The EMBO Journal vol.13 no.7 pp.1696–1705, 1994

An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum

Marie-Paule Schutze[1], Per A.Peterson and Michael R.Jackson[2]

Department of Immunology, IMM8, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, USA
[1]Present address: Unité d'Immunité Cellulaire Antivirale, Institut Pasteur, 28 rue du Dr Roux, Paris 75015, France
[2]Corresponding author

Communicated by P.A.Peterson

Use of alternative initiator methionines in human invariant (Ii) chain mRNA results in the synthesis of two polypeptides, Iip33 and Iip31. After synthesis both isoforms are inserted into the endoplasmic reticulum (ER) as type II membrane proteins. Subsequently, Iip31 is transported out of the ER, guiding MHC class II to the endocytic pathway, whereas Iip33, which differs by only a 16 residue extension at the N-terminus, becomes an ER resident. Mutagenesis of this extension showed that multiple arginines close to the N-terminus were responsible for ER targeting. The minimal requirements of this targeting motif were found to be two arginines (RR) located at positions 2 and 3, 3 and 4 or 4 and 5 or split by a residue at positions 2 and 4 or 3 and 5. Transplanting an RR motif onto transferrin receptor demonstrated that this motif can target other type II membrane proteins to the ER. The characteristics of this RR motif are similar to the KK ER targeting motif for type I membrane proteins. Indeed, RR-tagged transferrin receptor partially localized to the intermediate compartment, suggesting that like the KK motif, the RR motif directs the retrieval of membrane proteins to the ER via a retrograde transport pathway.

Key words: endoplasmic reticulum/membrane sorting/protein targeting motif/protein trafficking

Introduction

After synthesis integral membrane proteins destined for the secretory pathway are inserted across the membrane of the endoplasmic reticulum (ER). Current data support the idea that once these nascent proteins have correctly folded and assembled they are rapidly transported out of the ER along the exocytic pathway (Pfeffer and Rothman, 1987). Although proteins are probably concentrated prior to exit from the ER, specific signals within a protein are not thought to be required for this process. In contrast, protein targeting signals, typically consisting of short amino acid sequence motifs, have been shown to be important in directing proteins to, and maintaining them in, specific organelles of the exocytic and endocytic pathways. Examples of this type of signal are known for the ER (Munro and Pelham, 1987; Nilsson *et al.*, 1989; Lotteau *et al.*, 1990), the *trans*-Golgi network (Wilcox *et al.*, 1992; Humphrey *et al.*, 1993), early endosomes (see Trowbridge, 1991) and late endosomes/lysosomes (Bäkke and Dobberstein, 1990; Lotteau *et al.*, 1990; Harter and Mellman; 1992; Johnson and Kornfeld, 1992; Letourneur and Klausner, 1992). In each of these cases the signals are composed of a short discrete amino acid motif located in the cytoplasmic domain of the protein. However, other domains of a protein may also contain targeting information as exemplified by the importance of the sequence of the membrane spanning domain for maintaining proteins in the Golgi (see Machamer, 1993).

We have previously demonstrated that type I membrane proteins can be efficiently targeted and accumulated in the ER by the introduction of two lysine residues close to the C-terminus (Jackson *et al.*, 1990). A strict positional requirement for the two lysines of this KK motif was found and the consensus for this ER sorting signal, two lysine residues located either three and four or three and five residues from the C-terminus, is a common feature of ER-resident membrane proteins (Jackson *et al.*, 1990). Marker proteins tagged with a KK motif are rapidly, post-translationally modified in the ER–Golgi intermediate compartment and partially co-localize to this organelle (Jackson *et al.*, 1993). Furthermore, KK-tagged proteins slowly receive post-translational modifications characteristic of exposure to Golgi enzymes and lectin staining localized these Golgi modified proteins to the ER. These data suggest that the KK motif does not fix or retain molecules in the ER. Instead, it directs the retrieval to the ER of proteins from post-ER organelles via a retrograde transport route in a manner analogous to retrieval of KDEL-tagged proteins (Munro and Pelham, 1987) from the Golgi to the ER by the KDEL receptor (Lewis and Pelham, 1992). In order to provide further insight into the process of sorting of membrane proteins to the ER we were interested in characterizing the signal responsible for ER localization of type II integral membrane proteins.

Human invariant chain (Ii) is a type II membrane glycoprotein (Kvist *et al.*, 1982). Two distinct forms of the Ii protein, Iip31 and Iip33, are generated by the use of alternative initiator codons (Strubin *et al.*, 1986) and consequently differ only in their N-terminal sequence (see Figure 1). Indirect immunofluorescence and pulse–chase analysis have demonstrated that Iip31 is transported rapidly out of the ER after synthesis whereas Iip33 is maintained in the ER (Lotteau *et al.*, 1990). These observations are consistent with the presence of an ER sorting signal in the N-terminal 16 amino acid extension of Iip33. In this paper we have characterized the ER targeting motif in Iip33. Using a series of deletion and replacement mutants we have shown that this motif consists of arginine residues located close to the N-terminus. We show that transplanting two arginine residues into the appropriate positions in the N-terminus of a type II membrane marker protein, which is ordinarily

transported to the cell surface, rendered the chimera ER-resident. The similarity of this arginine based ER targeting motif to the double-lysine motif which targets type I membrane proteins to the ER is discussed with reference to the mechanism by which targeting is achieved.

Results

Truncation of the cytoplasmic tail of Iip33 abrogates ER retention

In order to allow analysis the ER targeting motif present in Iip33 without complications from co-expression of Iip31, we have used as the parent molecule a form of Iip33 in which the second methionine has been mutated to isoleucine, an alteration which does not affect ER targeting (Lotteau *et al.*, 1990). Our initial characterization of the ER sorting information present in the N-terminal 16 residues of Iip33 (Figure 1) made use of a series of N-terminal deletion mutants of Iip33 starting after the initiator methionine (referred to as Δ1, Δ2, Δ4, Δ6, Δ9 and Δ12Iip33). The constructs were inserted into the expression vector pCMUIV (Nilsson *et al.*, 1989) and co-transfected into HeLa cells with DNA encoding CD8/E19, a chimeric human CD8 protein retained in the ER by a double-lysine motif (Jackson *et al.*, 1990). Cells were fixed 72 h after transfection, permeabilized and co-stained for immunofluorescence using K2, a rabbit antiserum against a synthetic peptide derived from a portion of the lumenal domain of Ii (Lotteau *et al.*, 1990), and OKT8, a mouse monoclonal antibody against human CD8. At relatively low levels of expression Iip33, Δ1Iip33 and Δ2Iip33 were localized to an extensive tubular network characteristic of the ER (Figure 2A). Double staining for Ii and CD8 showed essentially coincident patterns for the subcellular distribution of Iip33, Δ1Iip33, Δ2Iip33 and the ER-resident protein CD8/E19 (Figure 2A and B). Interestingly, in cells expressing very high levels of the three transfected Ii proteins, a striking array of tubular membrane filaments was observed in the cytoplasm (Figure 2C).

Deletion of four amino acids from the N-terminus of Iip33 dramatically affected the subcellular distribution of Ii. Δ4Iip33 primarily localized to numerous vesicular structures of variable sizes typically clustered in a perinuclear location (Figure 2D), a pattern similar to that observed for Iip31 expressed at high levels in transiently transfected HeLa cells (Lotteau *et al.*, 1990). The exact nature of these large

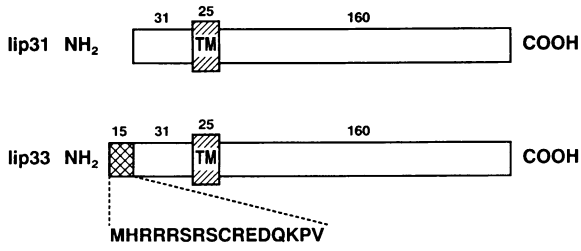

Fig. 1. Schematic representation of the relationship between Iip31 and Iip33. The numbers in the figure denote the lengths of the various Ii segments.

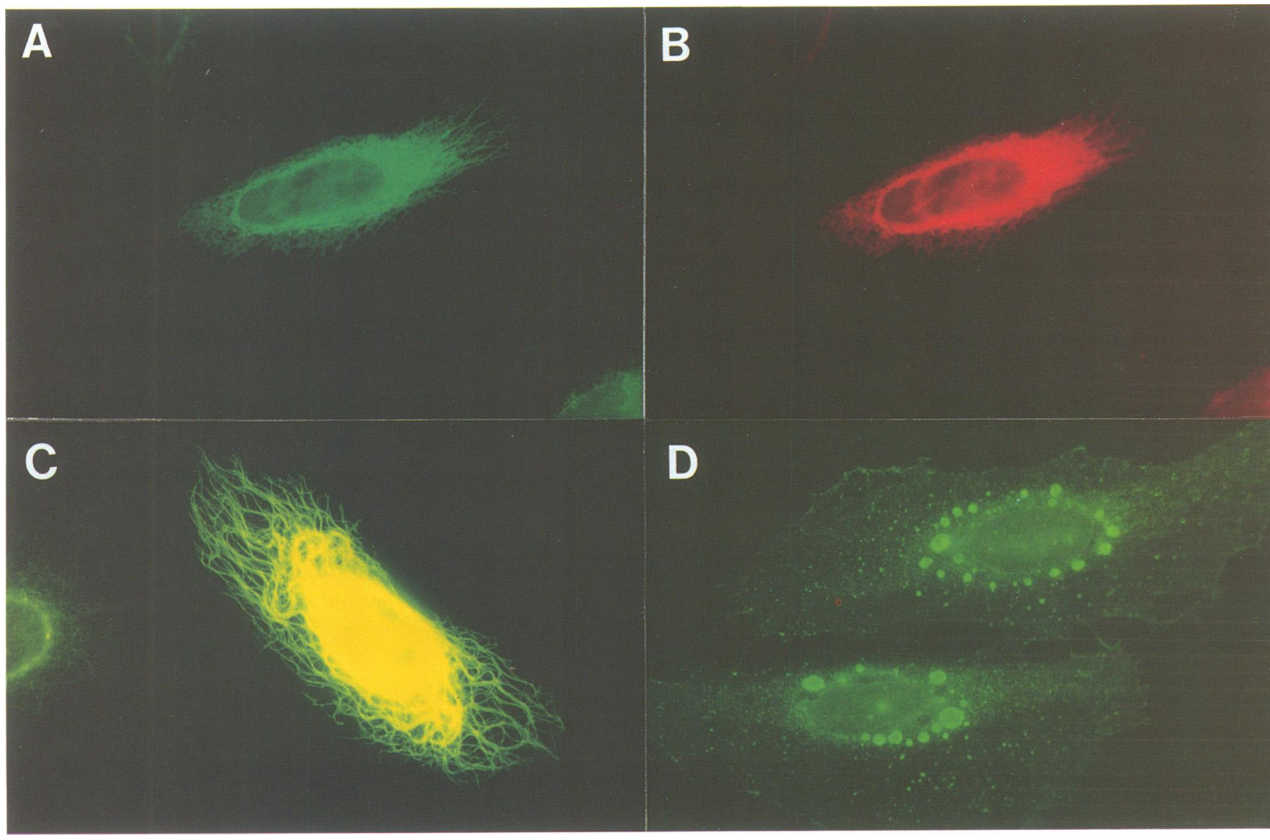

Fig. 2. Immunofluorescence microscopy of Δ1Iip33, Δ2Iip33 and Δ4Iip33. Transfected HeLa cells expressing Δ1Iip33 and CD8/E19 (**A** and **B**), or Δ2 Iip33 (**C**) or Δ4Iip33 (**D**) were fixed in 4% formaldehyde, permeabilized using 0.2% Triton X-100 and stained with rabbit anti-Ii antibody K2 (A, C and D) and OKT8, a mouse monoclonal antibody against CD8 (B). Fluorescein anti-rabbit or Texas Red anti-mouse secondary reagents were subsequently used to identify binding of the primary antibodies.

vesicles is unclear, previous work indicates that they are derived from the endosomal pathway (Romagnoli *et al.*, 1993). Further deletion of residues from the N-terminus of Iip33, e.g. Δ6, Δ9 and Δ12Iip33, resulted in mutants with subcellular distribution similar to that of Δ4Iip33 (data not shown).

In order to provide some biochemical analysis of Iip33 trafficking and turnover, we performed biosynthetic

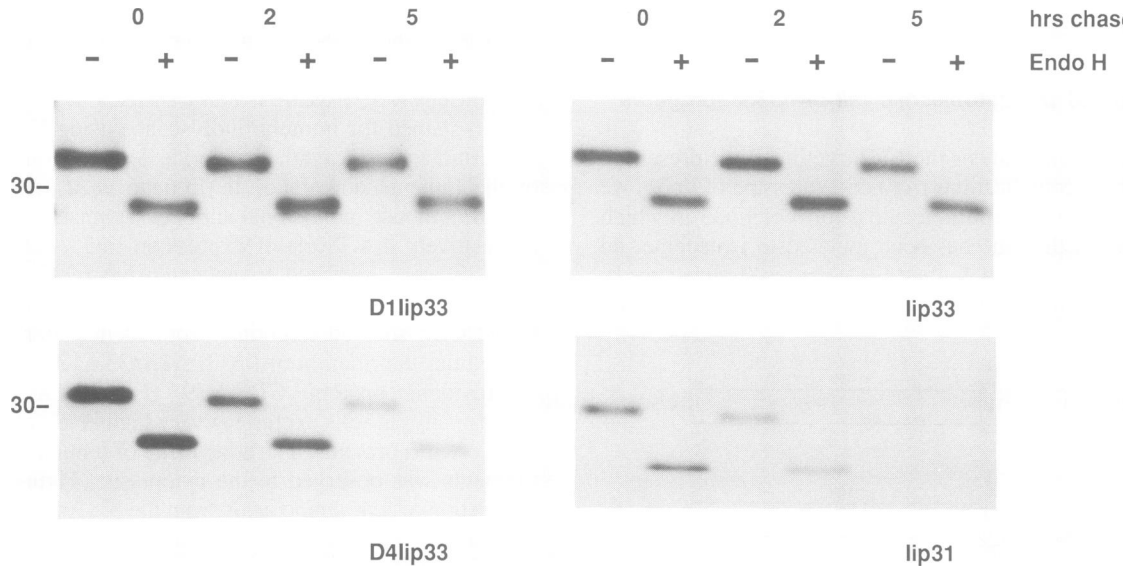

Fig. 3. Post-translational modification of Δ1Iip33 and Δ4Iip33 proteins. HeLa cells expressing Iip33, Δ1Iip33 (D1), Δ4Iip33 (D4) or Iip31 were labelled for 20 min with [^{35}S]methionine and then chased in the presence of medium containing cold methionine for 0, 2 or 5 h prior to lysis of the cells with 1% Triton X-100. Ii protein was immunoprecipitated from the cell lysates using the anti-Ii antiserum K2. The figure shows an autoradiograph of the immunoprecipitated Ii analyzed by SDS–PAGE. Each Ii precipitate is shown with (+) or without (−) prior digestion with endo H. The migration position of the 30 kDa marker protein carbonic anhydrase is indicated on the left of the figure.

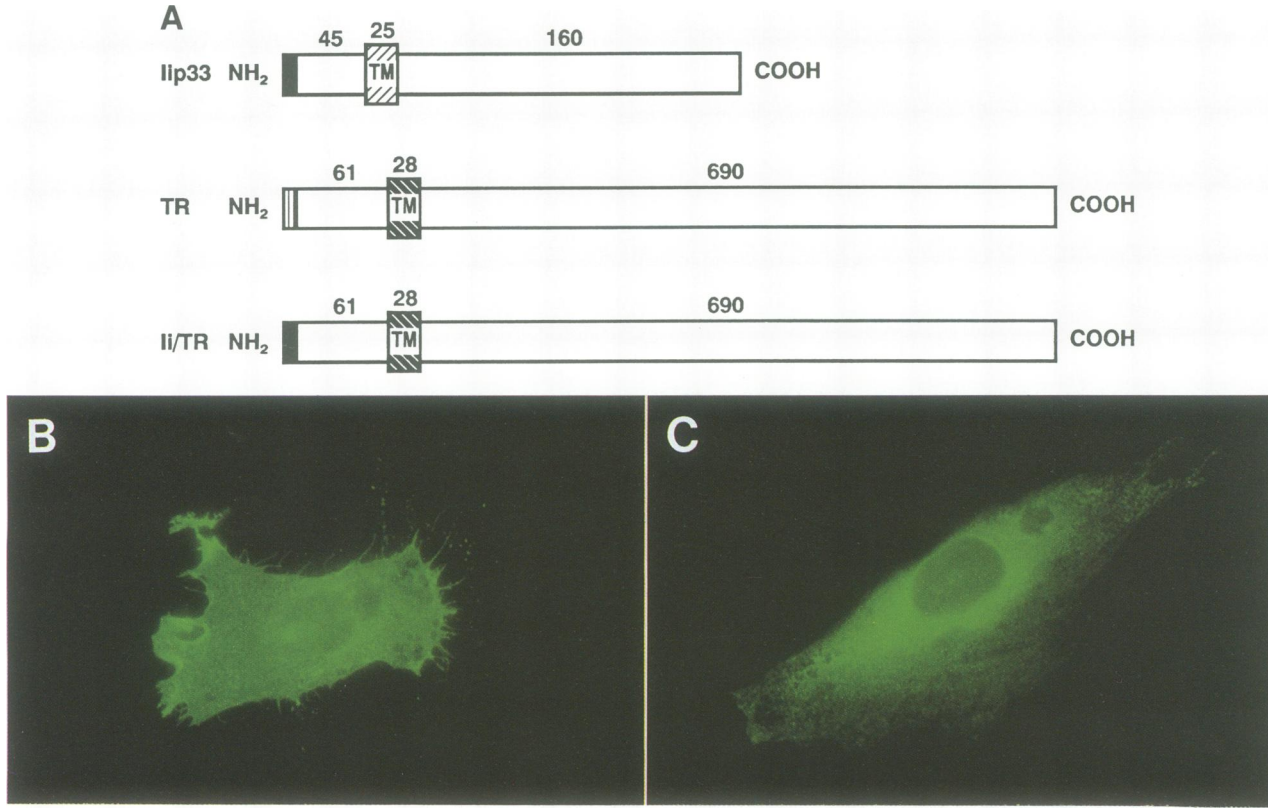

Fig. 4. Expression of TfR and Ii/TfR chimeric protein in transfected 3T3 cells. (**A**) Construction of a TfR chimera (Ii/TR) in which the N-terminal residues of TfR (MMDQA) were replaced by the N-terminal sequence of Iip33 (MHRRRS). The transmembrane region (TM) of TfR and Ii are shown boxed. (**B** and **C**) Immunofluorescence microscopy of mouse 3T3 cells transiently transfected with plasmids encoding wild-type TfR (B) or the Ii/TfR chimera (C). Permeabilized cells were stained with an anti-mouse antibody specific for the external domain of human TfR followed by fluorescein conjugated secondary antibody.

pulse–chase experiments on these Iip33 deletion mutants. HeLa cells transfected with the appropriate constructs were labelled with radioactive methionine for 20 min, chased for 0, 2 or 5 h in the presence of an excess of unlabelled methionine and subsequently lysed using 1% Triton X-100. Invariant chain was immunoprecipitated from the lysates using the K2 antiserum. Prior to analysis by SDS–PAGE, half of the immunoprecipitated protein was treated with endoglycosidase H (endo H). The results, presented in Figure 3, demonstrate that even after 5 h of chase Δ1Iip33 remained fully sensitive to endo H, suggesting that it had not been exposed to Golgi enzymes (Kornfeld and Kornfeld, 1985). Figure 3 also shows that Δ1Iip33 is not significantly degraded over the time course of the experiment. In contrast, the half-life of Δ4Iip33 is considerably shorter and significant degradation is observed after only 2 h of chase ($t_{½} < 2$ h). Since Iip31 has a short half-life relative to Iip33 (Figure 3), the long half-life of Δ1Iip33 ($t_{½} > 5$ h) is indicative of its retention in the ER. Interestingly, significant levels of endo H resistance for either Δ4Iip33 or Iip31 were not observed even after many hours of chase. This may be due to the short half-life of the protein once it has passed the Golgi or to the fact that the N-linked carbohydrates on Ii are a poor substrate for the Golgi enzymes, as discussed previously (Bäkke and Dobberstein, 1990; Lotteau *et al.*, 1990; Nilsson *et al.*, 1991). Taken together, the above results provide morphological and biochemical evidence that the ER sorting signal for Iip33 is located close to the N-terminus and that maintenance of Iip33 in the ER prevents its rapid degradation.

The N-terminal five residues of Ii retain human transferrin receptor in the endoplasmic reticulum

In order to determine if the sorting signal found in the N-terminal domain of Iip33 is sufficient to confer ER residency, we transplanted the five N-terminal residues of Iip33 (MHRRR) onto the N-terminus of the human transferrin receptor (TfR) (Figure 4A). The transferrin receptor is a type II integral membrane protein which is normally transported to the cell surface. The resulting chimera was expressed in mouse 3T3 cells as control experiments showed that the anti-TfR antibody did not recognize the endogenous mouse TfR (not shown). Figure 4B shows that cells transfected with wild type TfR expressed high levels at the cell surface. In contrast, expression of the chimeric molecule containing the N-terminal five residues transplanted from Iip33, resulted in a perinuclear and tubular network staining, typical of an ER-resident protein (Figure 4C). No cell surface expression of the chimeric TfR could be detected in non-permeabilized cells by either immunofluorescence microscopy or flow cytometry (not shown). These results indicate that the sorting signal MHRRR is both necessary and sufficient to cause ER localization of a type II integral membrane protein.

Arginines are the essential residues in Iip33 ER targeting motif

In order to determine the minimum requirements for the Iip33 ER targeting motif a series of single or double point mutations were introduced. Specific amino acids close to, or within, the putative signal were replaced with serine. The various mutants that were generated are summarized in Table I, and their subcellular localizations as determined by immunofluorescence microscopy are indicated. Immunofluorescence micrographs of six of the substitution mutants are shown in Figure 5. Replacement of the His2 residue or one of the arginine residues located at the N-terminal extremity with a serine residue did not disrupt the ER targeting of Ii (Figure 5). Similarly, if the N-terminal sequence of the wild type Ii33 (MHRRR) was replaced with the sequence MSSRR, MRRSS or MRSRS the resulting mutant proteins all localized to the ER (Figure 5A–D). In contrast, replacement of any two of R3, R4 or R5 with serine residues resulted in transport of the mutant molecule to the cell surface and to vesicular structures (Figure 5E and F). The above results showed that it is the presence of two or more arginine residues close to the N-terminus of Iip33, which targets this molecule to the ER. We henceforth refer to this targeting signal as the RR motif.

The position of the RR motif relative to the N-terminus is critical

The results presented in Table I show that, in order to achieve efficient targeting of Ii to the ER, the arginine residues need to be either adjacent to one another, as in

Table I. Characterization of the Iip33 ER targeting motif

	Sequences of N-terminal tail[a]																ER targeting[b]
Iip33	M	H	R	R	R	S	R	S	C	R	E	D	Q	K	P	V	+
p33+2S	M	**S**	R	R	R	S	R	S	C	R	E	D	Q	K	P	V	+
p33+3S	M	H	**S**	R	R	S	R	S	C	R	E	D	Q	K	P	V	+
p33+4S	M	H	R	**S**	R	S	R	S	C	R	E	D	Q	K	P	V	+
p33+5S	M	H	R	R	**S**	S	R	S	C	R	E	D	Q	K	P	V	+
p33+3S+5S	M	H	**S**	R	**S**	S	R	S	C	R	E	D	Q	K	P	V	−
p33+2S+3S+4S	M	**S**	**S**	**S**	R	S	R	S	C	R	E	D	Q	K	P	V	−
p33+2R+4S+5S	M	**R**	R	**S**	**S**	S	R	S	C	R	E	D	Q	K	P	V	+
p33+2R+3S+5S	M	**R**	**S**	R	**S**	S	R	S	C	R	E	D	Q	K	P	V	+
p33+2R+3S+4S	M	**R**	**S**	**S**	R	S	R	S	C	R	E	D	Q	K	P	V	−/+
p33+2S+3S	M	**S**	**S**	R	R	S	R	S	C	R	E	D	Q	K	P	V	+
p33+3S+4S	M	H	**S**	**S**	R	S	R	S	C	R	E	D	Q	K	P	V	−
p33+4S+5S	M	H	R	**S**	**S**	S	R	S	C	R	E	D	Q	K	P	V	−

[a]Bold letters correspond to residues altered from Iip33.
[b]Determined by immunofluorescence microscopy.
−/+ indicates partial retention of construct in the ER.

MHSRR (p33+3S), or separated by one residue, as in MHRSR (p33+4S) and MRSRS (p33+2R+3S+5S). When the two arginine residues were placed even further apart, e.g. in MHSRSSR (p33+3S+5S) or MRSSR (p33+2R+3S+4S), the resulting mutants were no longer retained in the ER, suggesting that the positioning of the arginines relative to one another is important for ER targeting. The precise location of the RR motif relative to the N-terminus was also critical, as for example MRSRSS (p33+2R+3S+5S) and MHRSRS (p33+4S) localized to the ER, whereas MSSSRSR (p33+2S+3S+4S) and MHSSRSR (p33+3S+4S) were transported to vesicles and the cell surface.

In order to analyze more precisely the importance of the location of the RR motif relative to the N-terminus, we determined the ability of a pair of adjacent arginines in various positions in an N-terminal environment composed of polyserine (nine residues) to target Ii to the ER. Table II summarizes the constructs and their subcellular localization when expressed in HeLa cells. This series of constructs showed that efficient targeting of this polyserine-containing Iip33 to the ER is only accomplished if two adjacent arginines are located in the four residues following the initiator methionine. Taken together with the data from Table I and Figure 1, we conclude that the consensus motif for an ER targeting motif for Ii is two arginine residues located in the N-terminal five amino acids but where the arginines are not separated by more than one residue.

The RR motif confers ER residency to human transferrin receptor

To investigate whether a motif composed of only two arginines is sufficient to target a reporter type II membrane

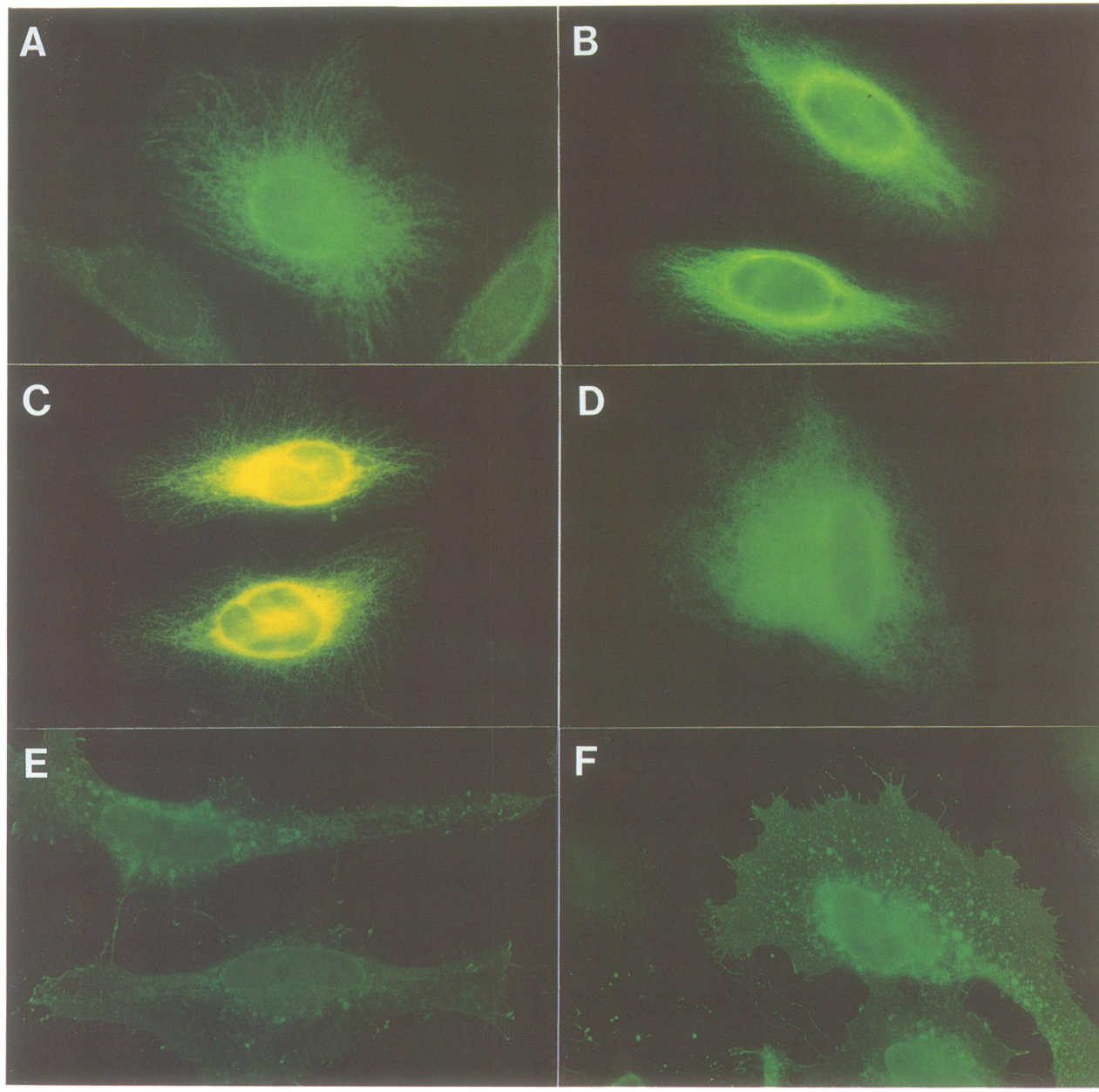

Fig. 5. Immunofluorescence microscopy of serine substitution mutants of Iip33. Immunofluorescence micrographs of HeLa cells expressing (**A**) p33+3S, (**B**) p33+4S, (**C**) p33+2R+4S+5S, (**D**) p33+2R+3S+5S, (**E**) p33+4S+5S and (**F**) p33+3S+4S, processed and stained for invariant chain as described in the legend to Figure 2. Details of the N-terminal sequences of the mutants are given in Table I.

protein to the ER, we transplanted such a signal onto the N-terminus of human TfR. The results presented in Table III demonstrate that the various RR motifs that were characterized above to maintain Ii in the ER functioned in a similar, but not quite identical manner when expressed in the context of the TfR cytoplasmic tail. Whereas two or three adjacent arginine residues at the appropriate position led to retention of the TfR chimera in the ER (Table III) a motif composed of for example, MRSR, which we found to maintain Ii in the ER (see Table III for summary), did not fully maintain TfR in the ER. These results suggest that the function of the RR motif is influenced either by other characteristics of the cytoplasmic tail, e.g. tail length or sequence, and/or by some other property of the reporter protein such as multimerization state (see Discussion).

Lysine residues can in some contexts functionally replace the critical arginine residues in the RR motif

One obvious question arising from the identification of arginine residues as the essential residues of the Iip33 ER targeting motif was whether conservative substitution of these arginines would disrupt the targeting motif. The results, summarized in Table IV, show that in some instances lysine can substitute for one of the arginines of the RR motif but never for both. For example, replacement of both arginines in the ER targeted constructs p33+5S (MHRRSS), p33+3S (MHSRRS) and Δ2p33 (MRRS) results in mutants p33+3K+4K+5S (MHKKS), p33+3S+4K+5K (MHSKKS) and Δ2p33+4K+5K, which were all transported out of the ER. However, if only the +4 position arginine in p33+5S (MHRRS) or the +5 arginine in p33+3S (MHSRRS) was replaced with lysine, the resulting molecules MHRKS (p33+4K+5S) and MHSRKS (p33+3S+5K), were targeted to the ER. Note, that if these same residues were replaced by serine, as in constructs p33+4S+5S (MHRSS) and p33+3S+5S (MHSRSS), the ER targeting motif (see Table I) was lost. The relative position between the lysine and arginine residues appears more important than previously observed for adjacent arginine residues. Thus, although the arginine-lysine motif functions when the two basic residues are placed adjacent to each other (MHSRK, MHRKS, MHKKR), it does not appear to work when the two residues are separated, e.g. in MHKSR. In summary, one of the two arginines, but not both, can in some sequence contexts be replaced with lysine without destroying the RR ER targeting motif.

RR-like ER targeting motifs in other ER-resident proteins

Searching of the protein sequence database for ER-resident proteins, which might be maintained in the ER by the presence of basic residues located close to their N-terminus, revealed two proteins, TRAM (Görlich *et al.*, 1992) and p63 (Schweizer *et al.*, 1993a), as potential candidates. Using

Table II. The position of the double-arginine motif at the N-terminus is required for ER targeting

Sequence of N-terminal tail[a]	ER targeting[b]
M **R R** S S S S S S R S C R E	+
M S **R R** S S S S S R S C R E	+
M S S **R R** S S S S R S C R E	+
M S S S **R R** S S S R S C R E	−/+
M S S S S **R R** S S R S C R E	−

[a]The N-terminal sequences transferred to Δ5Iip33 are indicated in bold letters.
[b]Determined by immunofluorescence microscopy.

Table III. Targeting of Ii and TfR to the ER by various RR motifs

	Sequence of N-terminal tail	Surface expression	Intracellular distribution
Wild type TfR	MMDQ/ARSAFS -TfR	yes	endosomes
p33+2S/TfR	MSRRR/SRSAFS -TfR	no	ER
p33+5S/TfR	MHRRS/SRSAFS -TfR	no	ER + intermediate compartment
p33+4S/TfR	MHRSR/SRSAFS -TfR	yes (weak)	ER + Golgi
p33+2S+3S+4S/TfR	MSSSR/SRSAFS -TfR	yes	ER + endosomes
p33+2S	MSRRR/SRSCRE -Ii	no	ER
p33+5S	MHRRS/SRSCRE -Ii	no	ER
p33+4S	MHRSR/SRSCRE -Ii	no	ER
p33+2S+3S+4S	MSSSR/SRSCRE -Ii	yes	ER + Golgi

Table IV. The critical arginine residues in the ER targeting motif can be partially replaced by lysine

	Sequences of N-terminal tail[a]																ER targeting[b]
Iip33	M	H	R	R	R	S	R	S	C	R	E	D	Q	K	P	V	+
Δ2p33+4K+5K			M	**K**	**K**	S	R	S	C	R	E	D	Q	K	P	V	−/+
p33+3S+5K	M	H	**S**	R	**K**	S	R	S	C	R	E	D	Q	K	P	V	+
p33+4K+5S	M	H	R	**K**	**S**	S	R	S	C	R	E	D	Q	K	P	V	+
p33+3K+4S	M	H	**K**	**S**	R	S	R	S	C	R	E	D	Q	K	P	V	−/+
p33+3S+4K+5K	M	H	**S**	**K**	**K**	S	R	S	C	R	E	D	Q	K	P	V	−
p33+3K+4K+5S	M	H	**K**	**K**	**S**	S	R	S	C	R	E	D	Q	K	P	V	−
p33+3K+4K	M	H	**K**	**K**	R	S	R	S	C	R	E	D	Q	K	P	V	−/+

[a]Bold letters correspond to substituted residues.
[b]Determined by immunofluorescence microscopy.

oligonucleotide cassette mutagenesis we replaced the DNA encoding the six N-terminal residues of Iip33 (MHRRRS) with that encoding the N-terminal sequence of either TRAM (MAIRKKS) or p63 (MPSAKQRGS). When expressed in HeLa cells, the Ii chimeras encoded by these constructs were observed by immunofluorescence microscopy to be maintained in the ER. In contrast, replacing the arginine residue in the TRAM sequence (MAISKKS) or the lysine in the p63 sequence (MPSASQRGS) resulted in disruption of the ER targeting motif and transport of the Ii chimera to the cell surface and to perinuclear vesicles characteristic of Iip31 (see Figure 2D). These data suggest that the ER targeting strategy utilized by Iip33 is not unique to this protein, but rather is shared by a family of ER-resident proteins, all maintained in this organelle by basic residues located close to the N-terminus of their cytoplasmic tails.

RR-tagged transferrin receptor partially localizes to the ER–Golgi intermediate compartment

Recently, evidence has been presented in favor of the existence of efficient retrieval pathways for type I membrane proteins tagged with a KK motif (Jackson *et al.*, 1993) and for soluble ER lumenal proteins with a KDEL motif (Pelham, 1989; Dean and Pelham, 1990; Lewis and Pelham, 1992).

In order to investigate whether proteins maintained in the ER by an RR motif are retrieved from post-ER compartments, we examined the possible co-localization of the RR-tagged TfR molecules with p58, a resident of the ER–Golgi intermediate compartment (Saraste *et al.*, 1987; Saraste and Svensson, 1991). The results presented in Figure 6 demonstrate that the TfR chimera with the N-terminal sequence MHRSR localized both to ER and to vesicular structures that co-stained for p58 (Figure 6A and B). Co-localization could only be observed at low levels of protein expression since at higher levels the bright ER staining obscured details of the perinuclear region. Furthermore, if transfected cells were treated with the microtubule disrupting agent nocodozole (De Brabander *et al.*, 1976), a procedure which accentuates the intermediate compartment by changing its structure (Lippincott-Schwartz *et al.*, 1989, 1990; Saraste and Svensson, 1991), co-localization of the TfR chimera and p58 is clearly apparent (Figure 6C). Using confocal immunofluorescence microscopy and image overlay a coincident staining pattern between the TfR chimera and p58 in nocodazole-treated cells was observed (Figure 6D). Co-localization in p58-staining vesicular structures was also observed between the TfR chimera and CD4/E19, a CD4 chimera bearing a KK ER targeting motif (data not shown), consistent with a common retrieval pathway for type I and type II ER-resident membrane proteins that have escaped from the ER.

Discussion

In this paper we show that the ER targeting of Iip33 is directed by multiple arginine residues close to the N-

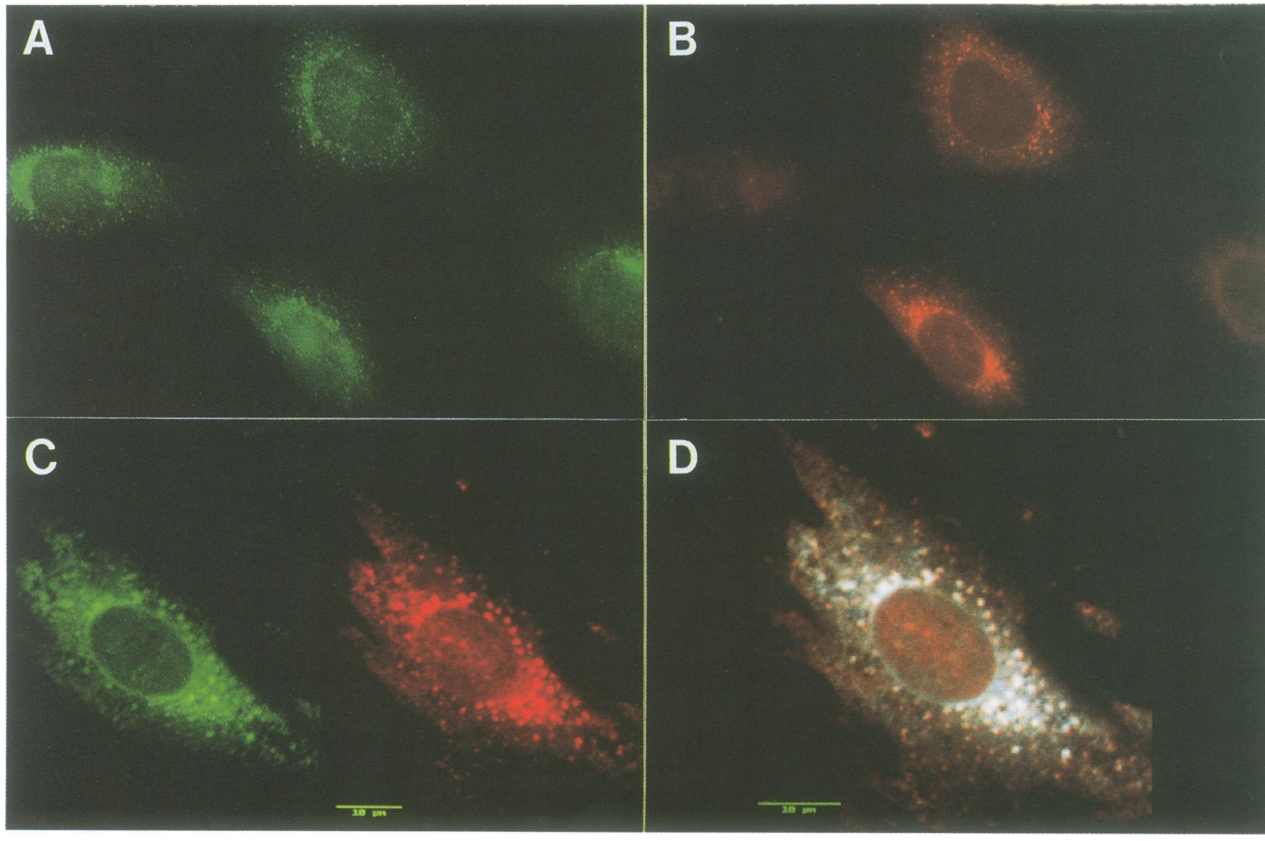

Fig. 6. Co-localization of RR-tagged TfR and p58 in post-ER compartment(s). 3T3 cells expressing the chimera p33+5S/TfR (see Table III for details) were fixed with formaldehyde, permeabilized with 0.1% Triton X-100 and co-stained with antibodies to p58 (**A**) and human TfR (**B**). Fluorescein anti-rabbit and Texas Red anti-mouse secondary anibodies were used to identify binding of the primary antibodies. (**C**) Analysis similar to that shown in panels A and B except that cells were treated with nocodazole for 2 h prior to staining, and fluorescence imaging was achieved using a confocal microscope. The left of panel C shows staining for p58 (green) while the right of the panel shows TfR staining (red). (**D**) Confocal images of panel C merged. Overlap of fluorescein and Texas Red staining is shown in white.

terminus. We have characterized this sorting motif using invariant chain as the reporter protein and arrived at a consensus which consists of two arginine residues located in the N-terminal five amino acids but where the arginines are not separated by more than one residue. In some cases arginine may be replaced with lysine although the precise position and combinations that are functional depends on the sequence context. It is presumably this phenomenon that explains why the N-terminal sequence of p63 (MPS-AKQRGS-) efficiently targets Ii to the ER even though this sequence does not quite fit the above RR consensus. Thus the consensus features we have determined for this RR motif represent only the essence of this motif and we fully expect that depending on the exact sequence context in which the basic residues are placed and the reporter protein analyzed, further combinations of basic residues not covered by the above RR consensus will be found that function as ER targeting motifs.

The characteristics of the RR motif we have defined show similarities to those of the KK motif demonstrated to target type I membrane proteins to the ER (Jackson *et al.*, 1990). The consensus for the KK motif is two lysine residues located at the positions −3 and −4 or −5 from the C-terminus. In discussing the similarities between these two motifs and in particular the positioning of the lysine and arginines with respect to the terminus it is worth pointing out that it is possible that the initiator methionine may be trimmed from Iip33. However, an initiator methionine followed by a histidine as is the case in p33 is considered a poor substrate for the trimming enzyme (Moerschell *et al.*, 1990). Nevertheless, as trimming is sensitive to the nature of the +2 amino acid, differences in the efficiency of this process may account for the finding that the RR motif appears to function further away from the N-terminus in some proteins, e.g. p63, than in others, e.g. p33. In order to avoid confusion the amino acid positions to which we have referred throughout this paper assume that the initiator methionine is present.

The similarities between the consensus motifs we have defined for these two targeting signals is nicely demonstrated by the protein TRAM, an ER-resident protein (Görlich *et al.*, 1992) which is proposed to span the membrane eight times and has both termini exposed in the cytoplasm. We show above that the N-terminal sequence of human TRAM, NH_2-MAIRKKST-, contains an ER targeting motif. Similarly, the C-terminal sequence of TRAM, -NKKEKSS-COOH, contains two lysine residues positioned appropriately to fit the consensus for a KK ER targeting motif (Jackson *et al.*, 1990). Although it is not known if either or both of these signals are responsible for maintaining TRAM in the ER, the characteristics of the two putative ER targeting motifs in this protein serve to make the point that despite strong similarities in the sequence requirements for these two motifs they are also different. Whereas the ER targeting motif for type I membrane protein is based essentially on lysine residues, that for type II membrane proteins is based essentially on arginines. Although some replacement of arginine is permitted by lysine in the RR motif and, similarly, some replacement of lysine by arginine is permissible in the KK motif (Shin *et al.*, 1992), we have found no example in which targeting of type I membrane proteins can be achieved with a motif made up only of arginines. Similarly, we have found no example of ER targeting of type II membrane proteins by a motif made up solely of lysines. These data suggest that if a common sorting machinery is used to achieve targeting by these motifs it is likely that more than one putative receptor molecule would be required to recognize these motifs. It is tempting to speculate that a family of putative receptor molecules exists that recognizes these motifs, where the binding site is a pocket which accommodates only termini of proteins, binding specificity being accomplished by charged interactions between the wall of the pocket and the basic residues of the motif.

Is the RR motif a retrieval signal?

The RR motif is unlikely to be involved in a simple receptor−ligand interaction where the receptor is somehow tethered and fixed in the ER as we have been unable to saturate this system. We favor a model in which the RR motifs function by directing sorting of proteins in post-ER compartments resulting in the retrieval of escaped molecules back to the ER (Pelham, 1989), in a manner analogous to that which we described for targeting by the KK motif (Jackson *et al.*, 1993). The only direct evidence we have that the RR motif may direct retrieval to the ER is the observation that TfR molecules tagged with the sequence MHRSRS partially co-localized to the ER−Golgi intermediate compartment, and that this co-localization increased if cells were treated with nocodazole. However, identification of this chimera in a post-ER location does not prove that such proteins are subsequently retrieved to the ER: they could for example be slowly leaking out of the ER. Nevertheless, a retrieval model is consistent with the finding that marker proteins with arginines close to their N-terminus but in a sub-optimal configuration are relatively slowly transported. We have previously suggested that inefficient retrieval may best explain the slow rate of transport of type I membrane proteins with sub-optimal configurations of lysine residues (Jackson *et al.*, 1990, 1993). Perhaps the strongest support for the RR motif directing retrieval is that it shows such similar characteristics to the KK motif which we have previously shown to be an ER retrieval motif (Jackson *et al.*, 1993). Nevertheless, without more direct evidence that RR-tagged proteins are actually retrieved to the ER, the idea that RR-tagged proteins recirculate between the ER and the Golgi remains a working hypothesis.

Correlation between the efficiency of ER targeting and the oligomeric state of the reporter protein

The observation that the sorting signal MHRSR efficiently maintained Ii in the ER but not TfR suggests that the nature of the marker protein is an important component in modulating the ability of a sequence to target. We have previously suggested that the oligomeric state of the reporter protein can determine the efficiency of targeting by the KK motif (Jackson *et al.*, 1993). Invariant chain is known to form trimers in the ER (Marks *et al.*, 1990), and probably forms even larger complexes, whereas TfR is a dimer. Thus it is quite plausible that MHRSR may function sufficiently well to maintain a trimer in the ER but not a dimer. Indeed, it is worth noting that despite the fact that detectable levels of the MHRSR/TfR chimera reach the cell surface the majority of the MHRSR/TfR protein is localized to the ER, indicating that the motif does function but not sufficiently well to prevent some leakage.

Interestingly, neither Iip33 nor any of the ER-maintained Iip33 mutants described in this work were observed to co-localize to the intermediate compartment. Furthermore, analysis of the post-translational modifications of Iip33 after long chase times by 2-D gel analysis showed no evidence to suggest that Iip33 had been exposed to Golgi enzymes. These latter findings contrast with a similar analysis of Iip33 in the B lymphoblastoid cell lines, Raji and Swei, in which a small but significant portion of Iip33 becomes sialylated (L.Karlsson, unpublished data; Newcomb and Cresswell, 1993). However, the sialylated forms of Iip33 observed in these B-cell lines were not observed in RJ 2.2.5, a cell line derived from Raji, in which MHC class II chains are expressed at very low levels. The explanation for these apparently conflicting observations may therefore lie in whether Iip33 alone or Iip33 in the presence of MHC class II chains is being analyzed. When Iip33 is analyzed alone, e.g. in transfected HeLa cells, it exists as trimers or most likely as even bigger complexes (Marks *et al.*, 1990), so each molecule contains three or more RR motifs. In contrast, the majority of Iip33 in B cells is most likely bound to MHC class IIα/β and/or Iip31 and, as such, would possess only one RR motif per molecule. As discussed above one would predict that a molecule possessing three or more RR motifs would be very much more efficiently targeted to the ER than one possessing only one motif. The more efficient the retrieval process, the shorter the dwell time in post-ER compartments and consequently the slower the rate of modification by enzymes located in these compartments. Thus, our inability to detect Iip33 in the intermediate compartment of transfected cells may reflect the fact that it is very rapidly retrieved and therefore not accumulated in such structures to levels high enough to be detected rather than the fact that the protein never leaves the ER.

Alteration of ER morphology by over-expression of Iip33

Although we did not detect Iip33 in vesicles of the intermediate compartment, the ER of cells expressing high levels of this protein appeared to be remodelled into tubular structures (see Figure 2C). A similar phenomenon has been previously reported in cells in which p63 is over-expressed (Schweizer *et al.*, 1993a). p63 is a 63 kDa membrane protein which has been characterized as a resident protein of a membrane network interposed between the ER and Golgi apparatus (Schweizer *et al.*, 1993a,b). In contrast to the intermediate compartment marker protein p53 and its putative human homolog, p58 (Schindler *et al.*, 1993), the subcellular distribution of p63 was unaffected by low temperature or brefeldin A. However, p63 and p53 were found to be enriched in the same subcellular fractions, distinct from typical ER markers, suggesting that they were part of a complex ER–Golgi intermediate.

Based on the observation that over-expression of p63 resulted in the development of these filamentous structures, Schweizer *et al.* (1993a) suggested that p63 is the determinant for the structure of the compartment marked by this protein. However, as we have shown above that the N-terminal sequence of p63 contains an RR-like motif which can functionally replace the motif present in Iip33 and that Iip33 also results in the proliferation of similar tubular membrane structures, an alternative explanation for this phenomenon is required. We propose that these tubular membrane structures may result from an increase in the steady state distribution of membrane to an ER–Golgi intermediate compartment due to over-expression and accumulation of resident proteins of this compartment. Given that the RR motif maintaining p33 in the ER is likely to be an ER retrieval motif it seems quite plausible that p33 might localize to and concentrate in the same ER–Golgi sub-compartment as p63 (see Hauri and Schweizer, 1992; Schweizer *et al.*, 1993b). Recent work (Hobman *et al.*, 1992) has shown that over-expression of unassembled rubella virus E1 glycoprotein results in amplification of a pre-Golgi compartment consisting of a tubular network of smooth membranes.

In conclusion we have characterized a general ER targeting motif for type II membrane proteins. The essential features of this motif are two arginine residues close to the N-terminus. The similarities to a double-lysine ER targeting motif utilized by type I membrane proteins suggest that both motifs probably function via a common retrograde transport pathway. The detailed characterization of this ER targeting motif we have carried out in this paper provides the necessary groundwork for further experiments aimed at understanding the molecular basis by which ER targeting is achieved.

Materials and methods

Recombinant DNAs

Construction of the expression plasmids utilized the polymerase chain reaction (PCR) and full-length cDNAs encoding Iip33 as the template. PCR was carried out essentially as previously described (Nilsson *et al.*, 1989). For construction of Δ2Iip33, two oligonucleotides were used as primers. The 5′ primer encoded a *Sac*I site followed by a *Bam*HI site, the Kozak consensus sequence, ATG and the sequence encoding residues 3–10 of Iip33. The 3′ primer consisted of the sequence encoding residues 26–29 followed by a *Sac*II site. The PCR product Δ2Iip33 encoded the cytoplasmic tail of Iip33. This PCR product was digested with *Sac*I and *Sac*II and subcloned between the *Sac*I and *Sac*II sites of a pBluescript vector containing a *Sac*II–*Bam*HI fragment corresponding to the remaining transmembrane and extracellular domain of Iip33. Following sequencing, the insert was excised with *Bam*HI and cloned into a *Bam*HI site of an expression vector pCMU IV (Nilsson *et al.*, 1989) for transient expression in HeLa cells. Similar protocols were followed for Δ6, Δ9 and Δ12Iip33.

Δ4Iip33 was constructed as described above using two primers: a 5′ primer encoding a *Sac*I site followed by the Kozak consensus, ATG, an *Xba*I site and the sequence encoding residues 8–13 of Iip33 with the previously described 3′ primer. The *Xba*I site was introduced at codon 6 without changing the encoded amino acid sequence (AGCTGT was changed to TCTAGA). All further tail mutants of Iip33 were constructed by inserting oligonucleotides between the unique *Sac*I and *Xba*I sites in this plasmid. All mutants were directly sequenced in the pCMU expression vector.

The recombinant TfR/Ii chimera was constructed using the PCR and the full-length TfR cDNA (McClelland *et al.*, 1984), a kind gift of I.Trowbridge (Salk Institute, La Jolla, CA) as the template and a 5′ primer containing a *Hin*dIII site, a *Sac*I site, the Kozak consensus sequence, CCACCATG, followed by an *Xba*I site and the sequence encoding residues 7–11 of wild type TfR (McClelland *et al.*, 1984) and a 3′ primer containing the reverse strand sequence encoding residues 269–274 of the wild type TfR (McClellan *et al.*, 1984) followed by a *Hin*dIII site. Introduction of this *Xba*I site at the 5′ extremity was such that the resulting amino acid sequence at codon five in the TfR sequence was unaltered (GCTAGA to TCTAGA). The generated PCR fragment encoded an initiator methionine followed by amino acids 5–276 of TfR corresponding to an almost complete cytoplasmic tail and transmembrane domain. In order to ensure that the *Xba*I site introduced into the PCR fragment was unique in the reconstructed TfR, we eliminated the only other *Xba*I site in TfR located at the C-terminal extremity by cleaving the TfR with *Xba*I and filling in the overhangs with Klenow and religating. Reconstruction of the recombinant TfR with the altered 5′ sequence was achieved by cleaving the PCR fragment with *Hin*dIII and subcloning this fragment into the *Hin*dIII site of a pCMU recombinant containing the remaining extracellular domain of TfR directionally cloned between the

*Hin*dIII site (5′) and the *Bam*HI site (3′) in the pCMU polylinker. The unique *Sac*I and *Xba*I sites at the 5′ extemity of the reconstructed TfR, i.e. just downstream from the initiator methionine, were used to clone directly the same series of *Sac*I–*Xba*I oligonucleotides used to mutagenize the N-terminus of Iip33 into the N-terminus of TfR.

Cell culture and transfection

HeLa (ATCC CCL185) and NIH 3T3 cells were grown in Dulbecco's modified Eagle's medium (DMEM; Gibco/BRL) supplemented with 10% fetal calf serum, 2 mM glutamine, and penicillin and streptomycin at 100 μg/ml. Transfections were carried out using the calcium phosphate method as described previously (Nilsson *et al.*, 1989) except that 25 μg of DNA were used to transfect 1×10^5 cells in a 100 mm Petri dish.

Metabolic labelling, immunoprecipitation and endo H treatment

Metabolic labelling of cells was carried out as follows, 72 h after transfection, cells were incubated for 20 min in methionine-deficient DMEM prior to the addition of pulse medium. Pulse medium containing [^{35}S]methionine (Trans ^{35}S-label; ICN Biochemicals) was added to a final concentration of 0.2 mCi/ml of methionine-deficient DMEM (Gibco) and the cells were incubated for 20 min followed by chase periods of 2 and 5 h in the presence of normal culture medium (see above). Subsequently, the cells were washed three times with ice-cold PBS and lysed with PBS containing 1% Triton X-100 and 20 μM PMSF. Following a 15 min incubation (4°C), the cell lysates were cleared by centrifugation for 15 min at 15 000 *g* and then incubated with protein A–Sepharose (Pharmacia) for 1 h at 4°C. After removal of the Sepharose beads by centrifugation the lysates were incubated with the appropriate antibody for 3 h at 4°C. Antibody precipitates were collected by addition of a second aliquot of protein A–Sepharose and the immunoprecipitates washed twice with 1 ml of low salt buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 2 mM EDTA and 0.2% Triton X-100), twice with 1 ml of high salt buffer (10 mM Tris–HCl pH 7.5, 500 mM NaCl, 2 mM EDTA and 0.2% Triton X-100) and twice with 1 ml of 10 mM Tris–HCl pH 7.5. Where appropriate, samples were treated with endo H for 12 h at 37°C, (Boehringer Mannheim) as described by the manufacturer prior to analysis by SDS–PAGE.

Nocodazole treatment of cells

In order to depolymerize microtubules, cells were incubated in ice-cold medium for 25 min. The medium was then replaced with 37°C DMEM for the control cells, or with DMEM containing 10 μg/ml nocodazole (Aldrich, Milwaukee, WI); after 2 h in this medium, cells were prepared for immunofluorescence as described below.

Immunofluorescence microscopy

Cells were fixed and stained for immunofluorescence as described previously (Jackson *et al.*, 1990). Briefly, 48 h after transfection, the cells were trypsinized and plated on coverslips precoated with Celtak as described by the manufacturer (Collaborative Research). Twelve hours later, cells were fixed with 4% formaldehyde in PBS for 20 min. After removal of the formaldehyde, coverslips were incubated for 10 min in 50 mM NH_4Cl/PBS. Cells were permeabilized for 3 min in 1% Triton X-100/PBS followed by 30 min incubation with 0.2% gelatin/PBS. Cells were incubated for 20 min with the specific antibody and then stained with fluorescein- (Cappell Laboratories, Malvern, PA) or Texas Red- (Molecular Probes, Eugene, OR) labelled secondary antibodies for 20 min. Fluorescence microscopy was performed using a Zeiss Axiophot microscope using a 63× objective. Kodak Ektachrome 400 film was used for photography. Confocal microscopy was performed using a Bio-Rad confocal laser-scanning microscope.

Antibodies

A polyclonal rabbit antiserum specific for the lumenal domain of the invariant chain was raised against the C-terminal peptide (KESLELEDPSSGLGV-TKQDL) of Ii. Ascites fluid containing monoclonal antibodies against CD8 (OKT8) was prepared using hybridoma CRL 8014 from ATCC. A monoclonal antibody against the C-terminal domain of the TfR was kindly provided by I.Trowbridge (Salk Institute, La Jolla, CA) and a polyclonal rabbit anti-p58 antibody (Saraste *et al.*, 1987) was the kind gift of J.Saraste (Ludwig Institute for Cancer Research, Stockholm, Sweden).

Acknowledgements

We are especially grateful to our colleagues in the laboratory in particular L.Teyton, V.Lotteau, L.Karlsson, R.Castaño and L.Pond for helpful comments and reagents. We thank J.Chambers for sequence analysis and oligonucleotide synthesis and I.Trowbridge and J.McClellan for TfR cDNA and anti-TfR antibodies. This work was supported by grants from the NIH to P.A.P. (GM46411) and M.R.J. (GM 49305). M.-P.S. was supported by an NIH Fogarty International Fellowship.

References

Bäkke,O. and Dobberstein,B. (1990) *Cell*, **63**, 707–716.
Burke,B., Griffith,G., Reggio,H., Louvard,D. and Warren,G. (1982) *EMBO J.*, **1**, 1621–1628.
Dean,N. and Pelham,H.R.B. (1990) *J. Cell Biol.*, **111**, 369–377.
De Brabander,M.J., Von de Veire,R.M.L., Aerts,R.M., Borgers,M. and Janssen,P.N.J. (1976) *Cancer Res.*, **36**, 905–916.
Görlich,D., Hartmann,E., Prehn,S. and Rapoport,T.A. (1992) *Nature*, **357**, 47–52.
Harter,C. and Mellman,I. (1992) *J. Cell Biol.*, **117**, 311–325.
Hauri,H.-P. and Schweizer,A. (1992) *Curr. Opin. Cell Biol.*, **4**, 600–608.
Hobman,T.C., Woodward,L. and Farquhar,M.G. (1992) *J. Cell Biol.*, **118**, 795–811.
Humphreys,J.S., Peters,P.J., Yuan,L.C. and Bonifacino,J.S. (1993) *J. Cell Biol.*, **120**, 1123–1135.
Jackson,M.R., Nilsson,T. and Peterson,P.A. (1990) *EMBO J.*, **9**, 3153–3162.
Jackson,M.R., Nilsson,T. and Peterson,P.A. (1993) *J. Cell Biol.*, **121**, 317–333.
Johnson,K.F. and Kornfeld,S. (1992) *J. Biol. Chem.*, **267**, 17110–17115.
Kornfeld,R. and Kornfeld,S. (1985) *Annu. Rev. Biochem.*, **54**, 631–664.
Kvist,S., Wiman,K., Claesson,L., Peterson,P.A. and Dobberstein,B. (1982) *Cell*, **29**, 61–69.
Letourneur,F. and Klausner,R.D. (1992) *Cell*, **69**, 1143–1157.
Lewis,M.J. and Pelham,H.R.B. (1992) *Cell*, **68**, 353–364.
Lippincott-Schwartz,J., Yuan,L.C., Bonifacino,J.S. and Klausner,R.D. (1989) *Cell*, **56**, 801–813.
Lippincott-Schwartz,J., Donaldson,J.G., Schweizer,A., Berger,E.G., Hauri,H.-P., Yuan,L.C. and Klausner,R.D. (1990) *Cell*, **60**, 821–836.
Lotteau,V., Teyton,L., Peleraux,A., Nilsson,T., Karlsson,L., Schmid,S.L., Quaranta,V. and Peterson,P.A. (1990) *Nature*, **348**, 600–605.
Machamer,C.E. (1993) *Curr. Opin. Cell Biol.*, **5**, 606–612.
Marks,M.S., Blum,J.S. and Cresswell,P. (1990) *J. Cell Biol.*, **111**, 839–855.
McClelland,A., Kuhn,L.C. and Ruddle,F.H. (1984) *Cell*, **39**, 267–274.
Moerschell,R.P., Hosokawa,Y., Tsunasawa,S. and Sherman,F. (1990) *J. Biol. Chem.*, **265**, 19638–19643.
Munro,S. and Pelham,H.R.B. (1987) *Cell*, **48**, 899–907.
Newcomb,J.R. and Cresswell,P. (1993) *J. Immunol.*, **151**, 4153–4163.
Nilsson,T., Jackson,M.R. and Peterson,P.A. (1989) *Cell*, **58**, 707–718.
Nilsson,T., Lucocq,J.M., Mackay,D. and Warren,G. (1991) *EMBO J.*, **10**, 3567–3575.
Pelham,H.R.B. (1989) *Annu. Rev. Cell Biol.*, **5**, 1–23.
Pfeffer,S.R. and Rothman,J.E. (1987) *Annu. Rev. Biochem.*, **56**, 829–852.
Romagnoli,P., Layet,C., Yewdell,J., Bäkke,O. and Germain,R.N. (1993) *J. Exp. Med.*, **177**, 583–596.
Saraste,J. and Svensson,K. (1991) *J. Cell Sci.*, **100**, 415–430.
Saraste,J., Palade,G.E. and Farquhar,M. (1987) *J. Cell Biol.*, **105**, 2021–2029.
Schindler,R., Itin,C., Zerial,M., Lottspeich,F. and Hauri,H.-P. (1993) *Eur. J. Cell Biol.*, **61**, 1–9.
Schwiezer,A., Rohrer,J., Jeno,P., DeMaio,A., Buchman,T.G. and Hauri,H.-P. (1993a) *J. Cell Sci.*, **104**, 685–694.
Schweizer,A., Ericsson,M., Bachi,T., Griffith,G. and Hauri,H.-P. (1993b) *J. Cell Sci.*, **104**, 671–683.
Shin,J., Dunbrack,R.L., Songjae,L. and Strominger,J.L. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 1918–1922.
Strubin,M., Long,E.O. and Mach,B. (1986) *Cell*, **47**, 619–625.
Trowbridge,I.S. (1991) *Curr. Opin. Cell Biol.*, **3**, 634–641.
Wilcox,C.A., Redding,K., Wright,R. and Fuller,R.S. (1992) *Mol. Biol. Cell*, **3**, 1353–1371.

Received on November 30, 1993; revised on January 17, 1994