phe Tyr Arg Leu Asp Val Val Gln
 165 170 175
 Ile Asn Glu Asn Asn Ser Asn Ser Ser Tyr Arg Leu Ile Asn Cys Asn
 180 185 190
 Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile
 195 200 205

 Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys
 210 215 220
 Asp Lys Glu Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val
 225 230 235 240
 Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu
 245 250 255
 Asn Gly Ser Leu Ala Glu Glu Glu Val Ile Ile Arg Ser Glu Asn Ile
 260 265 270
 Thr Asn Asn Ala Lys Thr Ile Ile Val Gln Leu Thr Lys Pro Val Lys

 275 280 285
 Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile
 290 295 300
 Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile
 305 310 315 320
 Arg Gln Ala His Cys Asn Val Ser Arg Ser Glu Trp Asn Lys Thr Leu
 325 330 335
 Gln Lys Val Ala Lys Gln Leu Arg Lys Tyr Phe Lys Asn Lys Thr Ile
 340 345 350

 Ile Phe Thr Asn Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser
 355 360 365
 Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe
 370 375 380
 Asn Ser Thr Trp Asn Asn Gly Thr Met Lys Asn Thr Ile Thr Leu Pro
 385 390 395 400
 Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Ala Gly Gln Ala

405 410 415
Met Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Glu Ser Asn Ile

420 425 430
Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Asn Glu

435 440 445
Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu

450 455 460
Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro

465 470 475 480
Thr Arg Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly

485 490 495

Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
500 505 510

Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
515 520 525

Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln
530 535 540

Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
545 550 555 560

Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly

565 570 575
Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp

580 585 590
Asn Ser Ser Trp Ser Asn Lys Ser Gln Asn Glu Ile Trp Asp Asn Met

595 600 605
Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr His Ile Ile

610 615 620
Tyr Asn Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln

625 630 635 640

Asp Leu Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp
645 650 655

Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly
660 665 670

Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Val Ile Asn
675 680 685

Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Thr Pro
690 695 700

Asn Pro Arg Gly Leu Asp Arg Pro Gly Arg Ile Glu Glu Glu Gly Gly

705 710 715 720

Glu Gln Gly Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala
725 730 735

Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg
740 745 750

Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly
755 760 765

His Ser Ser Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr
770 775 780

Leu Trp Asn Leu Leu Leu Tyr Trp Gly Arg Glu Leu Lys Ile Ser Ala
785 790 795 800

Ile Asn Leu Val Asp Thr Ile Ala Ile Ala Val Ala Gly Trp Thr Asp
805 810 815

Arg Val Ile Glu Ile Gly Gln Arg Ile Gly Arg Ala Ile Leu His Ile
820 825 830

Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu
835 840 845

<210> 69

<211> 845

<212> PRT

<213> Human immunodeficiency virus

<400> 69

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Met Ile Phe Gly Met Ile Ile Ile Cys Ser Ala Ala Glu Asn

20 25 30
 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Glu
 35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50 55 60

 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80
 Gln Glu Ile Asp Leu Glu Asn Val Thr Glu Glu Phe Asn Met Trp Lys
 85 90 95
 Asn Asn Met Val Glu Gln Met His Ala Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Ser Asn Val Asn Val Thr Asn Asn Thr Thr Asn Thr His Glu

 130 135 140
 Glu Glu Ile Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp
 145 150 155 160
 Lys Lys Gln Lys Val Tyr Ser Leu Phe Tyr Arg Leu Asp Val Val Pro
 165 170 175
 Ile Asn Glu Asn Asn Ser Asn Ser Ser Tyr Arg Leu Ile Asn Cys Asn
 180 185 190
 Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile
 195 200 205

 Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys
 210 215 220
 Asp Lys Glu Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val
 225 230 235 240
 Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu
 245 250 255
 Asn Gly Ser Leu Ala Glu Glu Glu Val Met Ile Arg Ser Glu Asn Ile
 260 265 270

Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Thr Glu Pro Val Lys

275 280 285

Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile

290 295 300

Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile

305 310 315 320

Arg Gln Ala His Cys Asn Val Ser Arg Thr Glu Trp Asn Lys Thr Leu

325 330 335

Gln Lys Val Ala Ala Gln Leu Arg Lys His Phe Asn Asn Lys Thr Ile

340 345 350

Ile Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser

355 360 365

Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe

370 375 380

Asn Ser Thr Trp Asn Asn Gly Thr Met Lys Asp Thr Ile Thr Leu Pro

385 390 395 400

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Val Gly Gln Ala

405 410 415

Met Tyr Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Glu Ser Asn Ile

420 425 430

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Asn Glu

435 440 445

Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu

450 455 460

Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro

465 470 475 480

Thr Arg Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly

485 490 495

Leu Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met

500 505 510

Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser

515 520 525
 Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln
 530 535 540
 Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
 545 550 555 560
 Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly

 565 570 575
 Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp
 580 585 590
 Asn Ser Ser Trp Ser Asn Lys Ser Gln Asp Glu Ile Trp Asp Asn Met
 595 600 605
 Thr Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr Asp Ile Ile
 610 615 620
 Tyr Asn Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln
 625 630 635 640

 Asp Leu Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp
 645 650 655
 Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly
 660 665 670
 Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Val Ile Asn
 675 680 685
 Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro
 690 695 700
 Asn Pro Glu Gly Pro Asp Arg Pro Gly Arg Ile Glu Glu Glu Gly Gly

 705 710 715 720
 Glu Gln Gly Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala
 725 730 735
 Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg
 740 745 750
 Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly
 755 760 765
 Arg Ser Ser Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr

770

775

780

Leu Trp Asn Leu Leu Leu Tyr Trp Gly Arg Glu Leu Lys Ile Ser Ala

785 790 795 800

Ile Asn Leu Leu Asp Thr Ile Ala Ile Ala Val Ala Gly Trp Thr Asp

805 810 815

Arg Val Ile Glu Ile Gly Gln Arg Ile Cys Arg Ala Ile Leu Asn Ile

820 825 830

Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu

835 840 845

<210> 70

<211> 2538

<212> DNA

<213> Human immunodeficiency virus

<400> 70

atgcgcgtga tgggcatcca gcgcaactgc cagcacctgc tgcgctgggg caccatgatac 60

ctgggcatga tcatcatctg ctccgccgcc gagaacctgt gggtgaccgt gtactacggc 120

gtgccctgtt ggaaggacgc cgagaccacc ctgttctgcg cctccgacgc caaggcctac 180

gagaccgaga tgcacaactg gtgggccacc cagcctgctg tgcaccacga ccccaacccc 240

caggagatcc acctggagaa cgtgaccgag gatttcaaca tgtggaagaa caacatggtg 300

gagcagatgc acaccgacat catctccctg tgggaccagt cctgaagcc ctgcgtgaag 360

ctgaccccc tgtgcgtgac cctgaactgc tccaactga acgtgaccaa caacaccacc 420

aacaccacg aggaggagat caagaactgc tccttcaaca tgaccaccga gctgcgcgac 480

aagaagcaga aggtgtactc cctgttctac cgcttgacg tgggtcagat caacgagaac 540

aactccaact cctcctaccg cctgatcaac tgcaaacact ccgccatcac ccaggcctgc 600

cccaagggtg ccttcgagcc catecccatc cactactgcg ccccgcggg cttegccatc 660

ctgaagtga aggacaagga gttcaacggc accggcccct gcaagaact gtccaccgtg 720

cagtgcacc accgcatcaa gcccggtgtg tccaccagc tgetgctgaa cggtccctg 780

gccgaggagg aggtgatcat ccgtccgag aacatcacca acaacgcaa gaccatcatc 840

gtgcagctga ccaagccgt gaagatcaac tgcaccgcc ccaacaaca caccgcaag 900

tccatccga tgggcccgg ccaggccttc tacgccaccg gcgacatcat cggcgacatc 960

cgccaggccc actgcaactg gtcccgctcc gattggaaca agaccctgca gaaggaggcc 1020

aagcagctgc gcaagtactt caagaacaag accatcatct tcaccaactc ctccggcggc 1080

gacctggaga tcaccacca ctcttcaac tgcggcggcg agttcttcta ctgcaacacc 1140

tccggcctgt tcaactccac ctggaacaac ggacacatga agaacacat caccctgccc 1200

tgccgcatca agcagatcat caacatgtgg cagcgcgccg gccaggccat gtacgcccc 1260

cccatccagg gcgtgatccg ctgcgagtc aacatcaccg gcctgtgtct gacccgcgac 1320

ggcggcaaca acaacaccaa cgagaccttc cgccccggcg gcggcgacat gcgcgacaac 1380

tggcgctccg agctgtacaa gtacaagggt gtgaagatcg agccctggg cgtggcccc 1440

acccgcgcca agcgcgcgt ggtggagcgc gagaagcgcg ccgtgggcat cggcgccgtg 1500

ttcctgggct tctggggcgc cgccggctcc accatgggcg ccgcctccat caccctgacc 1560

gtgcaggccc gccagctgt gtccggcatc gtgcagcagc agtccaacct gctgcgcgc 1620

atcgaggccc agcagcacct gctgaagctg accgtgtggg gcatcaagca gctgcaggcc 1680

cgctgtctgg ccgtggagcg ctacctgaag gaccagcagc tgctgggcat ctggggctgc 1740

tccggcaagc tgatctgcac caccaacgtg ccttggaaact cctcctgtc caacaagtcc 1800

cagaacgaga tctgggacaa catgacctgg ctgcagtggg acaaggagat ctccaactac 1860

acccacatca tctacaacct gatcgaggag tcccagaacc agcaggagaa gaacgagcag 1920

gacctgttg ccttggacaa gtgggccaac ctgtggaact ggttcgacat ctccaactgg 1980

ctgtggtaca tcaagatctt catcatgatc gtggcgggc tgatcggcct gcgcatctg 2040

ttcgccgtgc tgtccgtgat caaccgcgtg cgccagggt actccccct gtcttccag 2100

acccacaccc ccaacccccg cggcctggac cgccccggc gcatcgagga ggaggcggc 2160

gagcagggcc gcgaccgtc catccgcctg gtgtccggt tcttggccct ggcctgggac 2220

gacctgcgt cctgtgctt gttctctac caccgcctgc gcgacttcat cctgatcgcc 2280

gcccgcaccg tggagctgt gggccactcc tccctgaagg gcctgcgcct gggctgggag 2340

ggcctgaagt acctgtggaa cctgtgtctg tactggggc gcgagctgaa gatctccgc 2400

atcaacctgg tggacacat cgccatcgcc gtggccggct ggaccgaccg cgtgatcgag 2460

atcgccagc gcatcgccg cgccatcctg cacatcccc gccgcatcg ccagggcctg 2520

gagcgcgcc tgctgtaa 2538

<210> 71

<211> 2538

<212> DNA

<213> Human immunodeficiency virus

<400> 71

atgcgcgtga tgggcatcca gcgcaactgc cagcacctgt ggcgtgagg caccatgac	60
tccggcatga tcatcatctg ctccggccgc gagaacctgt gggtagcgt gtactacggc	120
gtgcccgtgt ggaaggacgc cgagaccacc ctgttctgcg cctccgacgc caaggcctac	180
gacaccgagg tgcacaactg gtgggccacc cagccctgcg tgcccaccga cccaacccc	240
caggagatcg acctggagaa cgtgaccgag gatttcaaca tgtggaagaa caacatggtg	300
gagcagatgc acgcccacat catctccctg tgggaccagt ccctgaagcc ctgctgaag	360
ctgaccccc tgtagctgac cctgaactgc tccaactga acgtgacaa caacaccacc	420
aacacccacg aggaggagat caagaactgc tccttcaaca tgaccaccga gctgcgcgac	480
aagaagcaga aggtgtactc cctgttctac cgcctggacg tggtagccat caacgagaac	540
aactccaact cctcctaccg cctgatcaac tgcaaacct ccgcatcac ccaggcctgc	600
cccaagggtg ctctcgagcc catcccatc cactactgcg ccccgccgg ctctgcatc	660
ctgaagtga aggacaagga gttcaacggc accggccct gcaagaact gtccaccgtg	720
cagtgcacc acggcatcaa gccctggtg tccaccagc tgctgtgaa cggctccctg	780
gccgaggagg aggtgatgat ccgctccgag aacatcacg acaacgcaa gaccatcatc	840
gtgcagctga ccgagccctg gaagatcaac tgacccgcc ccaacaaca caccgcaag	900
tccatccga tcggccccgg ccaggccttc tacgccaccg gcgacatcat cggcgacatc	960
cgccaggccc actgcaactg gtcccgacc gattggaaca agaccctgca gaaggtggcc	1020
gcccgctgc gcaagcactt caacaacaag accatcatc tcaactctc ctccggcggc	1080
gacctggaga tcaccacca ctcttcaac tgccggggc agtttttcta ctgcaacacc	1140
tccggcctgt tcaactcac ctggaacaac ggacccatga aggacaccat caccctgcc	1200
tgccgcatca agcagatcat caacatgtgg cagcgctgg gccaggccat gtacgcccc	1260
cccatccagg gcgtgatccg ctgaggtcc aacatcacg gcctgtgtgt gaccccgac	1320
ggcggaaca acaacacaa cgagacctc gccccggcg gcgagcat gcgagacaac	1380
tggcgctcc agctgtaca gtacaagggt gtgaagatc agccctggg cgtggcccc	1440
acccgcca agcgcgcgt ggtggagcg gagaagcgc ccgtgggcct ggccgctg	1500
ttctgggt tctgggcgc cgcggctcc accatgggc cgcctcat caccctgacc	1560
gtgagggcc gccagctgt gtccggcatc gtgagcagc agtccaact gtgagcgc	1620
atcgaggccc agcagacct gctgaagctg accgtgtgg gcatcaagca gctgagggc	1680
cgcgtgtgg ccgtggagcg ctacctgaag gaccagcagc tgctgggcat ctgggctgc	1740
tccggcaagc tgatctgac caccaactg ccttgaact cctcctgtc caacaagtc	1800
caggacgaga tctgggaca catgacctg ctgcagtggg acaaggagat ctccaactac	1860

accgacatca tctacaacct gatcgaggag tcccagaacc agcaggagaa gaacgagcag	1920
gacctgctgg ccttgacaa gtgggccaac ctgtggaact ggttcgacat ctccaactgg	1980
ctgtggtaca tcaagatctt catcatgac gtggcgccg tgcgagcct gcgcatcgtg	2040
ttcgcccgtg tgcctgtgat caaccgcgtg cgccagggt actccccct gtccttcag	2100
acctgaccc ccaacccga gggcccccac cggcccgcc gcacgagga ggaggcgcc	2160
gagcaggcc gcgaccgtc catcgccgt gtgtccggt tctggccct ggcctgggac	2220
gacctgcgt cctgtgcct gttctctac caccgcctgc gcgacttcat cctgatgcc	2280
gcccgcaccg tggagctgt gggccgctc tccctgaagg gctgcgcct gggtgggag	2340
ggcctgaagt acctgtgga cctgtgtgt tactggggc gcgagctgaa gatctccgc	2400
atcaacctgc tggacacat cgccatgcc gtggccggct ggaccgacc cgtgatcag	2460
atcgccagc gcacgtgcc cgccatcct aacatcccc gccgcatcc ccaggccctg	2520
gagcgcccc tgcgtgaa	2538
<210> 72	
<211> 2532	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 72	
atgcgcgtga tgggcacca gcgcaactac cagcacctgt ggcgtgggg catcctgatc	60
ctgggcatgc tgatcatgt caagccacc gacctgtggg tgaccgtgta ctacggcgtg	120
cccgtgtgga aggacgccga caccacctg ttctgcgcct ccgacgcaa ggcctacgac	180
accgaggtgc acaacgtgt gggcaccac gctgcgtgc ccaccgacc caacccccag	240
gaggtgaacc tggagaacgt gaccaggac ttcaacatgt ggaagaaca catggtggag	300
cagatgcacg aggacatcat ctccctgtgg gaccagtccc tgaagccctg cgtgaagctg	360
acccccctgt gcgtgacct gaactgtcc aacccaaca ccaccaaca ctccaccatg	420
gaggagatca agaactgtc ctacaacatc accaccgagc tgcgcgaaa gaccagaag	480
gtgtactccc tgttctaaa gctggacgtg gtgcagctgg acgagtcaa caagtccgag	540
tactactacc gctgatcaa ctgcaacacc tccgcatca cccaggcctg ccccaagggtg	600
tccttcgagc ccatcccat ccaactactg cccccgccg gcttcgcat cctgaagtgc	660
aaggaccccc gcttcaacgg caccggctcc tgcaacaac gtgtcctcgt gcagtgcacc	720
caggcatca agccgtggc ctccaccag ctgtgtgta acggtccct ggccgagggc	780
aaggtgatga tccgctcca gaacatcacc aacaaccca agaacatcat cgtgcagttc	840

aacaagcccg tgcccatcac ctgcatccgc cccaacaaca acacccgcaa gtccatccgc	900
ttcgccccg gccaggcctt ctacaccaac gacatcatcg gcgacatccg ccaggccac	960
tgcaacatca acaagaccaa gtggaacgcc accctgcaga aggtggccga gcagctgcgc	1020
gagcacttcc ccaacaagac catcatcttc accaactcct ccggcggcga cctggagatc	1080
accacccact ccttcaactg cggcggcgag ttcttctact gcaacaccac cggcctgttc	1140
aactccacct ggaagaacgg caccaccaac aacaccgagc agatgatcac cctgccctgc	1200
cgcatcaagc agatcatcaa catgtggcag cgcgtgggcc gcgccatgta cgcaccccc	1260
atcgccggcg tgatcaagtg cacctccaac atcaccggca tcctcctgac ccgcgacggc	1320
ggcaacaacg agaccgagac ctccgcccc ggcgggcgcg acatgcgcga caactggcgc	1380
tccgagctgt acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc cccacccgc	1440
gccaagcgcc gcgtggtgga gcgcgagaag cgcgccgtgg gcatgggcgc cgtgttcctg	1500
ggcttcctgg gcgcgcggc ctccaccatg ggcgccgcct ccatcacct gaccgtgcag	1560
gccccccagc tgctgtccgg catcgtgcag cagcagtcga acctgctgaa ggccatcgag	1620
gcccagcagc acctgctgaa gctgaccgtg tggggcatca agcagctgca ggccccgctg	1680
ctggccctgg agcgtacct gcaggaccag cagctgtctg gcatctgggg ctgctccggc	1740
aagctgatct gcgccaccac cgtgccctgg aactcctcct ggtccaaca gaccaggag	1800
gagatctgga acaacatgac ctggctgcag tgggacaagg agatctcaa ctacaccaac	1860
atcatctaca agctgctgga ggagtcccag aaccagcagg agaagaacga gcaggacctg	1920
ctggccctgg acaagtgggc caacctgtgg aactggttca acatcaccaa ctggctgtgg	1980
tacatccgca tcttcatcat gatcgtgggc ggctgatcg gcctgcgcat cgtgatcgcc	2040
atcatctccg tggatgaaccg cgtgcccag ggctactccc cctgtcctt ccagatcccc	2100
accccccaacc ccgagggcct ggaccgcccc ggccgcatcg aggagggcgg cggcgagcag	2160
ggccgcgacc gtccatccg cctggtgtcc ggcttcttgg cctggcctg ggacgacctg	2220
cgctccctgt gctgtttctc ctaccaccgc ctgcgcgact gcatcctgat cgcgccccgc	2280
accgtggagc tgctgggcca ctctccctg aagggcctgc gcctgggctg ggagggcctg	2340
aagtacctgt ggaacctgct gctgtactgg ggccgcgagc tgaagaactc cgccatctcc	2400
ctgctggaca ccatcgccgt ggccgtggcc gactggaccg accgctgat cgagatcggc	2460
cagcgcgct gccgcgcat cctgaacatc ccccgccga tccgccaggg cttcgagcgc	2520
gccctgctgt aa	2532
<210> 73	
<211> 843	

<212> PRT

<213> Human immunodeficiency virus

<400> 73

```

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

Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln
  65             70             75             80
Glu Val Asn Leu Glu Asn Val Thr Glu Asp Phe Asn Met Trp Lys Asn
      85             90             95
Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln
      100            105            110
Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn
      115            120            125
Cys Ser Asn Ala Asn Thr Thr Asn Asn Ser Thr Met Glu Glu Ile Lys

      130            135            140
Asn Cys Ser Tyr Asn Ile Thr Thr Glu Leu Arg Asp Lys Thr Gln Lys
  145            150            155            160
Val Tyr Ser Leu Phe Tyr Lys Leu Asp Val Val Gln Leu Asp Glu Ser
      165            170            175
Asn Lys Ser Glu Tyr Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala
      180            185            190
Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His
      195            200            205

Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Lys Asp Pro Arg
      210            215            220

```

Phe Asn Gly Thr Gly Ser Cys Asn Asn Val Ser Ser Val Gln Cys Thr
 225 230 235 240
 His Gly Ile Lys Pro Val Ala Ser Thr Gln Leu Leu Leu Asn Gly Ser
 245 250 255
 Leu Ala Glu Gly Lys Val Met Ile Arg Ser Glu Asn Ile Thr Asn Asn
 260 265 270
 Ala Lys Asn Ile Ile Val Gln Phe Asn Lys Pro Val Pro Ile Thr Cys
 275 280 285
 Ile Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Phe Gly Pro Gly
 290 295 300
 Gln Ala Phe Tyr Thr Asn Asp Ile Ile Gly Asp Ile Arg Gln Ala His
 305 310 315 320
 Cys Asn Ile Asn Lys Thr Lys Trp Asn Ala Thr Leu Gln Lys Val Ala
 325 330 335
 Glu Gln Leu Arg Glu His Phe Pro Asn Lys Thr Ile Ile Phe Thr Asn
 340 345 350
 Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly
 355 360 365
 Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gly Leu Phe Asn Ser Thr Trp
 370 375 380
 Lys Asn Gly Thr Thr Asn Asn Thr Glu Gln Met Ile Thr Leu Pro Cys
 385 390 395 400
 Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Arg Val Gly Arg Ala Met
 405 410 415
 Tyr Ala Pro Pro Ile Ala Gly Val Ile Lys Cys Thr Ser Asn Ile Thr
 420 425 430
 Gly Ile Ile Leu Thr Arg Asp Gly Gly Asn Asn Glu Thr Glu Thr Phe
 435 440 445
 Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr
 450 455 460
 Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Arg
 465 470 475 480

Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Met Gly
485 490 495

Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala
500 505 510

Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile
515 520 525

Val Gln Gln Gln Ser Asn Leu Leu Lys Ala Ile Glu Ala Gln Gln His
530 535 540

Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val
545 550 555 560

Leu Ala Leu Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu Gly Ile Trp
565 570 575

Gly Cys Ser Gly Lys Leu Ile Cys Ala Thr Thr Val Pro Trp Asn Ser
580 585 590

Ser Trp Ser Asn Lys Thr Gln Glu Glu Ile Trp Asn Asn Met Thr Trp
595 600 605

Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr Asn Ile Ile Tyr Lys
610 615 620

Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu
625 630 635 640

Leu Ala Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asn Ile Thr
645 650 655

Asn Trp Leu Trp Tyr Ile Arg Ile Phe Ile Met Ile Val Gly Gly Leu
660 665 670

Ile Gly Leu Arg Ile Val Ile Ala Ile Ile Ser Val Val Asn Arg Val
675 680 685

Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Ile Pro Thr Pro Asn Pro
690 695 700

Glu Gly Leu Asp Arg Pro Gly Arg Ile Glu Glu Gly Gly Gly Glu Gln
705 710 715 720

Gly Arg Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala

725 730 735
 Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg
 740 745 750
 Asp Cys Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly His Ser
 755 760 765
 Ser Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr Leu Trp
 770 775 780

 Asn Leu Leu Leu Tyr Trp Gly Arg Glu Leu Lys Asn Ser Ala Ile Ser
 785 790 795 800
 Leu Leu Asp Thr Ile Ala Val Ala Val Ala Glu Trp Thr Asp Arg Val
 805 810 815
 Ile Glu Ile Gly Gln Arg Ala Cys Arg Ala Ile Leu Asn Ile Pro Arg
 820 825 830
 Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu
 835 840
 <210> 74
 <211> 841
 <212> PRT

 <213> Human immunodeficiency virus
 <400> 74
 Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
 1 5 10 15
 Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Lys
 20 25 30
 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
 35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50 55 60

 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80
 Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
 85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asp Leu Met Asn Ala Thr Asn Thr Asn Thr Thr Ile Ile
 130 135 140
 Tyr Arg Trp Arg Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr
 145 150 155 160
 Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu
 165 170 175
 Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Arg Leu Ile Ser
 180 185 190
 Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu
 195 200 205
 Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys
 210 215 220
 Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser
 225 230 235 240
 Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu
 245 250 255
 Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Glu
 260 265 270
 Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser
 275 280 285
 Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile
 290 295 300
 His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly Glu Ile Ile Gly
 305 310 315 320
 Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Asn
 325 330 335
 Thr Leu Lys Gln Ile Val Lys Lys Leu Arg Glu Gln Phe Gly Asn Lys

340 345 350
 Thr Ile Val Phe Asn Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met
 355 360 365
 His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln
 370 375 380
 Leu Phe Asn Ser Thr Trp Asn Gly Thr Trp Asn Asn Thr Glu Gly Asn
 385 390 395 400
 Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu
 405 410 415
 Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys
 420 425 430
 Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn
 435 440 445
 Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp
 450 455 460
 Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly
 465 470 475 480
 Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg
 485 490 495
 Ala Val Gly Ile Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly
 500 505 510
 Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln
 515 520 525
 Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile
 530 535 540
 Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
 545 550 555 560
 Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln
 565 570 575
 Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala
 580 585 590

Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp
595 600 605

Asp Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr
610 615 620

Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
625 630 635 640

Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
645 650 655

Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met
660 665 670

Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser
675 680 685

Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr
690 695 700

Arg Leu Pro Ala Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu
705 710 715 720

Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser Gly Arg Leu Val Asp Gly
725 730 735

Phe Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser
740 745 750

Tyr His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu
755 760 765

Leu Leu Gly Arg Arg Gly Trp Glu Val Leu Lys Tyr Trp Trp Asn Leu
770 775 780

Leu Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu
785 790 795 800

Asn Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu
805 810 815

Val Val Gln Arg Ala Cys Arg Ala Ile Leu His Ile Pro Arg Arg Ile
820 825 830

Arg Gln Gly Leu Glu Arg Ala Leu Leu
835 840

<210> 75
 <211> 2532
 <212> DNA
 <213> Human immunodeficiency virus

<400> 75

atgcgcgtga tgggcaccca gcgcaactac cagcacctgt ggcgctgggg catcctgac	60
ctgggcatgc tgatcatgtg caaggccacc gacctgtggg tgaccgtgta ctacggcgtg	120
cccgtgtgga aggagccga caccacctg ttctgcgcct ccgacgcaa ggcctacgac	180
accgaggtgc acaactgtg ggccaccac gcctgcgtgc ccaccgacc caaccccag	240
gaggtgaacc tggagaactg gaccgaggac ttcaacatgt ggaagaaca catggtggag	300
cagatgcacg aggacatcat ctccctgtgg gaccagtccc tgaagccctg cgtgaagctg	360
accccctgt gcgtgacct gaactgctcc aacgccaaca ccaccaaaa ctccaccatg	420
gaggagatca agaactgctc ctacaacatc accaccgagc tgcgcgaaa gaccagaag	480
gtgtactccc tgttctaaa gctggacgtg gtgcagctgg acgagtcaa caagtccgag	540
tactactacc gcctgatcaa ctgcaacacc tccgcatca cccaggcctg cccaaggtg	600
tccttcgagc ccatcccat ccaactatgc gccccgccg gcttcgcat cctgaagtgc	660
aaggacccc gcttcaacgg caccggctcc tgcaacaacg tgtctccgt gcagtgcacc	720
cacggcatca agcccgtggc ctccaccag ctgctgtga acggctccct ggccgagggc	780
aaggtgatga tccgctcca gaacatcac aacaaccca agaacatcat cgtgcagttc	840
aacaagccc tgcccatcac ctgcatccgc cccaacaaca acaccgcaa gtccatccgc	900
ttcggcccc gccaggcctt ctacaccaac gacatcatc gcgacatccg ccaggccac	960
tgcaacatca acaagaccaa gtggaacgcc acctgcaga aggtggccga gcagctgcgc	1020
gagcacttc ccaacaagac catcatctt accaactcct ccggcggcga cctggagatc	1080
accacccact cttcaactg ggccggcgag ttcttctact gcaacaccac cggcctgttc	1140
aactccact ggaagaacgg caccaccaac aacaccgagc agatgatcac cctgccctgc	1200
cgcacatcagc agatcatcaa catgtggcag cgcgtgggcc gcgcatgta ccccccccc	1260
atgcgcggcg tgatcaagtg cactccaac atcaccggca tcatcctgac ccgcgacggc	1320
ggcaacaacg agaccgagac ctccgcccc ggccggcgcg acatgcgcga caactggcgc	1380
tccgagctgt acaagtacaa ggtgtgaag atcgagcccc tggcgctggc cccacccgc	1440
gccacggcc gcgtggtgga gcgcgagaag cgcgccgtgg gcatgggcgc cgtgttcctg	1500
ggcttcctgg gcgcccggc ctccaccatg ggccgcgcct ccatcacct gaccgtgcag	1560

gccccccagc tgctgtccgg catcgtgcag cagcagtcga acctgctgaa ggccatcgag 1620
gcccagcagc acctgctgaa gctgaccgtg tggggcatca agcagctgca ggccccgctg 1680

ctggccctgg agcgctacct gcaggaccag cagctgctgg gcatctgggg ctgctccggc 1740
aagctgatct gcgccaccac cgtgccctgg aactcctcct ggtccaacaa gaccaggag 1800
gagatctgga acaacatgac ctggctgcag tgggacaagg agatctccaa ctacaccaac 1860
atcatctaca agctgctgga ggagtcccag aaccagcagg agaagaacga gcaggacctg 1920
ctggccctgg acaagtgggc caacctgtgg aactgggtca acatcaccaa ctggctgtgg 1980
tacatccgca tcttcatcat gatcgtgggc ggctgctgct gcctgcgcat cgtgatcgcc 2040
atcatctccg tggatgaaccg cgtgcgccag ggctactccc cctgtcctt ccagatcccc 2100

acccccaacc ccgagggcct ggaccgcccc ggccgcatcg aggaggcgcg cgccgagcag 2160
ggccgcgacc gctccatccg cctggtgtcc ggcttctctg cctggcctg ggacgacctg 2220
cgctccctgt gcctgttctc ctaccaccgc ctgcgcgact gcatcctgat cgccgcccgc 2280
accgtggagc tgctgggcca ctctccctg aagggcctgc gcctgggctg ggagggcctg 2340
aagtacctgt ggaacctgct gctgtactgg ggccgcgagc tgaagaactc cgccatctcc 2400
ctgctggaca ccatcgccgt ggccgtggcc gagtggaccg accgcgtgat cgagatcggc 2460
cagcgcgcct gccgcgccat cctgaacatc ccccgccgca tccgccaggg cttcgagcgc 2520

gccctgctgt aa 2532

<210> 76
<211> 2526
<212> DNA
<213> Human immunodeficiency virus
<400> 76

atgcgcgtga agggcatccg caagaactac cagcacctgt ggcgctgggg caccatgctg 60
ctgggcatgc tgatgatctg ctccgccgcc gagaagctgt gggtgaccgt gtactacggc 120
gtgcccgctg ggaaggaggc caccaccacc ctgttctgcg cctccgacgc caaggcctac 180
gacaccgagg tgcacaacgt gtgggccacc cagcctgctg tgcccaccga cccaacccc 240
caggaggtgg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300

gagcagatgc acgaggacat catctccctg tgggaccagt cctgaagcc ctgcgtgaag 360
ctgaccccc tgtgcgtgac cctgaactgc accgacctga tgaacgccac caacaccaac 420
accacatca tctaccctg gcgcggcgag atcaagaact gctccttcaa catcaccacc 480
tccatccgcg acaaggtgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540

atcgacaacg acaacacctc ctaccgcctg atctcctgca acacctccgt gatcacccag	600
gcctgcccc aggtgtcctt cgagcccatc cccatccact actgcgcccc cgccggcttc	660
gccatcctga agtgcaacga caagaagttc aacggcaccc gccctgcac caacgtgtcc	720
accgtgcagt gcacccacgg catccgcccc gtggtgtcca cccagctgct gctgaacggc	780
tccttggccg aggaggaggt ggtgatccgc tccgagaact tcaccgaca cgccaagacc	840
atcatcgtgc agctgaacga gtccgtggag atcaactgca cccgccccaa caacaacacc	900
cgcaagtcca tcacatcgg ccccgccgc gccttctaca ccaccggcga gatcatcggc	960
gacatccgcc aggcccactg caacatctcc cgcgccaagt ggaacaacac cctgaagcag	1020
atcgtgaaga agctgcgca gcagttcggc aacaagacca tcgtgttcaa ccagtctcc	1080
ggcggcgacc ccgagatcgt gatgcactcc ttcaactgcg gcggcgagtt cttctactgc	1140
aacaccaccc agctgttcaa ctccacctgg aacggcacct ggaacaacac cgagggaac	1200
atcacctgc cctgccgat caagcagatc atcaacatgt ggcaggaggt gggcaaggcc	1260
atgtacgcc ccccatccg cggccagatc cgtgtctct ccaacatcac cggcctgctg	1320
ctgacccgcg acggcggcaa caacgagacc gagatcttc gcccggcgg cggcgacatg	1380
cgcgacaact ggcgctcca gctgtacaag tacaaggtgg tgaagatcga gcccctgggc	1440
gtggccccc ccaaggccaa gcgccgctg gtgcagcgc agaagcgcg cgtgggcatc	1500
ggcgccatgt tcttgggtt cctgggcgc gccggctcca ccatgggcgc cgcctccatg	1560
accctgaccg tgcaggcccg ccagctgctg tccggcatcg tgcagcagca gaacaacctg	1620
ctgcgcgcca tcgaggccca gcagcacctg ctgcagctga ccgtgtgggg catcaagcag	1680
ctgcaggccc gcgtgtggc cgtggagcgc tacctgaagg accagcagct gctgggcatc	1740
tggggctgct ccggcaagct gatctgcacc accgccgtgc cctggaacgc ctctgtgtcc	1800
aacaagtccc tggacgagat ctgggacaac atgacctgga tggagtggga gcgcgagatc	1860
gacaactaca cctccctgat ctacacctg atcgaggagt cccagaacca gcaggagaag	1920
aacgagcagg agctgtgga gctggacaag tgggcctccc tgtggaactg gtctgacatc	1980
accaactggc tgtggtacat caagatcttc atcatgatcg tggcgccct ggtgggcctg	2040
cgcactgtgt tcgccgtgt gtccatcgtg aaccgcgtgc gccagggcta ctccccctg	2100
tccttcaga ccgcctgcc gccecccgcc ggcgccgacc gcccagagg catcgaggag	2160
gaggcgggcg agcgcgaccg cgaccgtcc ggccgcctgg tggacggctt cctggccctg	2220
atctgggacg acctgcgtc cctgtgctg ttctctacc accgcctgcg cgacctgctg	2280
ctgatcgtga cccgatcgt ggagctgctg ggccgcgcg gctgggaggt gctgaagtac	2340

tgggtggaacc tgctgcagta ctgggtcccag gagctgaaga actccgccgt gtccctgctg 2400

aacgccaccg ccatacgcgt ggccgagggc accgaccgagc tgatcgaggt ggtgcagcgc 2460

gcctgccgagc ccatacctgca catecccccgc cgcatacgcgc agggcctgga gcgcgccctg 2520

ctgtaa 2526

<210> 77

<211> 840

<212> PRT

<213> Human immunodeficiency virus

<400> 77

Met Arg Val Lys Gly Ile Arg Lys Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Leu Leu Asn Ala Thr Asn Thr Asn Ser Thr Asn Met

130 135 140

Tyr Arg Trp Arg Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr

145 150 155 160

Ser Ile Arg Asp Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu

165 170 175

Asp Val Val Pro Ile Asp Asn Asn Thr Ser Tyr Arg Leu Ile Asn Cys
 180 185 190
 Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro
 195 200 205
 Ile Pro Ile His Tyr Cys Thr Pro Ala Gly Phe Ala Ile Leu Lys Cys
 210 215 220
 Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr
 225 230 235 240

 Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu
 245 250 255
 Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Glu Asn
 260 265 270
 Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val
 275 280 285
 Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His
 290 295 300
 Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Glu Ile Ile Gly Asp

 305 310 315 320
 Ile Arg Gln Ala His Cys Asn Leu Ser Arg Ala Lys Trp Asn Asn Thr
 325 330 335
 Leu Lys Gln Val Val Thr Lys Leu Arg Glu Gln Phe Asp Asn Lys Thr
 340 345 350
 Ile Val Phe Asn Pro Ser Ser Gly Gly Asp Pro Glu Ile Val Met His
 355 360 365
 Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln Leu
 370 375 380

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

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Glu
435 440 445

Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
450 455 460

Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
465 470 475 480

Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala
485 490 495

Val Gly Ile Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser
500 505 510

Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu
515 520 525

Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu
530 535 540

Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
545 550 555 560

Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Arg Asp Gln Gln Leu
565 570 575

Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Thr Val
580 585 590

Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asn
595 600 605

Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr Gly
610 615 620

Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
625 630 635 640

Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
645 650 655

Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile
660 665 670

Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile

675 680 685
Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Arg
690 695 700
Leu Pro Ala Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu
705 710 715 720
Gly Gly Glu Arg Asp Arg Asp Arg Ser Gly Arg Leu Val Asn Gly Phe
725 730 735
Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr

740 745 750
His Arg Leu Arg Asp Leu Leu Leu Ile Val Ala Arg Ile Val Glu Leu
755 760 765
Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu
770 775 780
Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn
785 790 795 800
Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val
805 810 815

Val Gln Arg Ala Cys Arg Ala Ile Leu His Ile Pro Arg Arg Ile Arg
820 825 830
Gln Gly Leu Glu Arg Ala Leu Leu
835 840

<210> 78
<211> 835
<212> PRT
<213> Human immunodeficiency virus
<400> 78

Met Arg Val Arg Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp
1 5 10 15
Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Val Val Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
35 40 45

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

Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile
305 310 315 320

Ser Glu Asp Lys Trp Asn Lys Thr Leu Gln Lys Val Ser Lys Lys Leu
325 330 335

Lys Glu His Phe Pro Asn Lys Thr Ile Lys Phe Glu Pro Ser Ser Gly
340 345 350

Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe
355 360 365

Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser Thr Tyr Asn Ser Thr
370 375 380

Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
385 390 395 400

Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn
405 410 415

Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
420 425 430

Gly Lys Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
435 440 445

Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys
450 455 460

Pro Leu Gly Ile Ala Pro Thr Lys Ala Lys Arg Arg Val Val Glu Arg
465 470 475 480

Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly
485 490 495

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
500 505 510

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
515 520 525

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
530 535 540

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys

545 550 555 560
 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys
 565 570 575
 Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Glu
 580 585 590
 Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile Ser
 595 600 605

 Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Asn Gln
 610 615 620
 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
 625 630 635 640
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
 645 650 655
 Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala
 660 665 670
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser

 675 680 685
 Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Pro Asp Arg Leu Gly Arg
 690 695 700
 Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile Arg Leu
 705 710 715 720
 Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys
 725 730 735
 Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg
 740 745 750

 Ala Val Glu Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly
 755 760 765
 Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp Gly Leu
 770 775 780
 Glu Leu Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala Ile Ala
 785 790 795 800
 Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys

805	810	815
Arg Ala Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala		
820	825	830
Ala Leu Gln		
835		
<210> 79		
<211> 2523		
<212> DNA		
<213> Human immunodeficiency virus		
<400> 79		
atgcgcgtga agggcatccg caagaactgc cagcacctgt ggcgctgggg caccatgctg	60	
ctgggcatgc tgatgatctg ctccgccgcc gagaacctgt gggtagacct gtactacggc	120	
gtgcccgctgt ggaaggaggc caccaccacc ctgttctgcg cctccgacgc caaggcctac	180	
gagaccgagg tgcacaactg gtgggccacc cagcctgctg tgcccaccga cccaacccc	240	
caggaggtgg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300	
gagcagatgc acgaggacat catctccctg tgggaccagt ccctgaagcc ctgcgtgaag	360	
ctgaccccc tgtagctgac cctgaactgc accgacctgc tgaacgccac caacaccaac	420	
tccaccaaca tgiaccgctg gcgcggcgag atcaagaact gctccttcaa catcaccacc	480	
tccatccgcg acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc	540	
atcgacaaca acacctcta ccgctgac aactgcaaca cctccgtgat caccagggcc	600	
tgcccaagg tgccttca gccatcccc atccactact gcacccccgc cggttcgcc	660	
atcctgaagt gcaacgacaa gaagtcaac ggcaccggcc cctgcaagaa cgtgtccacc	720	
gtgcagtgca cccacggcat ccgccccgtg gtgtccacc agctgctgct gaacggctcc	780	
ctggccgagg aggaggtggt gatccgctcc gagaacttca ccgacaacgc caagaccatc	840	
atcgtgcagc tgaacgagtc cgtggagatc aactgcaccc gcccaacaa caacacccgc	900	
aagtccatcc acatcgccc cggccgcgcc ttctaccca ccggcgagat catcgcgac	960	
atccgccagg cccactgcaa cctgtccgc gccaaagtga acaacacct gaagcaggtg	1020	
gtgaccaagc tgcgcgagca gttagacaac aagaccatcg tgttaaccc ctctccggc	1080	
ggcgacccc agatcgtgat gcactcttc aactgcggcg gcgagttctt ctactgcaac	1140	
accaccagc tgttcaactc cacctggaac ggcacctgga acaacaccga gggcaacatc	1200	
accctgccct gccgcatcaa gcagatcatc aacatgtggc aggaggtggg caaggccatg	1260	

tacgcccccc ccatccgcgg ccagatccgc tgctcctcca acatcaccgg cctgctgctg	1320
acccgcgacg gcggcaacaa cgagaccgag atcttccgcc ccggcggcgg cgacatgcgc	1380
gacaactggc gctccgagct gtacaagtac aaggtggtga agatcgagcc cctgggcgtg	1440
gccccacca aggccaagcg ccgctggtg cagcgcgaga agcgcgccgt ggcatcggc	1500
gccatgttcc tgggtttcct gggcgccgcc ggctccacca tggcgccgc ctccatgacc	1560
ctgaccgtgc agggccgcc gctgctgtcc ggcatcgtgc agcagcagaa caacctgctg	1620
cgcgccatcg agggccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg	1680
caggcccgcg tgctggccgt ggagcgctac ctgcgcgacc agcagctgct ggcatctgg	1740
ggctgctccg gcaagctgat ctgcaccacc accgtgccct ggaacgcctc ctggtccaac	1800
aagtccctgg acgagatctg gaacaacatg acctggatgg agtgggagcg cgagatcgac	1860
aactacaccg gcctgatcta cacctgatc gaggagtccc agaaccagca ggagaagaac	1920
gagcaggagc tgctggagct ggacaagtgg gcctccctgt ggaactggtt cgacatcacc	1980
aactggctgt ggtacatcaa gatcttcac atgatcgtgg gcggcctggt gggcctgcgc	2040
atcgtgttcg ccgtgctgtc catcgtgaac cgcgtgcgcc agggctactc cccctgtcc	2100
ttccagaccg gcctgccgc ccccgcggc cccgaccgcc ccgagggcat cgaggaggag	2160
ggcggcgagc gcgaccgca ccgctccggc cgcttggga acggttcct ggccctgatc	2220
tgggacgacc tgcgtccct gtgcctgttc tctaccacc gcctgcgca cctgctgctg	2280
atcgtggccc gcacgtgga gctgctgggc cgcgcggct gggaggccct gaagtactgg	2340
tggaacctgc tgcagtactg gtcccaggag ctgaagaact ccgccgtgtc cctgtgaac	2400
gccaccgcca tcgccgtggc cgagggcacc gaccgctga tcgaggtggt gcagcgcgcc	2460
tgccgcgcca tctgcacat ccccgcgcg atccgccagg gcctggagcg cgcctgctg	2520
taa	2523
<210> 80	
<211> 2508	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 80	
atgcgcgtgc gcggcatcct gcgcaactgc cagcagtggt ggatctgggg catcctgggc	60
ttctggatgc tgatgatctg caacgtggtg ggcaacctgt gggtagacct gtactacggc	120
tgcccggtgt ggaaggagge caagaccacc ctgtttctgcg cctccgacgc caaggcctac	180
gagaaggagg tgcacaactg gtgggccacc cagcctgcg tgcccaccga ccccaacccc	240

caggagatcg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg	300
gaccagatgc acgaggacat catctccctg tgggaccagt ccctgaagcc ctgcgtgaag	360
ctgaccccc tgctgcgtgac cctgaactgc accaacgcca ccaacgccac caacaccatg	420
ggcgagatca agaactgctc cttaacatc accaccgagc tgcgcgacaa gaagcagaag	480
gtgtacgccc tgttctaccg cctggacatc gtgccctga acgagaacaa ctctaccgc	540
ctgatcaact gcaacacctc cgccatcacc caggcctgcc ccaaggtgtc cttcgacccc	600
atccccatcc actactgcgc ccccgccggc tacgcatcc tgaagtcaa caacaagacc	660
ttcaacggca cgggccctg caacaactg tccaccgtgc agtgcacca cggcatcaag	720
cccgtggtgt ccaccagct gctgctgaac ggctccctgg ccgaggagga gatcatcatc	780
cgctccgaga acctgaccaa caacccaag accatcatcg tgcacctgaa cgagtccgtg	840
gagatcgtgt gcacccgcc caacaacaac acccgcaagt ccatccgcat cgccccggc	900
cagaccttct acgccaccgg cgacatcatc ggcgacatcc gccaggccca ctgcaacatc	960
tccgaggaca agtggaaaca gacctgcag aaggtgtcca agaagctgaa ggagcacttc	1020
cccaacaaga ccatcaagtt cgagccctcc tccggcggcg acctggagat caccaccac	1080
tccttcaact gccgcggcga gttcttctac tgcaaacact ccaagctgtt caactccacc	1140
tacaactcca ccaactccac catcaccctg cctgccgca tcaagcagat catcaacatg	1200
tggcaggagg tgggcgcgc catgtacgcc cccccatcg ccggcaacat cacctgcaag	1260
tccaacatca cggcctgct gctgacccgc gacggcggca agaacaacac cgagaccttc	1320
cgccccggcg gcggcgacat gcgcgacaac tggcgctccg agctgtacaa gtacaagggtg	1380
gtggagatca agccctggg catgcccc accaaggcca agcgccgct ggtggagcgc	1440
gagaagcgcg ccgtgggcat cggcgccgtg ttcttgggtt tcttgggcgc cgccggtcc	1500
accatgggcg ccgcctccat cacctgacc gtgcaggccc gccagctgct gtccggcatc	1560
gtgcagcagc agtccaact gctgcgcgc atcgaggccc agcagcacat gctgcagctg	1620
accgtgtggg gcatcaagca gctgcagacc cgcgtgctgg ccatcgagcg ctacctgaag	1680
gaccagcagc tgctgggcat ctggggctgc tccggcaagc tgatctgcac caccgccgtg	1740
ccctggaact cctctggtc caacaagtcc caggaggaca tctgggacaa catgacctgg	1800
atgcagtggg accgcgagat ctccaactac accgacacca tctaccgct gctggaggac	1860
tcccagaacc agcaggagaa gaacgagaag gacctgtgg ccctggactc ctggaagaac	1920
ctgtggaact ggttcgacat caccaactgg ctgtggtaca tcaagatctt catcatgatc	1980
gtgggcggcc tgatcgccct gcgcatcatc ttgccgtgc tgtccatcgt gaaccgcgtg	2040
cgccagggtt actccccct gtccttcag acctgaccc ccaacccccg cgccccgcac	2100

cgcttgggccc gcatcgagga ggagggcggc gagcaggacc ggcaccgctc catccgcctg 2160
 gtgtccgctt tcttggccct ggcctgggac gacctgcgtt ccctgtgcct gttctctac 2220
 caccgcctgc gcgacttcat cctgatcgcc gcccgcgccg tggagctgct gggccgctcc 2280

tccctgcgcg gcctgcagcg cggctgggag gcctgaagt acctgggctc cctggtgcag 2340
 tactggggcc tggagctgaa gaagtccgcc atctccctgc tggacaccat cgccatcgcc 2400
 gtggccgagg gcaccgaccg catcatcgag ctgatccage gcatctgccg cgccatccgc 2460
 aacatccccc gccgcatccg ccagggtctc gagggcgccc tgcagtaa 2508

<210> 81

<211> 835

<212> PRT

<213> Human immunodeficiency virus

<400> 81

Met Arg Val Met Gly Ile Leu Arg Asn Cys Gln Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Val Val Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Arg Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asn Ala Thr Asn Ala Thr Asn Thr Met Gly Glu Met Lys

130 135 140

Asn Cys Ser Phe Asn Ile Thr Thr Glu Leu Arg Asp Lys Lys Gln Lys

145 150 155 160
 Val Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asn Asp Asn
 165 170 175
 Asn Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala
 180 185 190
 Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro
 195 200 205
 Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr
 210 215 220
 Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys

 225 230 235 240
 Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu
 245 250 255
 Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Ala Lys Thr Ile
 260 265 270
 Ile Val His Leu Asn Glu Ser Val Glu Ile Val Cys Thr Arg Pro Asn
 275 280 285
 Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr
 290 295 300

 Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile
 305 310 315 320
 Ser Glu Glu Lys Trp Asn Lys Thr Leu Gln Arg Val Gly Glu Lys Leu
 325 330 335
 Lys Glu His Phe Pro Asn Lys Thr Ile Lys Phe Ala Pro Ser Ser Gly
 340 345 350
 Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg Gly Glu Phe
 355 360 365
 Phe Tyr Cys Asn Thr Ser Arg Leu Phe Asn Ser Thr Tyr Asn Ser Lys

 370 375 380
 Asn Ser Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
 385 390 395 400

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

 Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys
 450 455 460
 Pro Leu Gly Ile Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg
 465 470 475 480
 Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly
 485 490 495
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 500 505 510
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu

 515 520 525
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
 530 535 540
 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Lys
 545 550 555 560
 Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys
 565 570 575
 Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln Glu
 580 585 590

 Glu Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Arg Glu Ile Ser
 595 600 605
 Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Asn Gln
 610 615 620
 Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp Ser Trp Glu Asn
 625 630 635 640
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
 645 650 655

Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala

660 665 670
Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser
675 680 685
Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Pro Asp Arg Leu Gly Arg
690 695 700
Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile Arg Leu
705 710 715 720
Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser Leu Cys
725 730 735

Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg
740 745 750
Ala Val Glu Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly
755 760 765
Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp Gly Leu
770 775 780
Glu Leu Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala Ile Ala
785 790 795 800
Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys

805 810 815
Arg Ala Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala
820 825 830

Ala Leu Leu

835

<210> 82

<211> 840

<212> PRT

<213> Human immunodeficiency virus

<400> 82

Met Arg Val Arg Gly Ile Gln Arg Asn Tyr Gln His Leu Trp Arg Trp
1 5 10 15
Gly Ile Met Leu Leu Gly Met Leu Met Ile Cys Ser Val Ala Glu Asn

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

Thr Asn Asn Ala Lys Ile Ile Ile Val Gln Leu Asn Glu Ser Val Thr
 275 280 285
 Ile Asn Cys Thr Arg Pro Tyr Asn Asn Thr Arg Gln Arg Thr Pro Ile
 290 295 300
 Gly Pro Gly Gln Ala Leu Tyr Thr Thr Arg Ile Lys Gly Asp Ile Arg
 305 310 315 320

 Gln Ala His Cys Asn Ile Ser Arg Ala Glu Trp Asn Lys Thr Leu Gln
 325 330 335
 Gln Val Ala Lys Lys Leu Gly Asp Leu Leu Asn Lys Thr Thr Ile Ile
 340 345 350
 Phe Lys Pro Ser Ser Gly Gly Asp Pro Glu Ile Thr Thr His Ser Phe
 355 360 365
 Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Arg Leu Phe Asn
 370 375 380
 Ser Thr Trp Asn Asn Thr Lys Trp Asn Ser Thr Gly Lys Ile Thr Leu

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

 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
 465 470 475 480
 Ala Pro Thr Arg Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala
 485 490 495
 Ile Gly Leu Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser
 500 505 510
 Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu
 515 520 525

Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu

530

535

540

Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu

545

550

555

560

Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu

565

570

575

Leu Gly Ile Trp Gly Cys Ser Gly Lys His Ile Cys Thr Thr Thr Val

580

585

590

Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asn

595

600

605

Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr Gly

610

615

620

Leu Ile Tyr Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn

625

630

635

640

Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp

645

650

655

Phe Ser Ile Thr Gln Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile

660

665

670

Val Gly Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Leu

675

680

685

Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu

690

695

700

Leu Pro Ala Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu

705

710

715

720

Gly Gly Glu Gln Gly Arg Gly Arg Ser Ile Arg Leu Val Asn Gly Phe

725

730

735

Ser Ala Leu Ile Trp Asp Asp Leu Arg Asn Leu Cys Leu Phe Ser Tyr

740

745

750

His Arg Leu Arg Asp Leu Ile Leu Ile Ala Ala Arg Ile Val Glu Leu

755

760

765

Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Leu Trp Asn Leu Leu

770	775	780	
Gln Tyr Trp Ile Gln Glu Leu Lys Asn Ser Ala Ile Ser Leu Phe Asp			
785	790	795	800
Thr Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Ile			
805	810	815	
Val Gln Arg Ala Cys Arg Ala Ile Leu Asn Ile Pro Thr Arg Ile Arg			
820	825	830	
Gln Gly Leu Glu Arg Ala Leu Leu			
835	840		
<210>	83		
<211>	2508		
<212>	DNA		
<213>	Human immunodeficiency virus		
<400>	83		
atgcgcgtga tgggcatcct gcgcaactgc cagcagtggt ggatctgggg catcctgggc		60	
ttctggatgc tgatgatctg caacgtgggt ggcaacctgt gggtgaccgt gtactacggc		120	
gtgcccggtgt ggaaggagge caagaccacc ctgttctgcg cctccgacgc caaggcctac		180	
gagcgcgagg tgcacaacgt gtgggccacc cagcctgcg tgcccaccga ccccaacccc		240	
caggagatgg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg		300	
gaccagatgc acgaggacat catctccctg tgggaccagt cctgaagcc ctgcgtgaag		360	
ctgaccccc tgtgcgtgac cctgaactgc accaacgcca ccaacgccac caacaccatg		420	
ggcgagatga agaactgctc cttcaacatc accaccgagc tgcgcgacaa gaagcagaag		480	
gtgtacgcc tgttctaccg cctggacatc gtgccctga acgacaacaa ctctaccgc		540	
ctgatcaact gcaacacctc cgccatcacc caggcctgcc ccaaggtgtc cttcgacccc		600	
atccccatcc actactgcgc ccccgccggc tacgcatcc tgaagtgcaa caacaagacc		660	
ttcaacggca cgggccctg caacaacgtg tccaccgtgc agtgcacca cggcatcaag		720	
cccgtggtgt ccaccagct gctgtgaac ggctccctgg ccgaggagga gatcatcatc		780	
cgctccgaga acctgaccga caacgccaag accatcatcg tgcacctgaa cgagtccgtg		840	
gagatcgtgt gcacccggcc caacaacaac acccgcaagt ccatccgcat cggccccggc		900	
cagaccttct acgccaccgg cgacatcatc ggcgacatcc gccaggccca ctgcaacatc		960	
tccgaggaga agtggaaaca gacctgcag cgcgtgggcg agaagctgaa ggagcacttc		1020	

cccaacaaga ccatcaagtt cgccccctcc tccggcggcg acctggagat caccacccac 1080

 tccttcaact gccgcggcga gttctttctac tgcaacacct cccgcctgtt caactccacc 1140
 tacaactcca agaactccac catcaccctg ccttgccgca tcaagcagat catcaacatg 1200
 tggcagggcg tgggccgcgc catgtacgcc ccccccatcg ccggcaacat cacctgcaag 1260
 tccaacatca ccggcctgct gctgacccgc gacggcggca agaacaacac cgagaccttc 1320
 cgccccggcg gggcgacat gcgcgacaac tggcgctccg agctgtacaa gtacaaggtg 1380
 gtggagatca agccccctggg catcgcccc accgaggcca agcggccgct ggtggagcgc 1440
 gagaagcgcg ccgtgggcat cggcgccgtg ttcttgggct tcttgggcgc cgccggctcc 1500

 accatgggcg ccgcctccat caccctgacc gtgcaggccc gccagctgct gtccggcatc 1560
 gtgcagcagc agtccaacct gctgcgcgcc atcgaggccc agcagcacat gctgcagctg 1620
 accgtgtggg gcatcaagca gctgcagacc cgcgtgtctgg ccatcgagcg ctacctgaag 1680
 gaccagcagc tgctgggcat ctggggctgc tccggcaagc tgatctgcac caccgccgtg 1740
 ccctggaact cctcctggtc caacaagtcc caggaggaga tctgggacaa catgacctgg 1800
 atgcagtggg accgcgagat ctccaactac accgacacca tctaccgcct gctggaggac 1860
 tcccagaacc agcaggagaa gaacgagcag gacctgtctg ccctggactc ctgggagaac 1920

 ctgtggaact ggttcgacat caccaactgg ctgtggtaca tcaagatctt catcatgac 1980
 gtgggcggcc tgatcggcct gcgcatcatc ttgcgcgtgc tgtccatcgt gaaccgcgtg 2040
 cgccagggtt actccccctt gtccttcag accctgacct ccaacccccg cggccccgac 2100
 cgcttgggc gcatcgagga ggaggcggc gagcaggacc gcgaccgtc catccgcctg 2160
 gtgtccggtt tcttggcctt ggcctgggac gacctgcgt cctgtgcct gttctcctac 2220
 caccgcctgc gcgacttcct cctgatcgcc gcccgcccg tggagctgct gggccgctcc 2280
 tccctgcgcg gcctgcagcg cggctgggag gccctgaagt acctgggctc cctggtgcag 2340

 tactggggcc tggagctgaa gaagtccgc atctccctgc tggacacat cgccatcgcc 2400
 gtggccgagg gcaccgacc catcatcgag ctgatecage gcatctgccg cgccatccgc 2460
 aacatcccc gccgcatccg ccagggttc gaggcgccc tgctgtaa 2508

 <210> 84
 <211> 2523
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 84
 atgcgcgtgc gcggcatcca gcgcaactac cagcacctgt ggcgtgggg catcatgctg 60

ctgggcatgc tgatgatctg ctccgtggcc gagaacctgt gggtgaccgt gtactacggc	120
gtgccctgtt ggaaggaggg caccaccacc ctgtttctgc cctccgacgc caagtcttac	180
aagaccgagg ccacaacat ctgggccacc caegcctgcg tgeccaccga ccccaacccc	240
caggagatcg agctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300
gagcagatgc acgaggacat catctccctg tgggaccagt cctgaagcc ctgctgaag	360
ctgaccccc tgctgctgac cctgaactgc accgacgtga agcgcaaaa cacctccaac	420
gacaccaacg agggcgagat gaagaactgc tccttcaaca tcaccaccga gatccgcgac	480
aagaagaagc aggtgcacgc cctgtttctac aagctggacg tgggtcccat cgacgacaac	540
aactccaaca cctcctaccg cctgatcaac tgcaaacctt ccgcatcac ccaggcctgc	600
cccaaggtga ccttcgagcc catcccatc cactactgcg ccccgccgg cttcgccatc	660
ctgaagtga aggacaagaa gttcaacggc accggccctt gcaagaacgt gtccaccgtg	720
cagtgcacc acggcatccg ccccggtgtg tccaccagc tgcgtctgaa cggctccctg	780
gccgaggagg agatcatcat ccgctccgag aacctgacca acaacgcaa gatcatcatc	840
gtgcagctga acgagtcgt gaccatcaac tgcaccgccc cctacaacaa caccgcccag	900
cgcaccccc tgggccccgg ccaggccctg tacaccacc gcatcaaggg cgacatccgc	960
caggcccact gcaacatctc ccgcgccgag tggaacaaga ccctgcagca ggtggccaag	1020
aagctgggcg acctgctgaa caagaccacc atcatcttca agccctctc cggcggcgac	1080
cccgagatca ccaccactc cttcaactgc ggcggcgagt tcttctactg caacacctcc	1140
cgctgttca actccactg gaacaacacc aagtggaaact ccaccggcaa gatcacctg	1200
ccctgccga tcaagcagat catcaaatg tggcagggcg tgggcaaggc catgtacgcc	1260
cccccatcg agggcctgat caagtgtcc tccaacatca ccggcctgct gctgaccgc	1320
gacggcggcg ccaacaactc ccacaacgag accttcgcc ccggcggcgg cgacatgcgc	1380
gacaactggc gctccgagct gtacaagtac aaggtggtga agatcgagcc cctgggcgtg	1440
gccccaccc gcgccaagcg ccgctgtgtg gagcgcgaga agcgcgcat cggcctgggc	1500
gccatgttcc tgggcttcc gggcgccgc ggtccacca tggcgccgc ctccatgacc	1560
ctgaccgtgc agggccgcca gctgtgtcc ggcatcgtgc agcagcagaa caacctgctg	1620
cgcgccatcg agggccagca gcaactgctg cagctgaccg tgtggggcat caagcagctg	1680
caggccccga tcttgccgt ggagcgctac ctgaaggacc agcagctgct gggcatctgg	1740
ggctgtctcg gcaagcacat ctgcaccacc accgtgccct ggaactctc ctggtccaac	1800

aagtccttgg acgagatctg gaacaacatg acctggatgg agtgggagcg cgagatcgac 1860

aactacaccg gcctgatcta ctccctgata gaggagtccc agaaccagca ggagaagaac 1920

gagcaggagc tgcctggagct ggacaagtgg gcctccctgt ggaactggtt ctccatcacc 1980

cagtggctgt ggtacatcaa gatcttcata atgatcgtgg gcggcctgat cggcctgcgc 2040

atcgtgttcg ccgtgctgtc cctggtgaac cgcgtgcgcc agggctactc cccctgtcc 2100

ttccagaccg tgcctgccgc ccccgcggc cccgaccgcc ccgagggcat cgaggaggag 2160

ggcggcgagc agggcccgcg ccgctccatc cgctggtga acggcttctc cgccctgata 2220

tgggacgacc tgcgcaacct gtgcctgttc tcctaccacc gcctgcgcga cctgatactg 2280

atcgccgccc gcatcgtgga gctgctgggc cgccgcggct gggaggccct gaagtacctg 2340

tggaacctgc tgcagtactg gatccaggag ctgaagaact ccgccatctc cctgttcgac 2400

accaccgcca tcgccgtggc cgagggcacc gaccgcgtga tcgagatcgt gcagcgcgcc 2460

tgccgcgcca tcctgaacat cccacccgc atccgccagg gcctggagcg cgccctgctg 2520

taa 2523

<210> 85

<211> 832

<212> PRT

<213> Human immunodeficiency virus

<400> 85

Met Arg Val Arg Gly Met Gln Arg Asn Trp Gln His Leu Gly Lys Trp

1 5 10 15

Gly Leu Leu Phe Leu Gly Ile Leu Ile Ile Cys Asn Ala Ala Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ser Tyr Glu Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asp Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp

100 105 110
 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asp Val Asn Ala Thr Asn Asn Asp Thr Asn Asp Asn Lys
 130 135 140
 Thr Gly Ala Ile Gln Asn Cys Ser Phe Asn Met Thr Thr Glu Val Arg

 145 150 155 160
 Asp Lys Lys Leu Lys Val His Ala Leu Phe Tyr Lys Leu Asp Ile Val
 165 170 175
 Pro Ile Ser Asn Asn Asn Ser Lys Tyr Arg Leu Ile Asn Cys Asn Thr
 180 185 190
 Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Ser Trp Asp Pro Ile Pro
 195 200 205
 Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asp
 210 215 220

 Lys Arg Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln
 225 230 235 240
 Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn
 245 250 255
 Gly Ser Leu Ala Glu Glu Asp Ile Ile Ile Arg Ser Gln Asn Ile Ser
 260 265 270
 Asp Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Ser Val Gln Ile
 275 280 285
 Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Leu Gly

 290 295 300
 Pro Gly Gln Ala Phe Tyr Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg
 305 310 315 320
 Lys Ala His Cys Asn Ile Ser Gly Thr Gln Trp Asn Lys Thr Leu Glu
 325 330 335
 Gln Val Lys Ala Lys Leu Lys Ser His Phe Pro Asn Lys Thr Ile Lys
 340 345 350
 Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Met His Ser Phe

355 Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn
 370 375 380
 Asp Thr Gly Ser Asn Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln
 385 390 395 400
 Ile Val Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Ala Pro
 405 410 415
 Ile Ala Gly Asn Ile Thr Cys Asn Ser Asn Ile Thr Gly Leu Leu Leu
 420 425 430
 Thr Arg Asp Gly Gly Gln Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly
 435 440 445
 Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val
 450 455 460
 Val Glu Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Gln
 465 470 475 480
 Val Val Lys Arg Glu Arg Arg Ala Val Gly Ile Gly Ala Val Phe Leu
 485 490 495
 Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr
 500 505 510
 515 520 525
 Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln
 Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu
 530 535 540
 Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu
 545 550 555 560
 Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly
 565 570 575
 Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn
 580 585 590
 Lys Ser Gln Asp Glu Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu
 595 600 605

Lys Glu Ile Ser Asn Tyr Ser Asn Ile Ile Tyr Arg Leu Ile Glu Glu
 610 615 620
 Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp
 625 630 635 640
 Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp
 645 650 655

 Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg
 660 665 670
 Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Lys Gly Tyr
 675 680 685
 Ser Pro Leu Ser Leu Gln Thr Leu Ile Pro Ser Pro Arg Glu Pro Asp
 690 695 700
 Arg Pro Glu Gly Ile Glu Glu Gly Gly Gly Glu Gln Gly Lys Asp Arg
 705 710 715 720
 Ser Val Arg Leu Val Asn Gly Phe Leu Ala Leu Val Trp Asp Asp Leu

 725 730 735
 Arg Asn Leu Cys Leu Phe Ser Tyr Arg His Leu Arg Asp Phe Ile Leu
 740 745 750
 Ile Ala Ala Arg Ile Val Asp Arg Gly Leu Arg Arg Gly Trp Glu Ala
 755 760 765
 Leu Lys Tyr Leu Gly Asn Leu Thr Gln Tyr Trp Ser Gln Glu Leu Lys
 770 775 780
 Asn Ser Ala Ile Ser Leu Leu Asn Thr Thr Ala Ile Val Val Ala Glu
 785 790 795 800

 Gly Thr Asp Arg Val Ile Glu Ala Leu Gln Arg Ala Gly Arg Ala Val
 805 810 815
 Leu Asn Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu
 820 825 830
 <210> 86
 <211> 831
 <212> PRT
 <213> Human immunodeficiency virus

<400> 86

Met Arg Val Arg Glu Met Gln Arg Asn Trp Gln His Leu Gly Lys Trp

1 5 10 15

Gly Leu Leu Phe Leu Gly Ile Leu Ile Ile Cys Asn Ala Ala Asp Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Arg Glu Val

50 55 60

His Asn Val Trp Ala Thr Tyr Ala Cys Val Pro Thr Asp Pro Ser Pro

65 70 75 80

Gln Glu Leu Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Val Asn Val Thr Ile Asn Thr Thr Asn Val Thr Leu

130 135 140

Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Lys Asp

145 150 155 160

Lys Lys Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro

165 170 175

Ile Asn Asn Ser Ile Val Tyr Arg Leu Ile Ser Cys Asn Thr Ser Thr

180 185 190

Val Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His

195 200 205

Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys

210 215 220

Phe Asn Gly Thr Gly Leu Cys Arg Asn Val Ser Thr Val Gln Cys Thr

225 230 235 240

His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser
 245 250 255
 Leu Ala Glu Glu Asp Ile Ile Ile Arg Ser Glu Asn Ile Ser Asp Asn
 260 265 270
 Thr Lys Thr Ile Ile Val Gln Phe Asn Arg Ser Val Glu Ile Asn Cys
 275 280 285
 Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly
 290 295 300
 Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Lys Ala

 305 310 315 320
 Tyr Cys Asn Ile Asn Arg Thr Leu Trp Asn Glu Thr Leu Lys Lys Val
 325 330 335
 Ala Glu Glu Phe Lys Asn His Phe Asn Ile Thr Val Thr Phe Asn Pro
 340 345 350
 Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg
 355 360 365
 Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asp Leu Phe Asn Asn Thr Glu
 370 375 380

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

 450 455 460
 Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Gln Val
 465 470 475 480
 Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Leu Leu Gly
 485 490 495

Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu
 500 505 510
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
 515 520 525

 Asn Leu Leu Lys Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr
 530 535 540
 Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu Arg
 545 550 555 560
 Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys
 565 570 575
 Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys
 580 585 590
 Ser Gln Asp Glu Ile Trp Asp Asn Met Thr Trp Met Gln Trp Glu Lys

 595 600 605
 Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Ile Glu Asp Ala
 610 615 620
 Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp Lys
 625 630 635 640
 Trp Asp Asn Leu Trp Ser Trp Phe Thr Ile Thr Asn Trp Leu Trp Tyr
 645 650 655
 Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile
 660 665 670

 Val Phe Ala Val Leu Ser Val Val Asn Arg Val Arg Gln Gly Tyr Ser
 675 680 685
 Pro Leu Ser Leu Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Glu Arg
 690 695 700
 Pro Gly Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser
 705 710 715 720
 Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg
 725 730 735
 Ser Leu Cys Leu Phe Ser Tyr Arg His Leu Arg Asp Phe Ile Leu Ile

740	745	750	
Ala Ala Arg Thr Val Asp Met Gly Leu Lys Arg Gly Trp Glu Ala Leu			
755	760	765	
Lys Tyr Leu Trp Asn Leu Pro Gln Tyr Trp Gly Gln Glu Leu Lys Asn			
770	775	780	
Ser Ala Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val Ala Glu Gly			
785	790	795	800
Thr Asp Arg Ile Ile Glu Val Leu Gln Arg Ala Gly Arg Ala Val Leu			
805	810	815	
His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu			
820	825	830	
<210>	87		
<211>	2499		
<212>	DNA		
<213>	Human immunodeficiency virus		
<400>	87		
atgcgcgtgc gcggcatgca gcgcaactgg cagcacctgg gcaagtgggg cctgctgttc	60		
ctgggcatcc tgatcatctg caacgccgcc gagaacctgt gggtagacct gtactacggc	120		
gtgccctgtt ggaaggaggg caccaccacc ctgttctgcg cctccgacgc caagtcctac	180		
gagaaggagg tgcacaacgt gtgggccacc cagcctgcg tgeccaccga cccaacccc	240		
caggaggtgg tgctggagaa cgtgaccgag aacttcgaca tgtggaagaa caacatggtg	300		
gagcagatgc acaccgacat catctccctg tgggaccagt ccctgaagcc ctgcgtgaag	360		
ctgaccccc tgtagctgac cctgaactgc accgacgtga acgccaccaa caacgacacc	420		
aacgacaaca agaccggcgc catccagaac tgctccttca acatgaccac cgaggtgcgc	480		
gacaagaagc tgaaggtgca cgccctgttc tacaagctgg acatcgtgcc catctccaac	540		
aacaactcca agtaccgcct gatcaactgc aacacctcca ccatcaccca ggccctgcccc	600		
aaggtgtcct gggaccccat ccccatccac tactgcgccc ccgccggcta cgccatcctg	660		
aagtgcacg acaagcgctt caacggcacc ggccccgtga agaactgtc caccgtgcag	720		
tgcaccacg gcatcaagcc cgtggtgtcc acccagctgc tgctgaacgg ctccctggcc	780		
gaggaggaca tcatcatccg ctcccagaac atctccgaca acgccaagac catcatcgtg	840		
cacctgaacg agtccgtgca gatcaactgc acccgcccca acaacaacac ccgcaagtcc	900		
atccacctgg gccccggcca ggcccttctac gccaccggcg agatcatcgg cgacatccgc	960		

aaggccact gcaacatctc cggcaccag tggaacaaga ccctggagca ggtgaaggcc	1020
aagctgaagt ccacttccc caacaagacc atcaagtca actcctctc cggcggcgac	1080
ctggagatca ccatgcactc cttaactgc cggcgaggt tcttctactg caacacctcc	1140
ggcctgttca acgacaccgg ctccaacggc accatcacc tgccctgccg catcaagcag	1200
atcgtgaaca tgtggcagga ggtgggccc gccatgtacg cgcgcccat cgccggcaac	1260
atcacctgca actccaacat caccggcctg ctgctgaccc gcgacggcgg ccagaacaac	1320
accgagacct tcgccccgg cggcggcaac atgaaggaca actggcgctc cgagctgtac	1380
aagtacaagg tggtaggat cgagccctg ggcgtggccc ccaccaaggc caagcggcag	1440
gtggtgaagc gcgagcgccg cgccgtgggc atcgcgccg tgttcctggg ctctctgggc	1500
gccgccggt ccacatggg cgccgctcc atcacctga ccgtgcaggc ccgccagctg	1560
ctgtccggca tcgtgcagca gcagaacaac ctgctgcggc ccacgcaggc ccagcagcac	1620
ctgctgcagc tgaccgtgtg gggcatcaag cagctgcagg cccgcgtgct ggccgtggag	1680
cgctacctga aggaccagca gctgctgggc ctgtggggct gctccggcaa gctgatctgc	1740
accaccaacg tgccctggaa ctctcctgg tccaacaagt ccagagcga gatctggaac	1800
aacatgacct ggatggagt ggagaaggag atctccaact actccaacat catctaccgc	1860
ctgatcgagg agtcccagaa ccagcaggag aagaacgagc aggagctgct ggccctggac	1920
aagtgggcct ccctgtggaa ctggttcgac atctccaact ggctgtggta catcaagatc	1980
tcatcatga tcgtggcgcc cctgatcggc ctgcgcatcg tgttcgccg gctgtccatc	2040
gtgaaccgcg tgcgaagggt ctactcccc ctgtccctgc agaccctgat cccctcccc	2100
cgcgagccc accgccccga gggcatcgag gagggcggcg gcgagcagg caaggaccgc	2160
tccgtgcgcc tggtagacgg ctctctggcc ctggtgtggg acgacctgcg caacctgtgc	2220
ctgttctct accgccaact gcgcgacttc atcctgatcg ccgccgcat cgtggaccgc	2280
ggcctgcgcc gcggctggga ggccctgaag tacctgggca acctgacca gtactggtcc	2340
caggagctga agaactccgc catctccctg ctgaacacca ccgccatcgt ggtggccgag	2400
ggcaccgacc gcgtgatcga ggccctgcag cgcgccggcc gcgccgtgct gaacatcccc	2460
cgccgcatcc gccaggcct ggagcgcgcc ctgctgtaa	2499
<210> 88	
<211> 2496	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 88	

atgcgcgtgc gcgagatgca gcgcaactgg cagcacctgg gcaagtgggg cctgctgttc	60
ctgggcatcc tgatcatctg caacgccgcc gacaacctgt gggtgaccgt gtactacggc	120
gtgcccgtgt ggaaggaggc caccaccacc ctgttctgcg cctccgacgc caaggcctac	180
gagcgcgagg tgcacaacgt gtgggccacc tacgcctgcg tgcccaccga cccctcccc	240
caggagctgg tgctgggcaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300
gaccagatgc acgaggacat catctccctg tgggaccagt cctgaagcc ctgcgtgaag	360
ctgaccccc tgigcgtgac cctgaactgc accgacgtga acgtgaccat caacaccacc	420
aacgtgacc tgggcgagat caagaactgc tccttcaaca tcaccaccga gatcaaggac	480
aagaagaaga aggagtacgc cctgttctac cgcctggacg tggtgcccat caacaactcc	540
atcgtgtacc gcctgatctc ctgcaacacc tccaccgtga cccaggcctg cccaaggtg	600
tccttegagc ccatecccat ccactactgc gccccgccg gcttcgcat cctgaagtgc	660
aacgacaaga agttcaacgg caccggcctg tgccgcaacg tgtccaccgt gcagtgcacc	720
cacggcatcc gccccgtggt gtccaccag ctgctgtga acggctcct ggccgaggag	780
gacatcatca tccgctccga gaacatctcc gacaacacca agaccatcat cgtgcagttc	840
aaccgctccg tggagatcaa ctgcaccgc cccaacaaca acaccgcaa gtccatccgc	900
atcgccccg gccgcgctt ctacgccacc ggcgacatca tcggcgacat ccgcaaggcc	960
tactgcaaca tcaaccgcac cctgtggaac gagaccctga agaagggtggc cgaggagttc	1020
aagaaccact tcaacatcac cgtgaccttc aaccctcct ccggcggcga cctggagatc	1080
accaccact cttcaactg ccgcggcgag ttcttctact gcaacacctc cgacctgttc	1140
aacaacaccg aggtgaacaa caccaagacc atcacctgc cctgccgat ccgccagttc	1200
gtgaacatgt ggcagcgcgt gggccgcgcc atgtacgcc ccccatcgc cggccagatc	1260
cagtgcact ccaacatcac cggcctgtg ctgaccgcg acggcggcaa gaacggctcc	1320
gagaccctgc gccccggcgg cggcgacatg cgcgacaact ggcgctccga gctgtacaag	1380
tacaaggtgg tgaagatcga gccctgggc gtggcccca ccaaggcaa gcgccaggtg	1440
gtgcagcgcg agaagcgcgc cgtgggcac cgcgccgtgc tgctgggctt cctgggcgcc	1500
gccggctcca ccatgggcgc cgcctccatc accctgaccg tgcaggcccg ccagctgctg	1560
tccggcatcg tgcagcga gtccaactg ctgaaggcca tcgaggccca gcagcacctg	1620
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcatcctggc cgtggagcgc	1680
tacctgaagg accagcagct gctgggcac tggggctgct ccggcaagct gatctgcacc	1740
accaactgac cctggaactc ctctgggtcc aacaagtccc aggacgagat ctgggacaac	1800

atgacctgga tgcagtggga gaaggagatc tccaactaca ccgacaccat ctaccgcctg 1860

atcgaggacg cccagaacca gcaggagaag aacgagcagg acctgctggc cctggacaag 1920

tgggacaacc tgttgtcctg gttcaccatc accaactggc tgtgtacat caagatcttc 1980

atcatgatcg tgggcggcct gatcggcctg cgcacgtgtg tcgccgtgct gtccgtggtg 2040

aaccgcgtgc gccagggcta ctccccctg tcctgcaga ccctgatecc caacccccgc 2100

ggccccgagc gccccggcgg catcgaggag gagggcggcg agcaggaccg cgaccgtcc 2160

atccgcctgg tgtccggctt cctggccctg gcctgggacg acctgcgctc cctgtgcctg 2220

ttctcctacc gccacctgcg cgacttcac ctagtcgccg cccgcaccgt ggacatgggc 2280

ctgaagcgcg gctgggaggc cctgaagtac ctgtggaacc tgccccagta ctggggccag 2340

gagctgaaga actccgcat ctccctgctg gacaccaccg ccatcgccgt ggccgagggc 2400

accgaccga tcacgaggt gctgcagcgc gccggccgcg ccgtgctgca catccccgc 2460

cgcacccgc agggcttcga gcgcgccctg ctgtaa 2496

<210> 89

<211> 842

<212> PRT

<213> Human immunodeficiency virus

<400> 89

Met Arg Val Lys Gly Ile Gln Arg Asn Trp Gln His Leu Trp Lys Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asn Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Glu Asp Ala Asp

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Ser Thr Glu Arg

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Thr Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125
 Asn Cys Thr Asp Val Asn Val Thr Asn Asn Asn Thr Asn Asn Thr Lys
 130 135 140
 Lys Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp
 145 150 155 160

 Lys Lys Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro
 165 170 175
 Ile Asn Asp Asn Gly Asn Ser Ser Ile Tyr Arg Leu Ile Asn Cys Asn
 180 185 190
 Val Ser Thr Ile Lys Gln Ala Cys Pro Lys Val Thr Phe Asp Pro Ile
 195 200 205
 Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Arg
 210 215 220
 Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val

 225 230 235 240
 Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu
 245 250 255
 Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile
 260 265 270
 Thr Asp Asn Thr Lys Val Ile Ile Val Gln Leu Asn Glu Thr Ile Glu
 275 280 285
 Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile
 290 295 300

 Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile
 305 310 315 320
 Arg Gln Ala His Cys Asn Val Ser Arg Thr Lys Trp Asn Glu Met Leu
 325 330 335
 Gln Lys Val Lys Ala Gln Leu Lys Lys Ile Phe Asn Lys Ser Ile Thr
 340 345 350
 Phe Asn Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe
 355 360 365

Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn

370 375 380

Asn Ser Leu Leu Asn Ser Thr Asn Ser Thr Ile Thr Leu Pro Cys Lys

385 390 395 400

Ile Lys Gln Ile Val Arg Met Trp Gln Arg Val Gly Gln Ala Met Tyr

405 410 415

Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly

420 425 430

Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Glu Thr Phe Arg

435 440 445

Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys

450 455 460

Tyr Lys Ile Val Lys Ile Lys Pro Leu Gly Val Ala Pro Thr Arg Ala

465 470 475 480

Arg Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Leu Gly Ala

485 490 495

Val Leu Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala

500 505 510

Ser Ile Thr Leu Thr Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val

515 520 525

Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu

530 535 540

Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu

545 550 555 560

Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly

565 570 575

Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Thr Ser

580 585 590

Trp Ser Asn Lys Ser Tyr Asn Glu Ile Trp Asp Asn Met Thr Trp Ile

595 600 605

Glu Trp Glu Arg Glu Ile Ser Asn Tyr Thr Gln Gln Ile Tyr Ser Leu

610 615 620
 Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu
 625 630 635 640
 Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Lys
 645 650 655
 Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile

 660 665 670
 Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg
 675 680 685
 Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr His His Gln Arg
 690 695 700
 Glu Pro Asp Arg Pro Glu Arg Ile Glu Glu Gly Gly Gly Glu Gln Asp
 705 710 715 720
 Lys Asp Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp
 725 730 735

 Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
 740 745 750
 Phe Ile Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Ser Ser
 755 760 765
 Leu Lys Gly Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr Leu Trp Asn
 770 775 780
 Leu Leu Leu Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala Ile Asn Leu
 785 790 795 800
 Leu Asp Thr Ile Ala Ile Ala Val Ala Asn Trp Thr Asp Arg Val Ile

 805 810 815
 Glu Val Ala Gln Arg Ala Cys Arg Ala Ile Leu Asn Ile Pro Arg Arg
 820 825 830
 Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu
 835 840
 <210> 90
 <211> 839
 <212> PRT

<213> Human immunodeficiency virus

<400> 90

Thr Arg Val Met Glu Thr Gln Arg Asn Tyr Pro Ser Leu Trp Arg Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Leu Ile Cys Ser Ala Ala Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Thr Glu Lys

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Glu

85 90 95

Asn Asp Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asp Cys Ser Asn Val Asn Thr Thr Asn Ala Thr Asn Ser Arg Phe Asn

130 135 140

Met Gln Glu Glu Leu Thr Asn Cys Ser Phe Asn Val Thr Thr Val Ile

145 150 155 160

Arg Asp Lys Gln Gln Lys Val His Ala Leu Phe Tyr Arg Leu Asp Val

165 170 175

Val Pro Ile Asp Asp Asn Asn Ser Tyr Gln Tyr Arg Leu Ile Asn Cys

180 185 190

Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro

195 200 205

Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys

210 215 220

Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr

225						230						235						240
Val	Gln	Cys	Thr	His	Gly	Ile	Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu			
					245						250						255	
Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Gln	Val	Ile	Ile	Arg	Ser	Lys	Asn			
					260						265						270	
Ile	Ser	Asp	Asn	Thr	Lys	Asn	Ile	Ile	Val	Gln	Leu	Asn	Lys	Pro	Val			
					275						280						285	
Glu	Ile	Thr	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	His			
					290						295						300	
Leu	Gly	Pro	Gly	Gln	Ala	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp			
305						310						315						320
Ile	Arg	Gln	Ala	His	Cys	Asn	Ile	Ser	Gly	Lys	Lys	Trp	Asn	Lys	Thr			
					325						330						335	
Leu	His	Gln	Val	Val	Thr	Gln	Leu	Gly	Lys	Tyr	Phe	Asp	Asn	Arg	Thr			
					340						345						350	
Ile	Ile	Phe	Lys	Pro	His	Ser	Gly	Gly	Asp	Met	Glu	Val	Thr	Thr	His			
					355						360						365	
Ser	Phe	Asn	Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Gly	Leu			
					370						375						380	
Phe	Asn	Ser	Ser	Trp	Thr	Asn	Ser	Thr	Asn	Asp	Thr	Lys	Asn	Ile	Ile			
385						390						395						400
Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Val	Asn	Met	Trp	Gln	Arg	Val			
					405						410						415	
Gly	Gln	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Lys	Gly	Asn	Ile	Thr	Cys	Val			
					420						425						430	
Ser	Asn	Ile	Thr	Gly	Leu	Ile	Leu	Thr	Phe	Asp	Glu	Gly	Asn	Asn	Thr			
					435						440						445	
Val	Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser			
					450						455						460	
Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile	Glu	Pro	Leu	Gly	Val	Ala			
465						470						475						480

Pro Thr Glu Ala Arg Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val
485 490 495

Gly Met Gly Ala Phe Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr
500 505 510

Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
515 520 525

Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Gln Ala
530 535 540

Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
545 550 555 560

Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu
565 570 575

Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro
580 585 590

Trp Asn Ser Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asp Asn
595 600 605

Met Thr Trp Met Glu Trp Asp Lys Gln Ile Asn Asn Tyr Thr Glu Glu
610 615 620

Ile Tyr Arg Leu Leu Glu Val Ser Gln Thr Gln Gln Glu Lys Asn Glu
625 630 635 640

Gln Asp Leu Leu Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe
645 650 655

Ser Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val
660 665 670

Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val
675 680 685

Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Ile
690 695 700

Pro Asn Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly
705 710 715 720

Gly Glu Gln Asp Arg Asp Arg Ser Val Arg Leu Val Asn Gly Phe Leu
725 730 735

Pro Leu Val Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr Arg

740 745 750
 Leu Leu Arg Asp Leu Leu Leu Ile Val Val Arg Thr Val Glu Leu Leu
 755 760 765
 Gly Arg Arg Gly Arg Glu Ala Leu Lys Tyr Leu Trp Asn Leu Leu Gln
 770 775 780
 Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala Ile Asn Leu Leu Asn Thr
 785 790 795 800
 Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Ile Val
 805 810 815

Gln Arg Ala Trp Arg Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln
 820 825 830
 Gly Phe Glu Arg Thr Leu Leu

835
 <210> 91
 <211> 2529
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 91

atgcgcgtga agggcatcca gcgcaactgg cagcacctgt ggaagtgggg caccctgac	60
ctgggcctgg tgatcatctg ctccgcctcc aacaacctgt gggtgaccgt gtactacggc	120
gtgcccggtgt gggaggacgc cgacaccacc ctgttctgcg cctccgacgc caaggcctac	180
tccaccgagc gccacaactg gtgggccacc cagcctgcg tgcccaccga cccaacccc	240
caggagatca cctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300
gagcagatgc acgaggacat catctccctg tgggacgagt cctgaagcc ctgcgtgaag	360
ctgaccccc tgtgcgtgac cctgaactgc accgacgtga acgtgaccaa caacaacacc	420
aacaacacca agaaggagat caagaactgc tcttcaaca tcaccaccga gatccgcgac	480
aagaagaaga aggagtacgc cctgttctac cgcctggacg tggtgcccat caacgacaac	540
ggcaactcct ccatctaccg cctgatcaac tgcaactgtt ccaccatcaa gcaggcctgc	600
cccaaggatga ccttcgacc catecccatc cactactgcg ccccgcggg cttcgccatc	660
ctgaagtgcc gcgacaagaa gttcaacggc accggcccct gcaagaactg gtccaccgtg	720

cagtgacccc acggcatcaa gcccgtgggtg tccaccacgc tgctgctgaa cggctccctg	780
gccgaggagg agatcatcat ccgctccgag aacatcaccg acaacaccaa ggtgatcatc	840
gtgcagctga acgagaccat cgagatcaac tgcaccgcgc ccaacaacaa caccgcgaag	900
tccatccgca tcggccccgg ccaggccttc tacgccaccg gcgacatcat cggcgacatc	960
cgccaggccc actgcaacgt gtcccgcacc aagtggaaag agatgctgca gaaggtgaag	1020
gcccagctga agaagatctt caacaagtcc atcaccttca actcctctc cggcggcgac	1080
ctggagatca ccaccactc cttcaactgc cgcggcgagt tcttctactg caacacctcc	1140
ggcctgttca acaactccct gctgaactcc accaactcca ccatcaccct gccctgcaag	1200
atcaagcaga tcgtgcgcat gtggcagcgc gtgggccagg ccatgtacgc ccccccatc	1260
gccggcaaca tcacctgccg ctccaacatc accggcctgc tgctgaccgc cgacggcggc	1320
aacaacaaca ccgagacctt ccgccccggc ggcgccgaca tgcgcgacaa ctggcgctcc	1380
gagctgtaca agtacaagat cgtgaagatc aagccccctgg gcgtggcccc caccgcgcgc	1440
cgcgccgcgc tgggtggagcg cgagaagcgc gccgtgggcc tgggcgccgt gctgctgggc	1500
ttcttgggcg ccgccggctc caccatgggc gccgcctcca tcacctgac cgtgcaggtg	1560
cgccagctgc tgtccggcat cgtgcagcag cagtccaacc tgctgcgcgc catcgaggcc	1620
cagcagcacc tgctgcagct gaccgtgtgg ggcatcaagc agctgcaggc ccgcgtgctg	1680
gccgtggagc gctacctgaa ggaccagcag ctgctgggca tctggggctg ctccggcaag	1740
ctgatctgca ccaccaacgt gccctggaac acctcctggt ccaacaagtc ctacaacgag	1800
atctgggaca acatgacctg gatcgagtgg gagcgcgaga tctccaacta caccagcag	1860
atctactccc tgatcgagga gtcccagaac cagcaggaga agaacgagca ggacctgctg	1920
gccctggaca agtgggcctc cctgtggaac tggttcgaca tcaccaagtg gctgtggtac	1980
atcaagatct tcatcatgat cgtgggcggc ctgatcggcc tgcgcacgt gtctgccgtg	2040
ctgtccatcg tgaaccgcgt gcgccagggc tactcccccc tgtccttcca gacctgacc	2100
caccaccagc gcgagccccg ccgccccgag cgcacagagg agggcgccgg cgagcaggac	2160
aaggaccgt ccatccgct ggtgtccggc ttcttggccc tggcctggga cgacctgcgc	2220
tcctgtgccc tgttctccta ccaccgctg cgcgacttca tctgatcgc cgcgccacc	2280
gtggagctgc tgggcgcctc ctccctgaag ggctgcgcgc tgggctggga gggcctgaag	2340
tacctgtgga acctgctgct gtactggggc caggagctga agaactccgc catcaacctg	2400
ctggacacca tcgccatcgc cgtggccaac tggaccgacc gcgtgatcga ggtggcccag	2460
cgcgcctgcc gcgccatcct gaacatcccc cgcgcacatc gccagggcct ggagcgcgcc	2520
ctgctgtaa	2529

<210> 92
 <211> 2520
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 92

```

accgcgtga tggagacca gcgcaactac ccctccctgt ggcgctgggg caccctgac 60

ctgggcatgc tgctgatctg ctccgccgcc ggcaacctgt gggtgaccgt gtactacggc 120
gtgccctgtt ggaaggagge caagaccacc ctgttctgcg cctccgacgc caaggcctac 180
gagaccgaga agcacaacgt gtgggccacc cagcctgctg tgcccaccga ccccaacccc 240
caggagatgg tgctggagaa cgtgaccgag aacttcaaca tgtgggagaa cgacatggtg 300
gagcagatgc acaccgacat catctccctg tgggaccagt ccctgaagcc ctgctgaag 360
ctgaccccc tgctgctgac cctggactgc tccaactga acaccacaa cgccaccaac 420
tcccgttca acatgcagga ggagctgacc aactgctcct tcaactgac caccgtgac 480

cgcgacaagc agcagaaggt gcacgccctg ttctaccgcc tggacgtggt gcccacgac 540
gacaacaact cctaccagta ccgctgac aactgcaaca cctccgtgat caccaggcc 600
tgcccaagg tgtccttga gcccacccc atccactact gcgccccgc cggttcgcc 660
atcctgaagt gcaacaaca gacctcaac ggacccggcc cctgcacaa cgtgtccacc 720
gtgcagtga cccacggcat ccgccccgtg gtgtccacc agctgctgct gaacggctcc 780
ctggccgagg agcaggtgat catccgtcc aagaacatct ccgacaacac caagaacatc 840
atcgtgcagc tgaacaagcc cgtggagatc acctgcacc gccccaacaa caacacccgc 900

aagtccatcc acctggggccc cggccaggcc ttctacgcca ccggcgacat catcgcgac 960
atccgccagg ccactgcaa catctccggc aagaagtga acaagacct gcaccaggtg 1020
gtgaccagc tgggcaagta cttcgacaac cgcacatca tcttcaagcc cactccggc 1080
ggcgacatgg aggtgaccac ccaactcttc aactgccgcg gcgagttctt ctactgcaac 1140
acctccggc tgttcaactc ctctggacc aactccacca acgacacaa gaacatcatc 1200
acctgcctt gccgcatcaa gcagatctg aacatgtggc agcgctggg ccaggccatg 1260
tacgcccccc ccatcaaggg caacatcacc tgcgtgtcca acatcaccgg cctgatcctg 1320

accttcgacg agggcaacaa caccgtgacc ttccgccccg gcggcggcga catgcgcgac 1380
aactggcgt cagagctga caagtacaag gtggtgaaga tcgagccctt gggcgtggcc 1440
cccaccgagg cccgcccgcg cgtggtggag cgcgagaagc gcgccgtggg catgggcgcc 1500
ttcttctggt gcttcttggg cgccgccggc tccacatgg gcgccgctc catcacctg 1560
    
```


accgtgcagg cccgccagct gctgtccggc atcgtgcagc agcagtcctc cctgctgcgc 1620
 gccatccagg ccagcagca catgctgcag ctgacctgtt ggggcatcaa gcagctgcag 1680
 gcccgcgtgc tggccgtgga gcgtacctg aaggaccagc agctgctggg catctggggc 1740

tgctccggca agctgatctg caccaccaac gtgccctgga actcctcctg gtccaacaag 1800
 tccttgagc agatctggga caacatgacc tggatggagt gggacaagca gatcaacaac 1860
 tacaccgagg agatctaccg cctgctggag gtgtccaga cccagcagga gaagaacgag 1920
 caggacctgc tggccctgga caagtgggccc tcctgtgga actggttctc catcaccaac 1980
 tggctgtggt acatcaagat cttcatcatg atcgtgggag gcctgatcgg cctgcgcac 2040
 atcttcgccc tgctgtccat cgtgaaccgc gtgcgccagg gctactcccc cctgtccttc 2100
 cagacctga tcccaaccc ccgcggcccc gaccgcccc agggcatcga ggaggagggc 2160

ggcgagcagg accgcgaccg ctccgtgcgc ctggtgaacg gcttcctgcc cctggtgtgg 2220
 gacgacctgc gtcctctgtg cctgttctcc taccgcctgc tgcgcgacct gctgctgatc 2280
 gtggtgcgca ccgtggagct gctgggcccgc cgcggccgag aggcctgaa gtacctgtgg 2340
 aacctgctgc agtactgggg ccaggagctg aagaactccg ccatcaacct gctgaacacc 2400
 accgcatcgc ccgtggccga gggcaccgac cgcacatcgc agatcgtgca gcgcgcctgg 2460
 cgcgccatcc tgcacatccc ccgcgcac cgcagggtc tcgagcgac cctgctgtaa 2520
 2520

<210> 93

<211> 854

<212> PRT

<213> Human immunodeficiency virus

<400> 93

Met Arg Val Lys Glu Thr Gln Met Asn Trp Pro Asn Leu Trp Lys Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asp Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Asp

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala His Glu Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65				70				75				80			
Gln	Glu	Ile	His	Leu	Glu	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys
				85				90				95			
Asn	Asn	Met	Val	Glu	Gln	Met	Gln	Glu	Asp	Val	Ile	Ser	Leu	Trp	Asp
				100				105				110			
Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu
				115				120				125			
Asn	Cys	Thr	Asn	Ala	Asn	Leu	Thr	Asn	Val	Asn	Asn	Ile	Thr	Asn	Val
				130				135				140			
Ser	Asn	Ile	Ile	Gly	Asn	Ile	Thr	Asn	Glu	Val	Arg	Asn	Cys	Ser	Phe
145				150				155				160			
Asn	Met	Thr	Thr	Glu	Leu	Arg	Asp	Lys	Lys	Gln	Lys	Val	His	Ala	Leu
				165				170				175			
Phe	Tyr	Lys	Leu	Asp	Ile	Val	Gln	Ile	Glu	Asp	Asn	Asn	Ser	Tyr	Arg
				180				185				190			
Leu	Ile	Asn	Cys	Asn	Thr	Ser	Val	Ile	Lys	Gln	Ala	Cys	Pro	Lys	Ile
				195				200				205			
Ser	Phe	Asp	Pro	Ile	Pro	Ile	His	Tyr	Cys	Thr	Pro	Ala	Gly	Tyr	Ala
210				215				220							
Ile	Leu	Lys	Cys	Asn	Asp	Lys	Asn	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Lys
225				230				235				240			
Asn	Val	Ser	Ser	Val	Gln	Cys	Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser
				245				250				255			
Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile
				260				265				270			
Arg	Ser	Glu	Asn	Leu	Thr	Asn	Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu
				275				280				285			
Asn	Lys	Ser	Val	Glu	Ile	Asn	Cys	Thr	Arg	Pro	Ser	Asn	Asn	Thr	Arg
290				295				300							
Thr	Ser	Ile	Thr	Ile	Gly	Pro	Gly	Gln	Val	Phe	Tyr	Arg	Thr	Gly	Asp
305				310				315				320			
Ile	Ile	Gly	Asp	Ile	Arg	Lys	Ala	Tyr	Cys	Glu	Ile	Asn	Gly	Thr	Lys

325 330 335
Trp Asn Glu Val Leu Lys Gln Val Thr Glu Lys Leu Lys Glu His Phe

340 345 350
Asn Asn Lys Thr Ile Ile Phe Gln Pro Pro Ser Gly Gly Asp Leu Glu

355 360 365
Ile Thr Met His His Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn

370 375 380
Thr Thr Lys Leu Phe Asn Asn Thr Cys Ile Gly Asn Glu Thr Met Glu

385 390 395 400
Gly Cys Asn Gly Thr Ile Ile Leu Pro Cys Lys Ile Lys Gln Ile Ile

405 410 415

Asn Met Trp Gln Gly Ala Gly Gln Ala Met Tyr Ala Pro Pro Ile Ser

420 425 430
Gly Arg Ile Asn Cys Val Ser Asn Ile Thr Gly Ile Leu Leu Thr Arg

435 440 445
Asp Gly Gly Ala Asn Asn Thr Asn Glu Thr Phe Arg Pro Gly Gly Gly

450 455 460
Asn Ile Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val

465 470 475 480
Gln Ile Glu Pro Leu Gly Ile Ala Pro Thr Arg Ala Lys Arg Arg Val

485 490 495
Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Met Ile Phe Gly

500 505 510
Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu

515 520 525
Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser

530 535 540
Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr

545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg

565 570 575

Tyr Leu Lys Asp Gln Lys Phe Leu Gly Leu Trp Gly Cys Ser Gly Lys
 580 585 590
 Ile Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Thr Trp Ser Asn Arg
 595 600 605
 Ser Phe Glu Glu Ile Trp Asn Asn Met Thr Trp Ile Glu Trp Glu Arg
 610 615 620
 Glu Ile Ser Asn Tyr Thr Asn Gln Ile Tyr Glu Ile Leu Thr Glu Ser

 625 630 635 640
 Gln Asn Gln Gln Asp Arg Asn Glu Lys Asp Leu Leu Glu Leu Asp Lys
 645 650 655
 Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr
 660 665 670
 Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile
 675 680 685
 Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser
 690 695 700

 Pro Leu Ser Phe Gln Thr Pro Thr His His Gln Arg Glu Pro Asp Arg
 705 710 715 720
 Pro Glu Arg Ile Glu Glu Gly Gly Gly Glu Gln Gly Arg Asp Arg Ser
 725 730 735
 Val Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg
 740 745 750
 Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile
 755 760 765
 Ala Ala Arg Thr Val Glu Leu Leu Gly His Ser Ser Leu Lys Gly Leu

 770 775 780
 Arg Arg Gly Trp Glu Gly Leu Lys Tyr Leu Gly Asn Leu Leu Leu Tyr
 785 790 795 800
 Trp Gly Gln Glu Leu Lys Ile Ser Ala Ile Ser Leu Leu Asp Ala Thr
 805 810 815
 Ala Ile Ala Val Ala Gly Trp Thr Asp Arg Val Ile Glu Val Ala Gln
 820 825 830

Gly Ala Trp Arg Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln Gly
835 840 845

Leu Glu Arg Ala Leu Leu
850

<210> 94
<211> 844
<212> PRT
<213> Human immunodeficiency virus
<400> 94

Met Arg Val Met Gly Ile Gln Lys Asn Tyr Pro Leu Leu Trp Arg Trp
1 5 10 15

Gly Met Ile Ile Phe Trp Ile Met Ile Ile Cys Asn Ala Glu Asn Leu
20 25 30

Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Glu Thr
35 40 45

Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His
50 55 60

Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln
65 70 75 80

Glu Ile His Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn
85 90 95

Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln
100 105 110

Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asp
115 120 125

Cys His Asn Asn Ile Thr Asn Ser Asn Thr Thr Asn Asn Asn Ala Gly
130 135 140

Glu Ile Lys Asn Cys Ser Phe Asn Met Thr Thr Glu Leu Arg Asp Lys
145 150 155 160

Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Gln Ile
165 170 175

Asn Lys Asn Asn Ser Gln Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala

180	185	190	
Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His			
195	200	205	
Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Glu			
210	215	220	
Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr			
225	230	235	240
His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser			
245	250	255	
Leu Ala Glu Glu Glu Ile Val Ile Arg Ser Glu Asn Ile Thr Asn Asn			
260	265	270	
Ala Lys Thr Ile Ile Val Gln Leu Val Lys Pro Val Lys Ile Asn Cys			
275	280	285	
Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Val Arg Ile Gly Pro Gly			
290	295	300	
Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala			
305	310	315	320
His Cys Asn Val Ser Arg Thr Lys Trp Asn Asn Thr Leu Gln Gln Val			
325	330	335	
Ala Thr Gln Leu Arg Lys Tyr Phe Asn Lys Thr Ile Ile Phe Ala Asn			
340	345	350	
Pro Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Gly			
355	360	365	
Gly Glu Phe Phe Tyr Cys Asn Thr Ser Glu Leu Phe Asn Ser Thr Trp			
370	375	380	
Asn Ser Thr Trp Asn Asn Thr Glu Lys Cys Ile Thr Leu Gln Cys Arg			
385	390	395	400
Ile Lys Gln Ile Val Asn Met Trp Gln Lys Val Gly Gln Ala Met Tyr			
405	410	415	
Ala Pro Pro Ile Gln Gly Val Ile Arg Cys Glu Ser Asn Ile Thr Gly			
420	425	430	

Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Ser Thr Asn Glu Thr
 435 440 445
 Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu
 450 455 460
 Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr
 465 470 475 480

 Arg Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Leu
 485 490 495
 Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly
 500 505 510
 Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly
 515 520 525
 Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln
 530 535 540
 His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg

 545 550 555 560
 Val Leu Ala Leu Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile
 565 570 575
 Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn
 580 585 590
 Ser Ser Trp Ser Asn Lys Thr Tyr Asn Asp Ile Trp Asp Asn Met Thr
 595 600 605
 Trp Leu Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr Asp Ile Ile Tyr
 610 615 620

 Asn Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp
 625 630 635 640
 Leu Leu Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile
 645 650 655
 Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly
 660 665 670
 Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Thr Ile Ile Asn Arg
 675 680 685

Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr His His

690 695 700

Gln Arg Glu Pro Asp Arg Pro Glu Arg Ile Glu Glu Gly Gly Gly Glu

705 710 715 720

Gln Asp Arg Asp Arg Ser Val Arg Leu Val Ser Gly Phe Leu Ala Leu

725 730 735

Ala Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu

740 745 750

Arg Asp Phe Val Leu Ile Ala Ala Arg Thr Val Glu Leu Leu Gly His

755 760 765

Ser Ser Leu Lys Gly Leu Arg Leu Gly Trp Glu Ala Leu Lys Tyr Leu

770 775 780

Gly Asn Leu Leu Ser Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala Ile

785 790 795 800

Asn Leu Leu Asp Thr Ile Ala Ile Ala Val Ala Asn Trp Thr Asp Arg

805 810 815

Val Ile Glu Ile Gly Gln Arg Ala Gly Arg Ala Ile Leu Asn Ile Pro

820 825 830

Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu

835 840

<210> 95

<211> 2565

<212> DNA

<213> Human immunodeficiency virus

<400> 95

atgcgcgtga aggagaccca gatgaactgg cccaacctgt ggaagtgggg caccctgac 60

ctgggcctgg tgatcatctg ctccgcctcc gacaacctgt gggtgaccgt gtactacggc 120

gtgccctgt ggcgcgacgc cgacaccacc ctgtttctgcg cctccgacgc caaggccac 180

gagaccgagg tgcacaactgt gtgggccacc cagcctgcg tgcccaccga cccaacccc 240

caggagatcc acctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300

gagcagatgc aggaggacgt gatctccctg tgggaccagt cctgaagcc ctgcgtgaag 360

ctgaccccc tgtgctgac cctgaactgc accaacgcca acctgaccaa cgtgaacaac	420
atcaccaacg tgtccaacat catcggaac atcaccaacg aggtgcgcaa ctgctccttc	480
aacatgacca ccgagctgcg cgacaagaag cagaaggtgc acgccctgtt ctacaagctg	540
gacatcgtgc agatcgagga caacaactcc taccgcctga tcaactgcaa cacctccgtg	600
atcaagcagg cctgccccaa gatctccttc gaccccatcc ccatccacta ctgcaccccc	660
gccggctacg ccatcctgaa gtgcaacgac aagaacttca acggcaccgg cccctgcaag	720
aacgtgtcct ccgtgcagtg caccacggc atcaagcccg tgggtgccac ccagctgctg	780
ctgaacggct cctggccga ggaggagatc atcatccgt ccgagaacct gaccaacaac	840
gccaagacca tcctcgtgca cctgaacaag tccgtggaga tcaactgcac ccgccctcc	900
aacaacaccc gcacctccat caccatcggc cccggccagg tgttctaccg caccggcgac	960
atcatcggcg acatccgcaa ggcctactgc gagatcaacg gcaccaagtg gaacgaggtg	1020
ctgaagcagg tgaccgagaa gctgaaggag cacttcaaca acaagacat catcttcag	1080
ccccctccg gcggcgacct ggagatcacc atgcaccact tcaactgccg cggcgagttc	1140
ttctactgca acaccaccaa gctgttcaac aacacctgca tcggcaacga gaccatggag	1200
ggctgcaacg gcaccatcat cctgcctgc aagatcaagc agatcatcaa catgtggcag	1260
ggcgccggcc aggccatgta cgccccccc atctccggcc gcatcaactg cgtgtccaac	1320
atcacggca tcctcgtgac ccgcgacggc ggcgccaaca acaccaacga gaccttcgc	1380
cccgcgggcg gcaacatcaa ggacaactgg cgctccgagc tgtacaagta caagtggtg	1440
cagatcgagc cctgggcat cgccccacc cgcgccaagc gccgctggt ggagcgcgag	1500
aagcgcgccg tgggcatcgg cgccatgatc ttcggttcc tgggcgccgc cggctccacc	1560
atggcgccg cctccatcac cctgaccgtg caggcccgcc agctgctgtc cggcatcgtg	1620
cagcagcagt ccaacctgct gcgcgccatc gaggccagc agcacctgct gcagctgacc	1680
gtgtggggca tcaagcagt gcaggccgc gtgctggccg tggagcgcta cctgaaggac	1740
cagaagtcc tgggcctgtg gggctgctcc ggcaagatca tctgcaccac cgccgtgcc	1800
tggaactcca cctgggtcaa ccgtccttc gaggagatct ggaacaacat gacctggatc	1860
gagtgggagc gcgagatct ccaactacac aaccagatct acgagatcct gaccgagtc	1920
cagaaccagc aggaccgcaa cgagaaggac ctgctggagc tggacaagt ggcctccctg	1980
tggaactggt tcgacatcac caactggctg tggatcatca agatcttcat catgatcgtg	2040
ggcggcctga tcggcctgcg catcatcttc gccgtgctgt ccatcgtgaa ccgcgtgcgc	2100
cagggtact cccccgtc ctccagacc cccaccacc accagcgca gcccgaccgc	2160
cccgagcgca tcgaggagg cggcggcgag caggcccgcc accgtccgt gcgcctggtg	2220

tccggcttcc tggccctggc ctgggacgac ctgcgctccc tgtgcctgtt ctctaccac	2280
cgctgcgcg acttcatcct gatcgccgcc cgcaccgtgg agctgctggg ccactcctcc	2340
ctgaagggcc tgcgcgcgg ctgggagggc ctgaagtacc tgggcaacct gctgctgtac	2400
tggggccagg agctgaagat ctccgccatc tcctgtctgg acgccaccgc catcgccgtg	2460
gccggctgga ccgaccgct gatcgagggtg gccagggcg cctggcgcgc catcctgcac	2520
atccccgcc gcaccccca gggcctggag cgcgccctgc tgtaa	2565
<210> 96	
<211> 2535	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 96	
atgcgcgtga tgggcatcca gaagaactac cccctgctgt ggcgctgggg catgatcatc	60
ttctggatca tgatcatctg caacgccgag aacctgtggg tgaccgtgta ctacggcgtg	120
cccgtgtggc gcgagccga gaccacctg ttctgcgcct ccgacgcaa ggcctacgac	180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag	240
gagatccacc tggagaacgt gaccgagaac ttcaacatgt ggaagaaca catggtggag	300
cagatgcacg aggacatcat ctccctgtgg gaccagtccc tgaagccctg cgtgaagctg	360
acccccctgt gcgtgacct ggactgccac aacaacatca ccaactcaa caccaccaac	420
aacaacgcc gcgagatcaa gaactgtcc ttcaacatga ccaccgagct gcgcgacaag	480
aagcagaagg tgiacgcct gttctaccgc ctggacgtgg tgcagatcaa caagaacaac	540
tcccagtacc gcctgatcaa ctgcaacacc tccgccatca ccagggcctg ccccaagggtg	600
tccttcgagc ccaccccat ccactactgc gccccgccg gcttcgcat cctgaagtgc	660
aacgacaagg agttcaacgg caccggcccc tgcaagaac tgtccaccgt gcagtgcacc	720
cacggcatca agcccggtgt gtccaccag ctgctgtga acggctcct ggccgaggag	780
gagatcgtga tccgctccga gaacatcacc aacaaccca agaccatcat cgtgcagctg	840
gtgaagccc tgaagatcaa ctgcaccgc cccaacaaca acaccgcaa gtccgtgcgc	900
atcgccccg gccagacct ctacgccacc ggcgacatca tcggcgacat ccgccaggcc	960
cactgcaacg gtgccgcac caagtggaa aacacctgc agcaggtggc caccagctg	1020
cgcaagtact tcaacaagac catcatcttc gccaacccct ccggcggcga cctggagatc	1080
accacccact ccttcaactg cggcggcgag ttcttctact gcaacacctc cgagctgttc	1140
aactccacct ggaactccac ctggaacaac accgagaagt gcatcaccct gcagtgccgc	1200

atcaagcaga tcgtgaacat gtggcagaag gtgggccagg ccatgtacgc ccccccatc 1260
 cagggcgtga tccgctgcga gtccaacatc accggcctgc tgctgacctg cgacggcggc 1320
 aacaacaact ccaccaacga gaccttcgcg cccggcggcg gcgacatgcg cgacaactgg 1380
 cgctccgagc tgiacaagta caaggtgggtg aagatcgagc ccctgggcgt ggccccacc 1440

cgcgccaagc gccgcgtggt ggagcgcgag aagcgcgccg tgggcctggg cgccgtgttc 1500
 ctgggcttcc tgggcgcgcg cggtccacc atgggcgccg cctccatcac cctgacctg 1560
 caggcccccc agctgctgtc cggcatcgtg cagcagcagt ccaacctgct gcgcgccatc 1620
 gaggccccagc agcacctgct gaagctgacc gtgtggggca tcaagcagct gcaggccccg 1680
 gtgtggccc tggagcgcta cctgaaggac cagcagctgc tgggcatctg gggctgtccc 1740
 ggcaagctga tctgcaccac caccgtgccg tggaaactct cctgggccaa caagacctac 1800
 aacgacatct gggacaacat gacctggctg cagtgggaca aggagatctc caactacacc 1860

gacatcatct acaacctgat cgaggagtcc cagaaccagc aggagaagaa cgagcaggac 1920
 ctgttggccc tggacaagtg ggctccctg tggaaactgt tcgacatcac caactggctg 1980
 tggatcatca agatcttcat catgatcgtg ggcgccctga tcggcctgcg catcgtgttc 2040
 gccgtgctga ccatcatcaa ccgctgctgc cagggtact cccccctgtc cttccagacc 2100
 ctgacccacc accagcgcga gcccaccgc cccgagcgca tcgaggaggg cggcggcgag 2160
 caggaccgag accgctccgt gcgcctgggtg tccggcttcc tggccctggc ctgggacgac 2220
 ctgcgctccc tgtccctgtt ctctaccac cgctgctgcg acttcgtgct gatcgccgcc 2280

cgcacctgg agctgctggg ccaactctcc ctgaagggcc tgcgcctggg ctgggaggcc 2340
 ctgaagtacc tgggcaacct gctgtcctac tggggccagg agctgaagaa ctccgccatc 2400
 aacctgctgg acaccatcgc catcgccgtg gccaaactgga ccgaccgct gatcgagatc 2460
 ggccagcgcg ccggcgcgcg catctgaac atccccgcc gcatccgcca gggcctggag 2520
 cgcgccctgc tgtaa 2535

<210> 97

<211> 840

<212> PRT

<213> Human immunodeficiency virus

<400> 97

Met Arg Val Lys Glu Ile Arg Lys His Leu Trp Arg Trp Gly Thr Leu

1 5 10 15

Phe Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu Asn Leu Trp Val

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

 Pro Leu Glu Asn Val Thr Glu Asn Phe Asn Met Gly Lys Asn Asn Met
 85 90 95
 Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
 100 105 110
 Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr
 115 120 125
 Asp Leu Lys Lys Asn Val Thr Ser Thr Asn Thr Ser Ser Ile Lys Met
 130 135 140
 Met Glu Met Lys Asn Cys Ser Phe Asn Ile Thr Thr Asp Leu Arg Asp

 145 150 155 160
 Lys Val Lys Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Gln
 165 170 175
 Ile Asp Asn Asp Ser Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Val
 180 185 190
 Thr Gln Ala Cys Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr
 195 200 205
 Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe
 210 215 220

 Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His
 225 230 235 240
 Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu
 245 250 255
 Ala Glu Glu Glu Val Val Ile Arg Ser Val Asn Phe Thr Asp Asn Thr
 260 265 270
 Lys Thr Ile Ile Val Gln Leu Lys Glu Pro Val Glu Ile Asn Cys Thr

275 280 285
Arg Pro Asn Asn Asn Thr Arg Lys Gly Ile His Ile Gly Pro Gly Arg

290 295 300
Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His
305 310 315 320

Cys Asn Ile Ser Ile Thr Lys Trp Asn Asn Thr Leu Lys Gln Ile Val
325 330 335

Ile Lys Leu Arg Lys Gln Phe Gly Asn Lys Thr Ile Val Phe Asn Gln
340 345 350

Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly
355 360 365

Gly Glu Phe Phe Tyr Cys Asn Thr Thr Lys Leu Phe Asn Ser Thr Trp
370 375 380

Asn Gly Thr Glu Glu Leu Asn Asn Thr Glu Gly Asp Ile Val Thr Leu
385 390 395 400

Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys
405 410 415

Ala Met Tyr Ala Pro Pro Ile Ala Gly Gln Ile Arg Cys Ser Ser Asn
420 425 430

Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Gln Ser Asn Val

435 440 445
Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg

450 455 460
Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val

465 470 475 480
Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala

485 490 495
Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser

500 505 510
Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu

515 520 525

Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu
 530 535 540
 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
 545 550 555 560
 Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu
 565 570 575
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val
 580 585 590
 Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Leu Asp Glu Ile Trp Asn
 595 600 605
 Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asn Asn Tyr Thr Gly
 610 615 620
 Leu Ile Tyr Asn Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
 625 630 635 640
 Glu Gln Glu Ile Leu Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
 645 650 655
 Phe Asp Ile Ser Lys Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile
 660 665 670
 Val Gly Gly Leu Val Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile
 675 680 685
 Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Arg
 690 695 700
 Leu Pro Thr Gln Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu
 705 710 715 720
 Gly Gly Glu Arg Asp Arg Asp Thr Ser Ile Arg Leu Val Asn Gly Phe
 725 730 735
 Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ile Tyr
 740 745 750
 His His Leu Arg Asp Leu Leu Leu Ile Ala Ala Arg Ile Val Glu Leu
 755 760 765
 Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu
 770 775 780

Gln Tyr Trp Ile Gln Glu Leu Lys Ser Ser Ala Ile Asn Leu Ile Asp
785 790 795 800

Thr Ile Ala Ile Ala Val Ala Gly Trp Thr Asp Arg Val Ile Glu Ile
805 810 815

Gly Gln Arg Phe Cys Arg Ala Ile Arg Asn Ile Pro Arg Arg Ile Arg
820 825 830

Gln Gly Ala Glu Lys Ala Leu Gln
835 840

<210> 98

<211> 848

<212> PRT

<213> Human immunodeficiency virus

<400> 98

Met Arg Val Met Gly Ile Gln Arg Asn Tyr Pro His Leu Trp Glu Trp
1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Lys Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Glu
35 40 45

Thr Thr Pro Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Lys Glu Val
50 55 60

His Asn Ile Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile Ala Leu Lys Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Glu Gly Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ala Leu
115 120 125

Asn Cys Ser Asn Ala Thr Ile Asn Asn Ser Thr Lys Thr Asn Ser Thr
130 135 140

Glu Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp

145 150 155 160
 Lys Lys Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Ile Val Pro
 165 170 175
 Ile Asn Asp Ser Ala Asn Asn Asn Ser Ile Asn Ser Glu Tyr Met Leu
 180 185 190
 Ile Asn Cys Asn Ala Ser Thr Ile Lys Gln Ala Cys Pro Lys Val Thr
 195 200 205
 Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile
 210 215 220
 Leu Lys Cys Asn Asp Lys Asn Phe Thr Gly Leu Gly Pro Cys Thr Asn

 225 230 235 240
 Val Ser Ser Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr
 245 250 255
 Gln Leu Leu Leu Asn Gly Ser Leu Ala Thr Glu Gly Val Val Ile Arg
 260 265 270
 Ser Lys Asn Phe Thr Asp Asn Thr Lys Asn Ile Ile Val Gln Leu Ala
 275 280 285
 Lys Ala Val Lys Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys
 290 295 300

 Ser Val His Ile Gly Pro Gly Gln Thr Trp Tyr Ala Thr Gly Glu Ile
 305 310 315 320
 Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Asn Asp Trp
 325 330 335
 Asn Glu Thr Leu Gln Lys Ile Val Glu Glu Leu Arg Lys His Phe Pro
 340 345 350
 Asn Lys Thr Ile Ile Phe Ala Pro Ser Ala Gly Gly Asp Leu Glu Ile
 355 360 365
 Thr Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr

 370 375 380
 Ser Glu Leu Phe Asn Ser Thr Tyr Met Asn Ser Thr Asn Ser Thr Thr
 385 390 395 400

Ile Asn Lys Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Val Ser
405 410 415

Met Trp Gln Glu Val Gly Gln Ala Met Tyr Ala Pro Pro Ile Ala Gly
420 425 430

Ser Ile Asn Cys Ser Ser Asp Ile Thr Gly Ile Ile Leu Thr Arg Asp
435 440 445

Gly Gly Asn Asn Asn Thr Asn Asn Glu Thr Phe Arg Pro Gly Gly Gly
450 455 460

Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
465 470 475 480

Lys Ile Glu Pro Val Gly Val Ala Pro Thr Arg Ala Arg Arg Arg Val
485 490 495

Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly
500 505 510

Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu
515 520 525

Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser
530 535 540

Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Arg Leu Thr
545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Leu Glu Ser
565 570 575

Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys
580 585 590

Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys
595 600 605

Ser Tyr Asn Asp Ile Trp Asp Asn Met Thr Trp Leu Gln Trp Asp Lys
610 615 620

Glu Ile Asn Asn Tyr Thr Gln Ile Ile Tyr Glu Leu Leu Glu Glu Ser
625 630 635 640

Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp Lys
645 650 655

Trp Ala Asn Leu Trp Asn Trp Phe Asn Ile Ser Asn Trp Leu Trp Tyr

660 665 670
 Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile
 675 680 685
 Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser
 690 695 700
 Pro Leu Ser Leu Gln Thr Leu Ile Pro Thr Thr Gln Arg Gly Pro Asp
 705 710 715 720
 Arg Pro Glu Gly Thr Glu Glu Glu Gly Gly Glu Gln Asp Arg Ser Arg
 725 730 735

Ser Ile Arg Leu Val Asn Gly Phe Leu Pro Leu Ile Trp Asp Asp Leu
 740 745 750
 Arg Asn Leu Cys Leu Phe Ser Tyr Arg His Leu Arg Asn Leu Leu Leu
 755 760 765
 Ile Val Ala Arg Thr Val Glu Leu Leu Gly Ile Arg Gly Trp Glu Ala
 770 775 780
 Leu Lys Tyr Leu Trp Asn Leu Leu Leu Tyr Trp Gly Gln Glu Leu Arg
 785 790 795 800
 Asn Ser Ala Ile Asn Leu Leu Asp Thr Thr Ala Ile Ala Val Ala Glu

805 810 815
 Gly Thr Asp Arg Ile Ile Glu Ala Val Gln Arg Ala Cys Arg Ala Ile
 820 825 830
 Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala Leu Leu
 835 840 845

<210> 99
 <211> 2523
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 99

atgcgcgtga aggagatccg caagcacctg tggcgcctggg gcaccctgtt cctgggcatg 60
 ctgatgatct gctccgccac cgagaacctg tgggtgaccg tgtactacgg cgtgcccgtg 120

tggaaggagg ccaccaccac cctgttctgc gcctccgacg ccaaggccta ctccaaggag	180
gtgcacaacg tgtgggccac ctacgcctgc gtgcccaccg accctcccc ccaggagatc	240
cccttgaga acgtgaccga gaacttcaac atgggcaaga acaacatggt ggagcagatg	300
cacgaggaca tcatctcct gtgggaccag tcctgaagc cctgctgaa gctgaccccc	360
ctgtgctga cctgaactg caccgacctg aagaagaacg tgacctcac caacacctcc	420
tccatcaaga tgatggagat gaagaactgc tccttcaaca tcaccaccga cctgctcgac	480
aaggtgaaga aggagtacgc cctgttctac aagctggacg tggcgcagat cgacaacgac	540
tcctaccgcc tgatctcctg caacacctcc gtggtgaccc aggcctgccc caagatctcc	600
ttcgagccca tccccatcca ctactgcgc cccgccggct tcgcatcct gaagtgcac	660
gacaagaagt tcaacggcac cggccctgc accaacgtgt ccaccgtgca gtgcaccac	720
ggcatcaagc ccgtggtgc caccagctg ctgctgaacg gctccctggc cgaggaggag	780
gtggtgatcc gctccgtgaa cttaccgac aacaccaaga ccatcatcgt gcagctgaag	840
gagcccgtag agatcaactg caccgcccc aacaacaaca cccgaaggg catccacatc	900
ggccccggcc gcgccttcta cgccaccggc gacatcatcg gcgacatccg ccaggccac	960
tgcaacatct ccatcaccaa gtggaacaac accctgaagc agatcgtgat caagctgcgc	1020
aagcagttcg gcaacaagac catcgtgttc aaccagtcct ccggcggcga ccccgagatc	1080
gtgatgact cttcaactg cggcggcgag ttcttctact gcaacaccac caagctgttc	1140
aactccact ggaacggcac cgaggagctg aacaacaccg agggcgacat cgtgaccctg	1200
ccctgccga tcaagcagat catcaacatg tggcaggagg tgggcaaggc catgtacgcc	1260
cccccatcg ccggccagat ccgtgctcc tccaacatca ccggcctgct gctgaccgc	1320
gacggcggca accagtcaa cgtgaccgag atcttccgcc ccggcggcgg cgacatgcgc	1380
gacaactggc gctccgagct gtacaagtac aaggtggtga agatcgagcc cctgggcgtg	1440
gccccacca aggccaagcg ccgctggtg cagcgcgaga agcgcgccgt ggcatcggc	1500
gccgtgttc tgggttcct gggcgccgc ggctccacca tggcgccgc ctccatcacc	1560
ctgaccgtgc agggccgcca gctgctgtcc ggcatcgtgc agcagcagaa caacctgctg	1620
cgcgccatcg agggccagca gcacctgtg cagctgaccg tgtggggcat caagcagctg	1680
caggccccgc tctggtccgt ggagcgctac ctgaaggacc agcagctgct gggcatctgg	1740
ggctgtccg gcaagctgat ctgcaccacc gccgtgccct ggaacacctc ctggtccaac	1800
aagtccttg acgagatctg gaacaacatg acctggatgg agtgggagcg cgagatcaac	1860
aactacaccg gcctgatcta caacctgac gaggagtccc agaaccagca ggagaagaac	1920
gagcaggaga tcctggccct ggacaagtgg gcctccctgt ggaactggtt cgacatctcc	1980

aagtggctgt ggiacatcaa gatcttcac atgacgtgg gcggcctggt gggcctgcgc	2040
atcatcttcg ccgtgctgtc catcgtgaac cgcgtgcgcc agggctactc cccctgtcc	2100
ttccagacc gcctgccac ccagcgcgc cccgaccgcc ccgagggcat cgaggaggag	2160
ggcggcgagc gcgaccgca cacctccac cgcctggtga acggcttcct ggccctgac	2220
tgggacgacc tgcgtccct gtgcctgttc atctaccacc acctgcgca cctgctgtg	2280
atgcgcgcc gcacgtgga gctgctggc cgcgcggct gggaggccct gaagtactgg	2340
tggaacctgc tgcagtactg gatccaggag ctgaagtcct ccgccatcaa cctgatcgac	2400
accatcgcca tgcctgtgc cggctggacc gaccgcgtga tcgagatcg ccagcgttc	2460
tgccgcgcca tccgaacat ccccgccgc atccgccagg gcgcgagaa ggccctgcag	2520
taa	2523
<210> 100	
<211> 2547	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 100	
atgcgcgtga tgggcatcca gcgcaactac cccacctgt gggagtggg caccctgac	60
ctgggcctgg tgatcatctg ctccgcctcc aagaacctgt gggtagcgt gtactacggc	120
gtgccctgt ggcgcgagc cgagaccacc ccttctcg cctccgacgc caaggcctac	180
gacaaggagg tgcacaacat ctgggccacc cagcctgcg tgcccaccga ccccaaccc	240
caggagatcg cctgaagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300
gagcagatgc acgaggacat catctccctg tgggacgagg gcctgaagcc ctgcgtgaag	360
ctgaccccc tgtgcgtggc cctgaactgc tccaacgcca ccatcaaca ctccaccaag	420
accaactcca ccgaggagat caagaactgc tccttcaaca tcaccaccga gatccgcgac	480
aagaagaaga aggagtacgc cctgttctac cgcttgaca tctgtcccat caacgactcc	540
gccaacaaca atccatcaa ctccgagtac atgctgatca actgcaacgc ctccaccac	600
aagcaggcct gcccgaagg gacctcgag cccatccca tccactactg cgcgccgcc	660
ggcttcgcca tctgaagtg caacgacaag aacttcaccg gcctgggccc ctgcaccaac	720
gtgtctccg tgcagtgcac ccacggcac aagccgtgg tgtccacca gctgctgctg	780
aacggctccc tggccaccga gggcgtggtg atccgtcca agaacttac cgacaacacc	840
aagaacatca tctgacgtt ggccaaggcc gtgaagatca actgcaccg cccaacaac	900
aacaccgca agtccgtgca catcgcccc ggccagacct ggtacggcac cggcgagatc	960

atcggcgaca tccgccaggc ccactgcaac atctccggca acgactggaa cgagaccctg	1020
cagaagatcg tggaggagct gcgcaagcac ttccccaaca agaccatcat cttcgcccc	1080
tccgccggcg gcgacctgga gatcaccacc cactccttca actgcggcgg cgagtcttc	1140
tactgcaaca cctccgagct gttcaactcc acctacatga actccaccaa ctccaccacc	1200
atcaacaaga ccataccct gccctgccgc atcaagcaga tcgtgtccat gtggcaggag	1260
gtgggccagg ccattgacgc ccccccatc gccggctcca tcaactgtc ctccgacatc	1320
accggcatca tctgacccg cgacggcggc aacaacaaca ccaacaacga gaccttcgc	1380
cccggcggcg gcgacatcg cgacaactgg cgtccgagc tgtacaagta caagtggtg	1440
aagatcgagc ccgtgggcgt ggccccacc cgcgccgcc gccgcgtgt gcagcgcgag	1500
aagcgcgcg tgggcatcgg cgccgtgtt ctgggttcc tgggcgccgc cggctccacc	1560
atgggcgcg cctccatcac cctgaccgtg caggccccgc agctgtgtc cggcatcgtg	1620
cagcagcagt ccaacctgct gcgcgccatc gaggcccagc agcacctgct gcgcctgacc	1680
gtgtggggca tcaagcagct gcaggccgc gtgttgccc tggagtcta cctgaaggac	1740
cagcagctgc tgggcatctg gggtgtctc ggcaagctga tctgcaccac caacgtgcc	1800
tggaactcct cctggcca caagtctac aacgacatct gggacaacat gacctggctg	1860
cagtgggaca aggagatcaa caactacacc cagatcatct acgagctgct ggaggagtcc	1920
cagaaccagc aggagaagaa cgagcaggac ctgtggccc tggacaagt ggccaacctg	1980
tggaactggt tcaacatctc caactggctg tggatcatca agatcttcat catgatcgtg	2040
ggcgccctga tcggcctgc catcatctt gccgtgtgt ccatcgtgaa ccgcgtgcgc	2100
cagggtact cccccgtc cctgcagacc ctgatccca ccaccagcg cggccccgac	2160
cggcccgagg gcaccgagga ggaggcggc gaggcagacc gctccgctc catccgctg	2220
gtgaacggct tctgcccc gatctgggac gacctgcga acctgtgcct gttctctac	2280
cgccacctgc gcaacctgct gctgatcgt gcccgaccg tggagctgct ggcatccgc	2340
ggctgggagg cctgaagta cctgtggaac ctgtgtgt actggggcca ggagctgcgc	2400
aactcgcca tcaacctgct ggacaccacc gccatgccg tggccgaggg caccgaccgc	2460
atcatcgagg ccgtgcagcg gcctgccgc gccatccga acatccccg ccgatccgc	2520
cagggcctgg agcgcgcct gctgtaa	2547
<210> 101	
<211> 832	
<212> PRT	
<213> Human immunodeficiency virus	

<400> 101

Met Arg Val Lys Gly Ile Gln Lys Asn Trp Gln His Leu Trp Lys Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asn Asn

20 25 30

Met Trp Val Thr Val Tyr Tyr Gly Val Pro Ala Trp Glu Asp Ala Asp

35 40 45

Thr Ile Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Ser Ala Glu Lys

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Ala Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn His Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Glu Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asn Val Thr Lys Asn Asn Asn Thr Lys Ile Met Gly Arg

130 135 140

Glu Glu Ile Lys Asn Cys Ser Phe Asn Val Thr Thr Glu Ile Arg Asp

145 150 155 160

Lys Lys Lys Lys Glu Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro

165 170 175

Ile Asp Asp Asn Asn Asn Ser Tyr Arg Leu Ile Asn Cys Asn Ala Ser

180 185 190

Thr Ile Lys Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile

195 200 205

His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Arg Asp Lys

210 215 220

Asn Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys

225 230 235 240

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly

245										250										255																			
Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Ile	Lys	Ser	Glu	Asn	Leu	Thr	Asp																							
260										265										270																			
Asn	Thr	Lys	Thr	Ile	Ile	Val	Gln	Leu	Asn	Lys	Ser	Val	Glu	Ile	Arg																								
275										280										285																			
Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Ser	Phe	Gly	Pro																								
290										295										300																			
Gly	Gln	Ala	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln																								
305										310										315										320									
Ala	His	Cys	Asn	Val	Ser	Arg	Thr	Asp	Trp	Asn	Asn	Met	Leu	Gln	Asn																								
325										330										335																			
Val	Thr	Ala	Lys	Leu	Lys	Glu	Leu	Phe	Asn	Lys	Asn	Ile	Thr	Phe	Asn																								
340										345										350																			
Ser	Ser	Ala	Gly	Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys																								
355										360										365																			
Gly	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Gln	Leu	Phe	Asn	Ser	Thr																								
370										375										380																			
Arg	Pro	Asn	Glu	Thr	Asn	Thr	Ile	Thr	Leu	Pro	Cys	Lys	Ile	Lys	Gln																								
385										390										395										400									
Ile	Val	Arg	Met	Trp	Gln	Arg	Val	Gly	Gln	Ala	Met	Tyr	Ala	Pro	Pro																								
405										410										415																			
Ile	Ala	Gly	Asn	Ile	Thr	Cys	Thr	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu																								
420										425										430																			
Thr	Arg	Asp	Gly	Asn	Asn	Asn	Asp	Ser	Glu	Thr	Phe	Arg	Pro	Gly	Gly																								
435										440										445																			
Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val																								
450										455										460																			
Val	Lys	Ile	Lys	Pro	Leu	Gly	Ile	Ala	Pro	Thr	Arg	Ala	Arg	Arg	Arg																								
465										470										475										480									
Val	Val	Gly	Arg	Glu	Lys	Arg	Ala	Val	Gly	Leu	Gly	Ala	Val	Phe	Leu																								
485										490										495																			

Gly Phe Leu Gly Thr Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr
500 505 510

Leu Thr Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln
515 520 525

Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu
530 535 540

Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu
545 550 555 560

Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly
565 570 575

Lys Leu Ile Cys Pro Thr Asn Val Pro Trp Asn Ala Ser Trp Ser Asn
580 585 590

Lys Thr Tyr Asn Glu Ile Trp Asp Asn Met Thr Trp Ile Glu Trp Asp
595 600 605

Arg Glu Ile Asn Asn Tyr Thr Gln Gln Ile Tyr Ser Leu Ile Glu Glu
610 615 620

Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Asp Leu Leu Ala Leu Asp
625 630 635 640

Lys Trp Ala Ser Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp
645 650 655

Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg
660 665 670

Ile Val Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr
675 680 685

Ser Pro Leu Ser Leu Gln Thr Leu Ile Pro Asn Pro Thr Gly Ala Asp
690 695 700

Arg Pro Gly Glu Ile Glu Glu Gly Gly Gly Glu Gln Gly Arg Thr Arg
705 710 715 720

Ser Ile Arg Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu
725 730 735

Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Val Leu
740 745 750

Ile Ala Ala Arg Thr Val Glu Thr Leu Gly His Arg Gly Trp Glu Ile

755	760	765	
Leu Lys Tyr Leu Gly Asn Leu Val Cys Tyr Trp Gly Gln Glu Leu Lys			
770	775	780	
Asn Ser Ala Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val Ala Asn			
785	790	795	800
Trp Thr Asp Arg Val Ile Glu Val Val Gln Arg Val Phe Arg Ala Phe			
805	810	815	
Leu Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu			
820	825	830	

<210> 102
 <211> 858
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 102

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

50	55	60	
His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro			
65	70	75	80
Gln Glu Ile Val Met Glu Asn Val Thr Glu Asn Phe Asn Met Trp Asn			
85	90	95	
Asn Asp Met Val Asn Gln Met His Glu Asp Val Ile Ser Leu Trp Asp			
100	105	110	
Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu			
115	120	125	

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

Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly Thr Tyr Met Asn Gly
 385 390 395 400
 Thr Asn Asn Ser Ser Ser Ile Ile Thr Ile Pro Cys Arg Ile Lys Gln
 405 410 415

 Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro Pro
 420 425 430
 Ile Glu Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu
 435 440 445
 Val Arg Asp Gly Gly Arg Thr Glu Ser Asn Asn Thr Glu Ile Phe Arg
 450 455 460
 Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Asn Glu Leu Tyr Lys
 465 470 475 480
 Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Ala Ala

 485 490 495
 Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Leu Gly Ala
 500 505 510
 Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala
 515 520 525
 Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val
 530 535 540
 Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met
 545 550 555 560

 Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu
 565 570 575
 Ala Ile Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly
 580 585 590
 Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser
 595 600 605
 Trp Ser Asn Lys Ser Gln Gln Glu Ile Trp Asp Asn Met Thr Trp Met
 610 615 620
 Gln Trp Asp Lys Glu Ile Ser Asn Tyr Thr Asn Thr Ile Tyr Arg Leu

625 630 635 640
 Leu Glu Asp Ser Gln Asn Gln Gln Glu Arg Asn Glu Lys Asp Leu Leu
 645 650 655
 Ala Leu Asp Ser Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Thr Asn
 660 665 670
 Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile
 675 680 685
 Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg
 690 695 700

 Gln Gly Tyr Ser Pro Leu Ser Phe Gln Ile Leu Thr Pro Asn Pro Gly
 705 710 715 720
 Gly Pro Gly Arg Leu Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp
 725 730 735
 Lys Thr Arg Ser Ile Arg Leu Val Asn Gly Phe Leu Ala Leu Ala Trp
 740 745 750
 Asp Asp Leu Arg Asn Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp
 755 760 765
 Phe Ile Leu Leu Thr Ala Arg Gly Val Glu Leu Leu Gly Arg Asn Ser

 770 775 780
 Leu Arg Gly Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser
 785 790 795 800
 Leu Val Gln Tyr Trp Gly Leu Glu Leu Lys Lys Ser Thr Ile Ser Leu
 805 810 815
 Val Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile
 820 825 830
 Asn Ile Val Gln Gly Ile Cys Arg Ala Ile His Asn Ile Pro Arg Arg
 835 840 845

 Ile Arg Gln Gly Phe Glu Ala Ala Leu Gln
 850 855
 <210> 103
 <211> 2499
 <212> DNA

<213> Human immunodeficiency virus

<400> 103

atgcgcgtga agggcatcca gaagaactgg cagcacctgt ggaagtgggg caccctgac	60
ctgggccttg tgatcatctg ctccgcctcc aacaacatgt gggtgaccgt gtactacggc	120
gtgccgcctt gggaggacgc cgacaccatc ctgtttctgcg cctccgacgc caagcctac	180
tccgcccaga agcacaacgt gtgggccacc cagcctgcg tgcccaccga ccccaacccc	240
caggagatcg ccctggagaa cgtgaccgag aacttcaaca tgtggaagaa ccacatggtg	300
gagcagatgc acgaggacat catctccctg tgggacgagt ccctgaagcc ctgcgtgaag	360
ctgaccccc tggtcgtgac cctgaactgc accaacgtga ccaagaacaa caacaccaag	420
atcatgggcc gcgaggagat caagaactgc tccttcaacg tgaccaccga gatccgcgac	480
aagaagaaga aggagtacgc cctgtttctac cgcttgacg ttgtgcccac cgacgacaac	540
aacaactcct accgcctgat caactgcaac gcctccacca tcaagcaggc ctgcccacag	600
gtgtccttcg agcccatccc catccactac tgcgcccccg ccggcttcgc catcctgaag	660
tgcgcgcgaca agaacttcaa cggcaccggc ccctgcaaga acgtgtccac cgtgcagtgc	720
accacggca tcaagccctg ggtgtccacc cagctgtctg tgaacggctc cctggccgag	780
gaggagatca tcatcaagtc cgagaacctg accgacaaca ccaagaccat catcgtgcag	840
ctgaacaagt ccgtggagat ccgtctgcacc cgccccaaaca acaacaccgg caagtccatc	900
tccttcggcc ccggccaggc cttctacgcc accggcgaca tcatcggcga catccgccag	960
gcccactgca acgtgtcccg caccgactgg aacaacatgc tgcagaacgt gaccgccaag	1020
ctgaaggagc tgttcaacaa gaacatcacc ttcaactcct ccgcccggcg cgacctggag	1080
atcaccaccc actccttcaa ctgcggcggc gagtctctct actgcaacac ctcccagctg	1140
ttcaactcca cccgcccaca cgagaccaac accatcacc tgcctgcaa gatcaagcag	1200
atcgtgcgca tgtggcagcg cgtgggccag gccatgtacg cccccccat cgccggcaac	1260
atcacctgca cctccaacat caccggcctg ctgctgaccc gcgacggcaa caacaacgac	1320
tccgagacct tccgccccgg cggcggcgac atgcgcgaca actggcgctc cgagctgtac	1380
aagtacaagg tggatgaagat caagccctg ggcatcgccc ccaccgcgc ccgcccgcgc	1440
gtggtgggcc gcgagaagcg cgccgtgggc ctgggcgcg tgttcttggg cttcttgggc	1500
accgcccgtt ccacatggg cgccgcctcc atcacctga ccgtgcaggt gcgccagctg	1560
ctgtccggca tcgtgcagca gcagtccaac ctgctgcgcg ccatcgaggc ccagcagcac	1620
ctgctgcagc tgacctgtg gggcatcaag cagctgcagg cccgcgtgct ggccgtggag	1680
cgctacctga aggaccagca gctgtgggc atctggggct gctccggcaa gctgatctgc	1740

cccaccaacg tgccttgga cgcctcctgg tccaacaaga cctacaacga gatctgggac	1800
aacatgacct ggatcgagtg ggaccgcgag atcaacaact acaccagca gatctactcc	1860
ctgatcgagg agtcccagaa ccagcaggag aagaacgagc aggacctgct ggccctggac	1920
aagtgggcct ccctgtggtc ctggttcgac atctccaact ggctgtggta catcaagatc	1980
ttcatcatga tctggggcgg cctgatcggc ctgcgcatcg tgttcgccgt gctgtccatc	2040
gtgaaccgcg tgcgccaggg ctactcccc ctgtccctgc agacctgat ccccaacccc	2100
accggcgccg accgccccgg cgagatcgag gagggcggcg gcgagcaggg ccgcacccgc	2160
tccatccgcc tggatgaacg cttcctggcc ctggcctggg acgacctgcg ctccctgtgc	2220
ctgtttcctt accaccgctt gcgcgacttc gtgctgatcg ccgccgcac cgtggagacc	2280
ctgggccacc gcggctggga gatcctgaag tacctgggca acctggtgtg ctactggggc	2340
caggagctga agaactccgc catctccctg ctggacacca ccgccatcgc cgtggccaac	2400
tggaccgacc gcgtgatcga ggtggtgcag cgcgtgttcc gcgccttcct gaacatcccc	2460
cgccgcatcc gccagggtt cgagcgcgcc ctgctgtaa	2499
<210> 104	
<211> 2577	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 104	
atgcgcgtgc gcggcaccgc ccgcaactac cagcagtggt ggatctgggg cgtgctgggc	60
ttctggatgc tgatgatctg caacgtggag ggcaacctgt gggtgaccgt gtactacggc	120
gtgccctgtt ggaaggagcg caagaccacc ctgttctgcg cctccgacgc caaggcctac	180
gagaccgagg tgcacaacgt gtgggccacc cagcctgcg tgcccaccga ccccaacccc	240
caggagatcg tgatggagaa cgtgaccgag aacttcaaca tgtggaacaa cgacatggtg	300
aaccagatgc acgaggacgt gatctccctg tgggaccagt ccctgaagcc ctgcgtgaag	360
ctgaccccc tgtgcgtgac cctggagtgc accaactgtt cctccaacgg caacggcacc	420
tacaacgaga cctacaacga gtccgtgaag gagatcaaga actgctcctt caacgccacc	480
acctgctgc gcgaccgcaa gaagaccgtg tacgcctgtt tctaccgctt ggacatcgtg	540
ccctgaacg acgagaactc cggcaagaac tcctccgagt actaccgctt gatcaactgc	600
aacacctcgc ccatcaccca ggctgcccc aaggtgacct tcgaccccat ccccatccac	660
tactgcaccc ccgccggeta cgccatcctg aagtgcacg acaagaagtt caacggcacc	720
ggccagtgcc acaacgtgtc caccgtgcag tgcacccacg gcatcaagcc cgtggtgtcc	780

acccagctgc tgctgaacgg ctccctggcc gagcgcgaga tcatcatccg ctccgagaac 840
 ctgaccaaca acgtgaagac catcatcgtg cacctgaacc agtccgtgga gatcgtgtgc 900
 acccgcccca acaacaacac ccgcaagtcc atccgcatcg gccccggcca gaccttctac 960
 gccaccggcg acatcatcgg cgacatccgc caggcccaact gcaacatctc caaggacaag 1020
 tggtagaga cctgcagcg cgtgtccaag aagctggccg agcacttccc caacaagacc 1080

 atcaagttcg cctcctctc cggcggcgac ctggagatca ccaccactc cttcaactgc 1140
 cgcggcgagt tcttctactg caacacctcc ggctgttca acggcaccta catgaacggc 1200
 accaacaact cctcctccat catcaccatc cctgccgca tcaagcagat catcaacatg 1260
 tggcaggagg tggcgcgcg catgtacgcc cccccatcg agggcaacat cacctgcaag 1320
 tccaacatca ccggcctgct gctggtgcgc gacggcggcc gcaccgagtc caacaacacc 1380
 gagatcttcc gccccggcg ggcgacatg cgcaacaact ggcgcaacga gctgtacaag 1440
 tacaaggtag tggagatcaa gcccctgggc gtggcccca ccgccgcaa gcgccgctg 1500

 gtggagcgcg agaagcgcgc cgtgggcctg ggccgctgt tctgggctt cctggcgcc 1560
 gccgctcca ccatggcgcg cgctccatc accctgaccg tgcaggcccg ccagctgctg 1620
 tccggcatcg tgcagcaga gtccaacctg ctgcgcgcca tgcaggccca gcagcacatg 1680
 ctgcagctga ccgtgtgggg catcaagcag ctgcagacc gcgtgctggc catcgagcgc 1740
 tactgaagg accagcagct gctgggcac tggggctgct ccggcaagct gatctgcacc 1800
 accgcctgc cctggaactc ctctggtcc aacaagtccc agcaggagat ctgggacaac 1860
 atgacctgga tgcagtggga caaggagatc tccaactaca ccaacaccat ctaccgctg 1920

 ctggaggact ccagaacca gcaggagcgc aacgagaagg acctgctggc cctggactcc 1980
 tggagaacc tgtggtcctg gttcgacatc accaactggc tgtgtacat caagatctt 2040
 atcatgatcg tggcgccct gatcggcctg cgcatcatct tgcctgctgt gtccatcgtg 2100
 aaccgcgtgc gccagggcta ctccccctg tcttccaga tctgacccc caaccgcggc 2160
 ggccccggcc gcctgggccc catcgaggag gagggcggcg agcaggacaa gaccgcctcc 2220
 atccgcctgg tgaacggctt cctggccctg gcctgggacg acctgcgcaa cctgtgcctg 2280
 ttctcctacc accgcctgcg cgacttcac ctgctgaccg cccgcggcgt ggagctgctg 2340

 ggccgcaact cctgcgcgg cctgcagcgc ggtgggagg cctgaagta cctgggctcc 2400
 ctggtgcagt actgggcct ggagctgaag aagtccacca tctcctggt ggacaccatc 2460
 gccatgcgc tggccaggcg caccgaccgc atcatcaaca tctgcaggg catctgccgc 2520
 gccatcaca acatccccg ccgcatccgc cagggtctcg aggcgcct gcagtaa 2577

<210> 105

<211> 837

<212> PRT

<213> Human immunodeficiency virus

<400> 105

Met Arg Val Met Gly Ile Gln Arg Asn Cys Gln Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Met Ile Cys Asn Ala Thr Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Thr Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Lys Ala Glu Ala

50 55 60

His Asn Ile Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Gly Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Gly Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Ser Asp Val Asn Ala Thr Asn Ser Ala Thr Asn Thr Val Val

130 135 140

Ala Gly Met Lys Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp

145 150 155 160

Lys Lys Lys Gln Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Gln

165 170 175

Ile Asp Gly Ser Asn Thr Ser Tyr Arg Leu Ile Asn Cys Asn Thr Ser

180 185 190

Ala Ile Thr Gln Ala Cys Pro Lys Val Thr Phe Glu Pro Ile Pro Ile

195 200 205

His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys

210 215 220

Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys

225 230 235 240

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly

245 250 255

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp

260 265 270

Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val Thr Ile Asn

275 280 285

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro

290 295 300

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asn Ile Arg Gln

305 310 315 320

Ala Tyr Cys Asn Ile Ser Gly Thr Glu Trp Asn Lys Thr Leu Gln Gln

325 330 335

Val Ala Lys Lys Leu Gly Asp Leu Leu Asn Lys Thr Thr Ile Ile Phe

340 345 350

Lys Pro Ser Ser Gly Gly Asp Pro Glu Ile Thr Thr His Thr Phe Asn

355 360 365

Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys Leu Phe Asn Ser

370 375 380

Ser Trp Thr Ser Asn Asn Thr Gly Asn Thr Ser Thr Ile Thr Leu Pro

385 390 395 400

Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Gly Val Gly Lys Ala

405 410 415

Ile Tyr Ala Pro Pro Ile Ala Gly Leu Ile Asn Cys Ser Ser Asn Ile

420 425 430

Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ala Asn Asn Ser Glu Thr

435 440 445

Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu

450 455 460

Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Leu Ala Pro Thr

465 470 475 480
 Lys Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Ile Gly Leu
 485 490 495
 Gly Ala Val Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly
 500 505 510
 Ala Ala Ser Leu Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly

 515 520 525
 Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln
 530 535 540
 His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg
 545 550 555 560
 Val Leu Ala Val Glu Ser Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile
 565 570 575
 Trp Gly Cys Ser Gly Lys His Ile Cys Thr Thr Asn Val Pro Trp Asn
 580 585 590

 Ser Ser Trp Ser Asn Lys Ser Leu Glu Glu Ile Trp Asp Asn Met Thr
 595 600 605
 Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr Thr Gly Leu Ile Tyr
 610 615 620
 Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu
 625 630 635 640
 Leu Leu Gln Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Ser Ile
 645 650 655
 Thr Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly

 660 665 670
 Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Leu Val Asn Arg
 675 680 685
 Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Leu Pro Ala
 690 695 700
 Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu
 705 710 715 720
 Gln Gly Arg Gly Arg Ser Ile Arg Leu Val Asn Gly Phe Ser Ala Leu

725 730 735
 Ile Trp Asp Asp Leu Arg Asn Leu Cys Leu Phe Ser Tyr His Arg Leu
 740 745 750
 Arg Asp Leu Ile Leu Ile Ala Thr Arg Ile Val Glu Leu Leu Gly Arg
 755 760 765
 Arg Gly Trp Glu Ala Ile Lys Tyr Leu Trp Asn Leu Leu Gln Tyr Trp
 770 775 780
 Ile Gln Glu Leu Lys Asn Ser Ala Ile Ser Leu Leu Asp Thr Thr Ala
 785 790 795 800
 Ile Ala Val Ala Glu Gly Thr Asp Arg Ala Ile Glu Ile Val Gln Arg

 805 810 815
 Ala Val Arg Ala Val Leu Asn Ile Pro Thr Arg Ile Arg Gln Gly Leu
 820 825 830
 Glu Arg Ala Leu Leu
 835
 <210> 106
 <211> 839
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 106
 Met Arg Val Lys Glu Thr Gln Arg Asn Trp His Asn Leu Trp Arg Trp
 1 5 10 15
 Gly Leu Met Ile Phe Gly Met Leu Met Ile Cys Asn Ala Thr Glu Asn

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

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

Val Ala Lys Gln Leu Arg Glu Asn Phe Asn Lys Thr Ile Ile Phe Asn
340 345 350

Asn Pro Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys
355 360 365

Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Arg Leu Phe Asn Ser Thr
370 375 380

Trp Asn Asn Asp Thr Arg Asn Asp Thr Lys Gln Met His Ile Thr Leu
385 390 395 400

Pro Cys Arg Ile Lys Gln Ile Val Asn Met Trp Gln Arg Val Gly Gln
405 410 415

Ala Met Tyr Ala Pro Pro Ile Gln Gly Lys Ile Arg Cys Asn Ser Asn
420 425 430

Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Thr Asn
435 440 445

Glu Thr Phe Arg Pro Thr Gly Gly Asp Met Arg Asp Asn Trp Arg Ser
450 455 460

Glu Leu Tyr Lys Tyr Lys Val Val Glu Ile Lys Pro Leu Gly Val Ala
465 470 475 480

Pro Thr Arg Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val
485 490 495

Gly Ile Gly Ala Val Leu Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr
500 505 510

Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
515 520 525

Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Lys Ala Ile Glu Ala
530 535 540

Gln Gln His Leu Leu Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
545 550 555 560

Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu
565 570 575

Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro
580 585 590

Trp Asn Phe Ser Trp Ser Asn Lys Ser Tyr Asp Glu Ile Trp Asp Asn

595 600 605

Met Thr Trp Ile Glu Trp Glu Arg Glu Ile Asn Asn Tyr Thr Gln Thr

610 615 620

Ile Tyr Thr Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu

625 630 635 640

Gln Asp Leu Leu Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe

645 650 655

Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val

660 665 670

Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val

675 680 685

Asn Arg Cys Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Thr

690 695 700

Pro Asn His Lys Glu Ala Asp Arg Pro Gly Gly Ile Glu Glu Gly Gly

705 710 715 720

Gly Glu Gln Asp Arg Thr Arg Ser Ile Arg Leu Val Ser Gly Phe Leu

725 730 735

Ala Leu Ala Trp Asp Asp Leu Arg Asn Leu Cys Leu Phe Ser Tyr His

740 745 750

Arg Leu Arg Asp Phe Ile Leu Ile Ala Ala Arg Ile Val Glu Thr Leu

755 760 765

Gly Arg Arg Gly Trp Glu Ile Leu Lys Tyr Leu Gly Asn Leu Ala Gln

770 775 780

Tyr Trp Gly Gln Glu Leu Lys Asn Ser Ala Ile Ser Leu Leu Asn Ala

785 790 795 800

Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Val Val

805 810 815

His Arg Val Leu Arg Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln

820 825 830

Gly Phe Glu Arg Ala Leu Leu

835	
<210>	107
<211>	2514
<212>	DNA
<213>	Human immunodeficiency virus
<400>	107
atgcgcgtga tgggcatcca gcgcaactgc cagcagtggg ggatctgggg catcctgggc	60
ttctggatgc tgatgatctg caacgccacc ggcaacctgt gggtagacct gtactacggc	120
gtgcccgtgt ggaaggagac caccaccacc ctgttctgcg cctccgacgc caaggcctac	180
aaggccgagg ccacacaat ctgggccacc cagcctgcg tggccaccga ccccaacccc	240
caggagatcg tgcctggaga cgtgaccgag aacttcaaca tgtggaagaa cggcatggtg	300
gaccagatgc acgaggacat catctccctg tgggaccagg gcctgaagcc ctgcgtgaag	360
ctgaccccc tgtgcgtgac cctgaactgc tccgacgtga acgccacca ctccgccacc	420
aacacgtgg tggccggcat gaagaactgc tcttcaaca tcaccaccga gatccgcgac	480
aagaagaagc aggagtacgc cctgttctac aagctggacg tggtagcat cgacggctcc	540
aacacctct accgcctgat caactgcaac acctccgcca tcaccaggc ctgccccaa	600
gtgaccttcg agcccatccc catccactac tgcgccccg ccggtttcgc catcctgaag	660
tgcaacgaca agaagttaa cggcaccggc ccttgcaaga acgtgtccac cgtgcagtgc	720
accacggca tcaagccgt ggtgtccacc cagctgtgc tgaacggctc cctggccgag	780
gaggagatca tcatccgtc cgagaacctg accgacaacg ccaagacat catcgtgcag	840
ctgaacgagt ccgtgacct caactgcacc cgcccaaca acaacaccg caagtccatc	900
cgcatcggc ccggccagac cttctacgc accggcgaca tcatcggcaa catccgccag	960
gcctactgca acatctccg caccgagtgg aacaagacc tgcagcaggt ggccaagaag	1020
ctgggcgacc tgcgaacaa gaccaccatc atcttcaagc cctcctccg cggcgacccc	1080
gagatcacca ccacacctt caactgcggc ggcgagttct tctactgaa cacctccaag	1140
ctgttcaact cctcctggac ctccaacaac accggcaaca cctccaccat caccctgccc	1200
tgccgatca agcagatcat caacatgtgg caggcgctgg gcaaggccat ctacgcccc	1260
cccatcgccg gcctgatcaa ctgtcctcc aacatcccg gcctgctgt gacccgcgac	1320
ggcgcgcca acaactccga gacctccgc cccggcgcg gcgacatgc cgacaactgg	1380
cgctccgagc tgtacaagta caagtggtg aagatcgagc cctgggcct ggccccacc	1440
aaggccaagc gccgcgtggg ggagcgcgag aagcgcgcca tcggcctggg cgccgtgttc	1500

ctgggcttcc tgggcgcgcg cggctccacc atgggcgccc cctccctgac cctgaccgtg	1560
caggcccccc agctgctgtc cggcatcgtg cagcagcaga acaacctgct gcgcgccatc	1620
gaggccccagc agcacctgct gcagctgacc gtgtggggca tcaagcagct gcaggccccg	1680
gtgtggccg tggagtccca cctgaaggac cagcagctgc tgggcatctg gggctgctcc	1740
ggcaagcaca tctgcaccac caacgtgccc tggaactcct cctggtccaa caagtcctg	1800
gaggagatct gggacaacat gacctggatg gagtgggagc gcgagatcga caactacacc	1860
ggcctgatct actccctgat cgaggagtcc cagaaccagc aggagaagaa cgagcaggag	1920
ctgctgcagc tggacaagtg ggcctccctg tggaactggt tctccatcac caactggctg	1980
tggtacatca agatcttcat catgatcgtg ggcgccctga tcggcctgcg catcgtgttc	2040
gccgtgctgt ccttggtgaa ccgctgctgc cagggctact cccccctgtc cttccagacc	2100
ctgctgcccc cccccgcgg ccccgaccgc cccgagggca tcgaggagga gggcggcgag	2160
caggcgccgc gccgctccat ccgcttggtg aacggcttct ccgccctgat ctgggacgac	2220
ctgcgcaacc tgtgcctgtt ctctaccac cgctgcgcg acctgatcct gatcgccacc	2280
cgcatcgtgg agctgctggg ccgcgcggc tgggaggcca tcaagtacct gtggaacctg	2340
ctgcagtact ggatccagga gctgaagaac tccgcatct cctgctgga caccaccgcc	2400
atgccgtgg ccgagggcac cgaccgcgc atcgagatcg tgcagcgcgc cgtgcgcgc	2460
gtgtgaaca tccccaccg catccgccag ggcctggagc gcgccctgct gtaa	2514
<210> 108	
<211> 2520	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 108	
atgcgcgtga aggagaccga gcgcaactgg cacaacctgt ggcgtgagg cctgatgatc	60
ttcggcatgc tgatgatctg caacgccacc gagaacctgt gggtagacct gtactacggc	120
gtgccctgtt ggaaggacgc cgacaccacc ctgtttctgc cctccgacgc caaggcctac	180
tccaccgaga agcacaactg gtgggccacc cagcctgcg tgccaccga cccaacccc	240
caggagatcc ccttgagaaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg	300
gagcagatgc acgaggacat catctccctg tgggacgagt ccctgaagcc ctgcgtgaag	360
ctgaccccc tgtgcgtgac cctgaactgc accgacgtga agaacgccac caacaccacc	420
gtggaggccg ccgagatcaa gaactgctcc ttcaacatca ccaccgagat caaggacaag	480
aagaagaagg agtacgcctt gttctacaag ctggacgtgg tgcccatcaa cgacaacaac	540

aactccatct accgcctgat caactgcaac gtgtccaccg tgaagcaggc ctgccccaaag	600
gtgaccttcg agcccatccc catccactac tgcgcccccg ccggcttcgc catcctgaag	660
tgcaacgaca agaagttcaa cggcaccggc ccttgcaaga acgtgtccac cgtgcagtgc	720
accacaggca tcaagcccgt ggtgtccacc cagctgctgc tgaacggctc cctggccgag	780
ggcgagggtgc gcatccgctc cgagaacttc accaacaacg ccaagaccat catcgtgcag	840
ctgaactcct ccgtgcgcat caactgcacc cgcaccaaca acaacacccg caagtccatc	900
cacatcggcc ccggccaggc cttctacgcc accggcgaca tcatcggcga catccgccag	960
gcccactgca acatctcccg cgccgagtgg aacaacaccc tgcagcaggt ggccaagcag	1020
ctgcgcgaga acttcaacaa gaccatcatc ttcaacaacc cctccggcgg cgacctggag	1080
atcaccacc actccttcaa ctgcggcggc gaggttctct actgcaacac ctcccgcctg	1140
ttcaactcca cctggaacaa cgacaccgc aacgacacca agcagatgca catcacctg	1200
ccctgccga tcaagcagat cgtgaacatg tggcagcgcg tgggccaggc catgtacgcc	1260
cccccatcc agggcaagat ccgtgcaac tccaacatca ccggcctgct gctgaccgc	1320
gacggcggca acaacaacac caacgagacc ttccgcccc cggcgggcga catgcgcgac	1380
aactggcgct ccgagctgta caagtacaag gtggtggaga tcaagcccct gggcgtggcc	1440
cccccccg ccaagcgccg cgtggtggag cgcgagaagc gcgccgtggg catcggcgcc	1500
gtgctgctgg gcttcctggg cgccgcccgc tcacatgg gcgcgcctc catcacctg	1560
accgtgcagg cccgccagct gctgtccggc atcgtgcagc agcagtcaa cctgtgaag	1620
gccatcgagg ccagcagca cctgtgaag ctgaccgtgt ggggcatcaa gcagctgcag	1680
gcccgcgtgc tggccgtgga gcgtacctg aaggaccagc agctgctggg catctggggc	1740
tgctccggca agctgatctg caccaccaac gtgccctgga acttctctg gtccaacaag	1800
tcctacgacg agatctggga caacatgacc tggatcgagt gggagcgga gatcaacaac	1860
tacaccaga ccatctacac cctgtggag gaggccaga accagcagga gaagaacgag	1920
caggacctgc tggccctgga caagtgggc tcctgtgga actggttcga catctccaac	1980
tggtgtggt acatcaagat cttcatcatg atcgtggcg gcctgatcgg cctgcgcatc	2040
atcttcgccg tgctgtccat cgtgaaccgc tgccgccagg gctactccc cctgtccttc	2100
cagacctga ccccaacca caaggaggcc gaccgcccc gcggcatcga ggaggcggc	2160
ggcgagcagg accgcaccg ctccatccgc ctggtgtccg gcttcctggc cctggcctgg	2220
gacgacctgc gcaacctgtg cctgttctcc taccaccgcc tgcgcgactt catcctgate	2280
gcccgcgca tcgtggagac cctgggcccgc cgggctggg agatcctgaa gtacctgggc	2340
aacctggccc agtactgggg ccaggagctg aagaactccg ccatctccct gctgaacgcc	2400

accgccatcg ccgtggccga gggcaccgac cgcatcatcg aggtgggtgca cgcgtgctg 2460
 cgcgccatcc tgcacatccc ccgccgcacg cgccagggtc tcgagcgcgc cctgctgtaa 2520
 2520

<210> 109

<211> 838

<212> PRT

<213> Human immunodeficiency virus

<400> 109

Met Arg Val Arg Gly Met Gln Arg Asn Trp Gln His Leu Gly Lys Trp

1 5 10 15

Gly Leu Leu Phe Leu Gly Ile Leu Ile Ile Cys Asn Ala Thr Glu Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ser Tyr Glu Arg Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Val Asp Leu Glu Asn Val Thr Glu Asn Phe Asp Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Thr Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Ala Asn Ala Thr Ala Asn Ala Thr Lys Glu His Pro

130 135 140

Glu Gly Arg Ala Gly Ala Ile Gln Asn Cys Ser Phe Asn Met Thr Thr

145 150 155 160

Glu Val Arg Asp Lys Gln Met Lys Val Gln Ala Leu Phe Tyr Arg Leu

165 170 175

Asp Ile Val Pro Ile Ser Asp Asn Asn Ser Asn Glu Tyr Arg Leu Ile

180 185 190

Asn Cys Asn Thr Ser Thr Ile Thr Gln Ala Cys Pro Lys Val Ser Trp
 195 200 205

 Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu
 210 215 220
 Lys Cys Asn Asp Lys Lys Phe Asn Gly Thr Gly Pro Cys Lys Asn Val
 225 230 235 240
 Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln
 245 250 255
 Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser
 260 265 270
 Gln Asn Ile Ser Asp Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu

 275 280 285
 Ser Val Gln Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser
 290 295 300
 Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile
 305 310 315 320
 Gly Asp Ile Arg Lys Ala His Cys Asn Val Ser Gly Thr Gln Trp Asn
 325 330 335
 Lys Thr Leu Glu Gln Val Lys Lys Lys Leu Arg Ser Tyr Phe Asn Thr
 340 345 350

 Thr Ile Lys Phe Asn Ser Ser Ser Gly Gly Asp Pro Glu Ile Thr Met
 355 360 365
 His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Lys
 370 375 380
 Leu Phe Asn Asp Thr Val Ser Asn Asp Thr Ile Ile Leu Pro Cys Arg
 385 390 395 400
 Ile Lys Gln Ile Val Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr
 405 410 415
 Ala Ala Pro Ile Ala Gly Asn Ile Thr Cys Thr Ser Asn Ile Thr Gly

 420 425 430
 Leu Leu Leu Thr Arg Asp Gly Gly His Asn Glu Thr Asn Lys Thr Glu

435 440 445
 Thr Phe Arg Pro Gly Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu
 450 455 460
 Leu Tyr Lys Tyr Lys Val Val Glu Ile Glu Pro Leu Gly Val Ala Pro
 465 470 475 480
 Thr Arg Ala Lys Arg Gln Val Val Lys Arg Glu Lys Arg Ala Val Gly
 485 490 495

 Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met
 500 505 510
 Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser
 515 520 525
 Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln
 530 535 540
 Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala
 545 550 555 560
 Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly

 565 570 575
 Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp
 580 585 590
 Asn Ser Ser Trp Ser Asn Lys Ser Gln Glu Glu Ile Trp Glu Asn Met
 595 600 605
 Thr Trp Met Glu Trp Glu Lys Glu Ile Asn Asn Tyr Ser Asn Glu Ile
 610 615 620
 Tyr Arg Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln
 625 630 635 640

 Glu Leu Leu Ala Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp
 645 650 655
 Ile Ser Asn Trp Leu Trp Tyr Ile Arg Ile Phe Ile Met Ile Val Gly
 660 665 670
 Gly Leu Ile Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn
 675 680 685
 Arg Val Arg Lys Gly Tyr Ser Pro Leu Ser Leu Gln Thr His Ile Pro

690 695 700
 Ser Pro Arg Glu Pro Asp Arg Pro Glu Gly Ile Glu Glu Gly Gly Gly

 705 710 715 720
 Glu Gln Gly Lys Asp Arg Ser Val Arg Leu Val Asn Gly Phe Leu Ala

 725 730 735
 Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg

 740 745 750
 Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly

 755 760 765
 Arg Arg Gly Trp Glu Val Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr

 770 775 780

 Trp Ser Gln Glu Leu Lys Asn Ser Ala Ile Ser Leu Leu Asn Thr Thr
 785 790 795 800
 Ala Ile Val Val Ala Glu Gly Thr Asp Arg Val Ile Glu Ala Leu Gln

 805 810 815
 Arg Val Gly Arg Ala Ile Leu Asn Ile Pro Arg Arg Ile Arg Gln Gly

 820 825 830
 Leu Glu Arg Ala Leu Leu

 835
 <210> 110
 <211> 855
 <212> PRT
 <213> Human immunodeficiency virus

 <400> 110
 Met Lys Ala Lys Gly Thr Gln Arg Asn Trp Gln Ser Leu Trp Lys Trp

 1 5 10 15
 Gly Thr Leu Ile Leu Gly Leu Val Ile Ile Cys Ser Ala Ser Asn Asp

 20 25 30
 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

 35 40 45
 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Ala Glu Val

 50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Val Ala Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Glu

85 90 95

Asn Asn Met Val Asp Gln Met Gln Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Glu Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Phe Asn Asn Thr Thr Asn Asn Thr Thr Asn Thr Arg

130 135 140

Asn Asp Gly Glu Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr Thr

145 150 155 160

Ser Leu Arg Asp Lys Ile Lys Lys Glu Tyr Ala Leu Phe Tyr Asn Leu

165 170 175

Asp Val Val Gln Met Asp Asn Asp Asn Ser Ser Tyr Arg Leu Thr Ser

180 185 190

Cys Asn Thr Ser Ile Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Thr

195 200 205

Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Val Ile Leu Lys

210 215 220

Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser

225 230 235 240

Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu

245 250 255

Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Val Ile Arg Ser Lys

260 265 270

Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Asp Pro

275 280 285

Ile Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Arg Ile

290 295 300

Thr Met Gly Pro Gly Arg Val Leu Tyr Thr Thr Gly Gln Ile Ile Gly

305 310 315 320
 Asp Ile Arg Lys Ala His Cys Asn Ile Ser Lys Thr Lys Trp Asn Asn
 325 330 335
 Thr Leu Gly Gln Ile Val Lys Lys Leu Arg Glu Gln Phe Met Asn Lys
 340 345 350
 Thr Ile Val Phe Gln Arg Ser Ser Gly Gly Asp Pro Glu Ile Val Met

 355 360 365
 His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Thr Gln
 370 375 380
 Leu Phe Asn Ser Thr Trp Arg Ser Asn Ser Thr Trp Asn Asp Thr Thr
 385 390 395 400
 Glu Thr Asn Asn Thr Asp Leu Ile Thr Leu Pro Cys Arg Ile Lys Gln
 405 410 415
 Ile Val Asn Met Trp Gln Lys Val Gly Lys Ala Met Tyr Ala Pro Pro
 420 425 430

 Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu
 435 440 445
 Ile Arg Asp Gly Gly Ser Asn Asn Thr Glu Thr Phe Arg Pro Gly Gly
 450 455 460
 Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val
 465 470 475 480
 Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Arg Ala Lys Arg Arg
 485 490 495
 Val Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Leu Leu Phe

 500 505 510
 Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr
 515 520 525
 Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln
 530 535 540
 Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu
 545 550 555 560
 Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu

565 570 575
 Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly
 580 585 590
 Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn
 595 600 605
 Lys Ser Leu Asp Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu
 610 615 620
 Arg Glu Ile Asp Asn Tyr Thr Gly Leu Ile Tyr Thr Leu Ile Glu Gln
 625 630 635 640
 Ser Gln Asn Gln Gln Glu Arg Asn Glu Gln Glu Leu Leu Glu Leu Asp

 645 650 655
 Lys Trp Ala Ser Leu Trp Asn Trp Phe Asn Ile Thr Asn Trp Leu Trp
 660 665 670
 Tyr Ile Lys Ile Phe Ile Met Ile Ile Gly Gly Leu Ile Gly Leu Arg
 675 680 685
 Ile Val Phe Ala Val Leu Ser Ile Ile Asn Arg Val Arg Lys Gly Tyr
 690 695 700
 Ser Pro Leu Ser Phe Gln Thr Leu Thr His His Gln Arg Glu Pro Asp
 705 710 715 720

 Arg Pro Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp Lys Asp Arg
 725 730 735
 Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu
 740 745 750
 Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu
 755 760 765
 Ile Ala Ala Arg Thr Val Glu Leu Leu Gly Arg Ser Ser Leu Lys Gly
 770 775 780
 Leu Arg Leu Gly Trp Glu Gly Leu Lys Tyr Leu Trp Asn Leu Leu Leu

 785 790 795 800
 Tyr Trp Gly Arg Glu Leu Lys Asn Ser Ala Ile Asn Leu Leu Asp Thr
 805 810 815

Val Ala Ile Ala Val Ala Asn Trp Thr Asp Arg Ala Ile Glu Val Val

820 825 830

Gln Arg Val Gly Arg Ala Val Leu Asn Ile Pro Val Arg Ile Arg Gln

835 840 845

Gly Leu Glu Arg Ala Leu Leu

850 855

<210> 111

<211> 2517

<212> DNA

<213> Human immunodeficiency virus

<400> 111

atgcgcgtgc gcggcatgca gcgcaactgg cagcacctgg gcaagtgggg cctgctgttc	60
ctgggcatcc tgatcatctg caacgccacc gagaacctgt gggtgaccgt gtactacggc	120
gtgccctgtt ggaaggagge caccaccacc ctgttctgcg cctccgacgc caagtcttac	180
gagcgcgagg tgcacaactg gtgggccacc cagcctgcg tgcccaccga cccaacccc	240
caggaggtgg acctggagaa cgtgaccgag aacttcgaca tgtggaagaa caacatggtg	300
gagcagatgc acaccgacat catctccctg tgggaccagt cctgaagcc ctgcgtgaag	360
ctgaccccc tgctcgtgac cctgaactgc accgacgcca acgccaccgc caacgccacc	420
aaggagcacc ccgaggggcg cgccggcgcc atccagaact gctccttcaa catgaccacc	480
gaggtgcgcg acaagcagat gaaggtgcag gccctgttct accgcctgga catcgtgccc	540
atctccgaca acaactccaa cgagtaccgc ctgatcaact gcaacacctc caccatcacc	600
caggcctgcc ccaaggtgtc ctgggacccc atccccatcc actactgcgc ccccgccggc	660
tacgccatcc tgaagtcaa cgacaagaag ttcaacggca ccggccccctg caagaacgtg	720
tccaccgtgc agtgcaccca cggcatcaag cccgtggtgt ccaccagct gctgctgaac	780
ggctccctgg ccgaggagga gatcatcatc cgctcccaga acatctccga caacgccaag	840
accatcatcg tgcacctgaa cgagtccgtg cagatcaact gcacccgccc caacaacaac	900
acccgcaagt ccatccatc cggccccggc cgcgccttct acgccaccgg cgacatcatc	960
ggcgacatcc gcaaggccca ctgcaactgt tccggcacc agtggaaaca gaccctggag	1020
caggtgaaga agaagctgcg ctctacttc aacaccacca tcaagttcaa ctctctctcc	1080
ggcggcgacc ccgagatcac catgcactcc ttcaactgcc gcggcgagtt ctctactgc	1140
aacacctcca agctgttcaa cgacaccgtg tccaacgaca ccatcatcct gccctgccgc	1200

atcaagcaga tcgtgaacat gtggcaggag gtgggccgcg ccatgtacgc cgcccccatc 1260
gccggcaaca tcacctgcac ctccaacatc accggcctgc tgctgaccgc cgacggcggc 1320
cacaacgaga ccaacaagac cgagaccttc cgccccggcg gcggcaacat gaaggacaac 1380
tggcgctccg agctgtacaa gtacaagggtg gtggagatcg agccctggg cgtggcccc 1440
acccgcgcca agcgccaggt ggtgaagcgc gagaagcgc cgtgggcat cggcgccctg 1500
ttcttgggct tcttgggcgc cgccggctcc accatgggcg ccgcctccat caccctgacc 1560
gtgcaggccc gccagctgct gtccggcatc gtgcagcagc agtccaacct gctgcgcgcc 1620

atcgaggccc agcagcacct gctgcagctg accgtgtggg gcatcaagca gctgcaggcc 1680
cgctgctgg ccgtggagcg ctacctgaag gaccagcagc tgctgggcct gtggggctgc 1740
tccggcaagc tgatctgcac caccaacgtg ccttggaaact cctcctgtc caacaagtcc 1800
caggaggaga tctgggagaa catgacctgg atggagtggg agaaggagat caacaactac 1860
tccaacgaga tctaccgct gatcaggag tcccagaacc agcaggagaa gaacgagcag 1920
gagctgctgg ccttggacaa gtgggcctcc ctgtggaact ggttcgacat ctccaactgg 1980
ctgtggtaca tccgcatctt catcatgatc gtgggcggcc tgatcggcct gcgcatcgtg 2040

ttcgccgtgc tgtccatcgt gaaccgcgtg cgcaagggt actccccct gtccctgcag 2100
accacatcc cttcccccg cgagcccgac cgccccgagg gcatcgagga gggcggcggc 2160
gagcaggga aggaccgtc cgtgcgcctg gtgaacggct tcttggccct gatctgggac 2220
gacctgcgt ccttgtgcct gttctctac caccgcctgc gcgacctgct gctgatcgtg 2280
acccgcatcg tggagctgct gggccgccgc ggctgggagg tgctgaagta ctggtggaac 2340
ctgtgcagt actgggtcca ggagctgaag aactccgcca tctcctgct gaacaccacc 2400
gccatcgtgg tggccgaggg caccgaccgc gtgatcgagg cctgcagcg cgtgggccgc 2460

gccatcctga acatccccg ccgcatccgc cagggcctgg agcgcgcct gctgtaa 2517

<210> 112

<211> 2568

<212> DNA

<213> Human immunodeficiency virus

<400> 112

atgaaggcca agggcaccca gcgcaactgg cagtccctgt ggaagtggg caccctgatc 60
ctgggcctgg tgatcatctg ctccgcctcc aacgacctgt gggtgaccgt gtactacggc 120
gtgccgtgt ggaaggagge caccaccacc ctgttctgcg cctccgacgc caaggcctac 180
gacgccgagg tgcacaactg gtgggccacc cagcctgcg tgcccaccga ccccaacccc 240

caggaggtgg ccctggagaa cgtgaccgag aacttcaaca tgtgggagaa caacatggtg	300
gaccagatgc aggaggacat catctccctg tgggaccagt ccctgaagcc ctgcgtggag	360
ctgaccccc tgctcgtgac cctgaactgc accgacttca acaacaccac caacaacacc	420
accaacaccc gcaacgacgg cgaggcgag atcaagaact gctccttcaa catcaccacc	480
tccttgcgct acaagatcaa gaaggagtac gccctgttct acaacctgga cgtggtgcag	540
atggacaacg acaactctc ctaccgctg acctcctgca acacctccat catcaccag	600
gcctgcccc aggtgtcctt cacccecatc cccatccact actgcgcccc cgccggcttc	660
gtgatcctga agtgcaacaa caagacctt aacggcaccg gccctgcac caacgtgtcc	720
accgtgcagt gacccacgg catccgcccc gtggtgtcca cccagctgct gctgaacggc	780
tccttggcgg aggaggagat cgtgatccgc tccaagaact tcaccgacaa cgccaagacc	840
atcatcgtgc agctgaagga ccccatcgag atcaactgca cccgccccaa caacaacacc	900
cgcaagcgca tcacatggg ccccgccgc gtgctgtaca ccaccggcca gatcatcggc	960
gacatccga aggccactg caacatctcc aagaccaagt ggaacaacac cctgggccag	1020
atcgtgaaga agctgcgca gcagttcatg aacaagacca tcgtgttcca gcgtcctcc	1080
ggcgcgacc ccgagatcgt gatgactcc ttcaactgcg gcggcgagtt cttctactgc	1140
aacaccacc agctgttcaa ctccacctgg cgtccaact ccacctgga cgacaccacc	1200
gagaccaaca acaccacct gatcacctg cctgcccga tcaagcagat cgtgaacatg	1260
tggcagaagg tgggcaaggc catgtacgc ccccatct cggccagat ccgtgtctcc	1320
tccaacatca ccggcctgct gctgatccgc gacggcggct ccaacaacac cgagaccttc	1380
cgccccggcg gcggcaacat gaaggacaac tggcgctccg agctgtacaa gtacaagggtg	1440
gtgaagatcg agccctggg cgtggcccc acccgccca agcgccgct ggtgcagcgc	1500
gagaagcgcg ccgtgggcat cggcgccctg ctgttcggct tcctgggcgc cgccggctcc	1560
accatggcg ccgcctccat gacctgacc gtgcaggccc gccagctgct gtccggcatc	1620
gtgcagcagc agaacaacct gctgcgcgc atcgaggccc agcagcat gctgcagctg	1680
accgtgtggg gcatcaagca gctgcaggcc cgcgtgtgg ccgtggagcg ctacctgaag	1740
gaccagcagc tgctgggcat ctggggctgc tccggcaagc tgatctgcac caccaccgtg	1800
ccctggaacg cctcctggtc caacaagtcc ctggacgaca tctggaacaa catgacctgg	1860
atggagtggg agcgcgagat cgacaactac accggcctga tctacacct gatcgagcag	1920
tcccagaacc agcaggagcg caacgagcag gagctgtgg agctggacaa gtgggcctcc	1980

ctgtggaact gggtcaacat caccaactgg ctgtggtaca tcaagatctt catcatgata 2040
atcggcgccc tgatcgccct gcgcatcgtg ttcccggtgc tgtccatcat caaccgcgtg 2100
cgcaagggtt actccccct gtcttccag acctgaccc accaccagcg cgagcccgac 2160
cgcccgccc gcacgagga ggaggcggc gagcaggaca aggaccgctc catccgcctg 2220
gtgtccgctt tctggccct ggcctgggac gacctgcgtt ccctgtgcct gttctctac 2280
caccgcctgc gcgacttcat cctgatcgcc gccgcaccg tggagctgct gggccgctcc 2340
tcctgaagg gcctgcgctt gggctgggag ggctgaagt acctgtgaa cctgctgctg 2400

tactggggcc gcgagctgaa gaactccgcc atcaacctgc tggacaccgt ggccatcgcc 2460
gtggccaact ggaccgaccg cgccatcgag gtggtgcagc gcgtgggccc cgccgtgctg 2520
aacatccccg tgcgcatccg ccaggccctg gagcgcgccc tgctgtaa 2568

<210> 113

<211> 498

<212> PRT

<213> Human immunodeficiency virus

<400> 113

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Gln Gln Ile Ile Glu Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys

100 105 110

Gln Lys Thr Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Lys Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His

130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240

 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Ser Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr

 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350
 Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Thr Asn Thr Thr Ile Met Met Gln Arg Gly Asn Phe Lys Gly
 370 375 380

Gln Lys Arg Ile Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile
 385 390 395 400
 Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly
 405 410 415
 Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe
 420 425 430
 Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe Leu
 435 440 445
 Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Phe
 450 455 460
 Gly Glu Glu Ile Thr Pro Ser Pro Lys Gln Glu Pro Lys Asp Lys Glu
 465 470 475 480
 Leu Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu
 485 490 495
 Ser Gln

<210> 114

<211> 1497

<212> DNA

<213> Human immunodeficiency virus

<400> 114

atgggcgccc gcgcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtgggc ctccccgcgag 120

ctggagcgt tgcacctgaa ccccgccctg ctggagacct ccgagggtcg ccagcagatc 180

atcgagcagc tgcagcccg cctgcagacc ggctccgagg agctgcgtc cctgtacaac 240

accgtggcca cctgtactg cgtgcaccag cgcattcgagg tgaaggacac caaggaggcc 300

ctggacaaga tcgaggagga gcagaacaag tccaagcaga agaccagca ggccgccgcc 360

gacaccggca actcctcaa ggtgtcccag aactaccca tcgtgcagaa cctgcagggc 420

cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtggtggag 480

gagaaggcct tctccccga ggtgatcccc atgttctcgg ccctgtccga gggcgccacc 540

ccccaggacc tgaacacat gctgaacacc gtggcgcc accaggccgc catgcagatg 600

ctgaaggaca ccatcaacga ggaggccgcc gattgggacc gcctgcaccc cgtgcacgcc 660

ggcccatcc ccccgccca gatgcgcgag ccccgcggt cgcacatcg cggcaccacc 720
 tccacctgc aggagcagat cggctggatg acctccaacc ccccatccc cgtgggcgag 780
 atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840
 tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgcttc 900
 ttcaagacc tgcgcgccga gcaggccacc caggacgtga agaactggat gaccgacacc 960

ctgctggtgc agaagccaa ccccgactgc aagaccatcc tgaaggcctt gggccccggc 1020
 gccaccttg aggagatgat gaccgcctgc caggcgctgg gcggccctc ccacaaggcc 1080
 cgcgtgctgg ccgaggccat gtcccagtg accaacacca ccatcatgat gcagcgcggc 1140
 aacttcaagg gccagaagcg catcatcaag tgcttcaact gcggcaagga gggccacatc 1200
 gcccgaact gccgcgccc ccgcaagaag ggctgctgga agtgcggcaa ggagggccac 1260
 cagatgaagg actgcaccga gcgccaggcc aacttcttgg gcaagatctg gccctccaac 1320
 aaggcgccgc ccgcaactt cctgcagtcc cgccccgagc ccaccgccc ccccgccgag 1380

tccttcggct tcggcgagga gatcacccc tcccccaagc aggagcccaa ggacaaggag 1440
 ctgtacccc tgccctcctt gaagtcctg ttggcaacg acccctgtc ccagtaa 1497

<210> 115

<211> 498

<212> PRT

<213> Human immunodeficiency virus

<400> 115

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Met Gly Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln

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

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Thr Asn Ala Asn Ile Met Met Gln Arg Gly Asn Phe Lys Gly
 370 375 380
 Pro Arg Arg Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile
 385 390 395 400

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

465 470 475 480
 Leu Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Ser Asp Pro Leu
 485 490 495
 Ser Gln

<210> 116
 <211> 1497
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 116

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtggc ctcccgcgag	120
ctggagcgct tcgccctgaa ccccgccctg ctggagaccg ccgagggtg ccagcagatc	180
atgggccagc tgcagccgc cctgcagacc ggcaagcagg agctgcgctc cctgtacaac	240
accgtggcca cctgtactg cgtgcaccag cgcacgagg tgaaggacac caaggaggcc	300
ctggacaaga tcgaggagga gcagaacaag tcccagcaga agaccagca ggccgcccgc	360
gacaaggcgc actcctccca ggtgtcccag aactaccca tcgtgcagaa cctgcagggc	420

cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtggtggag 480
gagaaggcct tctccccga ggtgatcccc atgtttctcg cctgtccga gggcgccacc 540
ccccaggacc tgaacacat gctgaacacc gtgggcggcc accaggccgc catgcagatg 600

ctgaaggaca ccatcaacga ggaggccgcc gattgggacc gcctgcaccc cgtgcacgcc 660
ggccccatcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720
tccaccctgc aggagcagat cggctggatg acctccaacc ccccatccc cgtgggcgag 780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggacccttc 900
ttcaagacc tgcgcgcga gcaggccacc caggacgtga agaactggat gaccgacacc 960
ctgctggtgc agaacgcaa ccccgactgc aagaccatcc tgaaggccct gggccccggc 1020

gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc 1080
cgctgctgg ccgaggccat gtcccagggtg accaacgcca acatcatgat gcagcgcggc 1140
aacttcaagg gccccgcgc catcgtgaag tgcttcaact gcggcaagga gggccacatc 1200
gcccgaact gccgcgccc ccgcaagaag ggctgctgga agtgcggcaa ggaggccac 1260
cagatgaagg actgcaccga gcgccaggcc aacttctctg gcaagatctg gcctccaac 1320
aaggcgccc ccggcaactt cctgcagtcc cggccgagc ccaccgccc ccccgccgag 1380
tccttcggct tcggcgagga gatcacccc tccccaaagc aggagcccaa ggacaaggag 1440

ctgtacccc tggcctcct gaagtcctg ttgggtccg acccctgtc ccagtaa 1497

<210> 117

<211> 498

<212> PRT

<213> Human immunodeficiency virus

<400> 117

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Ser Leu Leu Glu Thr Thr Glu Gly Cys Gln Gln Ile Met Glu Gln Leu

50 55 60

Gln Pro Ala Leu Lys Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn
 65 70 75 80
 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp
 85 90 95
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Ile Gln Asn Lys Ser Lys
 100 105 110
 Gln Lys Thr Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Lys Val
 115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His
 130 135 140
 Gln Ser Leu Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Met Met Leu Asn Ile Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Pro Gln Glu Gln Ile Gly Trp Met Thr Gly Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Lys Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Ser Ile Leu Arg Ala
325 330 335

Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
340 345 350

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Val Gln His Thr Asn Ile Met Met Gln Arg Gly Asn Phe Arg Gly
370 375 380

Gln Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Leu Ala
385 390 395 400

Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys

405 410 415

Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu
420 425 430

Gly Lys Ile Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn Phe Pro Gln
435 440 445

Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ile Phe Gly Met Gly
450 455 460

Glu Glu Ile Thr Ser Pro Pro Lys Gln Glu Gln Lys Asp Arg Glu Gln
465 470 475 480

Asp Pro Pro Leu Val Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu
485 490 495

Ser Gln

<210> 118

<211> 1497

<212> DNA

<213> Human immunodeficiency virus

<400> 118

atgggcgcc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtgggc ctcccgcgag 120

ctggagcgt tcgccctgaa cccctccctg ctggagacca ccgagggtg ccagcagatc 180

atggagcagc tgcagcccg cctgaagacc ggcaccgagg agctgcgctc cctgtacaac 240

accgtggcca cctgtactg cgtgcaccag cgcacgcagc tgaaggacac caaggaggcc 300

ctggacaaga tgcaggagat ccagaacaag tccaagcaga agaccagca ggcggccgcc 360

gacaccggca actcctcaa ggtgtcccag aactaccca tctgtcagaa cggccagggc 420

cagatggtgc accagtcctt gtccccccgc accctgaacg cctgggtgaa ggtgatcgag 480

gagaaggcct tctccccga ggtgatcccc atgttctccg cctgtccga gggcgccacc 540

ccccaggacc tgaacatgat gctgaacatc gtgggcggcc accaggccgc catgcagatg 600

ctgaaggaca ccatcaacga ggaggccgcc gagtgggacc gcctgcaccc cgtgcacgcc 660

ggccccatcc ccccgcca gatgcgcgag ccccgcggt cgcacatgc cggcaccacc 720

tccaccccc aggagcagat cggctggatg accggcaacc ccccatccc cgtgggcgac 780

atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840

tccatcctgg acatcaagca gggccccaag gagcccttc gcgactacgt ggaccgcttc 900

ttcaagacc tgcgcgccga gcaggccacc caggaggtga agaactggat gaccgagacc 960

ctgctggtgc agaagccaa ccccgactgc aagtccatcc tgcgcgcctt gggccccggc 1020

gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc 1080

cgctgctgg ccgaggccat gtcccagggtg cagcacacca acatcatgat gcagcgcggc 1140

aacttccgcg gccagaagcg catcaagtgc ttcaactgcg gcaaggaggg ccacctggcc 1200

cgcaactgcc gcgcccccg caagaagggc tgctggaagt gcggaagga gggccaccag 1260

atgaaggact gcaccgagcg ccaggccaac ttctgggca agatctggcc ctctccaag 1320

ggccgccccg gcaacttccc ccagtccgc cccgagccca ccgcccccc cgccgagatc 1380

ttcgcatgg gcaggagat cactcccc ccgaagcagg agcagaagga ccgcgagcag 1440

gaccccccc ttgtgtcctt gaagtccctg ttggcaacg acccctgtc ccagtaa 1497

<210> 119

<211> 498

<212> PRT

<213> Human immunodeficiency virus

<400> 119

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Met Gly Gln Leu

50 55 60

Gln Pro Ala Leu Lys Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Ile Gln Asn Lys Ser Lys
100 105 110

Gln Lys Thr Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Lys Val
115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His
130 135 140

Gln Ser Leu Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu
145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Met Met Leu Asn Ile Val Gly
180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu

195 200 205

Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro
210 215 220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
225 230 235 240

Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Gly Asn Pro Pro Ile
245 250 255

Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
260 265 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly

275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Ser Ile Leu Arg Ala
 325 330 335
 Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly

340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Gln Asn Thr Asp Ile Met Met Gln Arg Gly Asn Phe Arg Gly
 370 375 380
 Pro Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Leu Ala
 385 390 395 400
 Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys
 405 410 415

Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu
 420 425 430
 Gly Lys Ile Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn Phe Pro Gln
 435 440 445
 Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Asn Phe Gly Met Gly
 450 455 460
 Glu Glu Met Ile Ser Ser Pro Lys Gln Glu Gln Lys Asp Arg Glu Gln
 465 470 475 480
 Tyr Pro Pro Leu Val Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu

485 490 495
 Ser Gln

<210> 120

<211> 1497

<212> DNA

<213> Human immunodeficiency virus

<400> 120

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgaggc ctcccgcgag	120
ctggagcgct tgcacctgaa ccccgccctg ctggagaccg ccgagggctg ccagcagatc	180
atgggcccagc tgcagccccg cctgaagacc ggaccgagg agctgcgctc cctgtacaac	240
accgtggcca cctgtactg cgtgcaccag cgcacgagg tgaaggacac caaggaggcc	300

ctggacaaga tgcaggagat ccagaacaag tccaagcaga agaccagca ggccgcccgc	360
gacaccggca actcctccaa ggtgtcccag aactaccca tcgtgcagaa cggccagggc	420
cagatggtgc accagtcctt gtccccccgc accctgaacg cctgggtgaa ggtgatcgag	480
gagaaggcct tctccccga ggtgateccc atgttctcgg cctgtccga gggcgccacc	540
ccccaggacc tgaacatgat gctgaacatc gtggcgggcc accaggccgc catgcagatg	600
ctgaaggaca ccatcaacga ggaggccgcc gaggggacc gcctgcaccc cgtgcacgcc	660
ggccccatcc cccccggcca gatgcgcgag ccccgcggtt ccgacatgc cggcaccacc	720

tccacctgc aggagcagat cggctggatg accggcaacc ccccatccc cgtgggcgac	780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg	840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc	900
ttcaagacc tgcgcgcga gcaggccacc caggaggtga agaactggat gaccgagacc	960
ctgctggtgc agaagccaa ccccgactgc aagtccatcc tgcgcgcctt gggccccggc	1020
gccacctgg aggagatgat gaccgcctgc caggcgctgg gcggccccgg ccacaaggcc	1080
cgcgtgctgg ccgaggccat gtcccagtg cagaacaccg acatcatgat gcagcgcggc	1140

aacttcgcg gcccgaagcg catcaagtgc ttcaactgcg gcaaggagg ccacctggcc	1200
cgcaactgcc gcgcccccg caagaagggc tgctggaagt gcggaagga ggccaccag	1260
atgaaggact gcaccgagcg ccaggccaac ttctgggca agatctggcc ctctccaag	1320
ggccgccccg gcaacttccc ccagtcctgc cccgagccca ccgcccccc cgcgagaac	1380
ttcgcatgg gcgaggagat gatctctcc cccaagcagg agcagaagga ccgagagcag	1440
taccccccc tgggtgctct gaagtccctg ttgggcaacg acccctgtc ccagtaa	1497

<210> 121

<211> 500

<212> PRT

<213> Human immunodeficiency virus

<400> 121

Met Gly Ala Arg Ala Ser Ile Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Lys Phe Ser Ile Asn Pro

35 40 45

Ser Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Ile Arg Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Lys Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Val Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Cys Lys

100 105 110

Gln Lys Thr Gln His Ala Ala Ala Asp Thr Gly Asn Ser Ser Ser Ser

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His

130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly

180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu

195 200 205

Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro

210 215 220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr

225 230 235 240

Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Ser Asn Pro Pro Ile

245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly

 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Asp Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Ser Ile Leu Arg Ala
 325 330 335
 Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350

 Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Gln Asn Thr Asn Thr Asn Ile Met Met Gln Arg Gly Asn Phe
 370 375 380
 Arg Gly Gln Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His
 385 390 395 400
 Leu Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
 405 410 415
 Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn

 420 425 430
 Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe
 435 440 445
 Pro Gln Ser Arg Thr Glu Pro Thr Ala Pro Pro Ala Glu Asn Leu Arg
 450 455 460
 Met Gly Glu Glu Ile Thr Ser Ser Leu Lys Gln Glu Leu Lys Thr Arg
 465 470 475 480
 Glu Pro Tyr Asn Pro Ala Ile Ser Leu Lys Ser Leu Phe Gly Asn Asp
 485 490 495

Pro Leu Ser Gln

500

<210> 122

<211> 1503

<212> DNA

<213> Human immunodeficiency virus

<400> 122

atgggcgccc gcgcctccat cctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtgggc ctcccgcgag	120
ctggagaagt tctccatcaa cccctccctg ctggagacct cggagggtg cgcgcagatc	180
atccgccagc tgcagccgc cctgcagacc ggcaccgagg agctgaagtc cctgtacaac	240
accgtggcgc tgctgtactg cgtgcaccag cgcctcgacg tgaaggacac caaggaggcc	300
ctggacaaga tcgaggagga gcagaacaag tgcaagcaga agaccagca cgcgcgcgc	360
gacaccggca actcctctc ctcctcccag aactaccca tctgtcagaa cgcgcagggc	420
cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtgggtggag	480
gagaaggcct tctccccga ggtgatccc atgttcaccg ccctgtccga gggcgccacc	540
ccccaggacc tgaacacat gctgaacacc gtgggcggcc accaggccgc catgcagatg	600
ctgaaggaca ccatcaacga ggaggccgcc gaggggacc gcctgcaccc cgtgcacgcc	660
ggccccatcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc	720
tccacctgc aggagcagat cggttgatg acctccaacc ccccatccc cgtgggcgag	780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg	840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc	900
ttcaagacc tgcgcgccga gcaggccacc caggaggtga agaactggat gaccgacacc	960
ctgctggtgc agaacgcaa ccccgactgc aagtccatcc tgcgcgcct gggccccggc	1020
gccacctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccct ccacaaggcc	1080
cgcgtgctgg ccgaggccat gtcccagggt cagaacacca acaccaacat catgatgcag	1140
cgcggaact tcccgggcca gaagcgcac aagtgttca actgcggcaa ggagggccac	1200
ctggccccga actgcgcgc ccccgcaag aagggtgct ggaagtgcgg caaggagggc	1260
caccagatga aggactgcac cgagcgccag gccaacttc tgggcaagat ctggccctcc	1320
aacaaggcc gccccgcaa ctccccccag tccgcaccg agccaccgc ccccccgcc	1380
gagaacctgc gcatgggcga ggagatcacc tcctccctga agcaggagct gaagaccgc	1440
gagccctaca accccgcat ctccctgaag tcctgttcg gcaacgacc cctgtcccag	1500

taa

1503

<210> 123

<211> 488

<212> PRT

<213> Human immunodeficiency virus

<400> 123

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys

20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys

100 105 110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His

130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly

180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu

195 200 205
 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270

 Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly

 340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg
 370 375 380
 Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
 385 390 395 400
 Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
 405 410 415

 Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
 420 425 430
 Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe
 435 440 445
 Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg

450	455	460	
Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp			
465	470	475	480
Lys Glu Leu Tyr Pro Leu Ala Ser			
	485		
<210>	124		
<211>	1467		
<212>	DNA		
<213>	Human immunodeficiency virus		
<400>	124		
atgggcgccc gcgcctccgt gctgtccggc ggcgagctgg accgctggga gaagatccgc		60	
ctgcgccccg gcggcaagaa gaagtacaag ctgaagcaca tcgtgtgggc ctcccgcgag		120	
ctggagcgct tcgccgtgaa ccccgccctg ctggagacct ccgagggtg cgcgcagatc		180	
ctgggccagc tgcagccctc cctgcagacc ggctccgagg agctgcgctc cctgtacaac		240	
accgtggcca ccctgtactg cgtgcaccag cgcacatcagg tgaaggacac caaggaggcc		300	
ctggagaaga tcgaggagga gcagaacaag tccaagaaga agggccagca ggccgcccgc		360	
gacaccggca actcctccca ggtgtcccag aactaccca tcgtgcagaa cctgcagggc		420	
cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtggtggag		480	
gagaaggcct tctccccga ggtgateccc atgttctcgc cctgtccga gggcgccacc		540	
ccccaggacc tgaacaccat gctgaacacc gtggcgggcc accaggccgc catgcagatg		600	
ctgaaggaga ccatcaacga ggaggccgcc gaggggacc gcctgcaccc cgtgcacgcc		660	
ggccccatcg cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc		720	
tccacctgc aggagcagat cggctggatg accaacaacc ccccatccc cgtgggcgag		780	
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccacc		840	
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc		900	
tacaagacc tgcgcgccga gcaggcctcc caggaggtga agaactggat gaccgagacc		960	
ctgctggtgc agaacgcaa ccccgactgc aagaccatcc tgaaggccct gggccccgcc		1020	
gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc		1080	
cgcgtgctgg ccgaggccat gtcccagggtg accaactccg ccaccatcat gatgcagcgc		1140	
ggcaacttc gcaaccagcg caagaccgtg aagtgttca actgcggcaa ggagggccac		1200	
atcgccaaga actgccgcgc cccccgaag aagggtgctt ggaagtgcgg caaggagggc		1260	

caccagatga aggactgcac cgagcgccag gccaaacttc tgggcaagat ctggccctcc 1320
 cacaagggcc gccccggcaa cttctgcag tcccgccccg agcccaccgc ccccccgag 1380
 gagtccctcc gcttcggcga ggagaccacc accccctccc agaagcagga gcccacgcac 1440
 aaggagctgt accccctggc ctcctaa 1467

<210> 125

<211> 500

<212> PRT

<213> Human immunodeficiency virus

<400> 125

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Lys Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys

20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys

100 105 110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His

130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly

180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
 195 200 205

 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Ile Ser Ile Leu Asp Ile Arg Gln Gly

 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Asp Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350

 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Thr Asn Ser Thr Thr Ile Met Met Gln Arg Gly Asn Phe Arg
 370 375 380
 Asp Gln Arg Lys Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
 385 390 395 400
 Ile Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
 405 410 415
 Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn

 420 425 430

Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe
 435 440 445
 Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
 450 455 460
 Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
 465 470 475 480
 Lys Glu Leu Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp
 485 490 495

Pro Ser Ser Gln
 500

<210> 126
 <211> 1503
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 126

atgggcgcgc ggcctccgt gctgtccggc ggcaagctgg acaagtggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtacaag ctgaagcaca tcgtgtgggc ctcccgcgag	120
ctggagcgct tcgccgtgaa ccccgccctg ctggagacct ccgagggctg ccgccagatc	180
ctggggccagc tgcagcccg cctgcagacc ggctccgagg agctgcgctc cctgtacaac	240
accgtggcca cctgtactg cgtgcaccag cgcacgagg tgaaggacac caaggaggcc	300
ctggacaaga tcgaggagga gcagaacaag tccaagaaga aggccagca ggccgcccgc	360
gacaccggca actcctccca ggtgtcccag aactaccca tcgtgcagaa cctgcagggc	420
cagatggtgc accaggccat ctcccccg cccctgaacg cctgggtgaa ggtggtggag	480
gagaaggcct tctccccga ggtgatccc atgttctccg cctgtccga gggcgccacc	540
ccccaggacc tgaacacat gctgaacacc gtggcgggc accaggccgc catgcagatg	600
ctgaaggaga ccatcaacga ggaggccgc gagtgggacc gcctgcaccc cgtgcacgcc	660
ggccccatcg ccccgccca gatgcgcgag ccccgcggt cgcacatgc cggcaccacc	720
tccacctgc aggagcagat cggctggatg accaacaacc ccccatccc cgtgggcgag	780
atctacaagc gctggatcat cctgggcctg aacaagatc tgccgatgta ctccccatc	840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc	900
tacaagacc tgcgcgccga gcaggcctcc caggacgtga agaactggat gaccgagacc	960
ctgctggtgc agaacgcaa ccccgactgc aagaccatcc tgaaggccct gggccccgc	1020

gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc 1080
 cgcgtgctgg ccgaggccat gtcccagggtg accaactcca ccaccatcat gatgcagcgc 1140
 ggcaacttcc gcgaccagcg caagatcgtg aagtgcctca actgcggcaa ggagggccac 1200
 atgccccga actgccgcgc cccccgaag aagggtgctt ggaagtgcgg caaggagggc 1260
 caccagatga aggactgcac cgagcgccag gccaaacttc tgggcaagat ctggccctcc 1320
 cacaagggcc gccccggcaa ctctctgcag tcccgccccg agcccaccgc ccccccgag 1380
 gagtcttcc gcttcggcga ggagaccacc accccctccc agaagcagga gcccatcgac 1440
 aaggagctgt accccctggc ctccctgaag tccctgttcg gcaacgaccc ctctcccag 1500
 taa 1503

<210> 127
 <211> 492
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 127

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

130 135 140
 Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala
 145 150 155 160
 Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala
 165 170 175
 Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
 180 185 190
 Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu

 195 200 205
 Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln
 210 215 220
 Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu
 225 230 235 240
 Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val Gly
 245 250 255
 Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
 260 265 270

 Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
 275 280 285
 Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu
 290 295 300
 Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
 305 310 315 320
 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
 325 330 335
 Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly

 340 345 350
 Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala Asn
 355 360 365
 Asn Thr Asn Ile Met Met Gln Arg Ser Asn Phe Lys Gly Pro Lys Arg
 370 375 380
 Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn

385 390 395 400
 Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly
 405 410 415

 His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys
 420 425 430
 Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Asn Arg
 435 440 445
 Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr
 450 455 460
 Thr Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser
 465 470 475 480
 Leu Lys Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln

 485 490
 <210> 128
 <211> 1479
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 128

atgggcgccc ggcctccat cctgcgcggc ggcaagctgg acaagtggga gaagatccgc	60
ctgcgccccg gcggcaagaa gcactacatg ctgaagcacc tgggtgtgggc ctcccgcgag	120
ctggagcgc tgcacctgaa ccccgccctg ctggagacct ccgagggtcg caagcagatc	180
atcaagcagc tgcagccgc cctgcagacc ggcaaccagg agctgcgctc cctgtacaac	240
accgtggcca ccctgtactg cgtgcacgag aagatcgagg tgcgcgacac caaggaggcc	300
ctggacaaga tcgaggagga gcagaacaag tcccagcaga agaccagca ggccaaggcc	360
gccgacggca aggtgtccca gaactacccc atcgtgcaga acctgcaggg ccagatggtg	420
caccaggcca tctcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc	480
ttctccccg aggtgatccc catgttcacc gccctgtccg agggcgccac cccccaggac	540
ctgaacacca tgetgaacac cgtgggcggc caccaggccg ccatgcagat gctgaaggac	600
accatcaacg aggaggccgc cgagtgggac cgctgcacc ccgtgcacgc cggccccatc	660
gccccgggc agatgcgcga gccccgggc tccgacatcg ccggcaccac ctccacctg	720
caggagcaga tcgcctggat gacctcaac cccccatcc ccgtgggcga catctacaag	780

cgctggatca tcctgggcct gaacaagatc gtgcgcatgt actccccgt gtccatcctg 840
gacatcaagc agggcccca ggagcccttc cgcgactacg tggaccgctt cttcaagacc 900
ctgcgcgcgc agcaggccac ccaggacgtg aagaactgga tgaccgacac cctgctggtg 960
cagaacgcca accccgactg caagaccatc ctgcgcgccc tgggccccgg cgccaccctg 1020
gaggagatga tgaccgcctg ccagggcgtg ggcgggccct cccacaaggc ccgcgtgctg 1080
gccgaggcca tgtcccaggc caacaacacc aacatcatga tgcagcgctc caacttcaag 1140

ggccccaagc gcatcgtgaa gtgcttcaac tgcggcaagg agggccacat cggccgcaac 1200
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggagggcca ccagatgaag 1260
gactgcaccg agcgcaggc caacttcctg ggcaagatct ggccctccca caagggccgc 1320
cccggcaact tcctgcagaa ccgccccgag ccaccgccc ccccgccga gtccttcgcg 1380
ttcgaggaga ccacccccgc cccaagcag gagcccaagg accgcgagcc cctgacctcc 1440
ctgaagtccc tgttcggtc cgacccccctg tcccagtaa 1479

<210> 129

<211> 495

<212> PRT

<213> Human immunodeficiency virus

<400> 129

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Thr Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Ile Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Met Lys Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Glu Arg Ile Glu Val Arg Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln

100 105 110

Gln Lys Thr Gln Gln Ala Glu Ala Ala Asp Gly Asp Asn Gly Lys Val

115 120 125
Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His

130 135 140
Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser
165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu
195 200 205

Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala
210 215 220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
225 230 235 240

Ser Thr Leu Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile
245 250 255

Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
260 265 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Lys Gln Gly

275 280 285
Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu

290 295 300
Arg Ala Glu Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr

305 310 315 320
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala

325 330 335
Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly

340 345 350
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser

355 360 365

Gln Ala Asn Asn Thr Asn Ile Met Met Gln Arg Ser Asn Phe Lys Gly
 370 375 380
 Pro Lys Arg Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile
 385 390 395 400
 Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly
 405 410 415
 Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe

420 425 430
 Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu
 435 440 445
 Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe
 450 455 460
 Glu Glu Thr Thr Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro
 465 470 475 480
 Leu Thr Ser Leu Lys Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln
 485 490 495

<210> 130
 <211> 1488
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 130

atgggcgccc ggcctccat cctgcgcggc ggcaagctgg acacctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gcactacatg atcaagcacc tgggtgtgggc ctcccgcgag	120
ctggagcgc tgcacctgaa ccccgccctg ctggagacct ccgagggctg caagcagatc	180
atgaagcagc tgcagccgc cctgcagacc ggcaccgagg agctgcgctc cctgtacaac	240
accgtggcca cctgtactg cgtgcacgag cgcacgcagg tgcgcgacac caaggaggcc	300
ctggacaaga tcgaggagga gcagaacaag tcccagcaga agaccagca ggccgaggcc	360
gccgacggcg acaacggcaa ggtgtcccag aactaccca tcgtgcagaa cctgcagggc	420
cagatggtgc accaggccat ctccccgc accctgaacg cctgggtgaa ggtggtggag	480
gagaaggcct tctccccga ggtgateccc atgttcaccg cctgtccga gggcgccacc	540
ccccaggacc tgaacacat gctgaacacc gtggcgcc accaggccgc catgcagatg	600
ctgaaggaca ccatcaacga ggaggccgcc gattgggacc gcctgcaccc cgtgcacgcc	660

ggccccgtgg cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720
tccaccctgc aggagcagat cgcttggatg acctccaacc ccccatccc cgtgggcgac 780

atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840
tccatcctgg acatcaagca gggccccaag gagcccttcc gcgactacgt ggaccgcttc 900
ttcaagacc tgcgcgcga gcagccacc caggacgtga agaactggat gaccgacacc 960
ctgtggtgc agaacgcaa ccccgactgc aagaccatcc tgcgcgcctt gggccccggc 1020
gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc 1080
cgcgtgctgg ccgaggccat gtcccaggcc aacaacacca acatcatgat gcagcgtcc 1140
aacttcaagg gcccgaagcg catcgtgaag tgcttcaact gcggcaagga gggccacatc 1200

gccccgaact gccgcgcccc ccgcaagaag ggctgctgga agtgcggcaa ggagggccac 1260
cagatgaagg actgcaccga gcgccaggcc aatttcttgg gcaagatctg gccctccac 1320
aagggccgcc ccggcaactt cctgcagtcc cgccccgagc ccaccgcccc ccccgccgag 1380
tccttcgct tcgaggagac ccccccgcc cccaagcagg agcccaagga ccgcgagccc 1440
ctgacctccc tgaagtcctt gttcggctcc gacccctgt cccagtaa 1488

<210> 131

<211> 499

<212> PRT

<213> Human immunodeficiency virus

<400> 131

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

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

Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
 100 105 110
 Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val
 115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His
 130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu

 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala
 210 215 220

 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Ser Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu

 290 295 300
 Arg Ala Glu Gln Ala Ser Gln Asp Val Lys Asn Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
 325 330 335
 Leu Gly Pro Glu Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350

Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Ala Thr Asn Ser Ala Ala Val Met Met Gln Arg Gly Asn Phe Lys
370 375 380

Gly Pro Arg Lys Ile Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His
385 390 395 400

Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
405 410 415

Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
420 425 430

Phe Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe

435 440 445

Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly
450 455 460

Phe Gly Glu Glu Ile Thr Pro Ser Gln Lys Gln Glu Gln Lys Asp Lys
465 470 475 480

Glu Leu Tyr Pro Leu Thr Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro
485 490 495

Leu Ser Gln

<210> 132

<211> 1500

<212> DNA

<213> Human immunodeficiency virus

<400> 132

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcaca tcgtgtgggc ctcccgcgag	120
ctggagcgct tcgccctgaa ccccgccctg ctggagacct ccgagggctg caagcagatc	180
atcgccagc tgcagccgc catccagacc ggctccgagg agctgcgctc cctgtacaac	240
accgtggcca cctgtactg cgtgcacgag cgcacgagg tgaaggacac caaggaggcc	300
ctggagaaga tcgaggagga gcagaacaag tccaagaaga aggcccagca ggccgcccgc	360

gacaccggca actcctccca ggtgtcccag aactaccca tcgtgcagaa cctgcagggc 420

cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtgatcgag 480

gagaaggcct tctccccga ggtgatcccc atgttctccg cctgtccga gggcgccacc 540

ccccaggacc tgaacacat gctgaacacc gtggcgccg accaggccgc catgcagatg 600

ctgaaggaga ccatcaacga ggaggccgcc gagtgggacc gcctgcaccc cgtgcacgcc 660

ggccccgtgg cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720

tccacctgc aggagcagat cggctggatg acctccaacc ccccatccc cgtggcgag 780

atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840

tccatcctgg acatccgcca gggccccaag gagcccttcc gcgactacgt ggaccgttc 900

tacaagacc tgcgcgccga gcaggcctcc caggacgtga agaactggat gaccgagacc 960

ctgctggtgc agaagccaa ccccgactgc aagaccatcc tgaaggcct gggccccgag 1020

gccacctg aggatgat gaccgcctgc caggcgctgg gcggcccctc ccacaaggcc 1080

cgctgctgg ccaggccat gtcccaggcc accaactccg ccgccgtgat gatgcagcgc 1140

ggcaacttca agggccccg caagatcatc aagtgttca actgcggcaa ggaggccac 1200

atcgccaaga atgcccgc ccccgcaag aagggtgct ggaagtgcgg caaggagggc 1260

caccagatga aggactgcac cgagcgccag gccaacttcc tgggcaagat ctggccctcc 1320

cacaaggcc gccccggcaa ctctctgag tccgccccg agccaccgc cccccgcc 1380

gagtccttc gttcggcga ggagatcacc cctcccaga agcaggagca gaaggacaag 1440

gagctgtacc cctgacctc cctgaagtcc ctgttcggca acgacccct gtcccagtaa 1500

1500

<210> 133

<211> 492

<212> PRT

<213> Human immunodeficiency virus

<400> 133

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

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

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

290 295 300
 Ala Thr Gln Glu Val Lys Gly Trp Met Thr Asp Thr Leu Leu Val Gln
 305 310 315 320
 Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Gly
 325 330 335
 Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro
 340 345 350
 Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala Thr Asn
 355 360 365

Thr Ala Ile Met Met Gln Lys Ser Asn Phe Lys Gly Gln Arg Arg Ile
 370 375 380
 Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Lys Asn Cys
 385 390 395 400
 Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Arg Glu Gly His
 405 410 415
 Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
 420 425 430
 Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro

435 440 445
 Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Phe Arg Glu Glu Ile
 450 455 460
 Thr Pro Ser Pro Lys Gln Glu Gln Lys Asp Glu Gly Leu Tyr Pro Pro
 465 470 475 480
 Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro
 485 490

<210> 134
 <211> 1479
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 134

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc atgaagcacc tggtgtgggc ctcccgcgag 120

ctggagcgct tcgcccctgga ccccgccctg ctggagacct cggagggtg ccagaagatc 180
atcgccagc tgcagccctc cctgcagacc ggctccgagg agctgcgctc cctgtacaac 240
accgtggcgg tgcgtactg cgtgcaccag aaggtggagg tgaaggacac caaggaggcc 300
ctggagaagc tggaggagga gcagaacaag tcccagcaga agaccagca ggccgcccgc 360
gacaaggcgg tgtcccagaa ctaccccatc gtgcagaacc tgcagggccca gatggtgcac 420
caggccatct cccccgcac cctgaacgcc tgggtgaagg tgatcgagga gaaggccttc 480

tccccgagg tgatcccat gttctccgc ctgtccgagg gcgccacccc ccaggacctg 540
aacaccatgc tgaacacctg gggcgccac caggccgcca tgcagatgct gaaggacacc 600
atcaacgagg aggcgcgga gtgggaccgc ctgcacccc tgcacgccgg ccccatcccc 660
cccggccaga tgcgcgagcc ccgcggctcc gacatcgccg gcaccacctc caccctgcag 720
gagcagatcc agtggatgac ctccaacccc cccgtgcccg tggcgcatat ctacaagcgc 780
tggatcatcc tgggcctgaa caagatcgtg cgcgtgtact ccccgctgc catcctggac 840
atccgcccagg gcccgaagga gcccttcgc gactacgtgg accgcttctt caagaccctg 900

cgcgccgagc aggcaccca ggaggtgaag ggctggatga ccgacacct gctggtgcag 960
aacgccaacc ccgactgcaa gaccatcctg aaggccctgg gcccggcgc caccctggag 1020
gagatgatga ccgctgcca gggcgtgggc ggccccggcc acaaggccc cgtgctggcc 1080
gaggccatgt ccaggccac caacaccgcc atcatgatgc agaagtcaa cttcaagggc 1140
cagcgccgca tcgtgaagtg cttcaactgc ggcaaggagg gccacatgc caagaactgc 1200
cgccccccc gcaagaaggc ctgctggaag tgcggccgcg agggccacca gatgaaggac 1260
tgaccgagc gccaggccaa cttcctgggc aagatctggc cctccaaca gggccgcccc 1320

ggcaacttc tgcagtccg ccccgagccc accgcccc cgcgcgagtc cttcggttc 1380
cgcgaggaga tcacccctc cccaagcag gagcagaagg acgagggcct gtacccccc 1440
ctggcctccc tgaagtcct gttcggaac gaccctaa 1479

<210> 135

<211> 497

<212> PRT

<213> Human immunodeficiency virus

<400> 135

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Met Lys

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

 Thr Lys Glu Ala Leu Glu Glu Val Glu Lys Ile Gln Lys Lys Ser Gln
 100 105 110
 Gln Lys Thr Gln Gln Ala Ala Met Asp Glu Gly Asn Ser Ser Gln Val
 115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His
 130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Met His Pro Gln Gln Ala Gly Pro Ile Pro
 210 215 220
 Pro Gly Gln Ile Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240

 Ser Thr Leu Gln Glu Gln Ile Arg Trp Met Thr Ser Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly

275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
 290 295 300
 Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Gly Trp Met Thr Asp Thr

 305 310 315 320
 Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala
 325 330 335
 Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350
 Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Ala Ser Gly Ala Ala Ala Ala Ile Met Met Gln Lys Ser Asn Phe
 370 375 380

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

 450 455 460
 Gly Phe Gly Glu Glu Ile Ala Pro Ser Pro Lys Gln Glu Gln Lys Glu
 465 470 475 480
 Lys Glu Leu Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Ser Asp
 485 490 495
 Pro

<210> 136

<211> 1494

<212> DNA

<213> Human immunodeficiency virus

<400> 136

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc atgaagcacc tgggtgggc ctcccgcgag 120

ctggagcgtc tcgccctgaa ccccgacctg ctggagaccg ccgagggtcg ccagcagatc 180

atgggccagc tgcagccgc cctgcagacc ggccaccagg agctgcgctc cctgttcaac 240

accgtggcca cctgtactg cgtgcaccag cgcacagagg tgaaggacac caaggaggcc 300

ctggaggagg tggagaagat ccagaagaag tcccagcaga agaccagca ggccgccatg 360

gacgagggca actcctccca ggtgtcccag aactaccca tcgtgcagaa cggccagggc 420

cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtggtggag 480

gagaaggcct tctccccga ggtgateccc atgttctcgg cctgtccga gggcgccacc 540

ccccaggacc tgaacacat gctgaacacc gtgggcggcc accaggccgc catgcagatg 600

ctgaaggaca ccatcaacga ggaggccgc gagtgggacc gcatgcacc ccagcaggcc 660

ggccccatcc cccccggcca gatecgcgag ccccgcggt cgcacatcg cggcaccacc 720

tccacctgc aggagcagat ccgtggatg acctccaacc ccccatccc cgtgggcgag 780

atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgatgta ctccccctg 840

tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc 900

ttcaagacc tgcgcgccga gcaggccacc caggaggtga agggctggat gaccgacacc 960

ctgctggtgc agaagccaa ccccgactgc aagaccatcc tgcgcgcct gggccccggc 1020

gccaccttg aggagatgat gaccgcctgc cagggcgtgg gcggccctc ccacaaggcc 1080

cgctgctgg ccaggccat gtcccaggcc tccggcgcg ccgccccat catgatgcag 1140

aagtccaact tcaaggcccc ccgcgcacc atcaagtgt tcaactgcgg caaggagggc 1200

cacctggccc gcaactgccg cgccccccgc aagaagggt gctggaagt cggcaaggag 1260

ggccaccaga tgaaggactg caccgagcgc caggccaact tcctgggcaa gatctggccc 1320

tccaacaagg gccgccccgg caacttctg cagaaccgcc ccgagccac cgccccccc 1380

gccgagtcct tcggcttcgg cgaggagatc gccccctccc ccaagcagga gcagaaggag 1440

aaggagctgt accccctggc ctccctgaag tcctgttcg gctccgacc ctaa 1494

<210> 137

<211> 499

<212> PRT

<213> Human immunodeficiency virus

<400> 137

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Leu Gln Ile Ile Glu Gln Leu

50 55 60

Gln Pro Ala Ile Lys Thr Gly Thr Glu Glu Leu Gln Ser Leu Phe Asn

65 70 75 80

Thr Val Ala Val Leu Tyr Cys Val His Gln Arg Ile Asp Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Gly Lys Ile Glu Glu Ile Gln Asn Lys Ser Gln

100 105 110

Gln Lys Thr Gln Gln Ala Ala Ala Asp Lys Glu Lys Asp Asn Lys Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His

130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Ala Met Leu Asn Thr Val Gly

180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu

195 200 205

Ala Ala Glu Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro

210 215 220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr

225 230 235 240

Ser Thr Leu Gln Glu Gln Ile Ala Trp Met Thr Gly Asn Pro Pro Ile

245										250					255				
Pro	Val	Gly	Asp	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys				
260					265					270									
Ile	Val	Arg	Met	Tyr	Ser	Pro	Val	Ser	Ile	Leu	Asp	Ile	Lys	Gln	Gly				
275					280					285									
Pro	Lys	Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Phe	Lys	Thr	Leu				
290					295					300									
Arg	Ala	Glu	Gln	Ala	Thr	Gln	Asp	Val	Lys	Asn	Trp	Met	Thr	Asp	Thr				
305					310					315					320				
Leu	Leu	Val	Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Arg	Ala				
325					330					335									
Leu	Gly	Gln	Gly	Ala	Ser	Ile	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly				
340					345					350									
Val	Gly	Gly	Pro	Ser	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser				
355					360					365									
Gln	Val	Thr	Asn	Ala	Asn	Ala	Ala	Ile	Met	Met	Gln	Lys	Gly	Asn	Phe				
370					375					380									
Lys	Gly	Pro	Arg	Lys	Ile	Val	Lys	Cys	Phe	Asn	Cys	Gly	Lys	Glu	Gly				
385					390					395					400				
His	Ile	Ala	Arg	Asn	Cys	Arg	Ala	Pro	Arg	Lys	Lys	Gly	Cys	Trp	Lys				
405					410					415									
Cys	Gly	Arg	Glu	Gly	His	Gln	Met	Lys	Asp	Cys	Thr	Glu	Arg	Gln	Ala				
420					425					430									
Asn	Phe	Leu	Gly	Lys	Ile	Trp	Pro	Ser	Ser	Lys	Gly	Arg	Pro	Gly	Asn				
435					440					445									
Phe	Leu	Gln	Ser	Arg	Pro	Glu	Pro	Thr	Ala	Pro	Pro	Ala	Glu	Ser	Phe				
450					455					460									
Gly	Phe	Gly	Glu	Glu	Met	Thr	Pro	Ser	Pro	Lys	Gln	Glu	Leu	Lys	Asp				
465					470					475					480				
Lys	Glu	Pro	Pro	Leu	Ala	Ser	Leu	Arg	Ser	Leu	Phe	Gly	Asn	Asp	Pro				
485					490					495									

Leu Ser Gln

<210> 138

<211> 1500

<212> DNA

<213> Human immunodeficiency virus

<400> 138

atgggcgccc gcgcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtggc ctcccgcgag 120

ctggagcgct tcgcccgtgaa ccccgccctg ctggagaccg ccgagggctg cctgcagatc 180

atcgagcagc tgcagccgc catcaagacc ggaccgagg agctgcagtc cctgttcaac 240

accgtggccg tgctgtactg cgtgcaccag cgcacgcagc tgaaggacac caaggaggcc 300

ctgggcaaga tcgaggagat ccagaacaag tcccagcaga agaccagca ggccgcccgc 360

gacaaggaga aggacaacaa ggtgtcccag aactaccca tcgtgcagaa cggccagggc 420

cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtgggtggag 480

gagaaggcct tctccccga ggtgatccc atgttctccg ccctgtccga gggcgccacc 540

ccccaggacc tgaacgcat gctgaacacc gtgggcggcc accaggccgc catgcagatg 600

ctgaaggaca ccatcaacga ggaggccgc gagtgggacc gcctgcaccc cgtgcacgcc 660

ggccccatcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720

tccacctgc aggagcagat cgcctggatg accggcaacc cccccatcc cgtgggcgac 780

atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840

tccatcctgg acatcaagca gggccccaag gagcccttc gcgactacgt ggaccgttc 900

ttcaagacc tgcgcgccga gcaggccacc caggacgtga agaactggat gaccgacacc 960

ctgctggtgc agaacgcaa ccccgactgc aagaccatcc tgcgcgcct gggccagggc 1020

gcctccatcg aggagatgat gaccgcctgc cagggcgtgg gcggccctc ccacaaggcc 1080

cgctgtctgg ccgaggccat gtcccagggtg accaacgcca acgcccctat catgatgcag 1140

aagggaact tcaagggcc ccgcaagatc gtgaagtgt tcaactgcgg caaggagggc 1200

cacatcgccc gcaactgccg cgccccccgc aagaagggt gctggaagtg cggccgcgag 1260

ggccaccaga tgaaggactg caccgagcgc caggccaact tcctgggcaa gatctggccc 1320

tcctccaagg gccgccccgg caacttctg cagteccgcc ccgagccac cgccccccc 1380

gccgagtcct tcggcttcgg cgaggagatg accccctccc ccaagcagga gctgaaggac 1440

aaggagcccc cctggcctc cctgcgctcc ctgttcggca acgacccct gtcccagtaa

1500

1500

<210> 139

<211> 495

<212> PRT

<213> Human immunodeficiency virus

<400> 139

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Thr Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Ser Leu Leu Glu Thr Thr Glu Gly Cys Arg Gln Ile Ile Arg Gln Leu

50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Lys Ser Leu Phe Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Val Arg Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Leu Glu Glu Glu Gln Asn Lys Ser Gln

100 105 110

Gln Lys Thr Gln Gln Glu Thr Ala Asp Lys Gly Val Ser Gln Asn Tyr

115 120 125

Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Leu Ser

130 135 140

Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala Phe

145 150 155 160

Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr

165 170 175

Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln Ala

180 185 190

Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu Trp

195	200	205	
Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro Pro Gly Gln Met			
210	215	220	
Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln			
225	230	235	240
Glu Gln Ile Thr Trp Met Thr Ser Asn Pro Pro Val Pro Val Gly Glu			
245	250	255	
Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met			
260	265	270	
Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro			
275	280	285	
Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu Gln			
290	295	300	
Ala Thr Gln Glu Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val Gln			
305	310	315	320
Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Gly			
325	330	335	
Ala Ser Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro			
340	345	350	
Gly His Lys Ala Arg Ile Leu Ala Glu Ala Met Ser Gln Val Thr Asn			
355	360	365	
Thr Ala Val Met Met Gln Arg Gly Asn Phe Lys Gly Gln Arg Lys Ile			
370	375	380	
Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys			
385	390	395	400
Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His			
405	410	415	
Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile			
420	425	430	
Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro			
435	440	445	
Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Phe Gly Glu Glu Ile			

450	455	460	
Thr Pro Ser Pro Arg Gln Glu Thr Lys Asp Lys Glu Gln Gly Pro Pro			
465	470	475	480
Leu Thr Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu Ser Gln			
	485	490	495
<210>	140		
<211>	1488		
<212>	DNA		
<213>	Human immunodeficiency virus		
<400>	140		
atgggcgccc gcgcctccgt gctgtccggc ggcaagctgg acacctggga gaagatccgc			60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgggc ctcccgcgag			120
ctggagcgtc tcgccctgaa cccctccctg ctggagacca ccgagggtg ccgccagatc			180
atccgccagc tgcagccctc cctgcagacc ggctccgagg agctgaagtc cctgttcaac			240
accgtggcca cctgtactg cgtgcaccag cgcacgagg tgcgcgacac caaggaggcc			300
ctggacaagc tggaggagga gcagaacaag tcccagcaga agaccagca ggagaccgcc			360
gacaaggcgc tgtcccagaa ctaccccatc gtgcagaacc tgcagggccca gatggtgcac			420
caggccctgt cccccgcac cctgaacgcc tgggtgaagg tgatcgagga gaaggccttc			480
tccccgagg tgatcccat gttctccgcc ctgtccgagg gcgccacccc ccaggacctg			540
aacaccatgc tgaacacgt gggcggccac caggccgcca tgcagatgct gaaggacacc			600
atcaacgagg aggccgccga gtgggaccgc ctgcacccc tgcacgccg ccccatcccc			660
cccggccaga tgcgcgagcc ccgcggtcc gacatgccg gcaccacctc caccctgcag			720
gagcagatca cctggatgac ctccaacccc cccgtgccg tggcgagat ctacaagcgc			780
tggatcatcc tgggcctgaa caagatcgt cgcacgtact ccccggtgc catcctggac			840
atccgccagg gcccgaagga gcccttcgc gactacgtgg accgtttctt caagacctg			900
cgcgccgagc aggccacca ggaggtgaag aactggatga ccgacacct gctggtgcag			960
aacgccaacc ccgactgaa gaccatcctg aaggccctgg gcccggcgc ctccctggag			1020
gagatgatga ccgcctgcc gggcgtgggc ggccccggc acaaggccc catcctggcc			1080
gaggccatgt ccaggtgac caacaccgcc gtgatgatgc agcgcggcaa cttcaagggc			1140
cagcgcaaga tcatcaagtg cttcaactgc ggcaaggagg gccacatcgc ccgcaactgc			1200
cgcgcccccc gcaagaagg ctgctggaag tgcggcaagg agggccacca gatgaaggac			1260

tgcaccgagc gccaggccaa ctctctgggc aagatctggc cctccaacaa gggccgcccc 1320
 ggcaacttcc tgcagtcccg ccccgagccc accgcccccc cgcgcgagtc cttcggcttc 1380
 ggcgaggaga tcacccctc cccccgccag gagaccaagg acaaggagca gggccccccc 1440

ctgacctccc tgaagtcctt gttcggcaac gacccctgtt cccagtaa 1488

<210> 141

<211> 498

<212> PRT

<213> Human immunodeficiency virus

<400> 141

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Met Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Ile Glu Gln Leu

50 55 60

Gln Ser Thr Leu Lys Thr Gly Ser Glu Glu Leu Lys Ser Leu Phe Asn

65 70 75 80

Thr Val Ala Thr Leu Trp Cys Val His Gln Arg Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Val Gln Asn Lys Ser Gln

100 105 110

Gln Lys Thr Gln Gln Ala Ala Ala Gly Thr Gly Ser Ser Ser Lys Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His

130 135 140

Gln Pro Leu Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160

Glu Lys Gly Phe Asn Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Met Met Leu Asn Ile Val Gly

180	185	190	
Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu			
195	200	205	
Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Pro			
210	215	220	
Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr			
225	230	235	240
Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile			
245	250	255	
Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys			
260	265	270	
Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly			
275	280	285	
Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu			
290	295	300	
Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Glu Thr			
305	310	315	320
Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Ser Ile Leu Lys Ala			
325	330	335	
Leu Gly Thr Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly			
340	345	350	
Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser			
355	360	365	
Gln Ala Gln His Ala Asn Ile Met Met Gln Arg Gly Asn Phe Lys Gly			
370	375	380	
Gln Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Leu Ala			
385	390	395	400
Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys			
405	410	415	
Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu			
420	425	430	

Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe Pro Gln
 435 440 445
 Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Asn Trp Gly Met Gly
 450 455 460
 Glu Glu Ile Thr Ser Leu Pro Lys Gln Glu Gln Lys Asp Lys Glu His
 465 470 475 480

Pro Pro Pro Leu Val Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu
 485 490 495
 Ser Gln

<210> 142

<211> 1497

<212> DNA

<213> Human immunodeficiency virus

<400> 142

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc atgaagcacc tgggtgtgggc ctcccgcgag	120
ctggagcgct tcgcccigaa ccccgccctg ctggagaccg ccgagggctg ccagcagatc	180
atcgagcagc tgcagtccac cctgaagacc ggctccgagg agctgaagtc cctgttcaac	240
accgtggcca cctgtgggtg cgtgcaccag cgcacgagg tgaaggacac caaggaggcc	300
ctggacaaga tcgaggaggt gcagaacaag tcccagcaga agaccagca ggccgcccgc	360
ggcacccggt cctcctcaa ggtgtcccag aactaccca tcgtgcagaa cggccagggc	420
cagatggtgc accagccctt gtccccccgc accctgaacg cctgggtgaa ggtggtggag	480
gagaagggt tcaacccca ggtgatcccc atgtttctcg cctgtccga gggcgccacc	540
ccccaggacc tgaacatgat gctgaacatc gtgggcggcc accaggccgc catgcagatg	600
ctgaaggaga ccatcaacga ggaggccgcc gaggggacc gcgtgcaccc cgtgcacgcc	660
ggccccatcc ccccgccca gatgcgcgag ccccgccggt cgcacatcgc cggcaccacc	720
tccaccctgc aggagcagat cggctggatg accaacaacc ccccatccc cgtgggcgac	780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg	840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc	900
tacaagacc tgcgcgccga gcaggccacc caggaggtga agaactggat gaccgagacc	960
ctgctggtgc agaacgcaa ccccgactgc aagtcaccc tgaaggccct gggcaccggc	1020

gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccctc ccacaaggcc 1080

cgcgtgctgg ccgaggccat gtcccaggcc cagcacgcca acatcatgat gcagcgcggc 1140

aacttcaagg gccagaagcg catcaagtgc ttcaactgcg gcaaggaggg ccacctggcc 1200

cgcaactgcc gcgcccccg caagaagggc tgctggaagt gcggaagga gggccaccag 1260

atgaaggact gcaccgagcg ccaggccaac ttctgggca agatctggcc ctccaacaag 1320

ggcgcgcccc gcaacttccc ccagtccgc cccgagccca ccgcccccc cgccgagaac 1380

tggggcatgg gcgaggagat caccitccctg cccaagcagg agcagaagga caaggagcac 1440

ccccccccc tgggtgccct gaagtcctg ttcggcaacg accccctgtc ccagtaa 1497

<210> 143

<211> 492

<212> PRT

<213> Human immunodeficiency virus

<400> 143

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Met Glu Gln Leu

50 55 60

Gln Ser Ala Leu Arg Thr Gly Ser Glu Glu Leu Lys Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Trp Cys Val His Gln Arg Ile Asp Ile Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Val Gln Asn Lys Ser Lys

100 105 110

Gln Lys Thr Gln Gln Ala Ala Ala Thr Gly Ser Ser Ser Gln Asn

115 120 125

Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Thr His Gln Ser Met

130 135 140

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

Lys Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
405 410 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
420 425 430

Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn Phe Pro Gln Ser Arg Pro
435 440 445

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Met Gly Glu Glu Ile
450 455 460

Thr Ser Ser Pro Lys Gln Glu Pro Arg Asp Lys Gly Leu Tyr Pro Pro
465 470 475 480

Leu Thr Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro

485 490

<210> 144

<211> 1479

<212> DNA

<213> Human immunodeficiency virus

<400> 144

atgggcgccc gcgcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtggc ctcccgcgag 120
ctggagcgct tcgcccgtgaa ccccgccctg ctggagaccg ccgagggctg ccagcagatc 180
atggagcagc tgcagtcgc cctgcgcacc ggctccgagg agctgaagtc cctgtacaac 240
accgtggcca cctgtgtgtg cgtgcaccag cgcacgcaca tcaaggacac caaggaggcc 300

ctggacaaga tcgaggaggt gcagaacaag tccaagcaga agaccagca ggccgcgcgc 360
gccaccggct cctcctccca gaactacccc atcgtgcaga acgcccagg ccagatgacc 420
caccagtcca tgtcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc 480
ttctccccg aggtgatccc catgttctcc gcctgtccg agggcgccac ccccaggac 540
ctgaacatga tctgaacat cgtgggcggc caccaggccg ccatgcagat gctgaaggac 600
accatcaacg aggaggccgc cgagtgggac cgcgtgcacc ccgtgcacgc cggccccatc 660
cccccgcc agatgcgcga gccccgccc tccgacatcg ccggcaccac ctccaccctg 720

caggagcaga tcggctggat gacctccaac ccccccatcc ccgtgggcga gatctacaag 780
cgctggatcg tctgtggcct gaacaagatc gtgcgcatgt actccccgt gtccatcctg 840

gacatccgcc agggcccca ggagcccttc cgcgactacg tggaccgctt cttcaagacc 900
ctgcgcgccg agcaggccac ccaggagggtg aagaactgga tgaccgagac cctgctggtg 960
cagaacgcca accccgactg caagtccatc ctgcgcgcc tgggccccgg cgccaccctg 1020
gaggagatga tgaccgctg ccagggcgtg ggcggccccg gccacaaggc ccgcgtgctg 1080
gccgaggcca tgtcccaggt gcagcagtcc aacatcatga tgcagcgagg caacttccgc 1140

ggccagcgca ccatcaagtg cttcaactgc ggcaaggagg gccacctggc ccgcaactgc 1200
aaggccccc gcaagaaggg ctgctggaag tgcggcaagg agggccacca gatgaaggac 1260
tgaccgagc gccaggccaa cttcctgggc aagatctggc cctcctcaa gggccgcccc 1320
ggcaacttcc ccagtcctcg ccccgagccc accgcccccc ccgcccagtc cttcgcatg 1380
ggcgaggaga tcacctctc cccaagcag gagccccgcg acaagggcct gtaccccccc 1440
ctgacctccc tgaagtccct gttcggcaac gacccctaa 1479

<210> 145

<211> 498

<212> PRT

<213> Human immunodeficiency virus

<400> 145

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Ile Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Ser Leu Leu Glu Thr Ser Glu Gly Cys Gln Gln Ile Leu Glu Gln Leu

50 55 60

Gln Pro Thr Leu Lys Thr Gly Ser Glu Glu Leu Lys Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Ile Gln Asn Lys Ser Lys

100 105 110

Gln Lys Thr Gln Gln Ala Ala Thr Gly Thr Gly Ser Ser Ser Lys Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Thr His

130

135

140

Gln Ser Met Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu

145

150

155

160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165

170

175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Met Met Leu Asn Ile Val Gly

180

185

190

Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu

195

200

205

Ala Ala Glu Trp Asp Arg Leu His Pro Ala Gln Ala Gly Pro Phe Pro

210

215

220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr

225

230

235

240

Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Ser Asn Pro Pro Ile

245

250

255

Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys

260

265

270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly

275

280

285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu

290

295

300

Arg Ala Glu Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Glu Thr

305

310

315

320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala

325

330

335

Leu Gly Ser Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly

340

345

350

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser

355

360

365

Gln Val Gln Asn Ala Asn Ile Met Met Gln Lys Ser Asn Phe Arg Gly

370 375 380
 Pro Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Asp Gly His Leu Ala
 385 390 395 400
 Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys
 405 410 415
 Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu

 420 425 430
 Gly Arg Ile Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn Phe Pro Gln
 435 440 445
 Ser Arg Pro Glu Pro Ser Ala Pro Pro Ala Glu Asn Phe Gly Met Gly
 450 455 460
 Glu Glu Ile Thr Pro Ser Leu Lys Gln Glu Gln Lys Asp Arg Glu Gln
 465 470 475 480
 His Pro Pro Ser Ile Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu
 485 490 495

Ser Gln

<210> 146

<211> 1497

<212> DNA

<213> Human immunodeficiency virus

<400> 146

atgggcgccc gcgctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc atcaagcacc tgggtgggc ctcccgcgag	120
ctggagcgct tcgcctgaa cccctccctg ctggagacct ccgagggtg ccagcagatc	180
ctggagcagc tgcagccac cctgaagacc ggctccgagg agctgaagtc cctgtacaac	240
accgtggcca cctgtactg cgtgcaccag cgcacgaga tcaaggacac caaggaggcc	300
ctggacaaga tcgaggagat ccagaacaag tccaagcaga agaccagca ggccgccacc	360
ggcaccggct cctcctcaa ggtgtcccag aactaccca tcgtgcagaa cgcccagggc	420
cagatgacc accagtccat gtccccccgc acctgaacg cctgggtgaa ggtgatcgag	480
gagaaggcct tctccccga ggtgatcccc atgttctccg ccctgtccga gggcgccacc	540
ccccaggacc tgaacatgat gctgaacatc gtggcgggcc accaggccgc catgcagatg	600

ctgaaggaca ccatcaacga ggaggccgcc gattgggacc gcctgcaccc cggccaggcc 660
ggccccttcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720

tccaccctgc aggagcagat cggctggatg acctccaacc ccccatccc cgtgggcgac 780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840
tccatcctgg acatccgcca gggccccaag gagcccttcc gcgactacgt ggaccgcttc 900
ttcaagacc tgcgcgcga gcagccacc caggacgtga agaactggat gaccgagacc 960
ctgctggtgc agaacgcaa ccccgactgc aagaccatcc tgcgcgcct gggtccggc 1020
gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc 1080
cgcgtgctgg ccgaggccat gtcccagggtg cagaacgcca acatcatgat gcagaagtcc 1140

aacttccgcg gcccgaagcg catcaagtgc ttcaactgcg gcaaggacgg ccacctggcc 1200
cgcaactgcc gcgcccccg caagaaggcg tcttggaagt gcggcaagga gggccaccag 1260
atgaaggact gcaccgagcg ccaggccaac ttctgggcc gcacttgcc ctctccaag 1320
ggccgccccg gcaacttccc ccagtccgc cccgagcct ccgcccccc cgcgagaaac 1380
ttcgcatgg gcgaggagat cacccttcc ctgaagcagg agcagaagga ccgcgagcag 1440
ccccccct ccacttccct gaagtccctg ttgggcaacg acccctgtc ccagtaa 1497

<210> 147

<211> 500

<212> PRT

<213> Human immunodeficiency virus

<400> 147

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Arg Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Leu Met Glu Gln Leu

50 55 60

Gln Ser Thr Leu Lys Thr Gly Ser Glu Glu Leu Lys Ser Leu Phe Asn

65 70 75 80

Thr Ile Ala Thr Leu Trp Cys Val His Gln Arg Ile Asp Val Lys Asp

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

340 345 350

Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Ala Ser Asn Ala Ala Ala Ala Ile Met Met Gln Lys Ser Asn Phe
370 375 380

Lys Gly Gln Arg Arg Ile Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly
385 390 395 400

His Leu Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys
405 410 415

Cys Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala
420 425 430

Asn Phe Leu Gly Arg Met Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn
435 440 445

Phe Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Leu
450 455 460

Glu Met Lys Glu Glu Thr Thr Ser Ser Pro Lys Gln Glu Pro Arg Asp
465 470 475 480

Lys Glu Leu Tyr Pro Leu Thr Ser Leu Lys Ser Leu Phe Gly Ser Asp
485 490 495

Pro Leu Ser Gln
500

<210> 148
<211> 1503
<212> DNA
<213> Human immunodeficiency virus
<400> 148

atggggcgccc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gcgcacccgc 60
ctgcgccccg gcgcaagaa gaagtaccgc ctgaagcacc tgggtgtgggc ctcccgcgag 120
ctggagcgt tgcacctgaa ccccgccctg ctggagaccg ccgagggtg ccagcagctg 180
atggagcagc tgcagtccac cctgaagacc ggtccgagg agctgaagtc cctgttcaac 240
accatcgcca ccctgtggtg cgtgcaccag cgcacgcagc tgaaggacac caaggaggcc 300

ctggacaagg tggaggagat gcagaacaag tccaagcaga agacccagca ggccgcccgc 360
gacaccggcg gtcctccaa cgtgtcccag aactaccca tcgtgcagaa cggccagggc 420
cagatggtgc accagtccat cccccccgc accctgaacg cctgggtgaa ggtgatcgag 480
gagaaggcct tccccccga ggtgatcccc atgttctccg cctgtccga gggcgccacc 540
ccccaggacc tgaacatgat gctgaacatc gtgggcggcc accaggccgc catgcagatg 600
ctgaaggaca ccatcaacga ggaggccgcc gagtgggacc gcgccaccc cgtgcacgcc 660
ggcccatcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720

tccacctgc aggagcagat cggctggatg acctccaacc ccccatccc cgtgggcgag 780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta cccccctg 840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgcttc 900
ttcaagtgcc tgcgcgccga gcagccacc caggaggtga agaactggat gaccgagacc 960
ctgctggtgc agaagccaa ccccgactgc aagtccatcc tgaaggcctt gggcaccggc 1020
gccacctgg aggagatgat gaccgcctgc cagggcgtgg gcggcccctc ccacaaggcc 1080
cgcgtgctgg ccgaggccat gtcccaggcc tccaacgccg ccgccccat catgatgcag 1140

aagtccaact tcaaggcca gcgcccatc atcaagtgt tcaactgcgg caaggagggc 1200
cacctggccc gcaactgccg cgccccccgc aagaagggt gctggaagt cggaaggag 1260
ggccaccaga tgaaggactg caccgagcgc caggccaact tcctgggccc catgtggccc 1320
tcctccaagg gccgccccgc caacttctg cagtcccgc ccgagccac cgccccccc 1380
gccgagtccc tggagatgaa ggaggagacc acctctccc ccaagcagga gccccgcgac 1440
aaggagctgt accccctgac ctccctgaag tcctgttcg gctccgacc cctgtcccag 1500
taa 1503

<210> 149

<400> 149

000

<210> 150

<211> 497

<212> PRT

<213> Human immunodeficiency virus

<400> 150

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Glu Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30
 His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
 35 40 45
 Gly Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Ile Glu Gln Leu

 50 55 60
 Gln Ser Ala Leu Lys Thr Gly Ser Glu Glu Leu Lys Ser Leu Tyr Asn
 65 70 75 80
 Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Lys Val Thr Asp
 85 90 95
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Ile Gln Asn Lys Ser Lys
 100 105 110
 Gln Lys Ala Gln Gln Ala Ala Ala Ala Thr Gly Asn Ser Ser Asn Leu
 115 120 125

 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His
 130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu
 145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Met Met Leu Asn Ile Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu

 195 200 205
 Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Pro
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Ser Asn Pro Pro Ile
 245 250 255
 Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
275 280 285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
290 295 300

Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Asp Thr
305 310 315 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
325 330 335

Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
340 345 350

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Ala Ser Gly Thr Glu Ala Ala Ile Met Met Gln Lys Ser Asn Phe
370 375 380

Lys Gly Pro Lys Arg Ser Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly
385 390 395 400

His Leu Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys
405 410 415

Cys Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala
420 425 430

Asn Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn
435 440 445

Phe Leu Gln Asn Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe
450 455 460

Gly Phe Gly Glu Glu Thr Ala Pro Ser Pro Lys Gln Glu Pro Lys Glu
465 470 475 480

Lys Glu Leu Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp
485 490 495

Pro

<210> 151

<211> 1494

<212> DNA

<213> Human immunodeficiency virus

<400> 151

atgggcgccc ggcctccgt gctgtccggc ggcaagctgg acgagtggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtgggc ctcccgcgag	120
ctggagcgct tcgccctgaa ccccgccctg ctggagaccg ccgagggtg ccagcagatc	180
atcgagcagc tgcagtccgc cctgaagacc ggctccgagg agctgaagtc cctgtacaac	240
accgtggcca cctgtactg cgtgcaccag cgcataagg tgaccgacac caaggaggcc	300
ctggacaaga tcgaggagat ccagaacaag tccaagcaga aggccagca ggccgcgcgc	360
gccaccggca actcctccaa cctgtcccag aactaccca tcgtgcagaa cgcccagggc	420
cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtgatcgag	480
gagaaggcct tctccccga ggtgatcccc atgttctccg cctgtccga gggcgccacc	540
ccccaggacc tgaacatgat gctgaacatc gtggcgggcc accaggccgc catgcagatg	600
ctgaaggaca ccatcaacga ggaggccgcc gagtgggacc gcgtgcaccc cgtgcacgcc	660
ggccccatcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc	720
tccaccctgc aggagcagat cggttgatg acctccaacc ccccatccc cgtgggcgag	780
atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg	840
tccatcctgg acatccgcca gggccccaag gagcccttc gcgactacgt ggaccgttc	900
ttcaagacc tgcgcgccga gcaggccacc caggaggtga agaactggat gaccgacacc	960
ctgctggtgc agaagccaa ccccgactgc aagaccatcc tgaaggccct ggcccccggc	1020
gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc	1080
cgcgtgctgg ccgaggccat gtcccaggcc tccggcaccg agcccgccat catgatgcag	1140
aagtccaact tcaagggcc caagcgctcc atcaagtgt tcaactgcgg caaggagggc	1200
cacctggccc gcaactgcgc gcccccccgc aagaagggt gctggaagt cggaaggag	1260
ggccaccaga tgaaggactg caccgagcgc caggccaact tcctgggcaa gatctggccc	1320
tccaacaagg gccgccccgg caacttctg cagaaccgcc ccgagccac gcccccccc	1380
gccgagtct tcggttcgg cgaggagacc gccccctccc ccaagcagga gccaaggag	1440
aaggagctgt accccctggc ctccctgaag tcctgttcg gcaacgacc ctaa	1494

<210> 152

<211> 495

<212> PRT

<213> Human immunodeficiency virus

<400> 152

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Thr Glu Ile Asp Val Arg Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ile Gln

100 105 110

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn

115 120 125

Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Ile

130 135 140

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala

145 150 155 160

Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala

165 170 175

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln

180 185 190

Ala Ala Met Gln Ile Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu

195 200 205

Trp Asp Arg Leu His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln

210 215 220

Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Asn Leu

225 230 235 240

Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Val Pro Val Gly
245 250 255

Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
260 265 270

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
275 280 285

Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu
290 295 300

Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
305 310 315 320

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
325 330 335

Gly Ala Ser Ile Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
340 345 350

Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn
355 360 365

Ser Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile
370 375 380

Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys
385 390 395 400

Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
405 410 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
420 425 430

Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro
435 440 445

Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg Phe Gly Glu Glu Thr
450 455 460

Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp Lys Glu Leu Tyr Pro
465 470 475 480

Leu Thr Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro Ser Ser Gln

485

490

495

<210> 153

<211> 1488

<212> DNA

<213> Human immunodeficiency virus

<400> 153

atgggcgccc gcgcctccat cctgcgcggc ggcaagctgg acaagtggga gaagatccgc 60
ctgcgccccg gcggcaagaa gcactacatg ctgaagcacc tgggtgtgggc ctcccgcgag 120
ctggagcgct tcgcccgtgaa ccccgccctg ctggagacct ccgagggctg caagcagatc 180
atcaagcagc tgcagccccg cctgcagacc ggccaccgagg agctgcgctc cctgttcaac 240
accgtggcca cctgtactg cgtgcacacc gagatcgacg tgcgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag atccagcaga agaccagca ggccaaggag 360

gccgacggca aggtgtccca gaactacccc atcgtgcaga acctgcaggg ccagatggtg 420
caccagccca ttccccccg caccctgaac gcctgggtga aggtgggtgga ggagaaggcc 480
tttcccccg aggtgatccc catgtttctc gcctgttccg agggcgccac ccccaggac 540
ctgaacacca tgctgaacac cgtgggcggc caccaggccg ccatgcagat cctgaaggac 600
accatcaacg aggaggccgc cgagtgggac cgcctgcacc ccgtgcacgc cggccccatc 660
gcccccgccc agatgcgcga gcccccggc tccgacatcg ccggcaccac ctccaacctg 720
caggagcaga tcgcctggat gacctccaac cccccgtgc ccgtgggcga catctacaag 780

cgctggatca tcttgggcct gaacaagatc gtgcgcatgt actccccac ctccatcctg 840
gacatcaagc agggccccc aaaggccctt cgcgactacg tggaccgctt cttcaagacc 900
ctgcgcgccc agcaggccac ccaggacctg aagaactgga tgaccgacac cctgctggtg 960
cagaacgcca accccgactg caagaccatc ctgcgcgccc tgggccccgg cgctccatc 1020
gaggagatga tgaccgctg ccaggggcgtg ggcgggccct cccacaagc ccgcgtgctg 1080
gccgaggcca tgtcccagac caactccacc atcctgatgc agcgtccaa cttcaagggc 1140
tccaagcgca tcgtgaagtg cttcaactgc ggcaaggagg gccacatcg ccgcaactgc 1200

cgcgcccc ccaagaaggc ctgctggaag tgcggcaagg agggccacca gatgaaggac 1260
tgaccagagc gccaggccaa cttcttgggc aagatctggc cctcccacaa ggcccgcccc 1320
ggcaacttcc tgcagtcccg ccccgagccc accgcccc ccgaggagtc cttccgttc 1380
ggcgaggaga ccaccacccc ctcccagaag caggagccca tcgacaagga gctgtacccc 1440
ctgacctccc tgaagtcctt gttcggcaac gaccttcct cccagtaa 1488

<210> 154
 <211> 491
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 154

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp

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

Thr Val Ala Thr Leu Tyr Cys Val His Ala Glu Ile Glu Val Arg Asp			
	85	90	95
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ile Gln			
	100	105	110
Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Glu Lys Val Ser Gln Asn			
	115	120	125
Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Leu			
	130	135	140
Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala			

145	150	155	160
Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala			
	165	170	175
Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln			
	180	185	190
Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu			
	195	200	205
Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala Pro Gly Gln			

210 215 220
 Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu
 225 230 235 240
 Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly
 245 250 255
 Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
 260 265 270
 Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
 275 280 285
 Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu

 290 295 300
 Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
 305 310 315 320
 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
 325 330 335
 Gly Ala Ser Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
 340 345 350
 Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn
 355 360 365

 Asn Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile
 370 375 380
 Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Lys Asn Cys
 385 390 395 400
 Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
 405 410 415
 Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
 420 425 430
 Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro

 435 440 445
 Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr
 450 455 460

Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu

465 470 475 480

Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln

485 490

<210> 155

<211> 1476

<212> DNA

<213> Human immunodeficiency virus

<400> 155

atgggcgccc gcgcctccat cctgcgcggc ggcaagctgg acaagtggga gaagatccgc 60

ctgcgccccg gcggcaagaa gcactacatg ctgaagcacc tgggtgtggc ctcccgcgag 120

ctggagcgct tcgccctgaa ccccgccctg ctggagacct ccgagggctg caagcagatc 180

atcaagcagc tgcagccgc cctgcagacc ggcaaccgagg agctgcgctc cctgttcaac 240

accgtggcca cctgtactg cgtgcacgcc gagatcgagg tgcgcgacac caaggaggcc 300

ctggacaaga tcgaggagga gcagaacaag atccagcaga agaccagca ggccaaggag 360

gccgacgaga aggtgtccca gaactaccc atcgtgcaga acctgcaggg ccagatggtg 420

caccagcccc tgtcccccg caccctgaac gcctgggtga aggtggtgga ggagaaggcc 480

ttctccccg aggtgatccc catgttcacc gccctgtccg agggcgccac cccccaggac 540

ctgaacacca tgctgaacac cgtgggcggc caccaggccg ccatgcagat gctgaaggac 600

accatcaacg aggaggccgc cgagtgggac cgcctgcacc ccgtgcacgc cggccccgtg 660

gcccccgcc agatgcgcga gcccccgcc tccgacatcg ccggcaccac ctccaccctg 720

caggagcaga tcggctggat gaccaacaac cccccatcc ccgtgggcga gatctacaag 780

cgctggatca tctgggcct gaacaagatc gtgcgcatgt actccccac ctccatcctg 840

gacatcaagc agggcccca ggagccctc cgcgactacg tggaccgctt cttcaagacc 900

ctgcgcgccc agcaggccac ccaggacgtg aagaactgga tgaccgacac cctgctggtg 960

cagaacgcca accccgactg caagaccatc ctgcgcgccc tgggccccgg cgcctcctg 1020

gaggagatga tgaccgctg ccaggcgctg ggcgccccct cccacaaggc ccgctgctg 1080

gccgaggcca tgtcccagac caacaacacc atcctgatgc agcgtccaa cttcaagggc 1140

tccaagcgca tcgtgaagtg cttcaactgc ggcaaggagg gccacatcgc caagaactgc 1200

cgcgcccc gcaagaagg ctgctggaag tgcggcaagg agggccacca gatgaaggac 1260

tgcaccgagc gccaggccaa cttcctgggc aagatctggc cctcccacaa ggccgcccc 1320

ggcaacttcc tgcagtcccg ccccgagccc accgcccccc ccgcccagtc ctcccgcttc 1380
gaggagacca ccccgcccc caagcaggag cccaaggacc gcgagcccct gacctccctg 1440
cgctccctgt tcggtccga cccctgtcc cagtaa 1476

<210> 156

<211> 499

<212> PRT

<213> Human immunodeficiency virus

<400> 156

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Glu Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Gly Gln Leu

50 55 60

Gln Pro Ala Ile Gln Thr Gly Ser Glu Glu Ile Lys Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Glu Arg Ile Lys Val Thr Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Thr Lys Ser Lys

100 105 110

Lys Lys Ala Gln Gln Ala Thr Ala Asp Thr Gly Asn Ser Ser Gln Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His

130 135 140

Gln Pro Leu Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu

145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly

180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
195 200 205

Ala Ala Glu Trp Asp Arg Leu His Pro Val Gln Ala Gly Pro Val Ala
210 215 220

Pro Gly Gln Ile Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
225 230 235 240

Ser Thr Leu Gln Glu Gln Ile Arg Trp Met Thr Ser Asn Pro Pro Ile
245 250 255

Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
260 265 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
275 280 285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
290 295 300

Arg Ala Glu Gln Ala Ser Gln Asp Val Lys Asn Trp Met Thr Glu Thr
305 310 315 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
325 330 335

Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
340 345 350

Val Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
355 360 365

Gln Ala Thr Ser Gly Asn Ala Ile Met Met Gln Arg Gly Asn Phe Lys
370 375 380

Gly Pro Lys Lys Ile Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His
385 390 395 400

Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
405 410 415

Gly Arg Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
420 425 430

Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe
435 440 445

Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly

450 455 460
Phe Gly Glu Glu Ile Thr Pro Ser Gln Lys Gln Glu Gln Lys Asp Lys
465 470 475 480
Glu Leu His Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro
485 490 495
Leu Ser Gln

<210> 157
<211> 1500
<212> DNA
<213> Human immunodeficiency virus
<400> 157

atgggcgcgc gcgcctccgt gctgtccggc ggcaagctgg acgagtggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tgggtgtgggc ctcccgcgag 120

ctggagcgct tcgccctgaa ccccgccctg ctggagacct ccgagggtctg caagcagatc 180

atcgcccgagc tgcagccgc catccagacc ggctccgagg agatcaagtc cctgtacaac 240

accgtggcca cctgtactg cgtgcacgag cgcataagg tgaccgacac caaggaggcc 300

ctggacaaga tcgaggagga gcagaccaag tccaagaaga agggccagca ggccaccgcc 360

gacaccggca actcctccca ggtgtcccag aactaccca tcgtgcagaa cctgcagggc 420

cagatggtgc accagcccct gtccccccgc accctgaacg cctgggtgaa ggtgatcgag 480

gagaaggcct tctccccga ggtgatcccc atgttctccg cctgtccga gggcgccacc 540

ccccaggacc tgaacacat gctgaacacc gtgggcggcc accaggccgc catgcagatg 600

ctgaaggaga ccatcaacca ggagccgcc gagtgggacc gcctgcaccc cgtgcaggcc 660

ggccccgtgg cccccggcca gatccgcgag ccccgccgct ccgacatcg cgccaccacc 720

tccacctgc aggagcagat ccgctggatg acctcaacc ccccatccc cgtgggcgag 780

atctacaagc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840

tccatcctgg acatccgcca gggcccgaag gagcccttc gcgactacgt ggaccgttc 900

tacaagacc tcgcgcccga gcaggcctcc caggacgtga agaactggat gaccgagacc 960

ctgctggtgc agaagccaa ccccgactgc aagacatcc tgaaggccct gggccccgcc 1020

gccaccctgg aggagatgat gaccgcctgc cagggcgtgg gcggcccctc ccacaaggcc 1080

cgcgtgctgg ccgaggccat gtcccaggcc acctccggca acgccatcat gatgcagcgc 1140
 ggcaacttca agggcccca gaagatcatc aagtgttca actgcggcaa ggagggccac 1200
 atcgccaaga actgccgcgc cccccgaag aagggtgct ggaagtgcgg ccgcgagggc 1260
 caccagatga aggactgcac cgagcgccag gccaacttcc tgggcaagat ctggccctcc 1320

aacaagggcc gccccggcaa ctctctgcag tccgccccg agcccaccgc ccccccgcc 1380
 gagtccttcg gcttcggcga ggagatcacc cctcccaga agcaggagca gaaggacaag 1440
 gagctgcacc cctggcctc cctgaagtcc ctgttcggca acgaccctt gtcccagtaa 1500
 1500

<210> 158
 <211> 497
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 158

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp
 1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys
 20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
 35 40 45

Ser Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Met Gly Gln Leu
 50 55 60

Gln Pro Ala Leu Gly Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn
 65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His His Arg Ile Glu Val Lys Asp

85 90 95
 Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Ile Gln Asn Lys Ser Lys

100 105 110
 Gln Lys Lys Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Lys Val

115 120 125
 Ser Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His

130 135 140
 Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160
 Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
 165 170 175
 Glu Gly Ala Thr Pro Gln Asp Leu Asn Met Met Leu Asn Ile Val Gly
 180 185 190
 Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu
 195 200 205
 Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Pro
 210 215 220
 Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr

 225 230 235 240
 Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Gly Asn Pro Pro Val
 245 250 255
 Pro Val Gly Glu Ile Tyr Arg Arg Trp Ile Ile Leu Gly Leu Asn Lys
 260 265 270
 Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
 275 280 285
 Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu
 290 295 300

 Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Ser Trp Met Thr Glu Thr
 305 310 315 320
 Leu Leu Ile Gln Asn Ala Asn Pro Asp Cys Lys Ser Ile Leu Arg Ala
 325 330 335
 Leu Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
 340 345 350
 Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
 355 360 365
 Gln Val Gln Gln Thr Asn Ile Met Met Gln Arg Ser Asn Phe Lys Gly

 370 375 380
 Gln Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Leu Ala
 385 390 395 400

Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys

405 410 415

Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu

420 425 430

Gly Lys Ile Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn Phe Leu Gln

435 440 445

Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Phe Gly

450 455 460

Glu Glu Ile Ala Pro Ser Pro Lys Gln Glu Pro Lys Glu Lys Glu Leu

465 470 475 480

Tyr Pro Leu Thr Ser Leu Lys Ser Leu Phe Gly Ser Asp Pro Leu Ser

485 490 495

Gln

<210> 159

<211> 1494

<212> DNA

<213> Human immunodeficiency virus

<400> 159

atgggcgccc gcgcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcacc tggtgtgggc ctcccgcgag 120

ctggagcgct tcgccctgaa cccctccctg ctggagaccg ccgagggctg ccagcagatc 180

atgggccagc tgcagcccg cctgggcacc ggccaccagg agctgcgctc cctgtacaac 240

accgtggcca cctgtactg cgtgcaccac cgcacagagg tgaaggacac caaggaggcc 300

ctggacaaga tcgaggagat ccagaacaag tccaagcaga agaagcagca ggccgcccgc 360

gacaccggca actcctccaa ggtgtcccag aactaccca tcgtgcagaa cgcccagggc 420

cagatggtgc accaggccat ctccccccgc accctgaacg cctgggtgaa ggtggtggag 480

gagaaggcct tctccccga ggtgatcccc atgttctccg cctgtccga gggcgccacc 540

ccccaggacc tgaacatgat gctgaacatc gtggcgggcc accaggccgc catgcagatg 600

ctgaaggaca ccatcaacga ggaggccgcc gagtgggacc gcgtgcaccc cgtgcacgcc 660

ggccccatcc cccccggcca gatgcgcgag ccccgcggt cgcacatcgc cggcaccacc 720

tccaccctgc aggagcagat cggctggatg accggcaacc cccccgtgcc cgtgggcgag 780

atctaccgcc gctggatcat cctgggcctg aacaagatcg tgcgcatgta ctccccctg 840
tccatcctgg acatccgcca gggccccaag gagcccttcc gcgactacgt ggaccgcttc 900

ttcaagacc tgcgcgcga gcagccacc caggaggtga agtcctggat gaccgagacc 960
ctgctgatcc agaacgcaa ccccgactgc aagtccatcc tgcgcgcctt gggccccggc 1020
gccaccttg aggagatgat gaccgcctgc cagggcgtgg gcggccccgg ccacaaggcc 1080
cgcgtgctgg ccgaggccat gtcccagggtg cagcagacca acatcatgat gcagcgctcc 1140
aacttcaagg gccagaagcg catcaagtgc ttcaactgcg gcaaggaggg ccacctggcc 1200
cgcaactgcc gcgcccccg caagaaggcg tgctggaagt gcggcaagga gggccaccag 1260
atgaaggact gcaccgagcg ccaggccaac ttcctgggca agatctggcc ctctccaag 1320

ggccgcccc gcaacttct gcagtccgc cccgagccca ccgcccccc cgcagagtc 1380
ttcggttcc gcgaggatg cgcctctcc cccaagcagg agcccaagga gaaggagctg 1440
tacccttga cctccctgaa gtccctgttc ggctccgacc cctgtccca gtaa 1494

<210> 160

<211> 492

<212> PRT

<213> Human immunodeficiency virus

<400> 160

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Lys Ile Ile Gly Gln Leu

50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Ile Ala Val Leu Tyr Phe Val His Gln Lys Val Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Asp Lys Leu Glu Glu Glu Gln Asn Lys Ser Gln

100 105 110

Gln Lys Thr Gln Gln Ala Ala Ala Asp Lys Gly Val Ser Gln Asn Tyr
115 120 125

Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Leu Ser
130 135 140

Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala Phe
145 150 155 160

Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr
165 170 175

Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln Ala
180 185 190

Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu Trp
195 200 205

Asp Arg Leu His Pro Val His Ala Gly Pro Ile Pro Pro Gly Gln Met
210 215 220

Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu Gln
225 230 235 240

Glu Gln Ile Gln Trp Met Thr Ser Asn Pro Pro Val Pro Val Gly Glu
245 250 255

Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met
260 265 270

Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro
275 280 285

Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu Gln
290 295 300

Ala Thr Gln Glu Val Lys Gly Trp Met Thr Asp Thr Leu Leu Val Gln
305 310 315 320

Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro Gly
325 330 335

Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro
340 345 350

Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Val Thr Asn
355 360 365

Thr Thr Val Met Met Gln Lys Ser Asn Phe Lys Gly Gln Arg Arg Ile
370 375 380

Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Lys Asn Cys
385 390 395 400

Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Arg Glu Gly His
405 410 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
420 425 430

Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe Leu Gln Asn Arg Pro
435 440 445

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Gly Phe Gly Glu Glu Ile

450 455 460

Thr Pro Ser Pro Lys Gln Glu Gln Lys Asp Glu Gly Leu Tyr Pro Pro
465 470 475 480

Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp Pro

485 490

<210> 161

<211> 1479

<212> DNA

<213> Human immunodeficiency virus

<400> 161

atgggcgccc gcgcctccgt gctgtccggc ggcgagctgg accgctggga gaagatccgc 60

ctgcgccccg gcggcaagaa gaagtaccgc ctgaagcaca tcgtgtgggc ctcccgcgag 120

ctggagcgct tcgccgtgaa ccccgccctg ctggagacct ccgagggtcg ccgcaagatc 180

atcgcccgagc tgcagccctc cctgcagacc ggctccgagg agctgcgctc cctgtacaac 240

accatcgccg tgctgtactt cgtgcaccag aaggtggagg tgaaggacac caaggaggcc 300

ctggacaagc tggaggagga gcagaacaag tcccagcaga agaccagca ggccgcccgc 360

gacaagggcg tgtcccagaa ctaccccatc gtgcagaacc tgcagggcca gatggtgcac 420

caggccctgt cccccgcac cctgaacgcc tgggtgaagg tgggtggagga gaaggccttc 480

tccccgagg tgatcccat gtctctcgcc ctgtccgagg gcgccacccc ccaggacctg 540

aacaccatgc tgaacaccgt gggcggccac caggccgcca tgcagatgct gaaggacacc 600

atcaacgagg aggccgccga gtgggaccgc ctgcaccccg tgcacgccgg ccccatcccc 660
 cccggccaga tgcgcgagcc ccgcggctcc gacatcgccg gcaccacctc caccctgcag 720
 gagcagatcc agtggatgac ctccaacccc cccgtgcccc tgggcgagat ctacaagcgc 780
 tggatcatcc tgggcctgaa caagatctgt cgcattgtact cccccgtgtc catcctggac 840
 atccgccagg gcccgaagga gcccttcgcg gactacgtgg accgcttctt caagaccctg 900
 cgcgccgagc aggccaccca ggaggtgaag ggctggatga ccgacaccct gctgggtgcag 960

aacgccaacc ccgactgcaa gaccatcctg aaggccctgg gccccggcgc caccctggag 1020
 gagatgatga ccgcctgcca gggcgtgggc ggccccggcc acaaggcccg cgtgctggcc 1080
 gaggccatgt cccaggtgac caacaccacc gtgatgatgc agaagtcaa cttcaagggc 1140
 cagcgccgca tcgtgaagtg cttcaactgc ggcaaggagg gccacatcgc caagaactgc 1200
 cgcgcccccc gcaagaaggg ctgctggaag tgcggccgcg agggccacca gatgaaggac 1260
 tgcaccgagc gccaggccaa cttcctgggc aagatctggc cctccaacaa ggcccgcccc 1320
 ggcaacttcc tgcagaaccg ccccgagccc accgcccccc ccgccgagtc cttcggttc 1380

ggcgaggaga tcacccctc cccaagcag gagcagaagg acgagggcct gtacccccc 1440
 ctggcctccc tgaagtcctt gttcggcaac gaccctaa 1479

<210> 162

<211> 496

<212> PRT

<213> Human immunodeficiency virus

<400> 162

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Lys Leu Asp Ala Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Met Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Asp Leu Leu Glu Thr Ala Glu Gly Cys Gln Gln Ile Met Gly Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Ile Arg Ser Leu Phe Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Lys Ile Glu Val Lys Asp

85 90 95

Thr Lys Glu Ala Leu Glu Glu Val Glu Lys Ala Gln Lys Lys Ser Gln
 100 105 110

 Lys Lys Gln Gln Ala Ala Met Asp Glu Gly Asn Asn Ser Gln Ala Ser
 115 120 125
 Gln Asn Tyr Pro Ile Val Gln Asn Ala Gln Gly Gln Met Val His Gln
 130 135 140
 Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu
 145 150 155 160
 Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu
 165 170 175
 Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly
 180 185 190
 His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala
 195 200 205
 Ala Glu Trp Asp Arg Met His Pro Gln Gln Ala Gly Pro Ile Pro Pro
 210 215 220
 Gly Gln Ile Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser
 225 230 235 240
 Thr Leu Gln Glu Gln Ile Arg Trp Met Thr Ser Asn Pro Pro Ile Pro
 245 250 255

 Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile
 260 265 270
 Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro
 275 280 285
 Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg
 290 295 300
 Ala Glu Gln Ala Thr Gln Glu Val Lys Gly Trp Met Thr Asp Thr Leu
 305 310 315 320
 Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu
 325 330 335
 Gly Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val

340 345 350
 Gly Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln
 355 360 365
 Ala Ser Gly Ala Thr Ile Met Met Gln Lys Ser Asn Phe Lys Gly Pro
 370 375 380
 Arg Arg Asn Ile Lys Cys Phe Asn Cys Gly Lys Glu Gly His Leu Ala
 385 390 395 400

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

465 470 475 480
 Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Ser Asp Pro Ser Gln
 485 490 495

<210> 163
 <211> 1494
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 163

atgggcgcc ggcctccgt gctgtccggc ggcaagctgg acgcctggga gaagatccgc	60
ctgcgccccg gcggcaagaa gaagtaccgc atgaagcacc tgggtgtgggc ctcccgcgag	120
ctggagcgct tcgccctgaa ccccgacctg ctggagaccg ccgagggctg ccagcagatc	180
atgggccagc tgcagccgc cctgcagacc ggcaccgagg agatccgctc cctgttcaac	240
accgtggcca cctgtactg cgtgcaccag aagatcgagg tgaaggacac caaggaggcc	300
ctggaggagg tggagaaggc ccagaagaag tcccagaaga agcagcaggc cgccatggac	360
gagggcaaca actcccaggc ctcccagaac taccccatcg tgcagaacgc ccagggccag	420
atggtgcacc aggccatctc cccccgacc ctgaacgcct gggtgaaggt ggtggaggag	480

aaggccttct ccccgaggt gatcccatg ttctccgcc tgtccgaggg cgccaccccc 540
caggacctga acaccatgct gaacaccgtg ggcggccacc aggccgcat gcagatgctg 600

aaggacacca tcaacgagga ggccgccgag tgggaccgca tgcaccccca gcaggccggc 660
cccatcccc ccggccagat ccgcgagccc cgcggctccg acatcgccgg caccacctcc 720
accctgcagg agcagatccg ctggatgacc tccaaccccc ccatccccgt gggcgagatc 780
tacaagcgt ggatcatcct gggcctgaac aagatcgtgc gcatgtactc ccccggtcc 840
atcctggaca tccgccaggg cccaaggag cccttcgcg actacgtgga ccgcttcttc 900
aagaccctgc gcgccgagca ggccaccag gaggtgaagg gctggatgac cgacaccctg 960
ctggtgcaga acgccaaccc cgactgcaag accatcctgc gcgcctggg ccccgcgcc 1020

accctggagg agatgatgac cgcctgccag ggcgtgggcg gccctccca caaggccgc 1080
gtgctggcgg aggccatgct ccaggcctcc ggcgccacca tcatgatgca gaagtccaac 1140
ttcaagggcc cccgccgcaa catcaagtgc ttcaactgcg gcaaggaggg ccacctggcc 1200
cgcaactgcc gcgcccccg caagaagggc tgctggaagt gcgcaagga gggccaccag 1260
atgaaggact gcaccgagtc caaggccaac ttctgggca agatctggcc ctccaacaag 1320
ggccgccccg gcaacttct gcagaaccgc cccgagccca ccgcccccc cgccgagtc 1380
ttcggttcg gcgaggagat cgccccctcc cccaagcagg agccaagga gaaggagatc 1440

tacccttg cctccctgaa gtcctgttc ggctccgacc cctaatecca gtaa 1494

<210> 164
<211> 206
<212> PRT
<213> Human immunodeficiency virus
<400> 164

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

Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
50 55 60
Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65					70					75					80	
Tyr	Lys	Gly	Ala	Phe	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	
					85						90					
Leu	Asp	Gly	Leu	Ile	Tyr	Ser	Lys	Lys	Arg	Gln	Glu	Ile	Leu	Asp	Leu	
					100						105					
Trp	Val	Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	

115					120					125					
Pro	Gly	Pro	Gly	Ile	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Phe	Lys
130					135					140					
Leu	Val	Pro	Val	Asp	Pro	Glu	Glu	Val	Glu	Glu	Ala	Asn	Glu	Gly	Glu
145					150					155					160
Asn	Asn	Cys	Leu	Leu	His	Pro	Met	Cys	Gln	His	Gly	Met	Glu	Asp	Glu
165					170					175					
Asp	Arg	Glu	Val	Leu	Met	Trp	Lys	Phe	Asp	Ser	Arg	Leu	Ala	Leu	Arg
180					185					190					

His Ile Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys

	195	200	205
<210>	165		
<211>	621		
<212>	DNA		
<213>	Human immunodeficiency virus		
<400>	165		

atgggcggca agtgggtccaa gtccctccatc gtgggctggc ccgccgtgcg cgagcgcac	60
cgccgcaccc ccccgccgc cgagggcggtg ggcgccgtgt cccaggacct ggacaagcac	120
ggcgccatca cctcctccaa caccgccgc accaacgccg actgcgcctg gctggaggcc	180
caggaggagg aggaggtggg ctccccgtg cgccccagg tgccctgcg ccccatgacc	240
tacaagggcg ccttcgacct gtcccacttc ctgaaggaga agggcgccct ggacggcctg	300
atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggtac	360
ttccccgact ggcagaacta ccccccggc cccggcatcc gctacccct gaccttcggc	420
tggtgcttca agctggtgcc cgtggacccc gaggaggtgg aggaggccaa cgagggcgag	480
aacaactgcc tgctgcaccc catgtgccag cacggcatgg aggacgagga ccgcgaggtg	540
ctgatgtgga agttcgactc ccgcctggcc ctgcgccaca tcgcccgcga gctgcacccc	600

gagttctaca aggactgcta a

621

<210> 166

<211> 206

<212> PRT

<213> Human immunodeficiency virus

<400> 166

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Val

1 5 10 15

Arg Glu Arg Met Arg Arg Thr Ala Pro Ala Ala Glu Gly Val Gly Ala

20 25 30

Val Ser Gln Asp Leu Asp Lys His Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu

50 55 60

Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65 70 75 80

Tyr Lys Ala Ala Phe Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly

85 90 95

Leu Asp Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu

100 105 110

Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr

115 120 125

Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys

130 135 140

Leu Val Pro Val Asp Pro Glu Glu Val Glu Glu Ala Asn Glu Gly Glu

145 150 155 160

Asn Asn Cys Leu Leu His Pro Met Cys Gln His Gly Met Glu Asp Glu

165 170 175

Glu Arg Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Leu Arg

180 185 190

His Ile Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys

195	200	205	
<210>	167		
<211>	621		
<212>	DNA		
<213>	Human immunodeficiency virus		
<400>	167		
atgggcggca agtgggtccaa gtcctccatc gtgggctggc cgcctgtgcg cgagcgcgatg		60	
cgccgcaccg cccccggcgc cgagggcggtg ggcgccgtgt cccaggacct ggacaagcac		120	
ggcgccatca cctcctccaa caccgccgcc accaacgccg actgcgcctg gctggaggcc		180	
caggaggagg aggaggtggg cttccccgtg cgccccagg tgccctgtcg ccccatgacc		240	
tacaaggccg ctttcgacct gtcccacttc ctgaaggaga agggcggcct ggacggcctg		300	
atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggtac		360	
ttccccgact ggcagaacta cccccggc cccggcatcc gctacccct gaccttcggc		420	
tggtgtttca agctgggtgcc cgtggacccc gaggaggtgg aggaggccaa cgagggcgag		480	
aacaactgcc tgctgcacc ccatgtgccag cacggcatgg aggacagga gcgcgaggtg		540	
ctgatgtgga agttcgactc ccgcctggcc ctgcgccaca tcgccgcga gctgcacccc		600	
gagttctaca aggactgeta a		621	
<210>	168		
<211>	207		
<212>	PRT		
<213>	Human immunodeficiency virus		
<400>	168		
Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Asp Ile			
1 5 10 15			
Arg Glu Arg Ile Arg Arg Thr Pro Pro Ala Ala Lys Gly Val Gly Ala			
20 25 30			
Val Ser Gln Asp Leu Asp Lys Tyr Gly Ala Val Thr Ile Asn Asn Thr			
35 40 45			
Ala Ala Thr Gln Ala Ser Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu			
50 55 60			
Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met			
65 70 75 80			

Thr Phe Lys Gly Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly
85 90 95
Gly Leu Asp Gly Leu Ile Tyr Ser Gln Lys Arg Gln Glu Ile Leu Asp
100 105 110
Leu Trp Val Tyr Asn Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr
115 120 125
Thr Pro Gly Pro Gly Thr Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe

130 135 140
Lys Leu Val Pro Val Asp Pro Asp Glu Val Glu Glu Ala Thr Glu Gly
145 150 155 160
Glu Asn Asn Cys Leu Leu His Pro Ile Cys Gln His Gly Met Asp Asp
165 170 175
Glu Glu Lys Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Arg
180 185 190
Arg His Ile Ala Leu Glu Met His Pro Glu Phe Tyr Lys Asp Cys
195 200 205

<210> 169
<211> 624
<212> DNA
<213> Human immunodeficiency virus
<400> 169

atgggcggca agtgggtccaa gtectccatc gtgggctggc cgcacatccg cgagcgcac	60
cgccgcaccc ccccgccgc caaggcggtg ggcgccgtgt cccaggacct ggacaagtac	120
ggcgccgtga ccatcaacaa caccgccgc acccaggcct cctgcgcctg gctggaggcc	180
caggaggagg aggaggaggt gggcttcccc gtgcgcccc aggtgccctt gcgccccatg	240
accttcaagg gcgccttcga cctgtccttc ttctgaagg agaagggcgg cctggacggc	300
ctgatctact ccagaagcg ccaggagatc ctggacctgt gggtgtacaa caccagggc	360
tacttccccg actggcagaa ctacaccccc ggccccggca cccgttccc cctgaccttc	420
ggctggtgct tcaagctggt gcccggtggac cccgacgagg tggaggaggc caccgagggc	480
gagaacaact gcctgtgca cccatctgc cagcacggca tggacgacga ggagaaggag	540
gtgctgatgt ggaagttcga ctcccgcctg gcccgccgc acatcgccct ggagatgcac	600
cccgagttct acaaggactg ctaa	624

<210> 170
 <211> 204
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 170

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Glu Val

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

Gly	Ala	Leu	Asp	Leu	Ser	His	Phe	Leu	Lys	Glu	Lys	Gly	Gly	Leu	Asp
85	90	95													
Gly	Leu	Ile	Tyr	Ser	Arg	Lys	Arg	Gln	Glu	Ile	Leu	Asp	Leu	Trp	Val
100	105	110													
Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn	Tyr	Thr	Pro	Gly
115	120	125													
Pro	Gly	Ile	Arg	Tyr	Pro	Leu	Thr	Phe	Gly	Trp	Cys	Phe	Lys	Leu	Val
130	135	140													
Pro	Val	Asp	Pro	Asp	Glu	Val	Glu	Lys	Ala	Thr	Glu	Gly	Glu	Asn	Asn

145	150	155	160
Ser	Leu	Leu	His
Pro	Ile	Cys	Gln
His	Gly	Met	Asp
Asp	Glu	Glu	Arg
165	170	175	
Glu	Val	Leu	Lys
Trp	Lys	Phe	Asp
Ser	Arg	Leu	Ala
Leu	Lys	His	Arg
180	185	190	
Ala	Gln	Glu	Leu
His	Pro	Glu	Phe
Tyr	Lys	Asp	Cys
195	200		

<210> 171

<211> 615
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 171

```

atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccgaggtgcg cgagcgcacg      60
cgccgcaccc cccccgcgc caccggcgtg ggcgccgtgt cccaggacct ggacaagcac      120
ggcgccgtga cctcctccaa catcaaccac cctcctgctg tgtggctgga ggcccaggag      180
gaggaggagg tgggcttccc cgtgcgcccc caggtgcccc tgcgccccat gacctacaag      240
ggcgccctgg acctgtccca ctctctgaag gagaaggcg gcctggacgg cctgatctac      300
tcccgaagc gccaggagat cctggacctg tgggtgtacc acaccaggg ctacttcccc      360
gactggcaga actacacccc cggccccggc atccgctacc cctgacctt cggtggtgc      420

ttcaagctgg tgcccgtgga ccccgcagag gtggagaagg ccaccgaggg cgagaacaac      480
tcctgtctgc acccatctg ccagcacggc atggacgacg aggagcgcga ggtgctgaag      540
tggaagtctg actccgcct ggccctgaag caccgcgccc aggagctgca ccccagttc      600
tacaaggact gctaa      615
    
```

<210> 172
 <211> 206
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 172

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Glu Val
 1 5 10 15

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

85 90 95
 Leu Asp Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu
 100 105 110
 Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 115 120 125
 Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys
 130 135 140
 Leu Val Pro Val Asp Pro Ala Glu Val Glu Glu Ala Thr Glu Gly Glu
 145 150 155 160

Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Met Asp Asp Glu
 165 170 175
 Glu Arg Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Leu Lys
 180 185 190
 His Arg Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys
 195 200 205

<210> 173
 <211> 621
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 173

atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccgaggtgcg cgagcgcacg 60

 cgccgcaccc cccccgccgc caagggcgtg ggcgccgtgt cccaggacct ggacaagcac 120
 ggcgccgtga cctcctccaa caccgccgcc aacaaccccg gctgcgcctg gctggaggcc 180
 caggaggagg aggaggtggg cttccccgtg cgccccagg tgcccctgcg ccccatgacc 240
 tacaagggcg ctttcgacct gtccacttc ctgaaggaga agggcggcct ggacggcctg 300
 atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggtac 360
 ttccccgact ggcagaacta caccgccgcc cccggcatcc gctacccct gaccttcggc 420
 tgggtgttca agctggtgcc cgtggacccc gccgaggtgg aggaggccac cgagggcgag 480

 aacaactccc tgctgcaccc catctgccag cacggcatgg acgacgagga gcgcgaggtg 540
 ctgatgtgga agttcgactc ccgcttgccc ctgaagcacc gcgccgcga gctgcacccc 600
 gagttctaca aggactgcta a 621
 <210> 174

<211> 207

<212> PRT

<213> Human immunodeficiency virus

<400> 174

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Ile

1 5 10 15

Arg Glu Arg Met Arg Lys Arg Thr Pro Pro Ala Ala Glu Gly Val Gly

20 25 30

Ala Val Ser Gln Asp Leu Ala Thr Arg Gly Ala Val Thr Ser Ser Asn

35 40 45

Thr Ala Ala Thr Asn Pro Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu

50 55 60

Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met

65 70 75 80

Thr Phe Lys Gly Ala Phe Asp Leu Ser His Phe Leu Lys Glu Lys Gly

85 90 95

Gly Leu Asp Gly Leu Ile Tyr Ser Gln Lys Arg Gln Asp Ile Leu Asp

100 105 110

Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr

115 120 125

Thr Pro Gly Pro Gly Thr Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe

130 135 140

Lys Leu Val Pro Val Asp Pro Ser Glu Val Glu Glu Ala Thr Glu Gly

145 150 155 160

Glu Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Ile Glu Asp

165 170 175

Pro Glu Arg Glu Val Leu Arg Trp Lys Phe Asp Ser Arg Leu Ala Leu

180 185 190

Arg His Arg Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys

195 200 205

<210> 175

<211> 624

<212> DNA

<213> Human immunodeficiency virus

<400> 175

atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccgccatccg cgagcgcacg 60

cgcaagcgca cccccccgc cgccgagggc gtgggcgccg tgtcccagga cctggccacc 120

cgcggcgcgc tgacctctc caacaccgcc gccaccaacc ccgactgcgc ctggctggag 180

gcccaggagg aggaggaggt gggcttcccc gtgcgcccc aggtgccct gcgccccatg 240

accttcaagg gcgccttcga cctgtccac ttcctgaagg agaagggcgg cctggacggc 300

ctgatctact ccagaagcg ccaggacatc ctggacctgt ggggtgtacca caccagggc 360

tacttcccc actggcagaa ctacacccc ggccccggca cccgctacc cctgacctc 420

ggctgggtgt tcaagctggt gcccgtagg ccctccgagg tggaggagc caccgagggc 480

gagaacaact ccctgtgca ccccatctgc cagcacggca tcgaggacc cgagcgcgag 540

gtgctgcgct ggaagttcga ctccgcctg gccctgcgc accgcgccg cgagctgcac 600

cccgagttct acaaggactg ctaa 624

<210> 176

<211> 206

<212> PRT

<213> Human immunodeficiency virus

<400> 176

Met Gly Gly Lys Trp Ser Lys Arg Ser Val Val Gly Trp Pro Thr Val

1 5 10 15

Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala

20 25 30

Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Ala Ala Asn Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu

50 55 60

Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65 70 75 80

Tyr Lys Gly Ala Leu Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly

85 90 95

Leu Glu Gly Leu Ile Tyr Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu

	100	105	110	
	Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr			
	115	120	125	
	Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys			
	130	135	140	
	Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Asn Glu Gly Glu			
	145	150	155	160
	Asn Asn Ser Leu Leu His Pro Met Ser Leu His Gly Met Asp Asp Pro			
	165	170	175	
	Glu Arg Glu Val Leu Val Trp Lys Phe Asp Ser Arg Leu Ala Phe His			
	180	185	190	
	His Met Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys			
	195	200	205	
<210>	177			
<211>	621			
<212>	DNA			
<213>	Human immunodeficiency virus			
<400>	177			
	atgggcggca agtgggtccaa gcgctccgtg gtgggctggc ccaccgtgcg cgagcgcgatg	60		
	cgccgcgcgcg agcccgccgc cgacggcgtg ggcgccgtgt cccgcgacct ggagaagcac	120		
	ggcgccatca cctcctccaa caccgccgcc aacaacgccg actgcgcctg gctggaggcc	180		
	caggaggagg aggaggtggg ctccccctg cgcgccagg tgccctgcg ccccatgacc	240		
	tacaaggcgc ccttgacct gtcccacttc ctgaaggaga agggcggcct ggagggcctg	300		
	atctactccc agaagcgcca ggacatcctg gacctgtggg tgtaccacac ccagggtac	360		
	ttccccgact ggcagaacta caccgccgc cccggcatcc gctacccct gaccttcggc	420		
	tggtgcttca agctggtgcc cgtggagccc gagaaggtgg aggaggcaa cgagggcgag	480		
	aacaactccc tgtgcaccc catgtccctg cacggcatgg acgacccga gcgcgaggtg	540		
	ctggtgtgga agttcgactc ccgctggcc ttccaccaca tggcccgcga gctgcacccc	600		
	gagtactaca aggactgcta a	621		
<210>	178			
<211>	206			

<212> PRT

<213> Human immunodeficiency virus

<400> 178

Met Gly Gly Lys Trp Ser Lys Ser Ser Met Gly Gly Trp Pro Ala Val

1 5 10 15

Arg Glu Arg Met Lys Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala

20 25 30

Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu

50 55 60

Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65 70 75 80

Tyr Lys Ala Ala Leu Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly

85 90 95

Leu Glu Gly Leu Ile Tyr Ser Gln Lys Arg Gln Asp Ile Leu Asp Leu

100 105 110

Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr

115 120 125

Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys

130 135 140

Leu Val Pro Val Glu Pro Glu Lys Val Glu Glu Ala Thr Glu Gly Glu

145 150 155 160

Asn Asn Ser Leu Leu His Pro Met Cys Gln His Gly Met Asp Asp Pro

165 170 175

Glu Lys Glu Val Leu Val Trp Lys Phe Asp Ser Arg Leu Ala Phe His

180 185 190

His Met Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys

195 200 205

<210> 179

<211> 621

<212> DNA

<213> Human immunodeficiency virus

<400> 179

atgggcggca agtgggtccaa gtcttccatg ggcggtggc cgcctgtgc cgagcgcag 60
aagcgcgccg agcccgccgc cgacggcgtg ggcgccgtgt cccgcgacct ggagaagcac 120
ggcgccatca cctcctccaa caccgcccgc accaacgccg actgcgcctg gctggaggcc 180
caggaggagg aggaggtggg ctccccgtg cgcgccagg tgccctgcg ccccatgacc 240
tacaaggccg ccctggacct gtcccattc ctgaaggaga agggcggcct ggagggcctg 300

atctactccc agaagcgcca ggacatcctg gacctgtggg tgtaccacac ccagggtac 360
ttccccgact ggagaaacta cccccccgc cccggcatcc gctacccct gaccttcggc 420
tggtgttca agctggtgcc cgtggagccc gagaaggtgg aggaggccac cgaggcgag 480
aacaactccc tgtgcaccc catgtgccag caggcatgg acgacccga gaaggagtg 540
ctggtgtgga agtgcactc ccgctggcc ttccaccaca tggccgcga gctgcacccc 600
gagtactaca aggactgcta a 621

<210> 180

<211> 206

<212> PRT

<213> Human immunodeficiency virus

<400> 180

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Lys Val

1 5 10 15

Arg Glu Arg Ile Arg Gln Thr Pro Pro Ala Ala Thr Gly Val Gly Ala

20 25 30

Ala Ser Gln Asp Leu Asp Arg His Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu

50 55 60

Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65 70 75 80

Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly

85 90 95

Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu

100 105 110

Trp Val Tyr His Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr Thr

115 120 125

Pro Gly Pro Gly Thr Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys

130 135 140

Leu Val Pro Met Asp Pro Ala Glu Val Glu Glu Ala Asn Glu Gly Glu

145 150 155 160

Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Met Glu Asp Glu

165 170 175

Asp Arg Glu Val Leu Val Trp Arg Phe Asp Ser Ser Leu Ala Phe Lys

180 185 190

His Arg Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys

195 200 205

<210> 181

<211> 621

<212> DNA

<213> Human immunodeficiency virus

<400> 181

atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccaaggtgcg cgagcgcac	60
cgccagaccc ccccgccgc caccggcgtg ggcgccgcct cccaggacct ggaccgccac	120
ggcgccatca cctcctccaa caccgccgc accaacgccg actgcgcctg gctggaggcc	180
caggaggagg aggaggtggg cttccccgtg cgccccagg tgcccctgcg ccccatgacc	240
tacaaggccg ccgtggacct gtcccattc ctgaaggaga agggcggcct ggagggcctg	300
atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggttc	360

ttccccgact ggcagaacta caccgccgc cccggcaccc gcttcccct gaccttcggc	420
tggtgcttca agctggtgcc catggacccc gccgaggtgg aggaggccaa cgaggcgag	480
aacaactccc tgctgcaccc catctgccag cacggcatgg aggacgagga ccgcgaggtg	540
ctggtgtggc gcttcgactc ctcctggcc ttcaagcacc gcgccgcga gctgcacccc	600
gagttctaca aggactgcta a	621

<210> 182

<211> 207

<212> PRT

<213> Human immunodeficiency virus

<400> 182

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Val

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

Thr Tyr Lys Ala Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly			
85	90	95	
Gly Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp			
100	105	110	
Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr			
115	120	125	
Thr Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe			
130	135	140	
Lys Leu Val Pro Val Asp Pro Arg Glu Val Glu Glu Ala Asn Glu Gly			

145	150	155	160
Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Met Glu Asp			
165	170	175	
Glu Asp Arg Glu Val Leu Lys Trp Lys Phe Asp Ser His Leu Ala Arg			
180	185	190	
Arg His Met Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys			
195	200	205	

<210> 183

<211> 624

<212> DNA

<213> Human immunodeficiency virus

<400> 183
atgggcggca agtgggtccaa gtcttccatc gtgggctggc cgcctgtgcg cgagcgcac 60
cgccgcaccg agccccccgc cgagggcgtg ggccgcgcct cccaggacct ggacaagcac 120
ggcgccctga cctctccaa caccgccacc aacaacgccg actgcgcctg gctggaggcc 180
caggaggagg aggaggaggt gggcttcccc gtgcgcccc aggtgccct gcgccccatg 240
acctacaagg ccgcttccga cctgtcttc ttctgaagg agaagggcgg cctggagggc 300
ctgatctact ccaagaagcg ccaggagatc ctggacctgt ggggttacca caccagggc 360
tacttccccg actggcagaa ctacacccc ggccccggcg tgcgtaccc cctgaccttc 420

ggctgtgtct tcaagctgtt gccctggac cccgcgagg tggaggaggc caacgagggc 480
gagaacaact gcctgtgca cccatgtcc cagcacggca tggaggacga ggaccgcgag 540
gtgtgaagt ggaagttcga ctcccactg gccgcgcgc acatggcccg cgagctgcac 600
cccagtgact acaaggactg ctaa 624

<210> 184
<211> 207
<212> PRT
<213> Human immunodeficiency virus
<400> 184

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Val
1 5 10 15

Arg Glu Arg Met Arg Arg Thr Glu Pro Ala Ala Glu Gly Val Gly Ala
20 25 30

Ala Ser Gln Asp Leu Asp Lys His Gly Ala Leu Thr Ser Ser Asn Thr
35 40 45

Ala Ala Asn Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
50 55 60

Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
65 70 75 80

Thr Tyr Lys Ala Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly
85 90 95

Gly Leu Asp Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp
100 105 110

Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr

115 120 125
 Thr Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe
 130 135 140
 Lys Leu Val Pro Val Asp Pro Arg Glu Val Glu Glu Ala Asn Glu Gly
 145 150 155 160

Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Met Glu Asp
 165 170 175
 Glu Asp Arg Glu Val Leu Lys Trp Lys Phe Asp Ser His Leu Ala Arg
 180 185 190
 Arg His Met Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys
 195 200 205

<210> 185
 <211> 624
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 185

atgggcggca agtgggtccaa gtcctccatc gtgggctggc cgcctgtgcg cgagcgcattg 60

cgccgcaccg agccccccgc cgagggcgtg ggcgcgcctt cccaggacct ggacaagcac 120
 ggcgccctga cctcctccaa caccgccgcc aacaacgccg actgcgcctg gctggaggcc 180
 caggaggagg aggaggaggt gggcttcccc gtgcgcccc aggtgccctt gcgccccatg 240
 acctacaagg ccgccttcga cctgtccttc ttctgaagg agaagggcgg cctggacggc 300
 ctgatctact ccaagaagcg ccaggagatc ctggacctgt ggggtgtacca caccagggc 360
 tacttccccg actggcagaa ctacaccccc ggccccggcg tgcgtaccc cctgaccttc 420
 ggctggtgct tcaagctggt gcccggtggac ccccgcgagg tggaggaggc caacgagggc 480

gagaacaact gcctgtgca ccccatgtcc cagcacggca tggaggacga ggaccgcgag 540
 gtgctgaagt ggaagttcga ctcccacctg gcccgccgcc acatggcccc cgagctgcac 600
 cccgagtact acaaggactg ctaa 624

<210> 186
 <211> 207
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 186

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Ile
1 5 10 15
Arg Glu Arg Ile Arg Arg Thr Glu Pro Ala Ala Asp Gly Val Gly Ala
20 25 30
Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr
35 40 45
Ala Ala Thr Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Asp
50 55 60
Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
65 70 75 80
Thr Tyr Lys Ala Ala Leu Asp Leu Ser His Phe Leu Lys Glu Lys Gly
85 90 95
Gly Leu Glu Gly Leu Val Trp Ser Gln Lys Arg Gln Glu Ile Leu Asp
100 105 110
Leu Trp Val Tyr Asn Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr
115 120 125
Thr Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe
130 135 140
Glu Leu Val Pro Val Asp Pro Glu Glu Val Glu Glu Ala Thr Glu Gly
145 150 155 160
Glu Asn Asn Cys Leu Leu His Pro Met Cys Gln His Gly Met Glu Asp
165 170 175
Pro Glu Arg Glu Val Leu Met Trp Arg Phe Asn Ser Arg Leu Ala Phe
180 185 190
Glu His Lys Ala Arg Val Leu His Pro Glu Phe Tyr Lys Asp Cys
195 200 205
<210> 187
<211> 624
<212> DNA
<213> Human immunodeficiency virus
<400> 187

atgggcggca agtgggtccaa gtcctccatc gtgggctggc cgcctatccg cgagcgcac

60

cgccgcaccg agccccccgc cgacggcgtg ggccgctgt cccgcgacct ggagaagcac 120

ggcgccatca cctcctccaa caccgccgcc accaacgcc actgcgcctg gctggaggcc 180

caggaggagg acgaggaggt gggcttcccc gtgcgcccc aggtgccctt gcgccccatg 240

acctacaagg ccgccctgga cctgtccac ttctgaagg agaagggcgg cctggagggc 300

ctggtgtggt ccagaagcg ccaggagatc ctggacctgt ggggtgtaca caccagggc 360

ttcttccccg actggcagaa ctacacccc ggccccggca tccgtacct cctgaccttc 420

ggctggtgct tcgagctggt gccctggac cccgaggagg tggaggaggc caccagggc 480

gagaacaact gcctgtgca ccccatgtgc cagcacggca tggaggacct cgagcgcgag 540

gtgctgatgt ggcgcttcaa ctcccgcctg gccttcgagc acaaggcccg cgtgctgcac 600

cccgagttct acaaggactg ctaa 624

<210> 188

<211> 205

<212> PRT

<213> Human immunodeficiency virus

<400> 188

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Val

1 5 10 15

Arg Glu Arg Met Arg Pro Thr Pro Pro Ala Ala Glu Gly Val Gly Ala

20 25 30

Val Ser Gln Asp Leu Glu Arg Arg Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Gly Ala Thr Asn Pro Asp Leu Ala Trp Leu Glu Ala Gln Glu Glu Glu

50 55 60

Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65 70 75 80

Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly

85 90 95

Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu

100 105 110

Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr

115 120 125

Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys
 130 135 140
 Leu Val Pro Val Asp Pro Glu Glu Val Glu Lys Ala Asn Glu Gly Glu
 145 150 155 160
 Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Met Glu Asp Glu
 165 170 175
 Asp Arg Glu Val Leu Ile Trp Lys Phe Asp Ser Arg Leu Ala Leu Arg

180 185 190
 His Ile Ala Arg Glu Arg His Pro Glu Phe Tyr Gln Asp
 195 200 205

<210> 189
 <211> 618
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 189

atgggcggca agtgggtccaa gtcctccatc gtgggctggc cgcctgtgcg cgagcgcgatg 60
 cgccccacc ccccgccgc cgagggcgtg ggcgccgtgt cccaggacct ggagcgccgc 120
 ggcgccatca cctcctccaa caccggcgcc accaaccctg acctggcctg gctggaggcc 180
 caggaggagg aggaggtggg cttccccgtg cgccccagg tgccctgcg ccccatgacc 240

 tacaaggcgc ccgtggacct gtcccacttc ctgaaggaga agggcggcct ggagggcctg 300
 atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggtac 360
 ttccccgact ggcagaacta caccgccgc cccggcatcc gctacccct gaccttcggc 420
 tgggtcttca agctggtgcc cgtggacccc gaggaggtgg agaaggccaa cgagggcgag 480
 aacaactgcc tgctgcacc catgtcccag caggcatgg aggacgagga ccgcgaggtg 540
 ctgatctgga agttcgactc ccgctggcc ctgcgccaca tcgccgcga gcgccaccc 600
 gagttctacc aggactaa 618

<210> 190
 <211> 205
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 190

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Thr Ile

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

 50 55 60
 Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr
 65 70 75 80
 Tyr Lys Ala Ala Phe Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 85 90 95
 Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu
 100 105 110
 Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 115 120 125

 Pro Gly Pro Gly Thr Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys
 130 135 140
 Leu Val Pro Val Asp Pro Glu Glu Val Glu Lys Ala Asn Glu Gly Glu
 145 150 155 160
 Asn Asn Cys Leu Leu His Pro Met Ser Leu His Gly Met Glu Asp Glu
 165 170 175
 Asp Arg Glu Val Leu Lys Trp Lys Phe Asp Ser Arg Leu Ala Leu Arg
 180 185 190
 His Ile Ala Arg Glu Arg His Pro Glu Tyr Tyr Lys Asp

 195 200 205
 <210> 191
 <211> 618
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 191

atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccaccatccg cgagcgcac 60
 cgccgcaccc ccgtggccgc cgagggcgtg ggcgccgtgt cccaggacct ggacaagcac 120

ggcgccatca cctcctccaa caccgcgcc accaacgccg acctggcctg gctggaggcc 180
caggaggacg aggaggtggg cttccccgtg cgccccagg tgcccctgcg ccccatgacc 240
tacaaggccg ctttcgacct gtcccattc ctgaaggaga agggcggcct ggagggcctg 300

atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggttac 360
ttccccgact ggcagaacta caccgccggc cccggcacc gctaccccct gaccttcggc 420
tggtgcttca agctggtgcc cgtggacccc gaggaggtgg agaaggccaa cgagggcgag 480
aacaactgcc tgctgcaccc catgtccctg cacggcatgg aggacgagga ccgcgaggtg 540
ctgaagtgga agttcgactc ccgcctggcc ctgcgccaca tcgccgcga gcgccacccc 600
gagtactaca aggactaa 618

<210> 192

<211> 207

<212> PRT

<213> Human immunodeficiency virus

<400> 192

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Glu Val

1 5 10 15

Arg Glu Arg Ile Arg Gln Thr Pro Pro Ala Ala Glu Gly Val Gly Ala

20 25 30

Val Ser Gln Asp Leu Ala Arg His Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Ala Ala Asn Asn Pro Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Asp

50 55 60

Ser Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met

65 70 75 80

Thr Tyr Lys Gly Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly

85 90 95

Gly Leu Asp Gly Leu Ile Tyr Ser Lys Lys Arg Gln Asp Ile Leu Asp

100 105 110

Leu Trp Val Tyr Asn Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr

115 120 125

Thr Pro Gly Pro Gly Thr Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe

130 135 140
 Lys Leu Val Pro Met Asp Pro Ala Glu Val Glu Glu Ala Asn Lys Gly
 145 150 155 160
 Glu Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Met Glu Asp
 165 170 175
 Glu Asp Arg Glu Val Leu Val Trp Arg Phe Asp Ser Ser Leu Ala Arg
 180 185 190
 Arg His Ile Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys
 195 200 205

<210> 193
 <211> 624
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 193
 atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccgaggtgcg cgagcgcac 60
 cgccagaccc cccccgccgc cgagggcggtg ggcgccgtgt cccaggacct ggccccccac 120
 ggcgccatca cctcctccaa caccgccgcc aacaaccccg actgcgcctg gctggaggcc 180
 caggaggagg actccgaggt gggcttcccc gtgcgcccc aggtgccctt gcgccccatg 240
 acctacaagg gcgccttcga cctgtccttc ttctgaagg agaagggcgg cctggacggc 300
 ctgatctact ccaagaagcg ccaggacatc ctggacctgt gggtgtacaa caccagggc 360
 ttcttccccg actggcagaa ctacacccc ggccccggca cccgcttccc cctgaccttc 420
 ggctggtgct tcaagctggt gcccattgac cccgccgagg tggaggaggc caacaagggc 480
 gagaacaact ccctgtgca ccccatctgc cagcacggca tggaggacga ggaccgcgag 540
 gtgctggtgt ggcgcttcga ctctccctg gcccgccgcc acatcgccc cgagctgcac 600
 cccgagtact acaaggactg ctaa 624

<210> 194
 <211> 207
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 194

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Gly Gly Trp Pro Ala Ile

Arg Glu Arg Ile Arg Arg Ala Glu Pro Ala Ala Glu Gly Val Gly Ala
 20 25 30
 Val Ser Arg Asp Leu Asp Arg Arg Gly Ala Val Thr Ile Asn Asn Thr
 35 40 45
 Ala Ser Thr Asn Pro Asp Ser Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 55 60
 Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
 65 70 75 80

Thr Tyr Lys Gly Ala Phe Asp Leu Ser His Phe Leu Lys Glu Lys Gly
 85 90 95
 Gly Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp
 100 105 110
 Leu Trp Val Tyr Asn Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr
 115 120 125
 Thr Pro Gly Pro Gly Glu Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe
 130 135 140
 Lys Leu Val Pro Val Asp Pro Gln Glu Val Glu Lys Ala Asn Glu Gly

145 150 155 160
 Glu Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Met Glu Asp
 165 170 175
 Glu Glu Arg Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Phe
 180 185 190
 Arg His Ile Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys
 195 200 205

<210> 195
 <211> 624
 <212> DNA
 <213> Human immunodeficiency virus

<400> 195
 atgggcggca agtgggtccaa gtctccatc ggcggtggc cgcctatccg cgagcgcac 60
 cgccgcgccg agcccgccgc cgaggcggtg ggcgccgtgt cccgcgacct ggaccgccgc 120
 ggcgccgtga ccatcaacaa caccgcctcc accaaccgcc actccgcctg gctggaggcc 180

caggaggagg aggaggaggt gggcttcccc gtgcgcccc aggtgccctt gcgccccatg 240
acctacaagg gcgccttcga cctgtccac ttctgaagg agaagggcgg cctggagggc 300
ctgatctact ccaagaagcg ccaggagatc ctggacctgt ggggtgtacaa cacccagggc 360
tacttccccg actggcagaa ctacaccccc ggccccggcg agcgctaccc cctgaccttc 420

ggctgtgtct tcaagctggt gcccgtggac ccccaggagg tggagaaggc caacgagggc 480
gagaacaact cctgtctga ccccatctgc cagcacggca tggaggacga ggagcgcgag 540
gtgctgatgt ggaagttcga ctccgcctg gccttcgcc acatcgccg cgagctgcac 600
cccgagttct acaaggactg ctaa 624

<210> 196
<211> 204
<212> PRT
<213> Human immunodeficiency virus
<400> 196

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Gln Val
1 5 10 15

Arg Glu Arg Ile Lys Gln Thr Pro Pro Ala Thr Glu Gly Val Gly Ala
20 25 30

Val Ser Gln Asp Leu Asp Lys His Gly Ala Val Thr Ser Ser Asn Met
35 40 45

Asn Asn Ala Asp Cys Val Trp Leu Arg Ala Gln Glu Glu Glu Glu Val
50 55 60

Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys
65 70 75 80

Gly Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly Gly Leu Asp

85 90 95

Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu Trp Val
100 105 110

Tyr Asn Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly
115 120 125

Pro Gly Ile Arg Tyr Pro Leu Cys Phe Gly Trp Cys Phe Lys Leu Val
130 135 140

Pro Val Asp Pro Arg Glu Val Glu Glu Asp Asn Lys Gly Glu Asn Asn

20 25 30
 Pro Val Ser Gln Asp Leu Asp Lys Tyr Gly Ala Val Thr Ser Ser Asn
 35 40 45
 Thr Ala Ala Asn Asn Ala Asp Cys Ala Trp Leu Glu Ala Gln Lys Glu
 50 55 60
 Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
 65 70 75 80
 Thr Tyr Lys Gly Ala Phe Asp Leu Ser His Phe Leu Lys Glu Lys Gly
 85 90 95
 Gly Leu Asp Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp
 100 105 110
 Leu Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr
 115 120 125
 Thr Pro Gly Pro Gly Ile Arg Phe Pro Leu Thr Phe Gly Trp Cys Tyr
 130 135 140
 Lys Leu Val Pro Val Asp Pro Asp Glu Val Glu Glu Ala Thr Glu Gly
 145 150 155 160
 Glu Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Met Asp Asp

165 170 175
 Glu Glu Lys Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Leu
 180 185 190
 Thr His Arg Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys
 195 200 205

<210> 199

<211> 624

<212> DNA

<213> Human immunodeficiency virus

<400> 199

atgggcggca agtgggtccaa gtcctccatc gtgggctggc cccaggtgcg cgagcgcac 60

cgccgcgccc ccgccccgc gcgccgaggc gtgggccccg tgtcccagga cctggacaag 120

tacggcgccg tgacctctc caacaccgcc gccaacaacg ccgactgcgc ctggctggag 180

gcccagaagg aggaggaggt gggcttcccc gtgcgcccc aggtgccct gcgccccatg 240

acctacaagg ggccttcga cctgtccac ttctgaagg agaagggcgg cctggacggc 300
ctgatctact ccaagaagcg ccaggagatc ctggacctgt gggtgtacca caccagggc 360
tatttccccg actggcagaa ctacaccccc ggccttcggc tccgttccc cctgaccttc 420
ggctgggtgct acaagctggg gccctggac cccgacgagg tggaggaggc caccgagggc 480
gagaacaact cctgtctgca ccccatctgc cagcacggca tggacgacga ggagaaggag 540

gtgctgatgt ggaagttcga ctccgcctg gccctgaccc accgcgccg cgagctgcac 600
cccgatttct acaaggactg ctaa 624

<210> 200
<211> 212
<212> PRT
<213> Human immunodeficiency virus
<400> 200

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Ile

1 5 10 15

Arg Glu Arg Met Arg Gln Arg Gly Pro Ala Gln Ala Glu Pro Ala Ala

20 25 30

Ala Gly Val Gly Ala Val Ser Gln Asp Leu Asp Lys His Gly Ala Ile

35 40 45

Thr Ser Ser Asn Thr Ala Ala Thr Asn Pro Asp Lys Ala Trp Leu Glu

50 55 60

Ala Gln Glu Glu Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val

65 70 75 80

Pro Leu Arg Pro Met Thr Phe Lys Ala Ala Leu Asp Leu Ser His Phe

85 90 95

Leu Lys Glu Lys Gly Gly Leu Asp Gly Leu Ile Tyr Ser Lys Lys Arg

100 105 110

Gln Glu Ile Leu Asp Leu Trp Val Tyr Asn Thr Gln Gly Tyr Phe Pro

115 120 125

Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Glu Arg Phe Pro Leu Cys

130 135 140

Phe Gly Trp Cys Phe Lys Leu Val Pro Val Asp Pro Gln Glu Val Glu

145 150 155 160

Glu Ala Thr Glu Gly Glu Asn Asn Cys Leu Leu His Pro Ile Ser Gln

165 170 175

His Gly Met Glu Asp Glu Glu Arg Glu Val Leu Lys Trp Lys Phe Asp

180 185 190

Ser Arg Leu Ala Tyr Lys His Ile Ala Arg Glu Leu His Pro Glu Phe

195 200 205

Tyr Lys Asp Cys

210

<210> 201

<211> 639

<212> DNA

<213> Human immunodeficiency virus

<400> 201

atgggcgcca agtgggtccaa gtctccatc gtgggctggc ccgcatccg cgagcgcatg 60

cgccagcgcg gccccgcca ggccgagccc gccgccgccc gcgtgggccc cgtgtcccag 120

gacctggaca agcacggcgc catcacctcc tccaacaccg ccgccaccaa ccccgacaag 180

gcctggctgg agggccagga ggaggaggag gaggtgggct tccccgtgcg ccccaggtg 240

cccctgcgcc ccatgacctt caaggccgcc ctggacctgt cccacttcct gaaggagaag 300

ggcggcctgg acggcctgat ctactccaag aagcggcagg agatcctgga cctgtgggtg 360

tacaacaccc agggctactt ccccgactgg cagaactaca ccccgggccc cggcgagcgc 420

ttccccctgt gcttcggctg gtgcttcaag ctggtgcccg tggaccccca ggaggtggag 480

gaggccaccg agggcgagaa caactgcctg ctgcacccca tctcccagca cggcatggag 540

gacgaggagc gcgaggtgct gaagtgggaag ttgcactccc gcctggccta caagcacatc 600

gccccgcgagc tgcaccccga gttctacaag gactgctaa 639

<210> 202

<211> 208

<212> PRT

<213> Human immunodeficiency virus

<400> 202

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Gln Val

1 5 10 15

Arg Glu Arg Met Arg Asn Pro Pro Thr Glu Gly Ala Ala Glu Gly Val

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

115 120 125
 Tyr Thr Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys
 130 135 140
 Tyr Lys Leu Val Pro Val Asp Pro Lys Glu Val Glu Glu Asp Thr Lys
 145 150 155 160
 Gly Glu Asn Asn Cys Leu Leu His Pro Met Cys Gln His Gly Val Glu
 165 170 175
 Asp Glu Glu Arg Glu Val Leu Met Trp Lys Phe Asp Ser Ser Leu Ala
 180 185 190

Arg Arg His Ile Ala Arg Glu Met His Pro Glu Phe Tyr Lys Asp Cys
 195 200 205

<210> 203

<211> 627

<212> DNA

<213> Human immunodeficiency virus

<400> 203

atgggcggca agtgggtccaa gtcctccatc gtgggctggc cccaggtgcg cgagcgcattg 60
 cgcaaccccc ccaccgaggg gcgcgccgag ggcgtgggcg ccgtgtccca ggacctggac 120
 aagcacggcg ccatcacctc ctccaacacc gccaccacca acgcgcgctg cgcctggctg 180

gaggcccaga ccgaggacga ggtgggcttc cccgtgcgcc cccaggtgcc cctgcgcccc 240

atgacctaca agggcgccctt cgacctgtcc ttcttcctga aggagaagg cgccctggac 300

ggcctgatct actccaagaa gcgccaggag atcctggacc tgtgggtgta ccacaccag 360

ggctttcttc ccgactggca gaactacacc cccggccccg gcatccgcta cccctgacc 420

ttcggtggt gctacaagct ggtgcccgtg gacccaagg aggtggagga ggacaccaag 480

ggcgagaaca actgcctgct gcaccccatg tgccagcacg gcgtggagga cgaggagcgc 540

gaggtgctga tgtggaagtt cgactcctcc ctggcccgcc gccacatcgc ccgcgagatg 600

caccccgagt tctacaagga ctgctaa 627

<210> 204

<211> 206

<212> PRT

<213> Human immunodeficiency virus

<400> 204

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Ile

1 5 10 15

Arg Glu Arg Ile Arg Arg Thr Glu Pro Ala Ala Asp Gly Val Gly Ala

20 25 30

Val Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr

35 40 45

Ala Asp Thr Asn Ala Asp Cys Ala Trp Leu Glu Thr Gln Glu Glu Glu

50 55 60

Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr

65 70 75 80

Phe Lys Gly Ala Leu Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly Gly

85 90 95

Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu

100 105 110

Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp His Asn Tyr Thr

115 120 125

Pro Gly Pro Gly Val Arg Phe Pro Leu Thr Phe Gly Trp Cys Phe Lys

130 135 140

Leu Val Pro Val Asp Pro Arg Glu Val Glu Glu Ala Asn Glu Gly Glu
 145 150 155 160
 Asp Asn Cys Leu Leu His Pro Val Cys Gln His Gly Met Glu Asp Glu
 165 170 175
 His Arg Glu Val Leu Lys Trp Lys Phe Asp Ser Gln Leu Ala His Arg
 180 185 190
 His Arg Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys

195 200 205
 <210> 205
 <211> 621
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 205
 atgggcggca agtgggtccaa gtctccatc gtgggctggc ccgcatccg cgagcgcac 60
 cgccgcaccg agcccgccgc cgacggcgtg ggcgccgtgt cccgcgacct ggagaagcac 120
 ggcgccatca cctcctccaa caccgccgac accaacgccg actgcgcctg gctggagacc 180
 caggaggagg aggaggtggg ctccccgtg cgccccagg tgcctctgcg ccccatgacc 240
 ttcaaggcg ccttgacct gtccttcttc ctgaaggaga agggcggcct ggagggcctg 300
 atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggtac 360
 ttccccgact ggcacaacta caccgccggc cccggcgtgc gcttccccct gaccttcggc 420
 tgggtgctta agctggtgcc cgtggacccc cgcgaggtgg aggaggccaa cgagggcgag 480
 gacaactgcc tgctgcacc cgtgtgccag cacggcatgg aggacgagca ccgcgaggtg 540
 ctgaagtgga agttcgactc ccagctggcc caccgccacc gcgcccgcga gctgcacccc 600
 gagttctaca aggactgcta a 621

<210> 206
 <211> 207
 <212> PRT

<213> Human immunodeficiency virus
 <400> 206

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Ala Val
 1 5 10 15
 Arg Glu Arg Ile Arg Arg Thr Asp Pro Ala Ala Glu Gly Val Gly Ala

20 25 30
 Ala Ser Arg Asp Leu Glu Lys Tyr Gly Ala Ile Thr Ser Ser Asn Thr
 35 40 45
 Ala Gln Thr Asn Pro Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 55 60

 Glu Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
 65 70 75 80
 Thr Tyr Lys Gly Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly
 85 90 95
 Gly Leu Glu Gly Leu Ile Tyr Ser Lys Arg Arg Gln Asp Ile Leu Asp
 100 105 110
 Leu Trp Val Tyr Asn Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr
 115 120 125
 Thr Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr

130 135 140
 Lys Leu Val Pro Val Asp Pro Arg Glu Val Glu Glu Ala Asn Glu Gly
 145 150 155 160
 Glu Asn Asn Ser Leu Leu His Pro Met Ser Leu His Gly Met Glu Asp
 165 170 175
 Pro His Gly Glu Val Leu Met Trp Lys Phe Asp Ser Asn Leu Ala His
 180 185 190
 Lys His Met Ala Arg Glu Leu His Pro Glu Tyr Tyr Lys Asp Cys
 195 200 205

<210> 207
 <211> 624
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 207

atgggcggca agtgggtccaa gtcctccatc gtgggctggc cgcgcgtgcg cgagcgcac 60
 cgccgcaccg acccgccgc cgaggcggtg ggcgcgcct cccgcgacct ggagaagtac 120
 ggcgcatca cctcctccaa caccgcccag accaaccg actgcgcctg gctggaggcc 180
 caggaggagg aggaggaggt gggcttccc gtgcgcccc aggtgccct gcgccccatg 240

acctacaagg gcgccttcga cctgtccttc ttctgaagg agaagggcgg cctggagggc 300
ctgatctact ccaagcgccg ccaggacatc ctggacctgt gggtgtacaa caccagggc 360

tttttccccg actggcagaa ctacaccccc ggccccggca tcegtaccc cctgaccttc 420
ggctgggtgct acaagctggg gcccggtggac ccccgcgagg tggaggaggc caacgagggc 480
gagaacaact cctgctgca ccccatgtcc ctgcacggca tggaggaccc ccacggcgag 540
gtgctgatgt ggaagtcca ctccaacctg gccacaage acatggcccc cgagctgcac 600
cccgagtact acaaggactg ctaa 624

<210> 208
<211> 208
<212> PRT
<213> Human immunodeficiency virus
<400> 208

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Glu Ile

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

Met	Thr	Tyr	Lys	Gly	Ala	Phe	Asp	Leu	Gly	Phe	Phe	Leu	Lys	Glu	Lys
85	90	95													
Gly	Gly	Leu	Asp	Gly	Leu	Ile	Tyr	Ser	Lys	Lys	Arg	Gln	Glu	Ile	Leu
100	105	110													
Asp	Leu	Trp	Val	Tyr	His	Thr	Gln	Gly	Tyr	Phe	Pro	Asp	Trp	Gln	Asn
115	120	125													
Tyr	Thr	Pro	Gly	Pro	Gly	Ile	Arg	Tyr	Pro	Leu	Cys	Phe	Gly	Trp	Cys
130	135	140													
Phe	Lys	Leu	Val	Pro	Val	Glu	Pro	Arg	Glu	Val	Glu	Glu	Ala	Asn	Glu

145 150 155 160
Gly Glu Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Met Asp

 165 170 175
Asp Glu Glu Arg Glu Val Leu Met Trp Lys Phe Asp Ser Ser Leu Ala

 180 185 190
Arg Arg His Ile Ala Arg Glu Leu His Pro Asp Phe Tyr Lys Asp Cys

 195 200 205

<210> 209

<211> 627

<212> DNA

<213> Human immunodeficiency virus

<400> 209

atgggcggca agtgggtccaa gtcctccatc gtgggctggc ccgagatccg cgagcgcctg 60
cgccgcaccc cccccaccgc cgccgccgag ggctgtggcg ccgtgtccaa ggacctggag 120
aagcacggcg ccgtgacctc ctccaacacc gccagacca acgccgcctg cgcctggctg 180
gaggcccagg aggaggagga ggtgggcttc cccgtgcgcc cccaggtgcc cctgcgcccc 240
atgacctaca agggcgccctt cgacctgggc ttcttctga aggagaaggc cggcctggac 300
ggcctgatct aciccaagaa gcgccaggag atcctggacc tgtgggtgta ccacaccag 360
ggctacttcc ccgactggca gaactacacc cccggccccg gcatccgcta cccctgtgc 420

ttcggctggt gcttcaagct ggtgcccgtg gagccccgcg aggtggagga ggccaacgag 480
ggcgagaaca actgcctgct gcaccccatg tccagcacg gcatggacga cgaggagcgc 540
gaggtgctga tgtggaagtt cgactcctcc ctggccccgc gccacatcgc ccgcgagctg 600
caccgccgact tctacaagga ctgctaa 627

<210> 210

<211> 206

<212> PRT

<213> Human immunodeficiency virus

<400> 210

Met Gly Gly Lys Trp Ser Lys Ser Ser Ile Val Gly Trp Pro Asp Ile

1 5 10 15

Arg Glu Arg Met Arg Arg Ala Pro Pro Ala Ala Glu Gly Val Gly Ala

20 25 30

Val Ser Gln Asp Leu Glu Asn Arg Gly Ala Ile Thr Ser Ser Asn Thr
 35 40 45
 Arg Ala Asn Asn Pro Asp Leu Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 55 60
 Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr
 65 70 75 80
 Tyr Lys Gly Ala Leu Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly

85 90 95
 Leu Glu Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu
 100 105 110
 Trp Val Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 115 120 125
 Pro Gly Pro Gly Ile Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys
 130 135 140
 Leu Val Pro Val Asp Pro Glu Glu Val Glu Lys Ala Asn Glu Gly Glu
 145 150 155 160

Asn Asn Cys Leu Leu His Pro Met Ser Gln His Gly Met Glu Asp Glu
 165 170 175
 Asp Arg Glu Val Leu Met Trp Lys Phe Asp Ser Arg Leu Ala Leu Arg
 180 185 190
 His Ile Ala Arg Glu Lys His Pro Glu Phe Tyr Gln Asp Cys
 195 200 205

<210> 211

<211> 621

<212> DNA

<213> Human immunodeficiency virus

<400> 211

atgggcggca agtgggtccaa gtccctccatc gtgggctggc ccgacatccg cgagcgcatg 60

cgccgcgccc cccccgccgc cgagggcggtg ggcgcctgtt cccaggacct ggagaaccgc 120

ggcgccatca cctcctccaa caccgcgcc aacaaccccg acctggcctg gctggaggcc 180

caggaggagg aggaggtggg cttccccgtg cgccccagg tgcccctgcg ccccatgacc 240

tacaaggcg ccttgacct gtccacttc ctgaaggaga agggcggcct ggagggcctg 300

atctactcca agaagcgcca ggagatcctg gacctgtggg tgtaccacac ccagggtac 360
 ttccccgact ggcagaacta caccctcggc cccggcatcc gctacccct gaccttcggc 420
 tgggtgttca agctggtgcc cgtggacccc gaggaggtgg agaaggcaa cgaggcgag 480

aacaactgcc tgcgcaccc catgtcccag cacggcatgg aggacgagga ccgcgaggtg 540
 ctgatgtgga agttcgactc ccgctggcc ctgcgccaca tcgccgcga gaagcacc 600
 gagttctacc aggactgcta a 621

<210> 212
 <211> 207
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 212

Met Gly Gly Lys Trp Ser Lys Cys Ser Ile Val Gly Trp Pro Glu Val
 1 5 10 15
 Arg Glu Arg Ile Arg Arg Thr Pro Pro Ala Ala Val Gly Val Gly Ala

20 25 30
 Val Ser Gln Asp Leu Ala Lys His Gly Ala Ile Thr Ser Ser Asn Thr
 35 40 45

Ala Ala Asn Asn Pro Asp Cys Ala Trp Leu Glu Ala Gln Glu Glu Asp
 50 55 60

Ser Glu Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met
 65 70 75 80
 Thr Tyr Lys Gly Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly
 85 90 95

Gly Leu Asp Gly Leu Ile Tyr Ser Lys Gln Arg Gln Asp Ile Leu Asp
 100 105 110

Leu Trp Val Tyr Asn Thr Gln Gly Phe Phe Pro Asp Trp Gln Asn Tyr
 115 120 125
 Thr Pro Gly Pro Gly Thr Arg Tyr Pro Leu Thr Phe Gly Trp Cys Phe
 130 135 140

Lys Leu Glu Pro Val Asp Pro Ala Glu Val Glu Glu Ala Thr Lys Gly
 145 150 155 160

Glu Asn Asn Ser Leu Leu His Pro Ile Cys Gln His Gly Met Glu Asp

165 170 175
Ala Asp Asn Glu Val Leu Ile Trp Arg Phe Asp Ser Ser Leu Ala Arg

180 185 190
Arg His Ile Ala Arg Glu Leu His Pro Asp Phe Tyr Lys Asp Cys

195 200 205

<210> 213

<211> 624

<212> DNA

<213> Human immunodeficiency virus

<400> 213

atgggcggca agtgggtccaa gtgctccatc gtgggctggc ccgagggtgcg cgagcgcac 60

cgccgcaccc cccccgccgc cgtgggcgtg ggcgccgtgt cccaggacct ggccaagcac 120

ggcgccatca ctctctccaa caccgccgcc aacaaccccg actgcgcctg gctggaggcc 180

caggaggagg actccgaggt gggcttcccc gtgcgcccc aggtgccctt gcgccccatg 240

acctacaagg gcgccttcga cctgtccttc ttctgaagg agaagggcgg cctggacggc 300

ctgatctact ccaagcagcg ccaggacatc ctggacctgt ggggtgtacaa caccagggc 360

ttcttccccg actggcagaa ctacacccc ggccccggca cccgctaccc cctgaccttc 420

ggctgggtgt tcaagctgga gcccgtggac cccgccgagg tggaggaggc caccaagggc 480

gagaacaact cctgtctgca ccccatctgc cagcacggca tggaggacgc cgacaacgag 540

gtgctgatct ggcgcttcga ctctccctg gcccgccgcc acatcgcccg cgagctgcac 600

cccgacttct acaaggactg ctaa 624

<210> 214

<211> 1002

<212> PRT

<213> Human immunodeficiency virus

<400> 214

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Glu Phe

1 5 10 15

Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Arg

20 25 30

Val Arg Gly Gly Asp Asn Pro Leu Ser Glu Ala Gly Ala Glu Arg Gln

35 40 45

Gly Thr Val Ser Leu Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro

50 55 60

Leu Val Thr Val Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu Asp

65 70 75 80

Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asn Leu Pro Gly Lys

85 90 95

Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg

100 105 110

Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly

115 120 125

Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Met

130 135 140

Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu

145 150 155 160

Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys

165 170 175

Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile Cys

180 185 190

Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn

195 200 205

Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys

210 215 220

Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp

225 230 235 240

Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys

245 250 255

Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val

260 265 270

Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser

275 280 285
 Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro
 290 295 300
 Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys
 305 310 315 320
 Ile Leu Glu Pro Phe Arg Thr Gln Asn Pro Glu Ile Val Ile Tyr Gln

 325 330 335
 Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His
 340 345 350
 Arg Thr Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly Phe
 355 360 365
 Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
 370 375 380
 Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu
 385 390 395 400

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

 465 470 475 480
 Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile Tyr
 485 490 495
 Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg
 500 505 510
 Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys
 515 520 525
 Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg

530 535 540
 Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp
 545 550 555 560
 Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu
 565 570 575
 Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu
 580 585 590
 Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys
 595 600 605
 Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu Thr

 610 615 620
 Glu Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala Leu
 625 630 635 640
 Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
 645 650 655
 Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val
 660 665 670
 Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser
 675 680 685

 Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
 690 695 700
 Leu Val Ser Thr Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
 705 710 715 720
 Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met
 725 730 735
 Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala
 740 745 750
 Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val

 755 760 765
 Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly
 770 775 780

Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala
 785 790 795 800
 Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu
 805 810 815
 Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly
 820 825 830

 Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly
 835 840 845
 Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
 850 855 860
 Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
 865 870 875 880
 Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
 885 890 895
 His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu

 900 905 910
 Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln
 915 920 925
 Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
 930 935 940
 Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
 945 950 955 960
 Gly Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro Arg
 965 970 975

 Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp
 980 985 990
 Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000

 <210> 215
 <211> 3009
 <212> DNA
 <213> Human immunodeficiency virus

<400> 215

ttcttccgcg agaacctggc cttccagcag ggcgaggccc gcgagttctc ctccgagcag	60
acccgcgcca actccccac ctcccgagag ctgcgcgtgc gcggcggcga caacccccctg	120
tccgaggccg gcgccgagcg ccagggcacc gtgtccctgt ccttccccca gatcacccctg	180
tggcagcgcc ccttggtagc cgtgaagatc ggcggccagc tgaaggaggc cctgctggac	240
accggcgccg acgacaccgt gctggaggag atcaacctgc ccggcaagtg gaagcccaag	300
atgatcgcg gcacggcggt cttcatcaag gtgcgccagt acgaccagat cctgatcgag	360
atctgcggca agaaggccat cggcacctg ctggtgggcc ccacccccgt gaacatcatc	420
ggccgcaaca tgctgacca gatcggtgc accctgaact tccccatctc ccccatcgag	480
accgtgcccg tgaagctgaa gcccggcatg gacggcccca aggtgaagca gtggcccctg	540
accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggcaag	600
atctcaaga tcggccccga gaaccctac aacaccccca tcttcgcat caagaagaag	660
gactccacca agtggcgcaa gctgtggtag ttccgcgagc tgaacaagcg caccaggac	720
ttctgggagg tgcagctggg cateccccac ccgcccggcc tgaagaagaa gaagtccgtg	780
accgtgctgg acgtggcgga cgcctacttc tccgtgcccc tggacgagga cttccgcaag	840
tacaccgct tcaccatccc ctccatcaac aacgagaccc ccggcatccg ctaccagtac	900
aacgtgctgc ccagggtg gaagggtcc ccgccatct tccagtcctc catgaccaag	960
atcctggagc ctttcgcac ccagaacccc gagatcgtga tctaccagta catggacgac	1020
ctgtacgtgg gtcccgacct ggagatcggc cagcacgca ccaagatcga ggagctgcgc	1080
gagcacctgc tgcgctgggg cttcaccacc cccgacaaga agcaccagaa ggagccccc	1140
ttctgtgga tgggctacga gctgcacccc gacaagtga cctgcagcc catccagctg	1200
cccgagaagg actcctggac cgtgaacgac atccagaagc tgggtgggcaa gctgaactgg	1260
gcctcccaga tctaccccg catcaaggta aagcagctgt gcaagctgct gcgcggcgcc	1320
aaggccctga ccgacatcgt gcccctgacc gaggaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg accctccaa ggacctgac	1440
gccgagatcc agaagcaggg ccaggaccag tggacctacc agatctacca ggagcccttc	1500
aagaacctga agaccggcaa gtacgccaag atgcgtccg cccacaccaa cgacgtgaag	1560
cagctgaccg aggcctgca gaagatgcc accgagtcca tcgtgatctg gggcaagacc	1620
cccaagtacc gctgcccac ccagaaggag acctgggaga cctggtggac cgagtactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacaccc cccccctgt gaagctgtgg	1740
taccagctgg agaaggagcc catcgtgggc gccgagacct tctacgtgga cggcgccgcc	1800

aaccgcgaga ccaagctggg caaggccggc tacgtgaccg accgcggcgg ccagaaggtg 1860

gtgtccctga ccgagaccac caaccagaag accgagctgc aggccatcca cctggccctg 1920

caggactccg gtcctgaggt gaacatcgtg accgactccc agtacgcctt gggcatcatc 1980

caggcccagc ccgacaagtc cgagtcggag ctggtgaacc agatcatcga gcagctgac 2040

aagaaggaga aggtgtacct gtcctgggtg cccgcccaagg ggcgaacgag 2100

caggtggaca agctggtgtc caccggcatc cgcaaggtgc tgttcctgga cggcatcgac 2160

aaggcccagg aggagcacga gaagtaccac tccaactggc gcgccatggc ctccgacttc 2220

aacctgcccc ccatcgtggc caaggagatc gtggcctcct gcgacaagt ccagctgaag 2280

ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc 2340

cacctggagg gcaagatcat cctggtggcc gtgcacgtgg cctccggcta catcgaggcc 2400

gaggtgatcc ccgccgagac cggccaggag accgcctact tcctcctgaa gctggccggc 2460

cgctggcccc tgaaggatgat ccacaccgac aacggctcca acttcacctc cgccgcctg 2520

aaggccgctt gctggtgggc cggcatccag caggagtctg gcatccccta caacccccag 2580

tcccaggcgg tgggtggagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc 2640

gaccaggccg agcacctgaa gaccgccgtg cagatggcgg tgttcatcca caacttcaag 2700

cgcaagggcg gcatcgccgg ctactccgcc ggcgagcgca tcacgacat catcgccacc 2760

gacatccaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgcgtgtac 2820

taccgcgact cccgcgaccc catctggaag gggccgccca agctgctgtg gaaggcgag 2880

ggcgccgtgg tgatccagga caactccgag atcaaggtgg tgccccgcg caaggccaag 2940

atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgg ccgccaggac 3000

gaggactaa 3009

<210> 216

<211> 1002

<212> PRT

<213> Human immunodeficiency virus

<400> 216

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Glu Phe

1 5 10 15

Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Arg

20 25 30

Val Arg Gly Gly Asp Asn Pro Leu Ser Glu Ala Gly Ala Glu Arg Gln

35 40 45
 Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro
 50 55 60

 Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Arg Glu Ala Leu Leu Asp
 65 70 75 80
 Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asn Leu Pro Gly Lys
 85 90 95
 Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg
 100 105 110
 Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly
 115 120 125
 Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Met

 130 135 140
 Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu
 145 150 155 160
 Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys
 165 170 175
 Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile Cys
 180 185 190
 Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn
 195 200 205

 Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys
 210 215 220
 Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp
 225 230 235 240
 Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys
 245 250 255
 Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val
 260 265 270
 Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser

 275 280 285

Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro
290 295 300

Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys
305 310 315 320

Ile Leu Glu Pro Phe Arg Thr Lys Asn Pro Glu Ile Val Ile Tyr Gln
325 330 335

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His
340 345 350

Arg Ala Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly Phe
355 360 365

Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
370 375 380

Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu
385 390 395 400

Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
405 410 415

Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys Gln
420 425 430

Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro
435 440 445

Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
450 455 460

Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
465 470 475 480

Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile Tyr
485 490 495

Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg
500 505 510

Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys
515 520 525

Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg
530 535 540

Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp
 545 550 555 560
 Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu

 565 570 575
 Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu
 580 585 590
 Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys
 595 600 605
 Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu Thr
 610 615 620
 Glu Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala Leu
 625 630 635 640

 Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
 645 650 655
 Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val
 660 665 670
 Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser
 675 680 685
 Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
 690 695 700
 Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp

 705 710 715 720
 Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met
 725 730 735
 Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala
 740 745 750
 Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
 755 760 765
 Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly
 770 775 780

 Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala

785 790 795 800
 Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu
 805 810 815
 Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly
 820 825 830
 Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly
 835 840 845
 Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val

 850 855 860
 Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
 865 870 875 880
 Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
 885 890 895
 His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
 900 905 910
 Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln
 915 920 925

 Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
 930 935 940
 Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
 945 950 955 960
 Gly Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro Arg
 965 970 975
 Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp
 980 985 990
 Asp Cys Val Ala Gly Arg Gln Asp Glu Asp

 995 1000
 <210> 217
 <211> 1003
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 217

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Lys Phe
 1 5 10 15
 Ser Ser Glu Gln Thr Gly Ala Asn Ser Pro Thr Ser Arg Asp Leu Trp
 20 25 30
 Asp Gly Gly Arg Asp Ser Leu Pro Ser Glu Ala Gly Ala Glu Arg Gln
 35 40 45

 Gly Thr Gly Pro Thr Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
 50 55 60
 Pro Leu Val Thr Val Arg Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
 65 70 75 80
 Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly
 85 90 95
 Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
 100 105 110
 Lys Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile

 115 120 125
 Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
 130 135 140
 Met Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
 145 150 155 160
 Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 165 170 175
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile
 180 185 190

 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 195 200 205
 Asn Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp Ser Thr
 210 215 220
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 225 230 235 240
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 245 250 255

Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser

260 265 270

Val Pro Leu Asp Glu Ser Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro

275 280 285

Ser Thr Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu

290 295 300

Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

305 310 315 320

Lys Ile Leu Glu Pro Phe Arg Ser Lys Asn Pro Glu Ile Ile Ile Tyr

325 330 335

Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln

340 345 350

His Arg Thr Lys Ile Glu Glu Leu Arg Ala His Leu Leu Ser Trp Gly

355 360 365

Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp

370 375 380

Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Glu

385 390 395 400

Leu Pro Glu Lys Glu Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val

405 410 415

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Lys

420 425 430

Gln Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val

435 440 445

Thr Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile

450 455 460

Leu Lys Asp Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu

465 470 475 480

Ile Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile

485 490 495

Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Lys

500 505 510
 Arg Ser Ala His Thr Asn Asp Val Lys Gln Leu Ala Glu Val Val Gln
 515 520 525
 Lys Val Val Met Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Met Asp Tyr

 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Asp Pro Ile Val Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
 610 615 620

 Thr Glu Thr Thr Asn Gln Lys Thr Glu Leu His Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Leu
 660 665 670
 Val Asn Gln Ile Ile Glu Lys Leu Ile Gly Lys Asp Lys Val Tyr Leu
 675 680 685
 Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp

 690 695 700
 Lys Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Arg Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln

755 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780
 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 805 810 815
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Val Val His Thr Asp Asn
 820 825 830
 Gly Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala
 835 840 845
 Asn Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
 Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 865 870 875 880
 Arg Glu Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910
 Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 915 920 925
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 930 935 940
 Ser Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 945 950 955 960
 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 965 970 975
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 980 985 990
 Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000

<210> 218
 <211> 3009
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 218

tttttccgcg agaacctggc cttccagcag ggcgaggccc gcgagtcttc ctccgagcag 60
 acccgcgcca actccccac ctcccgcgag ctgcgcgtgc gcggcggcga caaccacctg 120
 tccgaggccg gcgccgagcg ccagggcacc gtgtccttct ccttccccca gatcacctg 180
 tggcagcgcc cctggtgac catcaagatc ggccggccagc tgcgcgaggc cctgctggac 240

accggcgccg acgacaccgt gctggaggag atcaacctgc ccggcaagtg gaagcccaag 300
 atgatcgcg gcacgcgcg cttcatcaag gtgcgccagt acgaccagat cctgatcgag 360
 atctgcggca agaaggccat cggcaccgtg ctggtgggcc ccacccccgt gaacatcatc 420
 ggccgcaaca tgctgacca gatcggtgc accctgaact tccccatctc ccccatcgag 480
 accgtgccc tgaagctgaa gcccgcatg gacggcccca aggtgaagca gtggccctg 540
 accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggcaag 600
 atctccaaga tcggccccga gaaccctac aacacccccg tgttcgcat caagaagaag 660

gactccacca agtggcgcaa gctggtggac ttccgcgagc tgaacaagcg caccagggac 720
 ttctgggagg tgcagctggg catccccac ccgccggcc tgaagaagaa gaagtccgtg 780
 accgtgctgg acgtggcgga cgcctacttc tccgtgcccc tggacgagga ctccgcaag 840
 tacaccgct tcaccatccc ctccatcaac aacgagaccc ccggcatccg ctaccgtac 900
 aacgtgctgc ccagggtg gaagggtcc ccgccatct tccagtctc catgaccaag 960
 atctggagc cttccgcac caagaacccc gagatcgtga tctaccagta catggacgac 1020
 ctgtactgg gctccgacct ggagatcggc cagcaccgcg ccaagatcga ggagctgcgc 1080

gagcacctgc tgcgtgggg cttcaccacc ccgacaaga agcaccagaa ggagccccc 1140
 ttctgtgga tgggctacga gctgcacccc gacaagtga ccgtgcagcc catccagctg 1200
 ccgagaagg actcctggac cgtgaacgac atccagaagc tgggtgggcaa gctgaactgg 1260
 gcctcccaga tctacccgg catcaagtg aagcagctgt gcaagctgct gcgcggcgcc 1320
 aaggccctga ccgacatct gccctgacc gaggaggccg agctggagct ggccgagaac 1380
 cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg acccctcaa ggacctgatc 1440
 gccgagatcc agaagcagg ccaggaccag tggacctacc agatctacca ggagcccttc 1500

aagaacctga agaccggcaa gtacccaag atgcgctccg cccacaccaa cgactgaag 1560

cagctgaccg aggccgtgca gaagatcgcc accgagtcca tctgatctg gggcaagacc	1620
cccaagtcc gcctgccat ccagaaggag acctgggaga cctggtggac cgagtactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacaccc ccccccctgt gaagctgtgg	1740
taccagctgg agaaggagcc catcgtgggc gccgagacct tctacgtgga cggcgccgcc	1800
aaccgcgaga ccaagctggg caaggccggc tacgtgaccg accgcggccg ccagaaggtg	1860
gtgtccctga ccgagaccac caaccagaag accgagctgc aggccatcca cctggccctg	1920
caggactccg gctccgaggt gaacatcgtg accgactccc agtacgccct gggcatcatc	1980
caggcccagc ccgacaagtc cgagtcgag ctggtgaacc agatcatcga gcagctgac	2040
aagaaggaga aggtgtacct gtccctgggtg cccgccaca agggcatcgg cggcaacgag	2100
caggtggaca agctggtgtc ctccggcatc cgcaaggtgc tgttcctgga cggcatcgac	2160
aagggccagg aggagcacga gaagtaccac tccaactggc gcgccatggc ctccgacttc	2220
aacctgcccc ccgtggtggc caaggagatc gtggcctcct gcgacaagtg ccagctgaag	2280
ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc	2340
cacctggagg gcaaggtgat cctggtggcc gtgcacgtgg cctccggcta catcgaggcc	2400
gaggtgatcc ccgccgagac cggccaggag accgcctact tcctcctgaa gctggccggc	2460
cgctggcccc tgaaggtgat ccacaccgac aacggctcca acttcacctc cgccgccgtg	2520
aaggccgcct gctggtgggc cggcatccag caggagttcg gcatccccta caacccccag	2580
tcccaggcgc tgggtggagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc	2640
gaccaggccg agcacctgaa gaccgccgtg cagatggccg tgttcatcca caacttcaag	2700
cgcaaggcgc gcatcgccgg ctactccgcc ggcgagcgca tcctcgacat catcgccacc	2760
gacatccaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgctgtac	2820
taccgcgact ccgcgaccc catctggaag ggcccccca agctgctgtg gaaggcgag	2880
ggcgccgtgg tgatccagga caactccgag atcaaggtgg tgcccccg caaggccaag	2940
atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgg ccgccaggac	3000
gaggactaa	3009
<210> 219	
<211> 3012	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 219	
ttcttccgcg agaacctggc cttccagcag ggcgaggccc gcaagttctc ctccgagcag	60

accggcgcca actccccac ctcccgcgac ctgtgggacg gcggccgcga ctccctgccc	120
tccgaggcgg gcgccgagcg ccagggcacc ggccccacct tctccttccc ccagatcacc	180
ctgtggcagc gccccctggt gaccgtgcgc atcggcggcc agctgaagga ggccctgctg	240
gacaccggcg ccgacgacac cgtgctggag gacatcaacc tgcccggcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggcttcac c aaggtgaagc agtacgacca gatcctgac	360
gagatctgcg gcaagaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacac	420
atcgcccgca acatgctgac ccagatcggc tgcacctga acttccccat ctccccac	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcgcccc cgagaacccc tacaacaccc ccatcttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcacccag	720
gacttctggg aggtgcagct gggcatcccc cccccgcg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccctggacga gtccttccgc	840
aagtacaccg ctttaccat cccctccacc aacaacgaga cccccggcat ccgtaccag	900
tacaacgtgc tgccccaggg ctggaagggc tcccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg ctccaagaac cccgagatca tcatttacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgcgcccacc tgctgtcctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccacgag	1200
ctgcccgaga aggagtctg gaccgtgaac gacatccaga agctgggtggg caagctgaac	1260
tgggcctccc agatctacg cggcatcaag gtgaagcagc tgtgcaagct gctgcgcggc	1320
gccaaggccc tgaccgacat cgtgacctg accgaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga ccccgatcac ggctgtact acgaccctc caaggacctg	1440
atcgccgaga tccagaagca gggccaggac cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgc cgcaagcgt cgcgccacac caacgacgtg	1560
aagcagctgg ccgaggtggt gcagaagggt gtgatggagt ccatcgtgat ctggggcaag	1620
acccccaaat tcaagctgcc catccagaag gagacctggg agacctggtg gatggactac	1680
tggcaggcca cctggatccc cgagtgggag ttctgaaca cccccccct ggtgaagctg	1740
tggtaccagc tggagaagga ccccatcgtg ggcccgaga ccttctacgt ggacggcgcc	1800
gccaaccgcg agaccaagct gggcaaggcc ggctacgtga ccgaccgcg ccgccagaag	1860
gtggtgtccc tgaccgagac caccaaccag aagaccgagc tgcacgcat ccacctggcc	1920

ctgcaggact ccggctccga ggtgaacatc gtgaccgact cccagtacgc cctgggcatc 1980
atccaggccc agccccgaccg ctccgagtcc gagctgggtga accagatcat cgagaagctg 2040
atcggcaagg acaaggtgta cctgtcctgg gtgccccccc acaagggcat cggcggcaac 2100
gagcaggtgg acaagctggt gtcctccggc atccgcaagg tgctgttctt ggacggcatc 2160

gacaaggccc aggaggagca cgagcgctac cactccaact ggcgcgccat ggcctccgac 2220
ttcaacctgc ccccatcgt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg 2280
aaggcgagg ccatgcacgg ccaggtggac tgctccccg gcacatggca gctggactgc 2340
accacactgg agggcaaggt gatcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400
gccgaggtga tccccgccga gaccggccag gagaccgctt acttcctgct gaagctggcc 2460
ggccgctggc ccgtgaaggt ggtgcacacc gacaacggct ccaacttcac ctccgccgcc 2520
gtgaaggccg cctgctggtg ggccaacatc cagcaggagt tcggcatccc ctacaacccc 2580

cagtcccagg gcgtgggtga gtccatgaac aaggagctga agaagatcat cggccagggtg 2640
cgcgagcagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc 2700
aagcgcaagg gcggcatcgg cggtactacc gccggcgagc gcacatcga catcatgcc 2760
accgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa ctcccgctg 2820
tactaccgag actcccgca ccccatctgg aagggtcccg ccaagctgct gtggaagggc 2880
gaggcgccg ttgtgatcca ggacaactcc gacatcaagg tggcgcccg ccgcaaggcc 2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag 3000

gacgaggact aa 3012

<210> 220

<211> 1003

<212> PRT

<213> Human immunodeficiency virus

<400> 220

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Lys Phe

1 5 10 15

Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Trp

20 25 30

Asp Gly Gly Arg Asp Ser Leu Leu Ser Glu Ala Gly Ala Glu Arg Gln

35 40 45

Gly Thr Val Pro Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg

50 55 60
 Pro Leu Val Thr Val Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
 65 70 75 80
 Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly
 85 90 95
 Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
 100 105 110
 Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile

 115 120 125
 Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
 130 135 140
 Met Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
 145 150 155 160
 Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 165 170 175
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile
 180 185 190

 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 195 200 205
 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 210 215 220
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 225 230 235 240
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 245 250 255
 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser

 260 265 270
 Val Pro Leu Asp Glu Ser Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 275 280 285
 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 290 295 300
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

305 310 315 320
 Lys Ile Leu Glu Pro Phe Arg Ser Lys Asn Pro Glu Ile Val Ile Tyr
 325 330 335

 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 340 345 350
 His Arg Ala Lys Ile Glu Glu Leu Arg Ala His Leu Leu Ser Trp Gly
 355 360 365
 Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 370 375 380
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Lys
 385 390 395 400
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val

 405 410 415
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Lys
 420 425 430
 Gln Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val
 435 440 445
 Thr Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 450 455 460
 Leu Lys Asp Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 465 470 475 480

 Val Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile
 485 490 495
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Lys
 500 505 510
 Arg Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Val Val Gln
 515 520 525
 Lys Val Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Arg Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Met Glu Tyr

 545 550 555 560

Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ala Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
 610 615 620

 Thr Glu Thr Thr Asn Gln Lys Thr Glu Leu His Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Leu
 660 665 670
 Val Asn Gln Ile Ile Glu Lys Leu Ile Glu Lys Glu Lys Val Tyr Leu
 675 680 685
 Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp

 690 695 700
 Lys Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765

 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780
 Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
 805 810 815

Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Val Val His Thr Asp Asn

820 825 830

Gly Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala

835 840 845

Asn Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly

850 855 860

Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val

865 870 875 880

Arg Glu Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe

885 890 895

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly

900 905 910

Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu

915 920 925

Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp

930 935 940

Ser Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly

945 950 955 960

Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro

965 970 975

Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly

980 985 990

Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp

995 1000

<210> 221

<211> 3012

<212> DNA

<213> Human immunodeficiency virus

<400> 221

ttcttcgcgcg agaacctggc cttccagcag ggcgaggccc gcaagtcttc ctccgagcag 60

accgcgcgcca actccccac ctcccgcgag ctgtgggacg gcggccgcga ctccctgctg 120

tccgaggcgcg gcgccgagcg ccagggcacc gtgccctcct tctccttccc ccagatcacc 180

ctgtggcagc gccccctggt gaccgtgaag atcggcggcc agctgaagga ggccctgctg	240
gacaccggcg ccgacgacac cgtgctggag gacatcaacc tgcccggcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggcttcac c aaggtgcgcc agtacgacca gatcctgac	360
gagatctgcg gcaagaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacatc	420
atcggcgcca acatgctgac ccagatcggc tgcacctga acttccccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcggccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcacccag	720
gacttctggg aggtgcagct gggcatcccc caccgcccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccttggaaga gtccttccgc	840
aagtacaccg ccttcacat cccctccatc aacaacgaga ccccgccat ccgctaccag	900
tacaacgtgc tgccccaggg ctggaagggc tccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg ctccaagaac cccgagatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcgccaagat cgaggagctg	1080
cgcgcccacc tctgtcctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gccatcaag	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctggtggg caagctgaac	1260
tgggcctccc agatctacg cggcataag gtgaagcagc tgtgcaagct gctgcgggc	1320
gccaaagccc tgaccgacat cgtgacctg accgaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga cccgtgcac ggctgtact acgaccctc caaggacctg	1440
gtggccgaga tccagaagca gggccaggac cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgc aagaagcgt ccgcccacac caacgacgtg	1560
aagcagctga ccgaggtggt gcagaagggt gccaccgagt ccatcgtgat ctggggcaag	1620
acccccaagt tccgctgcc catccagaag gagacctggg agacctggtg gatggagtac	1680
tggcaggcca cctggatccc cgagtgggag ttcgtgaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaagga gcccatcgcc ggccgagaga ccttctacgt ggacggcgcc	1800
gccaaaccgc agaccaagct gggcaaggcc ggctacgtga ccgaccgagg ccgccagaag	1860
gtggtgtccc tgaccgagac caccaaccag aagaccgagc tgcacgcat ccacctggcc	1920

ctgcaggact ccggctccga ggtgaacatc gtgaccgact cccagtacgc cctgggcatc 1980
atccaggccc agcccgaccg ctccgagtcc gagctggtga accagatcat cgagaagctg 2040
atcgagaagg agaaggtgta cctgtcctgg gtgccccccc acaaggcat cggcggcaac 2100
gagcaggtagt acaagctggt gtcctccggc atccgcaagg tgctgttctt ggacggcatc 2160
gacaaggccc aggaggagca cgagaagtac cactccaact ggcgcccat ggctccgac 2220
ttcaacctgc ccccatcgt ggccaaggag atcgtggcct cctgcgaca gtgccagctg 2280
aaggcgagg ccatgcacgg ccaggtggac tgctccccg gcatctggca gctggactgc 2340

accacactgg agggcaaggt gatcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400
gccgaggtagt tccccgccga gaccggccag gagaccgctt acttctgct gaagctggcc 2460
ggccgctggc ccgtgaaggt ggtgcacacc gacaacggct ccaacttcac ctccgccgc 2520
gtgaaggccg cctgctggtg ggccaacatc cagcaggagt tcggcatccc ctacaacccc 2580
cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccagggtg 2640
cgcgagcagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc 2700
aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcatcatcga catcatcgcc 2760

accgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa cttccgcgtg 2820
tactaccgcg actccccgca ccccatctgg aaggccccg ccaagctgct gtggaagggc 2880
gagggcgccg tggtagatcca ggacaactcc gacatcaagg tggtagcccg ccgcaaggcc 2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag 3000
gacgaggact aa 3012

<210> 222

<211> 1003

<212> PRT

<213> Human immunodeficiency virus

<400> 222

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Arg Glu Ala Arg Lys Phe

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

Pro Leu Val Thr Val Lys Ile Glu Gly Gln Leu Arg Glu Ala Leu Leu
65 70 75 80

Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly
85 90 95

Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
100 105 110

Arg Gln Tyr Asp Gln Ile Ala Ile Glu Ile Cys Gly Lys Arg Ala Ile
115 120 125

Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
130 135 140

Met Leu Val Gln Leu Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile

145 150 155 160

Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
165 170 175

Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile
180 185 190

Cys Lys Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
195 200 205

Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
210 215 220

Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
225 230 235 240

Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
245 250 255

Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
260 265 270

Val Pro Leu His Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
275 280 285

Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu

290 295 300

Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

305 310 315 320
 Lys Ile Leu Glu Pro Phe Arg Ser Lys Asn Pro Glu Met Val Ile Tyr
 325 330 335
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 340 345 350
 His Arg Ala Lys Ile Glu Glu Leu Arg Ala His Leu Leu Arg Trp Gly
 355 360 365

 Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 370 375 380
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Lys
 385 390 395 400
 Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 405 410 415
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Lys
 420 425 430
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Asp Ile Val

 435 440 445
 Thr Leu Thr Lys Glu Ala Glu Leu Glu Leu Glu Glu Asn Arg Glu Ile
 450 455 460
 Leu Lys Asn Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 465 470 475 480
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile
 485 490 495
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Arg
 500 505 510

 Lys Ser Thr His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 515 520 525
 Lys Ile Ala Ile Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Arg Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro

565 570 575
Leu Val Lys Leu Trp Tyr Gln Leu Glu Thr Glu Pro Ile Ala Gly Ala

580 585 590
Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly

595 600 605
Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Ile Val Ser Leu

610 615 620
Thr Glu Thr Thr Asn Gln Lys Thr Glu Leu His Ala Ile Tyr Leu Ala

625 630 635 640
Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr

645 650 655
Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Leu

660 665 670
Val Asn Gln Ile Ile Glu Lys Leu Ile Glu Lys Glu Arg Val Tyr Leu

675 680 685
Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp

690 695 700
Lys Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile

705 710 715 720
Asp Lys Ala Gln Glu Glu His Glu Arg Tyr His Ser Asn Trp Arg Ala

725 730 735
Met Ala His Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val

740 745 750
Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln

755 760 765
Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu

770 775 780
Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu

785 790 795 800
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile

805 810 815
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile

Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn
 820 825 830
 Gly Pro Asn Phe Thr Ser Ala Thr Val Lys Ala Ala Cys Trp Trp Ala
 835 840 845
 Gly Val Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
 Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val

 865 870 875 880
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910
 Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 915 920 925
 Gln Lys Gln Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 930 935 940

 Ser Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 945 950 955 960
 Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
 965 970 975
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 980 985 990
 Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000
 <210> 223
 <211> 1003
 <212> PRT

 <213> Human immunodeficiency virus
 <400> 223
 Phe Phe Arg Glu Asp Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe
 1 5 10 15
 Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln

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

Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro

275 280 285

Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu

290 295 300

Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr

305 310 315 320

Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr

325 330 335

Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln

340 345 350

His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly

355 360 365

Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp

370 375 380

Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val

385 390 395 400

Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val

405 410 415

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Lys

420 425 430

Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile

435 440 445

Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile

450 455 460

Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu

465 470 475 480

Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile

485 490 495

Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met

500 505 510

Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln

515 520 525
 Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro

 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
 610 615 620
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala
 625 630 635 640

 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile

 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu

770	775	780
Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu		
785	790	795
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu		
805	810	815
Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn		
820	825	830
Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala		
835	840	845
Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly		
850	855	860
Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val		
865	870	875
Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe		
885	890	895
Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly		
900	905	910
Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu		
915	920	925
Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp		
930	935	940
Ser Arg Asp Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly		
945	950	955
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro		
965	970	975
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly		
980	985	990
Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp		
995	1000	
<210>	224	
<211>	3012	

<212> DNA

<213> Human immunodeficiency virus

<400> 224

ttcttccgcg agaacctggc cttccagcag cgcgaggccc gcaagttctc ctccgagcag	60
aaccgcgcca actccccac ctcccgcgag ctgcgcaacg gcggccgcga caacctgctg	120
tccgaggcgg gcgccgagga gcagggcacc gtgcactcct gcaacttccc ccagatcacc	180
ctgtggcagc gccccctggg gaccgtgaag atcgagggcc agctgcgca ggccctgctg	240
gacaccggcg ccgacgacac cgtgctggag gacatcaacc tgcccggcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggcttcac aaggtgcgcc agtacgacca gatcgccatc	360
gagatctgcg gcaagcgcg catcggcacc gtgctggtgg gccccacccc cgtgaacatc	420
atcgccgca acatgctggt gcagctgggc tgcacctga acttccccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca aggagatgga gaaggagggc	600
aagatctcca agatcgcccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcacccag	720
gacttctggg aggtgcagct gggcatcccc cccccgccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc cctgcacga ggacttccgc	840
aagtacaccg ccttcacat cccctccatc aacaacgaga ccccgccat ccgctaccag	900
tacaacgtgc tgccccaggg ctggaagggc tccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg ctccaagaac cccgagatgg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcgccaagat cgaggagctg	1080
cgcgcccacc tgctgcgctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccataag	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctggtggg caagctgaac	1260
tgggcctccc agatctacg cggcatacag gtgaagcagc tgtgcaagct gctgcgcggc	1320
accaaggccc tgaccgacat cgtgaccctg accaaggagg ccgagctgga gctggaggag	1380
aaccgcgaga tctgaagaa cccgtgcac ggcgtgtact acgaccctc caaggacctg	1440
atcgccgaga tccagaagca gggccaggac cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgcc aagcgcaagt ccaccacac caacgacgtg	1560
aagcagctga ccgaggccgt gcagaagatc gccatcgagt ccatcgtgat ctggggcaag	1620
acccccaaat tccgcctgcc catccagaag gagacctggg agacctggtg gaccgagtac	1680

tggcaggcca cctggatccc cgagtgggag ttcgtgaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaccga gcccatcgcc ggcgccgaga ctttctacgt ggacggcgcc	1800
gccaaaccgag agaccaagct gggcaaggcc ggtctacgtga ccgaccgcgg ccgccagaag	1860
atcgtgtccc tgaccgagac caccaaccag aagaccgagc tgcacgccat ctacctggcc	1920
ctgcaggact ccggcctgga ggtgaacatc gtgaccgact ccagtagc cctgggcatc	1980
atccaggccc agcccgaccg ctccgagtcc gagctggtga accagatcat cgagaagctg	2040
atcgagaagg agcgctgtga cctgtcctgg gtccccgcc acaaggcat cggcggcaac	2100
gagcaggtgg acaagctggt gtccctccgc atccgcaagg tgcgtttcct ggacggcatc	2160
gacaaggccc agggaggaga cgagcgctac cactccaact ggcgcccat ggcccacgac	2220
ttcaacctgc ccccatcgt ggccaaggag atcgtggcct cctgcgaca gtgccagctg	2280
aaggcgagg ccatgcagg ccaggtggac tgcctccccg gcatctggca gctggactgc	2340
accacactgg agggcaaggt gatcctggtg gccgtgcacg tggcctccgg ctacatcgag	2400
gccgaggtga tccccccga gaccggccag gagaccgct acttcatcct gaagctggcc	2460
ggcgcgtggc ccgtgaaggt gatecacacc gacaacggcc ccaacttcac ctccgccacc	2520
gtgaaggccg cctgctggtg ggccggcgtg cagcaggagt tcggcatccc ctacaacccc	2580
cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg	2640
cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc	2700
aagcgcaagg ggcgcacgg cggctactcc gccggcgagc gcatcatcga catcatcgcc	2760
accgacatcc agaccaagga gctgcagaag cagatcatca agatccagaa ctccgcgtg	2820
tactaccgag actcccgca ccccatctgg aaggccccg ccaagctgct gtggaagggc	2880
gagggcgccg tggtagtcca ggacaactcc gacatcaagg tgggtccccg ccgcaaggcc	2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag	3000
gacgaggact aa	3012
<210> 225	
<211> 3012	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 225	
ttcttccgag aggacctggc cttccccag ggcaaggccc gcgagttctc ctccgagcag	60
accgcgcca actccccac ccgccgcgag ctgcaggtgt ggggcgcgca caacaactcc	120
ctgtccgagg ccggcgccga ccgccagggc accgtgtcct tctcttccc ccagatcacc	180

ctgtggcagc gccccctggt gaccatcaag atcggcggcc agctgaagga ggccctgctg	240
gacaccggcg ccgacgacac cgtgctggag gagatgaacc tgcccggccg ctggaagccc	300
aagatgatcg gcggcatcgg cggcttcac c aaggtgcgc agtacgacca gatcctgatc	360
gagatctgcg gccacaaggc catcggcacc gtgctggtgg gcccacccc cgtgaacatc	420
atcggccgca acctgctgac ccagatcggc tgcacctga acttccccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctggtg gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcggccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcacccag	720
gacttctggg aggtgcagct gggcatcccc cccccgccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccctggacaa ggacttccgc	840
aagtacaccg ccttcacat cccctccatc aacaacgaga cccccggcat ccgctaccag	900
tacaacgtgc tgccccaggg ctggaagggc tccccgccca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg caagcagaac cccgacatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgccagcacc tgctgcgctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccacgtg	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctgggtgg caagctgaac	1260
tgggcctccc agatctacgc cggcatcaag gtgaagcagc tgtgcaagct gctgcgcggc	1320
accaaggccc tgaccgaggt gatccccctg accgaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tcctgaagga gcccgtgcac ggctgtact acgaccctc caaggacctg	1440
atgcgcgaga tcagaagca gggccagggc cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgcc cgcatgcgcg gcgcccacac caacgacgtg	1560
aagcagctga ccgaggccgt gcagaagatc gccaccgagt ccatcgtgat ctggggcaag	1620
acccccaaagt tcaagctgcc catccagaag gagacctggg aggcctggtg gaccgagtac	1680
tggcaggcca cctggatccc cgagtgggag ttctgtaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaagga gcccacgtg ggcgcgaga ccttctacgt ggacggcgcc	1800
gccaaccgcg agaccaagct gggcaaggcc ggctacgtga ccgaccgcgg ccgccagaag	1860
gtggtgtccc tgaccgacac caccaaccag aagaccgagc tgcaggccat ccacctggcc	1920
ctgcaggact ccggcctgga ggtgaacatc gtgaccgact ccagtacgc cctgggcatc	1980
atccaggccc agcccgacaa gtccgagtc gagctggtgt ccagatcat cgagcagctg	2040

atcaagaagg agaaggtgta cctggcctgg gtgccccccc acaagggcat cggcggcaac 2100
gagcaggtgg acaagctggt gtccgccggc atccgcaagg tgctgttcct ggacggcatc 2160
gacaaggccc aggaggagca cgagaagtac cactccaact ggcgcccat ggcctccgac 2220

ttcaacctgc cccccgtggt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg 2280
aagggcgagg ccatgcacgg ccaggtggac tgctcccccg gcatctggca gctggactgc 2340
accacactgg agggcaagat catcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400
gccgaggtga tccccgccga gaccggccag gagaccgcct acttcctgct gaagctggcc 2460
ggccgctggc ccgtgaagac catccacacc gacaacggt ccaacttcac ctccaccacc 2520
gtgaaggccg cctgctggtg ggccggcatc aagcaggagt tcggcatccc ctacaacccc 2580
cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg 2640

cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc 2700
aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcatcgtgga catcatcgcc 2760
accgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa ctccgcgtg 2820
tactaccgag actcccgca cccctgtgg aagggccccg ccaagctgct gtggaagggc 2880
gagggcgccg tggatgatcca ggacaactcc gacatcaagg tgggtccccg ccgcaaggcc 2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc ctccgccag 3000
gacgaggact aa 3012

<210> 226

<211> 1003

<212> PRT

<213> Human immunodeficiency virus

<400> 226

Phe Phe Arg Glu Asn Leu Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe

1 5 10 15

Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln

20 25 30

Val Trp Gly Arg Asp Asn Asn Pro Leu Ser Glu Ala Gly Ala Asp Arg

35 40 45

Gln Gly Thr Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg

50 55 60

Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu

65				70				75				80			
Asp	Thr	Gly	Ala	Asp	Asp	Thr	Val	Leu	Glu	Glu	Met	Asn	Leu	Pro	Gly
85				90				95							
Lys	Trp	Lys	Pro	Lys	Met	Ile	Gly	Gly	Ile	Gly	Gly	Phe	Ile	Lys	Val
100				105				110							
Arg	Gln	Tyr	Asp	Gln	Ile	Leu	Ile	Glu	Ile	Cys	Gly	His	Lys	Ala	Ile
115				120				125							
Gly	Thr	Val	Leu	Val	Gly	Pro	Thr	Pro	Val	Asn	Ile	Ile	Gly	Arg	Asn
130				135				140							
Leu	Leu	Thr	Gln	Ile	Gly	Cys	Thr	Leu	Asn	Phe	Pro	Ile	Ser	Pro	Ile
145				150				155				160			
Glu	Thr	Val	Pro	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	Lys	Val
165				170				175							
Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val	Glu	Ile
180				185				190							
Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu
195				200				205							
Asn	Pro	Tyr	Asn	Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr
210				215				220							
Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln
225				230				235				240			
Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys
245				250				255							
Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser
260				265				270							
Val	Pro	Leu	Asp	Lys	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro
275				280				285							
Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu
290				295				300							
Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	Met	Thr
305				310				315				320			
Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Glu	Ile	Val	Ile	Tyr

325 330 335
Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln

340 345 350
His Arg Thr Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly

355 360 365
Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp

370 375 380
Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val

385 390 395 400
Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val

405 410 415

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys Val Lys
420 425 430

Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Val
435 440 445

Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
450 455 460

Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
465 470 475 480

Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile

485 490 495
Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met

500 505 510
Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln

515 520 525
Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe

530 535 540
Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr Glu Tyr

545 550 555 560

Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
565 570 575

Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
 610 615 620
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala

 625 630 635 640
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700

 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu

 770 775 780
 Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile
 805 810 815
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn
 820 825 830

Gly Ser Asn Phe Thr Ser Thr Thr Val Lys Ala Ala Cys Trp Trp Ala
835 840 845

Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
850 855 860

Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
865 870 875 880

Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
885 890 895

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
900 905 910

Glu Arg Ile Val Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu

915 920 925

Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
930 935 940

Ser Arg Asp Pro Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945 950 955 960

Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
965 970 975

Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
980 985 990

Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
995 1000

<210> 227

<211> 3012

<212> DNA

<213> Human immunodeficiency virus

<400> 227

ttcttccgcg agaacctggc ctccccccag ggcaaggccc gcgagttctc ctccgagcag	60
accgcgcga actccccac ccgccgcgag ctgcaggtgt ggggccgcga caacaacccc	120
ctgtccgagg ccggcgccga ccgccagggc accgtgtcct tctccttccc ccagatcacc	180
ctgtggcagc gccccctggt gaccatcaag atcggcggcc agctgaagga ggccctgctg	240

gacaccggcg ccgacgacac cgtgctggag gagatgaacc tgcccggcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggcttcac aaggtgcgcc agtacgacca gatcctgac	360
gagatctgcg gccacaaggc catcggcacc gtgctggtgg gceccacccc cgtgaacac	420
atcgccgca acctgctgac ccagatcggc tgcaccctga acttcccat ctccccac	480
gagaccgtgc ccgtgaagct gaagcccgcc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctggtg gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcggccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcaccag	720
gacttctggg aggtgcagct gggcatcccc cccccgccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccctggacaa ggacttccgc	840
aagtacaccg ccttcacat cccctccac aacaacgaga ccccgccat ccgctaccag	900
tacaacgtgc tgccccaggc ctggaagggc tccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg caagcagaac cccgagatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgcgagcacc tgctgcgctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccacgtg	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctggtggg caagctgaac	1260
tgggcctccc agatctacg cggcatcaag gtgaagcagc tgtgcaagct gctgcgcggc	1320
accaaggccc tgaccgaggt ggtgccccctg accgaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga gcccgtgcac ggctgtact acgaccctc caaggacctg	1440
atcgccgaga tccagaagca gggccagggc cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgcc cgcattgcgc gcgccacac caacgacgtg	1560
aagcagctga ccgaggccgt gcagaagatc gccaccgagt ccatcgtgat ctggggcaag	1620
acccccaaat tcaagctgcc catccagaag gagacctggg aggcctggtg gaccgagtac	1680
tggcaggcca cctggatccc cgagtgggag ttcgtgaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaagga gcccacgtg ggcgcgaga ccttctacgt ggacggcgcc	1800
gccaaaccgc agaccaagct gggcaaggcc ggtacgtga ccgaccgcgg ccgccagaag	1860
gtggtgtccc tgaccgacac caccaaccag aagaccgagc tgcaggccat ccacctggcc	1920
ctgcaggact ccggcctgga ggtgaacatc gtgaccgact ccagtagc cctgggcatc	1980

atccaggccc agcccgacaa gtccgagtc gagctggtgt cccagatcat cgagcagctg 2040
 atcaagaagg agaaggtgta cctggcctgg gtgcccgcac acaagggcat cggcggcaac 2100
 gagcaggtgg acaagctggt gtccgccggc atccgcaagg tgctgttcct ggacggcatc 2160
 gacaaggccc aggaggagca cgagaagtac cactccaact ggcgcgccat ggctccgac 2220
 ttcaacctgc cccccgtggt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg 2280
 aagggcgagg ccatgcacgg ccaggtggac tgctcccccg gcatctggca gctggactgc 2340
 acccacctgg agggaagat catcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400

gccgaggtga tccccgccga gaccggccag gagaccgct acttcatcct gaagctggcc 2460
 ggccgctggc ccgtgaaggt gatccacacc gacaacggct ccaacttcac ctccaccacc 2520
 gtgaaggccg cctgctggtg ggccggcatc aagcaggagt tcggcatccc ctacaacccc 2580
 cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg 2640
 cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcac ccacaacttc 2700
 aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcatcgtgga catcatcgcc 2760
 accgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa ctccgcgtg 2820

tactaccgag actccccgca cccctgtggt aaggcccccg ccaagctgct gtggaagggc 2880
 gagggcgccg tggatgatcca ggacaactcc gacatcaagg tgggtccccg ccgcaaggcc 2940
 aagatcatcc gcgactacgg caagcagatg gccggcgagc actgcgtggc ctcccgccag 3000
 gacgaggact aa 3012

<210> 228
 <211> 999
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 228

Phe Phe Arg Glu Asn Leu Ala Phe Pro Gln Gly Glu Ala Arg Glu Phe
 1 5 10 15

Pro Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Gln
 20 25 30

Val Arg Gly Asp Asn Pro Arg Ser Glu Ala Gly Ala Glu Arg Gln Gly
 35 40 45

Thr Leu Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Ser
 50 55 60

Ile Lys Val Gly Gly Gln Ile Lys Glu Ala Leu Leu Asp Thr Gly Ala

65				70				75				80			
Asp	Asp	Thr	Val	Leu	Glu	Glu	Ile	Asn	Leu	Pro	Gly	Lys	Trp	Lys	Pro
				85				90				95			
Lys	Met	Ile	Gly	Gly	Ile	Gly	Gly	Phe	Ile	Lys	Val	Arg	Gln	Tyr	Asp
				100				105				110			
Gln	Ile	Leu	Ile	Glu	Ile	Cys	Gly	Lys	Lys	Ala	Ile	Gly	Thr	Val	Leu
				115				120				125			
Val	Gly	Pro	Thr	Pro	Val	Asn	Ile	Ile	Gly	Arg	Asn	Met	Leu	Thr	Gln
				130				135				140			
Leu	Gly	Cys	Thr	Leu	Asn	Phe	Pro	Ile	Ser	Pro	Ile	Glu	Thr	Val	Pro
145				150				155				160			
Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	Lys	Val	Lys	Gln	Trp	Pro
				165				170				175			
Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Thr	Ala	Ile	Cys	Glu	Glu	Met
				180				185				190			
Glu	Lys	Glu	Gly	Lys	Ile	Thr	Lys	Ile	Gly	Pro	Glu	Asn	Pro	Tyr	Asn
				195				200				205			
Thr	Pro	Val	Phe	Ala	Ile	Lys	Lys	Lys	Asp	Ser	Thr	Lys	Trp	Arg	Lys
				210				215				220			
Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln	Asp	Phe	Trp	Glu
225				230				235				240			
Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys	Lys	Lys	Lys	Ser
				245				250				255			
Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser	Val	Pro	Leu	Asp
				260				265				270			
Glu	Gly	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro	Ser	Ile	Asn	Asn
				275				280				285			
Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu	Pro	Gln	Gly	Trp
				290				295				300			
Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	Met	Thr	Lys	Ile	Leu	Glu
305				310				315				320			

Pro Phe Arg Ala Gln Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Asp
325 330 335

Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys
340 345 350

Ile Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro
355 360 365

Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu
370 375 380

Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys
385 390 395 400

Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn
405 410 415

Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys
420 425 430

Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu
435 440 445

Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro
450 455 460

Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile
465 470 475 480

Gln Lys Gln Gly His Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro
485 490 495

Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His
500 505 510

Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met
515 520 525

Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile
530 535 540

Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr
545 550 555 560

Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu
565 570 575

Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ala Gly Ala Glu Thr Phe Tyr
580 585 590

Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Ile Gly Lys Ala Gly Tyr
595 600 605

Val Thr Asp Arg Gly Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr
610 615 620

Asn Gln Lys Thr Glu Leu Gln Ala Ile Gln Leu Ala Leu Gln Asp Ser
625 630 635 640

Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile
645 650 655

Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile

660 665 670

Ile Glu Gln Leu Ile Lys Lys Glu Arg Val Tyr Leu Ser Trp Val Pro
675 680 685

Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser
690 695 700

Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln
705 710 715 720

Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Glu
725 730 735

Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp
740 745 750

Lys Cys Gln Leu Lys Gly Glu Ala Ile His Gly Gln Val Asp Cys Ser
755 760 765

Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Ile Ile
770 775 780

Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile
785 790 795 800

Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Tyr Ile Leu Lys Leu Ala

805 810 815

Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly Ser Asn Phe

820 825 830
 Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln
 835 840 845
 Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser
 850 855 860
 Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala
 865 870 875 880

Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe
 885 890 895
 Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile
 900 905 910
 Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile
 915 920 925
 Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro
 930 935 940
 Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val

945 950 955 960
 Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala
 965 970 975
 Lys Ile Ile Lys Asp Tyr Gly Lys Gln Met Ala Gly Ala Asp Cys Val
 980 985 990

Ala Gly Arg Gln Asp Glu Asp
 995

<210> 229
 <211> 3000
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 229

ttcttccgcg agaacctggc cttcccccag ggcgaggccc gcgagttccc ctccgagcag 60

accgcgcga actccccac ctcccgagag ctgcaggtgc gcggcgacaa cccccgtcc 120

gaggccggcg ccgagcgcca gggcaccttg aattccccc agatcaccct gtggcagcgc 180

ccctgtgtgt ccatcaaggt gggcggccag atcaaggagg ccctgctgga caccggcgcc 240

gacgacaccg tgctggagga gatcaacctg cccggcaagt ggaagcccaa gatgatcggc	300
ggcatcggcg gcttcatcaa ggtgcgccag tacgaccaga tcctgatcga gatctgcggc	360
aagaaggcca tcggcaccgt gctggtgggc cccacccccg tgaacatcat cggccgcaac	420
atgctgaccc agctgggctg caccctgaac ttccccatct ccccatcga gaccgtgccc	480
gtgaagctga agccccgcat ggacggcccc aaggtgaagc agtggcccct gaccgaggag	540
aagatcaagg cctgaccgc catctgcgag gagatggaga aggagggcaa gatcaccaag	600
atcgccccg agaacccta caacaccccc gtgttcgcca tcaagaagaa ggactccacc	660
aagtggcgca agctggtgga cttccgcgag ctgaacaagc gcaccagga cttctgggag	720
gtgcagctgg gcatcccca ccccgccggc ctgaagaaga agaagtccgt gaccgtgctg	780
gacgtgggcg acgcctactt ctccgtgccc ctggacgagg gcttccgca gtacaccgcc	840
ttcaccatcc cctccatcaa caacgagacc cccggcatcc gctaccagta caacgtgctg	900
ccccagggtt ggaagggtc ccccgccatc ttccagtcct ccatgaccaa gatcctggag	960
cccttccgcg ccagaaccc cgagatcgtg atctaccagt acatggacga cctgtacgtg	1020
ggctccgacc tggagatcgg ccagcacgc gccaatatcg aggagctgcg cgagcacctg	1080
ctgaagtggg gcttcaccac ccccgacaag aagcaccaga aggagcccc cttcctgtgg	1140
atgggtacg agctgcacc cgacaagtgg accgtgcagc ccatccagct gcccgagaag	1200
gactcctgga ccgtgaacga catccagaag ctggtgggca agctgaactg ggcctcccag	1260
atctacccc gcatcaaggt gcgccagctg tgcaagctgc tgcgcggcgc caaggccctg	1320
accgacatcg tgccccgac cgaggaggcc gagctggagc tggccgagaa ccgcgagatc	1380
ctgaaggagc ccgtgcacgg cgtgtactac gaccctcca aggacctgat cgccgagatc	1440
cagaagcagg gccacgacca gtggacctac cagatctacc aggagccctt caagaacctg	1500
aagaccgga agtacgcaa gatgcgcacc gccacacca acgacgtgaa gcagctgacc	1560
gaggccgtgc agaagatcg catggagtcc atcgtgatct ggggcaagac cccaagttc	1620
cgctgcca tccagaagga gacctgggag acctggtgga ccgactactg gcaggccacc	1680
tggatccccg agtgggagtt cgtgaacacc cccccctgg tgaagctgtg gtaccagctg	1740
gagaaggagc ccatgccgg cgccgagacc ttctacgtgg acggcgccgc caaccgcgag	1800
accaagatcg gcaaggccgg ctacgtgacc gaccgcgcc gccagaagat cgtgtccctg	1860
accgagacca ccaaccagaa gaccgagctg caggccatcc agctggccct gcaggactcc	1920
ggctccgagg tgaacatcgt gaccgactcc cagtacgcc tgggcatcat ccaggcccag	1980
cccgacaagt ccgagtcga gctggtgaac cagatcatcg agcagctgat caagaaggag	2040
cgctgtacc tgtcctgggt gcccggccac aagggcacg gcggcaacga gcaggtggac	2100

aagctggtgt cctccggcat ccgcaagggtg ctgttctctgg acggcatcga caaggcccag 2160

gaggagcacg agaagtacca ctccaactgg cgcgccatgg cctccgagtt caacctgccc 2220

cccctcgtgg ccaaggagat cgtggcctcc tgcgacaagt gccagctgaa gggcgaggcc 2280

atccacggcc aggtggactg ctccccggc atctggcagc tggactgcac ccacctggag 2340

ggcaagatca tcttgggtggc cgtgcacgtg gcctccggct acatcgaggc cgaggtgatc 2400

cccgcgaga cgggccagga gaccgcctac tacatcctga agctggccgg ccgctggccc 2460

gtgaagggtga tccacaccga caacggctcc aacttcacct ccgccgccgt gaaggccgcc 2520

tgctggtggg ccggcatcca gcaggagttc ggcatccctt acaaccccca gtcccagggc 2580

gtggtggagt ccatgaacaa ggagctgaag aagatcatcg gccagggtgcg cgaccaggcc 2640

gagcacctga agaccgccgt gcagatggcc gtgttcatcc acaacttcaa gcgcaagggc 2700

ggcatcggcg gctactccgc cggcgagcgc atcatcgaca tcatcgccac cgacatccag 2760

accaaggagc tgcagaagca gatcatcaag atccagaact tccgcgtgta ctaccgcgac 2820

tcccgcgacc ccatctggaa gggccccgcc aagctgctgt ggaagggcga gggcgccgtg 2880

gtgatccagg acaactccga catcaagggt gtgccccgcc gcaaggccaa gatcatcaag 2940

gactacggca agcagatggc cggcgccgac tgcgtggccg gccgccagga cgaggactaa 3000

3000

<210> 230

<211> 1001

<212> PRT

<213> Human immunodeficiency virus

<400> 230

Phe Phe Arg Glu Asn Leu Ala Phe Pro Gln Gly Glu Ala Arg Glu Phe

1 5 10 15

Pro Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Gln

20 25 30

Val Gly Arg Asp Asn Pro Arg Ser Glu Ala Gly Ala Glu Arg Gln Gly

35 40 45

Thr Leu Thr Leu Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu

50 55 60

Val Ser Ile Lys Val Gly Gly Gln Ile Lys Glu Ala Leu Leu Asp Thr

65 70 75 80

Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asn Leu Pro Gly Lys Trp
 85 90 95
 Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg Gln
 100 105 110
 Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly Thr
 115 120 125
 Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Met Leu
 130 135 140
 Thr Gln Leu Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr
 145 150 155 160
 Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln
 165 170 175
 Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Glu
 180 185 190

 Glu Met Glu Lys Glu Gly Lys Ile Thr Lys Ile Gly Pro Glu Asn Pro
 195 200 205
 Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp
 210 215 220
 Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe
 225 230 235 240
 Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys
 245 250 255
 Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro
 260 265 270
 Leu Asp Glu Gly Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile
 275 280 285
 Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln
 290 295 300
 Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile
 305 310 315 320
 Leu Glu Pro Phe Arg Ala Gln Asn Pro Glu Ile Val Ile Tyr Gln Tyr

				325				330				335			
Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln	His	Arg
340				345				350							
Ala	Lys	Ile	Glu	Glu	Leu	Arg	Glu	His	Leu	Leu	Lys	Trp	Gly	Phe	Thr
355				360				365							
Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp	Met	Gly
370				375				380							
Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Gln	Leu	Pro
385				390				395				400			
Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val	Gly	Lys
405				410				415							
Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	Val	Arg	Gln	Leu
420				425				430							
Cys	Lys	Leu	Leu	Arg	Gly	Ala	Lys	Ala	Leu	Thr	Asp	Ile	Val	Pro	Leu
435				440				445							
Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile	Leu	Lys
450				455				460							
Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu	Ile	Ala
465				470				475				480			
Glu	Ile	Gln	Lys	Gln	Gly	His	Asp	Gln	Trp	Thr	Tyr	Gln	Ile	Tyr	Gln
485				490				495							
Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Lys	Met	Arg	Thr
500				505				510							
Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln	Lys	Ile
515				520				525							
Ala	Met	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	Lys	Phe	Arg	Leu
530				535				540							
Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	Asp	Tyr	Trp	Gln
545				550				555				560			
Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	Pro	Pro	Leu	Val
565				570				575							

Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ala Gly Ala Glu Thr
 580 585 590
 Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Ile Gly Lys Ala
 595 600 605
 Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Ile Val Ser Leu Thr Glu
 610 615 620

 Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Gln Leu Ala Leu Gln
 625 630 635 640
 Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu
 645 650 655
 Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn
 660 665 670
 Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser Trp
 675 680 685
 Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu

 690 695 700
 Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys
 705 710 715 720
 Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala
 725 730 735
 Ser Glu Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala Ser
 740 745 750
 Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp
 755 760 765

 Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys
 770 775 780
 Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu
 785 790 795 800
 Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu Lys
 805 810 815
 Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly Ser
 820 825 830

Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly Ile

835	840	845	
Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val			
850	855	860	
Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp			
865	870	875	880
Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His			
885	890	895	
Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg			
900	905	910	

Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys			
915	920	925	
Gln Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg			
930	935	940	
Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly			
945	950	955	960
Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg			
965	970	975	
Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Ala Asp			

980	985	990
Cys Val Ala Gly Arg Gln Asp Glu Asp		
995	1000	

<210> 231

<211> 3006

<212> DNA

<213> Human immunodeficiency virus

<400> 231

ttcttccgcg agaacctggc ctccccccag ggcgaggccc gcgagttccc ctccgagcag	60
accgcgcga actccccac ctccgcgag ctgcaggtgg gccgcgaca cccccgtcc	120
gaggccggcg ccgagcgcca gggcacctg acctgaact tccccagat caccctgtgg	180
cagcgcccc tgggtgtccat caaggtgggc ggccagatca aggaggccct gctggacacc	240

ggcgccgacg acaccgtgct ggaggagatc aacctgcccc gcaagtggaa gccaagatg	300
atcggcggca tggcggtt catcaaggct cgccagtacg accagatcct gatcgagatc	360
tgcggcaaga aggccatcgg caccgtgctg gtgggcccc ccccggtgaa catcatcggc	420
cgcaacatgc tgaccagct gggctgcacc ctgaacttcc ccatctcccc catcgagacc	480
gtgccgtga agctgaagcc cgccatggac ggccccaagg tgaagcagt gcccctgacc	540
gaggagaaga tcaaggccct gaccgccatc tgcgaggaga tggagaagga gggcaagatc	600
accaagatcg gccccgagaa cccctacaac acccccggtg tcgcatcaa gaagaaggac	660
tccaccaagt ggcgcaagct ggtggacttc cgcgagctga acaagcgac ccaggacttc	720
tgggagggtgc agctgggcat cccccacccc gccggcctga agaagaagaa gtccgtgacc	780
gtgctggacg tggcgacgc ctacttctcc gtccccctgg acgagggtt ccgcaagtac	840
accgcttca ccatccctc catcaacaac gagaccccc gcacccgcta ccagtacaac	900
gtgctgcccc agggctggaa gggctcccc gccatcttcc agtcctccat gaccaagatc	960
ctggagccct tccgcgcca gaaccccgag atcgtgatct accagtacat ggacgacctg	1020
tacgtgggct ccgacctgga gatcgccag caccgcgcca agatcgagga gctgcgcgag	1080
cacctgctga agtggggctt caccaccccc gacaagaagc accagaagga gcccccttc	1140
ctgtggatgg gctacgagct gcaccccgac aagtggaccg tgcagcccat ccagctgccc	1200
gagaaggact cctggaccgt gaacgacatc cagaagctgg tgggcaagct gaactgggcc	1260
tcccagatct accccggcat caaggtgcgc cagctgtgca agctgctgcg cggcgccaag	1320
gcccctgaccg acatcgtgcc cctgaccgag gagccgagc tggagctggc cgagaaccgc	1380
gagatcctga aggagccgt gcacggcgtg tactacgacc cctccaagga cctgatgcc	1440
gagatccaga agcagggccga cgaccagtgg acctaccaga tctaccagga gcccttcaag	1500
aacctgaaga cgggcaagta cgccaagatg cgcaccgccc acaccaacga cgtgaagcag	1560
ctgaccgagg ccgtgcagaa gatcgccatg gattccatcg tgatctgggg caagaccccc	1620
aagtccgcc tggccatcca gaaggagacc tgggagacct ggtggaccga ctactggcag	1680
gccacctgga tccccagatg ggagttctg aacaccccc cctgggtgaa gctgtggtac	1740
cagctggaga aggagcccat cgccggcgcc gagaccttct acgtggacgg cgccgccaac	1800
cgcgagacca agatcgga gggcggtac gtgaccgacc gcggcgcca gaagatcgtg	1860
tccctgaccg agaccaccaa ccagaagacc gagctgcagg ccatccagct ggccctgcag	1920
gactccggt cggaggtgaa catcgtgacc gactcccagt acgccctggg catcatccag	1980
gcccagccc acaagtccga gtccgagctg gtgaaccaga tcatcgagca gctgatcaag	2040
aaggagaagg tgtacctgc ctgggtgccc gccacaagg gcacggcg caacgagcag	2100

gtggacaagc tgggtgtcctc cggcatccgc aaggtgctgt tcctggacgg catcgacaag 2160
 gcccaggagg agcacagagaa gtaccactcc aactggcgcg ccatggcctc cgagttcaac 2220
 ctgcccccca tcgtggccaa ggagatcgtg gcctcctgcg acaagtgccg gctgaagggc 2280
 gaggccatgc acggccaggt ggactgctcc cccggcatct ggcagctgga ctgcacccac 2340

ctggagggca agatcatcct ggtggccgtg cacgtggcct ccggctacat cgaggccgag 2400
 gtgatccccg ccgagaccgg ccaggagacc gcctacttca tcctgaagct ggccggccgc 2460
 tggcccgatga aggtgatcca caccgacaac ggctccaact tcacctccgc cgccgtgaag 2520
 gccgcctgct ggtgggcccgg catccagcag gatttcggca tcccctacaa cccccagtc 2580
 cagggcgtgg tggagtccat gaacaaggag ctgaagaaga tcatcggccg ggtgcgcgac 2640
 caggccgagc acctgaagac cgccgtgcag atggccgtgt tcattccaca cttcaagcgc 2700
 aaggcgccga tcggcgccga ctccgccggc gagcgcatca tcgacatcat cgccaccgac 2760

atccagacca aggagctgca gaagcagatc atcaagatcc agaacttccg cgtgtactac 2820
 cgcgactccc gcgaccccat ctggaagggc cccgccaagc tgctgtggaa gggcgagggc 2880
 gccgtggtga tccaggacaa ctccgacatc aaggtggtgc cccgccgcaa ggccaagatc 2940
 atccgcgact acggcaagca gatggccggc gccgactgcg tggccggccg ccaggacgag 3000
 gactaa 3006

<210> 232
 <211> 1002
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 232

Phe Phe Arg Glu Asn Leu Ala Phe Pro Gln Gly Lys Ala Gly Glu Leu

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

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

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His
 340 345 350
 Arg Thr Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly Phe
 355 360 365

 Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
 370 375 380
 Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Lys Leu
 385 390 395 400
 Pro Glu Lys Glu Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
 405 410 415
 Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln
 420 425 430
 Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro

 435 440 445
 Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
 450 455 460
 Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
 465 470 475 480
 Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr
 485 490 495
 Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg
 500 505 510

 Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys
 515 520 525
 Ile Ala Ile Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg
 530 535 540
 Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp
 545 550 555 560
 Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu
 565 570 575
 Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala Glu

580 585 590
 Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys
 595 600 605
 Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu Thr
 610 615 620
 Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Asn Leu Ala Leu
 625 630 635 640
 Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
 645 650 655

 Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val
 660 665 670
 Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala
 675 680 685
 Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
 690 695 700
 Leu Val Ser Asn Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
 705 710 715 720
 Lys Ala Gln Glu Glu His Glu Lys Tyr His Asn Asn Trp Arg Ala Met

 725 730 735
 Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala
 740 745 750
 Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
 755 760 765
 Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly
 770 775 780
 Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala
 785 790 795 800

 Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu
 805 810 815
 Lys Leu Ala Gly Arg Trp Pro Val Lys Val Val His Thr Asp Asn Gly
 820 825 830
 Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly

835 840 845
 Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
 850 855 860
 Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg

 865 870 875 880
 Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
 885 890 895
 His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
 900 905 910
 Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln
 915 920 925
 Lys Gln Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
 930 935 940

 Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
 945 950 955 960
 Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg
 965 970 975
 Arg Lys Val Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp
 980 985 990
 Asp Cys Val Ala Ser Arg Gln Asp Glu Asp
 995 1000

 <210> 233
 <211> 1003
 <212> PRT

 <213> Human immunodeficiency virus
 <400> 233
 Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Lys Phe
 1 5 10 15
 Pro Ser Glu Gln Thr Arg Ala Asn Ser Pro Ala Ser Arg Glu Leu Arg
 20 25 30
 Val Gln Arg Gly Asp Asn Pro Leu Ser Glu Ala Gly Ala Glu Arg Arg
 35 40 45

Gly Thr Val Pro Ser Leu Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg
50 55 60

Pro Leu Val Thr Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
65 70 75 80

Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly
85 90 95

Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
100 105 110

Lys Gln Tyr Asp His Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile
115 120 125

Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
130 135 140

Met Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
145 150 155 160

Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
165 170 175

Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile
180 185 190

Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
195 200 205

Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
210 215 220

Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
225 230 235 240

Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
245 250 255

Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
260 265 270

Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
275 280 285

Ser Val Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu

290 295 300
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr
 305 310 315 320
 Lys Ile Leu Glu Pro Phe Arg Thr Lys Asn Pro Asp Ile Val Ile Tyr
 325 330 335
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 340 345 350

 His Arg Thr Lys Ile Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly
 355 360 365
 Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 370 375 380
 Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln
 385 390 395 400
 Leu Pro Asp Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
 405 410 415
 Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys

 420 425 430
 Gln Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val
 435 440 445
 Pro Leu Thr Ala Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 450 455 460
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 465 470 475 480
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 485 490 495

 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met
 500 505 510
 Arg Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 515 520 525
 Lys Ile Ala Leu Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Arg Leu Pro Ile Leu Lys Glu Thr Trp Asp Thr Trp Trp Thr Asp Tyr

545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro

 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Thr Glu Pro Ile Val Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ser Asn Arg Glu Thr Lys Lys Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
 610 615 620
 Thr Glu Thr Thr Asn Gln Lys Ala Glu Leu Gln Ala Ile His Leu Ala
 625 630 635 640

 Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Asn Gln Ile Ile Glu Gln Leu Ile Gln Lys Glu Lys Val Tyr Leu
 675 680 685
 Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Lys Leu Val Ser Ala Gly Ile Arg Lys Ile Leu Phe Leu Asp Gly Ile

 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Asn Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780

 Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800

Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile
805 810 815
Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Ile Ile His Thr Asp Asn
820 825 830
Gly Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala
835 840 845
Gly Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly

850 855 860
Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
865 870 875 880
Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
885 890 895
Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
900 905 910
Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Arg Glu Leu
915 920 925

Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
930 935 940
Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945 950 955 960
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro
965 970 975
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
980 985 990
Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp

995 1000
<210> 234
<211> 3009
<212> DNA
<213> Human immunodeficiency virus
<400> 234

ttcttccgcg agaacctggc cttccccag ggcaaggccg gcgagctgtc ctccgagcag

60

acccgcgcca actccccac ctcccgag ctgcgcgtgt ggggcggcga caaccctg	120
tccgagaccg gcgcgagcg ccagggcacc gtgtccttca acttccccca gatcacctg	180
tggcagcgcc ccttggtgac catcaagatc ggcgccagc tgaaggagc cctgctggac	240
accggcgccg acgacacgt gctggaggag atcaacctgc ccggcaagt gaagccaag	300
atgatggcg gcatcgcggt cttcatcaag gtgcgccagt acgaccagat cctgatcgag	360
atctcgggc acaaggccat cggcaccgtg ctggtgggc ccaccccg gaacatcatc	420
ggccgaacc tgctgacca gatcggtgc accctgaact tccccatctc ccccatcgag	480
accgtcccc tgaagctgaa gcccgcatg gacggccca aggtgaagca tggccctg	540
accgaggaga agatcaagg cctgaccgag atctgcaccg agatggagaa ggagggaag	600
atctccgca tcggccccga gaaccctac aacacccca tcttcgcat caagaagaag	660
gactcacca agtggcgcaa gctggtggac ttccgcgagc tgaacaagc caccaggac	720
ttctgggagg tgcagctggg catccccac ccgcccggc tgaagaaga gaagtcctg	780
accgtgctgg acgtggcgga cgcctacttc tccgtgccc tggacgagga ctccgcaag	840
tacacgcct tcaccatccc ctccatcaac aacgagacc ccggcatccg ctaccgtac	900
aacgtgctgc ccagggtg gaagggtcc ccgcatct tccagtcctc catgaccaag	960
atctggagc cttccgca gcagaacccc gagatcgtga tctaccagta catggacgac	1020
ctgtacgtgg gtccgacct ggagatcggc cagcaccga ccaagatcga ggagctgcgc	1080
gagcacctgc tgcgtgggg cttcaccacc ccgacaaga agcaccagaa ggagcccc	1140
ttctgtgga tgggctacga gctgcacccc gacaagtga ccgtgcagcc catcaagctg	1200
cccgagaagg agtcctggac cgtgaacgac atccagaagc tgggggcaa gctgaactgg	1260
gcctcccaga tctaccccg catcaagtg cgccagctgt gcaagctgct gcgcggcacc	1320
aaggccctga ccgagtgat cccctgacc gaggaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagc cgtgcacggc gtgtactacg accctccaa ggacctgatc	1440
gccgagatcc agaagcagg ccagggccag tggacctacc agatctacca ggagccctc	1500
aagaacctga agaccggcaa gtacccccg atgcgcggcg ccacaccaa cgactgaag	1560
cagtgaccg aggcctgca gaagatgcc atcgagtcca tcgtgatctg gggcaagacc	1620
cccaagtcc gcctgccat ccagaaggag acctgggaga cctggtggac cgagtactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacaccc cccctggt gaagctgtgg	1740
taccagctgg agaaggagc catcatcggc gccgagacct tctacgtgga cggcgccgc	1800
aaccgcgaga ccaagctggg caaggccggc tacgtgaccg accgcggccg ccagaaggtg	1860
gtgcccctga ccgacaccac caaccagaag accgagctgc aggccatcaa cctggccctg	1920

caggactccg gcctggaggt gaacatcgtg accgactccc agtacgcctt ggcatcatc 1980

caggccccagc ccgacaagtc cgagtcagag ctggtgtccc agatcatcga gcagctgac 2040

aagaaggaga aggtgtacct ggcttgggtg cccgcccaca agggcatcgg cggcaacgag 2100

caggtggaca agctggtgtc caacggcatc cgcaaggtgc tgttcctgga cggcatcgac 2160

aaggccccagg aggagcacga gaagtaccac aacaactggc gcgccatggc ctccgacttc 2220

aacctgcccc ccgtggtggc caaggagatc gtggcctcct gcgacaagtg ccagctgaag 2280

ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc 2340

cacctggagg gcaaggtgat cctggtggcc gtgcacgtgg cctccggcta catcgaggcc 2400

gaggtgatcc ccgccgagac cggccaggag accgcctact tctgctgaa gctggccggc 2460

cgctggcccc tgaaggtggt gcacaccgac aacggctcca acttcacctc cgccgccgtg 2520

aaggccgcct gctggtgggc cggcatcaag caggagtctg gcaccccta caacccccag 2580

tccaggggcg tgggtggagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc 2640

gaccaggccg agcacctgaa gaccgccgtg cagatggccg tgttcacca caacttcaag 2700

cgcaaggcg gcacggcggt ctactccgcc ggcgagcgca tcacgacat catgccacc 2760

gacatccaga ccaaggagct gcagaagcag atcatcaaga tccagaactt ccgctgtac 2820

taccgcgact ccgcgaccc catctggaag ggcccccca agctgctgtg gaaggcgag 2880

ggcgccgtgg tgatccagga caactccgac atcaaggtgg tgcccccg caaggtgaag 2940

atcatccgag actacggcaa gcagatggc ggcgacgact gcgtggcctc ccgccaggac 3000

gaggactaa 3009

<210> 235

<211> 3012

<212> DNA

<213> Human immunodeficiency virus

<400> 235

ttcttcgagc agaacttggc ctccagcag ggcgaggccc gcaagttccc ctccgagcag 60

accgcgcca actccccgc ctccgcgag ctgcgcgtgc agcgcgcgca caacccctg 120

tccgaggcgc gcgccgagc cgcgggcacc gtgcctccc tgtcctccc ccagatcacc 180

ctgtggcagc gccccctggt gaccatcaag atcggcggcc agctgaagga ggccctgctg 240

gacaccggcg ccgacgacac cgtgtgtgag gacatcaacc tgcccggcaa gtggaagccc 300

aagatgatcg gcggcatcgg cggtttcatc aaggtgaagc agtacgacca catcctgac 360

gagatctgag gccacaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacatc 420

atcgccgca acatgctgac ccagatcggc tgcacctga acttcccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccgcc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcggccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctgggtg gacttccgcg agctgaacaa ggcgaccag	720
gacttctggg aggtgcagct gggcatcccc cccccgcg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccctggacaa ggacttccgc	840
aagtacaccg ccttcaccat cccctccgtg aacaacgaga cccccggcat ccgtaccag	900
tacaacgtgc tgccccaggc ctggaagggc tccccgcca tcttccagt ctccatgacc	960
aagatcctgg agcccttccg caccaagaac cccgacatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgcgagcacc tgctgaagtg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccattcag	1200
ctgccgaca aggactctg gaccgtgaac gacatccaga agctgggtggg caagctgaac	1260
tgggcctccc agatctaccc cggcatcaag gtgaagcagc tgtgcaagct gctgcgcggc	1320
gccaaggccc tgaccgacat cgtgcccctg accgccgagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga gcccgtgcac ggctgtact acgacccctc caaggacctg	1440
atcgccgaga tccagaagca gggccagggc cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgcc aagatgcgt ccgccacac caacgacgtg	1560
aagcagctga ccgaggccgt gcagaagatc gcctggagt ccatcgtgat ctggggcaag	1620
acccccaaat tccgctgcc catctgaag gagacctggg acacctggtg gaccgactac	1680
tggcaggcca cctggatccc cgagtgggag ttctgaaca cccccccct ggtgaagctg	1740
tgggtaccagc tggagaccga gcccacgtg ggccgagga ccttctacgt ggacggcgcc	1800
tccaaccgag agaccaagaa gggcaaggcc ggctacgtga ccgaccgagg ccgccagaag	1860
gtggtgtccc tgaccgagac caccaaccag aaggccgagc tgcaggccat ccacctggcc	1920
ctgcaggact ccggctccga ggtgaacatc gtgaccgact ccagtagc cctgggcac	1980
atccaggccc agcccgacaa gtccgagtcc gagctggtga accagatcat cgagcagctg	2040
atccagaagg agaaggtgta cctgtcctgg gtccccccc acaaggcat cggcggcaac	2100
gagcaggtgg acaagctggt gtccgccggc atccgcaaga tctgttcct ggacggcatc	2160

gacaaggccc aggaggagca cgagaagtac cacaacaact ggcgcgccat ggcctccgac 2220

ttcaacctgc ccccggtgtt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg 2280

aaggcgagg ccatgcacgg ccaggtggac tgcctccccg gcatctggca gctggactgc 2340

accacactgg agggcaagat catcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400

gccgaggtga tccccgccga gaccggccag gagaccgct acttcacct gaagctggcc 2460

ggcgcgtggc ccgtgaagat catecacacc gacaacggct ccaacttcac ctccgcgcc 2520

gtgaaggccg cctgctggtg ggccggcatc cagcaggagt tcggcatccc ctacaacccc 2580

cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg 2640

cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc 2700

aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcatcatcga catcatcgcc 2760

accgacatcc agaccgcga gctgcagaag cagatcacca agatccagaa ctccgcgtg 2820

tactaccgag actcccgca ccccggtgtg aagggtcccg ccaagctgct gtggaagggc 2880

gaggcgccg ttgtgatcca ggacaactcc gagatcaagg ttgtgccccg ccgcaaggcc 2940

aagatcatcc gcgactacgg caagcagatg gccggcgagc actgcgtggc cggccgccag 3000

gacgaggact aa 3012

<210> 236

<211> 1003

<212> PRT

<213> Human immunodeficiency virus

<400> 236

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Lys Phe

1 5 10 15

Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Ala Ser Arg Glu Leu Arg

20 25 30

Val Arg Arg Gly Asp Asn Ser Leu Pro Glu Ala Gly Ala Glu Arg Gln

35 40 45

Gly Thr Gly Ser Ser Leu Asp Phe Pro Gln Ile Thr Leu Trp Gln Arg

50 55 60

Pro Leu Val Thr Ile Lys Val Gly Gly Gln Leu Arg Glu Ala Leu Leu

65 70 75 80

Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly

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

Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln

340 345 350

His Arg Thr Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly

355 360 365

Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp

370 375 380

Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Ala Ile Gln

385 390 395 400

Leu Pro Asp Lys Ser Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val

405 410 415

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Arg Val Lys

420 425 430

His Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Val Val

435 440 445

Pro Leu Thr Ala Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile

450 455 460

Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu

465 470 475 480

Ile Ala Glu Ile Gln Lys Gln Gly His Asp Gln Trp Thr Tyr Gln Ile

485 490 495

Tyr Gln Glu Pro His Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Arg

500 505 510

Lys Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Val Val Gln

515 520 525

Lys Ile Ala Thr Glu Gly Ile Val Ile Trp Gly Lys Val Pro Lys Phe

530 535 540

Arg Leu Pro Ile Gln Lys Glu Thr Trp Glu Ile Trp Trp Thr Glu Tyr

545 550 555 560

Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro

565 570 575

Leu Val Lys Leu Trp Tyr Gln Leu Glu Thr Glu Pro Ile Val Gly Ala

580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Pro Leu
 610 615 620
 Thr Glu Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala

 625 630 635 640
 Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala His Pro Asp Lys Ser Glu Ser Glu Leu
 660 665 670
 Val Asn Gln Ile Ile Glu Gln Leu Ile Gln Lys Glu Arg Val Tyr Leu
 675 680 685
 Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700

 Lys Leu Val Ser Thr Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu

 770 775 780
 Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800
 Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile
 805 810 815
 Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Ile Ile His Thr Asp Asn
 820 825 830
 Gly Ser Asn Phe Thr Ser Thr Val Val Lys Ala Ala Cys Trp Trp Ala

835 840 845
 Gly Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
 Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
 865 870 875 880
 Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
 Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910
 Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
 915 920 925
 Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Phe Arg Asp
 930 935 940
 Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
 945 950 955 960
 Glu Gly Ala Val Val Ile Gln Asp Asn Asn Glu Ile Lys Val Val Pro
 965 970 975
 Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
 980 985 990
 Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000
 <210> 237
 <211> 1002
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 237
 Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Glu Phe
 1 5 10 15
 Ser Ser Glu Gln Ala Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Arg
 20 25 30
 Val Arg Arg Gly Asp Ser Pro Leu Pro Glu Ala Gly Ala Glu Gly Lys

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

Thr Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro
 290 295 300
 Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys
 305 310 315 320
 Ile Leu Glu Pro Phe Arg Thr Lys Asn Pro Glu Ile Val Ile Tyr Gln
 325 330 335
 Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His
 340 345 350
 Arg Ala Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly Phe
 355 360 365
 Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
 370 375 380
 Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu
 385 390 395 400
 Pro Asp Lys Glu Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
 405 410 415
 Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys Gln
 420 425 430
 Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro
 435 440 445
 Leu Thr Ala Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
 450 455 460
 Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Glu Leu Ile
 465 470 475 480
 Ala Glu Val Gln Lys Gln Gly Leu Asp Gln Trp Thr Tyr Gln Ile Tyr
 485 490 495
 Gln Glu Pro Tyr Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Arg Gly
 500 505 510
 Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Val Val Gln Lys
 515 520 525
 Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys
 530 535 540

Leu Pro Ile Arg Lys Glu Thr Trp Glu Val Trp Trp Thr Glu Tyr Trp

545 550 555 560

Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu

565 570 575

Val Lys Leu Trp Tyr Arg Leu Glu Thr Glu Pro Ile Pro Gly Ala Glu

580 585 590

Thr Tyr Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys

595 600 605

Ala Gly Tyr Val Thr Asp Lys Gly Lys Gln Lys Ile Ile Thr Leu Thr

610 615 620

Glu Thr Thr Asn Gln Lys Ala Glu Leu Gln Ala Ile His Leu Ala Leu

625 630 635 640

Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala

645 650 655

Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Leu Val

660 665 670

Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser

675 680 685

Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys

690 695 700

Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp

705 710 715 720

Lys Ala Gln Glu Glu His Glu Arg Tyr His Ser Asn Trp Arg Ala Met

725 730 735

Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala

740 745 750

Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val

755 760 765

Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly

770 775 780

Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala

785 790 795 800
 Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu
 805 810 815
 Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly
 820 825 830
 Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Asn

 835 840 845
 Ile Thr Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
 850 855 860
 Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
 865 870 875 880
 Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
 885 890 895
 His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
 900 905 910

 Arg Ile Ile Asp Ile Ile Ala Ser Asp Ile Gln Thr Lys Glu Leu Gln
 915 920 925
 Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
 930 935 940
 Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
 945 950 955 960
 Gly Ala Val Val Ile Gln Asp Asn Asn Glu Ile Lys Val Val Pro Arg
 965 970 975
 Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp

 980 985 990
 Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000
 <210> 238
 <211> 3012
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 238

ttcttccgcg agaacctggc cttccagcag ggcgaggccc gcaagtcttc ctccgagcag	60
acccgcgcca actccccgc ctcccgcgag ctgcgcgtgc gccgcggcga caactccctg	120
cccgaggccg gcgccgagcg ccagggcacc ggctcctccc tggacttccc ccagatcacc	180
ctgtggcagc gccccctggg gaccatcaag gtgggcggcc agctgcgcga ggccctgctg	240
gacaccggcg ccgacgacac cgtgctggag gacatcaacc tgcccggcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggttctatc aagggtgcgc agtacgacca gatccccatc	360
gagatctgcg gccagaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacatc	420
atcgccgcga acatgctgac ccagatcggc tgcacctga acttccccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcgcccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctgggt gacttccgcg agctgaacaa gcgcacccag	720
gacttctggg aggtgcagct gggcatcccc cccccgccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccttgacaa ggagtccgc	840
aagtacaccg ccttccatc cccctccatc aacaacgaga ccccgccat ccgctaccag	900
tacaacgtgc tgccccaggc ctggaagggc tccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg cgccaagaac ccgagatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgcgagcacc tgetgcgtg gggcttcacc accccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac ccgacaagt ggaccgtgca ggccatccag	1200
ctgccgaca agtctctctg gaccgtgaac gacatccaga agctggtggg caagctgaac	1260
tgggcctccc agatctacc cggcattcgc gtgaagcacc tgtgcaagct gctgcgcggc	1320
gccaaaggccc tgaccgacgt ggtgccccctg accgccgagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga gcccgatcac ggctgtact acgaccctc caaggacctg	1440
atcgccgaga tccagaagca gggccacgac cagtggacct accagatcta ccaggagccc	1500
cacaagaacc tgaagaccgg caagtacgcc cgccgcaagt ccgccacac caacgacgtg	1560
aagcagctga ccgaggtggg gcagaagatc gccaccgagg gcatcgtgat ctggggcaag	1620
gtgccaaagt tccgctgcc catccagaag gagacctggg agatctggtg gaccgagtac	1680
tggcaggcca cctggatccc cgagtgggag ttcgtgaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaccga gcccatctg ggccgagaga ccttctacgt ggacggcgcc	1800
gccaaaccgc agaccaagct gggcaaggcc ggctacgtga ccgaccgcgg ccgcagaag	1860

gtggtgcccc tgaccgagac caccaaccag aagaccgagc tgcaggccat ccacctggcc	1920
ctgcaggact cgggtccga ggtgaacatc gtgaccgact ccagtacgc cctgggcatc	1980
atccaggccc accccgacaa gtccgagtc gagctggtga accagatcat cgagcagctg	2040
atccagaagg agcgcgtgta cctgtcctgg gtccccccc acaaggccat cggcggcaac	2100
gagcaggtgg acaagctggt gtccaccggc atccgcaagg tgcgtttcct ggacggcatc	2160
gacaaggccc aggaggagca cgagaagtac cactccaact ggcgcccat ggcctccgac	2220
ttcaacctgc cccccgtggt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg	2280
aaggcgagg ccattgcacgg ccaggtggac tgcctccccg gcatttgga gctggactgc	2340
accacctgg agggcaagat catcctggtg gccgtgcacg tggcctccgg ctacatcgag	2400
gccgaggtga tccccccga gaccggccag gagaccgct acttcatcct gaagctggcc	2460
ggcgcgtggc ccgtgaagat catccacacc gacaacggct ccaacttcac ctccaccgtg	2520
gtgaaggccg cctgctggtg ggccggcatc cagcaggagt tcggcatccc ctacaacccc	2580
cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg	2640
cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc	2700
aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcattcatga catcatcgcc	2760
accgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa ctccgcgtg	2820
tacttccgcg actccccga cccgtgtggt aagggtcccg ccaagctgct gtggaagggc	2880
gaggcgccg tggatgacca ggacaacaac gagatcaagg tgggtccccc ccgcaaggcc	2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag	3000
gacgaggact aa	3012
<210> 239	
<211> 3009	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 239	
ttcttccgcg agaacctggc cttccagcag ggcgaggccc gcgagttctc ctccgagcag	60
gccccgcca actccccac ccgcccgcgag ctgcgcgtgc gccgcggcga ctccccctg	120
cccaggccg gcgcccaggg caaggcgcc atctccctgt cttccccca gatcacctg	180
tggcagcgc cctggtgac cgtgaagatc ggcgccagc tgatcgaggc cctgctggac	240
accggcgccg acgacacct gctggaggag atcaacctgc ccggcaagt gaagcccaag	300
atgatcgcg gcattcgcg cttcatcaag gtgcgccagt acgaccagat cctgatcgag	360

atctccggca agaaggccat cggcaccgtg ctggtgggcc ccaccccat caacatcatc	420
ggccgcaaca tgctgaccca gatcggctgc accctgaact tccccatctc ccccatcgag	480
accgtgcccc tgaagctgaa gcccggcatg gacggcccca aggtgaagca gtggccctg	540
accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggaag	600
atctccaaga tcggccccga gaacccttac aacaccccca tcttcgcat caagaagaag	660
gactccacca agtggcgcaa gctgtggac ttccgcgagc tgaacaagcg caccaggaac	720
ttctgggagg tgcagctggg catccccac cccgccggcc tgaagaaga gaagtccgtg	780
accgtgctgg acgtggcgga cgcctacttc tccgtgcccc tggacgagaa ctccgcaag	840
tacaccgcct tcaccatccc ctccaccaac aacgagaccc ccggcatccg ctaccgtac	900
aacgtgctgc ccagggtg gaagggtcc cccgccatct tcagtcctc catgaccaag	960
atcctggagc cttccgcaac caagaacccc gagatcgtga tctaccagta catggacgac	1020
ctgtactgtg gctccgacct ggagatcggc cagcacccg ccaagatcga ggagctgcgc	1080
gagcacctgc tgcgtgggg cttcaccacc cccgacaaga agcaccagaa ggagccccc	1140
ttctgtgga tgggctacga gctgcacccc gacaagtga cgtgcagcc catccagctg	1200
cccgacaagg agtcctggac cgtgaacgac atccagaagc tgggtggcaa gctgaactgg	1260
gcctcccaga tctaccccg catcaagggtg aagcagctgt gcaagctgct gcgcggcgcc	1320
aaggccctga ccgacatcgt gcccctgacc gccgaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg accctccaa ggagctgac	1440
gccgagggtc agaagcaggg cctggaccag tggacctacc agatctacca ggagccctac	1500
aagaacctga agaccggcaa gtacgccaag cgcggctccg cccacaccaa cgacgtgaag	1560
cagctgaccg aggtggtgca gaagatgcc accgagtcca tcgtgatctg gggcaagacc	1620
cccaagttca agctgccat ccgcaaggag acctgggagg tgtgtggac cgagtactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacacc ccccctggt gaagctgtgg	1740
taccgcctgg agaccgagcc cateccccgc gccgagacct actacgtgga cggcgccgcc	1800
aaccgcgaga ccaagctggg caaggccggc tacgtgaccg acaaggga gcagaagatc	1860
atcacctga ccgagaccac caaccagaag gccgagctgc agccatcca cctggccctg	1920
caggactccg gctccgaggt gaacatcgtg accgactccc agtacgcct gggcatcatc	1980
caggcccagc ccgaccgtc cgagtcgag ctggtgaacc agatcatcga gcagctgac	2040
aagaaggaga aggtgtacct gtcctgggtg cccgccaca agggcatcgg cggcaacgag	2100

caggtggaca agctggtgtc ctccggcatc cgcaaggtgc tgttcctgga cggcatcgac 2160

aaggccccagg aggagcacga gcgctaccac tccaactggc gcgccatggc ctccgacttc 2220

aacctgcccc ccacgtggc caaggagatc gtggcctcct gcgacaagtg ccagctgaag 2280

ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc 2340

cacctggagg gcaagatcat cctggtggcc gtgcacgtgg cctccgcta catcgaggcc 2400

gaggtgatec ccgccgagac cggccaggag accgcctact tcacctgaa gctggccggc 2460

cgctggcccc tgaaggtgat ccacaccgac aacggctcca acttcacctc cgccgccgtg 2520

aaggccgcct gctggtgggc caacatcacc caggagttcg gcaccccta caacccccag 2580

tcccaggcgc tgggtggagtc catgaacaag gagctgaaga agatcatcgg ccaggcgcgc 2640

gaccaggccg agcacctgaa gaccgccgtg cagatggccg tgttcacca caacttcaag 2700

cgcaaggcgc gcacggcgg ctactccgcc ggcgagcgca tcacgacat catgcctcc 2760

gacatccaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgcgtgtac 2820

taccgcgact cccgcgaccc catctggaag ggccccgcca agctgctgtg gaaggcgag 2880

ggcgccgtgg tgatccagga caacaacgag atcaaggtgg tgccccgccg caaggccaag 2940

atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgg ccgccaggac 3000

gaggactaa 3009

<210> 240

<211> 1001

<212> PRT

<213> Human immunodeficiency virus

<400> 240

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Arg Glu Ala Arg Lys Phe

1 5 10 15

Ser Pro Glu Gln Ala Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Arg

20 25 30

Val Arg Arg Gly Asp Asp Pro Leu Ser Glu Ala Gly Ala Glu Gly Gln

35 40 45

Gly Thr Ser Leu Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu

50 55 60

Val Thr Val Lys Ile Glu Gly Gln Leu Arg Glu Ala Leu Leu Asp Thr

65 70 75 80

Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asn Leu Pro Gly Lys Trp
 85 90 95
 Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg Gln
 100 105 110
 Tyr Glu Gln Val Ala Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly Thr
 115 120 125
 Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Ile Leu
 130 135 140
 Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr
 145 150 155 160
 Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln
 165 170 175
 Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile Cys Ile
 180 185 190

 Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro
 195 200 205
 Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp
 210 215 220
 Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe
 225 230 235 240
 Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys
 245 250 255
 Lys Ser Val Ser Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro
 260 265 270
 Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile
 275 280 285
 Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln
 290 295 300
 Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile
 305 310 315 320
 Leu Glu Pro Phe Arg Lys Gln Asn Pro Glu Met Ile Ile Tyr Gln Tyr

325								330								335															
Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln	His	Arg																
340								345								350															
Ala	Lys	Ile	Glu	Glu	Leu	Arg	Ala	His	Leu	Leu	Arg	Trp	Gly	Phe	Thr																
355								360								365															
Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp	Met	Gly																
370								375								380															
Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Val	Lys	Leu	Pro																
385								390								395								400							
Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val	Gly	Lys																
405								410								415															
Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	Val	Lys	Gln	Leu																
420								425								430															
Cys	Lys	Leu	Leu	Arg	Gly	Ala	Lys	Ala	Leu	Thr	Asp	Ile	Val	Pro	Leu																
435								440								445															
Thr	Lys	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile	Leu	Arg																
450								455								460															
Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu	Ile	Ala																
465								470								475								480							
Glu	Ile	Gln	Lys	Gln	Gly	Pro	Asp	Gln	Trp	Thr	Tyr	Gln	Ile	Tyr	Gln																
485								490								495															
Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Lys	Met	Arg	Thr																
500								505								510															
Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln	Lys	Ile																
515								520								525															
Ala	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Ile	Pro	Lys	Phe	Arg	Leu																
530								535								540															
Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr	Trp	Trp	Thr	Glu	His	Trp	Gln																
545								550								555								560							
Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	Pro	His	Leu	Val																
565								570								575															

Lys Leu Trp Tyr Gln Leu Glu Thr Glu Pro Ile Ala Gly Ala Glu Thr
 580 585 590
 Tyr Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Ile Gly Lys Ala
 595 600 605
 Gly Tyr Val Thr Asp Arg Gly Lys Gln Lys Val Val Ser Leu Thr Glu
 610 615 620

 Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln
 625 630 635 640
 Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu
 645 650 655
 Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn
 660 665 670
 Gln Ile Ile Glu Glu Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser Trp
 675 680 685
 Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu

 690 695 700
 Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys
 705 710 715 720
 Ala Gln Glu Glu His Glu Arg Tyr His Asn Asn Trp Arg Ala Met Ala
 725 730 735
 Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala Ser
 740 745 750
 Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp
 755 760 765

 Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys
 770 775 780
 Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu
 785 790 795 800
 Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu Lys
 805 810 815
 Leu Ala Gly Arg Trp Pro Val Lys Met Ile His Thr Asp Asn Gly Ser
 820 825 830

Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Asp Ile

835 840 845
Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val
850 855 860
Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp
865 870 875 880
Gln Ala Glu His Leu Arg Thr Ala Val Gln Met Ala Val Phe Ile His
885 890 895
Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg
900 905 910

Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys
915 920 925
Gln Ile Ser Lys Ile Gln Lys Phe Arg Val Tyr Tyr Arg Asp Ser Arg
930 935 940
Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly
945 950 955 960
Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro Arg Arg
965 970 975
Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp

980 985 990
Cys Val Ala Gly Arg Gln Asp Glu Asp
995 1000

<210> 241
<211> 1003
<212> PRT
<213> Human immunodeficiency virus
<400> 241

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Lys Ala Gly Glu Phe
1 5 10 15
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Lys Leu Gly
20 25 30

Asp Gly Gly Arg Asp Asn Leu Leu Thr Glu Ala Gly Ala Glu Arg Gln

35 40 45
Gly Thr Ser Ser Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg

50 55 60
Pro Leu Val Thr Val Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
65 70 75 80

Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly
85 90 95
Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
100 105 110

Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile
115 120 125

Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
130 135 140

Met Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
145 150 155 160

Asp Thr Val Pro Val Thr Leu Lys Pro Gly Met Asp Gly Pro Lys Val
165 170 175

Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile

180 185 190
Cys Lys Glu Met Glu Glu Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu

195 200 205
Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr

210 215 220
Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln

225 230 235 240
Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys

245 250 255

Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
260 265 270

Val Pro Leu Asp Glu Ser Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro

275 280 285
 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 290 295 300
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
 305 310 315 320
 Lys Ile Leu Glu Pro Phe Arg Ile Lys Asn Pro Glu Met Val Ile Tyr

 325 330 335
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 340 345 350
 His Arg Thr Lys Ile Glu Glu Leu Arg Ala His Leu Leu Ser Trp Gly
 355 360 365
 Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 370 375 380
 Met Gly Tyr Glu Leu His Pro Asp Arg Trp Thr Val Gln Pro Ile Glu
 385 390 395 400

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

 465 470 475 480
 Val Ala Glu Val Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile
 485 490 495
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Lys
 500 505 510
 Arg Ser Ala His Thr Asn Asp Val Arg Gln Leu Thr Glu Val Val Gln
 515 520 525
 Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe

530 535 540
 Arg Leu Pro Ile Gln Arg Glu Thr Trp Glu Thr Trp Trp Met Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Asp Pro Ile Val Gly Ala
 580 585 590
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Ser Arg Glu Thr Lys Leu Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu

 610 615 620
 Thr Glu Thr Thr Asn Gln Lys Thr Glu Leu His Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Val
 660 665 670
 Val Asn Gln Ile Ile Glu Glu Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685

 Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Lys Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Glu His Glu Arg Tyr His Ser Asn Trp Arg Thr
 725 730 735
 Met Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val
 740 745 750
 Ala Asn Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln

 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780

Gly Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
785 790 795 800
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu
805 810 815
Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn
820 825 830

Gly Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala
835 840 845
Asn Val Arg Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
850 855 860
Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
865 870 875 880
Arg Glu Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
885 890 895
Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
900 905 910
Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu
915 920 925
Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
930 935 940
Ser Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945 950 955 960
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
965 970 975

Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
980 985 990
Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
995 1000

<210> 242
<211> 3006
<212> DNA
<213> Human immunodeficiency virus

<400> 242

ttcttccgcg agaacctggc cttccagcag cgcgaggccc gcaagttctc ccccgagcag	60
gcccgcgcca actccccac ctcccgagag ctgcgcgtgc gccgcggcga cgacccccctg	120
tccgaggccg gcgccgaggg ccagggcacc tcctgtcct tccccagat caccctgtgg	180
cagcgccttc tggtagcgt gaagatcgag ggccagctgc gcgaggccct gctggacacc	240
ggcgccgacg acaccgtgct ggaggagatc aacctgcccg gcaagtggaa gccaagatg	300
atcggcggca tcggcggtt catcaagtg cgccagtacg agcaggtggc catcgagatc	360
tgcggcaaga aggccatcg caccgtgctg gtgggcccc ccccgtaga catcatcggc	420
cgcaacatcc tgacccagat cggtgcacc ctgaacttcc ccatctccc catcgagacc	480
gtgccctga agctgaagcc cggcatggac ggccccaaagg tgaagcagt gcccctgacc	540
gaggagaaga tcaaggccct gaccgagatc tgcacgaga tggagaagga gggcaagatc	600
tccaagatcg gccccagaa cccctacaac accccatct tcgcatcaa gaagaaggac	660
tccaccaagt ggcgcaagct ggtggacttc cgcgagctga acaagcgac ccaggacttc	720
tgggaggtgc agctgggcat cccccaccc gccggcctga agaagaagaa gtccgtgtcc	780
gtgtggacg tggcgacgc ctacttctcc gtccccctgg acaaggactt ccgaagtac	840
accgccttca ccatcccctc catcaacaac gagaccccc gcatccgcta ccagtacaac	900
gtgtgcccc agggctggaa gggctcccc gccatcttcc agtcctccat gaccaagatc	960
ctggagccct tccgaagca gaaccccgag atgatcatct accagtacat ggacgacctg	1020
tacgtgggct ccgacctgga gatcggccag caccgcgcca agatcgagga gctgcgcgc	1080
cacctgtgc gctggggctt caccaccccc gacaagaagc accagaagga gcccccttc	1140
ctgtggatgg gctacgagct gcaccccgac aagtggaccg tgcagcccgt gaagctgcc	1200
gagaaggact cctggaccgt gaacgacatc cagaagctgg tgggcaagct gaactgggcc	1260
tcccagatct accccggcat caagtggaag cagctgtgca agctgtgcg cggcgccaag	1320
gccctgaccg acatcgtgcc cctgaccaag gaggccgagc tggagctggc cgagaaccgc	1380
gagatcctgc gcgagccgt gcacggcgtg tactacgacc cctccaagga cctgatcgcc	1440
gagatccaga agcagggccc cgaccagtgg acctaccaga tctaccagga gcccttcaag	1500
aaactgaaga ccggcaagta gcacaagatg cgcaccgccc acaccaacga cgtgaagcag	1560
ctgaccgagg ccgtgcagaa gatcgccacc gattccatcg tgatctgggg caagatcccc	1620
aagtccgcc tgcccatcca gaaggagacc tgggagacct ggtggaccga gcaactggcag	1680
gccacctgga tccccgagtg ggagttcgtg aacaccccc acctgggtgaa gctgtgttac	1740
cagctggaga ccgagcccat cgccggcgcc gagacctact acgtggacgg cgccccaac	1800

cgcgagacca agatcggcaa ggccggctac gtgaccgacc gcggcaagca gaaggtggtg	1860
tccttgaccg agaccaccaa ccagaagacc gagctgcagg ccatctacct ggccctgcag	1920
gactccggcc tggaggtgaa catcgtgacc gactcccagt acgccctggg catcatccag	1980
gcccagcccc acaagtccga gtccgagctg gtgaaccaga tcatcgagga gctgatcaag	2040
aaggagaagg tgtacctgtc ctgggtgccc gcccaacaagg gcatcggcgg caacgagcag	2100
gtggacaagc tgggtgtctc cggcatccgc aaggtgctgt tcctggacgg catcgacaag	2160
gcccaggagg agcacgagcg ctaccacaac aactggcgcg ccatggcctc cgacttcaac	2220
ctgcccccca tcgtggccaa ggagatcgtg gcctcctgcg acaagtgccg gctgaagggc	2280
gaggccatgc acggccaggt ggactgctcc cccggcatct ggagctgga ctgcacccac	2340
ctggagggca aggtgatcct ggtggccgtg cacgtggcct ccggctacat cgaggccgag	2400
gtgatccccg ccgagaccgg ccaggagacc gcctacttca tctgaagct ggccggccgc	2460
tggcccgatga agatgatcca caccgacaac ggctccaact tcacctccgc cgccgtgaag	2520
gccgcctgct ggtgggcca catccagcag gatttcggca tcccctaca cccccagtc	2580
cagggcgctg tggagtccat gaacaaggag ctgaagaaga tcatcgcca ggtgcgcgac	2640
caggccgagc acctgcgcac cgccgtgcag atggccgtgt tcatccaca cttcaagcgc	2700
aaggcgccga tcggcggtc ctccgccggc gagcgcatca tcgacatcat cgccaccgac	2760
atccagacca aggagctgca gaagcagatc tccaagatcc agaagttccg cgtgtactac	2820
cgcgactccc gcgacccat ctggaagggc cccccaagc tgctgtgaa gggcgagggc	2880
gccgtggtga tccaggacaa ctccgagatc aaggtggtgc cccgccgca ggccaagatc	2940
atccgcgact acggcaagca gatggccggc gacgactgcg tggccggccg ccaggacgag	3000
gactaa	3006
<210> 243	
<211> 3012	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 243	
tttttcgcg agaacctggc ctccagcag ggcaaggccg gcgagttctc ctccgagcag	60
acccgcgcca actccccac ctccgcaag ctgggcgacg gcggccgca caacctgctg	120
accgaggccg gcgccgagcg ccagggcacc tctctctct tctcttccc ccagatcacc	180
ctgtggcagc gccccctggt gaccgtgaag atcgccggcc agctgaagga ggccctgctg	240
gacaccggcg ccgacgacac cgtgtgtggag gacatcaacc tgcccgcaa gtggaagccc	300

aagatgatcg gcggcatcgg cggcttcac c aaggtgccc agtacgacca gatcctgac	360
gagatctgcg gcaagaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacac	420
atcgcccgca acatgctgac ccagatcggc tgcacctga acttcccat ctccccac	480
gacaccgtgc ccgtgacct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca aggagatgga ggaggagggc	600
aagatctcca agatcggccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcaccag	720
gacttctggg aggtgcagct gggcatcccc cccccgcg gcctgaagaa gaagaagtc	780
gtgaccgtgc tggacgtggg cgagccctac ttctccgtgc ccttgacga gtccttccgc	840
aagtacaccg ccttcacat cccctccac aacaacgaga cccccgcat ccgtaccag	900
tacaacgtgc tgccccaggc ctggaagggc tccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg catcaagaac cccgagatgg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgccccacc tgctgtcctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgaccgt ggaccgtgca gcccacgag	1200
ctgccccgaga aggactcctg gaccgtgaac gacatccaga agctggtggg caagctgaac	1260
tgggcctccc agatctacg cggcacaaag gtgaagcagc tgtgcaagct gctgcgcggc	1320
gccaaggccc tgaccgacat cgtgccccctg accgaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagac cccgtgcac ggcgtgtact acgaccctc caaggacctg	1440
gtggccgagg tgcagaagca gggccaggac cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgc cgcaagcgt ccccccacac caacgacgtg	1560
cgccagctga ccgaggtggt gcagaagatc gccaccgagt ccatcgtgat ctggggcaag	1620
acccccaaat tccgcctgcc catccagcgc gagacctggg agacctggtg gatggagtac	1680
tggcaggcca cctggatccc cgagtgggag ttctgaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaagga ccccatcgtg ggcgccgaga ctttctacgt ggacggcgcc	1800
gcctcccgcg agaccaagct gggcaaggcc ggctacgtga ccgaccgcgg ccgccagaag	1860
gtggtgtccc tgaccgagac caccaaccag aagaccgagc tgcacgcat ccacctggcc	1920
ctgcaggact ccggctccga ggtgaacatc gtgaccgact cccagtacgc cctgggcatc	1980
atccaggccc agccccaccg ctccgagtc gaggtggtga accagatcat cgaggagctg	2040

atcaagaagg agaaggtgta cctgtcctgg gtgccccccc acaagggcat cggcggcaac 2100
gagcaggtgg acaagctggt gtccctccggc atccgcaagg tgctgttcct ggacggcatc 2160
gacaaggccc aggaggagca cgagcgctac cactccaact ggcgccacat ggccctccgac 2220
ttcaacctgc ccccatcgt ggccaaggag atcgtggcca actgcgacaa gtgccagctg 2280
aaggcgagg ccatgcacgg ccaggtggac tgctcccccg gcatctggca gctggactgc 2340
accacactgg agggcaaggt gatcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400
gccgaggtga tccccgccga gaccggccag gagaccgcct acttcctgct gaagctggcc 2460

ggccgctggc ccgtgaaggt gatccacacc gacaacggct ccaacttcac ctccgccgcc 2520
gtgaaggccg cctgctggtg ggccaacgtg cgccaggagt tcggcatccc ctacaacccc 2580
cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg 2640
cgcgagcagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcac ccacaacttc 2700
aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcatcatcga catcatcgcc 2760
accgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa ctccgcgtg 2820
tactaccgag actccccgga ccccatctgg aaggcccccg ccaagctgct gtggaagggc 2880

gaggcgccg tggatgatcca ggacaactcc gacatcaagg tgggtccccg ccgcaaggcc 2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag 3000
gacgaggact aa 3012

<210> 244
<211> 1003
<212> PRT
<213> Human immunodeficiency virus
<400> 244

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Lys Phe

1 5 10 15

Ser Ser Glu Gln Thr Gly Thr Asn Ser Pro Thr Ser Arg Glu Leu Trp

20 25 30

Asp Gly Gly Arg Asp Asn Leu Leu Ser Glu Ala Gly Thr Glu Gly Gln

35 40 45

Gly Thr Ile Ser Ser Phe Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg

50 55 60

Pro Leu Val Thr Val Arg Ile Gly Gly Gln Leu Ile Glu Ala Leu Leu

65 70 75 80

Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asn Leu Pro Gly
85 90 95

Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
100 105 110

Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile
115 120 125

Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
130 135 140

Met Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
145 150 155 160

Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val

165 170 175

Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Asp Ile
180 185 190

Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
195 200 205

Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
210 215 220

Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
225 230 235 240

Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
245 250 255

Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
260 265 270

Val Pro Leu Asp Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
275 280 285

Ser Val Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
290 295 300

Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ala Ser Met Thr

305 310 315 320

Lys Ile Leu Glu Pro Phe Arg Thr Lys Asn Pro Glu Ile Val Ile Tyr

325 330 335
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 340 345 350
 His Arg Ala Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly
 355 360 365
 Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
 370 375 380

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

 450 455 460
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Thr Lys Asp Leu
 465 470 475 480
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile
 485 490 495
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met
 500 505 510
 Arg Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Val Val Gln
 515 520 525

 Lys Val Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 530 535 540
 Arg Leu Pro Ile Gln Arg Glu Thr Trp Glu Ala Trp Trp Met Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Asp Pro Ile Val Gly Ala

580 585 590
Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly

595 600 605
Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
610 615 620

Thr Glu Thr Thr Asn Gln Lys Thr Glu Leu His Ala Ile His Leu Ala
625 630 635 640

Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
645 650 655

Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Leu
660 665 670

Val Asn Gln Ile Ile Glu Lys Leu Ile Glu Lys Asp Lys Val Tyr Leu
675 680 685

Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
690 695 700

Lys Leu Val Ser Asn Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
705 710 715 720

Asp Lys Ala Gln Glu Glu His Glu Arg Tyr His Ser Asn Trp Arg Ala
725 730 735

Met Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val

740 745 750
Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln

755 760 765
Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu

770 775 780
Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu

785 790 795 800
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile

805 810 815
Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn

820 825 830

Gly Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala
835 840 845

Asn Val Thr Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
850 855 860

Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val
865 870 875 880

Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
885 890 895

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
900 905 910

Glu Arg Ile Ile Asp Ile Ile Ala Ser Asp Ile Gln Thr Lys Glu Leu
915 920 925

Gln Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
930 935 940

Ser Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945 950 955 960

Glu Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro
965 970 975

Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly
980 985 990

Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
995 1000

<210> 245
<211> 1003
<212> PRT
<213> Human immunodeficiency virus
<400> 245

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Arg Glu Ala Arg Lys Phe
1 5 10 15

Ser Ser Glu Gln Thr Arg Ala Ile Ser Pro Thr Ser Arg Lys Leu Trp
20 25 30

Asp Gly Gly Arg Asp Asn Pro Leu Pro Glu Thr Gly Thr Glu Arg Gln

35 40 45
 Gly Thr Ala Ser Ser Phe Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg
 50 55 60
 Pro Leu Val Thr Val Arg Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu
 65 70 75 80
 Asp Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly

 85 90 95
 Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val
 100 105 110
 Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile
 115 120 125
 Gly Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn
 130 135 140
 Met Leu Thr Gln Leu Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile
 145 150 155 160

 Glu Thr Val Pro Val Thr Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 165 170 175
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Asp Ile
 180 185 190
 Cys Lys Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 195 200 205
 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 210 215 220
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln

 225 230 235 240
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 245 250 255
 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
 260 265 270
 Val Pro Leu Asp Gln Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 275 280 285
 Ser Thr Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu

290								295								300								
Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	Met	Thr									
305								310								315								320
Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Glu	Ile	Val	Ile	Tyr									
				325				330				335												
Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln									
				340				345				350												
His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Glu	His	Leu	Leu	Arg	Trp	Gly									
355								360				365												
Phe	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp									
370								375				380												
Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Val									
385								390								395								400
Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val									
				405				410				415												
Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Ala	Gly	Ile	Lys	Val	Arg									
				420				425				430												
Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Ala	Lys	Ala	Leu	Thr	Glu	Val	Ile									
435								440				445												
Pro	Leu	Thr	Ala	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile									
450								455				460												
Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	Asp	Leu									
465								470								475								480
Val	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	Gln	Ile									
				485				490				495												
Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Arg	Leu									
500								505				510												
Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	Val	Gln									
515								520				525												
Lys	Ile	Ala	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	Lys	Phe									
530								535								540								

Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 545 550 555 560
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 565 570 575
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 580 585 590

 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Ser Gly
 595 600 605
 Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu
 610 615 620
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile His Leu Ala
 625 630 635 640
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 645 650 655
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu

 660 665 670
 Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 675 680 685
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 690 695 700
 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile
 705 710 715 720
 Asp Lys Ala Gln Glu Ala His Glu Lys Tyr His Ser Asn Trp Arg Ala
 725 730 735

 Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val
 740 745 750
 Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln
 755 760 765
 Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu
 770 775 780
 Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu
 785 790 795 800

Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Val

805 810 815

Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Ile Ile His Thr Asp Asn

820 825 830

Gly Ser Asn Phe Ile Ser Thr Ala Val Lys Ala Ala Cys Trp Trp Ala

835 840 845

Gly Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly

850 855 860

Val Val Glu Ser Met Asn Lys Gln Leu Lys Gln Ile Ile Gly Gln Val

865 870 875 880

Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe

885 890 895

Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly

900 905 910

Glu Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu

915 920 925

Gln Lys Gln Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp

930 935 940

Ser Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly

945 950 955 960

Glu Gly Ala Val Val Ile Gln Asp Asn Asn Asp Ile Lys Val Val Pro

965 970 975

Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly

980 985 990

Asp Asp Cys Val Ala Ser Arg Gln Asp Glu Asp

995 1000

<210> 246

<211> 3012

<212> DNA

<213> Human immunodeficiency virus

<400> 246

ttcttccg cg agaacctggc cttccagcag ggcgaggccc gcaagttctc ctccgagcag	60
accggcacca actccccac ctcccgcgag ctgtgggacg gcggccgcga caacctgctg	120
tccgaggccg gcaccgaggg ccagggcacc atctctctct tcaacttccc ccagatcacc	180
ctgtggcagc gccccctggt gaccgtgcgc atcggcgccc agctgatcga ggccctgctg	240
gacaccggcg ccgacgacac cgtgctggag gagatcaacc tgcccggcaa gtggaagccc	300
aagatgatcg gggcatcgg cggttctatc aagggtgcgc agtacgacca gatcctgac	360
gagatctgcg gcaagaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacatc	420
atcgccgca acatgctgac ccagatcggc tgcacctga acttccccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gacatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcgcccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccg cgctgaacaa gcgcaccag	720
gacttctggg aggtgcagct gggcatcccc cccccgccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggagctggg cgacgcctac ttctccgtgc ccctggacaa ggacttccgc	840
aagtacaccg ccttaccat cccctccgtg aacaacgaga ccccgccat ccgtaccag	900
tacaacgtgc tgccccaggc ctggaagggc tccccgcca tcttccagc ctccatgacc	960
aagatcctgg agcccttccg caccaagaac cccgagatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcgccaagat cgaggagctg	1080
cgcgagcacc tgetgcgtg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccattcag	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctggtggg caagctgaac	1260
tgggcctccc agatctacg cggcattcaag gtgaagcagc tgtgcaagct gctgcgcggc	1320
gccaaaggccc tgaccgacat cgtgacctg accgaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga gcccggtgcac ggcgtgtact acgacccac caaggacctg	1440
atcgccgaga tccagaagca gggccaggac cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccg caagtacgc aagatgcgt cgcgccacac caacgacgtg	1560
aagcagctga ccgagtggt gcagaagggt gccaccgagt ccatcgtgat ctggggcaag	1620
accccaagt tccgctgcc catccagcgc gagacctggg aggcctggtg gatggagtac	1680
tggcaggcca cctggatccc cgagtgggag ttcgtgaaca cccccccct ggtgaagctg	1740

tggtaccagc tggagaagga ccccatcgtg ggcgccgaga ccttctacgt ggacggcgcc	1800
gccaaccgcg agaccaagct gggcaaggcc ggctacgtga ccgaccgcgg ccgccagaag	1860
gtggtgtccc tgaccgagac caccaaccag aagaccgagc tgcacgccat ccacctggcc	1920
ctgcaggact ccggctccga ggtgaacatc gtgaccgact cccagtacgc cctgggcatc	1980
atccaggccc agcccagacc ctccgagtcc gagctggtga accagatcat cgagaagctg	2040
atcgagaagg acaaggtgta cctgtcctgg gtgccccccc acaagggcat cggcggcaac	2100
gagcaggtgg acaagctggg gtccaacggc atccgcaagg tgctgttcct ggacggcatc	2160
gacaaggccc aggaggagca cgagcgctac cactccaact ggcgcgccat ggctccgac	2220
ttcaacctgc ccccatcgt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg	2280
aaggcgagg ccattgcacgg ccaggtggac tgctccccg gcatctggca gctggactgc	2340
accacctgg agggcaagat catcctggtg gccgtgcacg tggcctccgg ctacatcgag	2400
gccgaggtga tccccgccga gaccggccag gagaccgcct attcatcct gaagctggcc	2460
ggccgctggc ccgtgaaggt gatccacacc gacaacggct ccaacttcac ctccgccgcc	2520
gtgaaggccg cctgctggtg ggccaacgtg acccaggagt tcggcatccc ctacaacccc	2580
cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat cggccaggtg	2640
cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcac ccacaacttc	2700
aagcgcaagg gcggcatcgg ccgctactcc gccggcgagc gcatcatcga catcatgcc	2760
tccgacatcc agaccaagga gctgcagaag cagatcacca agatccagaa ctccgcgtg	2820
tactaccgcg actcccgcca ccccatctgg aaggcccccg ccaagctgct gtggaagggc	2880
gagggcgccg tggatgacca ggacaactcc gacatcaagg tggcgccccg ccgcaaggcc	2940
aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag	3000
gacgaggact aa	3012
<210> 247	
<211> 3012	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 247	
ttcttccgcg agaacctggc cttccagcag cgcgaggccc gcaagttctc ctccgagcag	60
accgcgcga tctccccac ctccgcaag ctgtgggacg gcggccgcga caaccctg	120
cccgagaccg gcaccgagcg ccagggcacc gcctctctct tcaacttccc ccagatcacc	180
ctgtggcagc gccccctggt gaccgtgcgc atcggcggcc agctgaagga ggccctgctg	240

gacaccggcg ccgacgacac cgtgctggag gacatcaacc tgcccggcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggcttcac aaggtgcgcc agtacgacca gatcctgac	360
gagatctgcg gcaagaaggc catcggcacc gtgctggtgg gcccacccc cgtgaacatc	420
atcgccgca acatgctgac ccagctgggc tgcacctga acttcccat ctccccatc	480
gagaccgtgc ccgtgacct gaagcccgcc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gacatctgca aggagatgga gaaggagggc	600
aagatctcca agatcggccc cgagaacccc tacaacccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctggtg gacttccgcg agctgaacaa gcgcaccag	720
gacttctggg aggtgcagct gggcatcccc cccccgccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccctggacca ggacttccgc	840
aagtacaccg ccttcacat cccctccacc aacaacgaga ccccgccat ccgtaccag	900
tacaacgtgc tgccccaggc ctggaagggc tccccgcca tcttccagtc ctccatgacc	960
aagatcctgg agcccttccg caagcagaac cccgagatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgcgagcacc tgctgcgctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac cccgacaagt ggaccgtgca gcccacgtg	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctgggtggg caagctgaac	1260
tgggcctccc agatctacg cggcatcaag gtgcgccagc tgtgcaagct gctgcgcggc	1320
gccaaggccc tgaccgaggt gateccccctg accgccgagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga gcccgtgcac ggctgtact acgaccctc caaggacctg	1440
gtggccgaga tccagaagca gggccagggc cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgcc cgcctgcgcg gcgccacac caacgacgtg	1560
aagcagctga ccgaggccgt gcagaagatc gccaccgagt ccatcgtgat ctggggcaag	1620
acccccaaat tcaagctgcc catccagaag gagacctggg agacctggtg gaccgagtac	1680
tggcaggcca cctggatccc cgagtgggag ttcgtgaaca cccccccct ggtgaagctg	1740
tggtagcagc tggagaagga gcccacgtg ggcccgaga ccttctacgt ggacggcgcc	1800
gccaaccgcg agaccaagtc cggcaaggcc ggctacgtga ccgaccgagg ccgccagaag	1860
gtggtgtccc tgaccgacac caccaaccag aagaccgagc tgcaggccat ccacctggcc	1920
ctgcaggact ccggcctgga ggtgaacatc gtgaccgact ccagtagc cctgggcatc	1980

atccaggccc agcccgacaa gtccgagtcc gagctggtgt cccagatcat cgagcagctg 2040
 atcaagaagg agaaggtgta cctggcctgg gtgcccgcac acaagggcat cggcggcaac 2100
 gagcaggtgg acaagctggt gtccgccggc atccgcaagg tgctgttcct ggacggcatc 2160
 gacaaggccc aggaggccca cgagaagtac cactccaact ggcgcgccat ggccctccgac 2220
 ttcaacctgc cccccgtggt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg 2280
 aagggcgagg ccatgcacgg ccaggtggac tgctcccccg gcatctggca gctggactgc 2340
 acccacctgg agggaagat catcctggtg gccgtgcacg tggcctccgg ctacatcgag 2400

gccgaggtga tccccgccga gaccggccag gagaccgct acttcgtgct gaagctggcc 2460
 ggccgctggc ccgtgaagat catccacacc gacaacggct ccaacttcat ctccaccgcc 2520
 gtgaaggccg cctgctggtg ggccggcatc aagcaggagt tcggcatccc ctacaacccc 2580
 cagtcccagg gcgtggtgga gtccatgaac aagcagctga agcagatcat cggccaggtg 2640
 cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc 2700
 aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcatcatcga catcatcgcc 2760
 accgacatcc agaccaagga gctgcagaag cagatcatca agatccagaa ctccgcgtg 2820

tactaccgcg actccccgga ccccatctgg aagggtcccg ccaagctgct gtggaagggc 2880
 gagggcgccg tggatgatcca ggacaacaac gacatcaagg tgggtccccc ccgcaaggcc 2940
 aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc ctcccgccag 3000
 gacgaggact aa 3012

<210> 248
 <211> 1002
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 248

Phe Phe Arg Glu Asn Val Ala Phe Gln Gln Arg Glu Ala Arg Lys Phe
 1 5 10 15

Ser Ser Glu Gln Ala Arg Ala Asn Ser Pro Ala Arg Arg Glu Leu Arg
 20 25 30

Asp Glu Arg Gly Asp Asn Leu Leu Ser Glu Ala Gly Thr Glu Gly Gln
 35 40 45

Gly Thr Ile Ser Phe Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro
 50 55 60

Leu Val Thr Ile Lys Ile Gly Gly Gln Ile Arg Glu Ala Leu Leu Asp

65						70						75						80
Thr	Gly	Ala	Asp	Asp	Thr	Val	Leu	Glu	Glu	Ile	Asn	Leu	Pro	Gly	Lys			
					85						90						95	
Trp	Lys	Pro	Lys	Met	Ile	Gly	Gly	Ile	Gly	Gly	Phe	Ile	Lys	Val	Arg			
					100						105						110	
Gln	Tyr	Asp	Gln	Ile	Pro	Ile	Glu	Ile	Cys	Gly	Lys	Lys	Ala	Ile	Gly			
					115						120						125	
Thr	Val	Leu	Val	Gly	Pro	Thr	Pro	Val	Asn	Ile	Ile	Gly	Arg	Asn	Met			
					130						135						140	
Leu	Thr	Gln	Leu	Gly	Cys	Thr	Leu	Asn	Phe	Pro	Ile	Ser	Pro	Ile	Glu			
					145						150						155	160
Thr	Val	Pro	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro	Lys	Val	Lys			
					165						170						175	
Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Thr	Glu	Ile	Cys			
					180						185						190	
Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly	Pro	Glu	Asn			
					195						200						205	
Pro	Tyr	Asn	Thr	Pro	Ile	Phe	Ala	Ile	Lys	Lys	Lys	Asn	Ser	Thr	Arg			
					210						215						220	
Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg	Thr	Gln	Asp			
					225						230						235	240
Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	Leu	Lys	Lys			
					245						250						255	
Lys	Lys	Ser	Val	Thr	Val	Leu	Asp	Val	Gly	Asp	Ala	Tyr	Phe	Ser	Val			
					260						265						270	
Pro	Leu	Asp	Pro	Glu	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	Ile	Pro	Ser			
					275						280						285	
Thr	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	Val	Leu	Pro			
					290						295						300	
Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Cys	Ser	Met	Thr	Lys			
					305						310						315	320

Ile Leu Glu Pro Phe Arg Thr Lys Asn Pro Glu Ile Val Ile Tyr Gln
325 330 335

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His
340 345 350

Arg Ala Lys Ile Glu Glu Leu Arg Glu His Leu Leu Arg Trp Gly Phe
355 360 365

Ser Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
370 375 380

Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu
385 390 395 400

Ala Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
405 410 415

Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys Gln
420 425 430

Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro
435 440 445

Leu Thr Thr Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
450 455 460

Lys Glu Pro Val His Gly Ala Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
465 470 475 480

Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr
485 490 495

Gln Glu Pro Tyr Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Thr Arg
500 505 510

Ser Ala His Thr Asn Asp Val Arg Gln Leu Thr Glu Ala Val Gln Lys
515 520 525

Ile Ala Met Glu Cys Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg
530 535 540

Leu Pro Ile Gln Lys Glu Thr Trp Asp Thr Trp Trp Thr Glu Tyr Trp
545 550 555 560

Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu
565 570 575

Val Lys Leu Trp Tyr Gln Leu Glu Thr Asp Pro Ile Ala Gly Ala Glu
580 585 590

Thr Phe Tyr Val Asp Gly Ala Ala Ser Arg Glu Thr Lys Gln Gly Lys
595 600 605

Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu Ser
610 615 620

Glu Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala Leu
625 630 635 640

Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
645 650 655

Ile Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Asp Leu Val
660 665 670

Asn Gln Ile Ile Glu Gln Leu Ile Gln Lys Asp Lys Val Tyr Leu Ser
675 680 685

Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
690 695 700

Leu Val Ser Asn Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
705 710 715 720

Lys Ala Gln Glu Glu His Glu Lys Tyr His Asn Asn Trp Arg Ala Met
725 730 735

Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala
740 745 750

Ser Cys Asn Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
755 760 765

Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly
770 775 780

Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala
785 790 795 800

Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu
805 810 815

Lys Leu Ala Gly Arg Trp Pro Val Lys Ile Ile His Thr Asp Asn Gly

820 825 830
Pro Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Asp
835 840 845
Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
850 855 860
Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
865 870 875 880

Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
885 890 895
His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
900 905 910
Arg Ile Ile Asp Ile Ile Ala Ser Asp Ile Gln Thr Lys Glu Leu Gln
915 920 925
Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
930 935 940
Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu

945 950 955 960
Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg
965 970 975
Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp
980 985 990
Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
995 1000

<210> 249
<211> 1002
<212> PRT
<213> Human immunodeficiency virus
<400> 249

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Glu Phe

1 5 10 15
Ser Ser Glu Gln Ala Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Arg
20 25 30

Val Arg Arg Gly Asp Ser Pro Leu Pro Glu Ala Gly Ala Glu Gly Gln
35 40 45

Gly Ala Ile Ser Leu Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro
50 55 60

Leu Val Thr Val Arg Ile Gly Gly Gln Leu Ile Glu Ala Leu Leu Asp
65 70 75 80

Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly Lys
85 90 95

Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg
100 105 110

Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly
115 120 125

Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Met
130 135 140

Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu
145 150 155 160

Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys
165 170 175

Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile Cys
180 185 190

Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn
195 200 205

Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys
210 215 220

Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp
225 230 235 240

Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys
245 250 255

Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val
260 265 270

Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser
275 280 285

Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro

290 295 300

Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Ile Lys

305 310 315 320

Ile Leu Glu Pro Phe Arg Ile Lys Asn Pro Glu Ile Val Ile Tyr Gln

325 330 335

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His

340 345 350

Arg Ala Lys Ile Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe

355 360 365

Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met

370 375 380

Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu

385 390 395 400

Pro Asp Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly

405 410 415

Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys Gln

420 425 430

Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro

435 440 445

Leu Thr Ala Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu

450 455 460

Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile

465 470 475 480

Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr

485 490 495

Gln Glu Pro His Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Ile Lys

500 505 510

Ser Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys

515 520 525

Ile Ala Leu Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg

530 535 540
 Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp
 545 550 555 560
 Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu
 565 570 575
 Val Lys Leu Trp Tyr Gln Leu Glu Thr Glu Pro Ile Val Gly Ala Glu

 580 585 590
 Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Lys Gly Lys
 595 600 605
 Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val Ser Leu Thr
 610 615 620
 Glu Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Asn Leu Ala Leu
 625 630 635 640
 Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
 645 650 655

 Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val
 660 665 670
 Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser
 675 680 685
 Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
 690 695 700
 Leu Val Ser Thr Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
 705 710 715 720
 Lys Ala Gln Glu Asp His Glu Arg Tyr His Ser Asn Trp Arg Ala Met

 725 730 735
 Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala
 740 745 750
 Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
 755 760 765
 Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly
 770 775 780
 Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala

785 790 795 800

Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu

805 810 815

Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly

820 825 830

Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Asn

835 840 845

Ile Thr Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val

850 855 860

Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg

865 870 875 880

Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile

885 890 895

His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu

900 905 910

Arg Ile Ile Asp Ile Ile Ala Ser Asp Ile Gln Thr Lys Glu Leu Gln

915 920 925

Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser

930 935 940

Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu

945 950 955 960

Gly Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro Arg

965 970 975

Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp

980 985 990

Asp Cys Val Ala Gly Arg Gln Asp Glu Asp

995 1000

<210> 250

<211> 3009

<212> DNA

<213> Human immunodeficiency virus

<400> 250

ttcttccgcg agaacgtggc cttccagcag cgcgaggccc gcaagttctc ctccgagcag	60
gcccgcgcca actccccgc cgcgcgcgag ctgcgcgacg agcgcggcga caacctgctg	120
tccgaggcgc gcaccgaggg ccagggcacc atctccttca acttccccca gatcacctg	180
tggcagcgcc ccttggtgac catcaagatc ggcgccaga tccgcgaggc cctgctggac	240
accggcgccg acgacaccgt gctggaggag atcaacctgc cgggcaagt gaagcccaag	300
atgatcggcg gcatcggcgg cttcatcaag gtgcgccagt acgaccagat ccccatcgag	360
atctgcggca agaaggccat cggcacctg ctggtgggccc ccacccccgt gaacatcatc	420
ggccgcaaca tgctgacca gctgggctgc accctgaact tccccatctc ccccatcgag	480
accgtgcccc tgaagctgaa gcccggcatg gacggcccca aggtgaagca gtggccctg	540
accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggcaag	600
atctccaaga tcggccccga gaaccctac aacaccccca tcttcgcat caagaagaag	660
aactccacc gctggcgcaa gctggtggac ttccgcgagc tgaacaagcg caccaggac	720
ttctgggagg tgcagctggg catccccac ccgcccggcc tgaagaagaa gaagtccgtg	780
accgtgctgg acgtggcgga cgcctacttc tccgtgcccc tggaccccga gttccgcaag	840
tacaccgct tcaccatccc ctccaccaac aacgagacc cggcatccg ctaccagtac	900
aacgtgctgc ccagggtg gaagggtcc ccgccatct tccagtgtc catgaccaag	960
atcctggagc ctttcgcac caagaacccc gagatcgtga tctaccagta catggacgac	1020
ctgtactggt gctccgacct ggagatcggc cagcaccgcg ccaagatcga ggagctgcgc	1080
gagcacctgc tgcgctgggg cttctccacc ccgacaaga agcaccagaa ggagccccc	1140
ttcctgtgga tgggctacga gctgcacccc gacaagtga cgtgcagcc catccagctg	1200
gccgagaagg actcctggac cgtgaacgac atccagaagc tgggtgggcaa gctgaactgg	1260
gcctcccaga tctaccccg catcaagtg aagcagctgt gcaagctgct gcgcgcgcc	1320
aaggccctga ccgacatcgt gcccctgacc accgaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gcctactacg accctccaa ggacctgatc	1440
gccgagatcc agaagcaggg ccagggccag tggacctacc agatctacca ggagccctac	1500
aagaacctga agaccggcaa gtacgccaag acccgctccg cccacaccaa cgacgtgcgc	1560
cagctgaccg aggccgtgca gaagatgcc atggagtga tctgatctg gggcaagacc	1620
cccaagttcc gcctgcccac ccagaaggag acctgggaca cctggtggac cgagtactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacaccc cccccctggt gaagctgtgg	1740
taccagctgg agaccgacc catcgccggc gccgagacct tctacgtgga cggcgccgc	1800

tcccgcgaga ccaagcaggg caaggccggc tacgtgaccg accgcggcgg ccagaaggtg	1860
gtgtccctgt ccgagaccac caaccagaag accgagctgc aggccatcta cctggccctg	1920
caggactccg gctccgaggt gaacatcgtg accgactccc agtacccat cggcatcatc	1980
caggcccagc ccgaccgctc cgagtcggac ctggtgaacc agatcatcga gcagctgac	2040
cagaaggaca aggtgtacct gtctgggtg cccgccaca agggcatcgg cggcaacgag	2100
caggtggaca agctggtgtc caacggcatc cgcaaggtgc tgttcctgga cggcatcgac	2160
aaggcccagg aggagcacga gaagtaccac aacaactggc gcgccatggc ctccgacttc	2220
aacctgcccc ccgtgggtgc caaggagatc gtggcctcct gcaacaagt ccagctgaag	2280
ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc	2340
cacctggagg gcaagatcat cctggtggcc gtgcacgtgg cctccggcta catcgaggcc	2400
gaggtgatec ccgccgagac cggccaggag accgcctact tcctcctgaa gctggccggc	2460
cgctggcccc tgaagatcat ccacaccgac aacggcccca acttcacctc cgccgcgtg	2520
aaggccgctt gctggtgggc cgacatccag caggagtctg gcatccccta caaccccag	2580
tcccagggcg tgggtgagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc	2640
gaccaggccg agcacctgaa gaccgccgtg cagatggcgg tgttcacca caacttcaag	2700
cgcaagggcg gcatcggcgg ctactccgcc ggcgagcgca tcacgacat catcgctcc	2760
gacatccaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgcgtgtac	2820
taccgcgact cccgcgaccc catctggaag ggccccgcca agctgctgtg gaaggcgag	2880
ggcgccgtgg tgatccagga caactccgac atcaaggtgg tgccccgccc caaggccaag	2940
atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgg ccgccaggac	3000
gaggactaa	3009
<210> 251	
<211> 3009	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 251	
ttttccgcg agaacctggc cttccagcag ggcgaggccc gcgagttctc ctccgagcag	60
gccccgcgca actccccac ccgccgcgag ctgcgcgtgc gccggcgga cttccccctg	120
cccgaggccg gcgccgaggg ccagggcgcc atctccctgt cttccccca gatcacctg	180
tggcagcgcc cctggtgac cgtgcgcac gccggccagc tgatcgaggc cctgctggac	240
accggcgccc acgacacctg gctggaggac atcaacctgc ccggcaagt gaagccaag	300

atgatcgcg gcatcgcggt cttcatcaag gtgcgccagt acgaccagat cctgatcgag	360
atctgcggca agaaggccat cggcaccgtg ctggtgggcc ccacccccgt gaacatcatc	420
ggccgcaaca tgctgaccca gatcggtgac accctgaact tccccatctc ccccatcgag	480
accgtgcccc tgaagctgaa gcccggcatg gacggcccca aggtgaagca gtggccctg	540
accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggaag	600
atctccaaga tcggccccga gaacccttac aacaccccca tcttcgcat caagaagaag	660
gactccacca agtggcgcaa gctggtggac ttccgcgagc tgaacaagcg caccaggac	720
ttctgggagg tgcagctggg catccccac cccgccggcc tgaagaagaa gaagtccgtg	780
accgtgctgg acgtggcgca cgcctacttc tccgtgcccc tggacgagga ctccgcaag	840
tacaccgct tcaccatccc ctccatcaac aacgagacc cggcatccg ctaccgtac	900
aacgtgctgc cccagggtg gaagggtcc cccgccatct tccagtctc catgatcaag	960
atcctggagc ctttcgcat caagaacccc gagatcgtga tctaccagta catggacgac	1020
ctgtacgtgg gctccgacct ggagatcggc cagcaccgcg ccaagatcga ggagctgcgc	1080
gagcacctgc tgaagtgggg cttcaccacc ccgacaaga agcaccagaa ggagccccc	1140
ttcctgtgga tgggctacga gctgcacccc gacaagtga cgtgcagcc catccagctg	1200
cccgacaagg actcctggac cgtgaacgac atccagaagc tgggtggcaa gctgaactgg	1260
gcctcccaga tctacccgg catcaagggt aagcagctgt gcaagctgct gcgcggcgcc	1320
aaggccctga ccgacatcgt gccctgacc gccgaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg acccctcaa ggacctgatc	1440
gccgagatcc agaagcaggg ccagggccag tggacctacc agatctacca ggagccccc	1500
aagaacctga agaccggcaa gtacggcgcc atcaagtccg cccacaccaa cgacgtgaag	1560
cagctgaccg aggccgtgca gaagatgcc ctggagtcca tctgatctg gggcaagacc	1620
cccaagtacc gcctgccccat ccagaaggag acctgggaga cctggtggac cgagtactgg	1680
caggccacct ggatccccga gtgggagtgc gtgaacaccc cccccctgt gaagctgtgg	1740
taccagctgg agaccgagcc catcgtgggc gccgagacct tctacgtgga cggcgccgcc	1800
aaccgcgaga ccaagaaggg caaggccggc tacgtgaccg accgcggccg ccagaagggtg	1860
gtgtccctga ccgagaccac caaccagaag accgagctgc aggccatcaa cctggccctg	1920
caggactccg gctccgaggt gaacatctg accgactccc agtacgcctt gggtcatcatc	1980
caggcccagc ccgacaagtc cgagtccgag ctggtgaacc agatcatcga gcagctgatc	2040
aagaaggaga aggtgtacct gtccctgggtg cccgccaca agggcatcgg cggcaacgag	2100
cagggtggaca agctggtgtc caccggcatc cgcaagggtc tgttcctgga cggcatcgac	2160

aaggcccagg aggaccacga gcgctaccac tccaactggc gcgccatggc ctccgacttc 2220
aacctgcccc ccatcgtagg caaggagatc gtggcctcct gcgacaagtg ccagctgaag 2280

ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc 2340
cacctggagg gcaagatcat cctggtaggc gtgcacgtgg cctccggcta catcgaggcc 2400
gaggtgatcc ccgccgagac cggccaggag accgcctact tcacctgaa gctggccggc 2460
cgctggcccc tgaaggatgat ccacaccgac aacggctcca acttcacctc cgccgccgtg 2520
aaggccgctt gctggtaggc caacatcacc caggagtctg gcatccccta caacccccag 2580
tcccaggcgc tggtaggagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc 2640
gaccaggcgc agcacctgaa gaccgccgtg cagatggcgc tgttcacca caacttcaag 2700

cgcaaggcgc gcatcgccgc ctactccgcc ggcgagcgca tcacgacat catgcctcc 2760
gacatccaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgctgtac 2820
taccgcgact cccgcgacct catctggaag ggccccgcca agctgctgtg gaaggcgag 2880
ggcgccgtgg tgatccagga caactccgag atcaaggtgg tgccccgcgc caaggccaag 2940
atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgc ccgccaggac 3000
gaggactaa 3009

<210> 252
<211> 999
<212> PRT

<213> Human immunodeficiency virus
<400> 252

Phe Phe Arg Glu Ile Leu Ala Phe Pro Gln Gly Glu Ala Arg Glu Phe
1 5 10 15
Pro Pro Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Gln
20 25 30
Val Arg Gly Asp Asn Pro Ser Ser Glu Ala Gly Thr Glu Arg Gln Gly
35 40 45
Thr Leu Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Ser
50 55 60

Ile Lys Val Gly Gly Gln Ile Lys Glu Ala Leu Leu Asp Thr Gly Ala
65 70 75 80
Asp Asp Thr Val Leu Glu Glu Val Asn Leu Pro Gly Lys Trp Lys Pro

85 90 95
 Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Glu
 100 105 110
 Gln Ile Pro Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly Thr Val Leu
 115 120 125
 Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Met Leu Thr Gln

 130 135 140
 Leu Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr Val Pro
 145 150 155 160
 Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro
 165 170 175
 Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Asp Glu Met
 180 185 190
 Glu Lys Glu Gly Lys Ile Thr Lys Ile Gly Pro Asp Asn Pro Tyr Asn
 195 200 205

 Thr Pro Ile Phe Ala Ile Arg Lys Lys Asp Ser Ser Lys Trp Arg Lys
 210 215 220
 Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu
 225 230 235 240
 Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser
 245 250 255
 Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp
 260 265 270
 Lys Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Val Asn Asn

 275 280 285
 Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp
 290 295 300
 Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr Lys Ile Leu Glu
 305 310 315 320
 Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp
 325 330 335
 Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys

340 345 350
 Ile Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro
 355 360 365
 Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu
 370 375 380
 Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys
 385 390 395 400
 Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn
 405 410 415
 Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys

 420 425 430
 Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu
 435 440 445
 Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro
 450 455 460
 Val His Gly Ala Tyr Tyr Asp Pro Ser Lys Glu Leu Ile Ala Glu Ile
 465 470 475 480
 Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro
 485 490 495

 Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His
 500 505 510
 Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met
 515 520 525
 Glu Ser Ile Val Ile Trp Gly Lys Ile Pro Lys Phe Arg Leu Pro Ile
 530 535 540
 Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr
 545 550 555 560
 Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu

 565 570 575
 Trp Tyr Gln Leu Glu Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr
 580 585 590

Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Ile Gly Lys Ala Gly Tyr
595 600 605

Val Thr Asp Arg Gly Arg Lys Lys Ile Val Ser Leu Thr Asp Thr Thr
610 615 620

Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Ile Ala Leu Gln Asp Ser
625 630 635 640

Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile
645 650 655

Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile
660 665 670

Ile Glu Gln Leu Ile Lys Lys Glu Arg Val Tyr Leu Ser Trp Val Pro
675 680 685

Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser
690 695 700

Asn Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln
705 710 715 720

Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp
725 730 735

Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp
740 745 750

Gln Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser
755 760 765

Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Ile Ile
770 775 780

Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile
785 790 795 800

Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu Lys Leu Ala
805 810 815

Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly Ser Asn Phe
820 825 830

Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln
835 840 845

Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser

850

855

860

Met Asn Lys Glu Leu Lys Lys Leu Ile Gly Gln Val Arg Asp Gln Ala

865

870

875

880

Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe

885

890

895

Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val

900

905

910

Asp Ile Ile Ala Thr Asp Ile Gln Thr Arg Glu Leu Gln Lys Gln Ile

915

920

925

Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro

930

935

940

Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val

945

950

955

960

Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala

965

970

975

Lys Ile Ile Lys Asp Tyr Gly Lys Gln Met Ala Gly Ala Asp Cys Val

980

985

990

Ala Gly Arg Gln Asp Glu Asp

995

<210> 253

<211> 1003

<212> PRT

<213> Human immunodeficiency virus

<400> 253

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Arg Lys Ala Arg Glu Leu

1

5

10

15

Pro Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Arg

20

25

30

Val Trp Gly Gly Asp Asn Thr Leu Ser Glu Thr Gly Ala Glu Arg Gln

35

40

45

Gly Ala Val Ser Leu Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro

50

55

60

Leu Val Thr Val Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu Asp

65

70

75

80

Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Met Asn Leu Pro Gly Lys

85

90

95

Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg

100

105

110

Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Tyr Lys Ala Ile Gly

115

120

125

Thr Val Leu Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Leu

130

135

140

Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu

145

150

155

160

Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys

165

170

175

Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Glu Ile Cys

180

185

190

Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Arg Ile Gly Pro Glu Asn

195

200

205

Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys

210

215

220

Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp

225

230

235

240

Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys

245

250

255

Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val

260

265

270

Pro Leu Tyr Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser

275

280

285

Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro

290 295 300
 Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys
 305 310 315 320
 Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Glu Met Val Ile Tyr Gln
 325 330 335
 Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His

 340 345 350
 Arg Ile Lys Ile Glu Glu Leu Arg Gly His Leu Leu Lys Trp Gly Phe
 355 360 365
 Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met
 370 375 380
 Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu
 385 390 395 400
 Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
 405 410 415

 Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln
 420 425 430
 Leu Cys Lys Leu Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro
 435 440 445
 Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
 450 455 460
 Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
 465 470 475 480
 Ala Glu Ile Gln Lys Gln Gly Gln Asp Gln Trp Thr Tyr Gln Ile Tyr

 485 490 495
 Gln Glu Pro His Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Arg Arg
 500 505 510
 Thr Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys
 515 520 525
 Ile Ala Gln Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg
 530 535 540
 Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp

545 550 555 560
 Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu
 565 570 575
 Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu
 580 585 590
 Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys
 595 600 605
 Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Ile Ser Ile Thr
 610 615 620
 Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Asn Leu Ala Leu

 625 630 635 640
 Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
 645 650 655
 Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val
 660 665 670
 Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser
 675 680 685
 Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
 690 695 700

 Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
 705 710 715 720
 Lys Ala Gln Glu Glu His Glu Lys Tyr His Asn Asn Trp Arg Ala Met
 725 730 735
 Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala
 740 745 750
 Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Leu His Gly Gln Val
 755 760 765
 Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly

 770 775 780
 Lys Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala
 785 790 795 800

Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu
805 810 815
Lys Leu Ala Gly Arg Trp Pro Val Lys Val Val His Thr Asp Asn Gly
820 825 830
Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Gly
835 840 845

Ile Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
850 855 860
Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
865 870 875 880
Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
885 890 895
His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
900 905 910
Arg Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln

915 920 925
Lys Gln Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
930 935 940
Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
945 950 955 960
Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg
965 970 975
Arg Lys Val Lys Ile Ile Lys Asp Tyr Gly Lys Gln Met Ala Gly Ala
980 985 990

Asp Cys Val Ala Ser Arg Gln Asp Glu Asp Gln
995 1000

<210> 254

<211> 3000

<212> DNA

<213> Human immunodeficiency virus

<400> 254

ttcttcgcg agatcctggc cttccccag ggcgaggccc gcgagttccc ccccagcag

60

acccgcgcca actccccac ctcccgcgag ctgcaggtgc gcggcgacaa cccctcctcc	120
gaggccggca ccgagcgcca gggcaccctg aacttcccc agatcacct gtggcagcgc	180
cccctggtgt ccatcaaggt gggcggccag atcaaggagg ccctgctgga caccggcgcc	240
gacgacaccg tgctggagga ggtgaacctg cccggcaagt ggaagccaa gatgatcggc	300
ggcatcggcg gcttcatcaa ggtgcgccag tacgagcaga tcccatcga gatctcggc	360
aagaaggcca tcggcacctg gctggtgggc cccaccccc tgaacatcat cggccgcaac	420
atgtgaccc agctgggctg caccctgaac ttcccatct ccccatcga gaccgtgcc	480
gtgaagtga agcccgcat ggacggcccc aaggtgaagc agtggccct gaccgaggag	540
aagatcaagg cctgaccgc catctgcgac gagatggaga aggagggcaa gatcaccaag	600
atcgccccc acaacccta caacacccc atcttcgcca tccgaagaa ggactcctcc	660
aagtggcgca agctggtgga ctcccgcgag ctgaacaagc gcaccagga ctctgggag	720
gtgcagctgg gcatcccca ccccgccgc ctgaagaaga agaagtcctg gaccgtgctg	780
gacgtggcg acgcctactt ctccgtgcc ctggacaagg acttcgcaa gtacaccgc	840
ttcacatcc cctccgtgaa caacgagacc cccgcatcc gctaccagta caacgtgctg	900
ccccagggt ggaagggtc ccccgccatc ttccagtgt ccatgacaa gatcctggag	960
cccttcgca agcagaacct cgacatctg atctaccagt acatggacga cctgtacgtg	1020
ggctccgacc tggagatcg ccagcaccgc accaagatcg aggagctgcg cgagcactg	1080
ctgaagtggg gcttcaccac ccccgacaag aagcaccaga aggagcccc ctctctgtgg	1140
atgggtacg agctgcacc cgacaagtgg accgtgcagc ccatccagct gcccgagaag	1200
gactcctgga ccgtgaacga catccagaag ctggtgggca agctgaactg ggcctcccag	1260
atctacccc gcatcaaggt gcgccagtgc tgcaagtgc tgcgcggcgc caaggccctg	1320
accgacatcg tgccctgac cgaggaggcc gagctggagc tggccgagaa ccgcgagatc	1380
ctgaaggagc ccgtgcacgg cgcctactac gacccctca aggagctgat cgccgagatc	1440
cagaagcagg gccaggacca gtggacctac cagatctacc aggagccctt caagaacctg	1500
aagaccggca agtacgcaa gatgcgcacc gccacacca acgacgtgaa gcagctgacc	1560
gaggccgtgc agaagatgc catggagtcc atcgtgatct ggggcaagat cccaagttc	1620
cgctgccc tccagaagga gacctgggag acctggtgga ccgactactg gcaggccacc	1680
tggatcccc agtgggagtt cgtgaacacc cccccctgg tgaagctgtg gtaccagctg	1740
gagaaggacc ccatcgccgg cgtggagacc ttctactgg acggcgccgc caaccgcgag	1800
accaagatcg gcaaggccgg ctacgtgacc gaccgcgcc gcaagaagat cgtgtccctg	1860
accgacacca ccaaccagaa gaccgagctg caggccatct acatcgccct gcaggactcc	1920

ggctccgagg tgaacatcgt gaccgactcc cagtacgccc tgggcatcat ccaggcccag 1980

cccgacaagt ccgagtcgga gctggtgaac cagatcatcg agcagctgat caagaaggag 2040

cgcggtgacc tgtcctgggt gcccgcacac aagggcacatc gcggcaacga gcaggtggac 2100

aagctggtgt ccaacggcat ccgcaagggt ctgttccttg acggcatcga caaggcccag 2160

gaggagcacg agaagtacca ctccaactgg cgcgccatgg cctccgactt caacctgccc 2220

cccatcgtgg ccaaggagat cgtggcctcc tgcgaccagt gccagctgaa gggcgaggcc 2280

atgcacggcc aggtggactg ctccccggc atctggcagc tggactgcac ccacctggag 2340

ggcaagatca tcttggtggc cgtgcacgtg gcctccggt acatcgaggc cgaggtgatc 2400

cccgcgaga ccggccagga gaccgcctac ttcacatcga agctggccgg ccgctggccc 2460

gtgaaggatga tccacaccga caacggctcc aacttcacct ccgccgccgt gaaggccgcc 2520

tgctggtggg ccggcatcca gcaggagttc ggcatccctt acaaccccca gtcccagggc 2580

tggtggagt ccatgaacaa ggagctgaag aagctgatcg gccaggtgcg cgaccaggcc 2640

gagcacctga agaccgccgt gcagatggcc gtgttcaccc acaacttcaa gcgcaagggc 2700

ggcatcggcg gctactccgc cggcgagcgc atcgtggaca tcatcgccac cgacatccag 2760

acccgcgagc tgcagaagca gatcatcaag atccagaact tccgctgtga ctaccgcgac 2820

tcccgcgacc ccatctggaa gggccccgcc aagctgctgt ggaaggcgga gggcgccgtg 2880

gtgatccagg acaactccga catcaagggt gtgccccgcc gcaaggccaa gatcatcaag 2940

gactacggca agcagatggc cggcgccgac tgcgtggccg gccgccagga cgaggactaa 3000

3000

<210> 255

<211> 3009

<212> DNA

<213> Human immunodeficiency virus

<400> 255

ttcttccgag agaactggc ctccagcag cgcaaggccc gcgagctgcc ctccgagcag 60

acccgcgcca actccccac ctccgcgag ctgcgcgtgt ggggcggcga caacacctg 120

tccgagaccg gcgccgagcg ccaggggccc gtgtccctgt cttccccca gatcacctg 180

tggcagcgcc ccttggtgac cgtgaagatc ggcgccagc tgaaggaggc cctgctggac 240

accggcgccg acgacaccgt gctggaggag atgaacctgc ccgcaagtga gaagcccaag 300

atgatcggcg gcatcgggcg cttcatcaag gtgcgccagt acgaccagat cctgatcgag 360

atctgcggct acaaggccat cggcaccgtg ctggtgggcc ccacccccgt gaacatcatc 420

ggccgcaacc tgctgaccca gatcggctgc accctgaact tccccatctc ccccatcgag	480
accgtgcccc tgaagctgaa gcccggcatg gacggcccca aggtgaagca gtggccctg	540
accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggaag	600
atctcccgca tcggccccga gaaccctac aacaccccca tcttcgcat caagaagaag	660
gactccacca agtggcgcaa gctggtggac ttccgcgagc tgaacaagcg caccaggac	720
ttctgggagg tgcagctggg cateccccac cccgccggcc tgaagaagaa gaagtccgtg	780
accgtgctgg acgtggcgca cgcctacttc tccgtgcccc tgtacgagga ctccgcaag	840
tacaccgct tcaccatccc ctccatcaac aacgagaccc ccggcatccg ctaccgtac	900
aacgtgctgc cccagggtg gaagggtcc cccgccatct tccagtctc catgaccaag	960
atctggagc cttccgcaa gcagaacccc gagatggtga tctaccagta catggacgac	1020
ctgtactgg gtcggacct ggagatcggc cagcacgca tcaagatcga ggagctgcgc	1080
ggccacctgc tgaagtggg cttcaccacc cccgacaaga agcaccagaa ggagccccc	1140
ttctgtgga tgggctacga gctgcacccc gacaagtga ccgtgcagcc catccagctg	1200
cccgagaagg actcctggac cgtgaacgac atccagaagc tggtaggcaa gctgaactgg	1260
gcctcccaga tctaccccg catcaaggct cgccagctgt gcaagctgct gcgcggcgcc	1320
aaggccctga ccgacatcgt gcccctgacc gaggaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg acccctcaa ggacctgac	1440
gccgagatcc agaagcaggc ccaggaccag tggacctacc agatctacca ggagccccac	1500
aagaacctga agaccggcaa gtacccaag cgccgcaccg cccacaccaa cgacgtgaag	1560
cagctgaccg aggccgtgca gaagatgcc caggagtcca tcgtgatctg gggcaagacc	1620
cccaagtcc gcctgccat ccagaaggag acctgggaga cctggtggac cgactactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacaccc ccccctggt gaagctgtgg	1740
taccagctgg agaaggagcc catcgtgggc gccgagacct tctacgtgga cggcgccgcc	1800
aaccgcgaga ccaagctggg caaggccggc tacgtgaccg accgcggccg ccagaaggtg	1860
atctccatca ccgacaccac caaccagaag accgagctgc aggccatcaa cctggccctg	1920
caggactccg gtcggaggt gaacatcgtg accgactccc agtacgcct ggcatcatc	1980
caggcccagc ccgacaagtc cgagtccgag ctggtgaacc agatcatcga gcagctgac	2040
aagaaggaga aggtgtacct gtccgtgggtg cccgcccaca agggcatcgg cggcaacgag	2100
caggtggaca agctggtgtc ctccggcatc cgcaaggtgc tgttcctgga cggcatcgac	2160

aaggcccagg aggagcacga gaagtaccac aacaactggc gcgccatggc ctccgacttc 2220

aacctgcccc ccgtgggtggc caaggagatc gtggcctcct gcgacaagtg ccagctgaag 2280

ggcgaggccc tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc 2340

cacctggagg gcaaggtgat cctgggtggc gtgcacgtgg cctccggcta catcgaggcc 2400

gaggtgatcc ccgccgagac cggccaggag accgcctact tcctgctgaa gctggccggc 2460

cgctggcccc tgaaggtggt gcacaccgac aacggctcca acttcacctc cgccgccgtg 2520

aaggccgcct gctgggtggc cggcatcaag caggagtctg gcatccccta caacccccag 2580

tcccaggcgc tgggtggagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc 2640

gaccaggccg agcacctgaa gaccgccgtg cagatggccg tgttcatcca caacticaag 2700

cgcaagggcg gcatcgccgg ctactccgcc ggcgagcgca tcatcgacat catcgccacc 2760

gacatccaga ccaaggagct gcagaagcag atcatcaaga tccagaactt ccgcgtgtac 2820

taccgcgact cccgcgacct catctggaag ggccccgcca agctgctgtg gaaggcgag 2880

ggcgccgtgg tgatccagga caactccgac atcaaggtgg tgccccgcg caaggtgaag 2940

atcatcaagg actacggcaa gcagatggc ggcgccgact gcgtggcctc ccgccaggac 3000

gaggaccag 3009

<210> 256

<211> 1002

<212> PRT

<213> Human immunodeficiency virus

<400> 256

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Glu Phe

1 5 10 15

Ser Pro Glu Gln Ala Arg Ala Asn Ser Pro Thr Ser Arg Glu Leu Arg

20 25 30

Val Arg Gly Gly Asp Ser Pro Leu Pro Glu Thr Gly Ala Glu Gly Glu

35 40 45

Gly Ala Ile Ser Phe Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro

50 55 60

Leu Val Thr Ile Lys Val Ala Gly Gln Leu Lys Glu Ala Leu Leu Asp

65 70 75 80

Thr Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asp Leu Pro Gly Arg

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

Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His

340 345 350
Arg Glu Lys Val Glu Glu Leu Arg Lys His Leu Leu Lys Trp Gly Phe

355 360 365
Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met

370 375 380
Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu
385 390 395 400

Pro Asp Lys Glu Cys Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
405 410 415

Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys Gln
420 425 430

Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Asp Ile Val Pro
435 440 445

Leu Thr Ala Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu
450 455 460

Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile
465 470 475 480

Ala Glu Val Gln Lys Gln Gly Leu Asp Gln Trp Thr Tyr Gln Ile Tyr

485 490 495
Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Arg Arg

500 505 510
Thr Ala His Thr Asn Asp Val Arg Gln Leu Ala Glu Val Val Gln Lys

515 520 525
Ile Ser Met Glu Ser Ile Val Ile Trp Gly Lys Ile Pro Lys Phe Arg

530 535 540
Leu Pro Ile Gln Arg Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp

545 550 555 560

Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu
565 570 575

Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Ile Gly Ala Glu

580 585 590
 Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys
 595 600 605
 Ala Gly Tyr Val Thr Asp Lys Gly Arg Gln Lys Val Val Thr Leu Thr
 610 615 620
 Glu Thr Thr Asn Gln Lys Thr Glu Leu Glu Ala Ile His Leu Ala Leu

 625 630 635 640
 Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
 645 650 655
 Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val
 660 665 670
 Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser
 675 680 685
 Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
 690 695 700

 Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
 705 710 715 720
 Lys Ala Gln Glu Glu His Glu Arg Tyr His Ser Asn Trp Arg Ala Met
 725 730 735
 Ala Ser Asp Phe Asn Leu Pro Pro Ile Val Ala Lys Glu Ile Val Ala
 740 745 750
 Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
 755 760 765
 Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly

 770 775 780
 Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala
 785 790 795 800
 Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu
 805 810 815
 Lys Leu Ala Gly Arg Trp Pro Val Lys Val Ile His Thr Asp Asn Gly
 820 825 830
 Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Asn

835 840 845
 Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
 850 855 860
 Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
 865 870 875 880
 Glu Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
 885 890 895
 His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
 900 905 910
 Arg Ile Val Asp Ile Ile Ala Thr Asp Leu Gln Thr Lys Glu Leu Gln
 915 920 925
 Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser
 930 935 940
 Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
 945 950 955 960
 Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg
 965 970 975
 Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp
 980 985 990
 Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000
 <210> 257
 <211> 1003
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 257
 Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Lys Phe
 1 5 10 15
 Pro Ser Glu Gln Ala Arg Ala Asn Ser Pro Ala Ser Arg Glu Leu Trp
 20 25 30
 Val Arg Arg Gly Asp Asn Pro Leu Ser Glu Ala Gly Ala Glu Arg Arg

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

Ser Val Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
290 295 300

Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
305 310 315 320

Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
325 330 335

Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
340 345 350

His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
355 360 365

Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp
370 375 380

Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
385 390 395 400

Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
405 410 415

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Lys
420 425 430

Gln Leu Cys Arg Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile
435 440 445

Pro Leu Thr Lys Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
450 455 460

Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
465 470 475 480

Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
485 490 495

Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
500 505 510

Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
515 520 525

Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
530 535 540

Arg Leu Pro Ile Leu Lys Glu Thr Trp Asp Thr Trp Trp Thr Glu Tyr

545 550 555 560

Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro

565 570 575

Leu Val Lys Leu Trp Tyr Gln Leu Glu Thr Glu Pro Ile Ala Gly Ala

580 585 590

Glu Thr Phe Tyr Val Asp Gly Ala Ser Asn Arg Glu Thr Lys Lys Gly

595 600 605

Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Ala Val Ser Leu

610 615 620

Thr Glu Thr Thr Asn Gln Lys Ala Glu Leu His Ala Ile Gln Leu Ala

625 630 635 640

Leu Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr

645 650 655

Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu

660 665 670

Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu

675 680 685

Ser Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp

690 695 700

Lys Leu Val Ser Ala Gly Ile Arg Lys Ile Leu Phe Leu Asp Gly Ile

705 710 715 720

Asp Lys Ala Gln Glu Glu His Glu Lys Tyr His Asn Asn Trp Arg Ala

725 730 735

Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val

740 745 750

Ala Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln

755 760 765

Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu

770 775 780

Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Leu Glu

785 790 795 800
Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile
 805 810 815
Leu Lys Leu Ala Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn
 820 825 830
Gly Pro Asn Phe Ser Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala

 835 840 845
Gly Ile Gln Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly
 850 855 860
Val Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Arg Gln Val
865 870 875 880
Arg Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe
 885 890 895
Ile His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly
 900 905 910

Glu Arg Ile Ile Asp Ile Ile Ser Thr Asp Ile Gln Thr Arg Glu Leu
 915 920 925
Gln Lys Gln Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp
 930 935 940
Ser Arg Asp Pro Val Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly
945 950 955 960
Glu Gly Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro
 965 970 975
Arg Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly

 980 985 990
Asp Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000
<210> 258
<211> 3009
<212> DNA
<213> Human immunodeficiency virus
<400> 258

ttcttccgcg agaacctggc cttccagcag ggcgaggccc gcgagttctc ccccagcag	60
gcccgcgcca actccccac ctcccgcgag ctgcgcgtgc gcggcggcga ctccccctg	120
cccagaccg gcgccaggg cgaggcgcc atctccttca acttccccca gatcacctg	180
tggcagcgcc ccctggtgac catcaagggtg gccggccagc tgaaggaggc cctgctggac	240
accggcgccg acgacaccgt gctggaggag atcgacctgc ccggccgctg gaagcccaag	300
atgatcgcg gcacggcgg cttcatcaag gtgcgccagt acgaggagat catcatcgag	360
atcgagggca agaaggccat cggcacctg ctggtgggccc ccacccccgt gaacatcatc	420
ggccgcaaca tgctgaccca gatcggtgac accctgaact tccccatctc ccccatcgac	480
accgtgcccc tgaagctgaa gcccggcatg gacggcccca aggtgaagca gtggccctg	540
accgaggaga agatcaaggc cctgaccgag atctgcaccg agatggagaa ggagggcaag	600
atctccaaga tcggccccga gaaccctac aacacccccg tgttcgcat caagaagaag	660
gactccacca agtggcgcaa gctggtggac ttccgcgagc tgaacaagcg caccagcac	720
ttctgggagg tgcagctggg catccccac ccgcccggcc tgaagaagaa gaagtccgtg	780
accgtgctgg acgtggcgca cgcctacttc tccgtgcccc tggacgagtc cttccgcaag	840
tacaccgct tcaccatccc ctccatcaac aacgagaccc ccggcatccg ctaccagtac	900
aacgtgctgc ccagggtg gaagggtcc ccgcatct tccagtctc catgaccaag	960
atctggagc cttccgcac ccagaacccc gagatctga tctaccagta catggacgac	1020
ctgtacgtgg gctccgacct ggagatcggc cagcaccgagc agaaggtgga ggagctgcgc	1080
aagcacctgc tgaagtggg cttcaccacc ccgacaaga agcaccagaa ggagccccc	1140
ttctgtgga tgggctacga gctgcacccc gacaagtga ccgtgcagcc catccagctg	1200
cccgacaagg agtctggac cgtgaacgac atccagaagc tgggtgggcaa gctgaactgg	1260
gcctcccaga tctaccgccg catcaagggtg aagcagctgt gcaagctgct gcgcggcacc	1320
aaggccctga ccgacatcgt gcccctgacc gccgaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg accctccaa ggacctgac	1440
gccgaggtgc agaagcaggg cctggaccag tggacctacc agatctacca ggagcccttc	1500
aagaacctga agaccggcaa gtacccaag cgccgacccg cccacaccaa cgacgtgcgc	1560
cagctggcgg aggtggtgca gaagatctcc atggagtcca tcgtgatctg gggcaagatc	1620
ccaagtctc gcctgcccac ccagcgcgag acctgggaga cctggtggac cgactactgg	1680
caggccacct ggatccccga gtgggagttc gtgaacaccc cccccctggt gaagctgtgg	1740
taccagctgg agaaggagcc catcatcggc gccgagacct tctacgtgga cggcgccgcc	1800
aaccgcgaga ccaagctggg caaggccggc tacgtgaccg acaagggccg ccagaaggtg	1860

gtgacctga ccgagaccac caaccagaag accgagctgg aggccatcca cctggccctg	1920
caggactccg gcctggaggt gaacatcgtg accgactccc agtacgccct gggcatcatc	1980
caggcccagc ccgacaagtc cgagtccgag ctggtgtccc agatcatcga gcagctgac	2040
aagaaggaga aggtgtacct gtcctgggtg cccgccaca agggcatcgg cggcaacgag	2100
caggtggaca agctggtgtc ctccggcatc cgcaaggtgc tgttcctgga cggcatcgac	2160
aaggcccagg aggagcacga gcgctaccac tccaactggc gcgcatggc ctccgacttc	2220
aacctgcccc ccatcgtggc caaggagatc gtggcctcct gcgacaagt ccagctgaag	2280
ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc	2340
cacctggagg gcaagatcat cctggtggcc gtgcacgtgg cctccggcta catcgaggcc	2400
gaggtgatcc ccgccgagac cggccaggag accgcctact tcatcctgaa gctggccggc	2460
cgctggcccc tgaaggtgat ccacaccgac aacggctcca acttcacctc cgccgccgtg	2520
aaggccgctt gctggtgggc caacatccag caggagtctg gcatccccta caaccccag	2580
tcccagggcg tggaggagtc catgaacaag gagctgaaga agatcatcgg ccaggtgcgc	2640
gagcaggccg agcacctgaa gaccgccgtg cagatggccg tgttcatcca caacttaag	2700
cgcaaggcg gcatcgccgg ctactccgcc ggcgagcgca tcgtggacat catcgccacc	2760
gacctgcaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgcgtgtac	2820
taccgcgact ccgcgaccc catctggaag ggcgccgcca agctgctgtg gaaggcgag	2880
ggcgccgtgg tgatccagga caactccgac atcaaggtgg tgcccccgcc caaggccaag	2940
atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgg ccgccaggac	3000
gaggactaa	3009
<210> 259	
<211> 3012	
<212> DNA	
<213> Human immunodeficiency virus	
<400> 259	
ttcttcgcg agaacctggc cttccagcag ggcgaggccc gcaagttccc ctccgagcag	60
gccccgcca actccccgc ctcccgag ctgtgggtgc gccgcggcga caacccctg	120
tccgaggccg gcgccgagcg ccgcggcacc gtgccctccc tgtccttccc ccagatcacc	180
ctgtggcagc gccccctggt gaccatcaag gtggcgggcc agctgaagga ggcctgctg	240
gacacggcg ccgacgacac cgtgtggag gacatcaacc tgccccgcaa gtggaagccc	300
aagatgatcg gcggcatcgg cggcttcac aaggtgaagc agtacgaaa catcctgac	360

gagatctgcg gccacaaggc catcggcacc gtgctggtgg gccccacccc cgtgaacatc	420
atcggccgca acctgctgac ccagctgggc tgcacctga acttccccat ctccccatc	480
gagaccgtgc ccgtgaagct gaagcccggc atggacggcc ccaaggtgaa gcagtggccc	540
ctgaccgagg agaagatcaa ggccctgacc gagatctgca ccgagatgga gaaggagggc	600
aagatctcca agatcgcccc cgagaacccc tacaacaccc ccgtgttcgc catcaagaag	660
aaggactcca ccaagtggcg caagctgggtg gacttccgcg agctgaacaa gcgcacccag	720
gacttctggg aggtgcagct gggcatcccc caccgcccg gcctgaagaa gaagaagtcc	780
gtgaccgtgc tggacgtggg cgacgcctac ttctccgtgc ccctggacaa ggacttccgc	840
aagtacaccg cttcaccat cccctccgtg aacaacgaga cccccggcat ccgtaccag	900
tacaacgtgc tgccccaggg ctggaagggc tccccgcca tcttcagtc ctccatgacc	960
aagatcctgg agcccttccg caagcagaac ccgcacatcg tgatctacca gtacatggac	1020
gacctgtacg tgggctccga cctggagatc ggccagcacc gcaccaagat cgaggagctg	1080
cgccagcacc tgctgcgctg gggcttcacc acccccgaca agaagcacca gaaggagccc	1140
cccttctgt ggatgggcta cgagctgcac ccgcacaagt ggaccgtgca gcccatcgtg	1200
ctgcccgaga aggactcctg gaccgtgaac gacatccaga agctgggtggg caagctgaac	1260
tgggcctccc agatctaccc cggcatcaag gtgaagcagc tgtgccgcct gctgcgcggc	1320
accaaggccc tgaccgaggt gatccccctg accaaggagg ccgagctgga gctggccgag	1380
aaccgcgaga tctgaagga gcccgatgcac ggcgtgtact acgacccctc caaggacctg	1440
atcgccgaga tccagaagca gggccagggc cagtggacct accagatcta ccaggagccc	1500
ttcaagaacc tgaagaccgg caagtacgcc cgcattgcgc gcgcccacac caacgacgtg	1560
aagcagctga ccgaggccgt gcagaagatc accaccgagt ccatcgtgat ctggggcaag	1620
acccccaaagt tccgcctgcc catcctgaag gagacctggg acacctggtg gaccgagtac	1680
tggcaggcca cctggatccc cgagtgggag ttctgaaca cccccccct ggtgaagctg	1740
tggtaccagc tggagaccga gcccatcgcc ggccgcgaga cttctacgt ggacggcgcc	1800
tccaaccgcg agaccaagaa gggcaaggcc ggctacgtga ccgaccgagg ccgccagaag	1860
gccgtgtccc tgaccgagac caccaaccag aaggccgagc tgcacgccat ccagctggcc	1920
ctgcaggact ccggctccga ggtgaacatc gtgaccgact cccagtacgc cctgggcac	1980
atccaggccc agcccgacaa gtccgagtcc gagctggtga accagatcat cgagcagctg	2040
atcaagaagg agaaggtgta cctgtcctgg gtccccgcc acaaggcat cggcggcaac	2100

gagcaggtgg acaagctggt gtccgccggc atccgcaaga tcctgttctt ggacggcatc 2160

gacaaggccc aggaggagca cgagaagtac cacaacaact ggcgcccat ggcctccgac 2220

ttcaacctgc ccccggtggt ggccaaggag atcgtggcct cctgcgacaa gtgccagctg 2280

aaggcgagg ccatgcacgg ccaggtggac tgctccccg gcattctggca gctggactgc 2340

acccacctgg agggcaagat catcctggtg gccgtgcacg tggcctccgg ctacctggag 2400

gccgaggtga tccccgcga gaccggccag gagaccgctt acttcatcct gaagctggcc 2460

ggccgctggc ccgtgaagac catccacacc gacaacggcc ccaacttctc ctccgccgcc 2520

gtgaaggccg cctgctggtg ggccggcatc cagcaggagt tcggcatccc ctacaacccc 2580

cagtcccagg gcgtggtgga gtccatgaac aaggagctga agaagatcat ccgccaggtg 2640

cgcgaccagg ccgagcacct gaagaccgcc gtgcagatgg ccgtgttcat ccacaacttc 2700

aagcgcaagg gcggcatcgg cggctactcc gccggcgagc gcattcatga catcatctcc 2760

accgacatcc agaccgcga gctgcagaag cagatcatca agatccagaa cttccgcgtg 2820

tactaccgcg actccgcga ccccggtggt aagggtcccg ccaagctgct gtggaagggc 2880

gaggcgccg tggatgatca ggacaactcc gagatcaagg tggcgcccg ccgcaaggcc 2940

aagatcatcc gcgactacgg caagcagatg gccggcgacg actgcgtggc cggccgccag 3000

gacgaggact aa 3012

<210> 260

<211> 1002

<212> PRT

<213> Human immunodeficiency virus

<400> 260

Phe Phe Arg Glu Asn Leu Ala Phe Gln Gln Gly Glu Ala Arg Glu Phe

1 5 10 15

Ser Pro Glu Gln Ala Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Trp

20 25 30

Val Arg Arg Gly Asp Ser Pro Leu Pro Glu Ala Arg Ala Glu Gly Lys

35 40 45

Gly Asp Ile Pro Leu Ser Leu Pro Gln Ile Thr Leu Trp Gln Arg Pro

50 55 60

Leu Val Thr Val Arg Ile Gly Gly Gln Leu Ile Glu Ala Leu Leu Asp

65 70 75 80

Thr Gly Ala Asp Asp Thr Val Leu Glu Asp Ile Asn Leu Pro Gly Lys
85 90 95

Trp Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg
100 105 110

Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala Ile Gly
115 120 125

Thr Val Leu Val Gly Pro Thr Pro Ile Asn Ile Ile Gly Arg Asn Met
130 135 140

Leu Thr Gln Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu
145 150 155 160

Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys
165 170 175

Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Asp Ile Cys
180 185 190

Thr Glu Met Glu Arg Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn
195 200 205

Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys
210 215 220

Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp
225 230 235 240

Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ser Gly Leu Lys Lys
245 250 255

Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val
260 265 270

Pro Leu Asp Glu Ser Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser
275 280 285

Thr Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro
290 295 300

Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys
305 310 315 320

Ile Leu Glu Pro Phe Arg Ile Lys Asn Pro Glu Ile Val Ile Tyr Gln

				325				330				335			
Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	Gly	Gln	His
				340				345				350			
Arg	Ala	Lys	Ile	Glu	Glu	Leu	Arg	Lys	His	Leu	Leu	Ser	Trp	Gly	Phe
				355				360				365			
Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	Leu	Trp	Met
				370				375				380			
Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	Ile	Gln	Leu
385				390				395				400			
Pro	Asp	Lys	Glu	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	Leu	Val	Gly
				405				410				415			
Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro	Gly	Ile	Lys	Val	Lys	Gln
				420				425				430			
Leu	Cys	Lys	Leu	Leu	Arg	Gly	Ala	Lys	Ala	Leu	Thr	Asp	Ile	Val	Pro
				435				440				445			
Leu	Thr	Ala	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	Glu	Ile	Leu
450				455				460							
Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Glu	Pro	Ser	Lys	Glu	Leu	Ile
465				470				475				480			
Ala	Glu	Val	Gln	Lys	Gln	Gly	Leu	Asp	Gln	Trp	Thr	Tyr	Gln	Ile	Tyr
				485				490				495			
Gln	Glu	Pro	Tyr	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	Lys	Arg	Gly
				500				505				510			
Ser	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Val	Val	Gln	Lys
				515				520				525			
Ile	Ala	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	Lys	Phe	Lys
530				535				540							
Leu	Pro	Ile	Arg	Lys	Glu	Thr	Trp	Glu	Val	Trp	Trp	Thr	Glu	Tyr	Trp
545				550				555				560			
Gln	Ala	Thr	Trp	Ile	Pro	Asp	Trp	Glu	Phe	Val	Asn	Thr	Pro	Pro	Leu
				565				570				575			

Val Lys Leu Trp Tyr Arg Leu Glu Thr Glu Pro Ile Ala Gly Ala Glu
580 585 590

Thr Tyr Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys
595 600 605

Ala Gly Tyr Val Thr Asp Lys Gly Lys Gln Lys Ile Ile Thr Leu Thr
610 615 620

Glu Thr Thr Asn Gln Lys Ala Glu Leu Gln Ala Ile His Ile Ala Leu
625 630 635 640

Gln Asp Ser Gly Ser Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala
645 650 655

Leu Gly Ile Ile Gln Ala Gln Pro Asp Arg Ser Glu Ser Glu Val Val
660 665 670

Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser
675 680 685

Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys
690 695 700

Leu Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
705 710 715 720

Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met
725 730 735

Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala
740 745 750

Ser Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val
755 760 765

Asp Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly
770 775 780

Lys Ile Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala
785 790 795 800

Glu Val Ile Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Ile Leu
805 810 815

Lys Leu Ala Gly Arg Trp Pro Val Lys Ile Ile His Thr Asp Asn Gly
820 825 830

Ser Asn Phe Thr Ser Ala Ala Val Lys Ala Ala Cys Trp Trp Ala Asn

835 840 845
 Ile Thr Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val
 850 855 860
 Val Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg
 865 870 875 880
 Asp Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile
 885 890 895
 His Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu
 900 905 910

Arg Ile Ile Asp Ile Ile Ala Ser Asp Ile Gln Thr Lys Glu Leu Gln
 915 920 925
 Lys Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Phe Arg Asp Ser
 930 935 940
 Arg Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu
 945 950 955 960
 Gly Ala Val Val Ile Gln Asp Asn Asn Glu Ile Lys Val Val Pro Arg
 965 970 975
 Arg Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp

980 985 990
 Asp Cys Val Ala Gly Arg Gln Asp Glu Asp
 995 1000

<210> 261
 <211> 3009
 <212> DNA
 <213> Human immunodeficiency virus
 <400> 261

ttcttccgcg agaacctggc cttccagcag ggcgaggccc gcgagttctc ccccgagcag	60
gcccgcgcca actccccac ccgccgcgag ctgtgggtgc gccgcggcga ctccccctg	120
cccgaggccc gcgccgagg caagggcgac atccccctgt ccctgcccc gatcacctg	180
tggcagcgcc ccctggtgac cgtgcgcac gccggccagc tgatcgaggc cctgctggac	240

accggcgccg acgacaccgt gctggaggac atcaacctgc ccggcaagt gaagcccaag	300
atgatcgccg gcatcgccgg cttcatcaag gtgcgccagt acgaccagat cctgatcgag	360
atctcgccga agaaggccat cggcaccgtg ctggtgggcc ccacccccat caacatcatc	420
ggccgcaaca tgctgaccca gatcggtcgc accctgaact tccccatctc ccccatcgag	480
accgtgcccc tgaagctgaa gcccgccatg gacggcccca aggtgaagca gtggccctg	540
accgaggaga agatcaaggc cctgaccgac atctgcaccg agatggagcg cgagggaag	600
atctccaaga tcggccccga gaaccctac aacaccccca tcttcgcat caagaagaag	660
gactccacca agtggcgcaa gctgtggac ttcgcgagc tgaacaagcg caccaggac	720
ttctgggagg tgcagctggg catccccac cctccggcc tgaagaagaa gaagtccgtg	780
accgtgctgg acgtggcgca cgcctacttc tccgtgcccc tggacgagtc cttccgaag	840
tacaccgct tcaccatccc ctccaccaac aacgagaccc ccggcatccg ctaccgtac	900
aacgtgctgc cccagggtcg gaagggtcc cccgccatct tccagtcctc catgaccaag	960
atcctggagc cttccgcat caagaacccc gagatcgtga tctaccagta catggacgac	1020
ctgtacgtgg gctccgacct ggagatcggc cagcaccgcg ccaagatcga ggagctcgc	1080
aagcacctgc tgtcctgggg cttcaccacc cccgacaaga agcaccagaa ggagccccc	1140
ttctgtgga tgggctacga gctgcacccc gacaagtga ccgtgcagcc catccagctg	1200
cccgacaagg agtcctggac cgtgaacgac atccagaagc tgggtggcaa gctgaactgg	1260
gcctcccaga tctaccccg catcaagggt aagcagctgt gcaagctgct gcgcggcgcc	1320
aaggccctga ccgacatcgt gccctgacc gccgaggccg agctggagct ggccgagaac	1380
cgcgagatcc tgaaggagcc cgtgcacggc gtgtactacg agccctccaa ggagctgac	1440
gccgaggtgc agaagcaggg cctggaccag tggacctacc agatctacca ggagccctac	1500
aagaacctga agaccggcaa gtacccaag cgcggctccg cccacaccaa cgacgtgaag	1560
cagctgaccg aggtggtgca gaagatgcc accgagtcca tcgtgatctg gggcaagacc	1620
cccaagtcca agctgccat ccgcaaggag acctgggagg tgtgtggac cgagtactgg	1680
caggccacct ggatccccga ctgggagttc gtgaacaccc cccccctggt gaagctgtgg	1740
taccgcctgg agaccgagcc catcgccggc gccgagacct actacgtgga cggcgccgcc	1800
aaccgcgaga ccaagctggg caaggccggc tacgtgaccg acaagggcaa gcagaagatc	1860
atcacctga ccgagaccac caaccagaag gccgagctgc aggccatcca catcgccctg	1920
caggactccg gctccgaggt gaacatcgtg accgactccc agtacgcctt gggcatcatc	1980
caggcccagc ccgaccgtc cgagtcgag gtggtgaacc agatcatcga gcagctgac	2040
aagaaggaga aggtgtacct gtccgtgggtg cccgccaca agggcatcgg cggaacgag	2100

caggtggaca agctggtgtc ctccggcatc cgcaaggtgc tgttcctgga cggcatcgac 2160
aaggcccagg aggagcacga gaagtaccac tccaactggc gcgccatggc ctccgacttc 2220
aacctgcccc ccgtggtggc caaggagatc gtggcctcct gcgacaagt ccagctgaag 2280
ggcgaggcca tgcacggcca ggtggactgc tccccggca tctggcagct ggactgcacc 2340

cacctggagg gcaagatcat cctggtggcc gtgcacgtgg cctccggcta catcgaggcc 2400
gaggtgatec ccgccgagac cggccaggag accgcctact tcctcctgaa gctggccggc 2460
cgctggcccc tgaagatcat ccacaccgac aacggctcca acttcacctc cgccgccgtg 2520
aaggccgcct gctggtgggc caacatcacc caggagtctg gcaccccta caacccccag 2580
tcccaggcgc tgggtggagt catgaacaag gagctgaaga agatcatcgg ccaggtgcgc 2640
gaccaggccg agcacctgaa gaccgccgtg cagatggccg tgttcacca caacttcaag 2700
cgcaaggcgc gcacggcgc ctactccgcc ggcgagcgca tcacgacat catgcctcc 2760

gacatccaga ccaaggagct gcagaagcag atcaccaaga tccagaactt ccgcgtgtac 2820
ttccgcgact ccgcgaccc catctggaag ggccccgcca agctgctgtg gaaggcgag 2880
ggcgccgtgg tgatccagga caacaacgag atcaaggtgg tgccccgccg caaggccaag 2940
atcatccgcg actacggcaa gcagatggcc ggcgacgact gcgtggccgg ccgccaggac 3000
gaggactaa 3009

<210> 262
<211> 15
<212> PRT
<213> Human immunodeficiency virus
<400> 262

Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro

1 5 10 15
<210> 263
<211> 15
<212> PRT
<213> Human immunodeficiency virus
<400> 263

Lys Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
1 5 10 15
<210> 264
<211> 15

<212> PRT
 <213> Human immunodeficiency virus
 <400> 264
 Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala Ile Thr Gln
 1 5 10 15

<210>
 265

<211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 265
 Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
 1 5 10 15

<210> 266
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 266
 Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
 1 5 10 15

<210> 267
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 267

Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val
 1 5 10 15
 <210> 268
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 268

Glu Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met
 1 5 10 15

<210> 269

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 269

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Gly Val

1 5 10 15

<210> 270

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 270

Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala

1 5 10 15

<210> 271

<211> 16

<212> PRT

<213> Human immunodeficiency virus

<400> 271

Val Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu

1 5 10 15

<

210> 272

<211> 17

<212> PRT

<213> Human immunodeficiency virus

<400> 272

Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu

1 5 10 15

Gln

<210> 273

<211> 16

<212> PRT

<213> Human immunodeficiency virus

<400> 273

Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu

1 5 10 15

<210> 274

<211> 15

<212> PRT

<213>

> Human immunodeficiency virus

<400> 274

Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr Gln

1 5 10 15

<210> 275

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 275

Asp Thr Glu Val His Asn Val Trp Ala Thr Gln Ala Cys Val Pro

1 5 10 15

<210> 276

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 276

Ala Cys Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys

1 5 10 15

<210> 277

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 277

Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly

1 5 10 15

<210> 278

<211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 278
 Arg Lys Arg Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr
 1 5 10 15

<210>
 279

<211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 279
 His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Lys Asn Ile Ile
 1 5 10 15

<210> 280

<211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 280

Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln
 1 5 10 15
 <210> 281
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus

<400> 281

Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu
 1 5 10 15
 <210> 282
 <211> 20
 <212> PRT
 <213> Human immunodeficiency virus

<400> 282

Val Pro Val Trp Lys Glu Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp

1 5 10 15

Ala Lys Ser Tyr

20

<210> 283

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 283

Gly Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr

1 5 10 15

Asp Pro Asn Pro

20

<210> 284

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 284

Ser Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr

1 5 10 15

Ser Ala Ile Thr

20

<210> 285

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 285

Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln

1 5 10 15

Leu Leu Leu Asn

20

<210> 286

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 286

Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu

1 5 10 15

Gln Leu Thr Val

20

<210> 287

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 287

Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu

1 5 10 15

Gln Thr Arg Val

20

<210> 288

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 288

Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp Gly Thr Met

1 5 10 15

<210> 289

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 289

Asn Cys Gln His Leu Trp Arg Trp Gly Thr Met Ile Leu Gly Met

1 5 10 15

<210> 290

<211> 15

<212> PRT

<213> Human immunodeficiency virus

<400> 290
 Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro
 1 5 10 15
 <210> 291
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 291
 Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
 1 5 10 15
 <210> 292
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 292
 Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Thr Trp Met Phe
 1 5 10 15
 <210> 293
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 293
 Phe Asn Ser Thr Trp Met Phe Asn Gly Thr Tyr Met Phe Asn Gly
 1 5 10 15
 <210> 294
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 294
 Gly Ile Arg Arg Asn Tyr Gln His Trp Trp Gly Trp Gly Thr Met
 1 5 10 15
 <210> 295
 <211> 15

<212> PRT
 <213> Human immunodeficiency virus
 <400> 295
 Asn Tyr Gln His Trp Trp Gly Trp Gly Thr Met Leu Leu Gly Leu
 1 5 10 15
 <210> 296
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 296
 Asn Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile
 1 5 10 15
 <210>
 297
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 297
 Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly
 1 5 10 15
 <210> 298
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 298
 Asn Ile Ile Gly Thr Ile Arg Gln Ala His Cys Asn Ile Ser Arg
 1 5 10 15
 <210> 299
 <211> 15
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 299
 Thr Ile Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn
 1 5 10 15

<210> 300

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 300

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Met Leu

20

<210> 301

<211> 18

<212> PRT

<213> Human immunodeficiency virus

<400> 301

Arg Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Ala

1 5 10 15

Glu Asn

<210> 302

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 302

Gly Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr

1 5 10 15

Asp Pro Asn Pro

20

<210> 303

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 303

Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Lys Asp Lys Trp Asn

1 5 10 15

Glu Thr Leu Gln

20

<210> 304

<211> 20

<212> PRT

<213> Human immunodeficiency virus

<400> 304

Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Thr

1 5 10 15

Val Pro Trp Asn

20

<210> 305

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic Oligonucleotide

<400> 305

ttcagtcgac ggccacc 17

<210> 306

<211> 11

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic Oligonucleotide

<400> 306

gccaccatgg a 11

<210> 307

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic Oligonucleotide

<400> 307

taaagatctt acaa 14

<210> 308

<211

> 16

<212> PRT

<213> Human immunodeficiency virus

<400> 308

Lys Gln Ile Ile Asn Met Trp Gln Val Val Gly Lys Ala Met Tyr Ala

1 5 10 15

<210> 309

<211> 36

<212> PRT

<213> Human immunodeficiency virus

<400> 309

Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr

1 5 10 15

Thr Asn Val Asn Val Thr Asn Thr Thr Asn Asn Thr Glu Glu Lys Gly

20 25 30

Glu Ile Lys Asn

35

<210> 310

<211> 49

<212> PRT

<213> Human immunodeficiency virus

<400> 310

Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr

1 5 10 15

Thr Glu Ile Arg Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg

20 25 30

Leu Asp Val Val Pro Ile Asp Asp Asn Asn Asn Asn Ser Ser Asn Tyr

35 40 45

Arg

<210> 311

<211> 38
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 311

Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly
 20 25 30
 Gln Ala Phe Tyr Ala Thr

35

<210> 312
 <211> 44
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 312

Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Asn Thr Ser Gly Leu Phe Asn Ser Thr Trp Ile Gly Asn Gly Thr Lys
 20 25 30
 Asn Asn Asn Asn Thr Asn Asp Thr Ile Thr Leu Pro
 35 40

<210> 313
 <211> 36
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 313

Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Arg Asp Gly Gly Asn Asn Asn Thr Asn Glu Thr Glu Ile Phe Arg Pro
 20 25 30
 Gly Gly Gly Asp
 35

<210> 314

<211> 35
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 314
 Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Asn Val Arg Asn Val Ser Ser Asn Gly Thr Glu Thr Asp Asn Glu Glu
 20 25 30
 Ile Lys Asn

35
 <210> 315
 <211> 56
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 315
 Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Thr Glu Leu Arg Asp Lys Lys Gln Lys Val Tyr Ala Leu Phe Tyr Arg
 20 25 30
 Leu Asp Val Val Pro Ile Asp Asp Lys Asn Ser Ser Glu Ile Ser Gly
 35 40 45
 Lys Asn Ser Ser Glu Tyr Tyr Arg

50 55
 <210> 316
 <211> 38
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 316
 Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly
 20 25 30
 Gln Ala Phe Tyr Ala Thr

35
 <210> 317
 <211> 47
 <212> PRT
 <213> Human immunodeficiency virus

<400> 317
 Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr
 1 5 10 15
 Asn Thr Ser Gly Leu Phe Asn Ser Thr Trp Met Phe Asn Gly Thr Tyr
 20 25 30
 Met Phe Asn Gly Thr Lys Asp Asn Ser Glu Thr Ile Thr Leu Pro
 35 40 45

<210> 318
 <211> 37
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 318
 Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr

1 5 10 15
 Arg Asp Gly Gly Asn Asn Ser Asn Lys Asn Lys Thr Glu Thr Phe Arg
 20 25 30
 Pro Gly Gly Gly Asp
 35

<210> 319
 <211> 4
 <212> PRT
 <213> Human immunodeficiency virus
 <400> 319

Arg Glu Lys Arg

1
 <210> 320
 <211> 26
 <212> PRT

<213> Human immunodeficiency virus

<400> 320

Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg

1 5 10 15

Ile Val Phe Ala Val Leu Ser Ile Val Asn

20 25

<210> 321

<211> 7

<212> PRT

<213> Human immunodeficiency virus

<400> 321

Glu Leu Asp Lys Trp Ala Ser

1 5