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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 December 2004 (02.12.2004)

PCT

(10) International Publication Number
WO 2004/103312 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number:
PCT/US2004/015825

(22) International Filing Date: 19 May 2004 (19.05.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/472,082 19 May 2003 (19.05.2003) US
60/473,330 23 May 2003 (23.05.2003) US

(71) Applicant *(for all designated States except US)*: **PROGENICS PHARMACEUTICALS, INC.** [US/US]; 777 Old Saw Mill River Road, Tarrytown, NY 10591 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **OLSON, William, C.** [US/US]; 21 Fawn Court, Ossining, NY 10562 (US). **SCHUELKE, Norbert** [DE/US]; 101 Ridge Road, New City, NY 10956 (US). **O'HARA, Bryan** [US/US]; 537 Broadway, Norwood, NJ 07648 (US).

(74) Agent: **WHITE, John, P.**; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PEPTIDES USEFUL AS HIV FUSION INHIBITORS

(57) Abstract: This invention provides polypeptides represented by the formula α-β-γ, which polypeptides are useful for inhibiting fusion between HIV-1 and CD4+ cells. This invention also provides related pharmaceutical compositions, nucleic acids, host-vector systems, articles of manufacture, and related methods of production, prophylaxis and therapy.

WO 2004/103312 A2

PEPTIDES USEFUL AS HIV FUSION INHIBITORS

Throughout this application, various publications are referenced. The disclosures of these publications are hereby incorporated by reference into this application to describe more fully the art to which this invention pertains.

Background of the Invention

The human immunodeficiency virus (HIV) is the agent that causes Acquired Immunodeficiency Syndrome (AIDS), a lethal disease characterized by deterioration of the immune system. The initial phase of the HIV replicative cycle involves the attachment of the virus to susceptible host cells followed by fusion of viral and cellular membranes.

These events are mediated by the exterior viral envelope glycoproteins, which are first synthesized as a fusion-incompetent precursor envelope glycoprotein (env) known as gp160. The gp160 glycoprotein is endoproteolytically processed to the mature envelope glycoproteins gp120 and gp41, which are noncovalently associated with each other in a complex on the surface of the virus. The gp120 surface protein contains the high affinity binding site for human CD4, the primary receptor for HIV, as well as domains that interact with fusion coreceptors, such as the chemokine receptors CCR5 and CXCR4. The gp41 protein spans the viral membrane and contains at its amino-terminus a sequence of amino acids important for the fusion of viral and cellular membranes.

The native, fusion-competent form of the HIV-1 envelope

glycoprotein complex is a trimeric structure composed of three gp120 and three gp41 subunits. The receptor-binding (CD4 and co-receptor) sites are located in the gp120 moieties, and the fusion peptides in the gp41 components (Chan, 1997, Cell 89:263-273; Kwong, 1998, Nature 393:648-659; Kwong, 2000, J. Virol. 74:1961-1972; Poignard, 2001, Annu. Rev. Immunol. 19:253-274; Tan, 1997, Proc. Natl. Acad. Sci. U.S.A 94:12303-12308; Weissenhorn, 1997, Nature 387:426-430; and Wyatt, 1998, Nature 393:705-711).

In the generally accepted model of HIV-1 fusion, the sequential binding of gp120 to CD4 and a co-receptor induces a series of conformational changes in the gp41 subunits, leading to the insertion of the fusion peptides into the host cell membrane in a highly dynamic process (Doms, 2000, Virol. 276:229-237; Jones, 1998, J. Biol. Chem. 273:404-409; Melikyan, 2000, J. Cell Biol. 151:413-423; Sattentau, 1991, J. Exp. Med. 174:407-415; Sullivan, 1998, J. Virol. 72:4694-4703; Trkola, 1996, Nature 384:184-187; Wu, 1996, Nature 384:179-183; Wyatt, 1998, Science 280:1884-1888; and Zhang, 1999, Biochemistry 38:9405-9416). The associations between the six components of the fusion-competent complex are maintained via non-covalent interactions between gp120 and gp41, and between the gp41 subunits (Poignard, 2001, Annu. Rev. Immunol. 19:253-274; and Wyatt, 1998, Science 280:1884-1888). These interactions are relatively weak, making the fusion-competent complex unstable. This instability perhaps facilitates the conformational changes in the various components that are necessary for the fusion reaction to proceed efficiently, but it greatly complicates the task of isolating the native complex in purified form.

Summary of the Invention

This invention provides a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment		1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

This invention further provides a composition comprising the instant polypeptide and a pharmaceutically acceptable carrier.

This invention further provides a nucleic acid encoding a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment		1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

This invention further provides a host-vector system comprising a host cell transfected with an expression vector comprising a nucleic acid encoding a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment		1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

This invention further provides a method for producing a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero, which method comprises (a) culturing the instant host-vector system under conditions permitting the expression of the polypeptide, and (b) recovering the polypeptide so expressed.

This invention further provides a method for inhibiting fusion of an HIV-1 virus and a CD4+ cell, comprising contacting the instant polypeptide with the HIV-1 virus and the CD4+ cell under conditions which, in the absence of the polypeptide, would permit fusion of the HIV-1 virus and CD4+ cell, thereby inhibiting fusion of the HIV-1 virus and CD4+ cell.

This invention further provides a method for reducing the likelihood of a subject's becoming infected with HIV-1 comprising administering to the subject a prophylactically effective amount of the instant polypeptide, thereby reducing the likelihood of the subject's becoming infected with HIV-1.

This invention further provides a method for preventing or delaying the onset of, or slowing the rate of progression of, an HIV-1-related disease in an HIV-1-infected subject which comprises administering to the subject a therapeutically effective amount of the instant polypeptide, thereby preventing or delaying the onset of, or slowing the rate of progression of, the HIV-1-related disease in the subject.

Finally, this invention provides an article of manufacture comprising a packaging material having therein the instant polypeptide, and a label indicating a use of the polypeptide for (i) reducing the likelihood of a subject's being infected with HIV-1 and/or (ii) preventing or delaying the onset of, or slowing the rate of progression of, an HIV-related disease in an HIV-1-infected subject.

Brief Description of the Figures

Figure 1

Helical wheel representation of the packing of the C-terminal ("C") and N-terminal ("N") peptides of HIV gp41 in the post-fusion conformation.

Figure 2

Schematic of gp41 protein and alignment of related polypeptides.

Figures 3A-3E

C-48 efficacy against various T-20-resistant viral strains.

Detailed Description of the Invention

Definitions

Each of the following terms when used herein, shall have the meaning set forth below unless stated otherwise.

As used herein, "exposed" to HIV-1 means contact with HIV-1 such that infection could result.

As used herein, "host cells" shall include, but are not limited to, bacterial cells (including gram-positive cells), yeast cells, fungal cells, insect cells and animal cells. Suitable animal cells include, but are not limited to HeLa cells, COS cells, CV1 cells and various primary mammalian cells. Numerous mammalian cells can be used as hosts, including, but not limited to, the mouse fibroblast cell NIH-3T3 cells, CHO cells, HeLa cells, Ltk- cells and COS cells. Mammalian cells can be transfected by methods well known in the art, such as calcium phosphate precipitation, electroporation and microinjection. Methods and conditions for culturing transfected cells and for recovering the polypeptide so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed.

As used herein, "HIV" shall mean the human immunodeficiency virus. HIV shall include, without limitation, HIV-1. The human immunodeficiency virus (HIV) may be either of the two known types of HIV (HIV-1 or HIV-2). The HIV-1 virus may represent any of the known major subtypes (Classes A, B, C, D E, F, G and H) or outlying subtype (Group O). The human immunodeficiency virus includes but is not limited to the JR-FL strain. "HIV-1_{JR-FL}" is a strain that was originally

isolated from the brain tissue of an AIDS patient taken at autopsy and co-cultured with lectin-activated normal human PBMCs (O'Brien, 1990, Nature 348:69-73). HIV-1_{JR-FL} is known to utilize CCR5 as a fusion coreceptor and has the ability to replicate in phytohemagglutinin (PHA)-stimulated PBMCs and blood-derived macrophages but does not replicate efficiently in most immortalized T cell lines.

"HIV-1_{DH123}" is a clone of a virus originally isolated from the peripheral mononuclear cells (PBMCs) of a patient with AIDS (Shibata, 1995, J. Virol 69:4453-4462). HIV-1_{DH123} is known to utilize both CCR5 and CXCR4 as fusion coreceptors and has the ability to replicate in PHA-stimulated PBMCs, blood-derived macrophages and immortalized T cell lines.

"HIV-1_{Gun-1}" is a cloned virus originally isolated from the peripheral blood mononuclear cells of a hemophilia B patient with AIDS (Takeuchi, 1987, Jpn. J. Cancer Res. 78:11-15). HIV-1_{Gun-1} is known to utilize both CCR5 and CXCR4 as fusion coreceptors and has the ability to replicate in PHA-stimulated PBMCs, blood-derived macrophages and immortalized T cell lines.

"HIV-$1_{89.6}$" is a cloned virus originally isolated from a patient with AIDS (Collman, 1992, J. Virol. 66: 7517-21). HIV-$1_{89.6}$ is known to utilize both CCR5 and CXCR4 as fusion coreceptors and has the ability to replicate in PHA-stimulated PBMCs, blood-derived macrophages and immortalized T cell lines.

"HIV-1_{HXB2}" is a TCLA virus that is known to utilize CXCR4 as a fusion coreceptor and has the ability to replicate in PHA-stimulated PBMCs and immortalized T cell lines but not blood derived macrophages. It is well known to those skilled in

the art that other HIV-1 strains could be substituted in the place of the this or any other strains defined herein.

As used herein, the term "nucleic acid" shall mean any nucleic acid including, without limitation, DNA, RNA and hybrids thereof. The nucleic acid bases that form nucleic acid molecules can be the bases A, C, T, G and U, as well as derivatives thereof. Derivatives of these bases are well known in the art and are exemplified in PCR Systems, Reagents and Consumables (Perkin-Elmer Catalogue 1996-1997, Roche Molecular Systems, Inc, Branchburg, New Jersey, USA).

"Pharmaceutically acceptable carriers" are well known to those skilled in the art and include, but are not limited to, 0.01-0.1M and preferably 0.05M phosphate buffer, phosphate-buffered saline, or 0.9% saline. Additionally, such pharmaceutically acceptable carriers may include, but are not limited to, aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers such as those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

The term "polypeptide", used equivalently herein with "peptide", means a polymer of amino acid residues. The amino acid residues can be naturally occurring or derivatives thereof. Polypeptides can also include modifications such as glycosylation, lipid attachment, sulfation, hydroxylation, and ADP-ribosylation. Derivatives include, without limitation, acetylated amino acids (e.g., N-terminal acetylated amino acid residues) and amidated amino acids (e.g., C-terminal amidated amino acid residues). Such N- and C-terminal modification is also referred to in the art as N- and C-terminal "blocking." Methods for blocking the termini of recombinant peptides are well-known to those skilled in the art (Mitsuda et al., Prot. Exp. Purif., 25:448, 2002; Cottingham et al., Nat. Biotechnol., 19:974, 2001; Hong et al., Biochim. Biophys. Res. Commun., 267:362, 2000; McKee et al., Nat. Biotechnol. 16:647, 1998; Ray et al., Biotechnology, 11:64, 1993). Routine methods exist for synthesizing peptides with amidated carboxy termini and acetylated amino termini using standard F-moc or related synthetic chemistries. These methods are generally known in the art.

As used herein "prophylactically effective amount" means amount sufficient to reduce the likelihood of a disorder from occurring. In the preferred embodiment, a prophylactically effective amount is an amount sufficient to prevent a disorder from occurring.

As used herein, "reducing the likelihood of a subject's becoming infected with a virus" means reducing the likelihood of the subject's becoming infected with the virus by at least two-fold. For example, if a subject has a 1% chance of becoming infected with the virus, a two-fold reduction in the likelihood of the subject becoming infected

with the virus would result in the subject having a 0.5% chance of becoming infected with the virus. In the preferred embodiment of this invention, reducing the likelihood of the subject's becoming infected with the virus means reducing the likelihood of the subject's becoming infected with the virus by at least ten-fold. In one embodiment, the subject has been exposed to HIV-1 prior to the prophylaxis.

As used herein, "subject" means any animal or artificially modified animal. Artificially modified animals include, but are not limited to, SCID mice with human immune systems. Animals include, but are not limited to, mice, rats, dogs, guinea pigs, ferrets, rabbits, and primates. In the preferred embodiment, the subject is a human.

As used herein, "therapeutically effective amount" means an amount effective to slow, stop or reverse the progression of a disorder. In the preferred embodiment, a therapeutically effective amount is an amount sufficient to eliminate the disorder.

As used herein, a "variant" of a given amino acid sequence is an amino acid sequence that differs from the given amino acid sequence in a specified manner.

As used herein, "vector" shall mean any nucleic acid vector known in the art or employing elements known in the art. Such vectors include, but are not limited to, plasmid vectors, cosmid vectors and bacteriophage vectors. For example one class of vectors utilizes DNA elements which are derived from animal viruses such as animal papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTC or MoMLV), Semliki Forest virus or SV40 virus.

As used herein, "virally infected" means having viral genetic information introduced into, for example, a target cell, such as by fusion of the target cell membrane with the virus or infected cell. The target may be a cell of a subject. In the preferred embodiment, the target cell is a cell in a human subject.

As used herein, the following standard abbreviations are used throughout the specification to indicate specific amino acids: A=ala=alanine; R=arg=arginine; N=asn=asparagine; D=asp=aspartic acid; C=cys=cysteine; Q=gln=glutamine; E=glu=glutamic acid; G=gly=glycine; H=his=histidine; I=ile=isoleucine; L=leu=leucine; K=lys=lysine; M=met=methionine; F=phe=phenylalanine; P=pro=proline; S=ser=serine; T=thr=threonine; W=trp=tryptophan; Y=tyr=tyrosine; V=val=valine; B=asx=asparagine or aspartic acid; Z=glx=glutamine or glutamic acid.

Embodiments of the Invention

This invention provides a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position	1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

In one embodiment of the instant polypeptide, in β, a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 3 or 4. In another embodiment, in β, a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 3 or 4. In another embodiment, in β, a proline residue resides at position 2, 3, 5, 6, or 7 of segment 3 or 4.

In another embodiment, in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, or 5. In another embodiment, in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 2, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, 5, or 6. In another embodiment, in β, (i) a phenylalanine residue

resides at position 2, 3, 5, 6, or 7 of segment 3, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 4, 5 or 6.

In another embodiment, in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, or 5. In another embodiment, in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 2, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, 5 or 6. In another embodiment, in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 3, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 4, 5 or 6.

In another embodiment, in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, or 5. In another embodiment, in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 2, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, 5 or 6. In another embodiment, in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 3, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 4, 5 or 6.

In another embodiment, in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5. In another embodiment, in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

In another embodiment, in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5. In another embodiment, in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

In another embodiment, in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5. In another embodiment, in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

In the instant polypeptides, the amino acid sequence of the *β* moiety can differ from the sequence of Formula I in numerous ways. In one embodiment, the *β* moiety varies from the sequence of Formula I by point mutations only (e.g. 2, 3, 4, 5, 6, 7, or 8 point mutations). In another embodiment, the *β* moiety varies by deletion mutations (e.g. the deletion of an entire segment). In another embodiment, the *β* moiety varies by insertion mutations (e.g. the insertion of a non-HIV α-helical segment).

Also in the instant polypeptides, the *β* moiety should retain an α-helical conformation, and the polypeptide should posses the ability to inhibit HIV-1 fusion with a CD4+ cell at least 10% as well as does C48 peptide. Preferably, the instant polypeptide inhibits such fusion 90%, or at least 50% as well as does C48 peptide.

Further, in the instant polypeptides, the α and γ moieties can possess sequences which facilitate expression (e.g. bacterial expression) and secretion, increase half-life, and serve as attachment points for functional groups (e.g. half-life-increasing groups). The α and γ moieties can be any length, such as 1-1000 amino acid residues. In one embodiment, the length of such moieties is about 30 residues.

This invention further provides a composition comprising the instant polypeptide and a pharmaceutically acceptable carrier.

This invention further provides a nucleic acid encoding a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment		1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such

consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

The instant nucleic acid can be DNA or RNA. In one embodiment, the nucleic acid is DNA. In another embodiment, the nucleic acid is an expression vector. The expression vector can be, for example, a plasmid, a cosmid, a bacteriophage or a eukaryotic virus. In the preferred embodiment, the expression vector is a bacteriophage.

This invention further provides a host-vector system comprising a host cell transfected with an expression vector comprising a nucleic acid encoding a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position	1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

In the preferred embodiment of the instant host-vector system, the expression vector is a bacteriophage.

This invention further provides a method for producing a polypeptide represented by the formula α-β-γ, wherein α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero; β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein (i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or (ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero, which method comprises (a) culturing the instant host-vector system under conditions permitting the expression of the polypeptide, and (b) recovering the polypeptide so expressed.

This invention further provides a method for inhibiting fusion of an HIV-1 virus and a CD4+ cell, comprising contacting the instant polypeptide with the HIV-1 virus and the CD4+ cell under conditions which, in the absence of the polypeptide, would permit fusion of the HIV-1 virus and CD4+ cell, thereby inhibiting fusion of the HIV-1 virus and CD4+ cell.

This invention further provides a method for reducing the likelihood of a subject's becoming infected with HIV-1 comprising administering to the subject a prophylactically effective amount of the instant polypeptide, thereby reducing the likelihood of the subject's becoming infected with HIV-1.

This invention further provides a method for preventing or delaying the onset of, or slowing the rate of progression

of, an HIV-1-related disease in an HIV-1-infected subject which comprises administering to the subject a therapeutically effective amount of the instant polypeptide, thereby preventing or delaying the onset of, or slowing the rate of progression of, the HIV-1-related disease in the subject.

Finally, this invention provides an article of manufacture comprising a packaging material having therein the instant polypeptide, and a label indicating a use of the polypeptide for (i) reducing the likelihood of a subject's being infected with HIV-1 and/or (ii) preventing or delaying the onset of, or slowing the rate of progression of, an HIV-related disease in an HIV-1-infected subject.

Polypeptide Structure, Properties and Compositions

This invention further provides an oligomeric polypeptide comprising a plurality of any of the instant polypeptides (i.e., a plurality of polypeptides represented by the formula α-β-γ). In one embodiment of the oligomeric polypeptide and the polypeptides represented by the formula α-β-γ, either or both of the α and γ moieties of the instant polypeptides improve solubility (e.g., via PEG attachment or via amino acid composition).

In one embodiment of the instant polypeptide the polypeptide has deleted from its β moiety 1, 2, 3, 4, 5, 6, 7 or 8 amino acid residues. In one embodiment, the deleted amino acids are contiguous. In another embodiment, the deleted amino acid residues (e.g., 8 residues) are all at the N-terminus. In another embodiment, the deleted amino acid residues (e.g., 8 residues) are all at the C-terminus. In a further embodiment, some deleted amino acid residues (e.g.,

4 residues) are at the C-terminus, and the remaining deleted amino acid residues are at the N-terminus. In a further embodiment, one of segments 1, 2, 3, 4, 5, 6 or 7 is deleted.

The instant polypeptides may contain additional substitutions as desired. For example, additional substitutions can be made to increase their potency against wild-type and T-20-resistant virus. In one embodiment, the glutamic acid at position 4 in segment 4 is replaced by isoleucine. In another embodiment, the glutamic acid at position 4 in segment 4 is replaced by leucine. In another embodiment, the asparagine at position 1 in segment 5 is replaced by isoleucine. In yet another embodiment, the glutamic acid at position 4 in segment 4 is replaced by isoleucine and the asparagine at position 1 in segment 5 is replaced by isoleucine. In yet another embodiment, the glutamic acid at position 4 in segment 4 is replaced by leucine and the asparagine at position 1 in segment 5 is replaced by isoleucine. In another embodiment, the serine at position 1 in segment 4 is replaced by alanine. As another example, additional substitutions can improve the solubility and stability of the peptide in a pharmaceutical carrier. In a preferred embodiment, the instant polypeptide peptide is soluble at ≥150 mg/mL in isotonic, pH-neutral pharmaceutical carriers that are compatible with subcutaneous or intramuscular administration. In another preferred embodiment, the instant polypeptide has low local reactogenicity when administered by subcutaneous or intramuscular injection.

The local and systemic reactogenicity and tolerability of the peptides in animals can be determined by routine methods. By way of example, 1.5 mL of peptide solution or

control vehicle is injected into male New Zealand White rabbits (3 per group) by intravenous, subcutaneous and/or intramuscular routes. Clinical observations for pharmacotoxicological signs and/or mortality are recorded at 1, 4, 24, 48, and 72 hours post-dose and daily through Day 15. Injection site observations are scored at similar times according the Draize method. Body weights and temperatures are recorded at pre-test and on Day 1, Day 8 and Day 15 post-treatment. At necropsy, representative portions of all injection sites are evaluated microscopically. In addition, tissues (e.g., adrenals, aorta, brain, colon, duodenum esophagus, eyes, gall bladder, heart, ileum, jejunum, kidneys, liver, lungs, mammary gland, mesenteric lymph node, pancreas, pituitary gland, prostate, rectum, salivary gland, sciatic nerve, skin, spinal cord, spleen, sternum, stomach, testes, thigh musculature, thymus, thyroid, tongue, trachea, urinary bladder) are harvested and fixed with 10% neutral buffered formalin or other fixative for histopathological analysis. The clinical and histopathological observations are compared for peptide solution and control vehicle. The dose, dose interval, gender, species, group size, observation period and other parameters can be varied as appropriate.

Additional substitutions may be made to reduce the immunogenicity of the polypeptides humans or other animals. In a preferred embodiment, the polypeptide does not elicit high levels of antibodies that recognize the polypeptide. In another preferred embodiment, the polypeptide does not elicit high levels of antibodies that neutralize the activity of the polypeptide in blocking HIV fusion.

T cell epitopes are important in the generation of a mature antibody response. An administered peptide can be taken up

by a B cell or other antigen-presenting cell and presented on the surface of the antigen-presenting cell in association with major histocompatibility (MHC) molecules. The peptide-MHC complex can be recognized by the T cell receptor on a naïve T cell, thus activating the T cell. The activated T cell can produce cytokines that drive the B cell to proliferate and secrete antibody. Antibody generation thus can be reduced or prevented by reducing or preventing activation of peptide-specific T cells.

Recent developments have led to a clearer understanding of the association between peptides and MHC molecules. It is now clear that the peptides presented by MHC class I or class II molecules follow stringent rules. The allele-specific interaction usually involves a sequence of nine amino acids spanning the MHC groove. For class I molecules, the entire peptide ligand is involved in allele-specific interaction with MHC but for class II, the peptides are longer and the nine amino acid sequence is roughly central to the peptide. Allele-specific interactions are brought about by anchoring peptide side chains in complementary pockets in the MHC groove. The sum of allele-specific peptide-MHC interaction requirements can be described as a motif, characterized by number, spacing and specificities of anchors, as well as the more degenerate preferences at non-anchor positions within the nonamer stretches (Rammensee et al., Curr. Opin. Immunol. 7:85, 1995). These principles can be applied to introduce substitutions into the peptides in order to reduce or eliminate the presentation of the peptides or fragments thereof to T cells, and thereby reduce or eliminate the generation of antibodies to the peptides.

Additional substitutions may also be made to improve recombinant expression of the peptide in prokaryotic or

eukaryotic systems. Additional substitutions may also modulate the self-aggregation properties of the peptides. In another example, additional substitutions may modulate peptide binding to proteins, such as albumin or other serum protein, for the purposes of improving the pharmacology of the peptide. As another example, additional substitutions can increase the serum half-life of the peptide in animal serum in vitro and in vivo, such as by decreasing the peptides' sensitivities to proteolytic degradation. These additional substitutions can be made at positions 1-7 within any of the segments.

The mutations described herein can be incorporated into the corresponding peptides derived from the homologous region of gp41 in other strains of HIV. Examples of "core sequences" (i.e., β moiety sequences) are provided in Table 1 below.

It is understood that the sequence substitutions listed in Table 1 can be introduced individually (e.g., as a single substitution) and in combination (e.g., more than one substitution derived from one or more virus isolates) into the instant peptide.

Table 1.
Example core C48 sequences derived from HIV variants. Sequence substitutions relative to standard C48 are indicated in bold.

SEQUENCE	GENBANK ID
TWMEWDREINNYTSLI**Y**TLIEESQNQQEKNEQELLELDKWASLWNWF**D**	AAC40593
TWMEWDREINNYTSLIHSLIEESQNQQEKNE**P**ELLELDKWA**R**LWNW**V**N	AAK38842
TWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWA**N**LWNW**L**N	AAB59873
TWMEWDREVNNYTSLI**Y**TL**L**EESQNQQEKNEQELLELDKWASLWNWF**N**	AAG22514
TWMEWDREINNYTS**I**I**Y**SLIEESQNQQEKNEQELLELDKWASLWNWF**D**	AAA75116
TWMEW**E**REI**D**NYTG**L**I**Y**SLI**E**ASQVQQEKNEQELLELDKWASLWNWF**N**	AAK53104

TWMEWDREINNYTSLIYSLLEESQNQQEKNEQELLALDKWASLWNWFD	AAG22509
TWMEWEREIDNYTSLIYTLIEESQNQQEKNEQELLELDKWASLWNWFN	AAA44073
TWMEWDREINNYTNLIYNLIEQSQTQQEKNEQELLELDKWASLWNWFD	AAC02519
TWIEWDREINNYTSIIYSLIEESQNQQEKNEQELLELDKWASLWNWFD	AAA44191
TWMEWEREIDNYTNLIYSLIEDSQIQQEKNEKELLELDKWASLWNWFN	AAA44221
TWLEWEREIDNYTSLIYSLIEESQSQQEKNEQELLELDKWASLWNWFN	AAB04067
TWMEWEREIDNYTSLIYTLIEESQNQQEKNEQELLELDKWASLWNWFN	VCLJSC
TWMEWEKEIENYTGLIYSLIEESQTQQEKNEQELLQLDKWASLWNWFS	AAN73697
TWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDKWASLWNWFD	S13289
TWMEWEREIDNYTGLIFSLIEESQVQQERNEKELLELDTWASLWNWFD	AAN73598
TWMEWEREIRNYTDLIYTLIEKSQNQQEKNEQELLELDKWASLWNWFD	AAM09794
TWMQWEREIDNYTGLIYNLIEESQIQQEKNEKELLELDKWASLWNWFS	AAC32651
TWMEWEREIDNYTGLIYSLIEDSQIQQEKNEQELLKLDQWASLWNWFN	AAB06240
TWMEWEKEIENYTSLIYTLIEKSQNQQEINEQELLALDKWASLWNWFD	AAB38304
TWMEWEREINNYTGEIYRLIEKSQNQQEKNEQELLELDKWASLWNWFD	CAD10893
TWMEWEREINNYTNLIYNLIEESQNQQEKNEQDLLALDKWDSLWNWFS	P31819
TWMEWEREIDNYTGMIYTLIEESQIQQEKNEKELLELDKWASLWNWFD	AAF13332
TWMDWEREIDNYTALIYSLLEESQNQQEKNEQELLELDKWASLWNWFN	CAD59643
TWMEWEREINNYTGLIYNLIEESQNQQEKNEQELLELDKWASLWTWFD	AAC32300
TWIQWDREINNYTQLIYTLLEESQTQQEKNEQELLSLDKWASLWNWFD	CAC51036
TWMEWDREISNYTQVIYGLLEDSQKQQEKSEKDLLELDKWASLWNWFD	AAK65968
TWMQWDKEIHNYTQIIYDLIEESQNQQEINEKDLLALDKWASLWNWFD	AAK94217
TWIEWDREINNYTGIIYDLIEKAQNQQESNEQDLLALDQWASLWNWFD	CAC15054
TWMEWNREINNYTNLIYTLLEESQNQQEKNEQDLLALDKWSSLWNWFD	AAC55524
TWMKWEKEINNYTDLIYTLIEESQNQQEKNEQELLKLDKWASLWNWFD	AAC40596
TWLQWDKEINNYTQLIYSLIEKSQTQQEINEQDLLALDKWANLWNWFD	AAM66210

In the preferred embodiment, one or more replacements are made at non-'a' (a.k.a. non-'1') or non-'d' (a.k.a. non-'4') positions, although replacements at 'a' and 'd' positions are also allowed. At non-'a' and non-'d' positions, Phe is the most preferred amino acid to be introduced, followed by Cys and Pro. If several amino acids are replaced, they could all be replaced with the same amino acid or with combinations of them. At 'a' and 'd' positions, aspartate is the preferred amino acid to be introduced, followed by histidine, glycine,

cysteine and proline being the least preferred one.

In addition, this invention also allows additional peptide sequences (termed "α" and "γ") ranging from zero to a plurality of amino acids (up to, but not restricted to, 30) in length to be added to the N-terminus, C-terminus, or to both as a macromolecular carrier and/or to introduce additional properties. These properties include, but are not limited to, facilitating recombinant expression in bacteria, yeast, or baculovirus, enabling simplified purification via affinity tags (including but not restricted to, six or more consecutive histidine residues (His-tag), the amino acid sequence DYKDDDDK (FLAG-tag), the amino acid sequence WSHPQFEK (Strep-tag), or the amino acid sequence KQTAAAKFQRQHMDS (S-tag)), providing serum stability, modulating the oligomeric properties of the peptide, or providing linkers to attach macromolecules (such as PEG, carbohydrates, MAP, and serum proteins) that increase pharmacokinetic properties or allow generation of oligomeric peptides. "α" and "γ" can also contain combinations of two or more of the above mentioned properties within one single carrier sequence. "α" and "γ" may or may not be identical.

This invention further comprises a composition comprising any of the instant polypeptides and at least one anti-viral agent. Antiviral agents include, without limitation, a viral fusion inhibitor, a viral integrase inhibitor, a viral budding inhibitor, an HIV protease inhibitor, a nucleoside reverse transcriptase inhibitor, a non-nucleoside reverse transcriptase inhibitor, and combinations thereof.

This invention further comprises a composition comprising any of the instant polypeptides conjugated to a moiety that increases the pharmacokinetics of the polypeptide. In one

embodiment, the moiety is at least one polymer. Such polymers include, without limitation, hydrophilic polyvinyl polymers, polyalkylene ethers, polyoxyalkylenes, polymethacrylates, carbomers, branched polysaccharides, unbranched polysaccharides, polymers of sugar alcohols, heparin, and heparon. In one embodiment, the polyalkylene ether is polyethylene glycol (PEG) or a derivative thereof. In another embodiment of the instant composition comprising any of the instant polypeptides conjugated to a moiety that increases the pharmacokinetics of the polypeptide, the composition has at least one of an increase in serum half-life, an increase in mean residence time in the circulation or a decrease in serum clearance rate, compared to a non-conjugated polypeptide.

Recombinant Expression

The instant polypeptides can be made using recombinant DNA techniques. Here, the nucleotide sequences encoding the polypeptides of the invention may be synthesized, and/or cloned, and expressed according to techniques well known to those of ordinary skill in the art. See, for example, Sambrook, et al., 1989, Molecular Cloning, A Laboratory Manual, Vols. 1-3, Cold Spring Harbor Press, NY.

One may obtain the DNA segment encoding the peptide of interest using a variety of molecular biological techniques, generally known to those skilled in the art. For example, polymerase chain reaction (PCR) may be used to generate the DNA fragment encoding the protein of interest. Alternatively, the DNA fragment may be obtained from a commercial source.

The DNA encoding the polypeptides of interest may be

recombinantly engineered into a variety of host vector systems that also provide for replication of the DNA in large scale. These vectors can be designed to contain the necessary elements for directing the transcription and/or translation of the DNA sequence encoding the peptide.

Vectors that may be used include, but are not limited to, those derived from recombinant bacteriophage DNA, plasmid DNA or cosmid DNA. For example, plasmid vectors such as pcDNA3, pBR322, pUC 19/18, pUC 118, 119 and the M13 mp series of vectors may be used. Bacteriophage vectors may include λgt10, λgt11, λgt18-23, λZAP/R and the EMBL series of bacteriophage vectors. Cosmid vectors that may be utilized include, but are not limited to, pJB8, pCV 103, pCV 107, pCV 108, pTM, pMCS, pNNL, pHSG274, COS202, COS203, pWE15, pWE16 and the charomid 9 series of vectors.

Alternatively, recombinant virus vectors including, but not limited to, those derived from viruses such as herpes virus, retroviruses, vaccinia viruses, adenoviruses, adeno-associated viruses or bovine papilloma viruses plant viruses, such as tobacco mosaic virus and baculovirus may be engineered.

Additional vector systems for expression of recombinant proteins are set forth as follows. In one embodiment, recombinant peptides are expressed in E. coli. BL21 (DE3) pLys S using the T7 expression system as described in Shu et al., Biochemistry 2000 39:1674-1642. In another embodiment, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or MoMLV), Semliki Forest virus or SV40 virus. Additionally, cells which have stably integrated the DNA

into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may provide, for example, prototropy to an auxotrophic host, biocide resistance, (e.g., antibiotics) or resistance to heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcriptional promoters, enhancers, and termination signals. The cDNA expression vectors incorporating such elements include those described by (Okayama and Berg, Mol Cell Biol 3:280, 1983).

Once the expression vectors or DNA sequences containing the constructs have been prepared for expression, the expression vectors may be transfected or introduced into an appropriate mammalian host cell, for example. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate precipitation, electroporation, retroviral transduction, or other conventional techniques. In the case of protoplast fusion, the cells are grown in media and screened for the appropriate activity.

Mammalian cell lines for use in making host-vector systems include, for example, monkey kidney CV1 line transformed by SV40 (COS-7); human embryonic kidney line 293; baby hamster kidney cells (BHK); Chinese hamster ovary-cells-DHFR$^+$ (CHO); Chinese hamster ovary-cells DHFR$^-$ (DXB11); monkey kidney cells (CV1); African green monkey kidney cells (VERO-76); human cervical carcinoma cells (HeLa); canine kidney cells (MDCK); human lung cells (W138); human liver cells (Hep G2); mouse mammary tumor (MMT 060562); mouse cell line (C127);

and myeloma cell lines.

Other eukaryotic expression systems utilizing non-mammalian vector/cell line combinations can be used to produce the mutant envelope proteins. These include, but are not limited to, baculovirus vector/insect cell expression systems and yeast shuttle vector/yeast cell expression systems.

In order to express a biologically active polypeptide, the nucleotide sequence coding for the protein may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequences. Methods which are well known to those skilled in the art can be used to construct expression vectors having the hybrid polypeptide coding sequence operatively associated with appropriate transcriptional/translational control signals. These methods include in vitro recombinant DNA techniques and synthetic techniques. See, for example, the techniques described in Sambrook, et al., 1992, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y. and Ausubel et al., 1989, Current Protocols in Molecular Biology, Greene Publishing Associates & Wiley Interscience, N.Y.

The nucleic acid molecule encoding the polypeptides of interest may be operatively associated with a variety of different promoter/enhancer elements. The promoter/enhancer elements may be selected to optimize for the expression of therapeutic amounts of protein. The expression elements of these vectors may vary in their strength and specificities. Depending on the host/vector system utilized, any one of a number of suitable transcription and translation elements may be used. The promoter may be in the form of the promoter

which is naturally associated with the gene of interest. Alternatively, the DNA may be positioned under the control of a recombinant or heterologous promoter, i.e., a promoter that is not normally associated with that gene. For example, tissue specific promoter/enhancer elements may be used to regulate the expression of the transferred DNA in specific cell types.

Examples of transcriptional control regions that exhibit tissue specificity which have been described and could be used include, but are not limited to, elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, Cell 38:639-646; Ornitz et al., 1986, Cold Spring Harbor Symp. Quant. Biol. 50:399-409; MacDonald, 1987, Hepatology 7:42S-51S); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, Nature 315:115-122); immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., 1984, Cell 38:647-658; Adams et al., 1985, Nature 318:533-538; Alexander et al., 1987, Mol. Cell. Biol. 7:1436-1444): albumin gene control region which is active in liver (Pinkert et al., 1987, Genes and Devel. 1:268-276) alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, Mol. Cell. Biol. 5:1639-1648; Hammer et al., 1987, Science 235:53-58); alpha-1-antitrypsin gene control region which is active in liver (Kelsey et al., 1987, Genes and Devel. 1:161-171); beta-globin gene control region which is active in myeloid cells (Magram et al., 1985, Nature 315:338-340; Kollias et al., 1986, Cell 46:89-94); myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Shani, 1985, Nature 314:283-286); and gonadotropic releasing

hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378). Promoters isolated from the genomes of viruses that grow in mammalian cells, (e.g., vaccinia virus 7.5K, SV40, HSV, adenoviruses MLP, MMTV, LTR and CMV promoters) may be used, as well as promoters produced by recombinant DNA or synthetic techniques.

In some instances, the promoter elements may be constitutive or inducible promoters and can be used under the appropriate conditions to direct high level or regulated expression of the nucleotide sequence of interest. Expression of genes under the control of constitutive promoters does not require the presence of a specific substrate to induce gene expression and will occur under all conditions of cell growth. In contrast, expression of genes controlled by inducible promoters is responsive to the presence or absence of an inducing agent.

Specific initiation signals are also required for sufficient translation of inserted protein coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where the entire coding sequence, including the initiation codon and adjacent sequences, are inserted into the appropriate expression vectors, no additional translational control signals may be needed. However, in cases where only a portion of the coding sequence is inserted, exogenous translational control signals, including the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the protein coding sequences to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The

efficiency of expression may be enhanced by the inclusion of transcription attenuation sequences and enhancer elements.

Specific export sequences at the N-terminus are also required if efficient secretion of encoded proteins or peptides is desired. These signal sequences include but are not restricted to the secretion signal of the yeast α-factor (Tettelin et al., Nature, 387:81, 1997) or of invertase (Carlson et al., Mol. Cell. Biol., 3:439, 1983). These signal sequences also include the ATG initiation codon and adjacent sequences for proper cleavage by the signal peptidases. In cases where an export sequence is already present, including the initiation codon, no additional secretion motifs may be needed. However, in all other cases where the protein needs to be secreted, exogenous export signals, including the ATG initiation codon must be provided. Furthermore, the initiation codon and adjacent export sequences must be in phase with the reading frame of the protein coding sequences to ensure translation of the entire insert. These exogenous export signals and initiation codons can be of a variety of origins, both natural and synthetic.

Signal sequences from the following bacterial proteins have been used to export heterologous proteins in *Escherichia coli*: Major outer membrane protein A (ompA; Beck & Bremer, Nucl. Acids Res., 8:3011, 1980); Maltose binding protein (malE; Bassford et al., J. Bact., 139:19, 1979); and Alkaline phosphatase (phoA; Inouye et al, J. Bact., 149:434, 1982).

Dosages and modes of administration

The peptides of the invention may be administered using

techniques well known to those in the art. Preferably, agents are formulated and administered systemically. Techniques for formulation and administration may be found in "Remington's Pharmaceutical Sciences", latest edition, Mack Publishing Co., Easton, Pa. Suitable routes may include oral, rectal, vaginal, lung (e.g., by inhalation), transdermal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as, intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections, just to name a few. For intravenous injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer to name a few. In addition, infusion pumps may be used to deliver the peptides of the invention. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

In instances wherein intracellular administration of the peptides of the invention or other inhibitory agents is preferred, techniques well known to those of ordinary skill in the art may be utilized. For example, such agents may be encapsulated into liposomes, or microspheres then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are effectively delivered into the cell cytoplasm. Additionally, due to their hydrophobicity, when

small molecules are to be administered, direct intracellular administration may be achieved.

Nucleotide sequences encoding the peptides of the invention which are to be intracellularly administered may be expressed in cells of interest, using techniques well known to those of skill in the art. For example, expression vectors derived from viruses such as retroviruses, vaccinia viruses, adeno-associated viruses, herpes viruses, or bovine papilloma viruses, may be used for delivery and expression of such nucleotide sequences into the targeted cell population. Methods for the construction of such vectors and expression constructs are well known. See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor N.Y., and Ausubel et al., 1989, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, New York.

Effective dosages of the peptides of the invention to be administered may be determined through procedures well known to those in the art which address such parameters as biological half-life, bioavailability, and toxicity. In particularly preferred embodiments, an effective peptide dosage range is determined by one skilled in the art using data from routine in vitro and in vivo studies well know to those skilled in the art. For example, in vitro cell culture assays of antiviral activity, such as the exemplary assays described below, will provide data from which one skilled in the art may readily determine the mean inhibitory concentration (IC) of the polypeptide necessary to block some amount of viral infectivity (e.g., 50%, IC_{50}; or 90%, IC_{90}). Appropriate doses can then be selected by one skilled in the art using pharmacokinetic data from one or more

routine animal models, so that a minimum plasma concentration (C_{min}) of the peptide is obtained which is equal to or exceeds the determined IC value.

Exemplary polypeptide dosages may be as low as 0.1 µg/kg body weight and as high as 10 mg/kg body weight. More preferably an effective dosage range is from 0.1-100 µg/kg body weight. Other exemplary dosages for peptides of the invention include 1-5 mg, 1-10 mg, 1-30 mg, 1-50 mg, 1-75 mg, 1-100 mg, 1-125 mg, 1-150 mg, 1-200 mg, or 1-250 mg of peptide.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (e.g., the concentration of the test compound which achieves a half-maximal inhibition of the fusogenic event, such as a half-

maximal inhibition of viral infection relative to the amount of the event in the absence of the test compound) as determined in cell culture., Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography (HPLC) or any biological or immunological assay capable of measuring peptide levels.

The polypeptides of the invention can be administered in a single administration, intermittently, periodically, or continuously. For example, the polypeptides of the invention can be administered in a single administration, such as a single subcutaneous, a single intravenous infusion or a single ingestion. The polypeptides of the invention can also be administered in a plurality of intermittent administrations, including periodic administrations. For example, in certain embodiments the polypeptides of the invention can be administered once a week, twice per week, three times per week, four times per week, once a day, twice a day (e.g., every 12 hours), every six hours, every four hours, every two hours, or every hour. The polypeptides of the invention may also be administered continuously, such as by a continuous subcutaneous or intravenous infusion pump or by means of a subcutaneous or other implant which allows the polypeptides to be continuously absorbed by the patient. In a further embodiment, the polypeptide is self administered subcutaneously via an article of manufacture for that purpose, preferably less frequently than once per a day.

The peptides of the invention can also be administered in combination with at least one other therapeutic agent. Although not preferred for HIV therapy, administration for other types of therapy (e.g., cancer therapy) can be performed concomitantly or sequentially, including cycling

therapy (that is, administration of a first compound for a period of time, followed by administration of a second antiviral compound for a period of time and repeating this sequential administration in order to reduce the development of resistance to one of the therapies).

In the case of viral, e.g., retroviral, infections, an effective amount of a peptide or a pharmaceutically acceptable derivative thereof can be administered in combination with at least one, preferably at least two, other antiviral agents.

Such antiviral agents can include, but are not limited to other fusion inhibitors such as PRO-542 and PRO-140 (Progenics Pharmaceuticals) DP-107 (T21), DP-178 (T20), cytokines, e.g., rIFN-α, rIFN-β, rIFN-γ; inhibitors of reverse transcriptase, including nucleoside and non-nucleoside inhibitors, e.g., AZT, 3TC, D4T, ddI, adefovir, abacavir and other dideoxynucleosides or dideoxyfluoronucleosides, or delaviridine mesylate, nevirapine, efavirenz; inhibitors of viral mRNA capping, such as ribavirin; inhibitors of HIV protease, such as ritonavir, nelfinavir mesylate, amprenavir, saquinavir, saquinavir mesylate, indinavir or ABT378, ABT538 or MK639; inhibitors of viral budding, such as PA-457 (Panacos Pharmaceuticals); amphotericin B as a lipid-binding molecule with anti-HIV activity; and castanospermine as an inhibitor of glycoprotein processing.

For all such treatments described above, the exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g. Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p. 1).

It should be noted that the attending physician would know how to and when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administered dose in the management of the viral infection will vary with the severity of the condition to be treated and the route of administration. The dose and perhaps dose frequency, will also vary according to the age, body weight, and response of the individual patient. A program comparable to that discussed above may be used in veterinary medicine.

Use of pharmaceutically acceptable carriers to formulate the polypeptides herein disclosed for the practice of the invention into dosages suitable for systemic administration is within the scope of the invention. With proper choice of carrier and suitable manufacturing practice, the compositions of the present invention, in particular, those formulated as solutions, may be administered parenterally, such as by subcutaneous injection, intravenous injection, by subcutaneous infusion or intravenous infusion, for example by pump. The compounds can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers enable the compounds of the invention to be formulated as tablets, pills, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated.

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve

its intended purpose. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. The preparations formulated for oral administration may be in the form of tablets, dragees, capsules, or solutions. For oral administration of peptides, techniques such of those utilized by, e.g., Emisphere Technologies well known to those of skill in the art and can routinely be used.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, spray drying, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, emulsions and suspensions of the active compounds may be prepared as appropriate oily injection mixtures. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, liposomes or other substances known in the art for making lipid or lipophilic emulsions. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also

contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical preparations for oral use can be obtained by combining the active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, trehalose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethyl-cellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Pharmaceutical preparations which can be used orally also include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may

optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Methods for modulating serum half-life

The present invention also provides polypeptide-polymer conjugates having an effective size or molecular weight that confers an increase in serum half-life, an increase in mean residence time in circulation (MRT) and/or a decrease in serum clearance rate over underivatized peptides.

The peptide-polymer conjugates of the invention can be made by derivatizing the desired peptide with an inert polymer. It will be appreciated that any inert polymer which provides the conjugate with the desired apparent size or which has the selected actual molecular weight is suitable for use in constructing the peptide-polymer conjugates of the invention.

Many inert polymers are suitable for use in pharmaceuticals. See, e.g., Davis et al., Biomedical Polymers: Polymeric Materials and Pharmaceuticals for Biomedical Use, pp. 441-451 (1980). In this invention, a non-proteinaceous polymer is preferably used. The non-proteinaceous polymer ordinarily is a hydrophilic synthetic polymer, i.e., a polymer not otherwise found in nature. However, polymers which exist in nature and are produced by recombinant or in vitro methods are also useful, as are polymers which are isolated from native sources. Hydrophilic polyvinyl polymers fall within the scope of this invention, e.g., polyvinylalcohol and

polyvinvypyrrolidone. Particularly useful are polyalkylene ethers such as polyethylene glycol (PEG); polyoxyalklyenes such as polyoxyethylene, polyoxypropylene and block copolymers of polyoxyethylene and polyoxypropylene (Pluronics); polymethacrylates; carbomers; branched or unbranched polysaccharides which comprise the saccharide monomers D-mannose, D- and L-galactose, fucose, fructose, D-xylose, L-arabinose, D-glucuronic acid, sialic acid, D-galacturonic acid, D-mannuronic acid (e.g., polymannuronic acid, or alginic acid), D-glucosamine, D-galactosamine, D-glucose and neuraminic acid including homopolysaccharides and heteropolysaccharides such as lactose, amylopectin, starch, hydroxyethyl starch, amylose, dextran sulfate, dextran, dextrins, glycogen, or the polysaccharide subunit of acid mucopolysaccharides, e.g., hyaluronic acid, polymers of sugar alcohols such as polysorbitol and polymannitol, heparin or heparon. The polymer prior to cross-linking need not be, but preferably is, water soluble but the final conjugate must be water soluble. Preferably, the conjugate exhibits a water solubility of at least about 0.01 mg/ml and more preferably at least about 0.1 mg/ml, and still more preferably at least about 1 mg/ml. In addition the polymer should not be highly immunogenic in the conjugate form, nor should it possess viscosity that is incompatible with intravenous infusion or injection if the conjugate is intended to be administered by such routes.

In one embodiment, the polymer contains only a single group which is reactive. This helps to avoid cross-linking of protein molecules. However it is within the scope of the invention to maximize reaction conditions to reduce cross-linking, or to purify the reaction products through gel filtration or ion-exchange chromatography to recover substantially homogeneous derivatives. In other embodiments

the polymer contains two or more reactive groups for the purpose of linking multiple peptides to the polymer backbone. Again, gel filtration or ion-exchange chromatography can be used to recover the desired derivative in substantially homogeneous form.

The molecular weight of the polymer can range up to about 500,000 D and preferably is at least about 20,000 D, or at least about 30,000 D, or at least about 40,000 D. The molecular weight chosen can depend upon the effective size of the conjugate to be achieved, the nature (e.g., structure such as linear or branched) of the polymer and the degree of derivitization, i.e., the number of polymer molecules per peptide fragment, and the polymer attachment site or sites on the peptide fragment.

The polymer can be covalently linked to the peptide through a multifunctional crosslinking agent which reacts with the polymer and one or more amino acid residues of the peptide to be linked. However, it is also within the scope of the invention to directly crosslink the polymer by reacting a derivatized polymer with the peptide, or vice versa.

The covalent crosslinking site on the peptide includes the N-terminal amino group and epsilon amino groups found on lysine residues, as well other amino, imino, carboxyl, sulfhydryl, hydroxyl or other hydrophilic groups. The polymer may be covalently bonded directly to the peptide without the use of a multifunctional (ordinarily bifunctional) crosslinking agent, as described in U.S. Patent No. 6,458,355.

The degree of substitution with such a polymer will vary depending upon the number of reactive sites on the peptide,

the molecular weight, hydrophilicity and other characteristics of the polymer, and the particular peptide derivitization sites chosen. In general, the conjugate contains from 1 to about 10 polymer molecules, but greater numbers of polymer molecules attached to the peptide fragments of the invention are also contemplated. The desired amount of derivitization is easily achieved by using an experimental matrix in which the time, temperature and other reaction conditions are varied to change the degree of substitution, after which the level of polymer substitution of the conjugates is determined by size exclusion chromatography or other means known in the art.

Functionalized PEG polymers to modify the peptides of the invention are available from Shearwater Polymers, Inc. (Huntsville, Ala.). Such commercially available PEG derivatives include, but are not limited to, amino-PEG, PEG amino acid esters, PEG-hydrazide, PEG-thiol, PEG-succinate, carboxymethylated PEG, PEG-propionic acid, PEG amino acids, PEG succinimidyl succinate, PEG succinimidyl propionate, succinimidyl ester of carboxymethylated PEG, succinimidyl carbonate of PEG, succinimidyl esters of amino acid PEGs, PEG-oxycarbonylimidazole, PEG-nitrophenyl carbonate, PEG tresylate, PEG-glycidyl ether, PEG-aldehyde, PEG-vinylsulfone, PEG-maleimide, PEG-orthopyridyl-disulfide, heterofunctional PEGs, PEG vinyl derivatives, PEG silanes and PEG phospholides. The reaction conditions for coupling these PEG derivatives will vary depending on the protein, the desired degree of PEGylation and the PEG derivative utilized. Some factors involved in the choice of PEG derivatives include: the desired point of attachment (such as lysine or cysteine R-groups), hydrolytic stability and reactivity of the derivatives, stability, toxicity and antigenicity of the linkage, suitability for analysis, etc.

Specific instructions for the use of any particular derivative are available from the manufacturer. The conjugates of this invention are separated from the unreacted starting materials by gel filtration or ion exchange HPLC.

This invention will be better understood from the Experimental Details which follow. However, one skilled in the art will readily appreciate that the specific methods and results discussed are merely illustrative of the invention as described more fully in the claims which follow thereafter.

Experimental Details

Introduction

The invention encompasses novel peptides that are derived from the transmembrane envelope glycoprotein gp41 of human immunodeficiency virus. The peptides are mutated from the native gp41 sequence at defined locations according to specific design criteria.

The instant peptides are superior to T-20, T-1249 and C34 (sequences provided below), in that they permit (1) enhanced interactions with HIV gp41, (2) optimized α-helical content, (3) limited self-association, and (4) improved pharmacology in vivo.

T-20: YTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF

T-1249: WQEWEQKITALLEQAQIQQEKNEYELQKLDKWASLWEWF

C34: WMEWDREINNYTSLIHSLIEESQNQQEKNEQELL

The peptides are derived from a region of the gp41 ectodomain that exhibits a 4-3 heptad repeat of hydrophobic amino acids. Such peptides have the potential to form α-helical coiled-coil structures. During HIV fusion and infection, this heptad repeat region is thought to pack in antiparallel fashion against a second, N-terminal heptad repeat region that forms a central, three-stranded coiled-coil. Synthetic peptides such as those described above are believed to block HIV fusion by binding to the three-stranded coiled-coil and preventing, in a dominant negative fashion, binding of the corresponding sequences from viral gp41 (reviewed in Chan and Kim, Cell, 93:881, 1998). The

resulting structure is illustrated schematically in Figure 1 using the helical wheel representation of α-helical peptide sequences.

In the α-helical representation, the letters a-g (a.k.a. positions 1-7) illustrate the positions of sequential amino acids. This structure repeats itself such that every eighth amino acid is aligned in the peptide coil. Notably, hydrophobic amino acids in the "a" and "d" (a.k.a. "1" and "4" respectively) positions mediate self-oligomerization of α-helical coils to form coiled-coils, and this juxtaposition of amino acids is illustrated above for the N-terminal heptad repeat peptide.

Structural studies reveal that the "a" and "d" positions (highlighted in bold for T-20, T-1249 and C34 in the above diagram) of the C-terminal peptides pack against the "e" and "g" positions of the N-terminal peptides, as illustrated above. Unlike typical "a" and "d" residues of α-helical peptides, the "e" and "g" positions of the N-terminal heptad repeat region are enriched in polar or weakly hydrophobic amino acids such as arginine, glycine and alanine. This atypical packing arrangement provides a remarkable opportunity to optimize peptide potency by optimizing C-terminal interactions with N-terminal sequences (e.g., via the atypical a-e and d-g interactions and secondarily by "g-c" and "e-b" interactions) while regulating oligomerization of the C-terminal peptides (via the typical a-d interactions). This is important since monomeric C-terminal peptides are highly active, and oligomerization of C-terminal peptides may blunt their antiviral activity in a dominant-negative fashion.

Improved therapies for treating HIV infection are needed.

The invention provides HIV-inhibitory compounds of improved potency with the potential for less rapid development of drug-resistant viruses and for less frequent dosing. In addition, production of the peptides in recombinant organisms (e.g., bacteria such as E. coli) results in reduced cost-of-goods over current-generation gp41-based peptide therapies, which are produced synthetically by standard solid-phase or solution-phase chemistry.

The gp41 sequences are further modified at the "a", "d" and/or adjacent positions to further improve their antiviral properties. The peptides may be further modified to enhance their pharmacology by the methods described herein. In addition, the peptides are produced by cost-efficient recombinant methods for commercial purposes.

Derivatives of a gp41 peptide (e.g., C34 by way of example) are produced by standard recombinant or synthetic means. The peptides contain point mutations in one or more of the "a", "d" and adjacent residues. High-throughput methods for creating modified peptides are well-known in the field. For example, it is simple to chemically synthesize peptides containing diverse point mutations (e.g., the other 19 amino acids that occur naturally in proteins as well as unnatural amino acids) at a given position by performing parallel syntheses and by adding a different amino acid building block to separate reaction vessels at the desired stage(s) of synthesis. The reaction and purification is then continued as for the unmodified peptide.

Mutated and non-mutated peptides are tested for the ability to block HIV membrane fusion in a robust, rapid, homogeneous, high-throughput assay (U.S. Patent No. 6,261,763 and Litwin et al., J. Virol., 70:6437, 1996).

Mutations that affect antiviral activity are documented in order to develop structure-activity relationships (SAR). All or a subset of the peptides are analyzed for α-helical content by standard methods such as circular dichroism and for oligomeric state by methods such as ultracentrifugation in order to further develop SAR. Mutations that enhance antiviral activity may be combined in multiply mutated peptides in order to further optimize the antiviral potency of the peptides. Optimized peptides may be tested for potency against HIV in standard *in vitro* or *in vivo* assays. *In vitro* assays include assays that measure the ability of wild-type HIV to replicate in primary human peripheral blood mononuclear cells (Trkola et al., J. Virol., 72:1876, 1998). *In vivo* assays include the hu-PBL-SCID mouse model (Poignard et al., Immunity, 10:431, 1999) the SHIV-macaque model (Baba et al., Nature Medicine, 6:200, 2000). The *in vitro* assays can examine a range of phenotypically and/or genotypically diverse viruses in order to establish the breadth of antiviral activity of a given peptide. In general, one will determine the pharmacokinetics of the peptides in the desired species prior to performing *in vivo* studies, in order to establish appropriate dosing routes and schedules.

Peptides that have been optimized for antiviral activity may be further optimized for pharmacokinetic properties. These enhancements include attachment of polymers such as polyethylene glycol (Witt et al., J. Pharmacol. Exp. Ther. 298:848, 2001), by further modifying the peptide to prevent degradation/proteolysis in vivo (Brinckerhoff et al., Int. J. Cancer, 83:326, 1999) or by genetically or chemically fusing the peptide to carriers (e.g., immunoglobulin sequences or serum albumin) with longer pharmacokinetic half-lives. These modified peptides and peptide fusion proteins are tested for antiviral activity and

pharmacokinetic properties as described above.

The following in this paragraph are preferred embodiments: (1) the instant peptides do not contain a proline residue; (2) the peptides contain one or more alanine residues in positions "a" and "d" of segments "3", "4", or "5; (3) no more than one alanine is present in any given segment; (4) the peptides contain one or more phenylalanine residues in positions, "b", "c", "e", "f" and "g" of one or more of the heptad repeats; (5) no more than one phenylalanine is present in any given heptad repeat; (6) one phenylalanine residue is present in the second heptad repeat; (7) one phenylalanine residue is present in the third heptad repeat; (8) one phenylalanine residue is present in the third heptad repeat; (9) one phenylalanine residue is present in the fourth heptad repeat; (10) one phenylalanine residue is present in each of two different heptad repeats; (11) the peptide contains one or more cysteine residues in positions, "b", "c", "e", "f" and "g"; (12) a glycine residue is present in one or more of the "a" or "d" positions; (13) a histidine residue is present in one or more of the "a" or "d" positions; (14) an aspartic acid residue is present in one or more of the "a" or "d" positions; and (15) a cysteine residue is present in one or more of the "a" or "d" positions.

The instant peptides inhibit HIV-1 viruses that are resistant to T-20. Compared to wild-type viruses, resistant viruses show at least 3-fold reduced susceptibility to T-20 in vitro. Mutations that confer T-20 resistance include the following mutations in gp41: G36D, V38M and V38A (Rimsky et al., J. Virol., 72:986, 1998). The T-20-resistant viruses show increased, unchanged or < 3-fold reduced susceptibility to the modified peptides of this invention.

As described above, in a preferred embodiment, the peptides are produced in recombinant organisms, such as the E. coli bacterium, for large-scale commercial use in order to minimize production costs.

Example I

A. Single substitutions at non-'a' or non-'d' positions

```
                a  d  a  d  a  d  a  d  a  d  a  d  a  d
Seq  1   X-TWMEWDREINNYTSLIHSLIEESFNQQEKNEQELLELDKWASLWNWFN-Z
Seq  2   X-TWMEWDREINNYTSLIHSLIEESQFQQEKNEQELLELDKWASLWNWFN-Z
Seq  3   X-TWMEWDREINNYTSLIHSLIEESQNQFEKNEQELLELDKWASLWNWFN-Z
Seq  4   X-TWMEWDREINNYTSLIHSLIEESQNQQFKNEQELLELDKWASLWNWFN-Z
Seq  5   X-TWMEWDREINNYTSLIHSLIEESQNQQEFNEQELLELDKWASLWNWFN-Z
                a  d  a  d  a  d  a  d  a  d  a  d  a  d
Seq  6   X-TWMEWDREINNYTSLIHSLIEESCNQQEKNEQELLELDKWASLWNWFN-Z
Seq  7   X-TWMEWDREINNYTSLIHSLIEESQCQQEKNEQELLELDKWASLWNWFN-Z
Seq  8   X-TWMEWDREINNYTSLIHSLIEESQNQCEKNEQELLELDKWASLWNWFN-Z
Seq  9   X-TWMEWDREINNYTSLIHSLIEESQNQQCKNEQELLELDKWASLWNWFN-Z
Seq 10   X-TWMEWDREINNYTSLIHSLIEESQNQQECNEQELLELDKWASLWNWFN-Z
                a  d  a  d  a  d  a  d  a  d  a  d  a  d
Seq 11   X-TWMEWDREINNYTSLIHSLIEESPNQQEKNEQELLELDKWASLWNWFN-Z
Seq 12   X-TWMEWDREINNYTSLIHSLIEESQPQQEKNEQELLELDKWASLWNWFN-Z
Seq 13   X-TWMEWDREINNYTSLIHSLIEESQNQPEKNEQELLELDKWASLWNWFN-Z
Seq 14   X-TWMEWDREINNYTSLIHSLIEESQNQQPKNEQELLELDKWASLWNWFN-Z
Seq 15   X-TWMEWDREINNYTSLIHSLIEESQNQQEPNEQELLELDKWASLWNWFN-Z
```

B. Multiple substitutions at non-'a' or non-'d' positions

Phe substitutions are preferred and are presented as sequence examples below. Other possibilities include substitutions at similar positions using Pro and/or Cys as amino acids of choice. The list contains only selected examples and is by no means complete.

```
                a  d  a  d  a  d  a  d  a  d  a  d  a  d
Seq 16   X-TWMEWDREINNYTSLIFSLIEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 17   X-TWMEWDREINNYTSLIHFLIEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 18   X-TWMEWDREINNYTSLIHSLFEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 19   X-TWMEWDREINNYTSLIHSLIFESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 20   X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNFQELLELDKWASLWNWFN-Z
```

Seq 21 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEFELLELDKWASLWNWFN-Z
Seq 22 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEFELLELDKWASLWNWFN-Z
Seq 23 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEFELLELDKWASLWNWFN-Z
Seq 24 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEFELLELDKWASLWNWFN-Z
Seq 25 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEFELLELDKWASLWNWFN-Z
a d a d a d a d a d a d a d
Seq 26 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 27 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 28 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 29 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 30 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQEFLELDKWASLWNWFN-Z
a d a d a d a d a d a d a d
Seq 31 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELFELDKWASLWNWFN-Z
Seq 32 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELFELDKWASLWNWFN-Z
Seq 33 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELFELDKWASLWNWFN-Z
Seq 34 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELFELDKWASLWNWFN-Z
Seq 35 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELFELDKWASLWNWFN-Z
Seq 36 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLFLDKWASLWNWFN-Z
Seq 37 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLFLDKWASLWNWFN-Z
Seq 38 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLFLDKWASLWNWFN-Z
Seq 39 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLFLDKWASLWNWFN-Z
Seq 40 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLFLDKWASLWNWFN-Z
a d a d a d a d a d a d a d
Seq 41 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 42 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 43 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 44 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 45 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 46 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELDFWASLWNWFN-Z
Seq 47 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELDFWASLWNWFN-Z
Seq 48 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELDFWASLWNWFN-Z
Seq 49 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELDFWASLWNWFN-Z
Seq 50 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELDFWASLWNWFN-Z
a d a d a d a d a d a d a d
Seq 51 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELDKWFSLWNWFN-Z
Seq 52 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELDKWFSLWNWFN-Z
Seq 53 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELDKWFSLWNWFN-Z
Seq 54 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELDKWFSLWNWFN-Z
Seq 55 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELDKWFSLWNWFN-Z
Seq 56 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELDKWAFLWNWFN-Z
Seq 57 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELDKWAFLWNWFN-Z
Seq 58 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELDKWAFLWNWFN-Z
Seq 59 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELDKWAFLWNWFN-Z
Seq 60 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELDKWAFLWNWFN-Z
a d a d a d a d a d a d a d
Seq 61 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELDKWASFWNWFN-Z
Seq 62 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELDKWASFWNWFN-Z
Seq 63 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELDKWASFWNWFN-Z
Seq 64 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELDKWASFWNWFN-Z
Seq 65 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELDKWASFWNWFN-Z
Seq 66 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELDKWASLWFWFN-Z

Seq 67 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELDKWASLWFWFN-Z
Seq 68 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELDKWASLWFWFN-Z
Seq 69 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELDKWASLWFWFN-Z
Seq 70 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELDKWASLWFWFN-Z

a d a d a d a d a d a d a d

Seq 71 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQELLELDKWASLWNFFN-Z
Seq 72 X-TWMEWDREINNYTSLIHFLIEESQNQQEKNEQELLELDKWASLWNFFN-Z
Seq 73 X-TWMEWDREINNYTSLIHSLFEESQNQQEKNEQELLELDKWASLWNFFN-Z
Seq 74 X-TWMEWDREINNYTSLIHSLIFESQNQQEKNEQELLELDKWASLWNFFN-Z
Seq 75 X-TWMEWDREINNYTSLIHSLIEFSQNQQEKNEQELLELDKWASLWNFFN-Z
Seq 76 X-TWMEWDREIFNYTSLIHSLIEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 77 X-TWMEWDREINFYTSLIHSLIEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 78 X-TWMEWDREINNYFSLIHSLIEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 79 X-TWMEWDREINNYTFLIHSLIEESQNQQEKNEQELLELFKWASLWNWFN-Z
Seq 80 X-TWMEWDREINNYTSFIHSLIEESQNQQEKNEQELLELFKWASLWNWFN-Z

a d a d a d a d a d a d a d

Seq 81 X-TWFEWDREINNYTSLIHSLIEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 82 X-TWMFWDREINNYTSLIHSLIEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 83 X-TWMEWFREINNYTSLIHSLIEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 84 X-TWMEWDFEINNYTSLIHSLIEESQNQQEKNFQELLELDKWASLWNWFN-Z
Seq 85 X-TWMEWDRFINNYTSLIHSLIEESQNQQEKNFQELLELDKWASLWNWFN-Z

a d a d a d a d a d a d a d

Seq 86 X-TWMFWDREINNYTSLIHSLIFESQNQQEKNEQELLFLDKWASLWNWFN-Z
Seq 87 X-TWMFWDREINFYTSLIHSLIFESQNQQFKNEQELLFLDKWASLWNWFN-Z
Seq 88 X-TWMFWDREINFYTSLIHSLIFESQNQQFKNFQELLFLDKWASLWNWFN-Z

C. Combinations of substitutions at non-'a' or non-'d' positions

a d a d a d a d a d a d a d

Seq 89 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQECLELDKWASLWNWFN-Z
Seq 90 X-TWMEWDREINNYTSLIFSLIEESQNQQEKNEQEPLELDKWASLWNWFN-Z
Seq 91 X-TWMEWDREINNYTSLIPSLIEESQNQQEKNEQECLELDKWASLWNWFN-Z
Seq 92 X-TWMEWDREINNYTSLIPSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 93 X-TWMEWDREINNYTSLICSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 94 X-TWMEWDREINNYTSLICSLIEESQNQQEKNEQEPLELDKWASLWNWFN-Z

a d a d a d a d a d a d a d

Seq 95 X-TWMFWDREINNYTSLIFSLIEESQNQQEKNEQECLELDKWASLWNWFN-Z
Seq 96 X-TWMFWDREINNYTSLICSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 97 X-TWMCWDREINNYTSLIFSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 98 X-TWMCWDREINNYTSLIFSLIEESQNQQEKNEQECLELDKWASLWNWFN-Z
Seq 99 X-TWMFWDREINNYTSLIPSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 100 X-TWMFWDREINNYTSLIFSLIEESQNQQEKNEQEPLELDKWASLWNWFN-Z
Seq 101 X-TWMPWDREINNYTSLIFSLIEESQNQQEKNEQEFLELDKWASLWNWFN-Z
Seq 102 X-TWMFWDREINNYTSLIPSLIEESQNQQEKNEQECLELDKWASLWNWFN-Z
Seq 103 X-TWMFWDREINNYTSLICSLIEESQNQQEKNEQEPLELDKWASLWNWFN-Z
Seq 104 X-TWMPWDREINNYTSLIFSLIEESQNQQEKNEQECLELDKWASLWNWFN-Z
Seq 105 X-TWMFWDREINNYTSLIPSLIEESQNQQFKNEQECLELDKWASLWNWFN-Z

D. Single substitutions at 'a' or 'd' positions

```
                a d a d a d a d a d a d a d
Seq 106  X-TWMEWDREINNYTSLIHSLIEEDQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 107  X-TWMEWDREINNYTSLIHSLIEESQNDQEKNEQELLELDKWASLWNWFN-Z
Seq 108  X-TWMEWDREINNYTSLIHSLIEEGQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 109  X-TWMEWDREINNYTSLIHSLIEESQNGQEKNEQELLELDKWASLWNWFN-Z
Seq 110  X-TWMEWDREINNYTSLIHSLIEEHQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 111  X-TWMEWDREINNYTSLIHSLIEESQNHQEKNEQELLELDKWASLWNWFN-Z
Seq 112  X-TWMEWDREINNYTSLIHSLIEEPQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 113  X-TWMEWDREINNYTSLIHSLIEESQNPQEKNEQELLELDKWASLWNWFN-Z
Seq 114  X-TWMEWDREINNYTSLIHSLIEECQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 115  X-TWMEWDREINNYTSLIHSLIEESQNCQEKNEQELLELDKWASLWNWFN-Z
```

E. Multiple substitutions at 'a' or 'd' positions

These substitutions are illustrated using Asp as an example. Other possibilities include amino acid substitutions at similar positions using glycine, histidine, proline and/or cysteine as amino acids of choice.

```
                a d a d a d a d a d a d a d
Seq 116  X-TWMEWDREDNNYTSLIHSLIEEDQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 117  X-TWMEWDREINNDTSLIHSLIEEDQNQQEKNEQELLELDKWASLWNWFN-Z
Seq 118  X-TWMEWDREDNNYTSLIHSLIEESQNDQEKNEQELLELDKWASLWNWFN-Z
Seq 119  X-TWMEWDREINNDTSLIHSLIEESQNDQEKNEQELLELDKWASLWNWFN-Z
Seq 120  X-TWMEWDREDNNYTSLIHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 121  X-TWMEWDREINNDTSLIHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 122  X-TWMEWDREDNNYTSLIHSLIEESQNQQEKNEQDLLELDKWASLWNWFN-Z
Seq 123  X-TWMEWDREINNDTSLIHSLIEESQNQQEKNEQDLLELDKWASLWNWFN-Z
Seq 124  X-TWMEWDREDNNYTSLIHSLIEESQNQQEKNEQELLEDDKWASLWNWFN-Z
Seq 125  X-TWMEWDREINNDTSLIHSLIEESQNQQEKNEQELLEDDKWASLWNWFN-Z
Seq 126  X-TWMEWDREDNNYTSLIHSLIEESQNQQEKNEQELLELDKDASLWNWFN-Z
Seq 127  X-TWMEWDREINNDTSLIHSLIEESQNQQEKNEQELLELDKDASLWNWFN-Z
                a d a d a d a d a d a d a d
Seq 128  X-TWMEWDREINNYTSLDHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 129  X-TWMEWDREINNYTSLIHSDIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 130  X-TWMEWDREINNYTSLDHSLIEESQNQQEKNEQDLLELDKWASLWNWFN-Z
Seq 131  X-TWMEWDREINNYTSLIHSDIEESQNQQEKNEQDLLELDKWASLWNWFN-Z
Seq 132  X-TWMEWDREINNYTSLDHSLIEESQNQQEKNEQELLEDDKWASLWNWFN-Z
Seq 133  X-TWMEWDREINNYTSLIHSDIEESQNQQEKNEQELLEDDKWASLWNWFN-Z
Seq 134  X-TWMEWDREINNYTSLDHSLIEESQNQQEKNEQELLELDKDASLWNWFN-Z
Seq 135  X-TWMEWDREINNYTSLIHSDIEESQNQQEKNEQELLELDKDASLWNWFN-Z
Seq 136  X-TWMEWDREINNYTSLDHSLIEESQNQQEKNEQELLELDKWASLDNWFN-Z
Seq 137  X-TWMEWDREINNYTSLIHSDIEESQNQQEKNEQELLELDKWASLDNWFN-Z
                a d a d a d a d a d a d a d
Seq 138  X-TWMEWDREDNNYTSLIHSDIEESQNQQEKDEQELLELDKWASLDNWFN-Z
```

Seq 139 X-TWMEWDREINNDTSLIHSDIEESQNQQEKDEQELLELDKWASLDNWFN-Z
Seq 140 X-TWMEWDREDNNYTSLDHSLIEESQNQQEKDEQELLELDKWASLDNWFN-Z
Seq 141 X-TWMEWDREDNNYTSLIHSDIEESQNQQEKNEQDLLELDKWASLDNWFN-Z
Seq 142 X-TWMEWDREDNNYTSLIHSDIEESQNQQEKNEQDLLELDKDASLWNWFN-Z

F. Combinations of substitutions at 'a' or 'd' positions

a d a d a d a d a d a d a d
Seq 143 X-TWMEWDREINNYTSLDHSLIEESQNQQEKGEQELLELDKWASLWNWFN-Z
Seq 144 X-TWMEWDREINNYTSLDHSLIEESQNQQEKHEQELLELDKWASLWNWFN-Z
Seq 145 X-TWMEWDREINNYTSLDHSLIEESQNQQEKPEQELLELDKWASLWNWFN-Z
Seq 146 X-TWMEWDREINNYTSLDHSLIEESQNQQEKCEQELLELDKWASLWNWFN-Z
Seq 147 X-TWMEWDREINNYTSLGHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 148 X-TWMEWDREINNYTSLGHSLIEESQNQQEKHEQELLELDKWASLWNWFN-Z
Seq 149 X-TWMEWDREINNYTSLGHSLIEESQNQQEKPEQELLELDKWASLWNWFN-Z
Seq 150 X-TWMEWDREINNYTSLGHSLIEESQNQQEKCEQELLELDKWASLWNWFN-Z
Seq 151 X-TWMEWDREINNYTSLHHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 152 X-TWMEWDREINNYTSLHHSLIEESQNQQEKGEQELLELDKWASLWNWFN-Z
a d a d a d a d a d a d a d
Seq 153 X-TWMEWDREINNYTSLHHSLIEESQNQQEKPEQELLELDKWASLWNWFN-Z
Seq 154 X-TWMEWDREINNYTSLHHSLIEESQNQQEKCEQELLELDKWASLWNWFN-Z
Seq 155 X-TWMEWDREINNYTSLPHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 156 X-TWMEWDREINNYTSLPHSLIEESQNQQEKGEQELLELDKWASLWNWFN-Z
Seq 157 X-TWMEWDREINNYTSLPHSLIEESQNQQEKHEQELLELDKWASLWNWFN-Z
Seq 158 X-TWMEWDREINNYTSLPHSLIEESQNQQEKCEQELLELDKWASLWNWFN-Z
Seq 159 X-TWMEWDREINNYTSLCHSLIEESQNQQEKDEQELLELDKWASLWNWFN-Z
Seq 160 X-TWMEWDREINNYTSLCHSLIEESQNQQEKGEQELLELDKWASLWNWFN-Z
Seq 161 X-TWMEWDREINNYTSLCHSLIEESQNQQEKHEQELLELDKWASLWNWFN-Z
Seq 162 X-TWMEWDREINNYTSLCHSLIEESQNQQEKPEQELLELDKWASLWNWFN-Z
a d a d a d a d a d a d a d
Seq 163 X-TWMEWDREDNNYTSLDHSLIEESQNQQEKGEQELLELDKDASLWNWFN-Z
Seq 164 X-TWMEWDREDNNYTSLDHSLIEESQNQQEKHEQELLELDKCASLWNWFN-Z
Seq 165 X-TWMEWDREDNNYTSLGHSLIEESQNQQEKPEQELLELDKCASLWNWFN-Z
Seq 166 X-TWMEWDREDNNYTSLGHSLIEESQNQQEKDEQELLELDKPASLWNWFN-Z
Seq 167 X-TWMEWDREDNNYTSLGHSLIEESQNQQEKDEQELLELDKCASLWNWFN-Z
Seq 168 X-TWMEWDREDNNYTSLGHSLIEESQNDQEKDEQELLELDKCASLWNWFN-Z
Seq 169 X-TWMEWDREDNNYTSLGHSLIEESQNDQEKPEQELLELDKCASLWNWFN-Z
Seq 170 X-TWMEWDREGNNYTSLPHSLIEESQNDQEKPEQELLELDKCASLWNWFN-Z
Seq 171 X-TWMEDDREGNNYTSLGHSLIEESQNDQEKPEQELLELDKCASLWNWFN-Z
Seq 172 X-TWMEWDREDNNYTSLIHSGIEESQNDQEKPEQELLELDKCASLWNWFN-Z

"X" and "Z" used in these examples are used equivalently to the "α" and "γ" moieties of the instant polypeptides represented by the formula α-β-γ.

Example II

A. Materials and Methods

Peptides were tested for inhibition of HIV-1_{JR-FL} envelope-mediated membrane fusion in a fluorometric resonance energy transfer assay (J. Virol. et al., 70:6437, 1996). Briefly, HeLa cells that stably express the envelope glycoproteins of the primary R5 virus HIV-1_{JR-FL} were membrane-labeled overnight with octadecyl fluorescein (F18; Molecular Probes, Eugene, OR), while PM1 cells were similarly labeled with octadecyl rhodamine (R18). PM1 is a T cell line that endogenously expresses CCR5 as well as CXCR4 and supports entry of R5 and X4 viruses. After labeling, the cells were washed in PBS containing 15% fetal bovine serum (PFBS buffer) and combined in equal numbers in 96-well microtiter plates (Becton-Dickinson, Franklin Lakes, NJ) in PFBS buffer. Serially diluted inhibitors were added at this time either individually or in combination in a fixed concentration ratio. The plates were incubated for 4h at 37°C and then read on a fluorescence plate reader (PE Biosystems, Foster City, CA).

Fluorescent RET from F18 to R18 occurs only when the dyes are placed in the same cellular membrane following fusion, and thus can be directly related to the extent of HIV-1 membrane fusion. RET observed in the presence of inhibitors was compared with that observed in their absence (0% inhibition) and in the presence of the anti-CD4 antibody Leu 3a (100% inhibition; Becton-Dickinson). The HeLa envelope cell lines can be prepared using the methods described in the original publication (Litwin et al., J. Virol. 70:6437, 1996). PM1 cells are available from the NIH AIDS Research and Reference Reagent Program (Cat. # 3038).

B. Results

Potency of gp41 Peptides in the RET Fusion Assay

Results are summarized in Table 2.

Table 2.

Peptide	Sequence	IC50 (nM)	IC90 (nM)
T-1249	Ac-WQEWEQK-------ITALLEQAQIQQEKNEYELQKLDKWASLWENF-NH2	16	49
PRO-1	TWLEWDREINNYTSLIHSLIEESQNQQEKIEQELLELDKWASLWNWFNITNWLWYIK		1120
PRO-2	NHTTWLEWDREINNYTSLIHSLIEESQNQQEKIEQELLELDKWASLWNWFNI		151
PRO-3	NHTTWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNI		754
PRO-4	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK		337
PRO-5	NHTTWLEWDREINNYTSLIHSLIEESQNIQEKNEQELLELDKWASLWNWFNI		127
PRO-6	NHTTWLEWDREINNYTSLIHSLIEESQNQQEKNEQILLELDKWASLWNWFNI		289
PRO-7	NHTTWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASL		1340
PRO-11	TWLEWDREINNYTSLIHSLIEEIQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	229	421
PRO-12	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQILLELDKWASLWNWFNITNWLWYIK	205	357
PRO-13	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	115	284
PRO-14	TWLEWDREINNYTSLIHSLIEESQNIQEKNEQELLELDKWASLWNWFNITNWLWYIK	154	290
PRO-15	TWLEWDREINNYTSLIHSLIEESQNQQEKIEQELLELDKWASLWNWFNITNWLWYIK	301	436
PRO-16	TWLEWDREINNYTSLIHSLIEEIQNIQEKNEQELLELDKWASLWNWFNITNWLWYIK	592	1447
C-2Q	TWLEWDREINNYTSLIHSLIEESQQQQEKAEQELLELDKWASLWNWFNITNWLWYIK	101	296
C-4Q	TWLEWDREIQQYSSLIASLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	71	232
C-6Q	TWLEWDREIQQYSSLIASLIEESQQQQEKAEQELLELDKWASLWNWFNITNWLWYI	96	278
W-34	WLEWDREINNYTSLIHSLIEESQNQQEKNEQELL	>4000	>4000
T-1249	Ac-WQEWEQK-------ITALLEQAQIQQEKNEYELQKLDKWASLWENF-NH2	1.6	8.5
C157S	TWLEWDREISSYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	103	300
C157D	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWDWFNITNWLWYIK	283	538
C157DS	TWLEWDREISSYTSLIHSLIEESQNQQEKNEQELLELDKWASLWDWFNITNWLWYIK	482	780
T-1249	Ac-WQEWEQK-------ITALLEQAQIQQEKNEYELQKLDKWASLWENF-NH2	7.9	34
T-1249	Ac-WQEWEQK-------ITALLEQAQIQQEKNEYELQKLDKWASLWENF-NH2	2.8	14.4
T-20	Ac-YTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF-NH2	9.4	47
C-57	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	58	1234
C-48	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFN	10	31
C-49	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNI	29	585
C-50	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNIT	20	40
C-51	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITN	7.3	40
C-52	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNW	33	58
C-53	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWL	42	429
C-55	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWY	93	221
C-56	TWLEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYI	151	358
YF628	TWLEWDREINNFTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	97	180
SA639	TWLEWDREINNYTSLIHSLIEEAQNQQEKNEQELLELDKWASLWNWFNITNWLWYIK	55	185
NQ651	TWLEWDREINNYTSLIHSLIEESQQQQEKNEQELLELDKWASLWNWFNITNWLWYIK	68	113
QL652	TWLEWDREINNYTSLIHSLIEESQNLQEKNEQELLELDKWASLWNWFNITNWLWYIK	89	122
NA656	TWLEWDREINNYTSLIHSLIEESQNQQEKAEQELLELDKWASLWNWFNITNWLWYIK	107	150
	a d a d a d a d a d a d a d a d		

Table 2 continued.

Peptide	Sequence	IC50 (µg/ml)	IC90 (µg/ml)
T-1249	Ac-WQEWEQK-------ITALLEQAQIQQEKNEYELQKLDKWASLWENF-NH2	0.04	0.13
T-20	Ac-YTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF-NH2	0.10	0.39
PRO-17	TWLEWDREIQQFSSLIASLIEEAQQLQEKAEQELLELDKWASLWNWFNITNWLWYIK	> 1	> 1
PRO-18	TWLEWDREIQQFSSLIASLIEEAQQLQEKAEQELLELDKWASLWNWFN	> 1	> 1
PRO-19	TWLEWDREIQQFSSLIASLIEEAQQLQEKAEQELLELDKWASLKKQFN	0.42	> 1
PRO-20	TWLEWDREIQQFSSLIASLIEEAQQLQEKAEQELLELDKWASLKKWFN	0.51	0.90
PRO-21	TWKEWDREIQQFSSLIASLIEEAQQLQEKAEQELLELDKWASLWNWFN	> 1	> 1
PRO-22	TWLEWDREIQQFSSKIASLIEEAQQLQEKAEQELLELDKWASLWNWFN	0.46	> 1
PRO-23	TWLEWDREIQQFSSLIASLKEEAQQLQEKAEQELLELDKWASLWNWFN	> 1	> 1
PRO-24	TWKEWDREIQQFSSKIASLKEEAQQLQEKAEQELLELDKWASLWNWFN	> 1	> 1
PRO-25	TWKEWDREIQQFSSKIASLKEEAQQLQEKAEQEKKELDKWASLWNWFN	> 1	> 1
PRO-26	TWLEWDREIQQFSSLIASLIEEAQQLQEKAEQELLELDKWASLKKWFNSSNKKW	0.52	> 1
JR-FL	TWLEWEREIDNYTSLIYTLIEESQNQQEKNEQELLELDKWASLWNWFD	0.19	> 1
	a d a d a d a d a d a d a d a d		

Example III

A. Materials and Methods

Virus-cell fusion was examined in an HIV-1 entry assay that employs single-cycle reporter viruses as described previously (Dragic et al., Nature, 381:667, 1996). Briefly, NL*luc*$^{+}$*env*$^{-}$ reporter viruses were complemented *in trans* with HIV-1 envelope glycoproteins by cotransfecting 293T cells with an NL4/3Δ*env*-luciferase plasmid and an HIV-1 *env*-expressing plasmid. Env-complemented viruses were harvested from the culture supernatants, normalized for p24 content, and used to infect U87-CD4-CCR5 (HIV-1_{JR-FL} and T-20-resistant variants) or U87-CD4-CXCR4 (HIV-1_{LAI}) cells in the presence of inhibitors. Following a 2 hour incubation at 37°C, cells were washed and supplied fresh inhibitor-containing media. Fresh media, without inhibitors, was added again after 12h. After 72h, cells were lysed and luciferase activity was measured as described (Dragic et al., Nature, 381:667, 1996). Peptides were tested against wild-type HIV-1_{JR-FL} and HIV-1_{LAI} reporter viruses (top panels) as well as mutant HIV-1_{JR-FL} reporter viruses containing known T-20-conferring resistance mutations. The T-20-resistant viruses contain point mutations in the ^{538}GIV^{540} region of HIV-1_{JR-FL} gp160 (Genbank AAB05624). The 36D, 38M and 38A reporter viruses contain G538D, V540M and V540A mutations, respectively. The point mutations were inserted into the HIV-1_{JR-FL} *env*-expressing plasmid using the QuickChange site-directed mutagenesis kit (Stratagene). The U87-CD4-CCR5 cells are available from the NIH AIDS Research and Reference Reagent Program (Cat. #4035). The U87-CD4-CCR5 cells are available from the NIH AIDS Research and Reference Reagent Program (Cat. #4036). C-48 was observed to be essentially equipotent against wild-type and T-20-resistant viruses.

B. Results

Results are summarized in Figures 2A-2E.

Example IV

A. Materials and Methods.

Peptides were tested in duplicate independent assays for inhibition of HIV-1_{JR-FL} envelope-mediated membrane fusion in a fluorometric resonance energy transfer assay (Litwin et al., J. Virol. 70:6437, 1996). Briefly, HeLa cells that stably express the envelope glycoproteins of the primary R5 virus HIV-1_{JR-FL} were membrane-labeled overnight with octadecyl fluorescein (F18; Molecular Probes, Eugene, OR), while PM1 cells were similarly labeled with octadecyl rhodamine (R18). PM1 is a T cell line that endogenously expresses CCR5 as well as CXCR4 and supports entry of R5 and X4 viruses. After labeling, the cells were washed in PBS containing 15% fetal bovine serum (PFBS buffer) and combined in equal numbers in 96-well microtiter plates (Becton-Dickinson, Franklin Lakes, NJ) in PFBS buffer. Serially diluted inhibitors were added at this time either individually or in combination in a fixed concentration ratio. The plates were incubated for 4h at 37°C and then read on a fluorescence plate reader (PE Biosystems, Foster City, CA).

Fluorescent RET from F18 to R18 occurs only when the dyes are placed in the same cellular membrane following fusion, and thus can be directly related to the extent of HIV-1 membrane fusion. RET observed in the presence of inhibitors was compared with that observed in their absence (0% inhibition) and in the presence of the anti-CD4 antibody Leu 3a (100% inhibition; Becton-Dickinson). The HeLa envelope cell lines can be prepared using the methods described in the original publication (Litwin et al., J. Virol. 70:6437, 1996). PM1 cells are available from the NIH AIDS Research

and Reference Reagent Program (Cat. # 3038).

B. Results

Activity of C85D Peptide in the RET Fusion Assay

Results are summarized in Table 3.

Table 3.

Peptide	Sequence	Potency, (μg/mL)	
		IC50	IC90
T-1249	WQEWEQKITALLEQAQIQQEKNEYELQKLDKWASLWENF	0.008	0.027
T-20	YTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF	0.12	0.55
C85D	NHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNIKIKQIEDKIK	0.059	0.254

Example V

A. Materials and Methods

Peptides were tested for inhibition of HIV-1_{JR-FL} envelope-mediated membrane fusion in a fluorometric resonance energy transfer assay (Litwin et al., J. Virol. 70:6437, 1996). Briefly, HeLa cells that stably express the envelope glycoproteins of the primary R5 virus HIV-1_{JR-FL} were membrane-labeled overnight with octadecyl fluorescein (F18; Molecular Probes, Eugene, OR), while PM1 cells were similarly labeled with octadecyl rhodamine (R18). PM1 is a T cell line that endogenously expresses CCR5 as well as CXCR4 and supports entry of R5 and X4 viruses. After labeling, the cells were washed in PBS containing 15% fetal bovine serum (PFBS buffer) and combined in equal numbers in 96-well microtiter plates (Becton-Dickinson, Franklin Lakes, NJ) in PFBS buffer. Serially diluted inhibitors were added at this time either individually or in combination in a fixed concentration ratio. The plates were incubated for 4h at 37°C and then read on a fluorescence plate reader (PE Biosystems, Foster City, CA). Fluorescent RET from F18 to R18 occurs only when the dyes are placed in the same cellular membrane following fusion, and thus can be directly related to the extent of HIV-1 membrane fusion. RET observed in the presence of inhibitors was compared with that observed in their absence (0% inhibition) and in the presence of the anti-CD4 antibody Leu 3a (100% inhibition; Becton-Dickinson). The HeLa envelope cell lines can be prepared using the methods described in the original publication (Litwin et al., J. Virol. 70:6437, 1996). PM1 cells are available from the NIH AIDS Research and Reference Reagent Program (Cat. # 3038).

B. Results

Activity of C85FL Peptide in the RET Fusion Assay

Results are summarized in Table 4.

Table 4.

JR-FL RET	Potency					
	IC50 (nM)			IC90 (nM)		
Peptide	C85FL	T-20	T-1249	C85FL	T-20	T-1249
Assay #1	*70.3*	*36.6*	*16.6*	*145.2*	*165.2*	*52.2*
Assay #2	*105.7*	*25.0*	*15.9*	*206.4*	*58.3*	*54.1*
Average	88.0	30.8	16.3	175.8	111.8	53.2

Example VI

Recombinant expression and purification of HIV gp41 peptides

Recombinant peptides are expressed in *Escherichia coli* strain BL21(DE3)/pLysS (Novagen, Madison, WI) by using the T7 expression system (Studier et al., Methods Enzymol., 185:60, 1990) and as described (Ji et al., J Virol., 73:8578, 1999). Cells, freshly transformed with the gp41 peptide encoding plasmid, are grown at 37°C in LB media to an optical density of 0.8 at 600 nm and induced with 0.5 mM isopropylthio-β-D-galactoside for 3-4 hours. Cells are harvested by centrifugation and lysed at 0°C by glacial acetic acid. The bacterial lysate is centrifuged (35,000 g for 30 minutes) to separate the soluble fraction from

inclusion bodies. The soluble fraction, containing the recombinant gp41 peptide, is dialyzed into 5% acetic acid overnight at room temperature. The peptides are purified from the soluble fraction to homogeneity by reverse-phase high-performance liquid chromatography (Waters Corp., Milford, MA) on a Vydac C-18 preparative column (Grace Vydac, Hesperia, CA), using a water-acetonitrile gradient in the presence of 0.1% trifluoroacetic acid, and lyophilized. The amount of peptide produced is determined by weighing the lyophilized powder and is confirmed by using tyrosine and tryptophan absorbance at 280 nm in 6 M guanidinium hydrochloride (Edelhoch, 1967, Biochemistry 6:1948-1954).

Example VII

Pharmacokinetics of gp41 peptides

250-300 gram male CD rats, double jugular catheter, obtained from Charles River Laboratories (Wilmington, MA) can be used to initially assess the plasma pharmacokinetic properties of the gp41 peptides administered by intravenous injection (IV) as a single dose. Other routes of administration include subcutaneous (SC) and intramuscular (IM) injections.

For the IV route, gp41 peptides, trace-labeled with ^{125}I, are injected in one jugular catheter in a volume of 200 µl of peptide solution (approximate concentrations: 3.75 mg/ml, 1.5 mg/ml, and 0.6 mg/ml). The concentration of the peptide solution is determined by absorbance at 280 nm in the presence of 6 M guanidinium hydrochloride (Edelhoch, 1967, Biochemistry 6:1948-1954) and adjusted based on animal weight such that each animal received a dose of 2.5 mg/kg, 1.0 mg/kg, or 0.4 mg/kg. Approximately 250-300 µl of blood is removed at predetermined time intervals post-

administration (0, 15, 30 min and 1, 2, 4, 6, 8, and 24 hours) and added to EDTA capiject tubes. Plasma is removed from pelleted cells upon centrifugation and either frozen or immediately processed for determination of the residual radioactivity in the serum by liquid scintillation counting. α and β plasma half-lives are calculated using a two-compartment model with first-order elimination from the central compartment (PCNONLIN version 4; SCI Software, Lexington, KY). Other calculated plasma parameters include maximum plasma concentration (C_{max}), area under the curve for plasma concentration plotted against time with extrapolation to infinity (AUC_{∞}).

Alternatively, a non-radioactive method can be used. Peptides are injected in one jugular catheter in a volume of 200 µl of peptide solution (approximate concentrations: 3.75 mg/ml, 1.5 mg/ml, and 0.6 mg/ml). The concentration of the peptide solution is determined by absorbance at 280 nm in the presence of 6 M guanidinium hydrochloride (Edelhoch, 1967, Biochemistry 6:1948-1954) and adjusted based on animal weight such that each animal received a dose of 2.5 mg/kg, 1.0 mg/kg, or 0.4 mg/kg. Approximately 250-300 µl of blood is removed at predetermined time intervals post-administration (0, 15, 30 minutes and 1, 2, 4, 6, 8, and 24 hours) and added to EDTA capiject tubes. Plasma is removed from pelleted cells upon centrifugation and either frozen or immediately processed for determination of the residual peptide concentration in the serum by fluorescence HPLC analysis as previously described (Lawless et al., J. Chrom. B. 707:213, 1998). In this assay, 100 µl of sample plasma is added to 900 µl of precipitation buffer (acetonitrile, 1.0% TFA, 1.0% *n*-nonyl-β-D-glycopyranoside) resulting in precipitation of the majority of plasma proteins. Following

centrifugation at 10,000 rpm for 10 minutes, 400 µl of the supernatant is removed and added to 600 µl of HPLC grade water. Serial dilutions are performed as dictated by the concentration of peptide present in each sample in dilution buffer comprised of 40% precipitation buffer and 60% HPLC water. In addition to sample dilutions, serial dilutions of dosing solution are performed in buffer as well as in plasma and used to generate a standard curve relating the peak area at a specific retention time to known concentrations of the peptide. This curve is then used to calculate the concentration of the peptide (eluting with the same retention time) in plasma taking into account all dilutions performed and quantity injected onto the column.

Example VII

Serum stability of gp41 peptides

A ^{125}I trace-labeled gp41 peptide stock solution (approximate concentration: 1.0 mg/ml) is diluted 100-fold with rat serum or human serum to obtain a peptide concentration of approximately 10 µg/ml. The concentration of the peptide stock solution is determined by absorbance at 280 nm in the presence of 6 M guanidinium hydrochloride (Edelhoch, 1967, Biochemistry 6:1948-1954). The mixture is sterile-filtered and incubated in a CO_2-enriched atmosphere at 37°C over a period of 20 days. At predetermined time intervals post-dilution (0, 1, 2, 4, 8, 14, and 20 days) samples are taken and the serum proteins and intact peptides are precipitated with 10% TCA for 1 hr on ice and separated from non-precipitated materials by centrifugation. Intact peptides will fractionate into the pellet, whereas degraded peptides and peptide derived free amino acids will remain in the supernatant. Analysis of pellet and supernatant fractions

for radioactivity will enable to quantify the serum stability of the peptide by plotting the amount of intact peptide as a function of time.

Alternatively, a non-radioactive method can be used. A gp41 peptide stock solution (approximate concentration: 1.0 mg/ml) is diluted 100-fold with rat serum or human serum to obtain a peptide concentration of approximately 10 μg/ml. The concentration of the peptide stock solution is determined by absorbance at 280 nm in the presence of 6 M guanidinium hydrochloride (Edelhoch, 1967, Biochemistry 6:1948-1954). The mixture is sterile-filtered and incubated in a CO_2-enriched atmosphere at 37°C over a period of 20 days. At predetermined time intervals post-dilution (0, 1, 2, 4, 8, 14, and 20 days) samples are taken and immediately processed for determination of the residual peptide concentration in the serum by fluorescence HPLC analysis as previously described (Lawless et al., J. Chrom. B. 707:213, 1998). In this assay, 100 μl of sample plasma is added to 900 μl of precipitation buffer (acetonitrile, 1.0% TFA, 1.0% *n*-nonyl-β-D-glycopyranoside) resulting in precipitation of the majority of plasma proteins. Following centrifugation at 10,000 rpm for 10 minutes, 400 μl of the supernatant is removed and added to 600 μl of HPLC grade water. Serial dilutions are performed as dictated by the concentration of peptide present in each sample in dilution buffer comprised of 40% precipitation buffer and 60% HPLC water. In addition to sample dilutions, serial dilutions of dosing solution are performed in buffer as well as in plasma and used to generate a standard curve relating the peak area at a specific retention time to known concentrations of the peptide. This curve is then used to calculate the concentration of the peptide (eluting with the same

retention time) in plasma taking into account all dilutions performed and quantity injected onto the column.

Example VIII

Cytotoxicity of gp41 peptides

In order to measure possible cytotoxic effects of the gp41 peptides, XTT assays (Weislow, O.S. et al., 1989, J. Natl. Cancer Inst. 81:577-586) are performed in the presence of varying concentrations of peptide in order to effectively establish a selective index (SI). A TC_{50} is determined in this assay by incubating cells in the presence and absence of serially diluted peptide followed by the addition of XTT. In surviving/metabolizing cells XTT is reduced to a soluble brown dye, XTT-formazan. Absorbance is read and comparisons made between readings in the presence and absence of peptide to determine a TC_{50} utilizing the Karber method (see. e.g., Lennette, E. H. et al., eds., 1969, "Diagnostic Procedures for Viral and Rickettsial Infections," American Public Health Association, Inc., fourth ed., pp. 47-52). Molt 4, HeLa, and CEM cells (80,000 cells/well) are plated and incubated with serially diluted peptide for 24 hours in a total volume of 100 µl. Following incubation, 25 µl of XTT working stock (1 mg/ml XTT, 250 µM PMS in complete medium containing 5% DMSO) is added to each well and the plates are incubated at 37°C. Color development is quantitated by absorbance and the values generated from peptide containing wells are expressed as a percentage of the untreated control wells.

What is claimed is:

1. A polypeptide represented by the formula α-β-γ, wherein

(a) α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero;

(b) β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment		1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein

(i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or

(ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue

is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and

(c) γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

2. The polypeptide of claim 1, wherein in β, a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 3 or 4.

3. The polypeptide of claim 1, wherein in β, a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 3 or 4.

4. The polypeptide of claim 1, wherein in β, a proline residue resides at position 2, 3, 5, 6, or 7 of segment 3 or 4.

5. The polypeptide of claim 1, wherein in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, or 5.

6. The polypeptide of claim 1, wherein in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 2, and (ii) a phenylalanine residue

resides at position 2, 3, 5, 6 or 7 of segment 3, 4, 5, or 6.

7. The polypeptide of claim 1, wherein in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 3, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 4, 5 or 6.

8. The polypeptide of claim 1, wherein in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, or 5.

9. The polypeptide of claim 1, wherein in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 2, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, 5 or 6.

10. The polypeptide of claim 1, wherein in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 3, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 4, 5 or 6.

11. The polypeptide of claim 1, wherein in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, or 5.

12. The polypeptide of claim 1, wherein in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 2, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4, 5 or 6.

13. The polypeptide of claim 1, wherein in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 3, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 4, 5 or 6.

14. The polypeptide of claim 1, wherein in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

15. The polypeptide of claim 1, wherein in β, (i) a phenylalanine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

16. The polypeptide of claim 1, wherein in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

17. The polypeptide of claim 1, wherein in β, (i) a cysteine residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a proline residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

18. The polypeptide of claim 1, wherein in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment 1, and (ii) a phenylalanine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

19. The polypeptide of claim 1, wherein in β, (i) a proline residue resides at position 2, 3, 5, 6, or 7 of segment

1, and (ii) a cysteine residue resides at position 2, 3, 5, 6 or 7 of segment 3, 4 or 5.

20. The polypeptide of claim 1, wherein in β, (i) a phenylalanine residue resides at position 2 or 6 of segment 3, and (ii) a phenylalanine residue resides at position 3 or 6 of segment 5.

21. The polypeptide of claim 1, wherein in β, an alanine residue resides at position 1 of segment 5.

22. The polypeptide of claim 1, wherein in γ, the amino acid sequence facilitates expression of the polypeptide in bacteria.

23. A composition comprising the polypeptide of claim 1 and a pharmaceutically acceptable carrier.

24. A nucleic acid encoding a polypeptide represented by the formula α-β-γ, wherein

(a) α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero;

(b) β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments

and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position	1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein

(i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or

(ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and

(c) γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

25. The nucleic acid of claim 24, wherein the nucleic acid is DNA or RNA.

26. The nucleic acid of claim 25, wherein the nucleic acid is DNA.

27. The nucleic acid of claim 24, wherein the nucleic acid is an expression vector.

28. The nucleic acid of claim 24, wherein the expression vector is selected from the group consisting of a plasmid, a cosmid, a bacteriophage and a eukaryotic virus.

29. The nucleic acid of claim 28, wherein the expression vector is a bacteriophage.

30. A host-vector system comprising a host cell transfected with an expression vector comprising a nucleic acid encoding a polypeptide represented by the formula α-β-γ, wherein

(a) α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero;

(b) β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein

(i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or

(ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and

(c) γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero.

31. The host-vector system of claim 30, wherein the expression vector is a bacteriophage.

32. A method for producing a polypeptide represented by the formula α-β-γ, wherein

(a) α represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero;

(b) β represents contiguous amino acid residues having a sequence which is a variant of the contiguous amino acid sequence having segments and amino acid residue position numbers as shown in "Formula I"

Formula I

Segment	1	2	3	4	5	6	7
	T	WMEWDRE	INNYTSL	IHSLIEE	SQNQQEK	NEQELLE	LDKWASL	WNWFN
Position		1234567	1234567	1234567	1234567	1234567	1234567	12345

wherein

(i) within any four consecutive segments, a cysteine, aspartate, glycine, histidine or proline residue is present at position 1 or 4 of at least one such consecutive segment; or

(ii) within any four consecutive segments, a cysteine, phenylalanine or proline residue is present at position 2, 3, 5, 6 or 7 of at least one such consecutive segment; and

(c) γ represents consecutive amino acids having a sequence $(aa)_n$, wherein "aa" represents any amino acid residue other than proline, each amino acid residue may be the same or different, and "n" represents the number of amino acid residues, wherein n may be greater than or equal to zero,

which method comprises (a) culturing the host-vector system of claim 30 under conditions permitting the expression of the polypeptide, and (b) recovering the polypeptide so expressed.

33. A method for inhibiting fusion of an HIV-1 virus and a CD4+ cell, comprising contacting the polypeptide of claim 1 with the HIV-1 virus and the CD4+ cell under conditions which, in the absence of the polypeptide, would permit fusion of the HIV-1 virus and CD4+ cell, thereby inhibiting fusion of the HIV-1 virus and CD4+ cell.

34. A method for reducing the likelihood of a subject's becoming infected with HIV-1 comprising administering to the subject a prophylactically effective amount of the polypeptide of claim 1, thereby reducing the likelihood of the subject's becoming infected with HIV-1.

35. A method for preventing or delaying the onset of, or slowing the rate of progression of, an HIV-1-related disease in an HIV-1-infected subject which comprises administering to the subject a therapeutically effective amount of the polypeptide of claim 1, thereby preventing or delaying the onset of, or slowing the

rate of progression of, the HIV-1-related disease in the subject.

36. An article of manufacture comprising a packaging material having therein the polypeptide of claim 1, and a label indicating a use of the polypeptide for (i) reducing the likelihood of a subject's being infected with HIV-1 and/or (ii) preventing or delaying the onset of, or slowing the rate of progression of, an HIV-related disease in an HIV-1-infected subject.

FIGURE 1

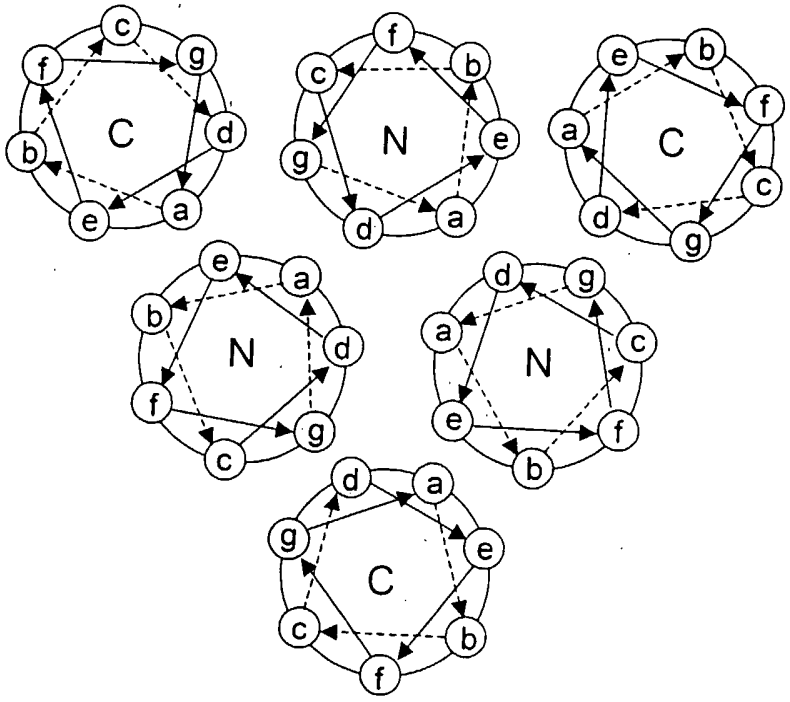


FIGURE 2

Location and Alignments of gp41 Peptides

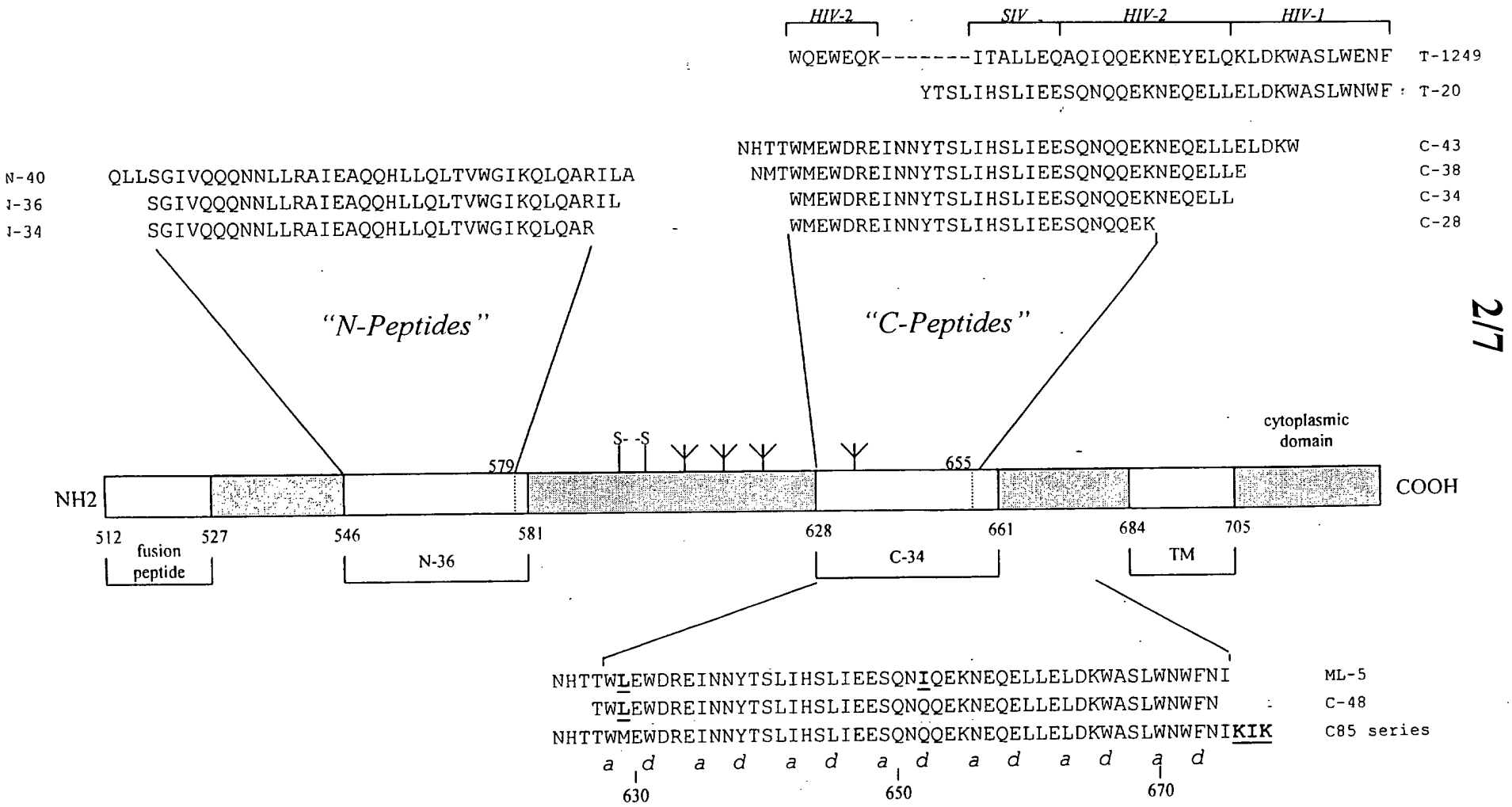


FIGURE 3A

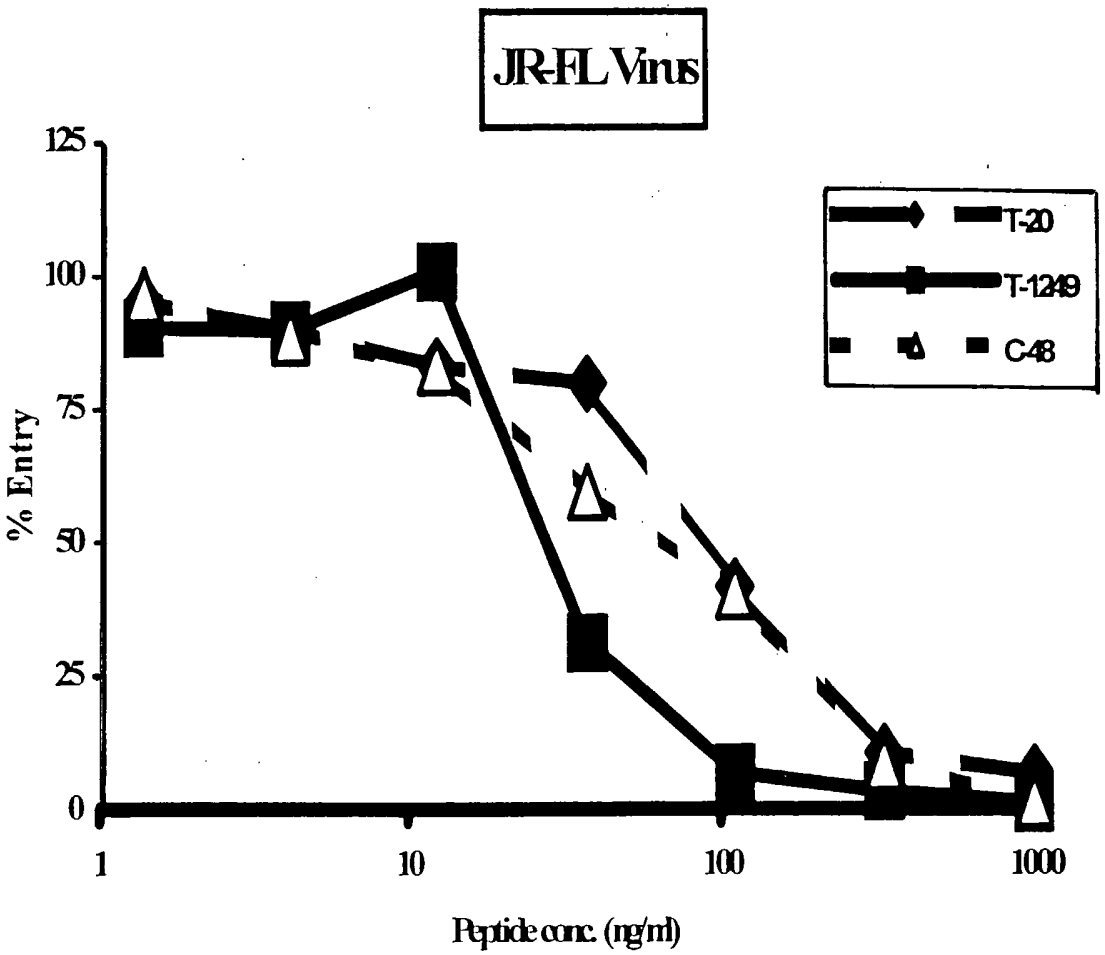
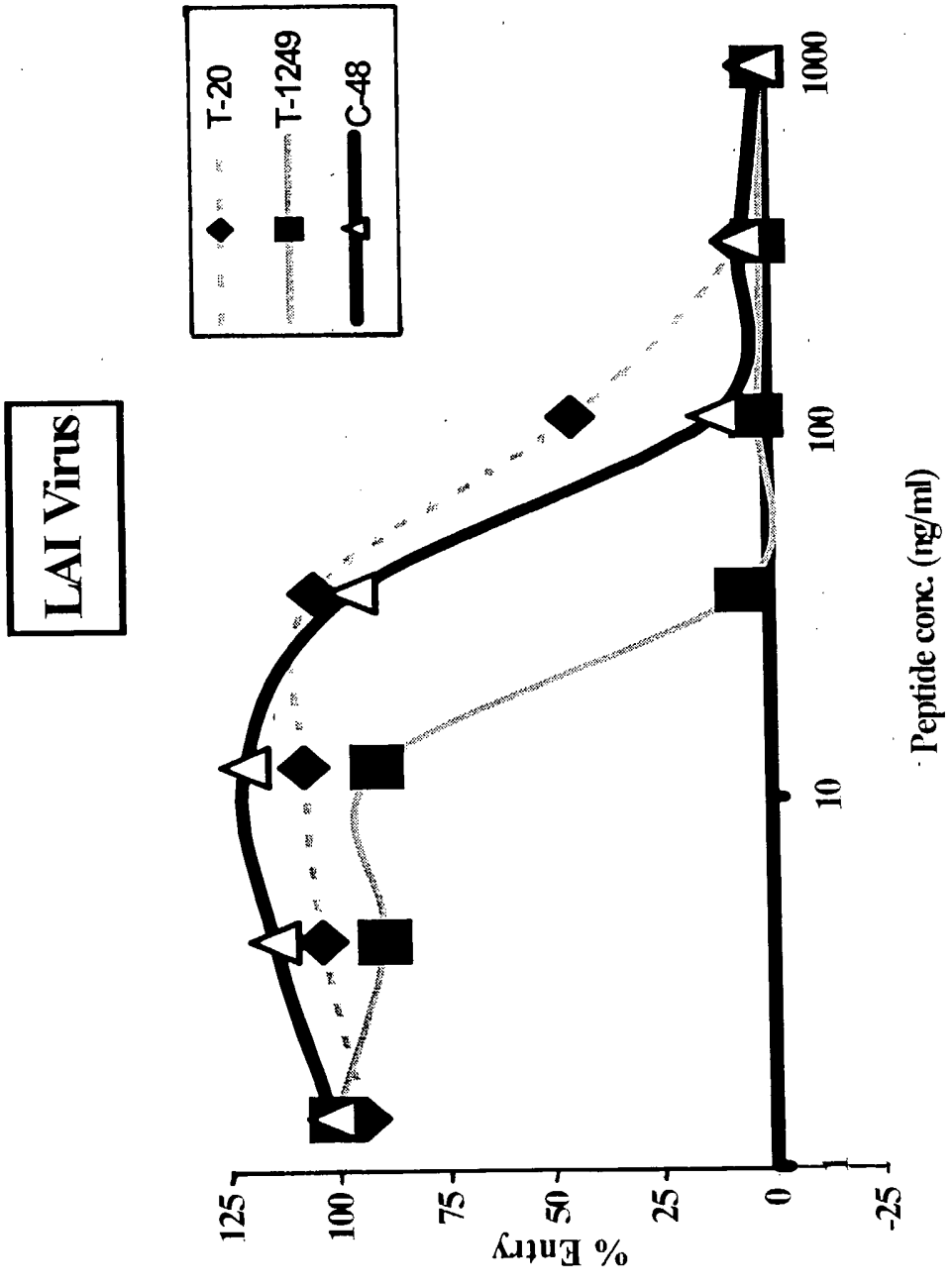

JR-FL Virus
T-20
T-1249
C-48
% Entry
125
100
75
50
25
0
1
10
100
1000
Peptide conc. (ng/ml)

FIGURE 3B


LAI Virus
T-20
T-1249
C-48
% Entry
125
100
75
50
25
0
-25
1
10
100
1000
Peptide conc. (ng/ml)

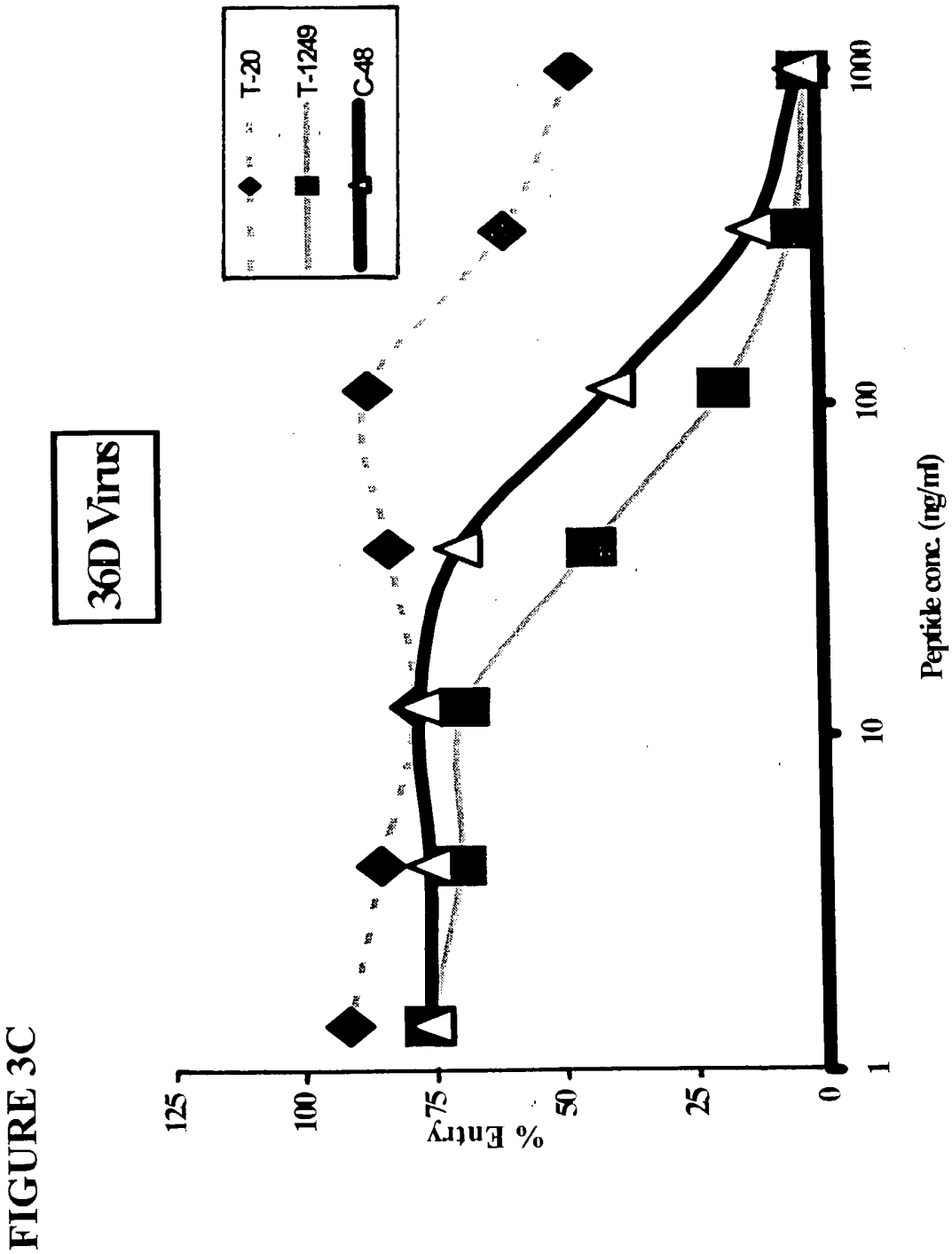

36D Virus
T-20
T-1249
C-48
% Entry
125
100
75
50
25
0
1
10
100
1000
Peptide conc. (ng/ml)

FIGURE 3C

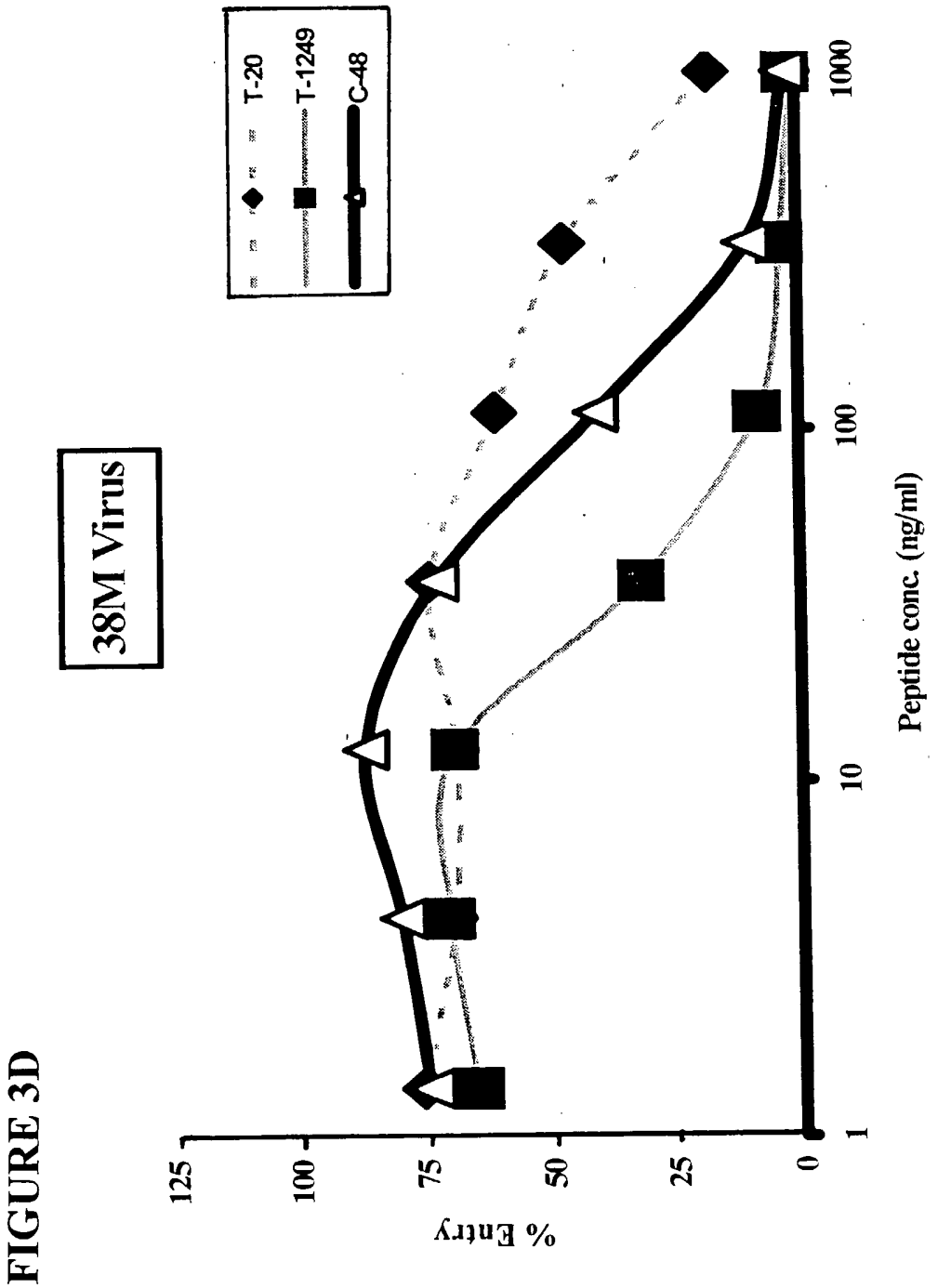

38M Virus
T-20
T-1249
C-48
% Entry
125
100
75
50
25
0
1
10
100
1000
Peptide conc. (ng/ml)

FIGURE 3D

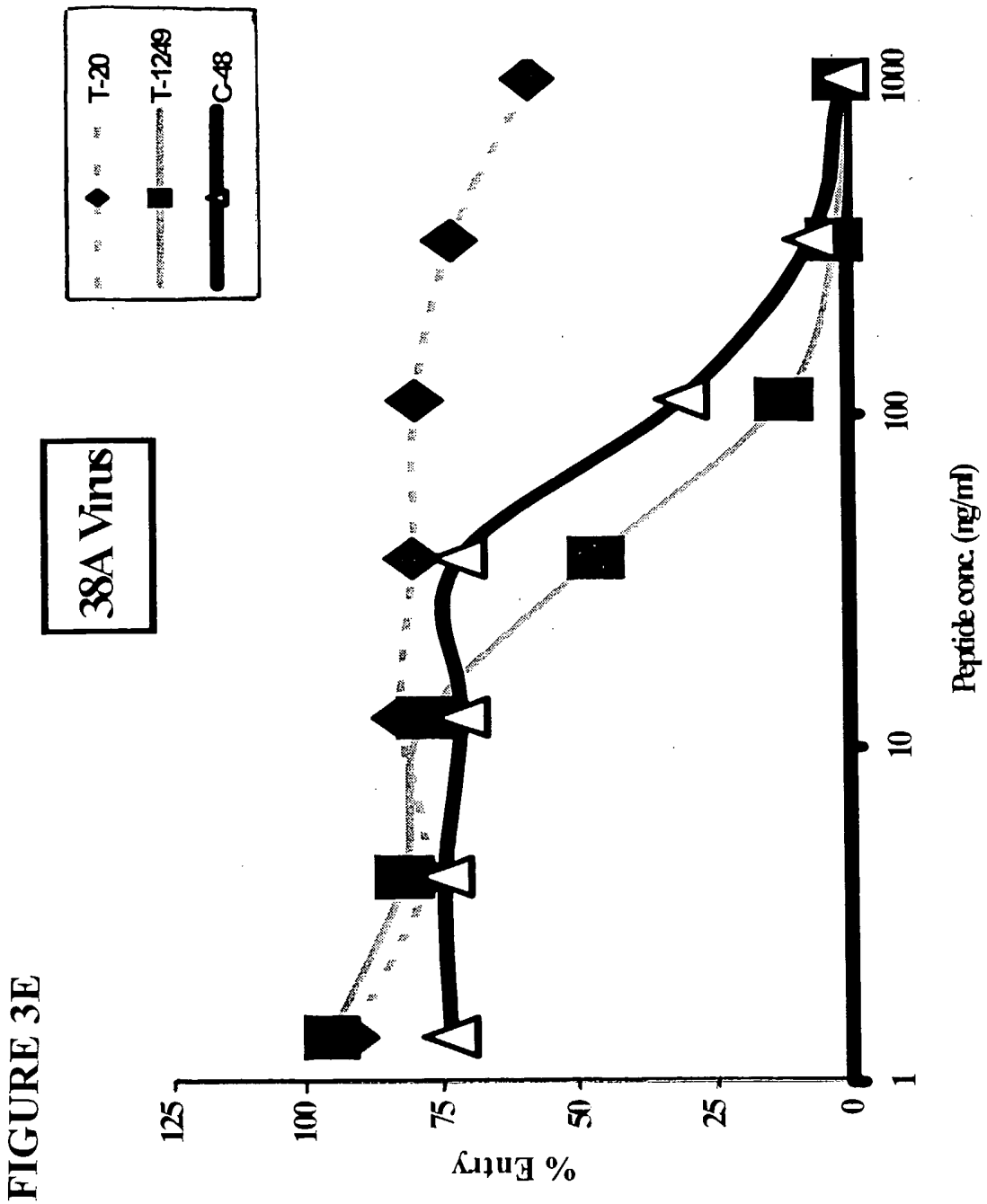

38A Virus
T-20
T-1249
C-48
% Entry
125
100
75
50
25
0
1
10
100
1000
Peptide conc. (ng/ml)

FIGURE 3E