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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 May 2003 (01.05.2003)

PCT

(10) International Publication Number
WO 03/035892 A2

(51) International Patent Classification[7]: C12Q

(21) International Application Number: PCT/US02/33305

(22) International Filing Date: 18 October 2002 (18.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/039,318 19 October 2001 (19.10.2001) US

(71) Applicant *(for all designated States except US)*: **THE SCRIPPS RESEARCH INSTITUTE** [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GHAZAL, Peter** [GB/GB]; 19 Warrender Park Crescent, Edinburgh EH9 1EA (GB). **BURTON, Dennis, R.** [GB/US]; 6044 Beaumont Avenue, La Jolla, CA 92037 (US). **HAHN, Klaus, M.** [US/US]; 10258 Saunders Drive, San Diego, CA 92131 (US).

(74) Agents: **ZIMMERMAN, Roger, P.** et al.; Bowditch & Dewey, LLP, 161 Worcester Road, P.O. Box 9320, Framingham, MA 01701-9320 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, 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, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

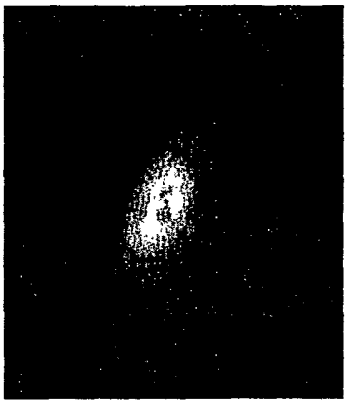
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, 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*

[Continued on next page]

(54) Title: NOVEL METHODS FOR INTRODUCING MOLECULES INTO CELLS AND VECTORS AND COMPOSITIONS FOR USE IN SUCH METHODS

A



B

(57) Abstract: Provided are chimeric molecules (i.e., antibody fusion or fusion protein) comprising a carboxy terminal protein import sequence and an animo terminal cargo region. A preferred chimeric molecule according to the invention is a chimeric molecule in which a protein import sequence (i.e., a HIV tat peptide) is linked to a Fab directed against a viral pathogen, such hCMV. Also provided are nucleic acid molecules encoding the chimeric molecules of the invention, as well as vectors and host cells comprising the nucleic acid molecules. In addition, methods of making, methods of using, and compositions comprising a chimeric molecule of the invention are provided. The vector constructs according to the invention can be employed for intracellular delivery, and particularly, for intracellular immunization.

WO 03/035892 A2

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.

NOVEL METHODS FOR INTRODUCING MOLECULES INTO CELLS AND VECTORS AND COMPOSITIONS FOR USE IN SUCH METHODS

GOVERNMENTAL RIGHTS

This invention was made with government support under Contract No. MDA 972-97-1-0002 by DARPA. The government has certain rights in the invention.

CROSS REFERENCES TO RELATED APPLICATIONS

The present application is a continuation of co–pending U.S. Patent Application No. 10/039,318, filed October 19, 2001. The entire contents of the above applications are incorporated herein by reference in entirety.

BACKGROUND OF THE INVENTION

Viral-encoded regulatory proteins and nucleic acids present attractive targets for antiviral therapy since interruption of their function should selectively impede (if not entirely abrogate) viral replication without affecting host cell machinery. Small molecules can penetrate cells to inhibit viral proteins or nucleic acid metabolism, but finding active small molecules with the required specificity is difficult, and drug resistance can be a problem. The inherent flexibility and diversity of the antibody system provides a wealth of molecules capable of recognizing any molecular shape with high specificity and affinity. Furthermore, the development of resistance by viral mutation can be countered by mutation in the appropriate antibody with minimal changes to pharmacokinetic properties. However, antibodies are extracellular molecules. A challenge thus remains to develop "invasive" antibodies, i.e., antibodies that are capable of entering cells and interrupting viral replication. In order for exogenously applied antiviral antibody to be active intracellularly, and thus, "invasive", it must be translocated with reasonable efficiency to the cytoplasm and nucleus of the host cell. This places further challenges on an antibody delivery system.

In particular, such intracellular immunity would prove useful with respect to human cytomegalovirus (hCMV), which is a double-stranded DNA virus of the herpes family. Generally, hCMV has low pathogenicity when infecting healthy

individuals. However, infection of the unborn and immunocompromised often leads to devastating consequences. Nearly 1% of all live births in the United States are associated with congenital hCMV infection, with approximately 10-20% of infections resulting in significant neurological defects. In immunosuppressed individuals, such as cancer and transplant patients, severe mononucleosis syndrome, pneumonia, and hepatitis are frequent complications with a mortality rate that can be as high as 40% (Fiala et al., *J. Infect. Dis.* 132, 421-433 (1975); Marker et al., *Surgery* 89, 660-671 (1981); Rubin et al., *Transplantation* 24, 458-464 (1977); Meyers et al., *Ann. Intern. Med.* 82,181-188 (1975)). Disseminated hCMV infection is also a common cause of morbidity and mortality in AIDS patients (Jacobson et al., *Ann. Intern. Med.* 108, 585-594 (1988)).

Drugs such as foscarnet and ganciclovir have been used for the treatment of hCMV infections (Collaborative DHPG Treatment Study Group, *N. Engl. J. Med.* 27, 801-805 (1986); Henderly et al., *Ophthalmology* 94, 425-434 (1987); Jacobson et al., *J. Infect. Dis.* 163, 1348-1351 (1991); Palestine et al., *Ann. Intern. Med.* 115, 665-673 (1991)). However, in immunocompromised patients, disease invariably recurs after cessation of treatment, necessitating long-term therapy (Gross et al., *Ophthalmology* 97, 681-686 (1990)). Drug toxicity and the emergence of resistant virus strains have limited the efficacy of these compounds and demonstrated the need for new drugs and treatment strategies (Collaborative DHPG Treatment Study Group, *supra;* Henderly et al., *supra;* Jacobson et al., *supra;* Palestine et al., *supra;* Gross et al., *supra*).

Intracellular immunity by using antibodies represents a novel alternative to currently available antiviral drugs. Inhibition of virus replication in cell culture by using gene transfer to express mouse antibodies inside human cells has recently been reported (Chen et al., *Proc. Natl. Acad. Sci. USA* 91, 5932-5936 (1994); Duan et al., *Proc. Natl. Acad. Sci.,* 91, 5075-5079 (1994)). However, there remains a need for a means by which antibodies, typically extracellular molecules, can be introduced inside of cells.

Work by Frankel et al. (*Cell*, 55, 1189-1193 (1988)) has shown that the tat protein of HIV-1 can enter a wide variety of cells efficiently when added exogenously. Tat protein added to medium at concentrations of 1 nM can induce biological responses in the nucleus of the target cell, although the mechanisms by

which the molecule is taken up and distributed within the cell are unknown. In addition, by chemically cross-linking tat peptides to a number of heterologous proteins otherwise impermeable to living cells, Fawell et al. (*Proc. Natl. Acad. Sci.*, 91, 664-668 (1994)) demonstrated delivery of these proteins to both the cytoplasm and nucleus. Of the different peptide constructs employed, the investigators found that the most successful was a peptide corresponding to tat residues 37-72 and 47-59 (Repinsky et al., *DNA and Cell Biol.*, 13, 1011-1019 (1994)).

Thus there remains a need for methods for introducing molecules into living cells, particularly for introducing antibody molecules into cells to effect intracellular immunization. The present invention is directed, amongst other things, to methods, vectors, and compositions to meet that need. These and other objects and advantages of the present invention, as well as additional inventive features, will be apparent from the description of the invention provided herein.

BRIEF SUMMARY OF THE INVENTION

The present invention provides a chimeric molecule comprising a carboxy terminal protein import sequence and an amino terminal cargo region. In a preferred embodiment, the present invention provides a protein import sequence, preferably an HIV tat protein import sequence, present as part of a chimeric molecule (e.g., an antibody fusion or fusion protein). Preferably, the present invention provides a chimeric molecule comprising a carboxy terminal protein import sequence and an amino terminal cargo region. The amino terminal cargo region can be an antibody, a non-antibody peptide, or a non-proteinaceous molecule. In one embodiment, the cargo region is an antibody raised against a viral antigen. Preferably, the cargo region is an antibody raised against an hCMV antigen, including, for example, an anti-UL44 antibody, an anti-UL83 antibody, and an anti-UL122 antibody. Also preferably, the cargo region is an antibody that is a Fab, including, for example, Fab GL14, Fab GL5, and Fab GL34.

The carboxy terminal protein import sequence is preferably an HIV tat peptide. Suitable HIV tat peptides include those comprising tat residues 37-72, tat residues 38-72, tat residues 47-59, and tat residues 49-57. Preferably, the protein import sequence is an HIV tat peptide comprising SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7. The protein import sequence can also be a lactoferrin or an antennapedia

peptide. A preferred antennapedia peptide is a *Drosophila* antennapedia peptide comprising from about 16 to about 20 amino acids.

In a preferred embodiment, the chimeric molecule comprises a protein import sequence linked to an antibody directed against a pathogen, such as a virus. In an especially preferred embodiment, the chimeric molecule comprises an HIV tat protein import sequence linked to a Fab directed against an hCMV antigen.

The chimeric molecule can further comprise a spacer sequence located between said protein import sequence and said cargo region. Preferably, the spacer sequence comprises from about 1 to about 50 amino acid residues.

The present invention also provides a composition comprising a chimeric molecule having a carboxy terminal protein import sequence and an amino terminal cargo region. In other preferred embodiments, the present invention provides for the use of the chimeric molecule of the present invention for the manufacture of a medicament for the treatment of viral infection.

In addition, the present invention provides a method of using any of the above-described chimeric molecules to effect entry into a cell by contacting the cell with a chimeric molecule such that the chimeric molecule is internalized.

The present invention further relates to nucleic acid molecules that encode any of the above-described chimeric molecules. In one embodiment, the nucleic acid encodes a chimeric molecule having a carboxy terminal protein import sequence and an amino terminal cargo region. Preferably, the nucleic acid encodes a chimeric molecule having an HIV tat protein import sequence linked to a Fab directed against an hCMV antigen. In another embodiment, the nucleic acid encodes a chimeric molecule that further has a spacer sequence between the protein import sequence and the cargo region.

In addition, the present invention relates to vectors and host cells comprising a nucleic acid that encodes any of the described chimeric molecules. Thus, in one embodiment, the invention provides a vector comprising a nucleic acid that encodes a chimeric molecule having a carboxy terminal protein import sequence and an amino terminal cargo region. Preferably, the vector comprises a nucleic acid that encodes a chimeric molecule having an HIV tat protein import sequence linked to a Fab directed against an hCMV antigen. In another embodiment, the vector comprises a nucleic acid that encodes a chimeric molecule that further has a spacer sequence between the protein import sequence and the cargo region. In addition, the present invention

provides a host comprising any of the described vectors, as well as a host cell comprising a chimeric molecule having a carboxy terminal protein import sequence and an amino terminal cargo region with or without a spacer sequence.

The present invention also relates to a method for making a chimeric molecule comprising a carboxy terminal protein import sequence and an amino terminal cargo region. The method comprises obtaining a host cell comprising a vector containing a nucleic acid that encodes a chimeric molecule having a carboxy terminal protein import sequence and an amino terminal cargo region, culturing the host cell, and recovering the chimeric molecule from the culture.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a graph of selective binding per cell density of a tagged Fab according to the invention relative to an untagged Fab.

Fig. 2A and Fig. 2B shows the uptake of FITC-conjugated tat-Fab by direct visualization of living, healthy glial U373 cells imaged at about 20 minutes post-treatment with tat-tagged Fab showing internalization of the FITC-conjugated tat-Fab. Fig. 2A, phase contrast image; Fig. 2B, fluorescence image.

Fig. 3 is a reproduction of a Coomassie-stained gel showing bacterially expressed wild-type IE86 (250 ng) protein following dialysis of the pooled fractions after nickel chelate chromatography.

Fig. 4 is a reproduction of the results of a DNase I protection analysis assay of the IE86 binding site (crs) in the hCMV major immediate early promoter by recombinant IE86 protein. *Lanes:* 1 and 4, controls showing cleavage patterns in the presence of 250 nM BSA; 2 and 3, footprinting pattern in the presence of 50 and 250 nM IE86 protein, respectively. *Symbols:* arrow, transcription start site; open box, the protected crs.

Fig. 5 is a reproduction of the results of an electrophoretic mobility shift assay (EMSA) revealing protein:DNA and protein:protein interactions. *Lanes:* 1, a nucleoprotein complex ("D") developed with the CMV TATA box probe and human TBP; 2, supershifting of complex D using a monoclonal antibody to hTBP; 3, a change in the mobility of the complex ("D86") due to protein:protein interactions of IE86 with TATA box bound hTBP; 4, supershifting of the D86 complex by monoclonal antibodies to hTBP; 5, supershifting of the D86 complex by monoclonal antibodies to IE86.

Fig. 6 is a chart that summarizes the activity of recombinant IE86 in a highly purified reconstituted *in vitro* transcription system, with the requirement for general transcription factors (indicated by a + sign) for initiating transcription from DNA templates containing the major immediate early promoter of hCMV (CMV-MIEP) and the adenovirus major late promoter (Ad-MLP).

Fig. 7 is a reproduction of a gel obtained from the reconstituted *in vitro* transcription system described in Fig. 6. *Lanes*: 1-3, transcription from the wild-type MIEP in the presence of increasing amounts of IE86 (Lanes 2 and 3 showing repression of transcription); 4-6, transcription reactions with a MIEP template in which the IE86 binding site (crs) has been mutated (Lanes 5 and 6 reveal the requirement for IE86 to interact with the crs element in order to effect inhibition); 7-9, transcription reactions using a mutant IE86 protein that is defective in DNA-binding and is also incapable of effecting inhibition of transcription from the wild-type MIEP template.

Figs. 8A-C are reproductions of the results of Western analysis of column fractions (concentrations of KCl used to elute UL44 from the column(s) indicated above the panels, and the location of molecular weight standards used for Western blotting shown on the right of each panel). *Panels:* Panel A, elution profile of purified UL44;. Panel B, elution profile of UL44 was subjected to thrombin cleavage; elution profile of UL44 after being incubated with an excess of a recombinant Fab (GL 5) that recognizes UL44.

Fig. 9 is a diagrammatic representation of the three thrombin cleavage sites in UL44.

Fig. 10 is a reproduction of autoradiograms produced in an electrophoresis mobility shift assay (EMSA) showing the interaction between Fab SD1 and DNA. *Lanes*: 0.08 pmol ^{32}P-labeled double-stranded probe (5'-AAT-GTA-TGC-GCG-CGC-GCT-TTA-GGG-GCC-CC-3' [SEQ ID NO:18]) (*lanes 1-6*) with 0.2 pmol of Fab SD1 (*lanes 2, 3,* and *4*), further including preincubation with an anti-Fab reagent (α-Fab IgG) (*lanes 3* and *6*), or further including 0.2 pmol of a control Fab reacting with HIV-1 surface glycoprotein gp120 (HIV Fab). *Symbols*: thick arrow, position of the specific Fab SD1 nucleoprotein complex; α_s arrow, position of the supershifted Fab SD1 nucleoprotein complex; 0, origin of migration; FP, position of free probe.

Fig. 11 is a Hofstee plot (% displaced/nm competitor versus % displaced) for

competition assays performed with double stranded MLP and MyoD as competitors.

Fig. 12 is a double reciprocal plot (1/free DNA versus 1/bound DNA) for competition assays performed with double-stranded MLP as the competitor.

Fig. 13 is a schematic representation of the protection pattern produced by TFEB binding [SEQ ID NO: 19]. *Symbols*: arrows, indicate the positions of DNase I hypersensitive sites; double-headed arrow, the E box, CACGTG.

Fig. 14 is a schematic representation of the protection pattern produced by Fab-E box binding [SEQ ID NO: 19]. *Symbols*: arrows, indicate the positions of DNase I hypersensitive sites; double-headed arrow, the E box, CACGTG.

Fig. 15 is a depiction of a consensus binding site for Fab-Ebox. *Abbreviation:* Pu, purine residue (A or G).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The invention provides, among other things, methods of making, methods of using, and compositions comprising a known protein import sequence present as part of a chimeric molecule (e.g., an antibody fusion or fusion protein). According to the invention, a chimeric molecule comprises a carboxy terminal protein import sequence and an amino terminal cargo region, as further described herein. In particular, the invention provides a method to import molecules into cells using such known protein import sequences present as part of a chimeric molecule. The invention further provides vectors, host cells, and compositions comprising a known protein import sequence.

In one embodiment, the present invention provides a method of selectively introducing a molecule into a cell with use of a human immunodeficiency (HIV) tat protein import sequence. Preferably this method is carried out in order to effect transfer into a cell of an exogenous moiety, i.e., "cargo" comprising a moiety such as an antibody (or portion thereof), a non-antibody polypeptide/protein, a nonproteinaceous molecule (e.g., oligosaccharides, oligonucloetides), or another suitable moiety, as further described herein. As used herein "antibody" means a complete antibody molecule or portions thereof, including Fab, Fab', $F(ab')_2$, and F(v) fragments as well as other portions known in the art.

In another embodiment, the invention further provides for the development of so-called "invasive antibodies", which preferably are antibodies directed against pathogens (e.g., viral antigens), and which are endowed with the ability to be taken up

from outside the cell and thereby effect intracellular immunity (e.g., against persistent/latent infections). An invasive antibody according to the invention need not necessarily effect complete immunity, but need only impair in some fashion the biological activity of the pathogen against which the antibody is directed such that there is a statistically significant reduction in biological activity (as compared to the untreated condition), as measured by any means acceptable to those working in the art area and as appropriate for a particular application (e.g., a reduction in DNA and/or RNA metabolism of the pathogen, a reduction in protein synthesis of the pathogen, inhibition of DNA binding function and activation domains, inhibition of DNA and polymerase binding, screening by microinjection, inhibition of viral replication, effect on infected cell growth and/or viability, etc., most of which are further exemplified herein).

Intracellular immunity by invasive antibodies presents a novel alternative, for instance, to currently available antiviral drugs. Intracellular antibody-based therapy optimally combines antibody specificity and selective inhibitory action with the pharmokinetic stability and administrative advantages of antibodies. The use of intracellular immunity by 'invasive' antibodies is exemplified as follows for human cytomegalovirus (hCMV), an important human pathogen. However, the concept is broadly applicable to any other immunizations, e.g., against other viruses or other pathogens, and is expressly so contemplated. The viral-encoded regulatory proteins and nucleic acids of hCMV (or other pathogen) offer attractive targets for intracellular immunity since interruption of their function ostensibly prevents viral replication without affecting host cell machinery.

Intracellular Protein Import Sequences

A number of naturally occurring proteins (or portions thereof) are effectively taken up by cells and are delivered to the nucleus, or other intracellular compartment, for subsequent intracellular localization. Thus, desirably, a protein "import sequence" or an "intracellular import sequence" according to the invention preferably is a known protein import sequence (or variation thereof) that is capable of mediating, either by commanding or facilitating, entry into a cell. Accordingly, a protein import sequence includes, but is not limited to a plasma membrane sequence, nuclear localization sequence, endoplasmic reticulum (ER) retention sequence, or a signal sequence, to name but a few. However, the protein import sequence

according to the invention also desirably allows merely for entry into a cell, and does not provide for delivery to a specific intracellular localization. It especially is preferable that a protein import sequence according to the invention is a sequence that is capable of recognizing a cell surface binding site and being internalized by receptor-mediated endocytosis upon binding at the cell surface.

Of course, the protein import sequence of the invention optionally can be present either in its peptide/protein form (e.g., in the chimeric molecule), or, in the form of its encoding nucleic acids (e.g., in certain of the vectors). To effect cell entry, the protein import sequence must be present in its peptide/protein form. The nucleic acid form of the protein import sequence is referred to herein as the "import coding sequence".

The use of the import systems of proteins in the method of the invention is highly advantageous, since these systems provide solutions to many of the problems that are likely to be encountered in designing invasive human antibodies, or in effecting cell entry of other chimeric molecules. One key feature of the protein import sequences employed according to the invention is that they preferably avoid the lysosomal compartment inside the cell, unlike a wide variety of alternative methods for the delivery of proteins into living cells (see, e.g., Prior et al., *Biochemistry* 31, 3555-3559 (1992); Prior et al., *Cell* 64, 1017-1023 (1991); Stenmark et al., *J. Cell Biol.* 113,1025-1032 (1991); Ishihara et al., *Pharm. Res.* 7, 542-546 (1990); Basu et al., *Biochem. Pharmacol.* 40, 1941-1946 (1990); Wu. et al., *Biochemistry* 27, 8867-892 (1988); Wilson et al., *J. Biol. Chem.* 267, 963-967 (1992); Leamon et al., *J. Biol. Chem.* 267, 24966-24971 (1992); Leamon et al., *Proc. Natl. Acad. Sci.* 88, 5572-5576 (1991)).

As an example of a preferred protein import sequence of the invention, the *Drosophila* transcription factor antennapedia can pass from a cell's exterior through to the nucleus, where this factor binds to regulatory DNA sequences and alters gene expression. Recently, it has been shown that antennapedia peptides ranging in length from between 16 and 20 amino acids can cross biological membranes by an energy-independent mechanism and are not targeted to the lysosomal compartment (Derossi et al., *J. Biolog. Chem.* 269, 10444-10450 (1994)). Thus, these *Drosophila* antennapedia peptides (particularly antennapedia peptides ranging in length from between about 16 and about 20 amino acids) desirably can be chemically and genetically linked to desired "cargo" (e.g., antibodies, non-antibody proteins, and

non-proteinaceous molecules including oligonucleotides, to name but a few), and employed for their ability to be taken up by cells.

Similarly, the lactoferrin protein can make its way across the plasma membrane of cells, presumably by binding to receptors or by interacting with sugar moieties on the cell surface. Lactoferrin plays a key role in the body's first line of defense against microbial invasion by sopping up extracellular oxidized iron. However, recent studies have indicated that lactoferrin also may play a more active role in the immune response to bacterial invasion by turning on macrophages and natural killer cells that are a part of the next wave of defense (Baeuerle, *Nature* 373, 661-662 (1995)). Once inside the cells, lactoferrin is transported to the nucleus where it can bind specific DNA sequences and may selectively alter gene expression (He et al., *Nature* 373, 721-724 (1995)). Accordingly, lactoferrin and antennapedia peptides, as well as other protein import sequences, desirably can be employed in the methods and constructs of the invention for effecting cell entry.

Viruses also encode proteins that can be taken up from outside the cell and delivered to the nucleus, or other intracellular locations. In particular, the tat protein from HIV is efficiently taken up by cells and, by avoiding the lysosomal compartment, can enter the nucleus. Thus, an especially preferred strategy according to the invention is to use the HIV tat protein import sequence attached to a Fab fragment (or, attached so as to comprise fusion proteins, or other chimeric molecules of the invention) to provide for tat protein import sequence-mediated delivery into cells and/or intracellular targeting (e.g., to the nucleus). For instance, in the case of a Fab fragment (i.e., a Fab), preferably a tat-containing construct is obtained by linkage of the tat protein import sequence to the C-terminus of the light chain using the vectors and methods further described herein.

In terms of protein import sequences derived from the HIV tat gene, the HIV tat gene encodes an 86-102 residue protein, depending on virus strain. The protein is characterized by a cysteine-rich region in its amino terminal portion, and by a cluster of highly basic residues in the carboxy terminal half of the molecule that are well conserved across the many strains of the virus. Mutations directed to the amino terminal half of the molecule confer a protein that is no longer able to transactivate HIV long terminal repeat-directed gene expression, although these molecules continue to be localized in the nucleus. In contrast, however, mutations directed at the cluster of basic residues, spanning amino acid residues 47-59, greatly reduce

nuclear accumulation of the protein (Ruben et al., *J. Virol.* 63, 1-8 (1989)).

This highly basic sequence is similar to nuclear localization sequences previously identified in other proteins (Burglin et al., *EMBO J.* 6, 2617-2625 (1987); Kalderon et al., *Cell* 39, 499-509 (1984)). Indeed, the amino acid sequence GRKKRRQRRRAHQ [SEQ ID NO:1] spanning residues 48-61, when fused to the amino terminus of ß-galactosidase and expressed in the cytoplasm, was able to direct the fusion to the nucleus of the cell (Siomi et al., *J. Virol.* 64, 1803-1807 (1990)). Using a similar system, other investigators further delineated the role of the basic amino acid sequences by fusing sequences GRKKR [SEQ ID NO:2] and QRRP [SEQ ID NO:3] to the amino terminus of ß-galactosidase for cytoplasmic expression (Ruben et al., *J. Virol.* 63, 1-8 (1989)). Whereas the tag GRKKR delivered the ß-galactosidase to the nucleus, the tag QRRP did not, indicating that the GRKKR sequence or part of this sequence can function as a nuclear localization signal.

This region of basic sequence between residues 47 and 59 also has been implicated in tat adhesion to the cell surface. Namely, a synthetic peptide composed of amino acid residues 49-57 was able to competitively inhibit cell attachment to both full length tat peptide and to recombinant tat protein, despite lacking the tat RGD sequence at residues 72-74 (Weeks et al., *J. Biological Chem.* 268, 5279-5284 (1993)). Similarly, a monoclonal antibody recognizing an epitope in the basic region of tat was able to inhibit tat binding and uptake (Mann et al., *EMBO J.* 10, 1733-1739 (1991)).

The studies suggest in the context of the invention that the basic region between residues 47-59 is an essential component of the binding, uptake, and nuclear delivery of the tat molecule, and suggest its potential to transport other molecules (i.e., so-called other cargo proteins and molecules from the medium) into the cell nucleus. When tat peptides were assessed for their ability to act as carrier proteins for the delivery of a number of heterologous proteins into cells, a peptide spanning residues 37-72 was consistently found to be successful (Fawell et al., *Proc. Natl. Acad. Sci. USA* 91, 664-668 (1994)). More recently, a peptide spanning residues 47-58 was shown to be efficacious for introducing cargo proteins (Pepinsky et al., *DNA and Cell Biol.* 13, 1011-1019 (1994)). In this study, specific inhibition of human papillomavirus E2 transactivation by intracellular delivery of a dominant negative E2 protein was demonstrated.

According to the invention, HIV tat residues 47-59 preferably comprise a core sequence for delivering an antibody, an antibody fragment or attached protein sequence when the tat residues form a fusion protein with the attached sequence to the cell interior (e.g., to the cell nucleus), and which can be manipulated by mutagenesis as further described herein in order to improve upon its function. It is advantageous that this preferred segment of the HIV tat protein according to the invention is incapable of activating the HIV-LTR, and is non-toxic to the cell (Fawell et al., *Proc. Natl. Acad. Sci. USA* 91, 664-668 (1994); Repinsky et al., *DNA and Cell Biol.* 13, 1011-1019 (1994)). Thus, a particularly preferred protein import sequence according to the invention is the tat protein import sequence, especially that of human immunodeficiency virus (HIV). The HIV protein import sequence desirably can be obtained from any strain of HIV (e.g., HIV-1, or HIV-2).

The protein import sequence can be obtained by any appropriate means, e.g., produced using genetic engineering techniques, chemically synthesized, or purchased from a commercial source. Also, desirably the tat protein import sequence can be from a virus other than HIV, such as equine infectious anemia virus (e.g., Carroll et al., *J. Virol.*, 65, 3460-67 (1991)), and simian immunodeficiency virus (e.g., Chakrabarti et al., *Nature*, 328, 543-47 (1987); Arya et al., *Nature*, 328 548-550 (1987). The tat protein import sequence generally is understood as comprising the tat basic region (i.e., residues 49-57). When a tat protein import sequence from a species other than HIV is employed according to the invention, then desirably the region of the sequence corresponding to that described herein for the HIV tat protein import sequence is employed.

An especially preferred tat protein import sequence according to the invention is a sequence selected from the group consisting of an HIV tat peptide comprising tat residues 37-72, CFITKALGISYGRKKRRQRRRPPQFSQTHQVSLSKQ (i.e., Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr His Gln Val Ser Leu Ser Lys Gln [SEQ ID NO:4]), an HIV tat peptide comprising comprising tat residues 38-72, FITKALGISYGRKKRRQRRRPPQFSQTHQVSLSKQ (i.e., Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr His Gln Val Ser Leu Ser Lys Gln [SEQ ID NO:5]), an HIV tat peptide comprising tat residues 47-59, YGRKKRRQRRRPP (i.e., Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro [SEQ ID NO:6]), and an HIV tat peptide comprising tat residues 49-

57, RKKRRQRRR (i.e., Arg Lys Lys Arg Arg Gln Arg Arg Arg [SEQ ID NO:7]). Of these different HIV tat peptide sequences that can be employed according to the invention, those comprising a tat protein import sequence corresponding to residues 38-72 [SEQ ID NO:5] and 47-59 [SEQ ID NO:6] are most preferred.

The conventional abbreviations for amino acids comprising proteins and peptides are used herein as generally accepted in the peptide art and as recommended by the IUPAC-IUB Commission on Biochemical Nomenclature (See, *European J. Biochem.*, 138, 9-37 (1984)). Similarly, protein and peptide sequences are written according to the standard convention wherein the N-terminal amino acid is on the left and the C-terminal amino acid is on the right (with corresponding nucleic acid sequences being written in a 5' to 3' direction). The term "peptide" as used herein refers to any length molecular chain of amino acids linked by peptide bonds, so long as the length of the peptide is less than that of a corresponding full-length protein of which the peptide forms a part. As used herein, "protein" refers to a full length (i.e., complete) protein having both amino and carboxy terminal ends. The peptides and protein of the invention desirably can be modified such as is known in the art.

The proteins of the present invention preferably comprise an amino end and a carboxyl end. Proteins with a modified amino- and/or carboxyl-terminus also desirably can be employed according to the invention. Such proteins and peptides with modified termini are expected to have longer *in vivo* half-lives because endopeptidases have reduced activity with respect to proteins and peptides with modified termini. The peptides and proteins can comprise D- or L- peptides, or a mixture of the D- and L-amino acid forms. Proteins and peptides comprising L-amino acids are preferred. However, the D-form of the amino acids are also desirable since proteins and peptides comprising D-amino acids are expected to have a greater retention of their biological activity *in vivo* given that the D-amino acids are not recognized by naturally occurring proteases. Of course, proteins comprising the D-form of amino acids, unlike those comprising the L-form of amino acids, cannot be synthesized in a living cell, and must be produced by chemical means.

Any appropriate expression vector (e.g., as described in Pouwels et al., *Cloning Vectors: A Laboratory Manual* (Elsevior, N.Y.: 1985)) and corresponding suitable host can be employed for production of peptides/proteins according to the invention. Expression hosts include, but are not limited to, bacterial species within the genera Escherichia, Bacillus, Pseudomonas, Salmonella, mammalian or insect

host cell systems including baculovirus systems (e.g., as described by Luckow et al., *Bio/Technology,* 6, 47 (1988)), and established cell lines such as the COS-7, C127, 3T3, CHO, HeLa, BHK cell line, and the like. The ordinary skilled artisan is, of course, aware that the choice of expression host has ramifications for the type of peptide/protein produced. For instance the glycosylation of peptides produced in yeast or mammalian cells (e.g., COS-7 cells) will differ from that of peptides produced in bacterial cells such as *Escherichia coli.*

Alternately, the protein import sequences of the invention, as well as other peptide/protein sequences, can be synthesized using standard peptide synthesizing techniques well known to those of skill in the art (e.g., as summarized in Bodanszky, *Principles of Peptide Synthesis,* (Springer-Verlag, Heidelberg: 1984)). In particular, the peptides can be synthesized using the procedure of solid-phase synthesis (see, e.g., Merrifield, *J. Am. Chem. Soc.,* 85, 2149-54 (1963); Barany et al., *Int. J. Peptide Protein Res.,* 30, 705-739 (1987); and U.S. Pat. No. 5,424,398). If desired, this can be done using an automated peptide synthesizer. Removal of the t-butyloxycarbonyl (t-BOC) or 9-fluorenylmethyloxycarbonyl (Fmoc) amino acid blocking groups and separation of the peptide from the resin can be accomplished by, for example, acid treatment at reduced temperature. The peptide-containing mixture can then be extracted, for instance, with dimethyl ether, to remove non-peptide organic compounds, and the synthesized peptides can be extracted from the resin powder (e.g., with about 25% w/v acetic acid). Following the synthesis of the peptide, further purification (e.g., using high performance liquid chromatography (HPLC)) optionally can be done in order to eliminate any incomplete peptides or free amino acids. Amino acid and/or HPLC analysis can be performed on the synthesized peptides to validate the identity of the peptide.

If desired, either the peptides or the proteins of the invention can be modified, for instance, by glycosylation, amidation, carboxylation, or phosphorylation, or by the creation of acid addition salts, amides, esters, in particular C-terminal esters, and N-acyl derivatives of the peptides of the invention. The peptides also can be modified to create peptide derivatives by forming covalent or noncovalent complexes with other moieties. Covalently bound complexes can be prepared by linking the chemical moieties to functional groups on the side chains of amino acids comprising the peptides, or at the N- or C-terminus. Such modifications can be particularly useful, for instance, in constructing bispecific molecules having a

ligand to a cell surface receptor attached to an antibody. Further modifications will be apparent to those of ordinary skill in the art.

A Chimeric Molecule

A suitable chimeric molecule according to the invention comprises a carboxy terminal protein import sequence and an amino terminal cargo region. Alternately, a suitable chimeric molecule desirably consists essentially of (or consists of) a carboxy terminal protein import sequence and an amino terminal cargo region.

Optionally, a chimeric molecule further comprises a spacer sequence. Such a spacer sequence optimally is located between the protein import sequence and the cargo region. Preferably, a spacer sequence comprises from about 1 to about 50 amino acid residues, and especially from about 1 to about 10 amino acids residues. The spacer sequence according to the invention must not interfere with the functionality of the remainder of the molecule (i.e., considered *in toto*, or considered in terms of its amino- and carboxy-terminal functionalities). Desirably the spacer sequence comprises glycine residues, proline residues, or a mixture of glycine residues and proline residues. Alternately, desirably the spacer sequence includes charged residues to assist with water solubility of the structure in which the spacer sequence is contained, and to minimally interfere with the amino and carboxy functionalities. Preferred residues include Arg, Asp, Glu, Ser, and Thr. Especially preferred residues are Arg, Asp, and Glu. Optimally, the spacer does not comprise Cys residues.

When a chimeric molecule is produced by means of a nucleic acid (e.g., transcribed and translated *in vitro*), then desirably the portion of the nucleic acid molecule that encodes the spacer sequence present in the chimeric molecule according to the invention (i.e., the "spacer coding sequence") is an amount of DNA that encodes from about 1 to about 50 amino acid residues in the resultant chimeric molecule, e.g., in the resultant polypeptide or protein. Preferably this means that the spacer coding sequence can span from about 3 to about 1000 base pairs. The spacer coding sequence can be longer than the number of residues encoding the amino acid sequence because the spacer coding sequence according to the invention can encompass an even number (e.g., particularly such as two, four, or six) of splice sites. Thus, some DNA present in the spacer coding sequence preferably may not contribute to the resultant chimeric molecule due to such splicing prior to translation

into nascent peptide/protein. The sequence of the spacer coding sequence optimally encodes a span of amino acid residues that does not interfere with the functionality of the remainder of the chimeric molecule of the invention.

The chimeric molecule according to the invention can be prepared by any of a number of conventional techniques. For instance, in the case of recombinant peptides, a DNA fragment encoding a desired peptide can be subcloned into an appropriate vector using well-known molecular genetic techniques (see, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual,* 2nd ed. (Cold Spring Harbor Laboratory, 1989)). The DNA fragment can be transcribed and the peptide subsequently translated *in vitro*. Commercially available kits can also be employed (e.g., such as manufactured by Clontech, Palo Alto, Calif.; Amersham Life Sciences, Inc., Arlington Heights, Ill.; InVitrogen, San Diego, Calif., and the like). The polymerase chain reaction optionally can be employed in manipulation of nucleic acids.

Alternatively, synthetic means (i.e., chemical synthesis) can be employed to obtain the peptide/protein and/or to synthesize the nucleic acid (e.g., that is used to obtain the peptide/protein). These are all standard experiments that are easily done by and well known to one skilled in the art. Automated equipment for polypeptide or DNA synthesis is commercially available. Host cells, cloning vectors, DNA expression controlling sequences, oligonucleotide linkers, and other reagents and components are also commercially available.

In particular, a chimeric molecule can comprise non-native protein/peptide components, such as a non-native protein import sequence or a non-native cargo region (i.e., when the cargo region is a protein/peptide), or any mixture of native or non-native sequences. A protein/peptide as isolated from or observed in nature, in the absence of any human intervention regarding its composition, is a "native" protein/peptide. A "non-native" protein/peptide is one that preferably comprises amino acid residues, or a sequence of amino acids, that are not found in the protein/peptide sequence as isolated from or observed in nature, and that generally results from human intervention regarding the composition of the protein sequence (i.e., either by manipulating the nucleic acid or amino acid sequence and/or their means of synthesis).

Preferably a non-native peptide/protein of the invention (e.g., a non-native protein import sequence and/or a non-native cargo region in the case where the cargo

region is a peptide/protein) differs from its native counterpart sequence in that it comprises one or more replacements (i.e., substitutions) in the amino acid sequence, or one or more additions or deletions in the amino acid sequence. In the case of a non-native protein import sequence, desirably such an addition, deletion, or replacement is made such that the protein import sequence comprises a binding domain, i.e., an attachment sequence or an epitope for an antibody that optimally differs from (i.e., comprises a different sequence than) the binding domain of a native protein import sequence. According to the invention, however, optimally the functionality of the non-native protein import sequence is about the same as, if not better than, that of the native protein import sequence. In particular, it is contemplated that the "RGD" peptide motif present in a tat protein import sequence according to the invention can be modified (e.g., desirably by either deletion or replacement).

Cargo Region

The chimeric molecule of the invention desirably contains an amino terminal cargo region. This cargo region preferably comprises a molecule or portion thereof, especially a protein/peptide molecule (or a modified protein/peptide) that is to be delivered intracellularly. Preferably the cargo region is covalently attached to the carboxy terminal protein import sequence, by means of a spacer sequence that comprises from about 1 to about 50 amino acid residues. Desirably, the cargo region comprises a moiety selected from the group consisting of antibodies (i.e., full length antibodies or portions thereof as previously described), peptides (i.e., non-antibody peptides which together with the protein import sequence comprise a fusion protein), and non-proteinaceous molecules (including, but not limited to, carbohydrates, DNA intercalating agents, spin probes, fluorescent indicators of intracellular pH, environmentally-sensitive fluorophores, and fluorescent indicators for specific ligands). Preferably, as further described below, the cargo region comprises an antibody (or portion thereof), such as can be employed in establishing intracellular immunity. When the cargo region comprises a peptide (i.e., either a non-antibody peptide or an antibody), desirably this can be further modified by one or more means that, for instance, allow for detection of the resultant chimeric molecule. In this instance, preferred means of modification include incorporation of a suitable

fluorophore, particularly fluorescein, substituted fluoroscein, and green fluorescent protein.

Desirably, a cargo region comprises either an antibody (or portion thereof) or a peptide/protein, which are appended to the protein import sequence so as to comprise, respectively, an antibody fusion or a fusion protein. In particular, according to the invention, a cargo region does not comprise a protein import sequence as described herein. In an especially preferred embodiment, the cargo region comprises an antibody (or portion thereof) so as to together with the protein import sequence (and optional spacer) comprise a so-called invasive antibody.

Antigens for Intracellular Immunization

The methods and products of the invention desirably are employed to effect intracellular immunization against particular pathogens through interaction with (e.g., interference with the function of) their antigens. Preferred antigens for intracellular immunization according to the invention are hCMV antigens, especially gene products encoded within the major immediate-early (IE) transcriptional unit and the UL44 gene product of hCMV (Hermiston et al., *J. Virol.* 61, 3214-3221 (1987); Pizzorno et al., *J. Virol.* 62, 1167-1179 (1988); Stenberg et al., *J. Virol.* 63, 2699-2708 (1989); Stenberg et al., *J. Virol.* 64, 1556-1565 (1990)), as further described below.

The IE region encodes several proteins responsible for regulation of virus gene expression and subsequent viral replication (Azad et al., *Antimicrobial Agents and Chemo.* 37, 1945-1954 (1993); Pari et al., *J. Virol.* 67, 2575-2582 (1993)). The products of this transcriptional unit have been shown by using specific antisense oligonucleotides to be essential for production of infectious hCMV (Azad et al., *Antimicrobial Agents and Chemo.* 37, 1945-1954 (1993)). Oligonucleotides complementary to RNA encoding for IE86 (IE2, UL122) exhibit the most potent antiviral activity (Azad et al., *Antimicrobial Agents and Chemo.* 37, 1945-1954 (1993)).

The UL44 gene product is a phosphorylated protein of about 52kDa, which is one of the most immunogenic proteins during natural infection with hCMV (Landini et al., *J. Clin. Microbiol.* 27, 2324-2327 (1989)). This protein is associated with the viral DNA polymerase (UL54). Antisense RNA strategies have demonstrated an essential role for the UL44 proteins in hCMV replication (Ripalti et al., *J. Virol.* 69,

2047-2057 (1995)).

Other hCMV antigens against which the invention desirably can be employed for effecting intracellular immunization include, but are not limited to, hCMV DNA, hCMV RNA, UL123-IE72, UL122-IE86, UL122-IE55, UL36-38, US3, TRSI-IRSI, UL54, UL57, UL105, UL70, UL-83, and UL44, wherein the aforementioned names refer to open reading frame (ORF) designations based on the genome sequencing project (Chee et al., *Current Topics in Microbiology and Immunology*. 154, 125-169 (1990)).

Also, preferably the invention can be employed to effect intracellular immunization against other antigens, including antigens other than hCMV antigens, and including antigens other than viral antigens.

Antibodies for Intracellular Immunization

As a means of effecting intracellular immunization according to the invention, preferably antibodies are obtained that specifically recognize and interact with the antigens against which immunity is being effected. Generally, such invasive antibodies desirably are obtained by first obtaining antibodies against the preferred antigens (e.g., antibodies which disrupt the functioning of these antigens), and then rendering the antibodies invasive by engineering the antibodies to permit their intracellular delivery, and particularly, to permit their cytoplasmic and/or nuclear targeting. This modification of the antibodies according to the invention desirably is accomplished by attaching to the antibody a protein import sequence (e.g., an HIV tat protein import sequence), as previously described.

Desirably, such an antibody "tagged" in this fashion with an intracellular protein import sequence exhibits an affinity for its target molecules that is unaltered. This is accomplished because, according to the invention, the protein import sequence is introduced into a region of the antibody sequence that is not responsible for epitope recognition. Generally, epitope recognition is confined to the N-terminus of an antibody or Fab fragment. Thus, for instance, desirably an HIV-2 tat protein import sequence (or another preferred protein import sequence according to the invention) is attached to the C-terminus of an antibody. This differs from other uses of tat for intracellular delivery which have been described, where tat is located at the amino terminus of a chimeric molecule (e.g., U.S. Patents 5,652,122, 5,674,980, PCT

International application WO 94/04686, and others), which requires in certain instances that a methionine residue be added in the construction.

An antibody for intracellular immunization according to the invention desirably includes, but is not limited to, immunoglobulin molecules and immunologically active portions of immunoglobulin molecules such as portions containing a paratope (i.e., an antigen binding site), such that the antibody comprises, for example, either intact immunoglobulin molecules or portions thereof, such as those known in the art as Fab (**F**ragment **a**ntigen **b**inding), Fab′, $F(ab')_2$ and F(v) fragments or a modified version thereof (e.g., one that contains a DNA recognition sequence). The antibody can be, for example, a monoclonal antibody, a polyclonal antibody, a single-chain antibody (e.g., that further can comprise a ligand or attachment sequence in addition to a paratope), and a bispecific antibody (e.g., that can have two of the same paratope, or two different paratopes, e.g., one paratope directed to a viral antigen epitope, and another paratope directed to another viral antigen epitope, or otherwise directed). Preferably the antibody according to the invention is a Fab. Some examples of chimeric molecules that include an antibody component are set in the following Examples. However, other antibody constructions that are known in the art can be employed in the invention and are contemplated herein.

Desirably an antibody is one that disrupts the viral functions of human cytomegalovirus (hCMV), e.g., by interfering with hCMV regulatory proteins. Especially preferred are antibodies capable of interfering with immediate-early protein function, and antibodies that impair viral nucleic acid metabolism (e.g., by selectively binding to viral nucleic acids). Particularly preferred antibodies according to the invention are those directed against the hCMV polymerase-associated protein UL44, the hCMV open reading frame UL122-IE86 (i.e., "IE86"), or the 65 kDa lower matrix phosphoprotein UL83, including, but not limited to, the antibodies Fab GL5, Fab GL14, and Fab GL34.

The antibody can be produced by any suitable technique, e.g., by conventional techniques for preparing monoclonal, polyclonal, single-chain, and bispecific antibodies, as well as more current recombinant DNA techniques that are familiar to those skilled in the art. In particular, preferably an antibody (e.g., a Fab) is produced as described in the further description and Examples, which follow. Chimeric molecules having a ligand component linked to an immunoglobulin constant region, and other immunoconjugates such as bispecific antibodies,

preferably can be made as described, for instance, in U.S. Pat. Nos. 4,816,567, 5,349,053, 5,332,567, and 5,443,953, and PCT International Applications WO 90/14424, WO 91/05805, WO 91/05871, WO 92/02553, and WO 95/16037; and Cook et al., *J. Immunol. Methods,* 171, 227-237 (1994). In particular, bispecific antibodies can be made by a variety of means, e.g., chemical techniques (see, e.g., Kranz et al., *Proc. Natl. Acad. Sci.,* 78, 5807 (1981)), for instance, disulfide cleavage reformation of whole IgG or, preferably, $F(ab')_2$ fragments; fusions of more than one clone to form polyomas that produce immunoglobulins having more than one specificity (see, e.g., Segal et al., *In Current Protocols in Immunology,* Coligan et al. (eds.), vol. 1, 2.13.1-2.13.16 (John Wiley & Sons, Inc. (1995))); or by genetic engineering (see, e.g., U.S. Pat. No. 4,816,567 and PCT International Patent Application WO 90/14424).

Considerations Regarding Efficacy of Intracellular Antibodies for Effecting Intracellular Immunity

There are multiple preferred embodiments of the method of using a chimeric molecule of the present invention. For instance, the method preferably is carried out to introduce a protein import sequence-containing construct (e.g., a tat protein import sequence containing construct) into any cell, even a cell that the protein import sequence binds and commands entry into with relatively low efficiency. This lattermost method can be carried out by modifying the cell to contain an increased number of protein import sequence binding sites.

The maximum achievable *in vivo* concentration of antibody is approximately 25 μg/ml $= 2 \times 10^{-7}$ M (Schacker et al., *J. Inf. Dis.* 169, 37-40 (1994)). If uptake efficiency is of the order of about 1%, then the intracellular concentration achieved is approximately 2×10^{-9} M. This is a conservative estimate since an efficiency of about 3-60% has been described for tat-mediated delivery of heterologous proteins (Fawell et al., *Proc. Natl. Acad. Sci. USA* 91, 664-668 (1994)). Therefore, generally antibodies with K_ds of the order of nM are needed to achieve intracellular immunity according to the invention.

Fabs derived from libraries typically have K_ds in this range. Whole bivalent antibody molecules typically have affinities an order of magnitude or more higher. Furthermore, the affinities of recombinant Fabs can be improved by more than three orders of magnitude by *in vitro* mutagenesis and selection (Barbas et al., *Proc. Natl.*

Acad. Sci. USA 91, 3809-3813 (1994); Yang et al., *J. Mol. Biol.* 254, 392-403 (1995)). Therefore, according to the invention, Fabs can be produced with sufficient affinities to permit *in vivo* intracellular activity if reasonable efficiency can be achieved in getting the molecules into cells. This invention, of course, addresses means of obtaining reasonable efficiencies for getting molecules into cells (i.e., desirably, by attachment to protein import sequences).

A related issue is the total number of intracellular antibody molecules compared to the total number of viral target molecules in the case where the antibody is employed for intracellular immunization against a viral pathogen. Antibody can be ineffective unless it can saturate the target molecule. Of course, this is not the case with all viral pathogens, and indeed, in some instances, even less than saturation of the target molecule will obtain a desirable effect (e.g., a therapeutic effect, or merely a reduction in viral load). An intracellular concentration of about 2 x 10^{-9} M as described above corresponds to about 10^{12} molecules/ml. Therefore, for cell densities in the range of about 10^{7}-10^{8} cells/ml, desirably there are about 10^{5}-10^{4} molecules of antibody/cell. Again, this is probably a conservative estimate since 10^{7} molecules per cell have been delivered in the tat system (Fawell et al., *Proc. Natl. Acad. Sci. USA* 91, 664-668 (1994)). For infected cells in culture, an estimated maximum of about 5 x 10^{4} IE86 molecules/cell has been observed based upon a quantitative western analysis (Klaus et al., unpublished results). Therefore target saturation *in vivo* appears to be feasible in the context of the invention.

Generation of Human Fabs to hCMV (or other Viral Epitopes) from Phage Display Libraries

In order to effect intracellular immunity, preferably a first step is to obtain panels of human Fabs to the desired viral epitopes, e.g., for hCMV, preferably to the epitopes UL44 and IE86. Since it routinely is not known in advance the epitopes to which Fabs should be directed for optimal functional activity, a library approach optimally can be employed to rapidly yield large numbers of Fabs to a given antigen. Functional screening, as described below, is preferable to determine the Fabs that will be characterized in greater detail (e.g., for effecting intracellular immunity).

For instance, preferably according to the invention, Fabs can be generated to

the viral polymerase-associated protein UL44 from, for instance, an HIV-1 seropositive donor library. Should antibodies that recognize IE86 or other desired viral epitope not be recovered from any existing libraries, then new libraries can be constructed using appropriate tissue donors. As an example, young and otherwise healthy adults suffering from acute hCMV mononucleosis should provide an excellent donor pool for constructing a library for obtaining human Fabs to IE86 since these patients frequently possess high serum antibody titers against a wide range of viral antigens. Serum samples from these patients can be examined for their IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2 reactivity against IE86, and against isolated domains of IE86.

Thus, desirably according to this invention, phage display libraries and subsequent screening are employed to generate and obtain Fabs to other epitopes, including other viral epitopes, optimally which further will be manipulated to effect intracellular immunity. The illustrative means by which the antibodies can be obtained and screened are further described below.

a) Cloning Human Fabs from Appropriate Human Donors

In the case where it is desirable of necessary to construct phage libraries for obtaining Fabs, the cloned human Fabs desirably can be obtained from the appropriate human donor (e.g., a hCMV-infected donor, or a donor infected with another pathogen against which immunity is sought) as follows.

From about 3 to 5 donors having high serum titers to the hCMV IE86 and UL44 antigen (or other viral antigen), routine venipuncture (about 20 cc) and bone marrow aspiration (about 5 cc) is carried out. About 5 ml of bone marrow is aspirated and reverse transcribed to cDNA using standard oligo dT priming. Both lambda and kappa light chains desirably are PCR amplified using appropriate primers (e.g., for hCMV antigen primers, *see*: Kang et al., "Combinatorial Immunoglobulin Libraries in Phage." *In*: Methods: A Companion to Methods in Enzymology (R.A. Lerner and D.R. Burton, eds.), Vol. 2 pp 111-118. Academic Press, Orlando (1991); Williamson et al., *Biochem. J.* 277, 561-563 (1991)). The choice of heavy chain class for amplification depends in each case upon the measured serum titer.

Following amplification with the various primer pairings, preferably the pooled light and heavy chain PCR products are cut with the appropriate restriction enzymes and cloned into an appropriate vector, e.g., pComb3 (Barbas et al., *Proc.*

Natl. Acad. Sci. 88, 7978-7982 (1991)), or other appropriate vector, to give a phage display library. The libraries desirably are panned against UL44, IE86, or other appropriate epitope (e.g., a viral or other antigen epitope). This panning optimally is done by layering a solution of the library onto immobilized antigen (e.g., immobilized on plastic, such as a tissue culture plate or wells), incubating, and removing non-specifically bound phage by repeated washings to enrich for specific phage-Fabs. Optionally, isolated domains of the epitope are coated on ELISA wells to facilitate screening.

Display phage derived from the final round of panning preferably are eluted (e.g., with acid or excess soluble antigen) and converted to phagemid expressing soluble Fabs (e.g., Barbas et al., *Proc. Natl. Acad. Sci.* 88, 7978-7982 (1991)), or other appropriate form. Of the obtained phagemids, preferably about 20-40 clones are initially assessed as bacterial supernates for reactivity in an ELISA assay against the antigen with which they were panned. Since the PCR and phage amplification can produce repeat clones, positive clones desirably are sequenced, or characterized by other appropriate means, to identify for further characterization and employ those which clearly are unique.

b) Generation of Cloned Human Fab Panels to Epitopes

The methods described above desirably should provide panels of Fabs to the various domains of the preferred epitopes, e.g., UL44 and IE86, or other preferred epitope. This is confirmed in the Examples that follow, which describe three independent human Fab clones which bind with high affinity and specificity to the hCMV UL44 protein and were obtained using these methods.

Thus, in particular, preferably according to the invention, the methods are employed to provide panels of Fabs to IE86 and UL44, to the DNA-binding domain of IE86, and to the transcriptional activation domains of IE86, as wells as to provide panels of Fabs to other epitopes (e.g., other viral epitopes). For these methods, preferably the whole protein is included as well as the individual domains, for instance, to capture by the method certain epitopes that are dependent upon the integrity of the whole molecule as opposed to merely a domain. If subsequently Fabs to other proteins (e.g., viral proteins), or to epitopes on IE86 not present in the above-described panels are desired, then the libraries optionally are selected with a different antigen, or with an antigen that is presented in a different form. Alternately, preferably selection is directed to one part of a molecule by blocking

other epitopes with existing antibodies (e.g., mouse antibodies or recombinant Fabs) prior to panning. The process optimally is carried out repeatedly to reveal responses to relatively minor epitopes. This should yield the desired specificities since it recently has been shown, for instance, that the epitope masking strategy can be used to access many different epitopes on a single protein (Ditzel et al., *J. Immunol.* 154, 895-908; Binley et al., *AIDS Res. Hum. Retroviruses* (1996); Burton, "Human antibodies to viral pathogens from phage display libraries." *In* Vaccines 95; Molecular approaches to the control of infectious diseases. R.M. Chanock, F. Brown, H.S. Ginsberg, and E. Norrby, Eds. (Plainview, NY: Cold Spring Harbor Press), pp. 1-11 (1995)). Also, desirably antibodies are generated according to the invention to functional sites from synthetic libraries using antigen masked at non-relevant epitopes. Such antibodies are likely to be of moderate affinity, and may require *in vitro* "affinity tuning" such as is known in the art (see, e.g., Yang et al., *J. Mol. Biol.* 254, 392-403 (1995)).

Design and Development of Molecularly Engineered "Invasive" Antibodies

For effecting intracellular immunity, the present invention desirably provides a chimeric molecule consisting essentially of a protein import sequence attached to the terminus of an antibody directed against a pathogen, such that the attachment does not interfere with the effectiveness of the antibody. In particular, the invention desirably provides a human immunodeficiency virus (HIV) tat protein import sequence, preferably residues 38-72 or 47-59 of such a tat protein import sequence, attached to the C-terminus of an antibody. Optimally the antibody is directed against hCMV, preferably is an anti-UL44 antibody, and even more desirably is a Fab selected from the group consisting of Fab GL CMV 14 ("Fab GL14"), Fab GL CMV 5 ("Fab GL5"), and Fab GL CMV 34 ("Fab GL34").

The Examples that follow describe the construction of chimeric molecules that contain the protein import sequence residues 38-72 of tat, or residues 47-59 of tat linked to the C-terminus of an anti-CMV (e.g., UL44) Fab. The constructs together with the original unmodified Fab, also have been "tagged" as described herein with green fluorescent protein (GFP). The GFP-fused constructs have many uses, including uses in experiments designed to determine the invasive and subcellular localization of cell-associated antibody by direct immunofluorescence. These GFP-tagged Fabs thus provide for optimizing conditions for cellular delivery.

Also, among other things, the original unmodified Fab and its GFP fusion desirably can be used as controls in experiments to assess antibody uptake.

In order to compare the relative efficiencies by which the unmodified and tat-tagged Fabs are bound to the cell surface, optimally the time course of binding using radioiodinated or biotinylated Fab is determined, for instance, using HeLa cells and/or human foreskin fibroblasts. Specifically, the cells desirably are treated with Fab for various times, washed in medium, and the amount of surface-bound Fab is determined using standard means. Results described in the Examples that follow confirm that the tat-tagged Fab shows dramatically enhanced binding to the cell surface. Labelled Fab also desirable is used to determine the affinity and number of binding sites involved in cell surface interaction.

The subcellular localization of cell-associated antibody preferably is followed, e.g., by direct immunofluorescence using GFP-Fab, FITC-conjugated Fab, and by indirect immunofluorescence, for instance, employing a mouse anti-human IgG $F(ab')_2$ specific anti-serum as the primary antibody and an anti-mouse $F(ab')_2$ specific FITC conjugate as the secondary antibody. Cells preferably are fixed and examined under a light fluorescence microscope at different times after addition of test Fab. Results described in the Examples that follow confirm the ability of the tagged antibodies to selectively bind the cell surface, and ability to penetrate the cell and translocate to the nucleus. However, it is not necessary according to the invention that the protein import sequence and/or or intracellular antibody physically bind the cell surface, i.e., other means of cell entry also are contemplated by the invention.

Recognition of and Interference With Viral Protein Function in Intracellular Immunity

The methods and products of the invention desirably are employed to effect intracellular immunity, especially viral immunity.

The immediate-early (IE) proteins of herpesviruses are a group of virally encoded nuclear phosphoproteins. They are the first proteins expressed in the infectious cycle and coordinate subsequent viral gene expression. The expression of these proteins are important events, not only in virus growth, but also in subtle processes such as viral latency. In this application a preferred target according to the invention (i.e., when employed for effecting intracellular immunity against hCMV)

is the major IE protein of hCMV termed IE86 (IE2-UL122). It is known that regulation of viral gene expression by this protein constitutes an important control mechanism of viral growth and perhaps latency. It should be noted that while the IE86 protein of hCMV provides a good candidate protein for antiviral targeting, in principle, any of the IE proteins or combinations thereof may well be equally effective targets. However, at present little information is available about the functional role of the other hCMV IE proteins.

Functional *in vitro* assays have been developed that faithfully reconstitute, in part, transcriptional regulation by purified recombinant IE86 protein (Jupp et al., *J. Virol.* 67, 7539-7546 (1993); Wu et al., *J. Virol.* 67, 7547-7555 (1993)). Using these systems, the mechanism by which the viral IE protein interacts with the host-cell transcriptional machinery associated with viral promoters desirably can be further investigated, as described in the Examples which follow. These assays include various DNA-binding and *in vitro* transcription systems and provide ideal secondary screening procedures for candidate Fabs identified in the primary screens. The IE86 protein encodes three distinct functional domains, a DNA-binding domain located between amino acid positions 290 to 579, and two transcriptional activation domains located between amino acid positions 1 to 89 and 544 to 579, respectively (Stenberg et al., *J. Virol.* 64, 1556-1565 (1990); Jupp et al., *J. Virol.* 67, 7539-7546 (1993); Wu et al., *J. Virol.* 67, 7547-7555 (1993); Baracchini et al., *Virology* 188, 518-529 (1992); Hermiston et al., *J. Virol.* 64, 3532-3536 (1990); Jupp et al., *J. Virol.* 67, 5595-5604 (1993); Macias et al., *Proc. Natl. Acad. Sci. USA* 90, 707-711 (1993); Pizzorno et al., *J. Virol.* 65, 3839-3852 (1991)). Since Fabs targeting these domains likely will disrupt IE protein function, it is valuable to identify and characterize Fabs binding to these regions of IE86. Since these domains account for approximately 67% of the IE86 coding sequence, it is possible to obtain inhibitory Fabs using the methods of the invention.

Another preferred target for effecting intracellular immunity (i.e., when the invention is employed for effecting intracellular immunity against hCMV) is the UL44 gene product, a phosphorylated DNA-binding protein of approximately 52 KDa (Landini et al., *J. Clin. Microbiol.* 27, 2324-2327 (1989)). This protein is expressed beginning in the early phase of infection and accumulates in the nuclei of infected cells throughout the late phase. Biochemical evidence suggests that UL44 is the hCMV homologue of HSV UL42, one of a subset of essential genes required

for viral DNA replication (Ertl et al., *J. Virol.* 66, 4126-4133 (1992)). The HSV UL42 protein is known to act as a catalytic subunit of the viral DNA polymerase by enhancing its processivity. The hCMV UL44 product is also highly associated with the viral polymerase and is believed to function in an analogous manner (Ertl et al., *J. Virol.* 66, 4126-4133 (1992)). Similar to the HSV UL42, the hCMV UL44 protein exhibits DNA binding characteristics that are thought to be independent of its polymerase-associated activity. The DNA-binding capacity of UL44 indicates that there may be other functions associated with this protein. Importantly, the hCMV UL44 protein has been shown to be essential for viral replication (Ripalti et al., *J. Virol.* 69, 2047-2057 (1995)). Thus, the design of intracellular antibodies which block nucleic acid or polymerase binding to UL44 accordingly is a very attractive approach to the development of antivirals which are likely to be highly viral-specific, and are especially preferred antibodies for use in the invention.

The method of the invention also preferably is employed for effecting intracellular immunity against other epitopes (e.g., other viral epitopes). The descriptions herein for hCMV are preferred embodiments that are merely exemplary of the use of the invention.

a) Cloning and Purification of Recombinant IE86 and UL44 Proteins.

In order to obtain antibodies according to the invention, desirably the antigens against which the antibodies are to be obtained are isolated, e.g., by any appropriate means, e.g., means as described herein.

The construction of the His-tagged IE86 expression clone (p86-6His) has been described and can be done based on the pDS56,6His vector system (Leuthardt et al., *Gene* 68, 35-42 (1988)). Similarly, a His-tagged UL44 expression clone also has been constructed based on the same vector system. This vector system contains six codons encoding histidine (His) residues immediately downstream of the ATG initiation codon. The generation of expression clones for the specific sub-domains of proteins can be prepared in a similar manner, e.g., by PCR-directed cloning. Briefly, the generation of the DNA-binding domain (or other appropriate domain) desirably is accomplished by fusing amino acid codons 290 to 579 of IE86 to the 6-His tagged residues. Similarly, the construction of the activation domains preferably involves fusing amino acid codons 2 to 90 (AF-1) and 540 to 579 (AF-2) to the 6-His residues, respectively. N-terminal and C-terminal deletion clones desirably also

are engineered for UL44. These domains preferably are expressed in *Escherichia coli* and purified by affinity chromatography as described below.

Namely, *Escherichia coli* harboring each of the expression plasmids desirably is grown to an optical density at 550 nm of about 0.7-0.8 prior to induction with IPTG. Cells generally are harvested after about 90 to about 120 minutes of induction, and then are lysed in buffer that preferably contains 50 mM sodium phosphate (pH 7.8), 1 mM PMSF, 1% Tween 20, 1M NaCl and 1 mg of lysozyme per ml. Lysates are cleared by centrifugation and subjected to Ni^{2+} chelate chromatography over a column equilibrated in buffer containing 50 mM sodium phosphate (pH 7.8), 500 mM NaCl, and 10% glycerol. After washing with this buffer, the column is washed in a similar buffer at pH 6.0. A final wash is performed with the latter buffer containing 75 mM imidazole prior to elution in the same buffer containing 500 mM imidazole. Fractions containing each of the proteins are pooled and dialyzed against storage buffer. At this stage of purification we have found, generally the protein products are more than about 90% pure.

To confirm that the correct protein products have been expressed, Western analysis desirably is performed, for instance, using a monoclonal antibody directed against AF-1 and an anti-peptide antibody that recognizes AF-2 of IE86, and cloned human Fabs that bind UL44. Alternate means of detecting expression also can be employed.

b) Antibody Inhibition of IE86 DNA-Binding Function and Activation Domains

Two lines of evidence indicate that interrupting the ability of IE86 to bind DNA impairs its function. First, mutations within the DNA-binding domain abrogate the function of IE86 as both an activator of early promoters and as a repressor of its own promoter (Stenberg et al., *J. Virol.* 64, 1556-1565 (1990); Jupp et al., *J. Virol.* 67, 7539-7546 (1993); Wu et al., *J. Virol.* 67, 7547-7555 (1993); Baracchini et al., *Virology* 188, 518-529 (1992); Hermiston et al., *J. Virol.* 64, 3532-3536 (1990); Macias et al., *Proc. Natl. Acad. Sci. USA* 90, 707-711 (1993)). Second, hCMV early promoters contain functional DNA-binding sites for IE86 (Arlt et al., *J. Virol.* 68, 4117-4125 (1994)). Therefore, preferably according to the invention, Fabs that inhibit DNA-binding desirably are selected from the previously described library panels.

Optimally such Fabs are selected by testing for their ability to disrupt IE86

binding in electrophoretic mobility shift assays (EMSA). Using this approach, desirably, bacterially expressed IE86 is pre-incubated for about 1 hour with increasing concentrations of individual Fabs prior to the addition of radiolabelled double-stranded oligonucleotides corresponding to the IE86 binding site. After an additional ~30 minutes of incubation time, reactions preferably are loaded onto a native polyacrylamide gel and subjected to electrophoresis. Fabs that do not bind IE86 can be used in control experiments to monitor for non-specific effects. To independently confirm that these Fabs inhibit DNA binding by IE86, DNase I footprinting experiments desirably can be performed. In these experiments, the same procedures are followed as described for the EMSA experiments, except the reactions preferably are challenged with DNase I, processed, and the products resolved on a sequencing gel. Finally, the Fabs also desirably are tested in a functional *in vitro* transcription assay. Previously, it has been shown that binding of IE86 to the crs element is an essential requirement for mediating transcriptional repression (Wu et al., *J. Virol.* 67, 7547-7555 (1993)). Using these *in vitro* systems (and as described in the following examples), selected Fabs preferably are pre-incubated with IE86 prior to the addition of test template (i.e., containing the crs element) and control template (i.e., having a mutated crs element) to the reactions. If Fabs interfere with DNA-binding, there accordingly should be observed an inhibition of IE86 repression.

In the event where there is failure to identify Fabs that disrupt DNA binding, preferably Fabs that recognize the two activation domains (AF-1 and-2) of IE86 can be identified and characterized. For these experiments, Fabs from the above-described panels desirably are selected for either AF-1 or AF-2 binders, e.g., by affinity chromatography. Briefly, the Fabs desirably are applied to a 200 microliters Ni^{2+} chelate columns equilibrated in a low salt buffer to which purified bacterially expressed His_6-tagged AF-1 or AF-2 domains have been previously bound. After extensive washing in the low salt buffer, the column is stepped-washed with about 400 microliter aliquots of the same buffer containing increasing concentrations of salt. An aliquot is removed from each fraction and subjected to SDS-PAGE. Following electrophoresis, the gel is transferred to a membrane and probed for Fabs by western analysis. Affinity chromatography with full-length IE86 bound to the columns can also be used to corroborate that Fab fragments bind holo-IE86, thereby eliminating Fab fragments that either recognize an incorrectly folded epitope or

those that might be blocked from binding due to inter-domain folding. Functional *in vitro* transcriptional assays are available for IE86 activation (Klucher et al., Mol. Cell. Biol. 13, 1238-1250 (1993)), however the precise mechanism and domains responsible for IE86 mediated *in vitro* activation is at present not known. Because these activation domains are relatively small, likely only a limited number of Fabs target these regions and therefore it may not be necessary to apply any additional selective screens. The ability of these Fabs to inhibit activation functions in co-transfection assays also can be performed, as is known in the art. In summary, these methods according to the invention should lead to the identification of Fabs capable of interrupting IE86 function.

c) Antibody Inhibition of UL44 DNA-and Polymerase binding

Studies on the HSV UL42 protein indicate that interrupting the ability of the protein to bind either DNA or polymerase impairs its function. Because the mechanism by which different members of the herpes virus family replicate appears to be the same, it is expected that the hCMV UL44 protein will behave in a similar manner. Therefore, desirably according to the invention, Fabs are obtained that inhibit DNA-binding.

A preferred selection system for accomplishing this goal is by testing for the Fabs' ability to disrupt UL44 binding during DNA cellulose chromatography. For these experiments, desirably, bacterially expressed UL44 is preincubated for about 1 hour with increasing concentrations of individual Fabs prior to loading on a DNA cellulose column. The columns are then desirably step- washed with increasing concentration of salt. The hCMV UL44 protein by itself elutes from such columns at about 400mM NaCl (see, e.g., Examples). If the antibodies disrupt DNA binding, an elution of UL44 in the presence of the test Fab should occur in the lower salt wash fractions. Antibodies that do not interact with UL44 or DNA can be used as controls.

The interaction of UL44 with DNA is thought to be non sequence-specific. However, it can be determined whether UL44 has any sequence preference by performing the CASTing (cyclic amplification and selection of targets) technique to fish through a set of randomly generated synthetic polynucleotides for any specific DNA sequences to which UL44 might bind. If UL44 is found to recognize DNA sequence-specifically then the ability of Fabs to disrupt DNA-binding can be further confirmed by performing the EMSA screens as described above for IE86.

In addition to identifying Fabs that disrupt DNA binding, Fabs that interfere with the ability of UL44 to associate with the viral DNA polymerase desirably can be identified and characterized according to the invention. For these experiments, Fabs desirably are selected for their ability to inhibit binding of polymerase to UL44 by affinity chromatography. Briefly, Fabs preferably are applied to Ni^{2+} chelate columns equlibrated in a low salt buffer to which purified bacterially expressed His-tagged UL44 (or sub-domains) have been previously bound. HCMV DNA polymerase (i.e., radio-Labelled in a cell-free translation system) desirably is applied to the column and washed in the equilibration buffer. A recombinant that expresses functional hCMV DNA polymerase in an *in vitro* transcription/translation system (Ye et al., *J. Virol.* 67: 6339-6347 (1993)) desirably can be applied in these studies. Basically, an aliquot is removed from each fraction and subjected to SDS-PAGE using conditions to ensure that about 100% of polymerase binds to the column. Following electrophoresis, the gel desirably is dried and exposed to autoradiographic film. If the test Fab disrupts polymerase association with UL44 then it is expected that polymerase will be observed in the eluate. However, if the Fabs fail to interfere with the association, then fractions will score negative for the presence of Labelled polymerase. Again, the Fabs that do not interact with UL44 or polymerase have practical utility as controls for specificity.

The interaction domain of UL44 with polymerase also can be mapped by using different deletion mutants of UL44 in the UL44 affinity chromatography described above. These associations can be confirmed, for instance, by co-immunoprecipitation studies (i.e., including pull-down experiments with GST fusion constructs), and by the yeast two hybrid-system. Mutants of UL44 that are incapable of interacting with polymerase serve as additional controls for the above experiments. In summary, these methods provide still further means leading to the identification of Fabs capable of interrupting UL44 function.

d) Screening of Potential Inhibitory Fabs by Microinjection

The above selection strategies desirably can be employed to identify a number of Fabs for inhibiting hCMV expression. Because viral late gene expression requires viral DNA replication, inhibiting either IE or UL44 protein function will block the ability of late genes (e.g., such as pp150 (UL32)) to be expressed.

To confirm that the above-identified Fabs have the potential to block full hCMV expression in cells, microinjection of Fabs into infected cells desirably is

performed. This can be done, for instance, using an Eppendorf microinjection system, or other appropriate means. Preferably, human foreskin fibroblast (HFF) cells (or other appropriate cells) are infected at a multiplicity of infection (m.o.i.) of from about 1 to about 10. After absorption of virus, infected cells desirably are injected (preferably immediately for IE inhibitory Fabs, and preferably about 10 hours post-infection for anti-UL44 Fabs) with test and control Fabs (e.g., about 2 mg/ml in PBS). At about 72 hours post-infection, the cells desirably are fixed and stained with an antibody against a true late protein (e.g., pp150, UL32). The percentage of viral inhibition preferably is calculated (using a minimum of about 100 injected cells) according to the formula: Viral inhibition (%) = (A-B)/A x 100, where A is the percentage of surrounding non-injected cells staining for late antigen, and B is the percentage of microinjected cells staining for late antigen. These experiments should confirm the ability of the Fabs to inhibit hCMV expression as well as provide as an additional screen for revealing the *in vivo* potency of candidate Fabs.

e) Invasive Antibody Inhibition of hCMV Gene Expression and Replication

The above selection strategies desirably can be employed to identify Fabs for reconstructing in the delivery system. Candidate Fabs desirably further are molecularly engineered with the tat-delivery sequence to provide for intracellular delivery. To confirm that these Fabs have been endowed with the ability to be taken up by cells, preferably, indirect immunofluorescence of cells treated with the Fabs is performed. Also, desirably studies are done to monitor the effects of Fabs on cell growth and viability.

Thus, experiments desirably are done to confirm that the intracellular Fabs interfere with IE86-induced activation of early promoters and repression of the MIEP in transient co-transfection assays using wild type and mutant IE86-responsive reporter constructs. To validate that these Fabs are capable of disrupting IE86 regulation of hCMV early promoters in the context of a viral infection, preferably human foreskin fibroblast (HFF) cells are infected with hCMV. After absorption of virus, infected cells desirably are treated with selected Fabs with and without transporter tags, and harvested for RNA extraction at different times (e.g., between about 8 and about 72 hours post-infection). Infected cells that have not been treated with Fabs desirably are employed as an additional control. The steady-state levels of two early gene transcripts (e.g., the 2.7 Kb transcript and the polymerase transcript which are known to be regulated by IE86) preferably also are monitored by

quantitative S1 nuclease protection experiments using standard protocols. These experiments accordingly should confirm the ability of the Fabs obtained according to the invention to inhibit IE86 from mediating transcriptional regulation of viral promoters in living cells. Further, these studies serve as an additional screen for revealing the *in vivo* potency of candidate Fabs.

To examine modulation of hCMV permissiveness by the so-called invasive Fabs, HFF cells desirably are infected with hCMV at m.o.i.s ranging from about 0.05 to about 5. Infected cells preferably are treated with varying concentrations of Fabs with and without the delivery sequence, and optimally, one-step growth analysis is performed. Cell lysates and supernatants then desirably are prepared at various times post-infection from Fab-treated and untreated cells, and monitored for hCMV replication by plaque assay (e.g., Wentworth et al., Proc. Soc. Exper. Biol. & Med. 135, 253-258 (1970)). Values for 50% effective concentration (EC_{50}) required to inhibit production of infectious virus for both extracellular and intracellular virus yields desirably are evaluated. These experiments accordingly can be employed to confirm that Fab treatment of infected cells results in an efficacious inhibition of hCMV replication. To verify that this effect is not cell-type restricted, experiments preferably are repeated with glial (U373) and retinoic-acid treated embryonal (NT2-D1) cells exposed to variable levels of Fabs. Also, to test that inhibition effected by the Fabs is not viral-strain specific, experiments optimally are conducted with not only different laboratory strains of hCMV (AD169 and Towne), but also with recent clinical isolates (Gerry).

Design of Antibodies Capable of Sequence Specific DNA Recognition

As previously described, a preferred target for invasive antibodies capable of reaching the nucleus is DNA. Antibodies capable of sequence-specific DNA recognition optimally can be targeted to viral promoters, for instance, and used to repress viral replication. It is generally believed that disrupting protein-DNA interactions is not particularly virus-specific because the same transcription factors are used normally by the host. Accordingly, targeting a viral factor binding to viral DNA tends to obviate toxicity issues (e.g., issues arising from interference with host cell transcription factors). Moreover, the precise target location of a repressor that functions by steric interference, such as a DNA-binding Fab, is absolutely critical for effecting inhibition (i.e., it must target a sequence between the TATA box and the

initiation site) (Wu et al., *J. Virol.* 67, 7547-7555 (1993)). Because repression occurs only at this location it further restricts the action of the inhibitor to a defined locus, since binding elsewhere would have no effect. The challenge is then to produce antibodies capable of DNA sequence-specific recognition. As described herein, the methods of the invention desirably can be employed along these lines, preferably using hCMV as a model system.

a) Selection and Variation of Sequence-Specific Recognition of DNA by Fabs

Desirably, experiments are done to confirm the DNA sequence specificity of the antibody Fab containing the TFEB recognition motif described above. Already some evidence suggests sequence-specific DNA recognition (McLane et al., *Proc. Natl. Acad. Sci.*, 92, 5214-5218 (1995)). The binding of the Fab-Ebox to the MLP probe has been further investigated in DNAse I protection and in methylation interference assays. The DNA sequence specificity of antibody recognition further desirably is examined by a target detection assay (Thiesen et al., *Nucleic Acids Rsch.* 18, 3203-3209 (1990)).

In this assay, preferably a purified putative DNA binding protein (e.g., here a Fab) and a pool of random double stranded oligonucleotides (preferably randomized at the position of the CACGTG sequence in the MLP probe) harboring PCR primer sites are included in a cycle which desirably consists of four separate steps: (1) a DNA-protein incubation step; (2) a protein-DNA complex separation step (i.e., using an anti-Fab affinity column); (3) a DNA elution step (e.g., with high salt); and (4) a PCR amplification step. The stringency of selection optimally can be increased in consecutive cycles. Since tiny amounts of retained DNA can be rescued by PCR, the buffer systems, salt concentrations and competitor concentrations can be varied to determine optimal binding oligonucleotides, following cloning and sequencing. The method has been successfully applied to the determination of the DNA binding specificities of mutated zinc finger domains (Thiesen et al., *FEBS Letters* 283, 23-26 (1991)).

These experiments optimally can be employed according to the invention to confirm the number of bases involved in Fab-Ebox recognition, and the specificity of the interaction. This specificity then desirably can be optimized, e.g., using a randomization/selection strategy. Namely, the positions around the central arginine residues in the TFEB motif, located within the heavy chain CDR3, of Fab-Ebox

which are currently alanines, desirably are randomized using random oligonucleotides and overlap PCR. This desirably will generate a library of phage display Fabs based on the Fab-Ebox. This library optimally will be panned against the MLP probe to select the variants with the highest affinity for the probe. Further, tight specificity can be encouraged by a number of strategies. For example, in the case where as sequence such CACGTG is the target recognition sequence but that cross-reactivity with CTCGTG is present, then the library optionally is pre-adsorbed on a probe bearing CTCGTG, and/or panning is carried out in the presence of excess soluble probe bearing CTCGTG. The overall strategy of randomization and selection for varying DNA specificity has been successfully employed for zinc finger domains (Rebar et al., *Science* 263, 671-673 (1994)).

Further illustrations of this strategy and generation of sequence specific DNA binding antibodies are set out in Le Blanc et al., *Biochemistry* 37 (17), 6015-22 (1998), and McLane et al., *Proc. Natl. Acad. Sci.* 92 (11), 5214-8 (1995), herein incorporated in their entireties by reference.

b) Generation of Bivalent Fabs as a Means for Enhancing the Complexity of Sequence Recognition

Accomplishment of specific recognition of the 6-base sequence using the TFEB motif will allow the potency of the corresponding invasive Fab to be assessed, for instance, in an *in vitro* system. In some cases, it is more likely that the best Fab obtained will show very tight recognition of 3-4 bases with some variation permitted in the others. This might be expected, for instance, for recognition in the major groove where 4 bases will be presented on one face of the double helix. In that case, desirably bivalent Fab recognition of a repeat sequence can be employed according to the invention. This should allow for more specific recognition of a target site. For example, in the case where CACG is the Fab recognition motif and this sequence occurs as a direct repeat with a spacing of 12 nucleotides, then a bivalent Fab recognizing the sequence CACG would have the advantage that the repeat is 12 base pairs apart which corresponds roughly to one turn of the double helix so that binding to the same face of the DNA helix could occur. The modular flexibility of the Fab arms of the antibody molecule should be of assistance in permitting bivalent recognition.

Thus, using the methods of the invention, binding of the most promising mutant Fab-Ebox as a bivalent molecule to its best recognition sequence repeated 12

base pairs apart desirably can be accomplished. The Fab preferably is rendered bivalent, for instance, by use of a redesigned oligonucleotide primer which extends the existing heavy chain C-terminus into the hinge region of the molecule. In so doing two extra cysteine residues optionally are incorporated into the sequence yielding Fab'. Following gentle reduction and reoxidation, the secreted Fab' will form $F(ab')_2$ via interchain hinge disulphide bonds. These hinge disulphides are more susceptible to reduction than those formed between the heavy and light chains in the Fab' molecule, hence $F(ab')_2$ may potentially be broken into monomer in the intracellular environment.

Alternately, preferably $F(ab')_2$ can be produced by chemically linking two Fab' molecules together. For instance, hinge sulphydryl (SH) groups are efficiently coupled via a stable thioether bond using bis-maleimide crosslinkers such as o-phenylenedimaleimide (o-PDM). In this instance, preferably the heavy chain of the Fab is extended to include only one extra cysteine residue at its C-terminus, thus avoiding the possibility of crosslinking SH groups on the same Fab' molecule. The length of the region connecting the folded Fab to this cysteine optionally is varied to optimize bivalent binding. The specificity of bivalent binding desirably is characterized, for instance, by competition experiments and by the target detection assay described above.

c) Regulation of Viral Expression by DNA-Binding Fabs

Desirably, at a next stage, the ability of the most sequence-specific DNA-targeted Fab to inhibit viral gene expression as an invasive antibody is confirmed. This preferably is accomplished by cloning the best recognition sequence (either 6 bases, or a repeated sequence of 3-4 residues) into the hCMV promoter (MIEP) by PCR mutagenesis of the IE86 (crs) element. Replacing the crs element with binding sites for other DNA binding proteins is sufficient for transcriptionally blocking the MIEP. By analogy, if an Fab recognition motif were to occur in the same location as the crs element then likely the Fab should be able to effect inhibition of transcription from the MIEP. This can be confirmed, for instance, by using the previously described *in vitro* transcription systems. If the Fab is found to inhibit transcription, an invasive antibody can be constructed, and inhibition of viral promoter function assessed by transfection assays.

This strategy preferably can be completed based not only on the TFEB recognition motif, but also, based on motif(s) occurring naturally in the hCMV

promoter. As an example of a potential target for this sort of approach, in the following sequence in the IE86 binding site (crs), the underlined bases have been identified as those which are absolutely required for functionality:

5'-AGCTCGTTTAGTAACCGTCAG-3' [SEQ ID NO: 16]

The sequence CGT (underlined) is appropriately spaced by 12 nucleotides and is conserved amongst other IE86 binding sites found in early promoters of hCMV (Arlt et al., *J. Virol.* 68, 4117-4125 (1994)). Therefore, it should also be possible to interfere with IE86 function by usurping or displacing its interaction with nucleic acids by directly competing with its binding site. If these experiments described above fail to provide a suitable Fab, a new Fab containing a different motif engineered into the heavy chain CDR3 desirably can be employed instead.

d) Animal model for testing *in vivo* efficacy

According to the invention, optionally, a transgenic animal model system can be employed for investigating tissue-specific expression of human CMV (Baskar et al., *J. Virol.* 70, 3207-3214 (1996); Baskar et al., *J. Virol.* 70, 3215-3226 (1996); see, procedures described in the Cold Spring Harbor Laboratory Manual "Manipulating a Mouse Embryo"). Preferably this model is based on transgenic mice containing the *lacZ* gene under the control of the MIEP of hCMV. In the transgenic mice, *lacZ* expression occurs in the brain, eye, spinal cord, esophagus, stomach, pancreas, kidney, bladder, gonads, spleen, salivary gland, thymus, bone marrow, skin, cartilage, and muscle. The cell types which demonstrated expression in these organs were predominantly epithelia and endothelial cells. The majority of tissues and cell types which display MIEP activity parallel tissues naturally infected by hCMV in the human host. These observations strongly suggest that the MIEP is a crucial determinant for directing the tissue-tropism and state of activation in the host. Somewhat surprisingly, in this animal model system a lack of expression occurs in the liver and lung, organs commonly associated with hCMV infection. Recent experiments have indicated that the initiation of an immune response in these organs can lead to activation of *lacZ* expression from their silent states. Thus, this animal model systems serves as a good model for studying *in vivo* aspects of hCMV, especially the acute and reactivated/latent states of viral expression as determined by the MIEP.

The system accordingly desirably can be employed for targeting the invasive antibodies *in vivo*. To date, the tat-delivery system has been shown to predominantly target heart, spleen, liver, lung, muscle with the major cell type targeted being endothelial cells. These tissues are all sites of infection by hCMV. Thus, in these experiments, first desirably will be examined the *in vivo* biodistribution and potential delivery activity of the tat-Fabs, preferably by injecting radiolabelled tat-Fabs and control Fabs intravenously into mice. After various time periods, desirably the animals will be sacrificed, their tissues harvested, and sections prepared for histological analysis and autoradiography. These experiments will confirm the *in vivo* targeting of the Fabs. It is anticipated that the tat-Fab biodistribution will not be markedly different from that observed in the study of Fawell et al. (Fawell et al., *Proc. Natl. Acad. Sci. USA* 91, 664-668 (1994)), since different cargo proteins appear to develop the same tissue pattern.

These experiments desirably are based on the DNA-binding antibodies that target the MIEPs' crs element described above. Different concentrations of these tat-Fabs and control Fabs optionally can be intravenously injected into the transgenic mice. After various times, animals desirably are sacrificed, and tissues harvested for histochemical analysis and RNase protection assays. ß-galactosidase activity in the kidney can be employed as a control since the antibody should not target this organ. If the antibody is capable of down regulating the MIEP, it is expected that a decrease in transgene expression in organs such as the spleen (but not the kidney) will be observed. These studies test an acute activation state of the MIEP. To confirm whether invasive antibody therapy may have some efficacy to maintain a latent state, organs such as the liver and lung desirably are examined for *lacZ* activity under conditions of reactivation in the presence of control and anti-crs Fabs.

For other Fabs employed according to the invention, it likely is desirable to produce a new line of transgenic animals containing a high affinity binding site for the best Fab in place of the viral crs.

The invention further contemplates selectively targeting infected cells. It is possible with an efficient system (i.e., having low to zero toxicicity) that it will not be necessary to specifically target those cells infected by hCMV. Alternatively, it may be possible to selectively target infected cells, for instance, based on the fact that hCMV expresses a number of cell surface proteins. These proteins could be used to deliver high local concentrations of antibodies to the infected cell. This

could be established, for instance, by generating a bivalent antibody consisting of the invasive Fab and an Fab that binds virally infected cells. Desirably, for targeting virally infected cells, viral gB glycoprotein is employed. In this case, an anti-gB Fab could also function to neutralize extracellullar virus, thus providing a two step inhibition protocol. Alternately, the tat delivery peptide preferably is modified to selectively bind CMV infected cells. In this case, desirably a library of randomized sequences is cloned in place of or in conjunction with the delivery peptide. This Fab-peptide library desirably is expressed on the surface of filamentous phage. The phage display library then optimally is panned against CMV-infected cells.

In summary, the development of novel molecularly targeted approaches to anti-viral therapy of persistent/latent infections necessitates two essential criteria to be fulfilled. These are: (1) the requirement to have the capacity to affect the intracellular nature of virus replication; (2) the requirement for tight selectivity such that virus inhibition can be effected without harming the host. The novel development of intracellular immunity by 'invasive' human antibodies as described herein appears to satisfy these criteria. First, these invasive antibodies are designed to combat viruses inside cells as opposed to their conventional external mode of action. Second, antibody phage display affords the possibility of obtaining human molecules of high affinity and tight selectivity. Third, since human antibodies are naturally occurring molecules in humans, they generally have low toxicity and, by targeting viral encoded functions, should avoid broad spectrum activity.

A Construct According to the Invention

Accordingly, the invention also provides a construct, e.g., a construct comprising a protein import sequence, which optimally is either a native or non-native protein import sequence according to the invention. A "construct" is any form of molecule in which a protein import sequence is operatively connected with a cargo region and that can be employed for intracellular delivery. For instance, the connection between the protein import sequence and the cargo region can be a noncovalent bond (e.g., as in antibody/antigen binding), or a covalent bond.

In this sense, a "construct" includes (but is not limited to) a chimeric molecule as defined herein. However, whereas a chimeric molecule in the case of a protein fusion or antibody fusion desirably is present in a proteinaceous form, a construct includes vectors (e.g., polynucleotide vectors) such as a polynucleotide

vector that encodes a protein fusion or antibody fusion according to the invention.

Along these lines, a "construct" includes, but is not limited to a vector (e.g., having genetic incorporation of a protein import coding sequence and/or perhaps another coding sequence into a polynucleotide vector), or a conjugate-type vector (i.e., wherein a coding sequence, polypeptide sequence, or other moiety is noncovalently associated with a vector), or other appropriate moiety that can be employed for cell entry. As used herein a "vector" is a vehicle capable of effecting gene transfer, and has the general meaning of that term as understood by those of skill in the art. Preferably a vector according to the invention comprises a nucleic acid sequence that encodes a protein import sequence and a cargo region as described herein, desirably which are so arranged on the vector as to form, upon translation, a fusion protein or antibody fusion, as discussed above. Along these lines, such a vector further preferably can comprise a spacer coding sequence.

The vectors according to the invention include, but are not limited to, plasmids, phages, and viruses. Preferably, a vector according to the invention is a viral or plasmid vector. In particular, desirably the vector comprises a nucleic acid sequence that encodes native or non-native HIV tat peptide sequence (i.e., SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, or SEQ ID NO:7). The vectors according to the invention are not limited to those that can be employed in the method of the invention of intracellular delivery, but also include intermediary-type vectors (e.g., "transfer vectors") that can be employed in the construction of other vectors, for instance, in the construction of other vectors, for instance, in the construction of those vectors that are actually employed to contact cells.

In terms of a viral vector (particularly a retroviral vector, especially a replication-deficient viral vector), such a vector can comprise either complete capsids (i.e., including a viral genome such as a retroviral genome) or empty capsids (i.e., in which a viral genome is lacking, is incomplete, or is degraded, e.g., by physical or chemical means). Preferably the viral vector comprises complete capsids, i.e., as a means of carrying one or more moieties (e.g., cargo, such as antibody peptides and non-antibody peptide; or their encoding nucleic acids). Since methods are available for transferring viruses, plasmids, and phages in the form of their nucleic acid sequences (i.e., RNA or DNA), a vector similarly can comprise RNA or DNA, in the absence of any associated protein such as capsid protein, and in the absence of any envelope lipid. Similarly, since liposomes effect cell entry by fusing

with cell membranes, a vector can comprise liposomes, with constitutive nucleic acids encoding the coat protein. Such liposomes are commercially available, for instance, from Life Technologies, Bethesda, Md., as well as from other vendors, and can be used according to the recommendations of the manufacturer. The soluble protein (as produced using methods described herein) can be added to the liposomes either after the liposomes are prepared according to the manufacturer's instructions, or during the preparation of the liposomes.

As stated previously, a chimeric molecule of the invention can be in the form of a fusion protein or antibody fusion, as well as in the form of another molecule. Such a fusion protein or antibody fusion can be produced by means of a vector, i.e., wherein the protein import sequence, optional spacer sequence, and further cargo region, are in their nucleic acid form, and are operably linked so as to form a gene that can be transcribed and translated.

A "nucleic acid" is a polynucleotide (DNA or RNA). A "gene" is any nucleic acid sequence encoding a protein or a nascent RNA molecule. A "gene product" is either an as yet untranslated RNA molecule transcribed from a given gene or coding sequence (e.g., mRNA or antisense RNA) or the polypeptide chain (i.e., protein or peptide) translated from the mRNA molecule transcribed from the given gene or coding sequence. Whereas a gene comprises coding sequences plus any non-coding sequences (e.g. promoter, introns, etc.), a "coding sequence" does not include any non-coding (e.g., regulatory) DNA. A gene or coding sequence is recombinant if the sequence of bases along the molecule has been altered from the sequence in which the gene or coding sequence is typically found in nature, or if the sequence of bases is not typically found in nature. According to this invention, a gene or coding sequence can be wholly or partially synthetically made, can comprise genomic or complementary DNA (cDNA) sequences, and can be provided in the form of either DNA or RNA.

Non-coding sequences or regulatory sequences include (but are not limited to) promoter sequences. A "promoter" is a DNA sequence that directs the binding of RNA polymerase and thereby promotes RNA synthesis. "Enhancers" are cis-acting elements of DNA that stimulate or inhibit transcription of adjacent genes. An enhancer that inhibits transcription is also termed a "silencer". Enhancers differ from DNA-binding sites for sequence-specific DNA binding proteins found only in the promoter (which are also termed "promoter elements") in that enhancers can function in either orientation, and over distances of up to several kilobase pairs, even from a position

downstream of a transcribed region. According to the invention, a coding sequence is "operably linked" to a promoter (e.g., when both the coding sequence and the promoter constitute a passenger gene) when the promoter is capable of directing transcription of that coding sequence.

A gene according to the invention (i.e., a gene that results from the operable linkage of nucleic sequences that encode the protein import sequence, cargo region, and optionally, the spacer sequence) can be any gene, and desirably is either a therapeutic gene, a gene that has an immunizing effect upon the cell (e.g., by encoding an antibody fusion), or a gene that includes a reporter coding sequence. Preferably such a gene is capable of being expressed in a cell in which the vector has been internalized. A reporter coding sequence is a nucleic acid sequence that encodes a protein/peptide (or a portion thereof) and can in some fashion be detected in a cell. A therapeutic gene is which exerts its effect at the level of DNA, RNA or protein and can be employed to obtain a therapeutic effect, e.g., in the treatment of a disease, disorder, or condition. The protein encoded by the therapeutic gene also desirably can exert its therapeutic effect by resulting in cell killing. Namely, expression of the gene in itself can lead to cell killing as with expression of the diphtheria toxin A gene, or the expression of the gene can render cells selectively sensitive to the killing action of certain drugs, e.g., expression of the HSV thymidine kinase gene renders cells sensitive to antiviral compounds including acyclovir, gancyclovir and FIAU (1-(2-deoxy-2-fluoro-β -D-arabinofuranosil)-5-iodouracil).

Moreover, the therapeutic gene can exert its effect at the level of RNA, e.g., by encoding an antisense message or ribozyme, a protein which affects splicing or 3′ processing (e.g., polyadenylation), or can encode a protein which acts by affecting the level of expression of another gene within the cell (i.e., where gene expression is broadly considered to include all steps from initiation of transcription through production of a processed protein), perhaps, among other things, by mediating an altered rate of mRNA accumulation, an alteration of mRNA transport, and/or a change in post-transcriptional regulation. Thus, the use of the term "therapeutic gene" is intended to encompass these and any other embodiments of that which is more commonly referred to as gene therapy and is known to those of skill in the art. In particular, preferably a therapeutic gene is one having therapeutic utility.

Method of Intracellular Transfer

By "selective introduction" is meant introduction into a particular cell rather than into another cell. According to the invention, a cell can be any cell, and, preferably, is a eukaryotic cell. A eukaryotic cell is a cell which possesses a nucleus surrounded by a nuclear membrane. Preferably the eukaryotic cell is of a multicellular species (e.g., as opposed to a unicellular yeast cell), and, even more preferably, is a mammalian (optimally human) cell. In particular, the cell is any cell in which a known protein import sequence according to the invention is known to effect (e.g., command, mediate, or facilitate) cell entry, or which has been modified such that the protein import sequence can effect cell entry. Thus, the method especially can be effectively carried out using any cell in which the HIV tat protein import sequence is known to command cell entry, or any cell that has been modified (e.g., modified to contain a cell surface binding protein recognized by the tat protein import sequence) such that the HIV tat sequence enters the cell.

Cells that can be employed thus include, but are not limited to, a wide variety of different cell types such as avian cells, and mammalian cells including but not limited to rodent, primate (such as chimpanzee, monkey, ape, gorilla, orangutan, or gibbon), feline, canine, ungulate (such as ruminant or swine), as well as, in particular, human cells. Desirably such a eukaryotic cell is one in which a chimeric molecule or construct according to the invention is stably maintained, or is at least maintained for a period of time (i.e., typically from anywhere up to three months, and potentially even after three months, including indefinitely) after entry into the cell.

A cell can be present as a single entity, or can be part of a larger collection of cells. Such a "larger collection of cells" can comprise, for instance, a cell culture (either mixed or pure), a tissue (e.g., muscle or other tissue), an organ (e.g., heart, lung, liver, gallbladder, urinary bladder, eye, and other organs), an organ system (e.g., skeletal system, circulatory system, respiratory system, gastrointestinal system, urinary system, nervous system, integumentary system or other organ system), or an organism (e.g., a bird, non-human mammal, human, or the like). Preferably, the organs/tissues/cells being targeted are muscle (e.g., heart and skeletal), spleen, liver, and lung. However, the organs/tissues/cells being targeted can be of the circulatory system (e.g., including, but not limited to heart, blood vessels, and blood), respiratory system (e.g., nose, pharynx, larynx, trachea, bronchi, bronchioles, lungs, and the like), gastrointestinal system (e.g., including mouth, pharynx, esophagus, stomach,

intestines, salivary glands, pancreas, liver, gallbladder, and others), urinary system (e.g., such as kidneys, ureters, urinary bladder, urethra, and the like), nervous system (e.g., including, but not limited to brain and spinal cord, and special sense organs such as the eye) and integumentary system (e.g., skin). Even more preferably, the cells being targeted are those which, without further modification, already are capable of being entered by virtue of the HIV-tat protein import sequence function (i.e., cells which, as isolated from nature, already possess a cell surface binding site capable of interacting with the HIV-tat protein import sequence).

Thus, in particular, a cell into which selective introduction is accomplished according to the invention differs from another cell (in which introduction of an HIV-tat construct is not accomplished) in that the cell so being targeted comprises a particular cell surface binding site for the HIV tat protein import sequence. By "particular cell surface binding site" is meant any site (i.e., molecule or combination of molecules) present on the surface of a cell which provides a site with which the HIV tat protein import sequence can interact to bind the cell as further described herein, and thereby promote cell entry. A particular cell surface binding site therefore encompasses a cell surface receptor, and preferably is a protein (including a modified protein), a carbohydrate, a glycoprotein, a proteoglycan, a lipid, a mucin molecule or mucoprotein, and the like. However, the present method of selectively introducing a HIV tat construct (e.g., including a tat protein import sequence-containing chimeric molecule) into a cell is not limited to any specific mechanism of cellular interaction (i.e., interaction with a given cell surface binding site), and is not to be so construed.

The method by which selective introduction of a HIV-tat containing construct (e.g., a chimeric molecule) into a cell is accomplished comprises contacting the cell with the construct (e.g., chimeric molecule), preferably so as to result in a cell having the construct (e.g., chimeric molecule) therein. Such "contacting" can be done by any means known to those skilled in the art, and described herein, by which the apparent touching or mutual tangency of the cell and the construct (e.g., the chimeric molecule) can be effected. For instance, contacting of the cell and the construct (e.g., chimeric molecule) can be done by mixing these elements in a small volume of the same solution. Alternately, the cell and the construct (e.g., the chimeric molecule) need not necessarily be brought into contact in a small volume, as, for instance, in cases where the construct (e.g., the chimeric

molecule) is administered to a host, and the construct (e.g., the chimeric molecule) travels by the bloodstream or other bodily fluid to the cell in which it selectively binds and enters.

This method preferably is carried out wherein the construct (e.g., the chimeric molecule) comprises either a non-native or native protein import sequence, particularly a HIV tat-2 protein import sequence, as described above, and as further described herein. Thus, the invention provides a method of using a chimeric molecule according to the invention to effect entry into a cell, wherein the method preferably comprises contacting the cell with the chimeric molecule in such a manner that the chimeric molecule is internalized.

The invention further provides host cells transformed with the constructs (e.g., the chimeric molecules) of the invention. The host cell comprising a nucleic acid species (e.g., present in a vector) that encodes a chimeric molecule according to the invention can be employed in a method of making a chimeric molecule comprising a carboxy terminal protein import sequence and an amino terminal cargo region. Namely, preferably this method comprising: (a) obtaining a host cell containing such a nucleic acid species, (b) culturing the host cell, and (c) recovering the chimeric molecule from the culture. These steps are well known in the art, and described in the Examples.

Means of Administration

The method of intracellular transfer of the present invention can be employed to contact cells that are located either *in vitro* or *in vivo*. According to the invention "contacting" comprises any means by which a product is introduced intracellularly; the method is not dependent on any particular means of introduction and is not to be so construed. Means of introduction are well known to those skilled in the art, and also are exemplified herein.

Accordingly, introduction of the products of the invention (e.g., vectors, compositions, and/or chimeric molecules) can be effected, for instance, either *in vitro* (e.g., in an *ex vivo* type method of gene therapy or in tissue culture studies) or *in vivo* by electroporation, transformation, transduction, conjugation or triparental mating, (co)transfection, (co-)infection, membrane fusion with cationic lipids, high velocity bombardment with DNA-coated microprojectiles, incubation with calcium phosphate-DNA precipitate, direct microinjection into single cells, and the like.

Similarly, the products can be introduced by means of cationic lipids, e.g., liposomes. Such liposomes are commercially available (e.g., Lipofectin®, Lipofectamine™, and the like, supplied by Life Technologies, Gibco BRL, Gaithersburg, Md., and other commercial vendors). Other methods also are available and are known to those skilled in the art.

One skilled in the art will appreciate that suitable methods of administering a product of the present invention to an animal (e.g., a human) for purposes of gene therapy (see, for example, Rosenfeld et al., *Science,* 252, 431-434 (1991); Jaffe et al., *Clin. Res.,* 39(2), 302A (1991); Rosenfeld et al., *Clin. Res.,* 39(2), 311A (1991)), chemotherapy, cell marking, and vaccination (i.e., effecting intracellular immunity) are available, and, although more than one route can be used for administration, a particular route can provide a more immediate and more effective reaction, or a more convenient or less invasive means, than another route.

Pharmaceutically acceptable excipients also are well-known to those who are skilled in the art, and are readily available. The choice of excipient will be determined in part by the particular method used to administer the product according to the invention. Accordingly, there is a wide variety of suitable formulations for use in the context of the present invention. The following methods and excipients are merely exemplary and are in no way limiting.

Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the compound dissolved in diluents, such as water, saline, or orange juice; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as solids or granules; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, mannitol, corn starch, potato starch, microcrystalline cellulose, acacia, gelatin, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, stearic acid, and other excipients, colorants, diluents, buffering agents, moistening agents, preservatives, flavoring agents, and pharmacologically compatible excipients. Lozenge forms can comprise the active ingredient in a flavor, usually sucrose and acacia or tragacanth, as well as pastilles comprising the active ingredient in an inert base; such as gelatin and glycerin, emulsions, gels, and the like containing, in addition to the active ingredient, such excipients as are known in the art.

A product of the present invention, alone or in combination with other suitable ingredients, can be made into aerosol formulations to be administered via inhalation.

These aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like. They also can be formulated as pharmaceuticals for non-pressured preparations such as in a nebulizer or an atomizer.

Formulations suitable for parenteral administration include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain anti-oxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid excipient, for example, water, for injections, immediately prior to use. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described.

Additionally, a product of the present invention can be made into suppositories by mixing with a variety of bases such as emulsifying bases or water-soluble bases.

Formulations suitable for vaginal administration can be presented as pessaries, tampons, creams, gels, pastes, foams, or spray formulas containing, in addition to the active ingredient, carriers such as are known in the art to be appropriate.

The dose administered to an animal, particularly a human, in the context of the present invention will vary with the coding sequence or antibody appended to the protein import sequence, the composition employed, the method of administration, the purpose of administration, and the particular site and organism being administered to. However, preferably a dose corresponding to an effective amount of a product (e.g., a product containing a protein import sequence-comprising construct according to the invention) is employed. An "effective amount" is one that is sufficient to produce the desired effect in a host, which can be monitored using several end-points known to those skilled in the art. For instance, one desired effect is nucleic acid transfer inside of a host cell, followed by intracellular expression of one or more sequences contained in the construct. Such transfer can be monitored by a variety of means, including, but not limited to, a therapeutic effect (e.g., alleviation of some symptom associated with the disease, condition, disorder or syndrome being treated), or by further evidence of the transferred gene or coding sequence or its expression within the host (e.g., using the polymerase chain reaction,

Northern or Southern hybridizations, or transcription assays to detect the nucleic acid in host cells, or using immunoblot analysis, antibody-mediated detection, or particularized assays to detect protein or polypeptide encoded by the transferred nucleic acid, or impacted in level or function due to such transfer). Particularized assay are described in the Examples which follow. These methods described are by no means all-inclusive, and further methods to suit the specific application will be apparent to the ordinary skilled artisan.

Generally, to ensure effective transfer of the products of the present invention, it is preferable that sufficient transfer be achieved such that antibodies (e.g., Fabs) can be introduced into cells at a level such that sufficient affinity is obtained to permit intracellular activity. For instance, desirably transfer will result from in an intracellular concentration of antibody (or fusion protein) of about 2 x 10^{-9} M, although an intracellular concentration from about 2 x 10^{-6} M to about 2 x 10^{-12} M is contemplated by the present invention. However, this is a general guideline which by no means precludes use of a higher or lower amount, as might be warranted in a particular application, either *in vitro* or *in vivo*. For example, the actual dose and schedule can vary depending on whether the composition is administered in combination with other pharmaceutical compositions, or depending on interindividual differences in pharmacokinetics, drug disposition, and metabolism. Similarly, amounts can vary in *in vitro* applications depending on the particular cell type utilized or the means by which the product is transferred. One skilled in the art easily can make any necessary adjustments in accordance with the necessities of the particular situation.

The general toxicity of antibody within any particular cell type and potential non-specific interactions of antibodies with normal cellular proteins are possible problems that have been encountered with other antibody usages, and should be considered using the invasive antibody approach of the invention. However, the studies carried out to date with intracellularly expressed antibodies give no indication that this will be major problems in the method of the invention (see, e.g., Chen et al., *Proc. Natl. Acad. Sci.,* 91, 5932-5936 (1994); Duan et al., *Proc. Natl. Acad. Sci.,* 91, 5075-5079 (1994); Mhashilkar et al., *EMBO J.* 14, 1542-1551 (1995); Maciejewski et al., *Nature Medicine* 1, 667-673 (1995); Marasco et al., *Proc. Natl. Acad. Sci. USA* 90, 7889-7893 (1993)). Toxicity problems have not been reported for the intracellular expression of antibodies against three different

antigens. Those studies also do not report problems with non-specific antibody reactivities. Given the expected high affinities and exquisite specificities of the preferred antibodies according to the invention, it is not anticipated that difficulties will be obtained from antibody cross-reactivity. In addition, no marked alterations in cell viability were observed after two weeks of chronic treatment of human fibroblasts with invasive test Fabs (see, Examples). Cell viability and cell-toxicity desirably can be monitored as described in the Examples.

EXAMPLES

The following examples further illustrate the present invention but, of course, should not be construed as in any way limiting its scope.

Example 1: Obtaining Human Fabs to hCMV

Human anti-hCMV Fab fragments were derived from phage libraries generated from immune donors. To prepare Fab fragments to a chosen antigen requires that the library donor has a serum titer to that antigen. In CMV mononucleosis, specific anti-viral IgG antibodies appear 2-3 weeks after the onset of symptoms, peak within 1-2 months and persist for years (Alford et al., "Cytomegalovirus". *In*: Virology, second edition, (Fields, B.N., Knipe, D.M. et al, eds.), New York: Raven Press, Ltd., pp1981-2010 (1990); Landini et al., *Prog. Med. Virl.* 35, 152-185 (1988); Rasmussen et al., *Current Topics in Microbiology and Immunology.* 154, 221-254 (1990)). The same pattern of antibody production occurs in subclinical infection but the response is quantitatively less than seen with symptomatic infection, at least in the early phase of the response. Both structural and non-structural hCMV proteins elicit strong humoral responses. The best characterized responses to non-structural proteins are to the major immediate-early proteins (IE72 (UL123) and IE86 (UL122)), the DNA polymerase accessory protein UL44 and the major DNA binding protein UL57 (Hayes et al., *J. Infect. Dis.* 156, 615-621 (1987); Gold et al., *J. Med. Virol.* 25, 237-243 (1988); Van Zanten et al., *Clin. Exp. Immunol.* 83, 102-107 (1991); Shiraki et al., *J. Med. Virol.* 34, 280-283 (1991); Mazeron et al., *J. General Virol.* 73, 2699-2703 (1992)). It is likely that human antibodies to further non-structural proteins could be characterized with the availability of purified antigen.

The human monoclonal Fabs were cloned from combinatorial libraries on the

surface of phage (Barbas et al., *Proc. Natl. Acad. Sci.* 88, 7978-7982 (1991); Burton et al., *Adv. in Immunol.*, (1994)). Antibody-producing cells, e.g., bone marrow and peripheral blood lymphocytes were isolated, and RNA was prepared and reverse transcribed into cDNA. Heavy (Fd portion) and light chains were then amplified using the PCR reaction. This provided the genetic information necessary to produce the Fab antigen binding fragment of the antibody molecule. The heavy and light chain DNA were then cut with restriction enzymes and cloned sequentially into a phage display vector. This vector displays an Fab molecule monovalently on the surface of the phage while containing the antibody genes on the inside. This recapitulates the main features of the B cell in a natural immune response. The library comprises an ensemble of phage-Fabs, and the typical library size is similar to that of natural antibody repertoires ($\geq 10^8$).

Specific antibodies were selected from the library making use of the displayed Fab in a process termed "panning". Typically the library contained 10^{12} phage-Fabs (about 10^4 copies of each member of a 10^8 initial library) in about 50 µl of solution. This solution was layered onto antigen immobilized on plastic, incubated, and then non-specifically bound phage were removed by repeated washings. The remaining phage, greatly enriched for specific phage-Fabs, were then eluted, e.g., by adding acid or excess soluble antigen. This process, a round of panning, generally produced an enrichment factor for specific phage of the order of 100. The panning was then repeated, usually 3 or 4 times, to reveal specific phage occurring only singly in the initial library, i.e., with frequencies of the order of 1 in 10^7.

Finally, the specific phage-Fabs were converted to plasmids expressing soluble Fabs by excision of the gene which otherwise directs the Fab to the phage surface, and religation. The plasmids were used to transform bacteria. Individual clones, carrying one plasmid and one set of antibody genes, were then grown up. The supernates or periplasmic cell extracts from each clone contained Fab, which was screened for reactivity with antigen, usually using a conventional ELISA assay.

Two libraries were constructed from HIV-1 seropositive individuals with antibody titers against hCMV. These libraries were constructed in pComb3 phagemids by isolating and restricting genomic DNA, size fractionating fragments, and ligating into restricted pComb3 using standard protocols. These libraries were

panned against lysates of hCMV-infected cells. Of 25 clones examined from a single library, 6 different heavy- and light chain sequence pairings were observed. These clones recognized infected cells, and were characterized with use of immunoprecipitation and Western blotting assays. One clone (i.e., Fab 11) immunoprecipitated a novel banding pattern. Of the remaining clones, two were found to recognize the 65 kDa lower matrix phosphoprotein, UL83, and three (i.e., Fabs GL5, 14 and 34) recognized the polymerase-associated protein (replication factor), UL44. The amino acid sequence of the essential heavy chain regions of the three anti-UL44 Fabs is shown below in Table 1.

Both of the UL83 and UL44 antigens are strongly immunogenic in humans, and are present in large quantities in viral lysates, so it was not surprising that antibodies with these specificities were recovered. The antibodies were able to strongly compete with donor serum for binding to viral lysate, confirming their high serum levels in the individuals from whom the libraries were constructed.

Table 1. Partial Heavy Chain Amino Acid Sequences Of hCMV UL44 Specific Fabs.

FAB	FRAMEWORK 3	CDR3	FRAMEWORK 4
GL CMV 14	RVTMTRDTSINTV YMELSRLRSDDTA VYWCAR [SEQ ID NO: 8]	EYRQPHNSAWSDD AFDI [SEQ ID NO: 9]	WGQGTMVTVSS [SEQ ID NO: 10]
GL CMV 5	RITISRDNSKNTLFL QMNSLRAEDTAIY YCAK [SEQ ID NO: 11]	DWYGDYGDAFDV [SEQ ID NO: 12]	WGQGTMVTVSS [SEQ ID NO: 10]
GL CMV 34	RVTMTRDTSLNTV FMELISLNLGDTAV YFCAR [SEQ ID NO: 13]	GDRQFGFRGTDWF DP [SEQ ID NO: 14]	WGQGTLVVVSS [SEQ ID NO: 15]

Example 2: Engineering an anti-hCMV Fab for Intracellular Targeting

A number of fusion constructs were created in which tat sequences were linked to an anti-UL44 recombinant Fab (GL14), as described in Example 1. The anti-UL44-tat fusion construct was prepared using the pComb3 phagemid in which the antibody was originally cloned and isolated (Barbas et al., *Proc. Natl. Acad. Sci.*, 89, 10164-10168 (1992)).

For one construct, a tat nucleic acid sequence that encodes tat peptide residues 38-72 [SEQ ID NO: 5] was PCR amplified and cloned into pComb3 via a unique *Xba*I site present at the carboxy terminal end of the antibody light chain sequence. A short peptide linker (i.e., Ser Gly Gly Gly) was inserted between the antibody and tat sequence to avoid steric interference with cysteine residue 214 of the light chain, which avoidance is crucial to the correct folding of the Fab molecule.

A second tat-Fab construct was constructed in which a nucleic acid sequence corresponding to tat amino acid residues 47-59 [SEQ ID NO:6] was fused, again via the short peptide linker, to the C- terminal ends of both the light and heavy antibody chains, affording bivalent tagging of tat peptide on a single Fab molecule. In this construct, the sequence corresponding to tat residues 47-59 was inserted via *Nco*I sites occurring in pComb3 at the C terminal sequence of each of the antibody chains. In the pComb3 system, following translation, the light and heavy (Fd) antibody chains are directed independently to the periplasm of the bacterium, where the light and heavy chain assemble to form a Fab. Antibody is then released from the cell by breaking the outer membrane of the bacterium, usually through a freeze/thawing procedure. Antibody-tat (i.e., "tat-tagged Fab") is then purified to homogeneity, if desired, for instance, by affinity chromatography. The antibodies then can be employed for intracellular delivery.

Example 3: The anti-hCMV Fab Binds the Cell Surface

It was investigated whether the tat-tag confers enhanced binding to the cell surface. Initial experiments were carried out with use of HeLa cells. For these experiments, biotinylated Fabs, control Fabs, and tat-tagged Fabs, each at a concentration of about 2ng/microliter, were incubated on ice for 20 minutes with 10^5 cells. The cells were then extensively washed to remove unbound ligand. The amount of bound Fab was quantitated by ELISA assay using HRP-conjugated streptavidin. The results of these studies are depicted in Fig. 1, which confirms the selective binding of the tagged Fab relative to the untagged version. Separate experiments confirmed that the majority of the cell surface binding occurs within 10 minutes.

Studies carried out to date with intracellularly expressed antibodies give no indication that there should be any problem with toxicity or non-specific interaction with normal cellular proteins (Duan et al., *Proc. Natl. Acad. Sci.* 91, 5075-5079

(1994); Maciejewski et al., *Nature Medicine* 1, 667-673 (1995); Marasco et al., *Proc. Natl. Acad. Sci,* 90, 7889-7893 (1993)). Also, no marked alterations in cell viability was observed after two weeks of chronic treatment of human fibroblasts with the test invasive Fabs.

Example 4: Internalization and Sub-Cellular Localization of the anti-hCMV Fab

To determine internalization and the sub-cellular localization of tat-Fab, direct visualization of FITC-conjugated tat-Fab was examined in living cells as these events occurred with use of live cell imaging. The cells were incubated with Fabs external to the cell and directly imaged at various times post-administration. In particular, healthy glial U373 cells imaged at about 20 minutes post-treatment with tat-tagged Fab contained the construct internal to the cell. See Fig. 2A, phase contrast image; Fig. 2B, fluorescence image. By contrast the control Fab without the tat transporter failed to target intracellularly.

In order to facilitate the subcellular localization of the original unmodified GL14 Fab and its tat fusion derivatives, additional constructs have been prepared in which the gene encoding the green fluorescent protein (GFP) from the bioluminescent jellyfish *Aequorea victoria*, is fused to the C-terminus of the antibody light chain at a unique *XbaI* site. The GFP emits green light maximally at 509 nm when excited with UV light at 395 nm. Bacteria that express Fab-GFP fusions appear green when exposed to UV light, and the purified Fab-GFP protein appears at its predicted size of 75 kDa following SDS-PAGE.

Example 5: *In vitro* Assays for hCMV IE86 Protein Function

This example describes *in vitro* assays that faithfully reconstitute functional activity by purified recombinant IE86 protein (e.g, modified from Jupp et al., *J. Virol.* 67, 7539-7546 (1993) and Wu et al., *J. Virol.* 67, 7547-7555 (1993)). These assays include various DNA-binding and protein-protein (as set out in Figs 3, 4, and 5) interaction assays and *in vitro* transcription systems (as set out in Figs 6 and 7), and provide ideal screening procedures for selecting inhibitory Fabs.

Namely, for these studies, the cDNA encoding IE86 was expressed in *E. coli* under the control of an IPTG-inducible promoter. To facilitate purification, a short linker sequence encoding six histidine residues was cloned onto the 5' end of the cDNA. Under non-denaturing conditions, recombinant proteins can be purified to

approximately 90% purity by nickel chelate chromatography. Accordingly, Fig. 3 shows a reproduction of a Coomassie-stained gel showing bacterially expressed wild-type IE86 (250 ng) protein following dialysis of the pooled fractions after nickel chelate chromatography. Fig. 4 shows a reproduction of a DNase I protection analysis assay of the IE86 binding site (crs) in the hCMV major immediate early promoter by recombinant IE86 protein, with Lanes 1 and 4 comprising controls showing cleavage patterns in the presence of 250 nM BSA, and Lanes 2 and 3 showing the footprinting pattern in the presence of 50 and 250 nM IE86 protein, respectively. In Fig. 4, the positions of the nucleotides, indicated by Maxam and Gilbert sequencing of the labeled probes, are shown relative to the transcription start site (arrow), and the protected crs (open box) is indicated.

By comparison, Fig. 5 shows an electrophoretic mobility shift assay (EMSA) revealing protein:DNA and protein:protein interactions. Lane 1 of Fig. 5 shows a nucleoprotein complex (D) developed with the CMV TATA box probe and human TBP. The presence of hTBP in complex D is confirmed by using a monoclonal antibody to hTBP that supershifts the complex (Lane 2). The association via protein:protein interactions of IE86 with TATA box bound hTBP is indicated by a change in the mobility of the complex (D86) (lane 3). The presence of hTBP and IE86 in the D86 complex is confirmed by using monoclonal antibodies to hTBP (Lane 4) and IE86 (Lane 5) which result in super-shifting the D86 complex. The reason for the partial disruption of the D86 complex by the IE86 monoclonal antibody is due to steric interference, since the epitope for the antibody occurs within the TBP contacting domain of IE86.

Fig. 6 confirms the activity of recombinant IE86 in a highly purified reconstituted *in vitro* transcription system. Fig. 6 shows the requirement for general transcription factors (indicated by a + sign) for initiating transcription from DNA templates containing the major immediate early promoter of hCMV (CMV-MIEP) and the adenovirus major late promoter (Ad-MLP).

Using the same *in vitro* transcription system in Fig. 7, Lanes 1-3 represent transcription from the wild-type MIEP in the presence of increasing amounts of IE86, with Lanes 2 and 3 showing repression of transcription. Lanes 4-6 are transcription reactions with a MIEP template in which the IE86 binding site (crs) has been mutated. Under identical conditions to those reactions shown in lanes 2 and 3, Lanes 5 and 6 reveal the requirement for IE86 to interact with the crs element in

order to effect inhibition. This requirement is confirmed by using a mutant IE86 protein that is defective in DNA-binding and is also incapable of effecting inhibition of transcription from the wild-type MIEP template (Lanes 7-9).

Example 6: *In vitro* assays for hCMV UL44 protein function

This Example describes *in vitro* assays for hCMV UL44 protein function. For these studies, UL44 was cloned into a 6His expression vector and purified from *E. coli* over a Ni^{2+}-NTA column. To test the DNA binding ability of purified UL44, the purified protein was bound to a ssDNA cellulose column (Sigma) in Buffer A with 60 mM KCl (25 mM HEPES pH 8.0, 0.5 mM EDTA, 2mM 2-mercaptoethanol, 10% glycerol, and 0.1% NP40) and the majority of UL44 was eluted at 400 mM KCl as determined by Western analysis of the column fractions (Figs 8A-C). Interestingly, a degraded fragment of UL44 (~30 kDa) was also able to stably bind the ssDNA column and was eluted at the same KCl concentration as the full-length protein.

To map the domain of UL44 responsible for DNA binding, the protein was subjected to thrombin cleavage and bound to the DNA cellulose column. The largest fragment left after protease digestion is an N-terminal 28 kDa peptide that flows through the column. This information, combined with previous mutagenesis studies of UL44 (Weiland et al., *Virus Research* 34, 191-206 (1994)), indicates that amino acid residues located between 274 and 310 are sufficient for DNA binding (Fig. 8B). Three recombinant Fabs (GL 14, GL 5, and GL 34), selected for their ability to recognize UL44, were tested for their capacity to inhibit DNA binding by the hCMV protein. Each of the three Fabs were incubated (in 10 molar excess) with UL44 before loading the reactions onto a ssDNA column. None of the recombinant Fabs were able to change the elution profile of the column, indicating that the antigenic moiety of UL44 recognized by the Fabs is probably located outside of the DNA binding domain (Fig. 8C).

In particular, Fig. 8A shows the elution profile of purified UL44 bound to a 0.2 mL ssDNA cellulose column as determined by Western blotting of the 0.5 mL fractions. The concentrations of KCl used to elute UL44 from the column(s) are indicated above panel A with the location of molecular weight standards used for Western blotting shown on the right of each panel. UL44 was subjected to thrombin cleavage and then loaded onto a ssDNA cellulose column with a elution profile as

shown in Panel B. The largest fragment of the protein, a 274 amino acid N-terminal peptide has much reduced affinity for DNA and is completely eluted from the column at 0.1 M KCl. Panel C shows the elution of UL44 from the ssDNA column after being incubated with an excess of a recombinant Fab (GL 5) selected for its ability to recognize UL44. The Fab does not appear to alter the ability of UL44 to associate with DNA. The three thrombin cleavage sites in UL44 are shown in Fig. 9.

Example 7: Incorporating DNA Binding Motifs into Human Antibodies

This example describes means for incorporating DNA binding motifs into human antibodies.

The generation of monoclonal antibodies to DNA by conventional immunization procedures has not been successful because of the very poor ability of double-stranded (ds) DNA to elicit an immune response in normal animals. A novel approach thus was employed in which DNA-binding antibodies were selected from semi-synthetic libraries expressed on the surface of phage. The semi-synthetic antibody libraries were constructed from a single human anti-tetanus toxoid binding antibody by randomizing over small defined regions which generally contact antigen, the complementarity determining regions (CDRs). These regions were completely randomized using synthetic oligonucleotides with a "NNS doping strategy", where N is any of the 4 nucleotides, and S is either G or C. Accordingly, NNS encodes all 20 amino acids and a single amber stop codon in a total of 32 codons.

Three combinatorial Fab libraries were prepared on the surface of phage using the pComb 3 system. Each library used the same heavy chain with the CDR3 of length 10 (library 1) or 16 amino acids (libraries 2 and 3) being randomized. The CDR3 of the heavy chain was chosen for randomization as it forms a loop which varies in sequence and length and generally provides the most significant contribution to binding antigen. The library sizes were each of the order of 108. The libraries were taken through rounds of panning against human placental double stranded DNA.

Two Fabs were isolated which bound double and single-stranded DNA with good affinity, approximately 10^7-10^8 M^{-1}. The Fabs did not bind the negatively-charged polyelectrolyte dextran sulfate or lipid A. The ability of one of the Fabs to perform as a naturally occurring DNA binding protein was investigated in electrophoretic mobility shift assays. The ability to form nucleoprotein complexes was clearly demonstrated (see, Fig. 10). The two Fabs isolated contained heavy chain

CDR3s (HCDR3s) of 10 and 16 residues in length. Interestingly, though the CDRs differed in length, the amino and carboxy ends of the HCDR3 were virtually identical. The clone with the 16 residue HCDR3 demonstrated a clear preference for poly(dGdC).poly(dGdC) relative to poly(dAdT).poly(dAdT) whereas the other clone bound the different oligonucleotides with similar affinity (data not shown).

These experiments thus confirm that DNA binding could be generated in an antibody with the appropriate heavy chain CDR3. Furthermore, the antibody framework was shown to be permissive for DNA binding. Accordingly, the experiments validate that antibodies against DNA can be employed for intracellular targeting/immunization. Portions of this work were published as Barbas et al., J. Am. Chem. Soc. 116, 2161-2162 (1994).

Example 8: Construction of an Antibody Containing a bHLH motif

This example describes the introduction of a specific DNA binding motif into the heavy chain CDR3 of a prototype antibody.

The basic helix-loop-helix (bHLH) motif of the TFEB protein was employed for these experiments since: (1) the basic binding region is relatively short (of the order of 25 amino acids) and can be accommodated in the heavy chain CDR3; (2) detailed crystallographic structural information on the DNA-bHLH interaction is available (Ferré-D'Amaré et al., *Nature* 363, 38-45 (1993); Ferré-D'Amaré et al., *EMBO J.*13, 180-189 (1994)); (3) alanine scanning mutagenesis has defined the amino acid residues in TFEB crucial for DNA binding (Fisher et al., *Cell*, 72, 467-476 (1993)); and (4) the TFEB recognition sequence has been well defined (Carr et al., *Molecular and Cellular Biology* 10, 4384-4388 (1990)). Tight binding is observed to the sequence CACGTG as found in the adenovirus major late promoter (MLP). Thus, the strategy was to try and mimic, in part, TFEB specificity in a phage-displayed antibody carrying the basic motif. The motif then desirably can be randomized in key positions and the resulting library panned against new DNA sequences to try and generate new specificities.

The heavy chain CDR3 of the prototype Fab was replaced by the basic domain of amino acid sequence GGAAKKAAHAAAERRRRAAINGGAAGG [SEQ ID NO:17] using overlap PCR cloning. Competition EMSAs were performed with double-stranded MLP and MyoD as competitors to compare class A bHLH versus class B bHLH DNA recognition sequences. The data are presented as a Hofstee plot, depicted in Fig. 11, where the slope of the line is the IC_{50} (nM) and the slopes can be

directly compared (Ferré-D'Amaré et al., *Nature* 363, 38-45 (1993); Ferré-D'Amaré et al., *EMBO J.*13, 180-189 (1994)). The IC_{50} and correlation coefficients (in parentheses) of the linear regression for these data were as follows: O, TFEB/MLP, 10.9 nM (0.98); and Δ, Fab-Ebox, 4.3 nM (0.96); Δ TFEB/MyoD, 10.9 nM (0.98); and Δ, Fab-Ebox/MyoD, 91.9 nM (0.94). The results of three IC_{50} determinations for each of the competing double-stranded oligodeoxyribonucleotides are summarized in Table 1.

The apparent dissociation constants were determined from competition EMSAs performed with double-stranded MLP as the competitor (data presented in Fig. 12). The apparent K_d values were determined by using double-reciprocal plots (1/[bound DNA] versus 1/[free DNA], where the K_d is the slope divided by the intercept of the axis 1/[bound DNA]). The reciprocal of the intercept is the apparent "B_{max}" of the EMSA and reflects the concentration of protein actively binding MLP under these conditions (5 pM and 15 pM for TFEBΔ265 and Fab-Ebox, respectively). In the experiment shown in Fig. 12, the apparent K_d values for TFEBΔ265 and Fab-Ebox were 2 nM and 24 nM, respectively. In this analysis, the apparent K_d values are tentative estimates of the true equilibrium constants.

This TFEB variant, containing alanines at a number of positions, was preferred over native TFEB because of its higher affinity for DNA. The resultant Fab, Fab-Ebox, was shown by competition EMSA to have an apparent affinity approximately 10-fold lower than TFEB for an MLP probe but still in a range consistent with high-affinity binding (apparent Kd of approximately 20 nM), as evidenced in Figs 11 and 12. Remarkably, although less sequence specific than the parent TFEB protein, the Fab-Ebox was found to interact with the major groove of the DNA helix and discriminate in its recognition of the CACGTG motif in a sequence-dependent manner.

These results confirm that the modular nature of immunoglobulins and transcription factors can be exploited to create unique DNA-binding proteins by molecular design. Portions of this work were published as McLane et al., *Proc. Natl. Acad. Sci.* 92, 5214-5218 (1995).

Example 9: Examination of the DNA Binding Properties of the Fab-Ebox

This example describes the DNA binding properties of the recombinant Fab-Ebox, assessed using the DNase I footprinting method to determine in greater detail

the actual binding site(s) of the Fab fragment, and the molecular details of the DNA binding site that govern its interactions with the Fab-Ebox.

For these studies, the Fab-Ebox was bound to a construction of the adenovirus major late promoter (Ad MLP), and the pattern of protection from DNase I was compared to TFEB. Increasing concentrations of TFEB and Fab-Ebox were added to ~1 ng of end labeled DNA fragments containing Ad MLP sequences from -72 to +10. To aid interpretation of results, a Maxam and Gilbert sequencing reaction for G residues was run on a separate lane of the same footprinting gel, as was a lane comprising DNase reactions without added protein. The amounts of TFEB used for assessing protection of the bottom strand of the MLP were 0.2, 0.4, 0.8, and 1.6 micrograms, while the amounts of Fab-Ebox used (i.e., in separate reactions) were 0.3, 0.7, 1.4, and 2.9 micrograms. The amounts of Fab-Ebox used for assessing protection of the bottom strand of the MLP were 0.7, 1.4, and 2.2 micrograms of Fab-Ebox. Binding reactions with TFEB were done in the presence of 50 micrograms/mL of poly (dI-dC), while reactions with Fab-Ebox did not contain poly (dI-dC).

The obtained data confirmed TFEB and Fab-Ebox protection of the bottom strand of the Ad MLP (results not shown). Unlike TFEB, which protected a single large site of ~17-19 bases on both strands of the Ad MLP around the E box sequence, Fab-Ebox binding to the same fragment resulted in protection of two smaller sites of about ~7-9 bases, including a sequence upstream of the Ebox (-64 to -74) that was not protected by TFEB binding. This is set out in Figs 13 and 14, respectively, which are schematic representations of the protection patterns of both proteins, with arrows indicating the positions of DNase I hypersensitive sites. The E box, CACGTG, is indicated by the double-headed arrow between positions -55 and -60.

Interestingly, the sequence was completely occupied by Fab-Ebox at the lowest concentration tested, indicating an enhanced affinity of binding in comparison with the E-box site. Portions of this work were published as Le Blanc et al., *Biochemistry* 37 (17), 6015-22 (1998).

Example 10: Examination of the DNA Binding Properties of the Fab-Ebox

This example describes an examination of the binding of the Fab-Ebox to other sequences outside of the E box to try to understand the differences in DNA binding between the parental TFEB transcription factor and the Fab.

The DNase I protection patterns of Fab-Ebox on a variety of DNA sequences other than the E box and MLP sequences was examined to identify a consensus binding site for Fab-Ebox. Ten sites protected from DNase I digestion by the binding of Fab-Ebox were aligned to determine the common features of these sequences. The downstream portion of the E box appears to be the prime determinant for Fab recognition as the internal TG(R) sequence is constant for each binding site examined. The sites that have been identified as are result of these studies are summarized in Fig. 15. The sites have been aligned along homologous sequences and the internal sequence of TGRR appears to be the most conserved portion of the Fab-Ebox binding site.

These studies thus confirm the feasibility of generating Fabs capable of sequence-specific recognition of DNA. The specificity of Fab-Ebox can be altered for binding to related and distinct target sequences using the methods described above. Portions of this work were published as Le Blanc et al., *Biochemistry* 37 (17), 6015-22 (1998).

Example 11: Use of Vertebrate Transgenic Animals

This example describes the protocol for use of vertebrate transgenic animals according to the invention, for instance, for examining the *in vivo* biodistribution and potential delivery activity of the tat-Fabs, for investigating tissue-specific expression of human CMV, for generation of transgenic animals containing a high affinity binding site for a Fab in place of the viral crs, and the like. Alternately, adult animals can be employed in certain of these studies.

Both C57BI/6 (BALB) transgenic and CD-1 outbred strains can be employed for the mating experiments. Approximately 25-30 mice of each strain are utilized in these experiments. Transgenic animals can be generated by microinjecting foreign DNA into the male pronucleus of fertilized eggs and subsequent reimplantation into pseudo-pregnant mothers, as described in the Cold Spring Harbor Laboratory manual "Manipulating a Mouse Embryo".

In order to obtain optimal numbers of unfertilized eggs from BALB females, 3-5 week old mice can be induced to superovulate by intraperitoneal injection of 5 IU of pregnant mare's serum (PMS) followed 48 hours later with an intraperitoneal injection of 6 IU of human chorionic gonadotropin. The BALB females are immediately placed with C57BI/6 fertile males overnight and the females checked for a copulation plug the following morning.

Females with fertilized eggs are sacrificed by cervical dislocation and their eggs collected from the oviducts and the cumulus cells removed by treatment with hyaluronidase. The isolated fertilized eggs are allowed to mature in Brinsters medium at 37°C and 5% CO_2. The male pronucleus of the fertilized egg is injected with 200 copies of the transgene.

Pseudo-pregnant female recipients (BALB) of the fertilized eggs are prepared by placing vasectomized C57BI/6 males in female cages overnight and checking for copulation plugs. The vasectomized males are prepared from adult mice. This procedure is performed by anesthetizing the male mouse with 0.015-0.017 ml of 2.5% Avertin per gram of body weight. A 1.0 cm transverse incision is made on the left and right flank of the body wall and the testicular fat pads, vas deferens and testes are pulled out. The vas deferens are cut, cauterized, and the testes placed back in the male and the incision closed with an autoclip. Vasectomized males are checked for sterility and about 10-20 males are maintained at any one time.

Six to eight week old pseudo-pregnant females (BALB) are prepared as recipients of the fertilized injected eggs by anesthetizing the mice with Avertin as described above. A 1 cm incision is made on the left and right flank of the body wall and the ovary and oviduct are exposed. The left and right bursa are opened and 15-20 microinjected eggs each are transfered to the ampulla with a transfer pipette. The incision is closed with a single autoclip.

Transgenic offspring are analyzed at 2-3 weeks of age to determine whether integration of viral DNA has occurred by analysis of tail DNA. Approximately 200 microinjected eggs produce 40-50 live mice, of which 10-20 will have integrated DNA and 5-10 will express the transgene. Those mice not expressing the transgene are sacrificed by exposure to CO_2. Expressing transgenic lines are maintained by breeding with the mouse species mentioned above.

All of the references cited herein are hereby incorporated in their entireties by reference.

While this invention has been described with an emphasis upon a preferred embodiment, it will be obvious to those of ordinary skill in the art that variations in the preferred composition and method may be used and that it is intended that the invention may be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the following claims.

CLAIMS

What is claimed:

1. A chimeric molecule comprising a carboxy terminal protein import sequence and an amino terminal cargo region.

2. The chimeric molecule of claim 1, wherein said cargo region is a moiety selected from the group consisting of an antibody, a non-antibody peptide, and a non-proteinaceous molecule.

3. The chimeric molecule of claim 2, wherein said cargo region is an antibody selected from the group consisting of an anti-UL44 antibody, an anti-UL83 antibody, and an anti-UL122 antibody.

4. The chimeric molecule of claim 2, wherein said cargo region is an antibody that is a Fab selected from the group consisting of Fab GL14, Fab GL5, and Fab GL34.

5. The chimeric molecule of claim 1, wherein said protein import sequence is selected from the group consisting of an

HIV tat peptide comprising tat residues 37-72 (SEQ ID NO:4);
HIV tat peptide comprising tat residues 38-72 (SEQ ID NO:5);
HIV tat peptide comprising tat residues 47-59 (SEQ ID NO:6); and
HIV tat peptide comprising tat residues 49-57(SEQ ID NO:7).

6. The chimeric molecule of claim 4, wherein said protein import sequence is an HIV tat peptide comprising tat residues 38-72 (SEQ ID NO:5).

7. The chimeric molecule of claim 4, wherein said protein import sequence is an HIV tat peptide comprising tat residues 47-59 (SEQ ID NO:6).

8. The chimeric molecule of claim 1, wherein said protein import sequence is selected from the group consisting of a lactoferrin and an antennapedia peptide.

9. The chimeric molecule of claim 1, wherein said protein import sequence is a Drosophila antennapedia peptide comprising from about 16 to about 20 amino acids.

10. The chimeric molecule of claim 1, wherein said chimeric molecule further comprises a spacer sequence located between said protein import sequence and said cargo region.

11. The chimeric molecule of claim 9, wherein said spacer sequence comprises from about 1 to about 50 amino acid residues.

12. A composition comprising the chimeric molecule of any of claims 1-11 and a pharmaceutically acceptable excipient.

13. The use of the chimeric molecule of any of claims 1-11 for the manufacture of a medicament for the treatment of viral infection.

14. A method of using the chimeric molecule according to any of claims 1-11 comprising contacting said cell with said chimeric molecule wherein the chimeric molecule is internalized.

15. An isolated and purified nucleic acid molecule that encodes the chimeric molecule of any of claims 1-11.

16. A vector comprising the nucleic acid molecule of claim 15.

17. A host cell comprising the chimeric molecule of any of claims 1-11.

18. A host cell comprising the vector of claim 16.

19. A method of making a chimeric molecule comprising a carboxy terminal protein import sequence and an amino terminal cargo region, said method comprising:

(a) obtaining the host cell of claim 17,

(b) culturing said host cell, and

(c) recovering said chimeric molecule from said culture.

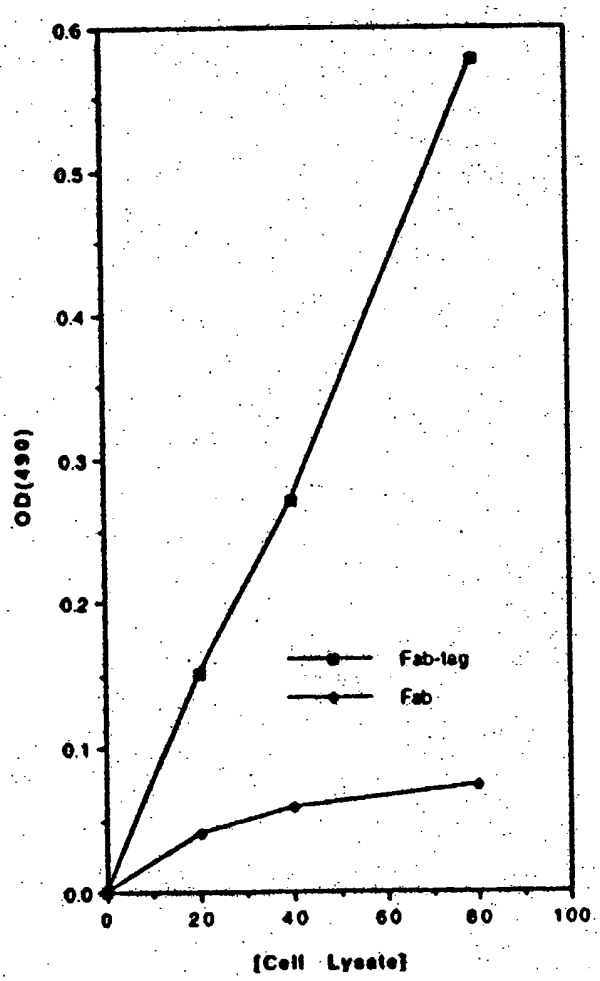

0.6
0.5
0.4
0.3
0.2
0.1
0.0
OD(490)
0
20
40
60
80
100
[Cell Lysate]
Fab-tag
Fab

Fig. 1

Fig. 2A

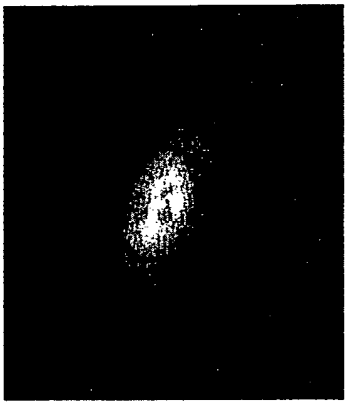


Fig. 2B

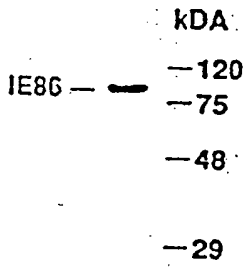


Fig. 3

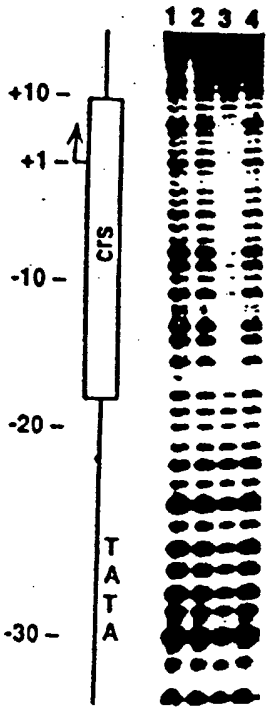


Fig. 4

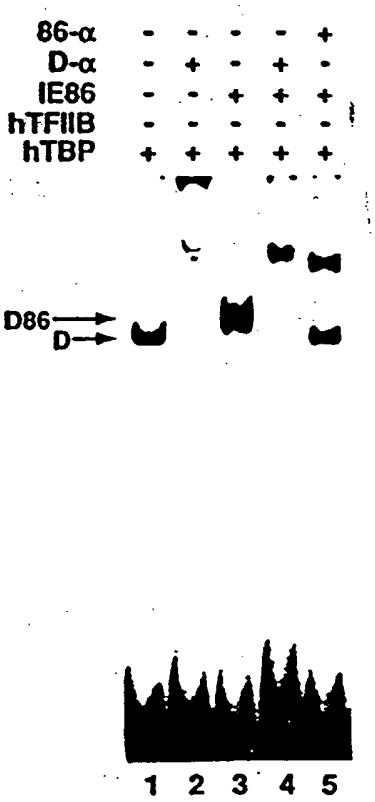


Fig. 5

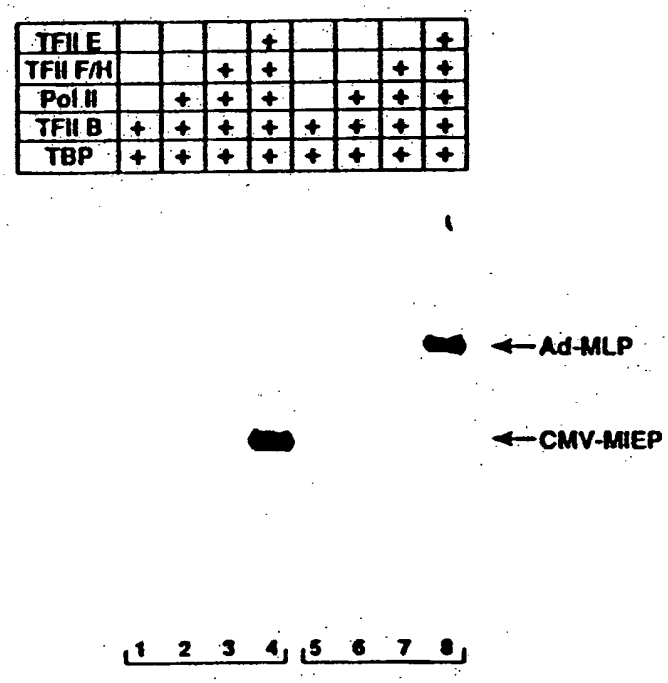


Fig. 6

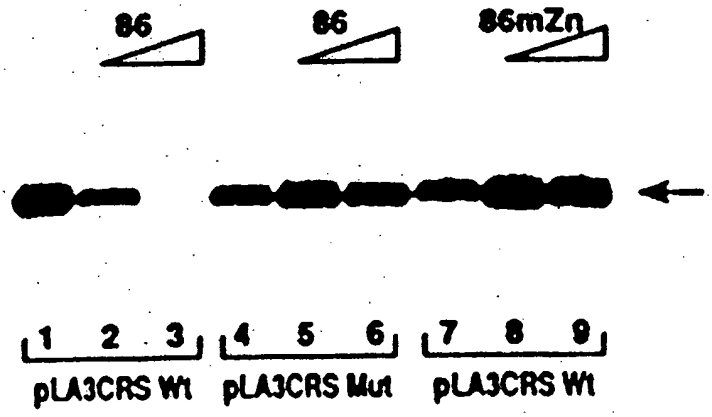


Fig. 7

Fig. 8A

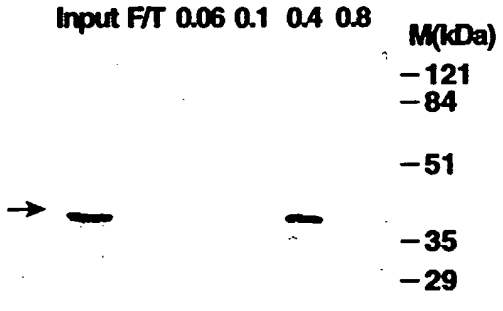


Fig. 8B

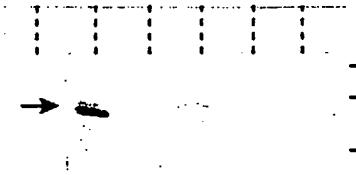
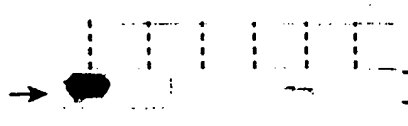


Fig. 8C



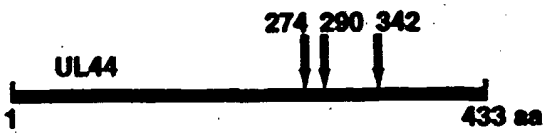


Fig. 9

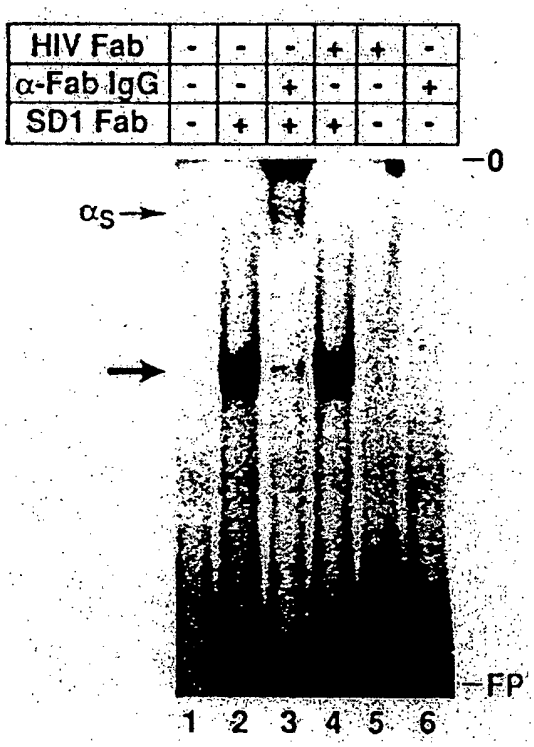


Fig. 10

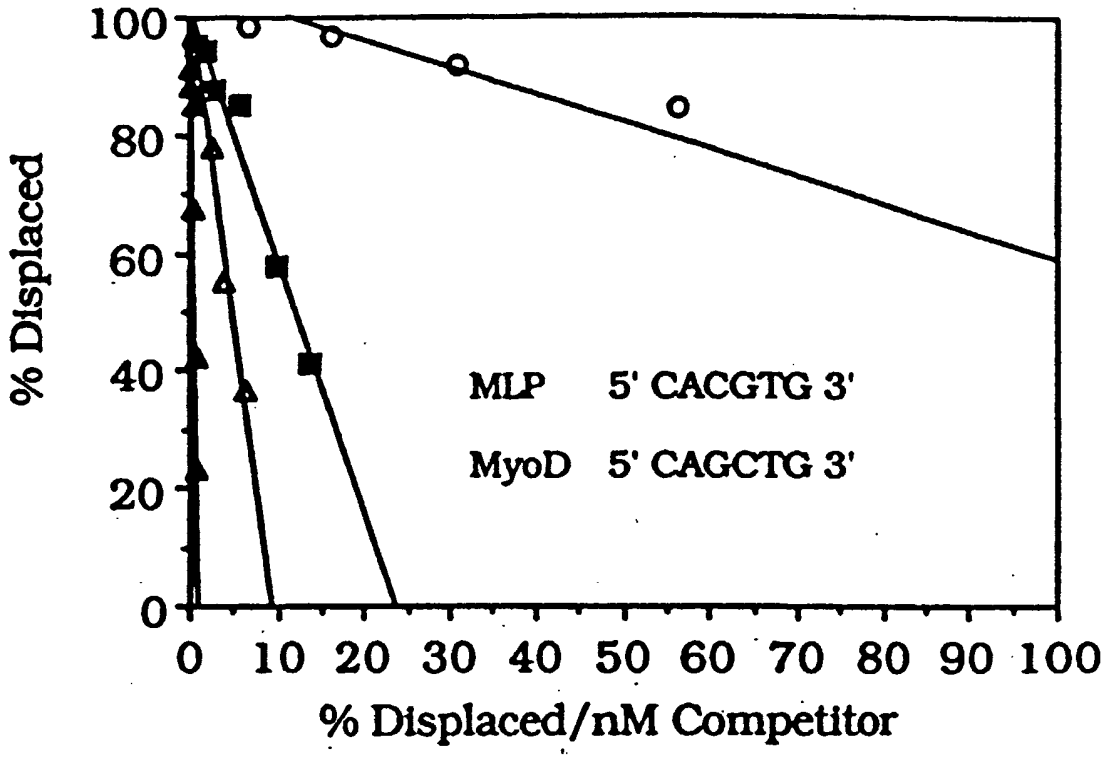

100
80
60
40
20
0
% Displaced
MLP 5' CACGTG 3'
MyoD 5' CAGCTG 3'
0 10 20 30 40 50 60 70 80 90 100
% Displaced/nM Competitor

Fig. 11

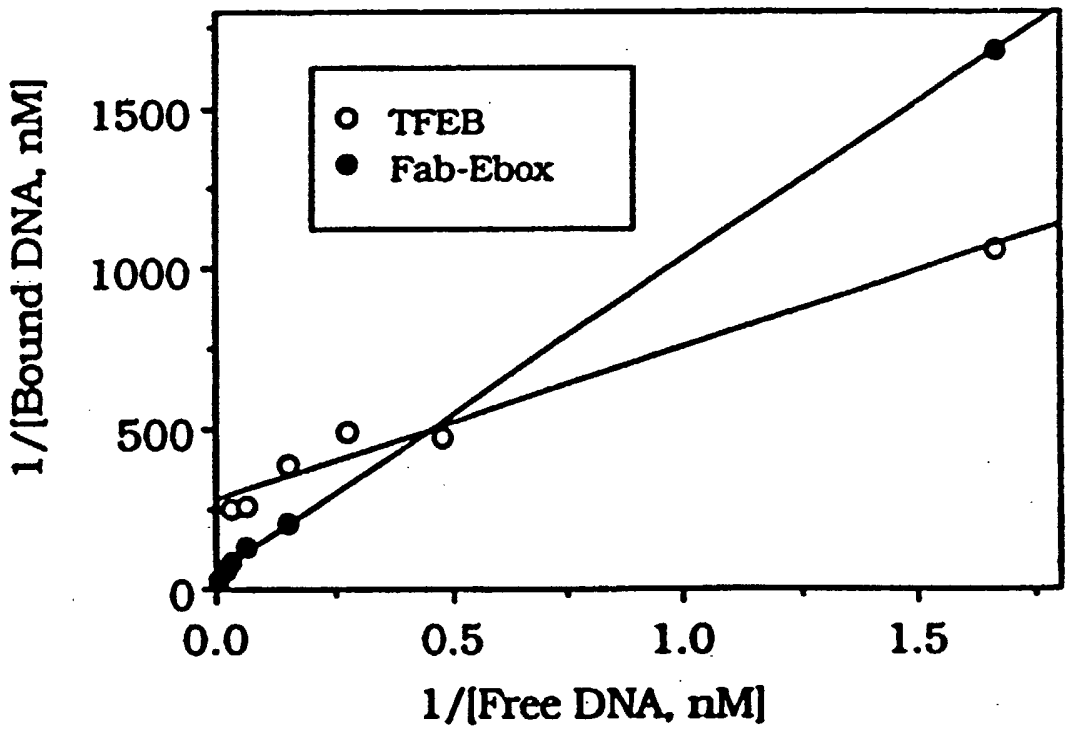

1/[Bound DNA, nM]
1500
1000
500
0
0.0
0.5
1.0
1.5
1/[Free DNA, nM]
TFEB
Fab-Ebox

Fig. 12

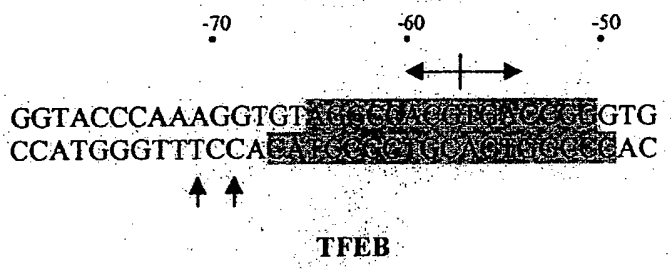


Fig. 13

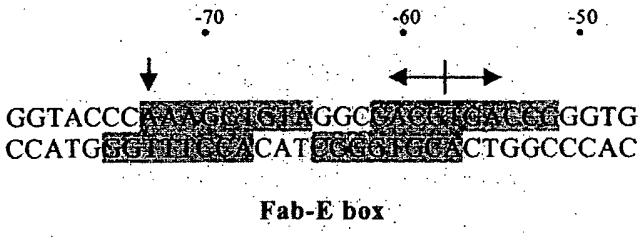


Fig. 14

E Box: CACGTG

CACGTGACCG
TAGATGGAGG
GATTTGGGAA
AGAGTGAATG
TGAATGATGA
CCTTTGGGTA
AGATTGAAAG
GGTATGGAGG
ATAATGAGGA
AATATAGAAG

Consensus: N N T/A T G Pu Pu N Pu

Fig.15

SEQUENCE LISTING

<110> Ghazal, Peter
Burton, Dennis R.
Hahn, Klaus M.

<120> NOVEL METHODS FOR INTRODUCING MOLECULES INTO CELLS AND VECTORS AND COMPOSITIONS FOR USE IN SUCH METHODS

<130> 302018.3002-101

<140>
<141> 2002-10-18

<150> US 10/039,318
<151> 2001-10-19

<160> 19

<170> PatentIn Ver. 2.0

<210> 1
<211> 13
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> nuclear localization sequence of HIV tat protein

<400> 1
Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala His Gln
1 5 10

<210> 2
<211> 5
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> partial sequence of nuclear localization peptide of HIV tat protein

<400> 2
Gly Arg Lys Lys Arg
1 5

<210> 3
<211> 4
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> partial sequence of nuclear localization of HIV tat protein

<400> 3
Gln Arg Arg Pro
1

<210> 4
<211> 36
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> residues 37-72 of HIV tat peptide

<400> 4
Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg
1 5 10 15

Arg Gln Arg Arg Arg Pro Pro Gln Phe Ser Gln Thr His Gln Val Ser
20 25 30

Leu Ser Lys Gln
35

<210> 5
<211> 35
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> residues 38-72 of HIV tat peptide

<400> 5
Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg Arg
1 5 10 15

Gln Arg Arg Arg Pro Pro Gln Phe Ser Gln Thr His Gln Val Ser Leu
20 25 30

Ser Lys Gln
35

<210> 6
<211> 13
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> residues 47-59 of HIV tat peptide

<400> 6
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro
1 5 10

<210> 7
<211> 9
<212> PRT
<213> Human immunodeficiency virus type 1

<220>
<223> residues 49-57 of HIV tat peptide

<400> 7
Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5

<210> 8
<211> 32
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL14

<400> 8
Arg Val Thr Met Thr Arg Asp Thr Ser Ile Asn Thr Val Tyr Met Glu
1 5 10 15

Leu Ser Arg Leu Arg Ser Asp Asp Thr Ala Val Tyr Trp Cys Ala Arg
20 25 30

<210> 9
<211> 17
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain seqeunce of Fab GL14

<400> 9
Glu Tyr Arg Gln Pro His Asn Ser Ala Trp Ser Asp Asp Ala Phe Asp
1 5 10 15

Ile

<210> 10
<211> 11
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL14

<400> 10
Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser
1 5 10

<210> 11
<211> 32
<212> PRT

<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL5

<400> 11
Arg Ile Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Phe Leu Gln
1 5 10 15

Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Ile Tyr Tyr Cys Ala Lys
20 25 30

<210> 12
<211> 12
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL5

<400> 12
Asp Trp Tyr Gly Asp Tyr Gly Asp Ala Phe Asp Val
1 5 10

<210> 13
<211> 32
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL34

<400> 13
Arg Val Thr Met Thr Arg Asp Thr Ser Leu Asn Thr Val Phe Met Glu
1 5 10 15

Leu Ile Ser Leu Asn Leu Gly Asp Thr Ala Val Tyr Phe Cys Ala Arg
20 25 30

<210> 14
<211> 15
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL34

<400> 14
Gly Asp Arg Gln Phe Gly Phe Arg Gly Thr Asp Trp Phe Asp Pro
1 5 10 15

<210> 15
<211> 11
<212> PRT
<213> Homo sapiens

<220>
<223> partial heavy chain sequence of Fab GL34

<400> 15
Trp Gly Gln Gly Thr Leu Val Val Val Ser Ser
1 5 10

<210> 16
<211> 21
<212> DNA
<213> Human cytomegalovirus

<220>
<223> consensus sequence for IE86 (crs) element in hCMV promoter

<400> 16
agctcgttta gtaaccgtca g 21

<210> 17
<211> 27
<212> PRT
<213> Homo sapiens

<220>
<223> basic helix-loop-helix sequence of human TFEB protein

<400> 17
Gly Gly Ala Ala Lys Lys Ala Ala His Ala Ala Ala Glu Arg Arg Arg
1 5 10 15

Arg Ala Ala Ile Asn Gly Gly Ala Ala Gly Gly
20 25

<210> 18
<211> 29
<212> DNA
<213> Homo sapiens

<220>
<223> synthetic probe for SD1 binding

<400> 18
aatgtatgcg cgcgcgcttt aggggcccc 29

<210> 19
<211> 33
<212> DNA
<213> Homo sapiens