



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12Q 1/68, 1/00, C12P 21/02 C12P 21/06, G01N 33/53 C07H 17/00, A01N 63/00 A61K 37/00	A1	(11) International Publication Number: WO 93/18185 (43) International Publication Date: 16 September 1993 (16.09.93)
(21) International Application Number: PCT/US93/01669 (22) International Filing Date: 1 March 1993 (01.03.93) (30) Priority data: 844,852 3 March 1992 (03.03.92) US (71) Applicants: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES [US/US]; 10010 North Torrey Pines Road, La Jolla, CA 92037 (US). THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10666 North Torrey Pines Road, La Jolla, CA 92037 (US). (72) Inventors: TROWBRIDGE, Ian, S. ; 6287 Lakewood, San Diego, CA 92122 (US). COLLAWN, James, F. ; 10312 Ridgewater Lane, San Diego, CA 92131 (US). TAINER, John, A. ; 4575 Niagara Avenue, San Diego, CA 92107 (US). KUHN, Leslie, A. ; 16500 South El Camino Real, Apt. F208, Encinitas, CA 92024 (US).		(74) Agents: WETHERELL, John, R., Jr. et al.; Spensley Horn Jubas & Lubitz, 1880 Century Park East, Fifth Floor, Los Angeles, CA 90067 (US). (81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: RECEPTOR INTERNALIZATION SIGNALS (57) Abstract Internalization signals and their use in modulating cell surface receptor endocytosis.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LI	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

RECEPTOR INTERNALIZATION SIGNALS

BACKGROUND OF THE INVENTION

1. Field of The Invention

5 The invention relates to internalization signals and the use of these signals to modulate the transport of ligand into a cell.

2. Related Art

Receptor-mediated endocytosis is the mechanism by which a variety of nutrients, hormones, and growth factors are specifically and efficiently transported into the cell. The process of receptor mediated endocytosis is complex and involves several distinct biochemical steps. Typically, the process proceeds by: (1) recruitment of soluble coat proteins to the cell membrane and nucleation of coated pit formation, (2) assembly of coat constituents and growth of the coated pit, (3) acquisition of specific receptors into the growing coated pit, (4) invagination of the cell membrane, (5) coat closure; and (6) membrane fusion wherein the coated pit buds in and pinches off to form a coated vesicle. The contents of the vesicles are ultimately delivered to the endosomes. Key to the entire process of receptor-mediator endocytosis is the internalization signal present in the cytoplasmic tail of the receptor. The cytoplasmic tail interacts with soluble coat proteins during formation of a coated pit.

10

15

20

-2-

Receptors such as the transferrin receptor (TR) and the low-density lipoprotein (LDL) receptor are constitutively clustered in coated pits and undergo rapid internalization in the presence and absence of ligand. Other receptors such as epidermal growth factor (EGF) receptor are only concentrated in coated pits and internalized after binding ligand. Previous studies have established that there are internalization signals in the cytoplasmic domains of constitutively recycling receptors that are believed to interact with adaptor proteins of coated pits and promote high-efficiency endocytosis.

It would be desirable to use heterologous internalization signals to regulate endocytosis. In so doing, it would be possible, for example, to stimulate the internalization of toxic substances by tumor cells or inhibit uptake of essential nutrients. However, prior to Applicants' invention, there has not been a successful transplantation of a heterologous internalization signal from one receptor to another. The present invention addresses this need and provides the means for utilizing heterologous internalization signals and predicting the sequence and structure of heretofore unknown signals.

SUMMARY OF THE INVENTION

5 In accordance with the present invention, it has been discovered that transport of ligand into a cell can be modulated by introducing a heterologous internalization signal into the cell. The signal is introduced as a nucleotide sequence or as the encoded peptide. The ability to transplant internalization signals from one receptor to another and retain activity has important implications for control of endocytosis for scientific and medical purposes, including drug delivery to cells.

10 Accordingly, the present invention provides a method of modulating receptor mediated transport of ligand into a cell, which method comprises introducing a heterologous internalization signal into the cell.

Further, the invention provides a method for identifying a sequence which modulates internalization of a cell surface receptor. This method comprises: (a) incubating cells having such receptors in the presence or absence of sequence and, optionally, in the presence of ligand for the cell surface receptor; and (b) measuring internalization of the cell receptor in the presence or absence of the sequence.

20 The present invention also provides compositions for modulating transport of ligand into a cell. Compositions embraced by the present invention comprise peptides having a tight turn conformation and a defined amino acid sequence described in greater detail below. Provided also are nucleotide sequences encoding such peptides.

-4-

In addition, the present invention provides a method of administering gene therapy to a host subject. This method can be accomplished, for example, by introducing into a host subject, cells derived from the subject which have been modified to contain heterologous internalization signal capable of modulating transport of ligand into a cell.

5

The present invention also provides a method of gene therapy comprising introducing into a host subject an expression vector comprising a nucleotide sequence encoding a heterologous internalization signal capable of modulating transport of ligand into a cell.

BRIEF DESCRIPTION OF THE DRAWINGS

The following figure legends describe the first successful transplantation of heterologous internalization signals.

FIGURE 1. Uptake of ^{59}Fe from human Tf by CEF expressing wild-type or mutant human TRs. Uptake of ^{59}Fe from human Tf by CEF expressing human TRs was determined by incubating cells with ^{59}Fe -Tf for the times indicated and then washing the cells and determining their radioactivity. The results shown are for two representative experiments, A and B, and each point represents the average values of ^{59}Fe uptake by triplicate cultures of CEF expressing wild-type or mutant TRs. Panel A shows the relative ^{59}Fe uptake for wild-type($\blacktriangle$), $^{20}\text{YTRF}^{23}\text{NPVY}$ ($\square$), $^{18}\text{LSYTRF}^{23}\text{FDNPVY}$ ($\blacksquare$), $^{20}\text{YTRF}^{23}\text{YKSV}$ ($\circ$), $^{18}\text{LSYTRF}^{23}\text{YKYSKV}$ ($\bullet$), and F23M($\triangle$) mutant TRs. Panel B shows the relative ^{59}Fe uptake for wild-type ($\blacktriangle$), $^{20}\text{YTRF}^{23}\text{NPVY}$ ($\square$), $^{18}\text{LSYTRF}^{23}\text{FDNPVY}$ ($\blacksquare$), F23W($\circ$), F23I($\bullet$), and $\Delta 3-59$ ($\triangle$) mutant human TRs.

FIGURE 2. Comparisons of steady-state distributions (■) and ^{59}Fe uptake (□) of human TR mutants expressed in CEFs (mutations at the carboxy-terminal aromatic residue of the TR internalization motif, YTRF). Phenylalanine 23 was changed to either methionine (F23M), isoleucine (F23I), tryptophan (F23W), alanine (F23A), or glycine (F23G). The data for the F23A and F23G mutants is from Collawn, et al., *Cell*, 63:1061-1072, 1990. For the steady-state analysis, CEFs expressing human TRs were incubated with ^{125}I -labeled Tf for 60 min at 37°C then washed with buffer. The acid wash technique described in Materials and Methods was used to distinguish surface-bound and internalized Tf. The internalization rates represent the average of three experiments and are given as percentages $\pm$ standard errors relative to the wild-type (Wt) TR.

FIGURE 3. Superimposed crystallographic turn structures for internalization motif analogs with sequences matching the six-residue mannose-6-phosphate receptor (Man-6-PR) and the low-density lipoprotein receptor (LDLR) patterns. (A) Man-6-PR internalization motif analogs. (B) LDLR internalization motif analogs. (C) Superimposed Man-6-PR and LDLR analogs.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates to a method of modulating receptor mediated transport of ligand into a cell wherein the method comprises introducing a heterologous internalization signal into the cell.

5 The term "heterologous" when used to describe the internalization signal of the invention refers to any internalization signal that is introduced into the cell. The term "ligand" refers to any substance capable of binding to or with a cell surface receptor. The term "tight turn" refers to the reverse or helical
10 turn conformation of amino acid residues within the heterologous internalization signal (Collawn, *et al.*, *Cell*, 63:1061, 1990; Collawn, *et al.*, *EMBO J.*, 10:3247, 1991). The term "cell-surface receptor" refers to any cell surface molecule whether or not it has a naturally occurring ligand. The term "sequence" refers to amino acid sequences as well as nucleic acid sequences. The term "internalization motif" refers to an amino acid
15 internalization signal having a tight turn structure. The terms internalization motif and internalization signal are used interchangeably. The term "core" refers to the smallest sequence of amino acid residues involved in cell surface internalization.

-8-

Amino acids referred to herein may be identified according to the following three-letter or one-letter abbreviations:

	<u>Amino Acid</u>	<u>Three-Letter Abbreviation</u>	<u>One-Letter Abbreviation</u>
5	L-Alanine	Ala.	A
	L-Arginine	Arg	R
	L-Asparagine	Asn	N
	L-Aspartic Acid	Asp	D
	L-Cysteine	Cys	C
10	L-Glutamine	Gln	Q
	L-Glutamic Acid	Glu	E
	L-Glycine	Gly	G
	L-Histidine	His	H
	L-Isoleucine	Ile	I
15	L-Leucine	Leu	L
	L-Lysine	Lys	K
	L-Methionine	Met	M
	L-Phenylalanine	Phe	F
	L-Proline	Pro	P
20	L-Serine	Ser	S
	L-Threonine	Thr	T
	L-Tryptophan	Trp	W
	L-Tyrosine	Tyr	Y
	L-Valine	Val	V

25 The heterologous internalization signals utilized herein may be the same as or different from internalization signal already present in the cell. When DNA encoding heterologous signal is introduced into a target cell, the introduced nucleotide sequence may further comprise additional nucleotide sequence encoding a cell surface receptor. This surface receptor may be the same

30 as or different from receptor already present in the target cell.

Thus, the present invention envisions embodiments wherein heterologous internalization signal is introduced into cells having receptors containing internalization signals as well as those having receptors without internalization signals. Further, when the internalization sequence is introduced as part of a cell surface receptor, the receptor may be the same as or different from a receptor already present in the target cell.

Introduction of heterologous internalization signal into target cells according to the invention can be accomplished by several means. For example, a signal can be introduced as a peptide or as a nucleotide sequence encoding such a peptide. Moreover, the introduced sequence can further comprise additional flanking sequence. Such additional flanking sequence can encode or can be a cell surface receptor.

Internalization signal peptide or nucleic acid sequence encoding the peptide may be introduced into a cell surface receptor by several methods including:

- 1) substituting residues in the receptor by residues in the introduced internalization sequence (one for one replacement of residues/bases);
- 2) inserting an internalization signal between two residues/bases in the resident receptor so that the receptor sequence becomes longer;
- 3) replacing some residues in the resident receptor with a greater number of internalization signal residues such that the receptor becomes longer.

The method selected for introduction of internalization signal peptide will be dictated by structural and experimental considerations on a case by case basis.

Preferred locations for introduction of internalization signal into a receptor sequence can be ascertained by locating endogenous internalization signal and by locating positions in the receptor sequence having tight turn

conformation and packing interactions that can accommodate introduced signal. Structure and packing information can be obtained from crystallography, NMR, or high electron resolution microscopy. In situations where such information is not available, commonly used and available secondary structure prediction algorithms may be used to locate sequence regions that likely fold as tight turns. In cases involving substitution of internalization sequence for receptor sequence, tight turn regions of the receptor which are most sequence similar to the internalization sequence are preferred sites for substitution since substitutions in such sites minimizes structural destabilization. Sequence similarity can be determined by mutation data matrices and other measures of sequence similarity, such as residue size and polarity.

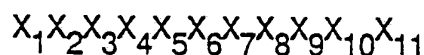
The introduction of heterologous internalization signal serves to modulate the transport of ligand into a cell having a surface receptor reactive with that ligand. This modulation can induce either an increase or a decrease in endocytosis, depending upon the choice of heterologous internalization signal. As discussed earlier, the key to receptor-mediated endocytosis is the internalization signal present in the cytoplasmic tail of the surface receptor. It is the internalization signal that regulates the uptake of cell surface receptor.

Identification of a heterologous internalization signal which can modulate internalization of cell surface receptor is accomplished by incubating cells having such receptors in the presence or absence of the sequence suspected of being an internalization signal. Cells are incubated according to the method of Jing, *et al.*, *J. Cell Biol.*, 110:283, 1990. Optionally, incubation is carried out in the presence of ligand and internalization of cell surface receptor, in the presence or absence of sequence, is measured.

-11-

The invention also provides a method for inhibiting internalization of cell surface receptor. This is accomplished by introducing into a cell having a receptor that clusters in coated pits, an effective amount of heterologous internalization signal peptide, wherein the heterologous peptide competes with internalization signal peptide of the receptor for binding with adaptor protein. The term "effective amount" refers to the amount of peptide which results in inhibition of endocytotic vesicle formation. Delivery of an effective amount of the internalization signal peptide can be accomplished by one of the mechanisms described herein, such as by encapsulation in liposomes, or other methods well known in the art.

Internalization signal peptides that can be utilized in the present invention are characterized by having a tight turn and the amino acid sequence:



where $X_5 - X_8$ constitutes the core sequence,

$X_1 - X_4$ residues and $X_9 - X_{11}$ residues are optional and wherein:

X_1 , when present, is leucine or glutamic acid;

X_2 , when present, is isoleucine, methionine or proline;

X_3 , when present, is any amino acid residue;

-12-

X_4 , when present, is selected from alanine, polar amino acids, or aromatic amino acids, and when X_3 is also present, at least one of the X_3 or X_4 residues is polar;

5 X_5 is an aromatic amino acid when residues $X_1 - X_4$ are not present, or is selected from aromatic amino acids or polar amino acids when at least residue X_4 or additional upstream residue(s) is present;

X_6 is a polar amino acid or alanine;

10 X_7 is selected from polar amino acids or alanine when residues $X_1 - X_4$ are not present, or is any amino acid residue when at least X_4 or additional upstream residue(s) is present;

X_8 is selected from aromatic amino acids or hydrophobic amino acids;

X_9 , when present, is serine or alanine;

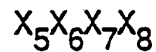
X_{10} , when present, is alanine or leucine; and

15 X_{11} , when present, is alanine or phenylalanine;

wherein at least one of residues X_3 , X_5 , and X_8 is an aromatic amino acid and further wherein residues X_1 , X_2 , X_3 , X_{10} , and X_{11} can only be present when the next adjacent residue(s) relative to the core is present.

-13-

A first preferred group of internalization signal peptides is characterized as having the amino acid sequence:



5 wherein $X_5 - X_8$ are as previously defined with the proviso that X_6 and X_7 cannot both be alanine.

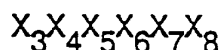
A more preferred group of internalization signal peptides having the sequence $X_5X_6X_7X_8$ are those wherein:

10 X_5 is phenylalanine or tyrosine;
 X_6 is alanine, arginine, glutamine, serine, or threonine;
 X_7 is alanine, arginine, aspartic acid, glycine, glutamic acid, histidine, lysine, or threonine; and
 X_8 is isoleucine, leucine, methionine, phenylalanine, valine, or tryptophan.

15 Most preferred internalization signal peptides having the $X_5X_6X_7X_8$ sequence are selected from the group consisting of YTRM, YARF, YTRI, YQDL, YTKF, YSKV, YTRW, YRHV, YSAF, YQTI, YTAF, YTGF, YTEF, FTRF, and YTRF.

-14-

A second preferred group of internalization signal peptides is characterized as having the amino acid sequence:



wherein $X_3 - X_8$ are as defined previously.

- 5 A more preferred group of internalization signal peptides having the sequence $X_3X_4X_5X_6X_7X_8$ are those wherein:

X_3 is asparagine, leucine, methionine, phenylalanine, proline, or tyrosine;

X_4 is alanine, aspartic acid, glutamine, lysine, phenylalanine, or serine;

- 10 X_5 is asparagine, glutamine, or tyrosine;

X_6 is arginine, glycine, proline, serine, or threonine;

X_7 is alanine, arginine, lysine, phenylalanine, or valine; and

X_8 is isoleucine, leucine, methionine, phenylalanine, tryptophan, tyrosine, or valine.

- 15 Most preferred internalization signal peptides having the sequence $X_3X_4X_5X_6X_7X_8$ are selected from the group consisting of YKYSKV, NFYRAL, LAYTRF, PQQGFF, FDNPVY, MSYTRF, or LSYTRF.

Novel compositions of the present invention comprise peptides having a tight turn and the amino acid sequence:

- 20 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}X_{11}$

-15-

where $X_5 - X_8$ constitutes the core sequence,

$X_1 - X_4$ residues and $X_9 - X_{11}$ residues are optional and wherein:

X_1 , when present, is leucine or glutamic acid;

X_2 , when present, is isoleucine, methionine or proline;

5 X_3 , when present, is any amino acid residue;

X_4 , when present, is selected from alanine, polar amino acids, or aromatic amino acids, and when X_3 is also present, at least one of the X_3 or X_4 residues is polar;

10 X_5 is an aromatic amino acid when residues $X_1 - X_4$ are not present, or is selected from aromatic amino acids or polar amino acids when at least residue X_4 or additional upstream residue(s) is present;

X_6 is a polar amino acid or alanine;

15 X_7 is selected from polar amino acids or alanine when residues $X_1 - X_4$ are not present, or is any amino acid residue when at least X_4 or additional upstream residue(s) is present;

X_8 is selected from aromatic amino acids or hydrophobic amino acids;

X_9 , when present, is serine or alanine;

-16-

X_{10} , when present, is alanine or leucine; and

X_{11} , when present, is alanine or phenylalanine;

5 wherein at least one of residues X_3 , X_5 , and X_8 is an aromatic amino acid and further wherein residues X_1 , X_2 , X_3 , X_{10} , and X_{11} can only be present when the next adjacent residue(s) relative to the core is present, provided that sequences selected from the group consisting of FXNPXY, GPLY, PPGY, and YXYXKV, where X stands for any amino acid, are excluded.

10 As employed herein, the term "polar amino acid" includes glycine, serine, threonine, histidine, tyrosine, proline, aspartic acid, asparagine, glutamic acid, glutamine, arginine, and lysine. (Rose, *et al.*, *Science*, 229:834, 1985).

The term "aromatic amino acid" refers to any amino acid containing an unsaturated carbocyclic ring including phenylalanine, tyrosine, and tryptophan.

15 The term "hydrophobic amino acid" includes cysteine, valine, isoleucine, leucine, methionine, phenylalanine, and tryptophan. (Rose, *et al.*, *Science*, 229:834, 1985). The term "large hydrophobic amino acid" refers to amino acids selected from the group consisting of valine, isoleucine, leucine, methionine, phenylalanine, and tryptophan.

20 In addition to increasing endocytosis by adding active internalization signal (where "active" is defined as having at least twice the internalization efficiency of a receptor lacking a signal), the present invention also embraces inhibition of endocytosis. Inhibition is accomplished by 1) introducing signal shown

-17-

to be inactive when substituted to an active resident signal or 2) using active internalization peptides as competitive inhibitors of receptor internalization (i.e., the peptide competes with intact receptor for binding to adaptor complex). Inactive signals which can be used according to the first method
5 include sequences corresponding to the core $X_5X_6X_7X_8$ and selected from the group consisting of ATRF, GTRF, ATRA, YTRG, FQDI, IGSY, NTLY, HLAF, PPGY and NPVY.

Generation of inactive internalization signals can be accomplished by mutating aromatic residues occurring in positions X_3 , X_4 , X_5 and X_8 to non-
10 aromatic residues.

The present invention can utilize any nucleotide sequence encoding peptides described above. These peptide encoding nucleotide sequences can further comprise flanking nucleotide sequence encoding a cell surface receptor.

In the present invention, nucleotide sequences encoding internalization
15 signal may be introduced into a host cell by means of a recombinant expression vector. The term "recombinant expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of internalization signal sequences. Such expression vectors typically contain a promotor sequence which facilitates
20 efficient transcription of the inserted sequence in the host. The expression vector also typically contains specific genes which allow phenotypic selection of the transformed cells.

Alternatively, nucleotide sequences encoding internalization signal can be introduced directly as a plasmid, for example, by microinjection, or

transfection.

The present invention also provides methods for the treatment of disease employing gene therapy that modulates receptor mediated transport of ligand into a cell or internalization of cell surface receptor. Such therapy can be effected by introduction of internalization signal into cells of a subject having the disease. Delivery of internalization signal can be achieved using techniques well known in the art. For example, when an internalization signal is introduced by means of a nucleotide sequence, a recombinant expression vector, such as a chimeric virus, or a colloidal dispersion system can be employed.

Various viral vectors which can be utilized for introduction of internalization signal according to the present invention, include adenovirus, herpes virus, vaccinia, or, preferably, an RNA virus such as a retrovirus. Preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of retroviral vectors in which a single foreign gene can be inserted include, but are not limited to: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), and Rous Sarcoma Virus (RSV). A number of additional retroviral vectors can incorporate multiple genes. All of these vectors can incorporate a gene for a selectable marker so that transduced cells can be identified and cultured.

By inserting a polynucleotide encoding the internalization signal of interest into a target specific viral vector, the internalization signal becomes targeted to specific cells. Retroviral vectors can be made target specific by including in the retroviral vector a polynucleotide encoding a substance which binds to cell surface target. Preferred targeting is accomplished by using an

antibody to target the retroviral vector. Those of skill in the art will know of, or can readily ascertain without undue experimentation, specific polynucleotide sequences which can be inserted into the retroviral genome to allow target specific delivery of the retroviral vector containing the internalization signal polynucleotide.

5

Since recombinant retroviruses are defective, they require assistance in order to produce infectious viral particles. This assistance can be provided, for example, by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences within the LTR (long terminal repeat). These plasmids are missing a nucleotide sequence which enables the packaging mechanism to recognize an RNA transcript for encapsidation. Helper cell lines which have deletions of the packaging signal include, but are not limited to, Ψ 2, PA317 and PA12, for example. These cell lines produce empty virions, since no genome is packaged. If a retroviral vector is introduced into such cells in which the packaging signal is intact, but the structural genes are replaced by other genes of interest, the vector can be packaged and vector virion produced. The vector virions produced by this method can then be used to infect a tissue cell line, such as NIH 3T3 cells, to produce large quantities of chimeric retroviral virions.

10

15

20

Alternatively, NIH 3T3 or other tissue culture cells can be directly transfected with plasmids encoding the retroviral structural genes gag, pol and env, by conventional calcium phosphate transfection. These cells are then transfected with the vector plasmid containing the genes of interest. The resulting cells release the retroviral vector into the culture medium.

25

-20-

An alternative use for recombinant retroviral vectors comprises the introduction of polynucleotide sequences into the host by means of skin transplants of cells containing the virus. Long term expression of foreign genes in implants, using cells of fibroblast origin, may be achieved if a strong housekeeping gene promoter is used to drive transcription. For example, the dihydrofolate reductase (DHFR) gene promoter may be used. Cells such as fibroblasts, can be infected with virions containing a retroviral construct containing the internalization signal gene of interest together with a gene which allows for specific targeting, such as tumor-associated antigen (TAA), and a strong promoter. The infected cells can be embedded in a collagen matrix which can be grafted into the connective tissue of the dermis in the recipient subject. As the retrovirus proliferates and escapes the matrix it will specifically infect the target cell population. In this way the transplantation results in increased amounts of internalization signal peptide being produced in cells manifesting the disease.

Another targeted delivery system for introduction of internalization signal peptides, or nucleotides encoding same, is a colloidal dispersion system. Colloidal dispersion systems include macromolecular complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. The preferred colloidal system of this invention is a liposome.

Since internalization signal peptide may be indiscriminant in its action with respect to cell type, a targeted delivery system offers a significant improvement in therapeutic applications over randomly injected non-specific liposomes. A number of procedures can be used to covalently attach either polyclonal or monoclonal antibodies to a liposome containing internalization

-21-

signals to render the liposome target specific. Antibody-targeted liposomes can include monoclonal or polyclonal antibodies or fragments thereof such as Fab, or F(ab')₂, as long as they bind efficiently to an epitope on the target cells. Liposomes may also be targeted to cells expressing receptors for hormones or other serum factors.

Liposomes are artificial membrane vesicles which are useful as in vitro and in vivo delivery vehicles. It has been shown that large unilamellar vesicles (LUV), which range in size from 0.2-4.0 μ m can encapsulate a substantial percentage of an aqueous buffer containing large macromolecules. RNA, DNA, intact virions and peptides can be encapsulated within the aqueous interior and be delivered to cells in a biologically active form (Fraley, *et al.*, *Trends Biochem. Sci.*, 6:77, 1981). In addition to mammalian cells, liposomes have been used for delivery of polynucleotides into plant, yeast and bacterial cells. In order for a liposome to be an efficient transfer vehicle, it should be capable of: (1) encapsulation of peptides or nucleotides of interest at high efficiency without compromising biological activity; (2) preferential and substantial binding to target cells relative to non-target cells; (3) delivery of aqueous contents of vesicle to the target cell cytoplasm at high efficiency; and (4) accurate and effective expression of genetic information (Mannino, *et al.*, *Biotechniques*, 6:682, 1988).

The composition of the liposome is usually a mixture of phospholipids, particularly high-phase-transition-temperature phospholipids, usually in combination with steroids, especially cholesterol. Other phospholipids or other lipids may also be used. The physical characteristics of liposomes depend on pH, ionic strength, and the presence of divalent cations.

-22-

Examples of lipids useful in liposome production include phosphatidyl compounds, such as phosphatidylglycerol, phosphatidylcholine, phosphatidylserine, phosphatidylethanolamine, sphingolipids, cerebrosides, and gangliosides. Particularly useful are diacylphosphatidylglycerols, where the lipid moiety contains from 14-18 carbon atoms, particularly from 16-18 carbon atoms, and is saturated. Illustrative phospholipids include egg phosphatidylcholine, dipalmitoylphosphatidylcholine and distearoylphosphatidylcholine.

The surface of the targeted delivery system may be modified in a variety of ways. In the case of a liposomal targeted delivery system, specific lipid groups can be incorporated into the liposome in order to maintain the target directed binding substance in stable association with the liposome. Various linking groups can be used for joining the lipid chains to the target directed binding substance.

In general, the targeted delivery system will be directed to cell surface receptors thereby allowing the delivery system to find and "home in" on the desired cells. Alternatively, the delivery system can be directed to any cell surface molecule preferentially found in the cell population for which treatment is desired provided that the cell surface molecule is capable of association with the delivery system. Antibodies can be used to target liposomes to specific cell-surface molecules. For example, certain antigens expressed specifically on tumor cells, referred to as tumor-associated antigens (TAAs), may be exploited for the purpose of targeting antibody-internalization signal-containing liposomes directly to a malignant tumor.

Because the present invention identifies sequences involved in the internalization of cell surface receptors, it is possible to design therapeutic or diagnostic protocols utilizing these sequences. Thus, where maintenance of a disease state is related to internalization of a cell surface receptor, the native internalization sequence can be utilized to design sequences which interfere with the function of the native internalization signal. This approach utilizes, for example, antisense nucleic acid and ribozymes to block translation of specific receptor mRNA, either by masking the mRNA with antisense nucleic acid or by cleaving it with ribozyme.

Antisense nucleic acids are DNA or RNA molecules that are complementary to at least a portion of a specific mRNA molecule (Weintraub, *Scientific American*, 262:40, 1990). In the cell, the antisense nucleic acids hybridize to the corresponding mRNA, forming a double-stranded molecule. The antisense nucleic acids interfere with the translation of the mRNA since the cell will not translate a mRNA that is double-stranded. Antisense oligomers of about 15 nucleotides are preferred, since they are easily synthesized and are less likely to cause problems than larger molecules when introduced into the target receptor-producing cell. The use of antisense methods to inhibit the in vitro translation of genes is well known in the art (Marcus-Sakura, *Anal.Biochem.*, 172:289, 1988).

Ribozymes are RNA molecules possessing the ability to specifically cleave other single-stranded RNA in a manner analogous to DNA restriction endonucleases. Through the modification of nucleotide sequences which encode these RNAs, it is possible to engineer molecules that recognize specific nucleotide sequences in an RNA molecule and cleave it (Cech, *J.Amer.Med.Assn.*, 260:3030, 1988). A major advantage of this approach is

that, because ribozymes are sequence-specific, only mRNAs with particular sequences are inactivated.

There are two basic types of ribozymes namely, tetrahymena-type (Hasselhoff, *Nature*, 334:585, 1988) and "hammerhead"-type. Tetrahymena-type ribozymes recognize sequences which are four bases in length, while
5 "hammerhead"-type ribozymes recognize base sequences 11-18 bases in length. The longer the recognition sequence, the greater the likelihood that that sequence will occur exclusively in the target mRNA species. Consequently, hammerhead-type ribozymes are preferable to tetrahymena-type
10 ribozymes for inactivating a specific mRNA species and 18-based recognition sequences are preferable to shorter recognition sequences.

Antisense sequences can be therapeutically administered by techniques as described above for the administration of heterologous internalization signal polynucleotides. Targeted liposomes are especially preferred for therapeutic
15 delivery of antisense sequences.

MATERIALS AND METHODS

The below described materials and methods were utilized in transplanting internalization signals and analyzing their structure.

A. Oligonucleotide Site-Directed Mutagenesis

5 A Cla I fragment containing the entire coding region of the human TR was cloned into the phagemid pBluescript SK (Stratagene, La Jolla, CA) (Jing, *et al.*, *J. Cell Biol.*, 110:283-294, 1990). Oligonucleotide site-directed mutagenesis was performed with pBluescript SK phagemid templates of the wild-type TR by the method of Kunkel, (*Proc. Natl. Acad. Sci. USA*, 82:488-
10 492, 1985) using the Muta-gene mutagenesis kit (Bio-Rad, Richmond, CA) as described previously (Collawn, *et al.*, *Cell*, 63:1061-1072, 1990). Mutants were selected by restriction mapping and Cla I fragments encoding the mutant receptors were then excised and cloned into the retroviral expression vector RCAS (Hughes, *et al.*, *J. Virology*, 61:3004-3012, 1987) or BH-RCAS.
15 The mutations were verified by dideoxynucleotide sequencing (Sanger, *et al.*, *Proc. Natl. Acad. Sci. USA*, 74:5463-5467, 1977; Tabor and Richardson, *Proc. Natl. Acad. Sci. USA*, 84:4767-4771, 1987) of the retroviral constructs using the Sequenase kit (USB Corp., Cleveland, Ohio) according to the manufacturer's directions.

B. Cell Culture and Expression of Human TRs in CEF

Primary chicken embryo fibroblasts (CEF) were obtained from nine day embryos (SPAFAS, Norwich, Conn.) and grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 1% (v/v) chicken serum, 1% (v/v) defined calf bovine serum (Hyclone, Logan, Utah), 2% (v/v) tryptose phosphate broth (Difco, Detroit, MI). CEF were transfected with retroviral construct prepared in (A) using 30 micrograms DNA per 10cm tissue culture plate of ~40% confluent cells using the polybrene-dimethyl sulfoxide method (Kawai and Nishizawa, *Mol. Cell. Biol.*, 4:1172-1174, 1984).

Stable expression of wild-type and mutant human TRs was achieved using a helper-independent retroviral vector, BH-RCAS, derived from Rous sarcoma virus (Hughes, *et al.*, *J. Virology*, 61:3004-3012, 1987). 1-2 wk after transfection, the CEF stably expressed wild-type and mutant receptors on their cell surface as a result of infection by recombinant virus. Cell surface expression was confirmed by ¹²⁵I-transferrin binding at 37° C under steady-state conditions.

To select for high expression of mutant human TRs, cells were grown in selection media (DMEM supplemented with 3% (v/v) horse serum, 2% tryptose phosphate broth and 50 micrograms/ml diferric human transferrin (Tf) (Miles Scientific, Naperville, IL)) starting 10-14 days after transfection (Jing, *et al.*, *J. Cell Biol.*, 110:283-294, 1990).

C. ^{125}I -Tf Binding

Diferric human transferrin (Tf) was labeled with ^{125}I to a specific activity of 2-4 $\mu\text{Ci}/\text{mg}$ using Enzymobeads (Bio-Rad, Richmond, CA) according to the manufacturer's directions. Cells were plated at a density of 7.5×10^4 cells/ cm^2 in 24-well Costar tissue culture plates 24 hr prior to the binding assay. Cells were incubated in serum-free DMEM for 1 hr at 37°C and then washed once with ice-cold 0.15M NaCl-0.01M Na phosphate buffer (pH 7.4) containing 0.1% bovine serum albumin (BSA-PBS). ^{125}I -Tf (4 micrograms/ml) in 0.15 ml BSA-PBS was added to triplicate wells and incubated at 4°C for 60 min. Cells were then washed three times with ice-cold 0.5 ml BSA-PBS, removed from the wells with 0.15 ml 1M NaOH, and the radioactivity counted in a gamma counter.

D. Analysis of Transferrin Internalization at Steady-State

For studies of the steady-state distribution of mutant and wild-type human TRs, CEFs expressing TR mutants were plated in triplicate wells as described for the binding studies. The cells were first incubated for 1 hr at 37°C in serum-free DMEM, and then incubated with 4 micrograms/ml ^{125}I -Tf in 0.1% BSA in DMEM for 1 hr at 37°C . The media was removed and the cells were washed three times with 1 ml of ice-cold BSA-PBS. The cells were then incubated twice for 3 min with 0.5 ml of 0.2 M acetic acid-0.5M NaCl (pH 2.4) to remove surface-bound ^{125}I -Tf (Hopkins and Trowbridge, *J. Cell. Biol.*, 97:508-521, 1983). Cells were then removed from the wells with 1M NaOH and radioactivity in the acid wash and the cell lysate was

-28-

determined. More prolonged incubation with the acid wash did not affect the radioactivity released (Jing, *et al.*, *J. Cell Biol.*, 110:283-294, 1990).

At steady-state $k_{\text{int}} [\text{TR}]_{\text{sur}} = k_{\text{ext}} [\text{TR}]_{\text{int}}$, the rate of internalization, k_{int} , of cell surface Tf-TR complexes, $[\text{TR}]_{\text{sur}}$, equals the rate of externalization of the internal pool of apoTf-TR complexes, $[\text{TR}]_{\text{int}}$, assuming an insignificant rate of intracellular degradation of receptors (Collawn, *et al.*, *Cell*, 63:1061-1072, 1990). The term "apoTf" refers to transferrin having no bound iron.

E. ⁵⁹Fe Uptake Measurements

Human apoTf was labelled with ⁵⁹Fe (FeCl₃; Amersham Corp., Arlington Heights, IL) to a specific activity of 5-10 $\mu\text{Ci}/\text{mg}$ using nitrilotriacetate (Bates and Schlabach, *J. Biol. Chem.*, 248:3228-3232, 1973). Cells were plated at a density of 7.5×10^4 cells/cm² in 24-well Costar tissue culture plates 24 hr before the assay. The following day, cells were washed twice in prewarmed (37°C) serum-free DMEM and then were incubated in DMEM containing 0.1% BSA and 4 $\mu\text{g}/\text{ml}$ ⁵⁹Fe-Tf at 37°C for 0, 1, 2, 3, and 4 hrs. After the indicated times, the media was removed, and the cells were washed three times with ice-cold 0.1% BSA in PBS. Cells from triplicate wells for each time point were removed in 0.5 ml 1M NaOH, and the radioactivity was counted in a gamma counter. The relative levels of human TRs expressed on the various CEF populations were determined in each experiment. After pre-incubation for 1 hr at 37°C in serum-free DMEM, triplicate wells of cells were incubated with 4 $\mu\text{g}/\text{ml}$ ¹²⁵I-labelled Tf on ice for 1 hr and then washed three times with 1 ml of ice-cold BSA-PBS, and the radioactivity bound to the cells was determined.

F. Comparison of Protein Structure Analogs of Six-Residue Internalization Motifs

5 Analogues of the Man-6-PR and LDLR six-residue signals were identified by searching the ~550 protein structures in the October 1990 Protein Data Bank (PDB) (Bernstein, et al., J. Mol. Biol., 193:775-791, 1977) for sequences matching the internalization patterns and having the final four residues in a surface-exposed tight turn (a helical turn or a reverse turn). The computer program Sequery (developed by Leslie Kuhn and Michael Pique as the successor to Searchwild, described in Collawn, et al., *Cell*, 63:1061-1072, 10 1990) searched PDB sequences for the six-residue Man-6-PR pattern (Y,F)(X)(Y,F)(X) (polar:K,R,H,N,Q,D,E) (large hydrophobic:V,I,L,M,F,Y,W) and the six-residue LDLR pattern (F,Y)(X)(N)(X)(X)(Y,F,W) with X being other than Gly for both patterns. Homologous matches and matches lacking side chain 3-dimensional coordinates were deleted, and those structures containing 15 surface-exposed tight turns in the final four positions were examined using molecular graphics to ascertain preferred positions for the side chains. The backbone atoms of the turn positions in each analog were superimposed on an ideal type I turn using the molecular graphics program Insight II (Biosym Technologies, San Diego, CA) on a Silicon Graphics Iris 4D310VGX 20 workstation.

The following examples are intended to illustrate but not limit the invention. While they are typical of those that might be used, other procedures known to those skilled in the art may alternatively be used.

EXAMPLE 1**TRANSPLANTATION OF HETEROLOGOUS INTERNALIZATION SIGNALS**

Studies were performed to determine whether internalization signals of cell surface receptors are interchangeable. These studies involved the
5 transplantation of the internalization signals from low density lipoprotein receptor (LDLR) and mannose-6-phosphate receptor (Man-6-PR) into TR. Oligonucleotide-directed mutagenesis was used to replace the four-residue YTRF internalization signal from the cytoplasmic tail of the wild-type TR with the four-residue sequences from LDLR and Man-6-PR signals and the six-
10 residue TR sequence LSYTRF by the six-residue LDLR and Man-6-PR signals (TABLE I).

TABLE I

INTERNALIZATION SEQUENCES
IN TRANSFERRIN RECEPTOR CONSTRUCTS

<u>Source</u>	<u>Length of Sequence</u>	<u>Sequence^a</u>					
		-2	-1	1	2	3	4
WtTR	4			²⁰ <u>Y</u>	T	<u>R</u>	<u>F</u> ²³
Man-6-PR	6	¹⁸ <u>Y</u>	K	<u>Y</u>	S	<u>K</u>	<u>V</u> ²³
Man-6-PR	4			²⁰ <u>Y</u>	S	<u>K</u>	<u>V</u> ²³
LDLR	6	¹⁸ <u>F</u>	D	<u>N</u>	<u>P</u>	V	<u>Y</u> ²³
LDLR	4			²⁰ <u>N</u>	<u>P</u>	V	<u>Y</u> ²³

- ^a For the four-residue substitutions, the TR signal YTRF, was changed to YSKV for the bovine Man-6-PR signal and NPVY for the human LDLR signal. For the six-residue substitutions, the TR sequence, LSYTRF, was changed to YKYSKV or to FDNPVY. Numbers above the sequences (-2 through 4) refer to positions relative to the turn region implicated in TR and LDLR (Collawn, *et al.*, *Cell*, **63**:1061-1072, 1990) with position 1 starting the turn. TR residue numbers appear as superscripts. Critical residues of internalization sequences identified by mutation to alanine are underlined (TR, Collawn, *et al.*, *Cell*, **63**:1061-1072, 1990; Man-6-PR, Canfield, *et al.*, *J. Biol. Chem.*, **266**:5682-5688, 1991; LDLR, Chen, *et al.*, *J. Biol. Chem.*, **265**:3116-3123, 1990), and aromatic and large hydrophobic residues are shown in bold type.

The four-residue sequence, NPVY, from the LDLR was tested because the NPXY pattern is also found in the cytoplasmic tails of other transmembrane proteins including the EGF and insulin receptors and therefore may play a role in the internalization of these proteins (Chen, *et al.*, *J. Biol. Chem.*, 265:3116-3123, 1990; Backer, *et al.*, *J. Biol. Chem.*, 265:16450-16454, 1990). The YSKV sequence from Man-6-PR was tested because mutational analysis of Man-6-PRs (Canfield, *et al.*, *J. Biol. Chem.*, 266:5682-5688, 1991) indicated that this tetrapeptide was almost as efficient as the complete internalization sequence in promoting high efficiency endocytosis.

Efficient expression of wild-type and mutant human TRs in chicken embryo fibroblasts (CEF) was achieved using a helper-independent retroviral vector, BH-RCAS, derived from Rous sarcoma virus (see Materials and Methods). The relative internalization efficiencies of wild-type and mutant receptors were determined from Fe uptake measurements and their intracellular accumulation under steady-state conditions.

Two representative Fe uptake experiments which include all the mutant TRs analyzed are shown in FIGURE 1, and the results of steady-state internalization studies and Fe uptake experiments are summarized in TABLE II. Both assays gave similar values for the internalization efficiencies of mutant TRs relative to wild-type receptors, and the capacity of mutant receptors to mediate Fe uptake was highly reproducible.

TABLE II

COMPARISONS OF STEADY-STATE DISTRIBUTIONS
AND ⁵⁹Fe UPTAKE EFFICIENCIES OF HUMAN TR MUTANTS

Human TR Constructs	% Internalized HTR at Steady-State	Internalization Efficiency (%)	Fe Uptake (atoms/receptor/h)	Internalization Efficiency (%)
Wt ^a	65.2 ± 1.9 ^b (7) ^c	100	34 ± 2.2 (6)	100
Δ3-59 "Tailless"	24.7 ± 1.6 (2)	17	6 ± 1.2 (5) ^d	18
LDLR 4mer	25.7 ± 0.9 (7)	18	8 ± 0.6 (5)	24
LDLR 6mer	40.9 ± 1.6 (7)	37	19 ± 1.9 (5)	56
Man-6-PR 4mer	61.7 ± 3.5 (4)	86	30 ± 3.4 (3)	88
Man-6-PR 6mer	68.5 ± 1.8 (4)	116	33 ± 4.0 (3)	97
F23M	70.9 ± 2.8 (3)	130	37 ± 2.7 (3)	109
F23I	60.5 ± 2.5 (3)	82	33 ± 3.8 (3)	97
F23W	58.7 ± 5.7 (3)	76	27 ± 4.1 (3)	79

^a Wild-type

^b Mean ± standard error

^c Number of independent experiments

^d From Collawn *et al.*, *Cell*, **63**:1061, 1990

Strikingly, mutant TRs containing the Man-6-PR six-residue signal, YKYSKV, were internalized as efficiently as wild-type receptors. Mutant receptors containing the Man-6-PR four-residue signal, YSKV, were also rapidly internalized and had an internalization efficiency of approximately 87% that of wild-type receptors (TABLE II, FIGURE 1A). Mutant TRs containing the LDLR six-residue signal, FDNPVY, were internalized at $\approx$ 50% the rate of wild-type receptors, indicating that the LDLR six-residue signal had significant activity when transplanted into the TR cytoplasmic tail. In contrast, mutant TRs containing the four-residue LDLR signal, NPVY, were internalized at essentially the same rate as tailless receptors (TABLE II, FIGURE 1A and B). The results of these experiments indicate that LDLR and Man-6-PR internalization signals can substitute for the TR signal and mediate high efficiency internalization of the TR. The lack of activity of the transplanted LDLR four-residue sequence, NPVY, implies that the complete LDLR internalization signal is the six-residue sequence, FDNPVY.

The analysis of mutant TRs with transplanted four- or six-residue Man-6-PR and LDLR internalization signals shows that an amino-terminal aromatic residue is required for activity, but that a large non-aromatic hydrophobic residue may substitute for an aromatic residue at the carboxy-terminal position. To investigate whether a carboxy-terminal aromatic residue is required for activity of the TR internalization signal, Phe-23 was altered to Met, then Ile, then Trp. Functional analysis of mutant TRs with Phe-23 altered to Met or Ile indicated that their internalization rate and ability to mediate Fe uptake were identical to wild-type TRs (TABLE II). Mutant TRs with Phe-23 altered to Trp were also rapidly internalized (FIGURES 1 and 2).

EXAMPLE 2

STRUCTURAL ANALYSIS OF INTERNALIZATION SIGNALS

The activity of LDLR and Man-6-PR six-residue internalization signals transplanted into the TR suggests that their active conformations are related to that of the native TR internalization motif. Given the tight turn structural preference of the four-residue TR and LDLR motifs (Collawn, *et al.*, *Cell*, 63:1061, 1990) and the TR internalization activity of the six-residue LDLR and Man-6-PR sequences, the three-dimensional arrangement of critical side chains was investigated in crystallographic protein structures containing six-residue sequences matching the internalization patterns, with the four carboxy-terminal residues in tight turns.

Within their native protein sequences, the TR, LDLR, and Man-6-PR internalization signals were all predicted by a secondary structure algorithm (Chou and Fasman, *Adv. Enzymol.*, 47:45, 1978) to be turns before or at the start of an alpha helix. Four structural analogs were identified for the Man-6-PR signal and five for the LDLR signal (Figure 3).

In Figure 3A, the Man-6-PR internalization motif analogs matching the sequence pattern YXYXKV (detailed in Material and Methods) are shown. Side chains are shown for critical residues at positions -2, 1, 3, and 4. Man-6-PR internalization analogs are: blue, FTFSDY, immunoglobulin 4-4-20 Fab fragment (PDB code 4FAB), residues H27-H32; red, FEFKFK, phosphoglycerate kinase (3PGK), residues 340-345; purple, FMFNQF, concanavalin A (1CN1), residues A128-A133; green, FAFIRL, D-xylose isomerase (4XIA), residues A377-A382. The similar, roughly parallel orientations of the

-36-

aromatic rings paired in positions -2 and 1 (see pairs of same-colored rings, lower left of Figure 3A) suggest how one of the aromatic rings could compensate for the loss of the other, as indicated by mutagenesis. All four functionally important side chains at positions -2, 1, 3, and 4 are simultaneously accessible from one side of the turn, requiring at most a single bond rotation.

In the four Man-6-PR analogs, the position -2 and 1 aromatic side chains (with position 1 being the first position in the turn) were in pairs with similar ring orientations (see like-colored rings, lower left). Of the four analogs, three were reverse turns and one was an initial turn of alpha helix. Distances between C_{β} 's of the aromatic side chains in positions -2 and 1 averaged 6.6Å and were comparable to those found between C_{β} 's in positions 1 and 4 of the turn (averaging 5.8Å), showing a similar spacing for the pairs of critical residues in positions -2 and 1 and in positions 1 and 4. As was also found for the sequence-similar TR motif (Collawn, *et al.*, *Cell*, 63:1061, 1990), the polar side chains in position 3 extended in front of the plane of the turn (projecting out of the page), and the position 4 hydrophobic side chains formed a fan of positions perpendicular to this plane.

It is noteworthy that a truncated mutant Man-6-PR with 5 alanine residues amino-terminal to the internalization signal is significantly more active than the wild-type receptor (Canfield, *et al.*, *J. Biol. Chem.*, 266:5682, 1991), as the alanine residues would be expected to stabilize a reverse or helical turn.

In Figure 3B, the LDLR internalization motif analogs matching the LDLR pattern FXNPXY are shown. Side chains at positions -2, 3, and 4 are shown. LDLR analogs identified by Sequery are: yellow, YANLVF, trypto-

-37-

phan synthetase (1WSY), residues A102-A107; red, FYNAAW, lectin (2LTN), residues A123-A128; orange, FTNEFY, cytochrome c peroxidase (2CYP), residues 198-203; blue, FKNSKY, chymotrypsinogen a (2CGA), residues A89-A94; and green, YDNKYW, calcium-free phospholipase A₂ (1PP2), residues L113-L118. The aromatic groups at positions -2 and 4 and the position 3 side chain were exposed on the same side of the structure.

The LDL analogs included three surface exposed reverse turns and two initial turns of alpha helix. For each of the analogs, the important aromatic groups in positions -2 and 4 were accessible from the same face of the structure. The Asn side chains in position 1 (shown in C) fell in two families, positioned in front or in back of the turn according to whether the incoming secondary structure was to the back or in front. The Asn side chain (position 1) can stabilize the turn by forming hydrogen bonds to adjacent main chain atoms, suggesting that this side chain has a structural role. Asn-Pro, the first two residues in the LDLR tetrapeptide, are the most favored residues for positions 1 and 2 in the most prevalent type of reverse turn (Wilmot, *et al.*, *J. Mol. Biol.*, 203:221, 1988) and Asn-Pro is also a potent helix initiator (Richardson, *et al.*, *Science*, 240:1648, 1988).

Figure 3C shows superimposed Man-6-PR and LDLR analogs. The Man-6-PR analogs from A are blue and the LDLR analogs from B are yellow with side chains shown for positions -2, 1 and 4. The -2 position aromatic groups extend to the left of the turn, and the position 1 Asn (LDLR) and aromatic (Man-6-PR) side chains form perpendicular fans like those found for the aromatic (LDLR) and hydrophobic (Man-6-PR) side chains in position

4. Position 2, Ser in Man-6-PR and Pro in LDLR, appears primarily structural, with the Ser side chains in Man-6-PR in YSKV analogs (L.A.K. and J.A.T., data not shown) forming backbone hydrogen bonds (Richardson, *et al.*, *Prediction of Protein Structure and the Principles of Protein Conformation*, G.D. Fasman, ed., (New York:Plenum Publishing Corp.), pp., 1-98, 1989) mimicking Pro conformations (Wilmot, *et al.*, *Protein Engineering*, 3:479, 1990), and the Pro side chains covalently stabilizing the turn. The structural role of this position is supported by the tendency of these side chains in the six-residue analogs to be somewhat behind the turn, away from the critical side chains. In both the Man-6-PR and LDLR analogs, the side chains in positions -2, 1, and 4 formed a line (.....) across the base of the turn, with the line of critical side chains in LDLR analogs at an angle of $\sim 35^\circ$ to that for Man-6-PR.

An overall structural pattern emerged from superposition of the Man-6-PR and LDLR analogs (Figure 3C): the critical amino-terminal aromatic groups were localized to the left and somewhat to the front of the turn, the important carboxy-terminal aromatic and hydrophobic side chains formed a single cluster, and all the critical side chains in each structure could be recognized simultaneously from one side of the turn. The analogs predicted a structural role for the position 2 side chain in stabilizing the turn, with direct recognition of this side chain unlikely because it was oriented away from the critical side chains. Whereas a tight turn accounts for the pattern of critical residues by being a self-contained structure that places critical residues so that they are simultaneously accessible, it is difficult to see how the motifs could be reconciled with positions within a β sheet. Extended β conformation would place critical residue side chains on opposite sides of the secondary structure, making their simultaneous recognition difficult, and

would require more extensive structure for stability; this has been ruled out by previous deletion mutants.

EXAMPLE 3

TRANSPLANTATION OF MULTIPLE INTERNALIZATION SIGNALS

Studies were performed to examine the effect of multiple internalization signals on TR endocytosis. Structural analysis of the TR cytoplasmic domain using the Chou-Fasman algorithm (Chou and Fasman, *Adv. Enzymol.*, 47:45, 1978) predicts that this domain is composed of a series of helices interrupted by turns with the wild-type internalization signal in one of these turns (turn 2, see Table III).

-40-

TABLE IIITR CYTOPLASMIC TAIL SEQUENCE SHOWING PREDICTED TURN POSITIONS

Turn 1	Turn 2	Turn 3	Turn 4
MMDQARSAF <u>SNL</u>	FGGEPLSY <u>TRF</u>	SLARQVDGDN <u>SHV</u>	EMKLAVDEEEN <u>NADN</u> NTKANVTKPKR-TM*

MUTANTS

Turn 1	<u>YTRF</u>	<u>YTRF</u>	
Turn 3		<u>YTRF</u>	<u>YTRF</u>
Turn 4		<u>YTRF</u>	<u>YTRF</u>

* Transmembrane region

As shown in TABLE III, mutant TRs containing an additional copy of the TR internalization signal YTRF at putative turn positions 1, 3, or 4 were prepared using oligonucleotide-directed mutagenesis. Each mutant retained the wild-type signal at turn 2, and therefore contained two internalization signals. Mutant TRs were tested in steady-state distribution assays in chicken embryo fibroblasts as described. The results of the analysis are shown in Table IV and indicate that additional signals have variable effects and appear to be positionally dependent.

TABLE IV
COMPARISONS OF STEADY-STATE DISTRIBUTIONS OF HUMAN
TR MUTANTS CONTAINING TWO INTERNALIZATION SIGNALS

<u>Human TR Constructs</u>	<u>% HTR Internalized at steady-state</u>	<u>Relative Internalization Efficiency (%)</u>
Wt TR [¹⁸ LSYTRF ²³]	66.0±1.3 ^a (11) ^b	100
Turn 1 [⁹ YTRF ¹²]	65.1±3.6 (3)	96
Turn 3 [³¹ YTRF ³⁴]	79.9±0.8 (3)	205
Turn 4 [⁴⁷ YTRF ⁵⁰]	73.8±1.4 (3)	145

^a Mean + S.E.

^b Number of independent experiments.

The sequence shown in brackets represents the additional internalization signal that was transplanted into the TR using oligonucleotide-directed mutagenesis. Each of the mutants contains the wild-type signal at Turn 2 [²⁰YTRF²³]. The superscripts represent the sequence position in the TR where the changes were made. The distribution of mutant TRs was determined from the surface bound and internal pools of radiolabeled Tf under steady-state conditions.

Addition of an internalization signal at turn 1 creates a TR that is indistinguishable from a wild-type receptor (96% internalization rate compared to wild-type). Addition at turn 3, however, creates a mutant with twice the

-42-

internalization rate of the wild-type ($206\% \pm 30\%$; 3 experiments, mean $\pm$ S.E.). Addition of a signal at turn 4 has a less dramatic effect, but still creates a TR mutant with 145% activity. The results demonstrate that internalization rates can be modulated by transplantation of additional internalization signals.

EXAMPLE 4

IDENTIFICATION OF HETEROLOGOUS INTERNALIZATION SIGNALS

Potential signals from heterologous receptors were tested by transplantation into the TR. In these studies putative internalization signals from other receptors were transplanted into TR using oligonucleotide-directed mutagenesis and tested as described above. The signals examined by this method are shown in Table V.

-43-

TABLE V
COMPARISONS OF STEADY-STATE DISTRIBUTIONS OF HUMAN
TR MUTANTS CONTAINING PUTATIVE INTERNALIZATION
SEQUENCES FROM OTHER RECEPTORS

<u>Human TR Constructs</u>	<u>% HTR Internalized at steady-state</u>	<u>Relative Internalization Efficiency (%)</u>
Wt TR [¹⁸ LSYTRF ²³]	66.0±1.3 ^a (11) ^b	100
ASGPR H1 [²⁰ YQDL ²³]	54.6±2.9 (4)	62
ASGPR H2 [²⁰ FQDI ²³]	23.6±2.1 (5)	16
Poly Ig R [²⁰ YSAF ²³]	52.6±3.5 (5)	57
EGFR [¹⁸ NFYRAL ²³]	69.8±1.6 (5)	119
EGFR [¹⁸ LIPQQGFFS ²⁴]	54.2±1.5 (4)	59
Mannose R [¹⁸ FENTLY ²³]	31.5±2.2 (4)	24
LAP [²⁰ YRHV ²³]	60.3±1.3 (5)	78
LAP [²⁰ PPGY ²³]	27.8±0.9 (3)	20
LAMP-1 [²⁰ YQTI ²³]	48.0±1.5 (6)	48
Tailless TR [Δ 3-59]	24.7±1.6 (2)	17

^a Mean + S.E.

^b Number of independent experiments.

In TABLE V, the sequences shown in brackets represent the putative signals that were transplanted into the TR using oligonucleotide-directed mutagenesis. The superscripts represent the sequence position in the TR where the changes were made. The distribution of mutant TRs was determined from the surface bound and internal pools of radiolabeled Tf under steady-state conditions.

Within this system, some sequences function very well as internalization signals, e.g., sequences from the asialoglycoprotein receptor (ASGPR H1 subunit, sequence YQDL), polymeric immunoglobulin (poly Ig, sequence YSAF), epidermal growth factor receptor (EGFR, sequence NFYRAL), lysosomal acid phosphatase (LAP, sequence YRHV), lysosome-associated membrane glycoprotein (LAMP-1, sequence YQTI), while others, e.g., ASGPR H2 subunit (sequence FQDI), mannose receptor (sequence FENTLY), and LAP (sequence PPGY) function poorly, with activity similar to that of tailless receptors (TABLE V). Once a signal has been localized to a region within the cytoplasmic tail, this method is useful for mapping the exact site. For example, analysis of the region near the tyrosine of LAP, sequence PPGYRHV, shows that one sequence, YRHV, functions well and another, PPGY, does not.

This strategy is also useful for mapping internalization signals from ligand-induced internalized receptors such as the EGFR. This is significant since none of these internalization signals have been clearly identified simply because the large cytoplasmic domains of these receptors makes the interpretation of deletional analysis more difficult. Identifying an internalization signal in a ligand-induced receptor may require additional transplantation experiments with the TR, and parallel mutagenesis experiments within

-45-

the receptor in question may have to be performed. The putative EGFR signal, NFYRAL, which functions very well in the context of the TR, appears to have a sequence related to the cation-independent mannose-6-phosphate receptor. The implication from these studies is that potential signals can be tested in the TR and that ligand-induced receptors may have internalization signals similar to those of the constitutively recycling receptors.

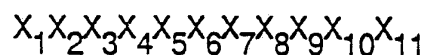
The foregoing is meant to illustrate, but not to limit, the scope of the invention. Indeed, those of ordinary skill in the art can readily envision and produce further embodiments, based on the teachings herein, without undue experimentation.

CLAIMS

1. A method of modulating receptor mediated transport and into a cell, said method comprising introducing heterologous internalization signal into the cell.
2. A method according to claim 1 wherein the internalization signal is introduced as a nucleotide sequence.
3. A method according to claim 2 wherein the nucleotide sequence further comprises flanking sequence encoding a cell surface receptor.
4. A method according to claim 3 wherein the receptor encoded by introduced nucleotide sequence is different from cell surface receptor already present in the cell.
5. A method according to claim 1 wherein at least two internalization signals are introduced.
6. A method according to claim 3 wherein the receptor encoded by introduced nucleotide has substantially the same ligand specificity as cell surface receptor already present in the cell.
7. A method according to claim 6 wherein the internalization signal of receptor encoded by introduced nucleotide differs from the internalization signal of cell surface receptor already present in the cell.

-47-

8. A method according to claim 1 wherein modulation results in increased transport of ligand into the cell.
9. A method according to claim 2 wherein the nucleotide sequence encoding the internalization signal is incorporated in an expression vector.
10. A method according to claim 9 wherein the expression vector is incorporated in a carrier system.
11. A method according to claim 1 wherein the internalization signal is introduced as a peptide.
12. A method according to claim 11 wherein the peptide is incorporated in a carrier system.
13. A method according to claim 1 wherein the internalization signal is characterized by having a tight turn and the amino acid sequence:



where $X_5 - X_8$ constitutes the core sequence,

5 $X_1 - X_4$ residues and $X_9 - X_{11}$ residues are optional and wherein:

X_1 , when present, is leucine or glutamic acid;

X_2 , when present, is isoleucine, methionine or proline;

-48-

X_3 , when present, is any amino acid residue;

10

X_4 , when present, is selected from alanine, polar amino acids, or aromatic amino acids, and when X_3 is also present, at least one of the X_3 or X_4 residues is polar;

X_5 is an aromatic amino acid when residues $X_1 - X_4$ are not present, or is selected from aromatic amino acids or polar amino acids when at least residue X_4 or additional upstream residue(s) is present;

15

X_6 is a polar amino acid or alanine;

X_7 is selected from polar amino acids or alanine when residues $X_1 - X_4$ are not present, or is any amino acid residue when at least X_4 or additional upstream residue(s) is present;

X_8 is selected from aromatic amino acids or hydrophobic amino acids;

20

X_9 , when present, is serine or alanine;

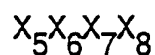
X_{10} , when present, is alanine or leucine; and

X_{11} , when present, is alanine or phenylalanine;

25

wherein at least one of residues X_3 , X_5 , and X_8 is an aromatic amino acid and further wherein residues X_1 , X_2 , X_3 , X_{10} , and X_{11} can only be present when the next adjacent residue(s) relative to the core is present.

14. A method according to claim 13 wherein the internalization signal is characterized by having the amino acid sequence:



wherein X_6 and X_7 cannot both be alanine.

15. A method according to claim 14 wherein:

X_5 is phenylalanine or tyrosine;

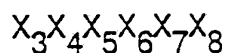
X_6 is alanine, arginine, glutamine, serine, or threonine;

X_7 is alanine, arginine, aspartic acid, glycine, glutamic acid, histidine, lysine, or threonine; and

X_8 is isoleucine, leucine, methionine, phenylalanine, valine, or tryptophan.

16. A method according to claim 15 wherein the amino acid sequence is YTRM, YARF, YTRI, YQDL, YTKF, YSKV, YTRW, YRHV, YSAF, YQTI, YTAF, YTGf, YTEF, FTRF, or YTRF.

17. A method according to claim 13 wherein the internalization signal is characterized by having the amino acid sequence:



-50-

18. A method according to claim 17 wherein:

5 X₃ is asparagine, leucine, methionine, phenylalanine, proline, or tyrosine;

 X₄ is alanine, aspartic acid, glutamine, lysine, phenylalanine, or serine;

 X₅ is asparagine, glutamine, or tyrosine;

 X₆ is arginine, glycine, proline, serine, or threonine;

10 X₇ is alanine, arginine, lysine, phenylalanine, or valine; and

 X₈ is isoleucine, leucine, methionine, phenylalanine, tryptophan, tyrosine or valine.

19. A method according to claim 18 wherein the amino acid sequence is:

 YKYSKV, NFYRAL, LAYTRF, PQQGFF, FDNPVY, MSYTRF, or LSYTRF.

20. A method for identifying a sequence which modulates internalization of a cell surface receptor, the method comprising:

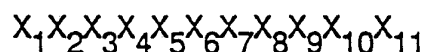
5 (a) incubating cells having such receptors in the presence or absence of the sequence and, optionally, in the presence of ligand for the cell surface receptor; and

 (b) measuring internalization of cell surface receptor in the presence

-51-

or absence of the sequence.

21. A composition comprising a peptide having a tight turn and the amino acid sequence:



where $X_5 - X_8$ constitutes the core sequence,

5 $X_1 - X_4$ residues and $X_9 - X_{11}$ residues are optional and wherein:

X_1 , when present, is leucine or glutamic acid;

X_2 , when present, is isoleucine, methionine or proline;

X_3 , when present, is any amino acid residue;

10 X_4 , when present, is selected from alanine, polar amino acids, or aromatic amino acids, and when X_3 is also present, at least one of the X_3 or X_4 residues is polar;

X_5 is an aromatic amino acid when residues $X_1 - X_4$ are not present, or is selected from aromatic amino acids or polar amino acids when at least residue X_4 or additional upstream residue(s) is present;

15 X_6 is a polar amino acid or alanine;

-52-

X₇ is selected from polar amino acids or alanine when residues X₁ - X₄ are not present, or is any amino acid residue when at least X₄ or additional upstream residue(s) is present;

X₈ is selected from aromatic amino acids or hydrophobic amino acids;

20 X₉, when present, is serine or alanine;

X₁₀, when present, is alanine or leucine; and

X₁₁, when present, is alanine or phenylalanine;

25 wherein at least one of residues X₃, X₅, and X₈ is an aromatic amino acid and further wherein residues X₁, X₂, X₃, X₁₀, and X₁₁ can only be present when the next adjacent residue(s) relative to the core is present. provided that sequences selected from the group consisting of FXNPXY, GPLY, PPGY, and YXYXKV, where X stands for any amino acid, are excluded.

22. A composition comprising a nucleotide sequence encoding a peptide according to claim 21.

23. A method of gene therapy comprising introducing into a host subject, cells derived from the subject and modified to contain heterologous internalization signal capable of modulating transport of ligand into the cells.

-53-

24. A method of gene therapy comprising introducing into the cells of a host subject an expression vector comprising a nucleotide sequence encoding a heterologous internalization signal capable of modulating transport of ligand into the cells.

1/3

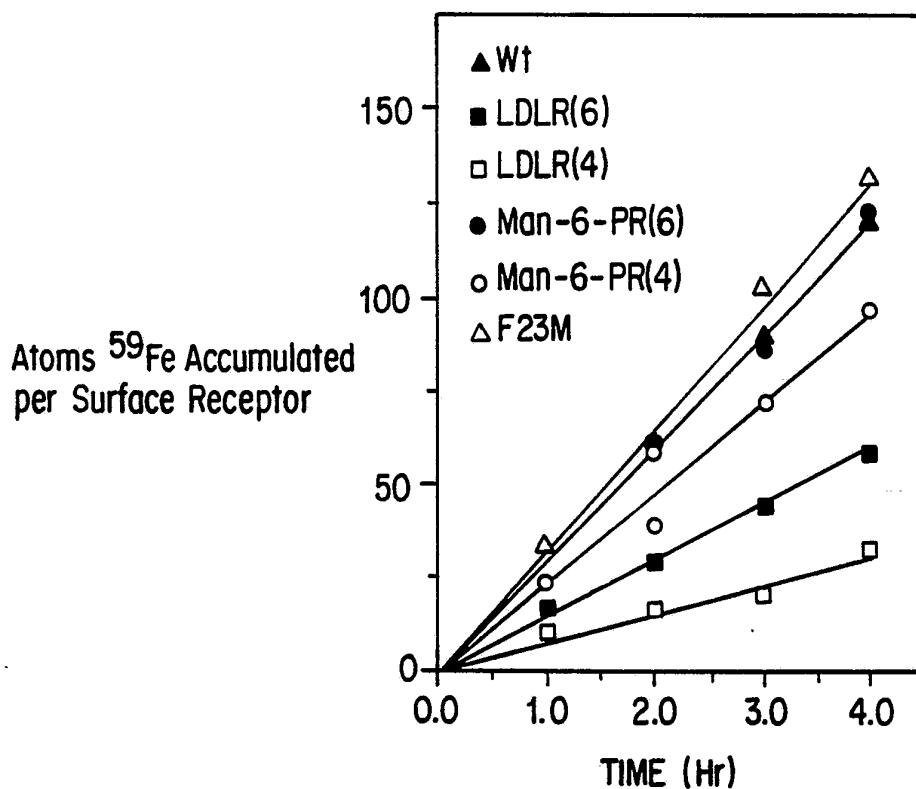


FIG. 1A

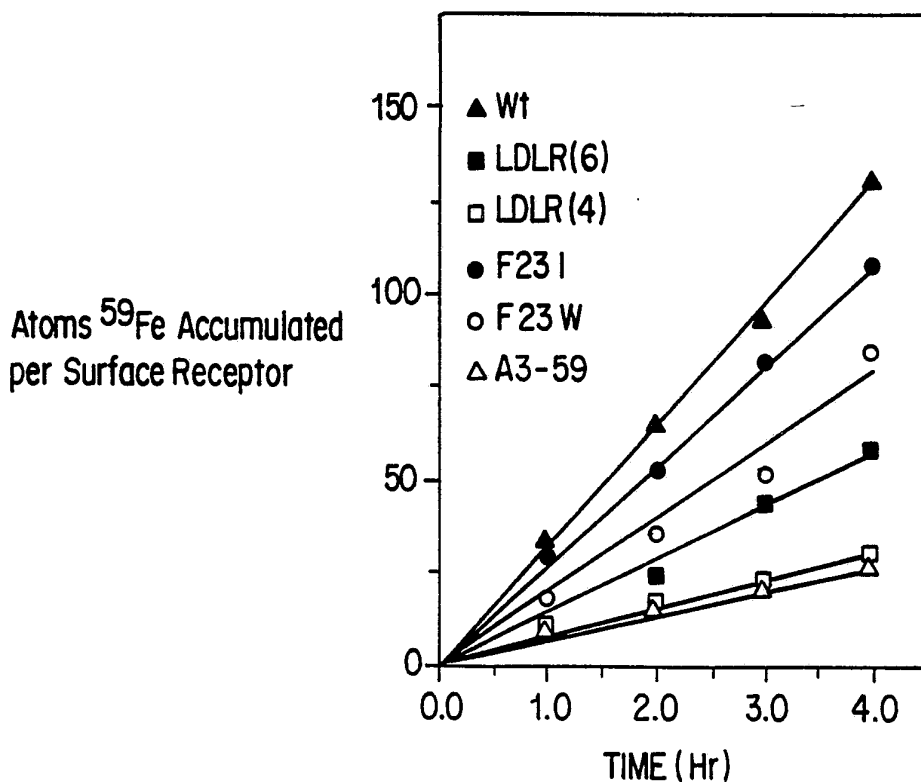


FIG. 1B

SUBSTITUTE SHEET

2/3

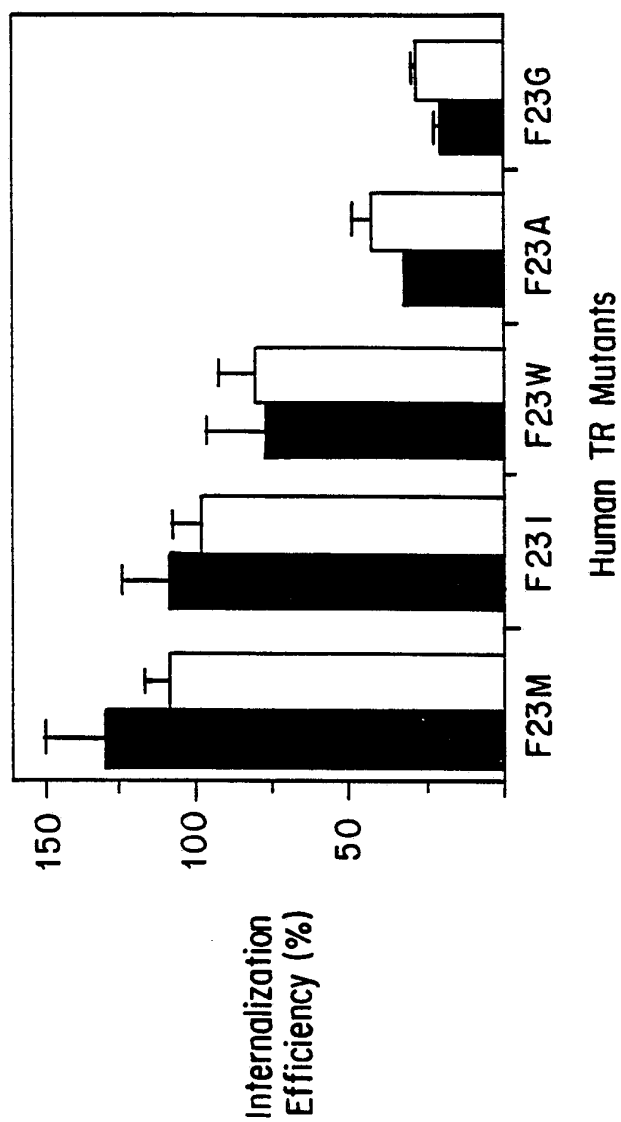


FIG. 2

3/3

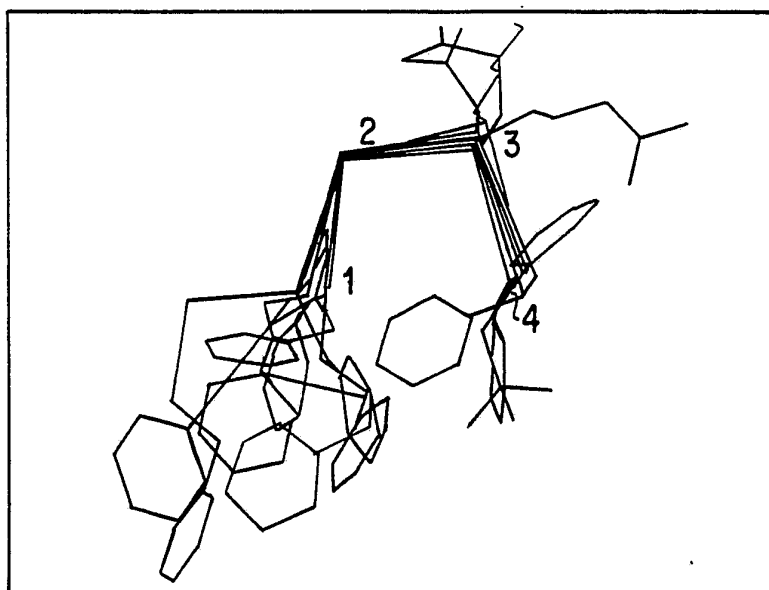


FIG. 3A

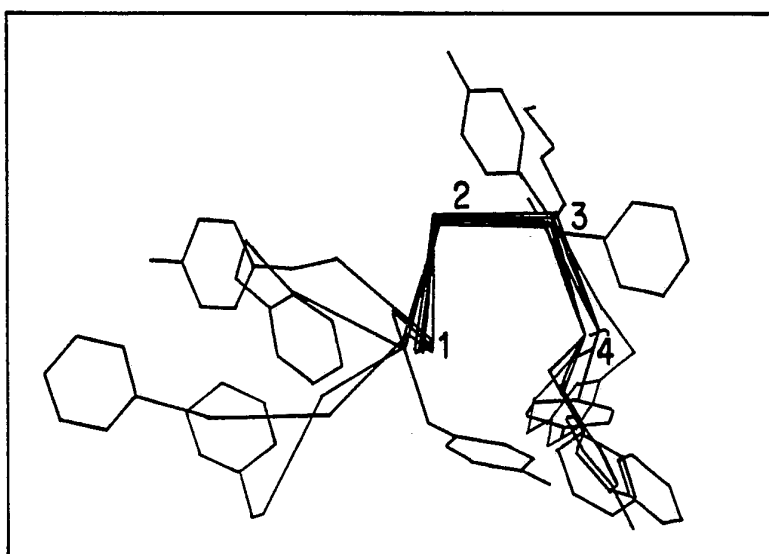


FIG. 3B

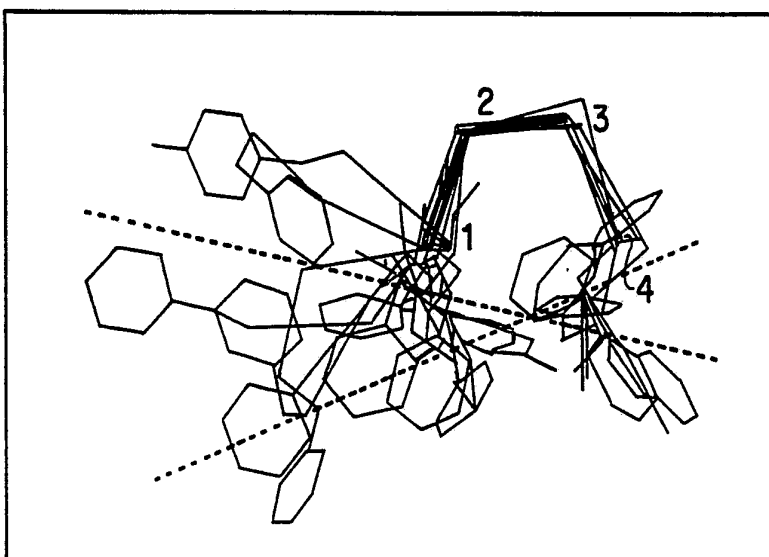


FIG. 3C

SUBSTITUTE SHEET

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/01669**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(5) : Please See Extra Sheet.

US CL : 435/6, 7.1, 7.21, 69.8, 70.1; 536/23.1, 23.2; 424/93

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 7.1, 7.21, 69.8, 70.1; 536/23.1, 23.2; 424/93

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog, Medline

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<u>X</u> Y	The European Molecular Biology Organization Journal, Volume 10, No. 11, issued 1991, J.F. Collawn et al., "Transplanted LDL and mannose-6-phosphate Receptor Internalization Signals Promote High-efficiency Endocytosis of the Transferrin Receptor", pages 3247-3253, see page 3247, table I and page 3248, Table II.	<u>1-19</u> 1-24
<u>X</u> Y	Cell, Volume 63, issued November 1990, J.F. Collawn et al., "Transferrin Receptor Internalization Sequence YXRF Implicates a Tight Turn as the Structural Recognition Motif for Endocytosis", pages 1061-1072, see page 1063, Figure 1 and page 1064, Table III.	<u>1-19</u> 1-24

☒ Further documents are listed in the continuation of Box C.	☐ See patent family annex.
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be part of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family
Date of the actual completion of the international search 26 APRIL 1993	Date of mailing of the international search report 04 MAY 1993
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer KATHLEEN L. CHOI Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/01669

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	A. Meister, "Advances in Enzymology", published 1978 by John Wiley & Sons (N.Y.), see pages 45-148, especially pages 60-64 and 70-74.	1-24
Y	Trends in Biochemical Sciences, issued March 1981, R. Fraley et al., "New Generation Liposomes: the Engineering of an Efficient Vehicle for Intracellular Delivery of Nucleic Acids", pages 77-80, see entire document.	23-24
Y	BioTechniques, Volume 6, no. 7, issued July/August 1988, R.J. Mannino et al., "Liposome Mediated Gene Transfer", pages 682-690, see entire document.	23-24

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/01669

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (5):

C12Q 01/68, 01/00; C12P 21/02, 21/06; G01N 33/53; C07H 17/00; A01N 63/00; A61K 37/00