Volume 253, number 1,2, 93–98 FEB 07461 August 1989

Signal peptidase can cleave inside a polytopic membrane protein

James P. Beltzer, Hans Peter Wessels and Martin Spiess

Department of Biochemistry, Biocenter, University of Basel, CH-4056 Basel, Switzerland

Received 14 June 1989

The signal peptides of most proteins targeted to the endoplasmic reticulum are specifically cleaved by signal peptidase. Although potential cleavage sites occur frequently in polytopic proteins after membrane-spanning segments, processing is restricted to the first hydrophobic domain, suggesting that signal peptidase might not have access to subsequently translocated, internal domains. To test this hypothesis, we replaced the third transmembrane segment of an artificial threefold membrane-spanning protein by a sequence which is normally an amino-terminal signal. Upon in vitro translation and insertion into microsomes, efficient cleavage at this sequence was observed, thus demonstrating the ability of signal peptidase to cleave within polytopic membrane proteins.

Asialoglycoprotein receptor; Signal peptidase; Polytopic membrane protein; Hemagglutinin

1. INTRODUCTION

In eukaryotic cells, most secretory and membrane proteins are targeted to the endoplasmic reticulum (ER) by amino-terminal signal peptides of typically 16–26 residues [1,2]. They contain a characteristic stretch of 8–12 apolar residues and are recognized by signal recognition particle (SRP) [3]. SRP in combination with its receptor in the ER membrane is responsible for directing the translation complex to the correct organelle. As translocation of the polypeptide across the membrane is initiated, the signal peptide interacts with a transmembrane protein termed the signal sequence receptor [4], presumed to be a component of the translocation machinery. Finally, most signal peptides are proteolytically removed from the nascent polypeptide by a highly specific endoprotease on the luminal side of the ER membrane.

The high specificity of this signal peptidase contrasts with the diversity of sequences found at the sites of signal peptide cleavage. The enzyme was shown to have a strong preference for small uncharged residues at positions −1 and −3 with respect to the cleaved peptide bond [1,5,6]. Statistical analysis of known secretory and membrane proteins revealed an additional but weaker bias for certain groups of amino acids at adjacent positions and a preferred proximity of the cleavage site to the hydrophobic domain of the signal sequence [1,7]. Based on these statistical data, an algorithm has been developed by Von Heijne [7] to calculate the most probable processing sites in a protein with a success rate of 75–80%. Whether or not a sequence is cleaved, however, cannot be predicted by this method. Within the sequences of membrane proteins, for example, sites with 'high processing probability' are found after almost every membrane-spanning domain. Transmembrane segments which have their carboxy-terminal end exposed to the ER lumen are potential substrates for signal peptidase. In two cases of uncleaved, internal signal sequences of type II membrane proteins (amino-terminus cytoplasmic, carboxy-terminus exoplasmic) [8] it has been shown that cryptic cleavage sites could be 'activated' by changing the cytoplasmic flanking sequences [9,10]. What features contribute to the

Correspondence address: M. Spiess, Department of Biochemistry, Biocenter, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

specificity of signal peptidase, however, is not completely understood.

Signal peptidase was found to be associated in a complex of several proteins suspected of being directly or indirectly related to protein translocation [11,12]. A close association with the translocation machinery is also consistent with the notion that signal cleavage occurs co-translationally [13]. In all natural and artificial proteins analyzed, cleavage is restricted to the first hydrophobic domain inserted into the membrane. A simple explanation could be that the enzyme has no access to subsequently inserted membrane-spanning domains. Here, we have tested the ability of signal peptidase to cleave also inside a polytopic membrane protein by inserting a normally cleaved signal sequence in place of the third potential transmembrane domain.

2. MATERIALS AND METHODS

2.1. *Materials*

SP6 RNA polymerase and endo-β-*N*-acetylglucosaminidase H (Endo H) were purchased from Boehringer, Mannheim. Other modifying and restriction enzymes and linkers were purchased from Boehringer and New England Biolabs. Proteinase K was purchased from Sigma; RNase inhibitor from Promega Biotech; and protein A-Sepharose from Pharmacia.

2.2. *Plasmid construction*

Standard techniques were used for DNA manipulations [14]. The plasmid pSAAA, which has been described earlier [15], was linearized with *Cla*I upstream of the sequence encoding the third transmembrane segment, digested with exonuclease *Bal*31, and provided with *Cla*I linkers (CATCGATG). In one of the resulting plasmids, pB5, 155 bp have been deleted upstream of the original *Cla*I site. The plasmid pSB6 corresponds to the wild-type cDNA of ASGP receptor H1 (pSA1) [16] in which nucleotides 267–283 had been deleted by *Bal*31 digestion and replaced by a *Cla*I linker. The sequence upstream of the *Cla*I site of pB5 was ligated to the sequence downstream of the *Cla*I site of pSB6 to yield pSAAA50. In summary, the recombinant DNA encodes the following protein sequence (amino acids generated by linker sequences are indicated in single-letter code): residues 1–156 of H1, RC, 38–104 of H1, PSM, 38–291 of H1.

To construct pSAAH50, the plasmid pSNHC (containing the cDNA of the ASGP receptor H1 in which the transmembrane segment had been replaced by the signal sequence of influenza virus hemagglutinin [10]) was linearized at the *Nru*I site preceding the hemagglutinin segment, and a *Sal*I(*Acc*I) linker (CCGGTCGACCGG) linker was inserted. The sequence downstream of the new *Acc*I site of this plasmid was ligated to the sequence upstream of the *Cla*I site of pB5. The resulting plasmid pSAAH50 encodes the following protein sequence: residues 1–156 of H1, RC, 38–104 of H1, PSTGD, 2–18 of hemagglutinin, R, 60–291 of H1.

2.3. *In vitro transcription and translation*

For in vitro transcription, 2 μg of plasmid DNA linearized with *Eco*RI were incubated in 40 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 0.5 mM each of ATP, CTP, GTP, UTP, and the cap analogue ^{7m} GpppG, 10 mM dithiothreitol, 50 U RNase inhibitor, and 20 U of SP6 RNA polymerase, in a total volume of 50 μl for 1 h at 40°C. Nucleic acids (containing approx. 2–3 μg of RNA) were then precipitated with ethanol and redissolved in 25 μl H_2O.

Translation was performed using rabbit reticulocyte lysate as described by Pelham and Jackson [17] for 1 h at 30°C using 1 μl (approx. 100 ng) of RNA in a standard 25 μl reaction, with or without 1 μl of dog pancreas microsomes (isolated and column-washed according to Walter and Blobel [18]). Protease protection was assayed by incubating the samples on ice for 20 min with 50 μg/ml proteinase K in the absence or presence of 0.5% Nonidet P-40. Digestion was terminated by addition of 100 μl of phosphate-buffered saline (PBS, 12.5 mM Na-phosphate, pH 7.5, 125 mM NaCl) containing 2 mM phenylmethylsulfonyl fluoride (PMSF). After addition of 500 μl PBS containing 1% Triton X-100, 0.5% deoxycholate, 0.5% SDS, 1 mM PMSF, 1 mg/ml bovine serum albumin, the samples were split in half and immunoprecipitated using antiserum raised against isolated human ASGP receptor and antiserum against a synthetic peptide, GCETELDKASQE, which (except for the amino-terminal glycine) corresponds to residues 277–287 at the carboxy-terminus of the ASGP receptor H1. The immune complexes were isolated with protein A-Sepharose (15 μl per sample) and released by boiling for 2 min in 100 μl of 50 mM Na-citrate, pH 6, 1% SDS. Some of the samples were deglycosylated by incubation with 3 mU of Endo H for 1 h at 37°C. Finally, all samples were boiled in SDS-sample buffer and analyzed by 15% SDS-polyacrylamide gel electrophoresis [19]. The gels were fixed, soaked in 1 M Na-salicylate containing 1% glycerol, dried and fluorographed on Kodak XAR-5 film for 1–2 days at −70°C.

3. RESULTS

The human asialoglycoprotein (ASGP) receptor H1 is a membrane protein containing a single hydrophobic segment of approx. 20 apolar amino acids that functions both as an internal, uncleaved signal sequence and as a transmembrane anchor of the mature protein [16,20,21] (schematically shown in fig.1A, pSA1). Previously we have described the construction of an artificial polytopic membrane protein on the cDNA level by repeating the signal-anchor domain as well as the adjacent sites for *N*-linked glycosylation three times [15]. Upon in vitro translation in the presence of microsomes, this protein was inserted sequentially into the target membranes: the first hydrophobic domain functioned as a signal and the second as a stop-transfer sequence, while the third initiated a second translocation process [15]

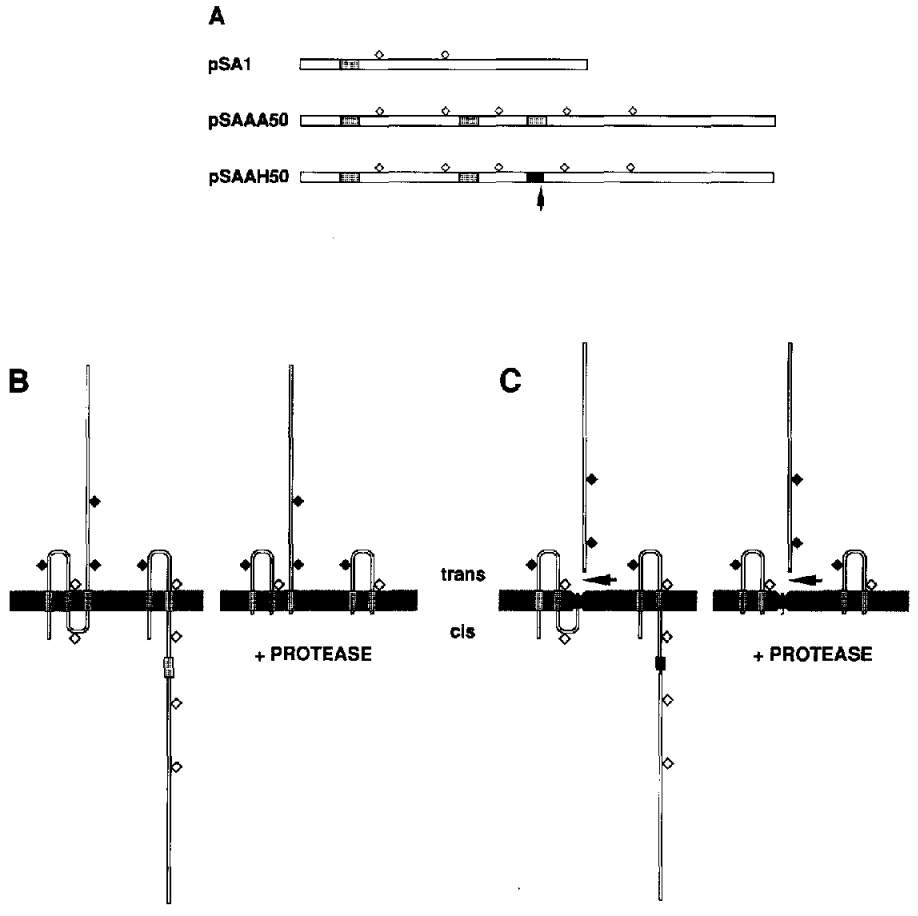

Fig.1. Summary of the membrane protein constructs. (A) The polypeptides encoded by the cDNA constructs are shown schematically. Sequences derived from the ASGP receptor H1 are shown as open bars with the hydrophobic domains stippled and the potential glycosylation sites indicated as diamonds. The sequence derived from influenza virus hemagglutinin is shown in black with the hydrophobic segment striped. An arrow points at the potential cleavage site. (B and C) Schematic representation of the expected topologies of the proteins encoded by the plasmids pSAAA50 (B) and pSAAH50 (C) in the microsomal membrane before and after digestion with exogenous protease. Potential sites for *N*-linked glycosylation are shown as diamonds; the ones that are actually used are shown in black. Signal peptidase cleavage is indicated by arrows. (Whether the internalized signal of hemagglutinin remains associated with the membrane after cleavage is not known.)

(illustrated in fig.1B). The topology of the translation products could be deduced from their glycosylation patterns and from the sizes of the peptide fragments resistant to exogenous protease.

To test the ability of signal peptidase to cleave inside a polytopic membrane protein, we exchanged the third copy of the signal-anchor sequence by the normally amino-terminal, cleavable signal sequence of influenza virus hemagglutinin (fig.1A). If signal cleavage occurs in this internal position, the polypeptide will be cut into two fragments upon membrane insertion (fig.1C). Because the fragments expected by cleavage at the third transmembrane segment would be very similar in size, a variant of the original threefold membrane-spanning construct pSAAA [15] was used for this study. In pSAAA50 the hydrophilic segment between the second and the third hydrophobic domain is shortened to 50 residues. Replacement of the third signal-anchor domain the hemagglutinin signal sequence resulted in pSAAH50.

Translation of mRNA transcribed from the plasmid pSAAA50 is analyzed in fig.2. The translation products were immunoprecipitated

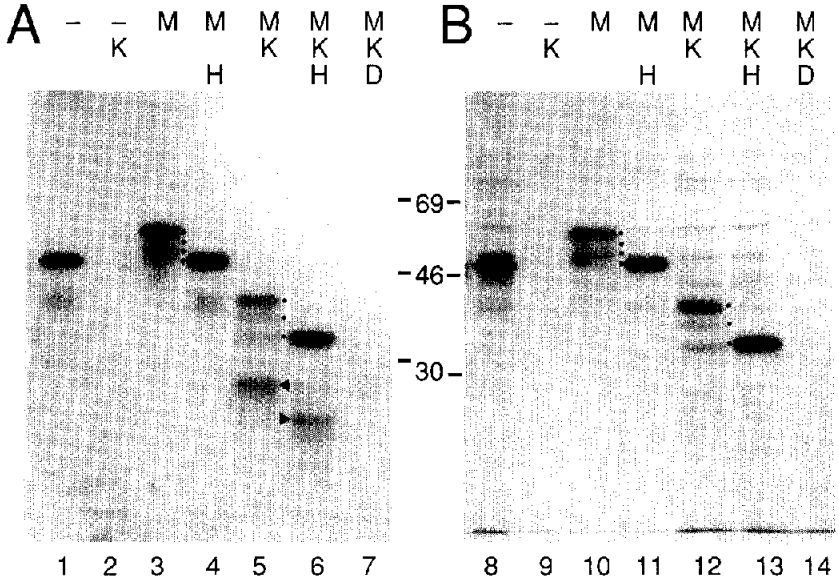

Fig.2. In vitro translation and membrane insertion of pSAAA50. RNA derived from the plasmid pSAAA50 was translated in a reticulocyte lysate system in the absence (–) or presence (M) of microsomes. The products were immunoprecipitated either with a polyclonal antiserum raised against the isolated human ASGP receptor (A) or with an antiserum raised against an oligopeptide corresponding to the carboxy-terminus of the protein (B). Fluorographs after SDS-polyacrylamide gel electrophoresis are shown. *Abbreviations:* M, microsomes present during translation; K, posttranslational incubation with proteinase K; D, detergent NP-40 present during incubation with proteinase K; H, Endo H digestion of the final products. The position of marker proteins and their molecular masses in kDa are indicated. The different glycosylated products are indicated by dots (carboxy-terminal polypeptides) and triangles.

either by an antiserum raised against the isolated ASGP receptor (panel A), or by an antiserum directed against a synthetic peptide corresponding to the carboxy-terminus of the protein (panel B). In the absence of membranes a polypeptide of approx. 50 kDa was synthesized (fig.2, lanes 1 and 8), which was completely sensitive to proteinase K (lanes 2 and 9). If translation was performed in the presence of dog pancreas microsomes, the majority of the products were glycosylated (lanes 3 and 10). It has been shown previously that each oligosaccharide transferred to the polypeptide results in an increase of the apparent molecular mass of approx. 3 kDa [15], which is reversed upon deglycosylation with endo-β-*N*-acetylglucosaminidase H (Endo H; lanes 4 and 11). Most of the products were threefold glycosylated, indicative of the threefold membrane-spanning topology illustrated in fig.1B: one oligosaccharide was transferred to the translocated 'loop' segment (the second potential glycosylation site is located too close to the membrane surface to be accessible to the oligosaccharide transferase [15]), and two to the translocated 'tail' segment. Only a small fraction of the polypeptides had not performed the second translocation step and was thus glycosylated only once.

The threefold membrane-spanning structure of the protein is confirmed by the size and glycosylation of the fragments resistant to posttranslational protease digestion. The larger twice-glycosylated one of approx. 38 kDa corresponds to the tail because it was precipitated by both the anti-receptor antiserum and the antiserum specific for the carboxy-terminal oligopeptide. The smaller fragment was glycosylated once, was recognized only by the anti-receptor antiserum, and thus corresponds to the loop (lanes 5, 6, 12 and 13). Disruption of the membranes by detergent resulted in complete digestion of the translation products by proteinase K (lanes 7 and 14).

Translation of mRNA transcribed from pSAAH50, in which the signal sequence of hemagglutinin serves as the third hydrophobic domain, also produced a protein of 50 kDa (fig.3, lanes 1 and 8). In the presence of microsomes, however,

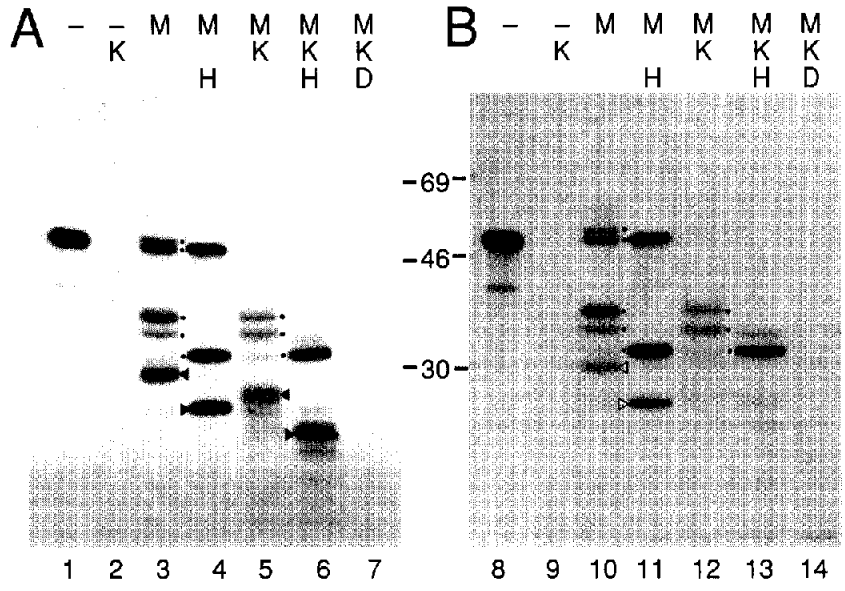

Fig.3. In vitro translation and membrane insertion of pSAAH50 containing an internalized hemagglutinin signal sequence. RNA derived from the plasmid pSAAH50 was translated and the products were analyzed as described in the legend to fig.2, using an antiserum against total ASGP receptor (A) and an antiserum specific for the carboxy-terminus of the protein (B). The abbreviations are the same as in fig.2. Open triangles point at the small amount of amino-terminal fragments that are coprecipitated with the carboxy-terminal fragments in lanes 10 and 11.

the only full-size products synthesized are either not glycosylated at all and are therefore most likely not inserted into the microsomes, or carry only one oligosaccharide and correspond to polypeptides that have performed only the first translocation step (see fig.1C). No threefold glycosylated protein was detectable (fig.3, lanes 3 and 10). Instead two fragments were generated. The larger one of approx. 38 kDa was glycosylated twice (lanes 3 and 4, dots), completely resistant to exogenous protease (lanes 5 and 6), and corresponds to translocated tail fragments, since it is efficiently immunoprecipitated by the antiserum against the carboxy-terminal oligopeptide (fig.3B). The smaller fragment was glycosylated once (lanes 3 and 4, triangles) and is only partially resistant to exogenous protease (lanes 5 and 6). It thus corresponds to the amino-terminal half of the protein that is partly exposed on the outside of the microsomes. Surprisingly, approx. 40% of this fragment (as determined by densitometric scanning of the fluorograph) is co-precipitated by the antiserum against the carboxy-terminal peptide (lanes 10 and 11, open triangles), suggesting that some of the cleavage products remained noncovalently associated with one another. This interaction might be similar to that responsible for the oligomerization of the wild-type ASGP receptor H1 to dimers and trimers [22]. Upon protease digestion of the microsomes (lanes 12 and 13), no association of the two fragments is detectable.

4. DISCUSSION

Previously it has been shown that signal peptides do not have to be exposed at the very amino-terminus to be a substrate of signal peptidase: normal amino-terminal signal sequences in an internal position, preceded by up to more than 100 amino acids, were still found to be efficiently cleaved [10,23,24]. However, the hemagglutinin signal sequence was not cleaved when, as part of a translocating polypeptide, it was transported through the membrane with the inverted orientation (amino to carboxy-terminus) with respect to that of cleaved signal sequences [25].

In this work, we have demonstrated that signal cleavage is not restricted to the first insertion sequence in a protein, but that signal peptidase can also act on a cleavable sequence inside a polytopic

membrane protein provided it is located on the luminal side of a membrane-spanning domain. The normally cleaved signal of influenza virus hemagglutinin inserted downstream of a first (uncleaved) signal and a stop-transfer sequence (two copies of the signal-anchor sequence of the ASGP receptor H1) is efficiently cleaved as it functions as a secondary insertion sequence. According to all available criteria, this processing is due to the ER signal peptidase: it does not occur in the absence of microsomes; it is restricted to polypeptides that are inserted into the microsomes and that have performed the second translocation process; and it is specific for the protein containing the hemagglutinin signal sequence.

The length of the hydrophilic segment between the second hydrophobic domain (the stop-transfer sequence) and the third (the second insertion signal) in our constructs is 50 residues well within the range found in natural multi-spanning proteins. Based on their predicted membrane topology, naturally occurring cytoplasmic loops contain from less than 5 to more than 100 amino acids. It is therefore likely that our findings also apply to natural polytopic membrane proteins. Signal peptidase is thus not specific for the first insertion sequence in a membrane protein.

According to our present knowledge the high specificity of signal peptidase is based on the primary sequence near the actual cleavage site, whereby the consensus for this sequence is very limited. Furthermore, there is evidence that the subtle positioning of a potentially cleavable sequence close to the exoplasmic surface of the ER membrane is a structural requirement for processing: deletion of the amino-terminal domains preceding the normally uncleaved, internal signals of the invariant chain of the HLA-DR histocompatibility antigen [9] and of the ASGP receptor H1 [10] resulted in cleavage at cryptic sites that are not accessible or not recognized in the wild-type proteins. If these are indeed the only determinants of signal peptidase activity, an evolutionary pressure must exist against the formation of cleavable sites on the exoplasmic side of transmembrane segments in membrane proteins.

Acknowledgements: We want to thank C. Handschin for her excellent technical assistance and F. Theiler for the kind gift of reticulocyte lysate. This study was supported by grant 3.366-0.86 from the Swiss National Science Foundation.

REFERENCES

[1] Von Heijne, G. (1983) Eur. J. Biochem. 133, 17–21.
[2] Watson, M.E.E. (1984) Nucleic Acids Res. 12, 5145–5164.
[3] Walter, P., Gilmore, R. and Blobel, G. (1984) Cell 38, 5–8.
[4] Wiedmann, M., Kurzchalia, T.V., Hartmann, E. and Rapoport, T.A. (1987) Nature 328, 830–833.
[5] Kuhn, A. and Wickner, W. (1985) J. Biol. Chem. 260, 15907–15913.
[6] Folz, R.J., Nothwehr, S.F. and Gordon, J.I. (1988) J. Biol. Chem. 263, 2070–2078.
[7] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[8] Garoff, H. (1985) Annu. Rev. Cell Biol. 1, 403–445.
[9] Lipp, J. and Dobberstein, B. (1986) Cell 46, 1103–1112.
[10] Schmid, S.R. and Spiess, M. (1988) J. Biol. Chem. 263, 16886–16891.
[11] Baker, R.K., Bentivoglio, G.P. and Lively, M.O. (1986) J. Cell. Biochem. 32, 193–200.
[12] Evans, E.A., Gilmore, R. and Blobel, G. (1986) Proc. Natl. Acad. Sci. USA 83, 581–585.
[13] Blobel, G. and Dobberstein, B. (1975) J. Cell Biol. 67, 835–851.
[14] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[15] Wessels, H.P. and Spiess, M. (1988) Cell 55, 61–70.
[16] Spiess, M. and Lodish, H.F. (1986) Cell 44, 177–185.
[17] Pelham, H.R.B. and Jackson, R.J. (1976) Eur. J. Biochem. 67, 247–256.
[18] Walter, P. and Blobel, G. (1983) Methods Enzymol. 96, 84–93.
[19] Laemmli, U.K. (1970) Nature 227, 680–685.
[20] Spiess, M., Schwartz, A.L. and Lodish, H.F. (1985) J. Biol. Chem. 260, 1979–1982.
[21] Spiess, M. and Handschin, C. (1987) EMBO J. 6, 2683–2691.
[22] Shia, M.A. and Lodish, H.F. (1989) Proc. Natl. Acad. Sci. USA 86, 1158–1162.
[23] Perara, E. and Lingappa, V.R. (1985) J. Cell Biol. 101, 2292–2301.
[24] Rottier, P.J.M., Florkiewicz, R.Z., Shaw, A.S. and Rose, J.K. (1987) J. Biol. Chem. 8889–8895.
[25] Finidori, J., Rizzolo, L., Gonzales, A., Kreibich, G., Adesnik, M. and Sabatini, D.D. (1987) J. Cell Biol. 104, 1705–1714.