Characterization of the Signal for Rapid Internalization of the Bovine Mannose 6-Phosphate/Insulin-like Growth Factor-II Receptor*

Michel Jadot†, William M. Canfield‡, Walter Gregory, and Stuart Kornfeld¶

From the Department of Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

The signal for rapid internalization of the mannose 6-phosphate/insulin-like growth factor II receptor has been localized to the sequence Tyr-Lys-Tyr-Ser-Lys-Val in positions 24–29 of its 163-residue cytoplasmic tail. Most of the activity of this signal is mediated by the carboxyl 4 amino acids, especially Tyr25 and Val28 (Canfield, W. M., Johnson, K. F., Ye, R. D., Gregory, W. and Kornfeld, S. (1991) J. Biol. Chem. 266, 5682–5688). In this study, we have tested the effect of a series of mutations on the internalization rate of a mutant receptor that contains a 29-amino acid cytoplasmic tail terminating with the 4-amino acid internalization sequence Tyr-Ser-Lys-Val. Replacement of Tyr28 with Phe or Trp gave rise to mutant receptors that were internalized at 10% the wild-type rate, while receptors with Ala, Leu, Ile, Val, or Asn at this position were totally inactive. Val29 could be replaced by other large hydrophobic residues (Phe, Leu, Ile, or Met) with no loss of activity, but the presence of Ala, Gly, Arg, Gln, or Tyr in this position inactivated the signal. Ser27 could be effectively replaced by many different amino acids, but not by Pro or Gly. However, Gly27 could be tolerated if the residues at positions 28 and 29 were also changed. A change in the 2-residue spacing between Tyr28 and Val29 destroyed the signal. These data show that the essential elements of this signal are an aromatic residue, especially a Tyr in the first position, separated from a large hydrophobic residue in the last position by 2 amino acids. The residues in positions 2 and 3 of the signal may have a modulating effect on its activity.

The Tyr-Ser-Lys-Val signal could be moved to a more proximal region of the cytoplasmic tail with only a modest loss of activity. In addition, the signal could be effectively replaced by the putative 4-residue signals of seven other receptors and membrane proteins known to undergo rapid endocytosis, including the Tyr-Thr-Arg-Phe sequence of the transferrin receptor, a Type II membrane protein. These results are compatible with the 4-residue signals of this type being interchangeable, even among Type I and Type II membrane proteins.

The 275-kDa cation-independent mannose 6-phosphate/insulin-like growth factor-II receptor (M6P/IGF-II receptor)1 is a Type I transmembrane protein that plays a key role in the biogenesis of lysosomes (1, 2). This receptor functions in the shuttling of newly synthesized lysosomal hydrolases from the Golgi to prelysosomal compartments and in the retrieval of lysosomal enzymes from the extracellular medium. The receptor also serves to internalize extracellular IGF-II (1, 2). In the endocytic pathway, the receptor with its bound ligand is initially concentrated in clathrin-coated pits at the plasma membrane and is then internalized via clathrin-coated vesicles, the first organelle involved in the vesicular transport of the ligand to its final destination, the lysosome (3). The ability of the M6P/IGF-II receptor to be efficiently concentrated in the clathrin-coated pits depends on the presence of an internalization signal in its cytoplasmic tail which is presumed to interact with the adaptin proteins of the coated pits (4, 5). In previous studies we localized this signal to the sequence Tyr24, Lys-Tyr-Ser-Lys-Val29 of the 163-residue cytoplasmic tail (6, 7). Alanine scanning mutagenesis identified Tyr24 and Val29 as the most important residues for rapid receptor internalization. Tyr24 and Lys28 also contributed to the signal, while the other amino acids were not critical.

The requirement for a tyrosine residue as a component of the internalization signal has been demonstrated for a number of receptors and membrane proteins (8–16), and in a few of these cases it has been shown that the tyrosine must be in the proper context relative to the surrounding amino acids to be functional (8, 10–12, 17).

When we compared the YKYSKV sequence with the sequences neighboring the critical tyrosines in the cytoplasmic tails of other proteins known to undergo rapid internalization, we noted that the crucial elements of the internalization sequence were present in a number of these proteins (7). In particular, the common features were a 4- or 6-amino acid motif with an aromatic residue, usually a Tyr, in the first position and a large hydrophobic residue (Val, Phe, or Ile) in the last position. Taken together, these data were consistent with the notion that the requirement for an effective internalization signal is a general motif rather than a specific amino acid sequence. Collawn et al. (10, 18) and Ktistakis et al. (17) studying the internalization signals of the transferrin receptor and the influenza virus hemagglutinin, respectively, came to a similar conclusion.

* This work was supported in part by United States Public Health Service Grant CA08759 and a Monsanto/Washington University biomedical research grant. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

‡ Supported by a NATO fellowship and a Fulbright grant. Current address: Laboratoire de Chimie Physiologique, Facultes Universitaires Notre-Dame de la Paix, 61, rue de Bruxelles, B-5000 Namur, Belgium.

§ Supported by National Institutes of Health Training Grant T32-HLO7088. Current address: University of Oklahoma Health Sciences Center, P. O. Box 26801, BSEB 302, Oklahoma City, OK 73190.

† To whom correspondence should be addressed: Div. of Hematology-Oncology, Washington University School of Medicine, 660 South Euclid Ave., Box 8125, St. Louis, MO 63110.

1 The abbreviations used are: M6P/IGF-II receptor, mannose 6-phosphate/insulin-like growth factor-II receptor; Man-6-P, mannose 6-phosphate, MEM–ø, minimal essential medium–ø; bp, base pair(s); PBS, phosphate-buffered saline.
In this study we have analyzed the effect of a series of mutations on the internalization rate of a mutant M6P/IGF-II receptor that contains a 29-amino acid cytoplasmic tail terminating with the 4-amino acid internalization sequence Y*SKVy. The goal was to define the structural requirements at each position of the 4-amino acid internalization motif and to determine if the putative internalization motifs of other recycling receptors and membrane proteins would function when transplanted onto the M6P/IGF-II receptor. Our results indicate that the internalization signal must contain an aromatic residue (preferably a Tyr) in the first position and a large hydrophobic residue in the last position in order to be functional. In addition, the spacing between these two residues is critical. We also demonstrate that the YSKV sequence is specific for a target residue (preferably a Tyr) in the first position and a large hydrophobic residue in the last position in order to be functional.

**EXPERIMENTAL PROCEDURES**

**Materials**—Enzymes used in molecular cloning were obtained from New England Biolabs, Bethesda Research Laboratories, or Promega. Lactoperoxidase (bovine milk), H2O2, Sephadex G-150 (120 mesh), and bovine serum albumin were obtained from Sigma. Lipofectin and G418 (neomycin sulfate) were purchased from Gibco-BRL, and [35S]ATP was obtained from Amerah Corp. Other reagent chemicals were obtained from standard suppliers.

**Cassette Mutagenesis of M6P/IGF-II Receptor cDNA**—Oligonucleotides were synthesized with an Applied Biosystems model 380A solid phase synthesizer by the Protein Chemistry Facility of Washington University, St. Louis, MO. In most cases the 5' dimethoxy-trityl oligonucleotides were synthesized with an Applied Biosystems model 380A solid phase synthesizer by the Protein Chemistry Facility of Washington University, St. Louis, MO. In most cases the 5' dimethoxy-trityl oligonucleotides were synthesized, annealed, and 20 fmol were ligated to the 13,356- and 1302-base pair fragments (8 fmol each) to generate plasmids (pSFFV-neo-M6P/IGF-II receptor), which were transfected into COS7 cells expressing M6P/IGF-II receptors. The strategy allowed the generation of receptors containing mutations beyond amino acid 26 of the cytoplasmic domain (the 163-amino acid cytoplasmic domain begins with lysine 2337). The remaining receptor sequence is wild-type, except for an alanine that has been substituted for tyrosine 24 of the cytoplasmic domain. Two additional plasmids were constructed to allow mutagenesis in other regions of the cytoplasmic or transmembrane domains by an analogous strategy. In pBC 223, an SfiI site (GGGCCCCAAAAGTCGGGA) is inserted at nucleotide 7065 which allows mutagenesis beyond amino acid 22 of the cytoplasmic domain. In plasmid pBC 224, an SfiI site (GGGGGAATAAACATTCATG) is inserted at nucleotide 6981 which allows mutagenesis beyond amino acid 7 of the transmembrane domain. The 13-kilobase SfiI fragment from each of these plasmids was ligated with the 1302-bp fragment from pBC 222 and appropriate oligonucleotides to generate plasmids encoding mutant receptors.

All sequences derived from synthetic oligonucleotides were confirmed by sequencing. Dideoxy sequencing (21) was performed with [35S]ATP and T7 DNA polymerase utilizing the 7-deaza-dGTP protocol for double-stranded plasmid templates as recommended by the manufacturer (Pharmacia LKB Biotechnology, Inc.).

**Purification and Iodination of 0-Glucuronidase**—The cell line 13.2.1 (a gift from Dr. W. Sly, St. Louis University) which overexpresses and secretes human β-glucuronidase (parental cell line mouse L reincorporation) was maintained in growth medium consisting of Dulbecco's modified Eagle medium supplemented with 0.9 mM sodium pyruvate, 26 mM sodium bicarbonate, 1.2 mM glutamine, 5% dialyzed fetal bovine serum, 3.2 µM methotrexate, 100 units/ml penicillin G, and 100 µg/ml streptomycin sulfate at 37°C in 5% CO2 (20). The cells were transfected with 40 µg of pSFFV-neo XbaI-linearized mutant pSFFV-neo M6P/IGF-II receptor and 30 µl of lipofectin as recommended by the manufacturer. After 2-3 weeks of selection with 500 µg/ml G418, cell colonies were isolated, expanded, and screened for M6P/IGF-II receptor expression as described previously (7).

**Fig. 1. Cassette mutagenesis strategy.** Plasmids encoding mutant receptors were constructed by ligation of mutant-specific oligonucleotides, a fragment of ~13,300 bp from pBC 222, 223, or 224 and a fragment of 1302 bp from pBC 222. The structures of the 3' termini of the ~13,300-bp fragments following SfiI digestion are illustrated. The 3' termini of these fragments are all the same and contain a 5' overhang with the sequence GGA which is complementary to the CCT overhang found on the 1302-bp fragment. The 5' terminus of the 1302-bp fragment provides the stop codon. For example, the mutant oligonucleotides used to prepare construct 344 are illustrated. Two stop codons are used to decrease the possibility of read-through.

15,094 base pairs in length and lacked unique restriction sites, we inserted sites for the restriction enzyme SfiI. The recognition sequence of SfiI is GGGCCNNNN/NNGGC, defining specificity for 8 nucleotides while allowing a half restriction site of 4 nucleotides to be placed without disrupting the coding sequence. The overhang sequence from different SfiI sites may be utilized, multiple SfiI sites may be utilized in a single ligation reaction while retaining the advantages of directional cloning.

A plasmid (pBC 222) was constructed by insertion of an oligonucleotide encoding the sequence GACIGGACGCCACAGCTGGTCGCAA-GTAGCCCTACTGCGTAAGG between the SfiI site at nucleotide 7053 and a Not1 site at nucleotide 7652 in the expression plasmid by the strategy previously described (7). The oligonucleotide retains the M6P/IGF-II receptor coding sequence up to amino acid 26 of the cytoplasmic domain, where the first SfiI site is located. The second SfiI site provides the stop codon and adjoins the non-coding sequence in the M6P/IGF-II receptor expression plasmid. The two SfiI sites have distinct overhang sequences to permit mutant-specific oligonucleotides to be inserted by directional cloning. The vector also contains a third SfiI site with a CCT overhang sequence. SfiI digestion of pBC 222 generates three fragments of ~13356, 1302, and 19 base pairs. The two larger fragments were isolated by agarose gel electrophoresis (19) and binding and elution from Silicic (20). Mutant-specific oligonucleotides were synthesized, annealed, and 20 fmol were ligated to the 13,366- and 1302-base pair fragments (8 fmol each) to generate plasmids (pSFFV-neo-M6P/IGF-II receptor), which were transfected into COS7 cells expressing M6P/IGF-II receptors. The strategy allowed the generation of receptors containing mutations beyond amino acid 26 of the cytoplasmic domain (the 163-amino acid cytoplasmic domain begins with lysine 2337). The remaining receptor sequence is wild-type, except for an alanine that has been substituted for tyrosine 24 of the cytoplasmic domain. Two additional plasmids were constructed to allow mutagenesis in other regions of the cytoplasmic or transmembrane domains by an analogous strategy. In pBC 223, an SfiI site (GGGGCAAACTCGGCC) is inserted at nucleotide 7065 which allows mutagenesis beyond amino acid 22 of the cytoplasmic domain. In plasmid pBC 224, an SfiI site (GGGGGAATAAACATTCATG) is inserted at nucleotide 6981 which allows mutagenesis beyond amino acid 7 of the transmembrane domain. The 13-kilobase SfiI fragment from each of these plasmids was ligated with the 1302-bp fragment from pBC 222 and appropriate oligonucleotides to generate plasmids encoding mutant receptors.

All sequences derived from synthetic oligonucleotides were confirmed by sequencing. Dideoxy sequencing (21) was performed with [35S]ATP and T7 DNA polymerase utilizing the 7-deaza-dGTP protocol for double-stranded plasmid templates as recommended by the manufacturer (Pharmacia LKB Biotechnology, Inc.).
β-Glucuronidase (30 μg) was iodinated with 1 μCi of Na\(^{125}\)I using soluble lactoperoxidase as described (22). The iodination reaction was gel filtered on a column (1.0 x 110 cm) of G-150 Sephadex equilibrated with 0.05 M NaPO\(_4\), pH 7.6, 0.15 M NaCl, 1 mg/ml bovine serum albumin. The column was eluted at a flow rate of 0.2 ml/min, and the material in the first peak (V/\(V_0\) = 0.40) was pooled, avoiding the leading edge of aggregated material. The specific radioactivity of the pooled β-glucuronidase was 8–16 μCi/μg assuming complete recovery.

**Short Internalization Assay**—Cells were seeded into 24-mm wells of a 12-well tissue culture plate and grown in MEM-a, 10% fetal calf serum for 36–48 h until confluent. The cells were rinsed 2 times with the same medium, and then 0.5 ml of MEM-a/10% fetal calf serum containing approximately 0.1 μg (between 1 and 2 x 10\(^{10}\) cpm) of \(^{125}\)I-β-glucuronidase was added. After a 30-min incubation on ice, the cells were rapidly washed 3 times with ice-cold PBS, 1% bovine serum albumin and 3 times with ice-cold PBS. The plate was then quickly transferred to a 37 °C water bath, where it was floated without its lid. A 0.5-ml portion of a mixture of 0.035% trypsin, 0.013% EDTA, 20 mM Man-6-P in 15 mM citrate-phosphate buffer, pH 5.5, was immediately added to each of two wells that were used for the measurement of the total surface binding of ligand (see below), while the other wells received an addition of 0.5 ml of 37 °C MEM-a. After an incubation period that usually ranged from 0.25 to 3 min, the medium was collected and 0.5 ml of the pH 5 trypsin-Man-6-P mixture was added to each well. At the end of an additional 5-min incubation at 37 °C, 0.8 ml of MEM-a/10% fetal calf serum was added to each well. The cells were then harvested in a 1.5-ml microcentrifuge tube and sedimented in an Eppendorf centrifuge 3200 for 1 min. After aspirating the supernatant with a pulled-out Pasteur pipette, the radioactivity in the cell pellet was measured in a \(\gamma\)-counter.

The initial rate of internalization was measured by calculating the maximal slope of the uptake curve following the short lag of 15–30 s. In general, two determinations were made with two different clones of cells expressing each mutant receptor. The "internalization ratio" was calculated by taking the ratio of the initial rate of internalization of the mutant construct relative to the initial rate of internalization of construct 344 (Tyr\(^{25}\) → Ala\(^{25}\), STOP\(^{25}\) in the same experiment.

**Long Internalization Assay**—The long internalization assays were performed by a slight modification of the previously described procedure (7). Cells were seeded into the wells of a 12-well tissue culture plate in triplicate and maintained for 36–48 h until confluent. To initiate uptake, the cells were rinsed 3 times with medium (MEM-a, 10% fetal calf serum), and then 1 ml of medium containing 1–2 x 10\(^{10}\) cpm of \(^{125}\)I-β-glucuronidase was added to each well. The cells were incubated at 37 °C for 4 h in a 5% CO\(_2\) atmosphere. The medium was then collected, and aliquots were precipitated with 10% trichloroacetic acid. The cells were then washed 5 times with cold PBS-bovine serum albumin and the surface-bound ligand was eluted with two washes of cold 12 mM citrate-phosphate buffer, pH 5.0, 0.16 M NaCl. The cell pellets were then solubilized in two 1-ml aliquots of 0.1 N NaOH. The radioactivity in the acid-eluted fraction, the NaOH-solubilized cells, and the trichloroacetic acid-soluble media were measured in a \(\gamma\)-counter to determine surface-bound, internalized, and degraded \(^{125}\)I-β-glucuronidase, respectively. For each measurement, the values obtained with wells containing L(rec-) cells were extremely slowly (Fig. 2B). These data also show that a significant proportion of the pre-bound \(^{125}\)I-glucuronidase is released into the media during the 3-min warming period. About one-third of the pre-bound ligand is released from the wild-type receptor, whereas more than 50% of the ligand is released from the tail-less receptor. This observation explains why the maximal internalization of ligand by the wild-type receptor is 50–60% of the surface-bound material.

The release of the surface-bound \(^{125}\)I-glucuronidase into the media is a temperature-dependent process. At 37 °C 80% of the pre-bound ligand is released from the tail-less receptor after 4 min, whereas the cells retain 85% of the ligand when kept on ice. At 21 °C the ligand is released at an intermediate rate.

**Analysis of a Truncated Receptor Containing a Tyr\(^{25}\) → Ala\(^{25}\) Mutation**—Since the goal of this study was to analyze the 4-amino acid internalization motif TYSK\(^{25}\) in detail, we designed a mutant receptor in which Tyr\(^{25}\) was substituted with an Ala residue and the cytoplasmic tail was truncated from 163 amino acids to 29 amino acids (construct 344: Tyr\(^{25}\) → Ala\(^{25}\), STOP\(^{25}\)). By using this mutant receptor as the reference construct for subsequent mutagenesis, we hoped to avoid secondary effects of mutations on other regions of the cyto-
plasmic tail. When this construct was analyzed in the internalization assay, it was found to be rapidly endocytosed with an initial rate of uptake similar to that observed with the wild-type receptor (Fig. 3). The only difference between the two receptors was that the mutant receptor internalized a smaller fraction of the pre-bound ligand (45% versus 62% for the wild-type receptor). This suggests that the off-rate for the bound ligand is somewhat greater for the mutant receptor than the wild-type receptor. Nevertheless, the rapid internalization of this mutant receptor is consistent with the critical element of the internalization signal being the YSKV sequence.

Since construct 344 also contains lysine 26, which is part of the YKYSKV 6-residue signal, we prepared a construct in which this residue was changed to alanine (construct 360; Tyrz4Lys26 → Ala24, 29; STOP39). As shown in Fig. 4, this mutant receptor was internalized at essentially the same rate as construct 344, confirming that the critical element of the internalization signal is the YSKV sequence.

Analysis of Mutant Receptors with Val29 Substitutions—
Previously we reported that a Val29 → Ala29 substitution in the cytoplasmic tail of the M6P/IGF-II receptor results in a drastic reduction in the rate of receptor internalization (7). In addition, when we compared the cytoplasmic tail sequences of a number of membrane proteins known to undergo rapid endocytosis, using the essential tyrosine as the basis of the alignment, we noted six other instances where a bulky hydrophobic residue (either a valine, phenylalanine, or isoleucine) was located 3 amino acids to the carboxyl side of the tyrosine. Based on these findings we reasoned that the presence of a bulky hydrophobic residue in this position may be an essential component of the internalization signal. To test this prediction, a series of constructs were generated in which valine 29 was replaced by residues with or without this characteristic. These cDNAs were transfected into receptor-negative mouse L cells and stable clones expressing the mutant receptors were selected and analyzed for their ability to bind and internalize β-glucuronidase using the rapid internalization assay. The results of typical experiments are shown in Fig. 5, and a summary of all the determinations is given in Table I (section A). These values are expressed as the ratio of the initial rate of internalization of the mutant receptor compared with the initial rate obtained with the Y29SKV-containing receptor in the same experiment. When valine 29 was substituted with either phenylalanine, leucine, isoleucine, or methionine, the

**Fig. 3.** Internalization of a truncated receptor containing the Tyr29 → Ala24 mutation. Cells expressing the wild-type M6P/IGF-II receptor (Cc2, O) or a truncated receptor with a Tyr29 → Ala24 substitution (construct 344; Tyr29 → Ala29; STOP39, •) were tested for their ability to internalize 125I-β-glucuronidase upon incubation at 37 °C. The internalized radioactivity is plotted as the percentage of the total surface-bound radioactivity prior to incubation at 37 °C. The data presented are the average ± standard deviations of 12 (Cc2) and 5 (construct 344) separate determinations.

**Fig. 4.** Internalization of a truncated receptor containing a Tyrz4Lys26 → Ala24,29 mutation. Cells expressing construct 344 and construct 360 (Tyrz4Lys26 → Ala24,29; STOP39) were assayed for their ability to internalize 125I-β-glucuronidase upon incubation at 37 °C. The values of internalized 125I-β-glucuronidase are expressed as the percentage of the total surface-bound radioactivity prior to incubation at 37 °C (construct 344 (×): 56,893 cpm; construct 360 (O): 38,605 cpm).

**Fig. 5.** Effect of Val29 substitutions on M6P/IGF-II receptor internalization. The values of internalized 125I-β-glucuronidase are expressed as the percentage of the total surface-bound radioactivity prior to incubation at 37 °C. **Panel A:** AKYSKV (construct 344; 41,503 cpm), AKYSKL (construct 342; 15,088 cpm), AKYSKQ (construct 384; 4699 cpm), AKYSKR (construct 395; 3186 cpm), AKYSKG (construct 368; 3063 cpm). **Panel B:** AKYSKV (construct 344; 34,265 cpm), AKYSKF (construct 341; 4725 cpm), AKYSKM (construct 380; 3704 cpm), AKYSKI (construct 343; 3380 cpm), AKYSKY (construct 419; 31,423 cpm).
resultant mutant receptors were internalized as rapidly as, or even faster than, the receptor with valine in position 29 (Fig. 5, A and B). The average internalization ratios were 2.22, 1.76, 1.36, and 1.32 for the phenylalanine, leucine, isoleucine, and methionine-containing mutants, respectively, compared to 1.00 for valine (Table I). By contrast, mutant receptors containing a glutamine, glycine, arginine (Fig. 5A), or a tyrosine (Fig. 5B) at position 29 were internalized very slowly (internalization ratios of 0.05, 0.02, 0.03, and 0.12, respectively). These results demonstrate that bulky, hydrophobic residues (Phe, Leu, Ile, and Met) can substitute for Val29 without any impairment of the signaling function of the Y29SKV29 sequence, whereas small (Gly) or polar (Tyr, Gln, Arg) residues cannot.

Analysis of Mutant Receptors with Tyr29 Substitutions—In our previous study we showed that a Tyr29→Ala29 substitution results in a striking decrease in the rate of receptor internalization (7). To further examine the amino acid requirement at this position of the signal, we replaced tyrosine 26 with 6 different amino acids and determined the effect on the rate of receptor internalization. The results are shown in Fig. 6 and summarized in Table I (section B). Replacement of tyrosine 26 with either a leucine, isoleucine, valine or asparagine gave rise to receptors that were internalized at the basal rate (internalization ratios of 0.00 to 0.01). Mutant receptors with either a phenylalanine or a tryptophan residue at position 26 were internalized more rapidly than the negative control, but much slower than construct 344 (internalization ratios of 0.06 and 0.13, respectively). These results indicate a specific requirement for tyrosine in this position although other aromatic residues can substitute to some extent.

The slow rate of internalization of the mutant receptor with the Tyr26→Phe26 mutation was surprising since our previous experiments had shown that a mutant receptor with a Tyr24,25→Phe24,25 substitution in the context of a full-length cyto-

### TABLE I

**Summary of internalization ratios of cytoplasmic tail mutants**

| Construct no. | Cytoplasmic tail structure | Internalization ratio* |
|---------------|-----------------------------|------------------------|
| A             |                             |                        |
| 341           | KKERR EMVMS RLTC CRRSA NVSAK YSKV | 1.86; 2.59          |
| 342           | KKERR EMVMS RLTC CRRSA NVSAK YSKL | 1.74; 1.79          |
| 343           | KKERR EMVMS RLTC CRRSA NVSAK YSKI | 1.00; 1.78          |
| 380           | KKERR EMVMS RLTC CRRSA NVSAK YSQM | 1.14; 1.51          |
| 419           | KKERR EMVMS RLTC CRRSA NVSAK YSKY | 0.00; 0.18          |
| 395           | KKERR EMVMS RLTC CRRSA NVSAK YSKQ | 0.02; 0.08          |
| 385           | KKERR EMVMS RLTC CRRSA NVSAK YSKR | 0.03; 0.03          |
| 368           | KKERR EMVMS RLTC CRRSA NVSAK YSKG | 0.01; 0.03          |
| B             |                             |                        |
| 395           | KKERR EMVMS RLTC CRRSA NVSAK VSKV | 0.11; 0.15          |
| 364           | KKERR EMVMS RLTC CRRSA NVFSK FSKV | 0.05; 0.08          |
| 333           | KKERR EMVMS RLTC CRRSA NVSKV NSKV | 0.04; 0.08          |
| 367           | KKERR EMVMS RLTC CRRSA NVSKV NSKV | 0.01                |
| 354           | KKERR EMVMS RLTC CRRSA NVSKV LSKV | 0.00                |
| 355           | KKERR EMVMS RLTC CRRSA NVSKV ISKV | 0.00                |
| 356           | KKERR EMVMS RLTC CRRSA NVSKV TPKV | 0.00                |
| C             |                             |                        |
| 405           | KKERR EMVMS RLTC CRRSA NVSAK YIKV | 1.55                |
| 407           | KKERR EMVMS RLTC CRRSA NVSAK YIKV | 0.85; 1.08          |
| 418           | KKERR EMVMS RLTC CRRSA NVSAK YARV | 0.82                |
| 406           | KKERR EMVMS RLTC CRRSA NVSAK YMSV | 0.71; 0.91          |
| 350           | KKERR EMVMS RLTC CRRSA NVSAK VYKV | 0.61; 0.76          |
| 403           | KKERR EMVMS RLTC CRRSA NVSAK YFKV | 0.32; 0.63          |
| 422           | KKERR EMVMS RLTC CRRSA NVSAK YGVF | 0.30; 0.50          |
| 345           | KKERR EMVMS RLTC CRRSA NVSAK YSVK | 0.33; 0.37          |
| 352           | KKERR EMVMS RLTC CRRSA NVSAK YGKV | 0.01; 0.01          |
| 354           | KKERR EMVMS RLTC CRRSA NVSAK YPKV | 0.01; 0.01          |
| D             |                             |                        |
| 346           | KKERR EMVMS RLTC CRRSA NVSAK YKV | 0.00                |
| 347           | KKERR EMVMS RLTC CRRSA NVSAK YSAK | 0.00                |
| 348           | KKERR EMVMS RLTC CRRSA NVSAK YSAK | 0.00                |
| 349           | KKERR EMVMS RLTC CRRSA NVSAK YSAK | 0.00                |
| E             |                             |                        |
| 377           | KKERR EMVMS RLTC CRRSA NVSAK YTRF | 1.34; 1.98          |
| 378           | KKERR EMVMS RLTC CRRSA NVSAK YSAF | 1.57; 1.70          |
| 356           | KKERR EMVMS RLTC CRRSA NVSAK YQTL | 1.07; 1.30          |
| 376           | KKERR EMVMS RLTC CRRSA NVSAK YQOF | 1.07; 1.10          |
| 375           | KKERR EMVMS RLTC CRRSA NVSAK YRHV | 0.80; 0.95          |
| 374           | KKERR EMVMS RLTC CRRSA NVSAK YRGV | 0.14; 0.37          |

* The value obtained for construct 344 in each experiment was set to 1.0, and all other values are expressed relative to this. Each construct was assayed in one or two independent experiments as noted.
plasmic tail was rapidly internalized. To test the possibility that a phenylalanine could substitute for tyrosine 26 provided that a phenylalanine was also located at position 24, we constructed a mutant receptor containing a F24KFSKV29 sequence in the truncated 29-amino acid cytoplasmic tail (construct 345). This mutant receptor had an internalization ratio of 0.07, similar to the value of 0.06 obtained with the Tyr26→Phe29 mutant (Fig. 6B, Table I). This demonstrates that phenylalanines will function in place of tyrosines at positions 24 and 26 only in the context of a full-length cytoplasmic tail.

Analysis of Mutant Receptors with Substitutions Involving Ser27 and Lys29 — The consequence of replacing serine 27 with a variety of amino acids is shown in Fig. 7 and summarized in Table I (section C). Serine 27 can be substituted with isoleucine, alanine, phenylalanine, valine, threonine, or methionine with little or no effect on the rate of internalization of the mutant receptor. The result with the Ser27→Ala27 mutant confirms our previous finding. However, not all residues are allowed in this position since the replacement of serine 27 with a glycine or a proline results in a drastic impairment in endocytosis (Fig. 7B). Interestingly, the effect of a glycine at position 27 varies depending on the residues at positions 28 and 29. Thus a mutant receptor with the sequence Y26GVF29 is internalized relatively well, with an average internalization ratio of 0.40 (Fig. 7B). These data demonstrate that the residue at position 27 can modulate or even abolish the functioning of the internalization signal.

We also tested the effect of reversing the Ser27Lys29 sequence in the Y26SKV29 signal. As shown in Fig. 7B, the resulting mutant receptor (construct 345) containing a Y26KSV29 sequence was internalized about one-third as well as the control receptor (internalization ratio of 0.35).

Importance of the Spacing between Residues 26 and 29 of the Cytoplasmic Tail — To determine if the spacing between tyrosine 26 and valine 29 is critical for the functioning of the internalization signal, four mutant receptors were constructed in which the spacing between these two residues was altered. In construct 346, serine 27 was removed to generate the sequence Y26KTV29, whereas in constructs 347–349, 2, 2, or 3 alanines were introduced between serine 27 and the lysine initially present at position 29. As summarized in Table I (section D), none of the mutant receptors were internalized over the basal rate. Clearly the correct spacing between tyrosine 26 and valine 29 is critical for the functioning of the internalization signal.

The YSKV Sequence Remains Functional When Moved to a Different Location in the Cytoplasmic Tail — Construct 420 was prepared to determine whether the YSKV internalization signal would still function when placed closer to the transmembrane segment and adjacent to a different set of amino acids. This construct contains an 11-amino acid cytoplasmic tail with the sequence K′KERRAKYSKV31. The first 5 amino acids.
acids represent the initial residues of the M6P/IGF-II receptor cytoplasmic tail. Surprisingly we found that the clones expressing this mutant receptor bound the \(^{131}I\)-labeled glucuronidase very poorly at 4 °C, making it impossible to perform the standard rapid internalization assay.\(^2\) The basis for the poor binding at 4 °C is not clear. However, the mutant receptor did bind the ligand quite well at 37 °C, and therefore it was possible to evaluate its rate of endocytosis using the long internalization assay described previously (7). Table II shows that construct 420 was internalized about 40–60% as well as the wild-type receptor with a full-length cytoplasmic tail and 4–5 times better than a receptor (Dd) with a 7-amino acid cytoplasmic tail. These data demonstrate that the YSKV sequence remains functional in endocytosis, albeit at a somewhat reduced rate, when transplanted to a different region of the cytoplasmic tail.

**Table II**

| Cytoplasmic tail sequence | Cell line | Surface-bound | Intracellular | Degraded | Internalization index* |
|--------------------------|-----------|---------------|---------------|----------|-----------------------|
| **Expt. 1**              |           |               |               |          |                       |
| Full-length              | Cc        | 12,983        | 134,298       | 33,350   | 18.9                  |
| Wild-type               | Dd        | 28,750        | 55,162        | 4,775    | 2.1                   |
| –KKERREM\(^a\) (Val\(^b\) – STOP\(^c\)) | 420  | 2,642         | 20,776        | 7.9      |
| –KKERAKYSKV             | 420  | 6,884         | 55,353        | 9.5      |
| **Expt. 2**              |           |               |               |          |                       |
| Full-length              | Cc        | 18,762        | 397,415       | 76,231   | 25.2                  |
| Wild-type               | Dd        | 75,646        | 134,416       | 19,344   | 2.0                   |
| –KKERREM\(^a\) (Val\(^b\) – STOP\(^c\)) | 420  | 6,884         | 242,385       | 14.0     |
| –KKERAKYSKV             | 420  | 6,831         | 38,625        | 9.0      |

* The internalization index is calculated by dividing the sum of the intracellular and degraded radioactivity by the surface-bound radioactivity (23).

**Mannose 6-P Receptor Internalization Signal**

In our previous study we presented evidence that the signal for the rapid internalization of the M6P/IGF-II receptor is contained within the YSKKYKV sequence of the 163-amino acid cytoplasmic tail (7). The most important elements of this signal were shown to be Tyr\(^5\) and Val\(^7\) with Lys\(^8\) and Tyr\(^24\) having lesser roles. Since the YSKKYKV sequence was almost as potent as the YKYSKV sequence in promoting rapid endocytosis, this internalization signal could be considered as either a 6- or a 4-amino acid motif. When we compared the YKYSKV sequence with the sequences neighboring the critical tyrosines in the cytoplasmic tails of eight other proteins known to undergo rapid internalization, we noted six instances where the tyrosine was separated from a valine, phenylalanine, or isoleucine by 2 intervening amino acids. Based on this information, we suggested that a common internalization motif may be a sequence of 4 amino acids with an aromatic residue, especially a tyrosine, in the first position separated from a bulky hydrophobic residue in the fourth position.

**DISCUSSION**

The current experiments provide further evidence to support this proposal. Using a construct with a 29-amino acid cytoplasmic tail that terminates with the sequence AKYSKV, we have systematically tested the effect of numerous amino acid substitutions at the first, second, and fourth position of the YSKV internalization sequence. Our results clearly show a requirement for an aromatic residue in the first position of the signal, particularly a tyrosine. Tryptophan and phenylalanine were about 10% as effective as tyrosine in promoting rapid endocytosis while valine, isoleucine, leucine, and asparagine were totally inactive. The finding that a Tyr\(^5\) \(\rightarrow\) Phe\(^5\) substitution resulted in a marked decrease in the rate of receptor internalization was unexpected since a Tyr\(^24\) \(\rightarrow\) Phe\(^24\) replacement in the context of a full-length cytoplasmic tail did not impair receptor endocytosis (7). The addition of a second phenylalanine at position 24 in the mutant receptor with a truncated cytoplasmic tail (construct 364) did not enhance the ability of the receptor to be rapidly endocytosed. The reason for this discrepancy is unclear, but we presume that the presence of the full-length cytoplasmic tail allows the phenylalanines to function better than they do in the context of a truncated cytoplasmic tail. One possibility is that elements in the full-length cytoplasmic tail enable the FPKSKV sequence to achieve a structure that is required for interaction with the adaptin proteins of the coated pits. Alternatively, there could be additional elements in the full-length cytoplasmic tail that interact directly with the adaptin proteins of the coated pits.
proteins and thereby complement the weak signal generated by the FKFSKV sequence. In this regard, it is of interest that phenylalanine will substitute for tyrosine in some internalization signals (12, 25, 26), but not in others (16, 24).

The analysis of the substitutions at the fourth position of the signal shows that valine 29 can be replaced with leucine, isoleucine, methionine, or phenylalanine with no impairment in receptor internalization. However, when the phenylalanine was replaced with an alanine or a glycine, rapid endocytosis was lost.

One function of the residues at positions 2 and 3 of the internalization signal is to provide the correct spacing between the tyrosine in position one and the hydrophobic residue in position 4. This is most clearly shown by construct 346, which has a deletion of serine 25 to produce a 3-amino acid signal, and construct 347, which contains an alanine inserted between the serine and lysine to produce a 5-amino acid signal. Neither of these mutant receptors was endocytosed above the base-line value, showing an absolute requirement for the correct spacing of the 4-amino acid signal.

While many amino acid substitutions were well tolerated in position 2 of the YSKV signal, proline and glycine were not. In addition, it is apparent that the nature of the residues in positions 2 and 3 may modulate the activity of the signal. For example, reversing the serine and lysine residues to produce a signal with a YKSV sequence results in a 65% decrease in the rate of internalization. An even more striking finding was made with mutant receptors containing glycine substitutions in position 2. Construct 352 with a YGVF sequence was internalized at the baseline value, whereas construct 422 with the sequence YGKF was internalized reasonably well (internalization ratio of 0.35). Interestingly the YGKF sequence is present in the cytoplasmic tail of the P-selectin molecule which is known to be translocated to the cell surface of platelets and endothelial cells upon activation and then cleared from the plasma membrane over a 10–20-min period (27). In these examples, the consequence of having a glycine in position 2 of the signal is dependent, in part, on the nature of the residues in positions 3 and 4. Since the same residue at a particular position of the signal may have a different effect on the activity depending on the nature of the other residues, it may be difficult to predict whether a putative internalization signal will be functional or not. However, this will only be a problem when certain amino acids, such as glycine and proline, are present in the signal.

The finding that the YSKV sequence can be moved to position 8–11 of the cytoplasmic tail and still mediate rapidly endocytosis shows that this 4-amino acid motif can function when placed in a different context. This result is consistent with our previous alanine scanning data which showed that the 6 residues preceding the YSKV signal in the cytoplasmic tail were not essential for rapid endocytosis (7).

One prediction from these studies is that an putative 4-amino acid internalization sequences of other recycling receptors and proteins should function when transplanted onto the cytoplasmic tail of the M6P/IGF-II receptor. This proved to be the case with the tyrosine-containing sequences from six different proteins, including the transferrin receptor, a type II membrane protein. This latter finding shows that the internalization signals work regardless of whether they are in the context of a type I or a type II membrane protein. Collawn et al. (18) have performed the converse experiment (18). These investigators substituted the 6- and the 4-residue internalization signal of the M6P/IGF-II receptor for the 4-residue signal of the transferrin receptor and found that both sequences promote the rapid internalization of the transferrin receptor. In addition, they found that the putative tetrapeptide internalization sequences of the poly(Ig) receptor, YSAF (15), and the human asialoglycoprotein receptor, YQDL (28), as well as the hexapeptide signal of the LDL receptor (FDNPVY) (8) are functional when transplanted into the transferrin receptor. Taken together with our results, these findings are consistent with the internalization signals being interchangeable motifs that function in the context of a type II membrane protein.
I or type II membrane protein. It should be noted that both of these studies involved transplanting putative internalization signals onto receptors that normally undergo constitutive recycling. Thus it is possible that these receptors have other properties that allow them to be rapidly endocytosed when they contain the internalization signals. For this reason it will be of interest to determine if these signals function in the same way when transplanted onto the cytoplasmic tails of membrane proteins that do not normally undergo rapid endocytosis.

Several recent studies have suggested that the internalization sequences have a propensity to form tight turn structures. Collawn et al. (10, 18) have presented indirect evidence that the 6-residue internalization signals of the M6P/IGF-II and the LDL receptors as well as the 4-residue signal of the transferrin receptor adopt tight turn structures based on their analysis of similar sequences in known protein crystal structures. In the case of the M6P/IGF-II receptor, they identified four structural analogs of the YKYSKV sequence. The four carboxyl-terminal residues of these analogs were noted to be in tight turns and the four functionally important side chains at positions −2, 1, 3, and 4 were simultaneously accessible from one side of the turn. (The numbering refers to the positions relative to the turn region.) In addition, the two aromatic rings in positions −2 and 1 were oriented in a roughly parallel manner, providing an indication of how one of the aromatic rings could compensate for the loss of the other. Collawn et al. (18) suggested that the similarity in three-dimensional placement of the critical residues in the 6- and 4-residue signals would allow both types of signals to interact with the same recognition structure in coated pits. Kitstakis et al. (17) have postulated that the tyrosine recognition signals may form a small surface loop structure, but this structure differs from the one proposed by Collawn et al. in terms of the positioning of the tyrosine in the loop. Bansal and Gierasch (29) have obtained more direct evidence that the NPVY sequence of the LDL receptor forms a β-turn structure by performing NMR analysis of nonapeptides containing this sequence. They showed that peptides containing the NPVY sequence assume a reverse-turn conformation with the Asn in position 1 and the Tyr in position 4 of the turn. Substitution of either the Asn, Pro, or Tyr with residues known to be inactive in endocytosis resulted in a disruption of the turn conformation. Eberle et al. (30) used a similar approach to obtain evidence that the PPGY sequence of the acid phosphatase cytoplasmic tail forms a type I β turn with the Tyr in position 4 of the turn.

All of these studies indicate that the critical tyrosine of the internalization signal is presented to the adaptin proteins in the context of a tight turn motif. However, the LDL receptor internalization motif differs from the signal of the M6P/IGF-II receptor and the transferrin receptor in that its essential tyrosine is in the carboxyl-terminal position of the signal, whereas the critical tyrosine of the other two receptors is in the amino-terminal position of the signal. In fact, our data show that a mutant receptor with a Val9 → Tyr9 substitution (construct 419) is internalized very poorly.

The finding that the PPGY motif of the acid phosphatase cytoplasmic tail achieves a tight turn conformation must be reconciled with the results of our transplantation experiments showing that the YRLHV sequence of this cytoplasmic domain functions to mediate rapid endocytosis in the context of the M6P/IGF-II receptor (construct 375). One possibility is that the internalization signal of acid phosphatase comprises a more extended structure, involving residues 5-11, with the RHV sequence providing additional interactions with the adaptin proteins over those achieved by the PPGY turn structure. If this is the case, then the YRLHV sequence may function well enough in the context of the M6P/IGF-II receptor to allow rapid endocytosis. The possibility that some internalization signals may have more extended structures has been suggested by several investigators (11, 12, 17, 30). Kitstakis et al. (17) have proposed a "tyrosine internalization signal" that extends from −6 to +2 relative to the essential tyrosine with polar or basic residues preferred at certain positions on both sides of the tyrosine. However, there are several exceptions to these rules, including the M6P/IGF-II receptor.

In summary, we have presented evidence that a tetrapeptide sequence with an aromatic residue, especially a tyrosine, in the first position and a large hydrophobic residue in the fourth position, functions as an effective signal for the rapid internalization of the M6P/IGF-II receptor and possibly many other receptors and membrane proteins as well. The relationship of this 4-residue signal to other internalization motifs needs further exploration, particularly at the structural level.

REFERENCES
1. Kornfeld, S., and Mellman, I. (1980) Annu. Rev. Cell Biol. 5, 483-525
2. Kornfeld, S. (1992) Annu. Rev. Biochem., in press
3. Willingham, M. C., Pasani, J. H., Sahagian, G. G., Jourdian, G. W., and Neufeld, E. F. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 6967-6971
4. Overbeek, R. (1985) EMBO J. 4, 2457-2460
5. Glickman, J. N., Combeert, E., and Overbeek, R. M. F. (1989) EMBO J. 8, 1041-1047
6. Obele, P., Fujimoto, K., Ye, R. D., Griffiths, G., and Kornfeld, S. (1989) Cell 57, 787-796
7. Canafax, W. M., Johnson, K. F., Ye, R. L., Gregory, W., and Kornfeld, S. (1991) J. Biol. Chem. 266, 5862-5868
8. Chen, W.-J., Goldstein, J. L., and Brown, M. S. (1990) J. Biol. Chem. 265, 3116-3123
9. Johnson, K. F., Chen, W., and Kornfeld, S. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 10010-10014
10. Collawn, J. F., Stangel, M., Kuhn, L. A., Esseghou, V., Jing, S., Trowbridge, I. S., and Tainer, J. A. (1990) Cell 63, 1061-1072
11. McGraw, T. E., Pytowski, B., Arzt, J., and Ferrone, C. (1991) J. Cell Biol. 112, 853-861
12. Girod, M., Alvarez, E., Seth, A., Lin, I.-M., Lateur, D., and Davis, R. J. (1991) J. Biol. Chem. 266, 19006-19012
13. Howe, C. L., Granger, B. L., Hull, M., Green, S. A., Gabel, C. A., Helenius, A., and Mellman, I. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 5707-5711
14. Granger, B. L., Green, S. A., Gabel, C. A., Howe, C. L., Mellman, I., and Helenius, A. (1990) J. Biol. Chem. 265, 12006-12043
15. Britfeld, P. F., Casanova, J. E., McKinnon, W. C., and Mostov, K. E. (1990) J. Biol. Chem. 265, 13750-13757
16. Myers, C. R., Braun, M., Weber, B., Wendland, M., Schmidt, R., Pohlmann, R., Wahled, A., and von Figura, K. (1990) EMBO J. 9, 3497-3506
17. Kitstakis, N. Y., Thomas, D., and Roth, M. G. (1990) J. Biol. Chem. 111, 1393-1407
18. Collawn, J. F., Kuhn, L. A., Liu, L.-P., Tainer, J. A., and Trowbridge, I. S. (1991) EMBO J. 10, 3247-3253
19. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467
20. Gabel, C. A., Goldberg, D. E., and Kornfeld, S. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1779-1782
21. Hovekamp, B., Fujimoto, K., and Kornfeld, S. (1987) J. Biol. Chem. 262, 15674-15679
22. Tait, J. F., Weinman, S. A., and Bradshaw, R. A. (1961) J. Biol. Chem. 236, 11086-11092
23. Goldstein, J. L., Basu, S. K., and Brown, M. S. (1985) Methods Enzymol. 105, 241-260
24. Williams, M. A., and Fukuda, M. (1990) J. Biol. Chem. 111, 955-966
25. Davis, C. G., van Driel, I. R., Russell, D. W., Brown, M. S., and Goldstein, J. L. (1987) J. Biol. Chem. 262, 4075-4082
26. McGraw, T. E., and Maxfield, F. R. (1990) Cell 69, 1369-1377
27. Latt, H., Hamilton, K. K., Busalacchi, A., and Spiess, P. J. (1989) J. Biol. Chem. 264, 7776-7781
28. Fuhrer, C., Geffen, I., and Spiess, P. J. (1991) J. Biol. Chem. 1107, 431-437
29. Bansal, A., and Gierasch, L. M. (1991) Cell 67, 1195-1203
30. Eberle, W., Sander, C., Kiesewetter, M., Schmidt, B., von Figura, K., and Peters, C. (1991) Cell 67, 1203-1209