Role of the hydrophobic region of signal sequences in the targeting of proteins to membranes and translocation across the hydrophobic membrane barrier

R NAGARAJ

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India

Abstract. Proteins destined for regions other than the cytoplasm in cells have to cross at least one membrane barrier before reaching their proper destination. Almost all such proteins are initially biosynthesized as precursors with signal sequences at the amino terminus. Signal sequences are essential and also sufficient for proteins to be targeted to membranes and also for translocation across membranes. One striking feature that is clearly evident amongst signal sequences of secretory proteins is a positively charged amino terminus followed by a region comprising 10–12 very hydrophobic amino acids. The structural and physico-chemical properties of signal sequences have been analysed. On the basis of the analyses it is proposed that the structural feature of a positively charged amino terminal region followed by a hydrophobic stretch of amino acids, rather than a conformational one, is recognised by components of the cells export machinery. It is also postulated that signal sequences insert in the lipid bilayer of the translocation competent membrane after targeting. The presence of the signal sequence results in the formation of local 'defects' in the bilayer which have a role in translocation of proteins across membranes.

Keywords. Signal sequence; hydrophobicity; targeting to membranes; translocation across membranes; signal sequence receptors; signal-sequence membrane interactions.

1. Introduction

Proteins destined for regions other than the cytoplasm in cells have to cross at least one membrane barrier before reaching their proper destination. Almost all such proteins are initially biosynthesized as precursors with signal sequences at the amino terminus (Kreil 1981; Michaelis and Beckwith 1982; Hay *et al* 1984). Peptidases present in the endoplasmic reticulum, the inner membrane in *E. coli* (Wolfe *et al* 1983) and in the matrix space in mitochondria (Boehni *et al* 1980) cleave off the signal sequences from the precursor proteins. Gene fusion and *in vitro* reconstitution experiments indicate that signal sequences are essential and also sufficient for proteins to be targeted to the appropriate membrane site (endoplasmic reticulum and mitochondria in eukaryotes and inner membrane in prokaryotes) and also translocation across these membranes (Lingappa *et al* 1984; Horwich *et al* 1984, Van Loon *et al* 1986; Perara and Lingappa 1985).

The primary structures of a large number of signal sequences have been determined and compiled recently (Watson 1984). Interestingly, there is no primary structure homology amongst signal sequences. However, one striking feature is clearly evident amongst signal sequences of secretory proteins. All of them have a positively charged amino terminus region followed by a region comprising 10–12 very hydrophobic amino acids like Leu, Ile, Val, Phe. There is
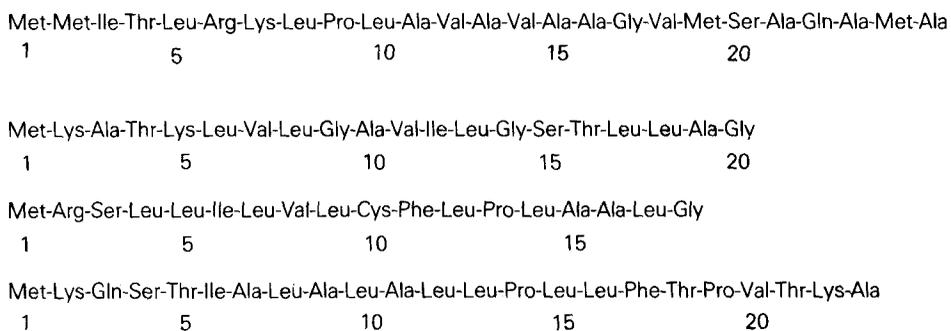
considerable evidence that the hydrophobic region is essential for signal sequences to be able to initiate export of proteins efficiently (Benson and Silhavy 1983; Lee and Beckwith 1986). This article analyses the structural and physico-chemical properties of signal sequences destined for secretion and attempts to answer the two following questions:

(1) Is it the hydrophobic region in signal sequences that is recognized by components of the cells export machinery?

(2) What is the role of the hydrophobic region in the ability of signal sequences to initiate translocation of proteins across membranes?

2. The hydrophobic nature of signal sequences

The primary structures of some representative signal sequences and a schematic sketch of the distribution of amino acids is shown in figure 1. All signal sequences (Watson 1984) have one or two positively charged amino acids in the amino terminal region, approximately 5-residue long. This is followed by a stretch of highly hydrophobic amino acids like Leu, Ile, Val or Phe. The carboxy terminal region amino acids have short side chains like Gly, Ala, Ser or Cys. Signal sequences have been analysed with respect to the mean hydrophobicity at each position in the peptide segment (von Heijne 1982). The analysis reveals a non-uniform hydrophobic distribution with maximum hydrophobicity at the mid-point of the peptide segment. This distribution was not observed in the hydrophobic regions of intrinsic membrane proteins.

The total hydrophobic contribution to the free energy of interaction between signal sequences and the hydrophobic interior of membranes have been computed (von Heijne 1981, 1985; Engelman and Steitz 1981). A free energy of ~120 kJ/mol was obtained indicating that the signal sequences would spontaneously partition into the hydrophobic interior of membranes.

Met-Met-Ile-Thr-Leu-Arg-Lys-Leu-Pro-Leu-Ala-Val-Ala-Val-Ala-Ala-Gly-Val-Met-Ser-Ala-Gln-Ala-Met-Ala
1   5   10   15   20

Met-Lys-Ala-Thr-Lys-Leu-Val-Leu-Gly-Ala-Val-Ile-Leu-Gly-Ser-Thr-Leu-Leu-Ala-Gly
1   5   10   15   20

Met-Arg-Ser-Leu-Leu-Ile-Leu-Val-Leu-Cys-Phe-Leu-Pro-Leu-Ala-Ala-Leu-Gly
1   5   10   15

Met-Lys-Gln-Ser-Thr-Ile-Ala-Leu-Ala-Leu-Ala-Leu-Leu-Pro-Leu-Leu-Phe-Thr-Pro-Val-Thr-Lys-Ala
1   5   10   15   20

Figure 1. Primary structures of some representative signal sequences. The areas A, B and C in the line diagram correspond to the positively charged amino terminal (~5 amino acids), hydrophobic core segment (~10–15 amino acids) and carboxy terminal (~6 amino acids) regions. The sequences have been taken from Watson (1984).

3. Conformations of signal sequences

The conformations of several signal sequences have been analysed by the method of Chou and Fasman (Austen 1979). However, a common conformational feature does not emerge from this analysis. The conformations of some synthetic signal peptides have been studied by circular dichroism (CD) spectroscopy (Rosenblatt *et al* 1980; Briggs and Gierasch 1984; Shinnar and Kaiser 1984; Katakai and Iizuka 1984; Laxma Reddy and Nagaraj 1985, 1986). In aqueous or polar environments, the peptides studied had very little ordered conformation. In hydrophobic environments the peptides adopted ordered conformations. Exclusively α-helical conformations were observed for 3 signal sequences (Rosenblatt *et al* 1980; Shinnar and Kaiser 1984; Laxma Reddy and Nagaraj 1985). One signal sequence showed both α-helical and β-sheet conformations (Laxma Reddy and Nagaraj 1986). The conformations of shorter fragments of signal sequences have also been studied (Briggs and Gierasch 1984; Katakai and Iizuka 1984; Laxma Reddy and Nagaraj 1985). The amino terminal regions were found to be unordered whereas the central hydrophobic portion adopted α-helical conformations.

Recently the conformations of a signal peptide induced by lipids have been studied (Briggs *et al* 1986). When the signal peptide was bound only to the charged head groups of phospholipids, a β-sheet conformation was observed. Under experimental conditions in which the peptide was associated with the hydrocarbon region of phospholipids, an α-helical conformation was discernible.

4. Features in signal sequences essential for proper sorting of proteins

Extensive genetic studies in *E. coli* (Benson and Silhavy 1983; Lee and Beckwith 1986) have indicated how mutations resulting in change or deletion of amino acids in the amino terminal and the hydrophobic regions of signal sequences affect their ability to initiate the export of proteins. Figure 2 shows the effect of changes in the amino terminus of the signal sequence of *E. coli* lipoprotein (Inouye *et al* 1982;

Wild type	Charge	
Met-Lys-Ala-Thr-Lys-Leu	+ 2	
Met-*Lys*-Ala-Thr-*Asn*-Leu	+ 1	
Met-Ala-Thr-*Asn*-Leu	0	
Met-*Glu-Asp*-Thr-*Asn*-Leu	− 2	non functional
Met-Lys-*Asp*-Thr-Lys-Leu	+ 1	
Met-*Ala*-Thr-Lys-Leu	+ 1	
Met-*Asp*-Thr-Lys	0	non functional
Met-*Glu-Asp*-Thr-Lys	− 1	

Figure 2. Amino terminal region of the wild type and mutant signal sequences of *E. coli* lipoprotein (Inouye *et al* 1982; Vlasuk *et al* 1983). The amino acids replaced by mutation are italicized.

Vlasuk *et al* 1983). When the amino terminus is positively charged or even devoid of any charge, the mutant signal sequences are able to initiate export just as the 'wild-type' sequence. However, when the charge is *negative* the mutant signal sequences are no longer able to initiate the export of proteins. It is thus clear that a positively charged amino terminus plays an important role in the function of a signal sequence. In figure 3 the effect of changes and deletions of amino acids in the hydrophobic region on the ability of signal sequences to initiate the export of

E. coli alkaline phosphatase:

Met-Lys-Gln-Ser-Thr-Ile-Ala-Leu-Ala-Leu-Leu-Pro-Leu-Leu-Phe-Thr-Pro-Val-Thr-Lys-Ala
1   5   10   15   20
                          Glu
                              His

E. coli Maltose binding protein:

Met-Lys-Ile-Lys-Thr-Gly-Ala-Arg-Ile-Leu-Ala-Leu-Ser-Ala-Leu-Thr-Thr-Met-Met-Phe-Ser-Ala
1   5   10   15   20
                       Glu   Lys

E. coli Lipoprotein:

Met-Lys-Ala-Thr-Lys-Leu-Val-Leu-Gly-Ala-Val-Ile-Leu-Gly-Ser-Thr-Leu-Leu-Ala-Gly
1   5   10   15   20
                     Asp

E. coli λ-receptor:

Met-Met-Ile-Thr-Leu-Arg-Lys-Leu-Pro-Leu-Ala-Val-Ala-Val-Ala-Ala-Gly-Val-Met-Ser-Ala-Gln-Ala-Met-Ala
1   5   10   15   20   25
                  Asp Gln Arg/Ly

deletion mutant

Met-Met-Ile-Thr-Leu-Arg-Lys-Leu-Leu-Pro-1   1-Val-Ala-Ala-Gly-Val-Met-Ser-Ala-Gln-Ala-Met-Aln
1   5   15   20   25

R1 Met-Met-Ile-Thr-Leu-Arg-Lys-Leu-Pro-1   1-Val-Ala-Ala-Cys-Val-Met-Ser-Ala-Gln-Ala-Met-Ala
1   5   15   20   25

R2 Met-Met-Ile-Thr-Leu-Arg-Lys-Leu-Leu-1   1-Val-Ala-Ala-Gly-Val-Met-Ser-Aln-Gln-Ala-Met-Ala
1   5   15   20   25

Figure 3. Wild-type and mutant signal sequences of some *E. coli* proteins. Changes as a result of mutation are indicated by arrows (↓) indicates that the mutational change results in a defective signal sequence that cannot initiate export of proteins. (↕) indicates that the mutational change does *not* result in a defective signal sequence. R1 and R2 are pseudo revertant sequences. The alkaline phosphatase wild type and mutant signal sequences are from Michaelis *et al* (1986). The wild type and mutant maltose binding protein, lipoprotein and λ-receptor sequences are from Benson and Silhavy (1983). The deletion mutant and pseudo revertant signal sequences of λ-receptor are from Emr and Silhavy (1983).

proteins is depicted. Mutant signal sequences in which a charged residue is present in the hydrophobic region resulting in the *reduction* of the length of the hydrophobic stretch are non-functional. In pseudo-revertant signal sequences, the reduction of length of the hydrophobic stretch in mutants is removed, resulting in functional signal sequences. Deletions of hydrophobic amino acids also render signal sequences non-functional. In the case of the *E. coli* λ-receptor 'wild-type' mutant and pseudo-revertant signal sequences (Emr and Silhavy 1983) the observations have been rationalized in terms of the conformation of these sequences. Peptides corresponding to these sequences were synthesized and their secondary structures analysed by circular dichroism spectroscopy (Briggs and Gierasch 1984). An α-helical conformation was observed only in the 'wild-type' and pseudo-revertant sequences, whereas the deletion mutant sequence was largely unordered. Hence it was concluded that an α-helical conformation in the central hydrophobic region is essential for the function of signal sequences. It is likely that a positively charged amino terminus and a 'critical size' of the hydrophobic region is essential for eukaryotic signals as well. In fact it has been demonstrated that replacement of Leu by its polar analog β-hydroxy leucine renders a eukaryotic signal sequence non-functional (Hortin and Boime 1980).

5. Targeting of proteins destined for secretion to the membrane site

Extensive biochemical studies particularly *in vitro* reconstitution experiments in eukaryotes have shown how ribosomes synthesizing secretory proteins are targeted to the endoplasmic reticulum (Walter *et al* 1984; Walter and Lingappa 1986). A protein-RNA complex, the signal recognition particle (SRP), recognizes and binds to the signal sequence of the secretory protein as it emerges from the ribosome. This binding results in the arrest of translation of the secretory protein from its m-RNA. The SRP-nascent secretory protein-ribosome complex then moves to the endoplasmic reticulum where SRP binds to its 'receptor', the docking protein. Thus binding results in the release of translation arrest. Although molecules like SRP and docking protein have not been characterized in prokaryotes, several genetic loci that may specify components of the cellular export machinery have been identified (Benson *et al* 1985). While it is not clear whether regions other than the signal sequence are recognized by SRP in precursor proteins, the presence of the signal sequence is essential for SRP induced translation arrest. Removal of the signal sequence of pre-proinsulin by recombinant DNA techniques abolished interaction between the polypeptide and SRP (Weidmann *et al* 1986) indicating the absolute requirement of signal sequence for recognition by SRP.

6. Translocation of proteins across membranes

Once targeting to the membrane site is accomplished the next step is translocation of proteins across the hydrophobic membrane barrier. Widely differing views as to how this happens exist. According to one school (Inouye and Halegoua 1979; Engelman and Steitz 1981; von Heijne 1985) translocation is initiated by spontaneous partitioning of the signal sequence into the lipid bilayer of membranes. This is followed by the rest of the polypeptide chain. There is no direct

experimental evidence supporting this view. Recently *E. coli* precursor proteins have been shown to translocate across membrane vesicles comprising only phospholipids (Geller and Wickner 1985). It would be of interest to examine whether precursor secretory proteins of eukaryotes can be translocated across a model membrane system devoid of proteins.

Another school favours translocation through a protein channel in the membrane (Walter *et al* 1984; Singer *et al* 1987). Translocation is initiated by binding of the signal sequence to the protein which results in the opening of an aqueous protein channel in the membrane. However no such protein or proteins have been identified so far.

7. The role of the hydrophobic region in signal sequences in targeting of proteins to the translocation competent membrane and in transfer of proteins across such membranes

A positively charged amino terminus followed by a stretch of at least 7–8 hydrophobic amino acids is absolutely essential for a signal sequence to be able to initiate the export of proteins. Any change in these regions which alters the charge or reduces the length of the hydrophobic stretch results in defective signal sequences. The conformations that signal sequences adopt are clearly environment dependent and both α-helical and β-sheet conformations are observed. Hence, I feel that it is the structural feature of positively charged amino terminus followed by a hydrophobic region that is recognized by components of the cells export machinery. An α-helical conformation may not be as stringent a requirement as the distribution of charged and hydrophobic amino acids for recognition.

The role of signal sequences in initiating translocation of proteins across membranes differs widely in the models that have been proposed to explain this step. In the 'direct transfer' models, translocation is initiated by spontaneous partitioning of signal sequences in the hydrophobic interior of membranes, whereas in models where translocation is postulated to occur through protein channels, the signal sequences interact with a protein receptor in the membrane. In any event, it is the initial interaction of signal sequences with membranes that opens up a pathway for translocation, as these sequences are cleaved off precursor proteins by membrane-bound signal peptidases before the translocation of the entire protein across membranes is completed. We (Nagaraj 1984; Nagaraj *et al* 1987) and some others (Briggs *et al* 1985, 1986) have studied the interaction of synthetic signal peptides with model membranes and have observed that these peptides partition spontaneously into the lipid bilayer. We have also demonstrated that the association of signal sequences with model membranes results in extensive perturbation of the lipid bilayer (Nagaraj *et al* 1987). Taking these results into account and also the observation that some precursor proteins can translocate across the lipid-bilayer of membrane vesicles devoid of proteins, I propose that *in vivo*, signal sequences insert into the lipid bilayer of the translocation competent membrane after targeting to the membrane site has occurred. The presence of the signal sequence results in the formation of 'local defects' in the lipid bilayer. Translocation of the mature portion of the protein can then conceivably occur through 'defects' in the lipid bilayer. Alternatively, the perturbation of the lipid

bilayer could result in the rearrangement of membrane proteins so as to open up aqueous channels for translocation.

References

Austen B M 1979 *FEBS Lett.* **103** 308

Benson S A and Silhavy T J 1983 *Gene function in prokaryotes* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory) p. 253

Benson S A, Hall M N and Silhavy T J 1985 *Annu. Rev. Microbiol.* **54** 101

Boehni P, Gasser S, Leaver C and Schatz G 1980 *The organization and expression of the mitochondrial genome* (eds) A M Kroon and C Saccone p. 423 (Amsterdam: Elsevier)

Briggs S M and Gierasch L M 1984 *Biochemistry* **23** 3111

Briggs S M, Gierasch L M, Zlotnick A, Lear J D and De Grado W F 1985 *Science* **228** 1096

Briggs S M, Cornell D G, Dluhy R A and Gierasch L M 1986 *Science* **233** 206

Emr S D and Silhavy T J 1983 *Proc. Natl. Acad. Sci. USA* **80** 4599

Engelman D M and Steitz T A 1981 *Cell* **23** 411

Evans E A, Gilmore R and Blobel G 1986 *Proc. Natl. Acad. Sci. USA* **83** 581

Geller B L and Wickner W 1985 *J. Biol. Chem.* **260** 13281

Hay R, Boehni P and Gasser S 1984 *Biochim. Biophys. Acta* **65** 779

Hortin G and Boime I 1980 *Proc. Natl. Acad. Sci. USA* **77** 1356

Horwich A L, Kalousek F, Mellman I and Rosenberg L 1985 *EMBO J.* **4** 1129

Inouye M and Halegoua S 1979 *Crit. Rev. Biochem.* **7** 339

Inouye S, Soberon X, Franceschini T, Nakamura K, Itakura K and Inouye M 1982 *Proc. Natl. Acad. Sci. USA* **258** 7141

Katakai R and Izuka Y 1984 *J. Am. Chem. Soc.* **106** 5715

Kreil G 1981 *Annu. Rev. Biochem.* **50** 317

Laxma Reddy G and Nagaraj R 1985 *Biochim. Biophys. Acta* **831** 340

Laxma Reddy G and Nagaraj R 1986 *FEBS Lett.* **202** 349

Lee C and Beckwith J 1986 *Annu. Rev. Cell Biol.* **2** 315

Lingappa V R, Chaidez J, Yost C S and Hedgpeth J 1984 *Proc. Natl. Acad. Sci. USA* **81** 456

Michaelis S and Beckwith J 1982 *Annu. Rev. Microbiol.* **36** 345

Michaelis S, Hunt J F and Beckwith J 1986 *J. Bacteriol.* **167** 160

Nagaraj R 1984 *FEBS Lett.* **165** 79

Nagaraj R, Joseph M and Laxma Reddy G 1987 *Biochim. Biophys. Acta* (submitted)

Perara E and Lingappa V R 1985 *J. Cell Biol.* **101** 2292

Rosenblatt M, Beaudette N V and Fasman G D 1980 *Proc. Natl. Acad. Sci. USA* **77** 3983

Shinnar A E and Kaiser E T 1984 *J. Am. Chem. Soc.* **106** 5006

Singer S J, Maher P A and Yaffe M P 1987 *Proc. Natl. Acad. Sci. USA* **84** 1015

Van Loon A P E M, Brandl A W and Schatz G 1986 *Cell* **44** 501

Vlasuk G P, Inouye S, Ito H, Itakura K and Inouye M 1983 *J. Biol. Chem.* **258** 7141

von Heijne G 1981 *Eur. J. Biochem.* **116** 419

von Heijne G 1982 *J. Mol. Biol.* **159** 537

von Heijne G 1985 *Curr. Top. Membr. Transp.* **24** 151

Walter P, Gilmore R and Blobel G 1984 *Cell* **38** 5

Walter P and Lingappa V R 1986 *Annu. Rev. Cell Biol.* **2** 499

Watson M E E 1984 *Nucleic Acids Res.* **12** 5145

Wiedmann M, Huth A and Rapoport T A 1986 *Biochim. Biophys. Res. Commun.* **134** 790

Wolfe P B, Wickner W and Goodman J 1983 *J. Biol. Chem.* **258** 12073