Transplanted LDL and mannose-6-phosphate receptor internalization signals promote high-efficiency endocytosis of the transferrin receptor

James F. Collawn¹, Leslie A. Kuhn²,
Li-Fung Sue Liu¹, John A. Tainer² and
Ian S. Trowbridge¹

¹Department of Cancer Biology, The Salk Institute, PO Box 85800, San Diego, CA 92186-5800 and ²Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

Communicated by P.A. Peterson

The internalization signals of several constitutively recycling receptors have recently been identified as regions of four or six amino acids that include an aromatic residue, usually tyrosine. Here, we show that transplanted signals from the low density lipoprotein receptor (LDLR) and cation-independent mannose-6-phosphate receptor (Man-6-PR) promote rapid internalization of the transferrin receptor (TR), directly establishing that recognition signals are interchangeable, self-determined structural motifs and that signals from type I membrane proteins are active in a type II receptor. We also show that the chemical and spatial patterns of critical residues in both four- and six-residue internalization motifs are consistent with a tight turn structure. A six-residue LDLR signal is needed for activity in TR, suggesting that an amino-terminal aromatic side chain is obligatory. In contrast, the carboxy-terminal aromatic side chain in the TR signal can be replaced by a large hydrophobic residue. Thus, internalization signals apparently require an aromatic amino-terminal residue and either an aromatic or large hydrophobic carboxy-terminal residue rather than a conserved tyrosine *per se*. Consistent with this conclusion, predicted internalization signals from the poly-Ig receptor, YSAF, and asialoglycoprotein receptor (ASGPR) subunit H1, YQDL, also promote internalization of TR.

Key words: asialoglycoprotein receptor/internalization motif/polymeric immunoglobulin receptor/transferrin receptor/turn conformation

Introduction

High efficiency endocytosis of receptors that constitutively cluster in coated pits is mediated by recognition signals in their cytoplasmic tails (for reviews see Goldstein *et al.*, 1985; Pearse and Robinson, 1990; Trowbridge, 1991). Recently, the internalization signals of the LDLR, TR, and the cation-independent Man-6-PR receptors have been identified as four or six contiguous amino acids including at least one aromatic residue, usually tyrosine (Chen *et al.*, 1990; Collawn *et al.*, 1990; Canfield *et al.*, 1991). We have previously shown that analogs of YTRF, the four-residue TR internalization sequence, and NPVY, the LDLR sequence required for internalization, have a high propensity to be in tight turns in proteins of known structure, implicating

an exposed tight turn as the structural recognition motif for high efficiency endocytosis (Collawn *et al.*, 1990).

To directly determine whether internalization signals are interchangeable, self-determined structural motifs and to investigate the structural and chemical requirements of both four- and six-residue signals independent of their original environment, we replaced the TR internalization signal with the internalization signals of the LDLR and cation-independent Man-6-PR. Our results indicate that transplanted signals from the LDLR and cation-independent Man-6-PR promote rapid internalization of the TR, and establish that all three recognition signals are functionally equivalent and encode interchangeable, self-determined structural motifs that are active in both type I and type II membrane proteins. Furthermore, examination of analogs of the six-residue LDLR and Man-6-PR sequences in proteins of known structure shows that the overall pattern of critical residue positions in both four- and six-residue motifs is consistent with a tight turn structure, indicating a common structural chemistry for endocytosis. We also show that putative internalization signals from the poly-Ig and ASGPRs are active when transplanted into the cytoplasmic tail of the TR.

Results

LDLR and Man-6-PR internalization signals promote TR endocytosis

To determine if internalization signals of recycling receptors are interchangeable, we transplanted the signals from the LDLR and the Man-6-PR into the TR. Oligonucleotide-directed mutagenesis was used to replace the four-residue YXRF internalization signal from the cytoplasmic tail of the wild type TR with the four-residue sequences from the LDLR and Man-6-PR signals and the six-residue TR

Table I. Internalization sequences in transferrin receptor constructs

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

^aFor 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 FDNVY. Numbers above the sequences (–2 through 4) refer to positions relative to the turn region implicated in TR and LDLR (Collawn *et al.*, 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.*, 1990; Man-6-PR, Canfield *et al.*, 1990; LDLR, Chen *et al.*, 1990), and aromatic and large hydrophobic residues are shown in bold type.

Fig. 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 (▲), $^{20}\text{YTRF}^{23}$ -NPVY(□), $^{18}\text{LSYTRF}^{23}$ -FDNPVY(■), $^{20}\text{YTRF}^{23}$ -YSKV(○), $^{18}\text{LSYTRF}^{23}$ -YKYSKV(●), and F23M (△) mutant TRs. **Panel B** shows the relative ^{59}Fe uptake for wild type (▲), $^{20}\text{YTRF}^{23}$ -NPVY(□), $^{18}\text{LSYTRF}^{23}$ -FDNPVY(■), F23W(○), F23I(●), and $\Delta 3$ -59 (△) mutant human TRs.

Table II. Comparisons of steady-state distributions and ^{59}Fe 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
$\Delta 3$ -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)	109	33 ± 3.8 (3)	97
F23W	58.7 ± 5.7 (3)	76	27 ± 4.1 (3)	79

^aWild type

^bMean ± standard error

^cNumber of independent experiments

^dFrom Collawn *et al.* (1990)

sequence LSYTRF by the six-residue LDLR and Man-6-PR signals (Table I). 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.*, 1990; Backer *et al.*, 1990). We tested the YSKV sequence from Man-6-PR, because mutational analysis of Man-6-PRs (Canfield *et al.*, 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 then determined from iron 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 iron uptake was highly reproducible (Table II and Figure 1). 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 ~87% that of wild type receptors (Table II, Figure 1A). Mutant TRs containing the LDLR six-residue signal, FDNPVY, were internalized at ~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.

Carboxy-terminal hydrophobic residues within the internalization motif maintain high efficiency endocytosis

The analysis of mutant TRs with transplanted four- or six-residue Man-6-PR and LDLR internalization signals strongly suggests that an amino-terminal aromatic residue is required for activity, but that a large 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, Phe23 was altered to Met, then Ile, then Trp. Functional analysis of mutant TRs with Phe23 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 Phe23 altered to Trp were also rapidly internalized (Figures 1 and 2).

Six-residue internalization signals are compatible with a tight turn structure

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.*, 1990) and the TR internalization activity of the six-residue LDLR and Man-6-PR sequences, we investigated the three-dimensional arrangement of critical side chains 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, 1978) to be turns before or at the start of an α helix. Four structural analogs were identified for the Man-6-PR signal and five for the LDLR signal (Figure 3). For the Man-6-PR structural analogs (Figure 3A), the similar, roughly parallel orientations of the aromatic rings paired in positions -2 and 1 (see pairs of same-coloured 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. 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.*, 1991), as the alanine residues would be expected to stabilize a reverse or helical turn. For the LDLR analogs (Figure 3B), the aromatic groups at positions -2 and 4 and the position 3 side chain were exposed on the same side of the structure. 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. An overall structural pattern emerged from superposition of the Man-6-PR and LDLR analogs (Figure 3C): the critical amino-terminal

Fig. 2. Comparisons of steady-state distributions (■) and ⁵⁹Fe 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.* (1990). For the steady-state analysis, CEFs expressing human TRs were incubated with ¹²⁵I-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.

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 or α helix. Both internal α helical turns and extended β conformations would place critical residue side chains on opposite sides of the secondary structure, making their simultaneous recognition difficult, and both would require more extensive structure for stability; this has been ruled out by previous deletion mutants.

Activity of putative internalization signals of Poly-Ig receptor and ASGPR

The internalization signals of other receptors that participate in high efficiency endocytosis have not yet been defined. It is known, however, that specific tyrosine residues in the cytoplasmic tails of the poly-Ig receptor (Breitfeld *et al.*, 1990), human ASGPR (Spiess, 1990) and two lysosomal membrane glycoproteins, lysosomal acid phosphatase (Peters *et al.*, 1990) and human Lamp-1 (Lgp-A) (Williams and Fukuda, 1990) are important for rapid internalization. Based on the sequence patterns of the internalization signals of the LDLR, Man-6-PR and TR, it is possible to identify putative tetrapeptide sequences that include the tyrosine residues critical for internalization in the cytoplasmic domains of these receptors (Canfield *et al.*, 1991; Trowbridge, 1991). We have transplanted the putative internalization signals of the poly-Ig receptor, YSAF, and the human asialoglycoprotein

Fig. 3. Superimposed crystallographic turn structures for internalization motif analogs with sequences matching the six-residue Man-6-PR and LDLR patterns. (A) Man-6-PR internalization motif analogs. Side chains are shown for critical residues at positions -2 , 1 , 3 , and 4 . In the four analogs matching the Man-6-PR YXYXKV sequence pattern (detailed in Materials and methods), 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 same-colored rings, lower left). Of the four analogs, three were reverse turns and one was an initial turn of α helix. Distances between C_{β} s of the aromatic side chains in positions -2 and 1 averaged $6.6 \text{ \AA}$ and were comparable to those found between C_{β} s in positions 1 and 4 of the turn (averaging $5.8 \text{ \AA}$), 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.*, 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. Man-6-PR internalization analogs were: blue, FTFSKY, immunoglobulin 4-4-20 Fab fragment (PDB code 4FAB), residues H27–H32; red, FEFKVF, phosphoglycerate kinase (3PGK), residues 340–345; purple, FMFNQF, concanavalin A (1CN1), residues A128–A133; green, FAFIRL, D-xylose isomerase (4XIA), residues A377–A382. (B) LDLR internalization motif analogs. Side chains in positions -2 , 3 and 4 are shown. Three surface-exposed reverse turns and two initial turns of α helix were identified for sequences similar to the LDLR pattern FXNPXY. 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 positions 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. 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 and Thornton, 1988) and Asn-Pro is also a potent helix initiator (Richardson and Richardson, 1988). LDLR analogs identified by Sequery were: yellow, YANLVF, tryptophan 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 A_2 (1PP2), residues L113–L118. (C) Superimposed Man-6-PR and LDLR analogs. The Man-6-PR analogs (from A) appear in blue and the LDLR analogs (from B) in yellow, with side chains shown for positions -2 , 1 , and 4 . The -2 position aromatic groups extended to the left of the turn, and the position 1 Asn (LDLR) and aromatic (Man-6-PR) side chains formed 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 and Richardson, 1989) mimicking Pro conformations (Wilmot and Thornton, 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 ($\cdots$) 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.

receptor, YQDL, into the TR to test their internalization activity. The results of steady-state internalization assays indicate that both signals have significant activity; mutant receptors containing the poly-Ig receptor internalization signal were internalized with an efficiency of $60.1 \pm 5.4\%$ of wild type TRs and mutant receptors containing the ASGPR internalization signal were internalized with an efficiency of $95.2 \pm 9.1\%$ (mean $\pm$ standard error of five determinations).

Discussion

Our results demonstrate that LDLR and Man-6-PR internalization sequences can substitute for the TR internalization sequence and promote high efficiency endocytosis of human TRs and establish that these two six-residue recognition signals are functionally equivalent to the four-residue TR signal. The ability to transplant functional internalization signals from one receptor to another further demonstrates that these short sequences encode interchangeable, self-determined structural motifs. Our results also establish that internalization signals from type I membrane proteins are active in a type II membrane protein. This observation is consistent with the proposal that internalization signals are self-determined tight turns, as the orientation of a turn in three-dimensional space is independent of the polypeptide chain's orientation with respect to the cell membrane. This result is not surprising once structure is taken into account, as the internalization signal is seen as a three-dimensional motif (with no preferred direction) by the recognition molecule with which it interacts.

The activity of the transplanted signals in the TR shown here, as well as previous data showing that residues 19–28 of the human TR cytoplasmic domain are sufficient for high efficiency endocytosis (Jing *et al.*, 1990), argues against the suggestion that residues in other regions of the TR cytoplasmic tail are required to generate an internalization signal (McGraw *et al.*, 1991), and indicates that such results may reflect indirect or secondary effects. It is not clear whether the lower activity of the LDLR six-residue signal transplanted into TR is because this signal is intrinsically less active than the TR signal or because the structural environment of the TR cytoplasmic tail is not optimal. Given the failure of the NPVY tetrapeptide to promote internalization of the TR, the functional significance of the NPXY sequences found in the cytoplasmic tails of other receptors requires further study. As previously noted (Chen *et al.*, 1990), deletion of the NPXY sequences in the C-terminal region of the EGF receptor cytoplasmic domain does not affect ligand-induced internalization (Livneh *et al.*, 1986).

With regard to the structural chemistry of internalization motifs, following the observation that Tyr807 is a critical element of the LDLR internalization signal (Davis *et al.*, 1986, 1987), tyrosine residues have been implicated in the high efficiency endocytosis of various other proteins (Lazarovits and Roth, 1988; Lobel *et al.*, 1989; Breitfeld *et al.*, 1990; Jing *et al.*, 1990; Johnson *et al.*, 1990; Peters *et al.*, 1990; Spiess, 1990; Williams and Fukuda, 1990). However, the tyrosine residue of the LDLR signal is in the carboxy-terminal position, whereas the tyrosine residues of the TR and Man-6-PR signals are in the amino-terminal position. Further, critical tyrosine residues can often be altered to other aromatic residues without significant loss of activity (Davis *et al.*, 1987; McGraw and Maxfield, 1990; Canfield *et al.*, 1991). Our transplantation results indicate that, rather than a conserved tyrosine, the essential requirements for TR internalization are a tight turn with an amino-terminal aromatic residue and either a carboxy-terminal aromatic or large hydrophobic residue. The LDLR internalization signal differs from the TR and Man-6-PR signals in the apparent requirement for a carboxy-terminal aromatic residue (Davis *et al.*, 1987). The four- or six-residue size of internalization signals and requirement for an amino-terminal aromatic residue and either a carboxy-terminal aromatic or large hydrophobic residue provides sufficient information that the putative internalization signals of other constitutively internalized proteins can be predicted (Canfield *et al.*, 1991; Trowbridge, 1991). Our results with the signals from the poly-Ig receptor and ASGPR show that such putative internalization signals can be directly tested experimentally by transplantation into TR independently of whether they are derived from type I or type II membrane proteins. Several tyrosine kinase growth factor receptors also contain similar four-residue sequences in their cytoplasmic domains that may play a role in ligand-induced internalization (Trowbridge, 1991).

The results presented here demonstrate that a tight turn conformation for the six-residue LDLR and Man-6-PR signal analogs accounts for the accessibility of critical residues in both four-residue and six-residue internalization signals. The two additional residues in the six-residue internalization sequences may place an aromatic side chain near the beginning of the tight turn and compensate for the absence of an aromatic side chain at the first position of the turn.

The similarity in three-dimensional placement of critical residues suggests that both types of signals can interact with the same recognition structure in coated pits. Recently, the β -turn structure of the LDLR internalization signal has been confirmed by NMR analysis of nonapeptides containing the NPVY sequence, and the effect of alanine substitutions at different positions of the turn on internalization efficiency were related to their disruption of turn structure (Bansal and Gierasch, 1991). An alternative model has recently been described in which internalization signals are proposed to be small surface loops of no more than 8–10 amino acids with polar or basic residues preferred at certain positions relative to a tyrosine residue (Ktistakis *et al.*, 1990). This loop model is useful in suggesting that internalization signals do not require extensive secondary or tertiary structure; however, the loop hypothesis does not explain the observed four- to six-residue size of these signals or the placement and accessibility of their critical residues. Nor does the loop model provide a basis for the interchangeability of the TR, LDLR, and Man-6-PR signals demonstrated here or identify any specific structural similarity underlying these four- and six-residue signals having different sequence patterns. In contrast, the tight turn internalization motif hypothesis provides a specific structural basis for the four- and six-residue size of these internalization signals, for their interchangeability, for the location of their critical residues, and for their relative independence from other secondary and tertiary structure within the cytoplasmic tail of the receptor.

The ability to transplant internalization signals from one protein to another and retain activity has important implications for control of endocytosis for scientific and medical purposes, including drug delivery to cells. The efficiency of known internalization signals can now be ranked within the same receptor and assay system allowing a quantitative comparison of their activities within a common structural context. Further, as demonstrated for the signals of the poly-Ig receptor and ASGPR, predicted internalization signals from both type I and type II membrane proteins can be assayed for activity in the TR system. The hypothesis that a tight turn forms the structural recognition motif for endocytosis predicted the interchangeability of internalization signals from different receptors, which we have demonstrated here. Our current results provide the essential, detailed knowledge of the conformational chemistry of internalization motifs that will enable the design of small molecules to mimic the structure of internalization motifs and to block endocytosis. These new results further suggest that it will also be possible to transplant internalization motifs into the cytoplasmic domains of cell surface molecules that lack a signal and thereby enhance their endocytosis.

Materials and methods

Oligonucleotide site-directed mutagenesis

A *Clal* fragment containing the entire coding region of the human TR was cloned into the phagemid pBluescript SK (Stratagene, La Jolla, CA) (Jing *et al.*, 1990). Oligonucleotide site-directed mutagenesis was performed with pBluescript SK phagemid templates of the wild type TR by the method of Kunkel (1985) using the Muta-gene mutagenesis kit (Bio-Rad, Richmond, CA) as described previously (Collawn *et al.*, 1990). Mutants were selected by restriction mapping and *Clal* fragments encoding the mutant receptors were then excised and cloned into the retroviral expression vector BH-RCAS (Hughes *et al.*, 1987; S.H. Hughes, unpublished results). The mutations were verified by dideoxynucleotide sequencing (Sanger *et al.*, 1977; Tabor and

Richardson, 1987) of the BH-RCAS constructs using the Sequenase kit (USB Corp., Cleveland, OH) according to the manufacturer's directions.

Cell culture and expression of human TRs in CEF

C/E CEF were prepared from fertilized eggs (SPAFAS, Norwich, CT) and grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 1% (v/v) chicken serum, 1% (v/v) defined bovine calf serum (Hyclone, Logan, UT), 2% (v/v) tryptose phosphate broth (Difco, Detroit, MI). CEF were transfected with 30 µg retroviral construct DNA per 10 cm tissue culture plate of ~40% confluent cells using the polybrene-dimethyl sulfoxide method (Kawai and Nishizawa, 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.*, 1987; S.H. Hughes, unpublished). 1–2 weeks 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 µg/ml diferric human transferrin (Tf) (Miles Scientific, Naperville, IL)] starting 10–14 days after transfection (Jing *et al.*, 1990).

¹²⁵I-Tf binding

Diferric human Tf was labeled with ¹²⁵I to a specific activity of 2–4 µCi/mg using Enzymobeads (Bio-Rad, Richmond, CA) according to the manufacturer's directions. Cells were plated at a density of 7.5 × 10⁴ cells/cm² in 24-well Costar tissue culture plates 24 h prior to the binding assay. Cells were incubated in serum-free DMEM for 1 h at 37°C and then washed once with ice-cold 0.15 M NaCl–0.01 M Na phosphate buffer (pH 7.4) containing 0.1% bovine serum albumin (BSA-PBS). ¹²⁵I-Tf (4 µg/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 1 M NaOH, and the radioactivity counted in a gamma counter.

Analysis of Tf internalization at steady-state

For studies of the steady-state distribution of mutant and wild type human TRs, CEF were plated in triplicate wells as described for the binding studies. The cells were first incubated for 1 h at 37°C in serum-free DMEM, and then incubated with 4 µg/ml ¹²⁵I-Tf in 0.1% BSA in DMEM for 1 h 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 0.2 M acetic acid–0.5 M NaCl (pH 2.4) to remove surface-bound ¹²⁵I-Tf (Hopkins and Trowbridge, 1983). Cells were then removed from the wells with 1 M NaOH and radioactivity in the acid wash and the cell lysate was determined. More prolonged incubation with the acid wash did not affect the radioactivity released (Jing *et al.*, 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.*, 1990).

⁵⁹Fe uptake measurements

Human apoTf was labeled with ⁵⁹Fe (FeCl₃; Amersham Corp., Arlington Heights, IL) to a specific activity of 5–10 µCi/mg using nitrilotriacetate (Bates and Schlabach, 1973). Cells were plated at a density of 7.5 × 10⁴ cells/cm² in 24-well Costar tissue culture plates 24 h 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 µg/ml ⁵⁹Fe-Tf at 37°C for 0, 1, 2, 3 and 4 h. 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 1 M 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 h at 37°C in serum-free DMEM, triplicate wells of cells were incubated with 4 µg/ml ¹²⁵I-labeled Tf on ice for 1 h and then washed three times with 1 ml of ice-cold BSA-PBS, and the radioactivity bound to the cells was determined.

Comparison of protein structure analogs of six-residue internalization motifs

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.*, 1977) for sequences matching the internalization patterns and having the final four residues in a surface-exposed tight turn (a turn of α helix 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.*, 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 ≠ Gly for both patterns. Homologous matches and matches lacking side chain 3-dimensional coordinates were deleted, and those structures containing 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 workstation.

Acknowledgements

We thank Stuart Kornfeld for discussion and for communicating results prior to their publication, Sandra Schmid and Elizabeth Getzoff for discussion, Gary Liao and Michael Pique for computer program development, and Joan Stewart for word processing. This work was supported by the NIH grant CA 34787 (I.S.T.) and the NSF grant 8822385 (J.A.T.), and by an Arthritis Foundation Fellowship to J.F.C., and an NIH Fellowship CA 08939 to L.A.K.

References

- Backer, J.M., Kahn, C.R., Cahill, D.A., Ullrich, A. and White, M.F. (1990) *J. Biol. Chem.*, **265**, 16450–16454.
- Bansal, A. and Gierasch, L.M. (1991) *J. Cell Biochem. Suppl.* **15G**, 171.
- Bates, G.W. and Schlabach, M.R. (1973) *J. Biol. Chem.*, **248**, 3228–3232.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr, Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.*, **193**, 775–791.
- Breitfeld, P.P., Casanova, J.E., McKinnon, W.C. and Mostov, K.E. (1990) *J. Biol. Chem.*, **265**, 13750–13757.
- Canfield, W.M., Johnson, K.F., Ye, R.D., Gregory, W. and Kornfeld, S. (1991) *J. Biol. Chem.*, **266**, 5682–5688.
- Chen, W.-J., Goldstein, J.L. and Brown, M.S. (1990) *J. Biol. Chem.*, **265**, 3116–3123.
- Chou, P.Y. and Fasman, G.D. (1978) *Adv. Enzymol.*, **47**, 45–148.
- Collawn, J.F., Stangel, M., Kuhn, L.A., Esekogwu, V., Jing, S., Trowbridge, I.S. and Tainer, J.A. (1990) *Cell*, **63**, 1061–1072.
- Davis, C.G., van Driel, I.R., Russell, D.W., Brown, M.S. and Goldstein, J.L. (1987) *J. Biol. Chem.*, **262**, 4075–4082.
- Davis, C.G., Lehrman, M.A., Russell, D.W., Anderson, R.G.W., Brown, M.S. and Goldstein, J.L. (1986) *Cell*, **45**, 15–24.
- Goldstein, J.L., Brown, M.S., Anderson, R.G.W., Russell, D.W. and Schneider, W.J. (1985) *Annu. Rev. Cell Biol.*, **1**, 1–39.
- Hopkins, C.R. and Trowbridge, I.S. (1983) *J. Cell Biol.*, **97**, 508–521.
- Hughes, S.H., Greenhouse, J.J., Petropoulos, C.J. and Suttrave, P. (1987) *J. Virology*, **61**, 3004–3012.
- Jing, S., Spencer, T., Miller, K., Hopkins, C. and Trowbridge, I.S. (1990) *J. Cell Biol.*, **110**, 283–294.
- Johnson, K.F., Chan, W. and Kornfeld, S. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 10010–10014.
- Kawai, S. and Nishizawa, M. (1984) *Mol. Cell. Biol.*, **4**, 1172–1174.
- Ktistakis, N.T., Thomas, D. and Roth, M.G. (1990) *J. Cell. Biol.*, **111**, 1393–1407.
- Lazarovits, J. and Roth, M. (1988) *Cell*, **53**, 743–752.
- Livneh, E., Prywes, R., Kashles, O., Reiss, N., Sasson, I., Mory, Y., Ullrich, A. and Schlessinger, J. (1986) *J. Biol. Chem.*, **261**, 12490–12497.
- Lobel, P., Fujimoto, K., Ye, R.D., Griffiths, G. and Kornfeld, S. (1989) *Cell*, **57**, 787–796.
- McGraw, T.E. and Maxfield, F.R. (1990) *Cell Regul.*, **1**, 369–377.
- McGraw, T.E., Pytowski, B., Arzt, J. and Ferrone, C. (1991) *J. Cell Biol.*, **112**, 853–861.
- Pearse, B.M.F. and Robinson, M.S. (1990) *Annu. Rev. Cell Biol.*, **6**, 151–171.
- Peters, C., Braun, M., Weber, B., Wendland, M., Schmidt, B., Pohlmann, R., Waheed, A. and Von Figura, K. (1990) *EMBO J.*, **9**, 3497–3506.
- Richardson, J.S. and Richardson, D.C. (1988) *Science*, **240**, 1648–1652.
- Richardson, J.S. and Richardson, D.C. (1989) In Fasman, G.D. (ed.), *Prediction of Protein Structure and the Principles of Protein Conformation*. Plenum Publishing Corp., New York, pp. 1–98.

- Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Spiess,M. (1990) *Biochemistry*, **29**, 10009–10018.
- Tabor,S. and Richardson,C.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4767–4771.
- Trowbridge,I.S. (1991) *Curr. Opin. Cell Biol.*, **3**, 634–641.
- Williams,M.A. and Fukuda,M. (1990) *J. Cell Biol.*, **111**, 955–966.
- Wilmot,C.M. and Thornton,J.M. (1988) *J. Mol. Biol.*, **203**, 221–232.
- Wilmot,C.M. and Thornton,J.M. (1990) *Protein Engng.*, **3**, 479–493.

Received on May 17, 1991; revised on July 22, 1991