Export of Proteins Across Membranes: The Helix Reversion Hypothesis

U. P. Shinde,[1] T. N. Guru Row[2] and Y. R. Mawal[3]

A model is presented which explains the biological role of the leader peptide in protein export. Along the lines of this model, the conformational changes of a protein with environment serves as a general mechanism for translocation. The leader peptide in the cytoplasm takes a hairpin like conformation which reverts to an extended helix upon integration into the membrane. The essential features of this model are in accord with recent results of protein export.

INTRODUCTION

The transient N-terminal sequence found on most secretory proteins serves to initiate export across the inner membrane (in prokaryotes) or the endoplasmic reticulum (in eukaryotes). The initial events in the synthesis of the secretory protein prior to its translocation across the membrane have been well elucidated (1, 2). The existing models for the export process incorporate the aspects of protein conformation in many ways. Engelman and Steitz (3) have postulated the formation of a hairpin structure, one arm of which is contributed by the leader peptide and the other by the protein. Von Heijne and Blomberg (4) have proposed that the transfer takes place in a completely lipophilic atmosphere. Another school of thought is that certain membrane proteins may play an important role in export (5).

Although leader peptides have common structural features, there is no primary sequence homology among various such sequences. Gierasch and co-workers (6, 7) have suggested that the hydrophobicity and conformation of the peptide are very crucial for its interaction with the proteins of the translocation

---

[1] To whom correspondence should be addressed Ujwal P. Shinde, Division of Chemical Engineering, National Chemical Laboratory, Pune 411 008, India.
[2] Division of Physical Chemistry, Indian Institute of Science, Bangalore.
[3] Division of Biochemical Sciences, Division of Chemical Engineering, National Chemical Laboratory, Pune 411 008, India.

apparatus and for its proper insertion into the lipid bilayer. Furthermore, the leader peptide also modulates the rate of folding of the precursor polypeptide to allow the initial interaction with the components of the export apparatus (8). We propose that, in the cytoplasm the leader peptide alone takes a hairpin like conformation which is stabilized in the vicinity of the ribosome. This conformation, as it passes from the aqueous cytoplasm into the hydrophobic membrane, undergoes a conformational change. This conformational "reversion" could play a very vital role in the initiation of export of the precursor across the membrane.

THE MODEL

The main basis for the proposal is provided by a statistical analysis of the positional preference of amino acids in 248 eukaryotic and 70 prokaryotic leader peptides whose primary sequences and cleavage sites are precisely known. These were chosen from our collection totalling at the present time about 300 eukaryotic and 75 prokaryotic entries (available on request). As the length of leader peptides is variable, they were normalized to a standard length of 100 to facilitate the analysis. The preference of helix inducers (leucine and alanine) and helix destabilizers (valine, proline glycine, methionine, cysteine, phenylalanine and isoleucine) to occur at a position "i" along the normalized length is calculated as follows;

$$Po(a, i) = \frac{F(a, i)T}{NF(a)}$$

where $F(a, i)$ is the frequency with which an amino acid or a group of amino acids "a" occurs at a position "i", "$F(a)$" is the total frequency of occurrence of that amino acid or group of amino acids at all positions, "T" is the total number of amino acid residues studied and "N" is the total number of residues occurring at any position "i".

It may be argued that the normalization or averaging of the sequences may disguise several signal sequences where the pattern of residues is quite different. This possibility was eliminated by repeatedly drawing out sequences, randomly, thirty at a time, and plotting the positional preference of the residues in these sequences and comparing them with that of the total data bank. The trend observed in all the cases were similar to each other and the overall trend for all the sequences is depicted in Fig. 1. The amino acids V, P, G, M, C, F and I have been classified as helix destabilizers mainly because the ratio of their alpha to beta forming potentials as calculated by Chou and Fasman (9) is less than one.

Figure 1 depicts plots of "PO (a, i)" versus "i" for helix inducers and helix destabilizers. Np denotes the theoretically expected preference value if there was no positional bias in the occurrence of an amino acid. Any value of "Po (a, i)" greater than Np implies that the amino acid or group of amino acids ("a"), show a preference to occur at that position ("i"), while any value less than Np implies a bias of that amino acid or a group of amino acids not to occur at that position. The percentage occurrence of amino acids in leader peptides is shown in Table 1.
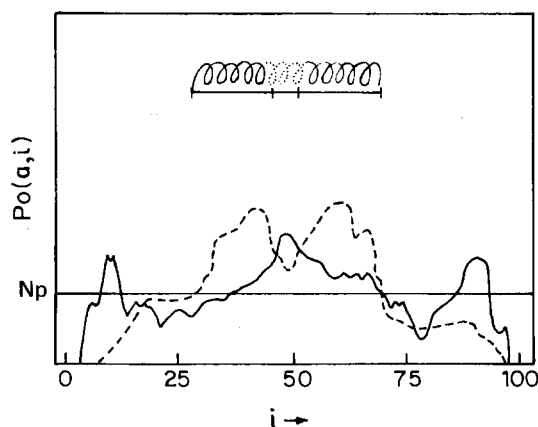
Fig. 1. Preferential occurrence of helix formers (-----)
and helix destabilizers (———) along the normalized length
of signal sequences.

We observed that certain helical conformation inducers like leucine and alanine
(9), which constitute 38% of the total composition (Table 1), show a high
preference to occur in the interior of the leader peptide (Fig. 1). Leucine, due to
its abundance could be postulated to play an important role in inducing a
conformation that is important in the interaction with the SRP and the
membrane. Furthermore the amino acids valine, isoleucine, phenylalanine,

Table 1. Percentage Composition of Amino Acid in various leader peptides

Amino Acid	Total Percentage occurrence	Total Occurrence at the N-terminal	Total Occurrence at the C-terminal	Total Occurrence in the Hydrophobic Core
Leucine	23.17	2.77	2.20	18.20
Alanine	14.53	1.95	4.00	8.58
Serine	8.22	1.22	1.85	5.15
Valine	7.63	0.96	1.10	5.57
Methionine	6.80	5.44	0.21	1.15
Glycine	5.18	0.63	1.62	2.93
Phenylalanine	5.07	0.83	0.56	3.68
Threonine	4.42	0.56	0.73	3.13
Cysteine	3.04	0.25	0.71	2.08
Isoleucine	4.99	0.98	0.64	3.37
Proline	3.08	0.56	0.69	1.83
Arginine	2.31	1.99	0.32	0.00
Tryptophan	1.72	0.37	0.20	1.15
Lysine	3.46	3.06	0.40	0.00
Glutamine	1.72	0.22	0.39	1.11
Tyrosine	1.24	0.32	0.12	0.80
Glutamic acid	0.87	0.33	0.54	0.00
Aspargine	1.14	0.44	0.26	0.44
Aspartic acid	0.54	0.34	0.20	0.00
Histidine	0.87	0.28	0.59	0.00
Total	100.00	23.50	17.33	59.17

methionine, cysteine, proline and glycine, which are helix destabilizers, show a bias to occur at 3 particular positions (Fig. 1). The centrally weak zone is sandwiched by two potentially strong helical zones and this positional specificity can give rise to the "molten globule" state as postulated by Bychkova *et al.* (10).

An analysis using the chi-square test was performed to check the statistical significance for the difference in frequency of occurrence of helix formers and destabilizers for every 5% of the normalized length of the leader peptide. These results indicate that with 19 degrees of freedom and X^2 0.05, there is significant variation in the occurrence of helix formers and destabilizers along the normalized length of leader peptides. Furthermore, this trend is observed clearly in both prokaryotes and eukaryotes separately (data not shown).

The leader peptides are all known to be highly hydrophobic, which means that the conformation taken will be one such that the hydrophobic forces act to reduce the surface area exposed to the aqueous cytoplasm. This conformation will be only marginally stable, devoid of any stable tertiary structure mainly due to the lack of any disulphide bonds and will depend entirely on the environment (11). Furthermore, strong evidence exists that protein translocation requires the involvement of a non-native or denatured state essentially in the form of a compact but flexible "molten globule" state. This state could be a hairpin like structure, with the bend occurring at the helix destabilizing zones in the interior of the hydrophobic core. Both the arms are, however, contributed by the leader peptide alone.

Blobel and coworkers (12) have shown the involvment of a cytoplasmic protein called the signal recognition particle (SRP) in the initial stages of export to form a 70 mer translation arrested ribosome complex. No significant primary sequence homology exists between the various leader peptides implying that the SRP binds to the leader peptide and the ribosome in an unconventional way possibly via conformational recognition (1). As the average length of leader peptides is about 22 amino acids and about 40 amino acid residues are shielded in the ribosome (13) this indicated that in the 70 mer complex, the leader is fully exposed to the cytoplasm and will try to attain conformation that is stable in the aqueous cytoplasmic phase. The only possible conformation with the least exposure of the hydrophobic surface will be the hairpin like helix. This is because, compared to an extended helix, approximately one fourth of the total surface area exposed to the cytoplasm is reduced when the conformation taken is a hairpin like one (3). We propose that this is the conformation that is recognized by the SRP. The stability of this conformation will further increase by intramolecular hydrophobic interactions and bifurcating hydrogen bonds (Fig. 2). From the evidences of Blobel and coworkers and Lodish and coworkers (11, 14–18) we can conclude that the SRP binds to the leader sequence and the ribosome in a way so as to (i) sterically hinder the ribosome-membrane binding site located on the exit domain, (ii) sterically hinder the A site of the ribosome so as to arrest chain elongation, (iii) hinder the charged N-terminal sequence of the nascent polypeptide from binding to the membrane, (iv) position the signal sequence in an hairpin structure so as to facilitate penetration and (v) to be in a
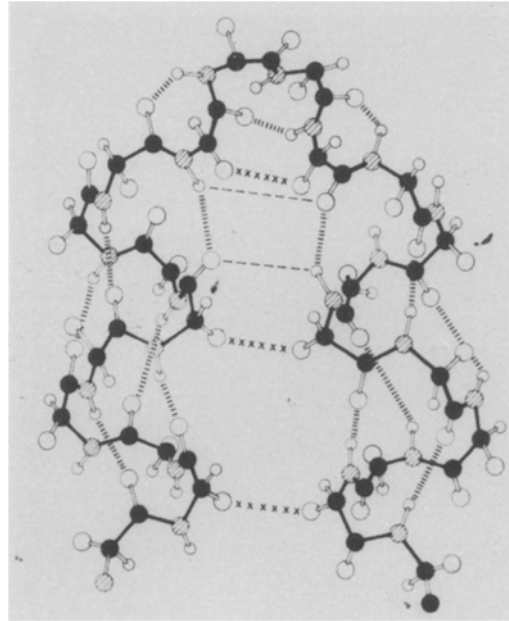
Fig. 2. Schematic representation of the proposed hairpin like conformation. × × × × Hydrophobic interactions; ||||||||| hydrogen bonding; ———— Bifurcating hydrogen bonding.

position so as to attach to the SRP receptor on the membrane. High resolution X-ray techniques have shown the exit and the translational domain are well separated on the ribosomal surface. The distance between the exit domain and the A site is estimated to be about 150 Å (13). This means that the SRP should take dimensions to span this gap so as to satisfy the above conditions. Electron microscopy studies of SRP have shown the length to be approximately 240 Å (19). When the SRP of the 70-mer complex binds to the SRP receptor it would relax the arrest on the elongation probably by a conformational change in the SRP, effecting the release of the signal sequence-ribosome complex. However, with this release, SRP would in turn lose its conformation for which the SRP receptor has high affinity, and hence dissociate away from the receptor.

After the release of the elongation arrest, the ribosome binding site as well as the charged N-terminal of the sequence would be free, thus effecting an immediate binding of the ribosome-polypeptide chain to the membrane. The model is schematically shown in Fig. 3. The hydrophobic hairpin is now exposed to the cytoplasm. However, due to the strong tendency for apolar amino acids to partition into the nonaqueous interior of the lipid bilayer, the peptide will tend to adopt a structure that maximizes the clustering of its apolar residues. This indicates that the hairpin conformation which it had acquired in the cytoplasm to reduce exposure of the hydrophobic groups will now essentially change to an extended helix. This structure reversion could be the key factor in the initiation of export. Hence the leader peptide in this state can accommodate to both polar and
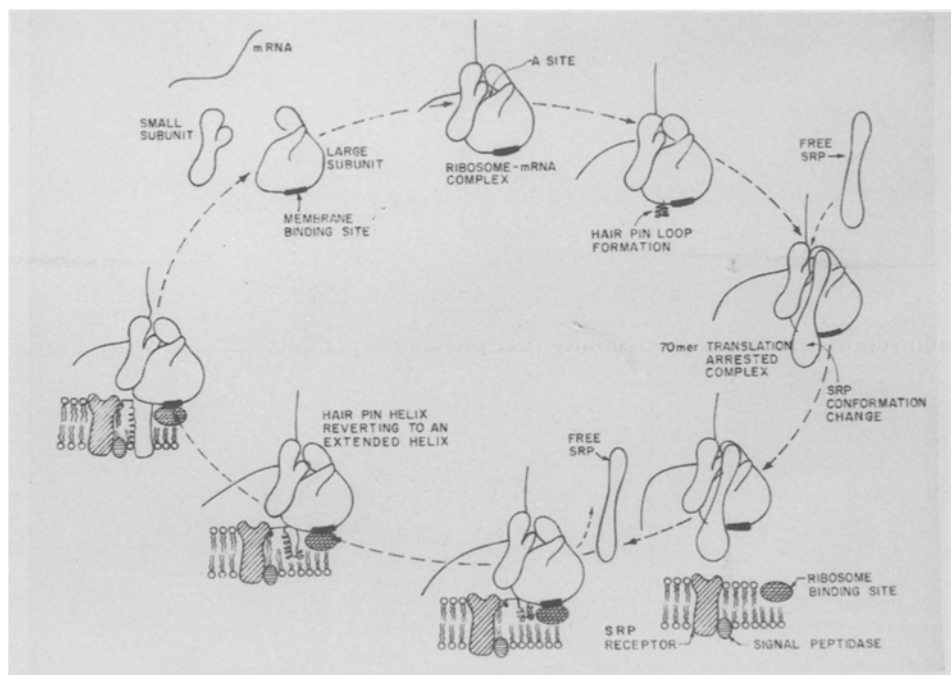
Fig. 3. Schematic diagram of the signal hypothesis as it is currently formulated.

nonpolar environments. Furthermore, the hairpin like conformation, with a less polar surface, would be translocated through the bilayer with a small free energy of activation. Energy studies have shown that such a reversion could be thermodynamically favoured due to the increase in the hydrophobic interactions (unpublished results). The subsequent transfer could then take place in a completely lipophilic atmosphere (4). Alternatively, the transport could continue via the translocator protein as postulated by Singer and co-workers (20). For post translational translocation the precursor should be devoid of a stable tertiary structure. Hence in most cases post translational translocation does not occur. However, if the stable tertiary structure can change to an unstable one, possibly via energy supplied by NTP, translocation can occur (1).

It may be doubted that leader peptides in eukaryotes are actually exposed to water. They may simply be buried into a hydrophobic pocket of SRP as soon as they emerge from the ribosome. This would imply that the SRP and ribosome would be physically associated with the nascent polypeptide chain even before the translational arrest. However, as per our knowledge, no such experimental evidence is available. It has also been shown by Rapoport and co-workers (21) that a time lag exists between the synthesis of the leader peptide and its binding to the SRP. Furthermore, by *in vitro* studies Ainger and Meyer (22) have shown that the SRP can arrest nascent chain elongation at various times during a synchronous translation, thus allowing the generation of nascent chains of increasing length. The SRP is added to the system after a time lag, indicating that

the signal sequence is completely exposed to the aqueous cytoplasm. Even so the conformation adopted in the cytoplasm is still recognized by the SRP. If the signal sequence was to enter directly inside the SRP, the conformation which it would acquire in the hydrophobic pocket of the SRP would be expected to be radically different from that which it would acquire in the aqueous cytoplasm. This is because the conformation adopted by a peptide is substantially affected by the environment (11). This advocates a proposal that the SRP can recognize not only different leader peptides but also different conformations, which implies that the SRP could interact not only with leader peptides, but also with certain other proteins. However, as per our knowledge no such experimental evidence has been reported. Thus the possibility that the leader peptide may be simply buried into the hydrophobic pocket of the SRP seems remote.

The essential features of this model are in accordance with recent ex-perimental results on protein export. Experimental work is now being carried out to prove the hairpin like conformation of the leader peptide in the vicinity of the ribosome.

BIOCHEMICAL EVIDENCE

Leucine is Essential in Conformation Induction, Protein–Protein and Protein–Lipid Interactions

Circular dichroism studies on random copolymers in aqueous solutions have shown that an increased helical stability is coincident with the incorporation of a larger amount of L-leucine in these polymers (9). Furthermore it is shown that leucine is thermodynamically a stronger helix former than most other amino acids. Alanine also shows the ability to increase the helical content of peptides. In view of this point the 27% occurrence of these amino acids in leader peptides would be expected to induce a helical conformation.

Incorporation of B-hydroxy leucine in place of leucine abolishes *in vitro* translocation, which could occur due to a high change in hydrophobicity and inability to induce a helical conformation (15). This could probably be due to an increase in the steric hindrance coincident with the incorporation of a bulky hydroxyl group in case of B-hydroxy leucine.

The Hairpin Like Conformation of the Leader Peptide is Stabilized in the Vicinity of the Ribosome and the SRP Probably Recognizes this Conformation

The SRP has been shown to interact with the leader peptide alone with a weak affinity (21). If the hairpin like conformation is acquired by this peptide in the absence of the ribosome, it would then be expected to bind strongly to the SRP. However, this is not the case. Furthermore, the binding of the SRP with the ribosomes alone has alone been shown to be very weak (21). In the context of the ribosome, the nascent polypeptide-SRP affinity is increased by approximately 6000 times. This could probably imply that the hairpin like conformation of the

leader peptide is stabilized in the context of the ribosome. Secondly, the post translational translocation of polypeptides has been shown to depend on the concentration of ribosomes (23). This indirectly implies that the export conducive conformation of the leader peptide is attained only in the presence of the ribosome.

Generally, a specific interaction between two proteins involve complementary steric fit of the surfaces, thus requiring particular amino acids at precise positions. However, as no primary sequence homology is observed in various leader peptides, it is postulated that the conformation could play an important role in recognition. In many cases, it has been shown that most amino acids in proteins are only structural in nature and their replacement by residues of similar lipophilicity do not impair biological activity (24). Thus various leader peptides, with absolutely no primary sequence homology whatsoever, could by virtue of their composition give rise to similar conformations. The SRP, therefore, may recognize the conformational backbone, irrespective of the side chains. Although such interactions may seem improbable, a precedent for this proposal is the binding by the protease La (product of the lon gene) by its denatured substrate (1, 25).

Conformational Change with Change in Environment

The helical content of signal sequences has been shown to increase with an increase in hydrophobicity of the medium (26). Circular dichroism studies of the leader peptide in an aqueous medium have also shown that conformational changes are possible upon the association and integration with the hydrophobic membrane (27).

PERSPECTIVES

The hypothesis proposed in this paper attributes a general mechanism for both co- and post-translational translocation in prokaryotes and eukaryotes. It can also explain the insertion of transmembrane proteins.

ACKNOWLEDGEMENTS

REFERENCES

1. Randall, L. L., Hardy, S. J. S. and Thom, J. R. (1987) *Ann. Rev. Microbiol.* **41**:507–541.
2. Von Heijne, G. (1988) *Biochem. Biophys. Acta.* **947**:307–333.
3. Engelman, D. M. and Steitz, T. A. (1981) *Cell* **23**:411–422.
4. Von Heijne, G. and Blomberg, C. (1979) *Eur. J. Biochem.* **97**:175–181.
5. Blobel, G. and Dobberstein, B. (1975) *J. Cell Biol.* **67**:835–851.

6. Briggs, M. and Gierasch, L. (1984) *Biochemistry* **23**:3111–3114.
7. Briggs, M., Gierasch, L., Zlotnick, A., Lear, J. and DeGrado, W. (1985) *Science* **228**:1096–1099.
8. Park, S., Liu, G., Topping, T. B., Cover, W. H. and Randall, L. L. (1988) *Science,* **269**:1033–1035.
9. Chou, P. Y. and Fasman, G. D. (1973) *J. Mol. Biol.* **74**:263–281.
10. Bychkova, V. E., Pain, R. H and Ptitsyn, O. B. (1988) *FEBS Lett.* **238**:231–234.
11. Kaiser, E. T. and Kerdy, F. J. (1984) *Science* **223**:249–255.
12. Walter, P., Ibrahimi, I. and Blobel, G. (1981) **91**:545–550.
13. Bernbeu, C. and Lake, J. A. (1982) *Proc. Natl. Acad. Sci. USA,* **79**:3111–3115.
14. Gilmore, R. and Blobel, G. (1982) *J. Cell Biol.* **96**:470–477.
15. Walter, P., Ibrahim, I. and Blobel, G. (1981) *J. Cell Biol.* **91**:551–556.
16. Walter, P. and Blobel, G. (1981) *J. Cell Biol.* **91**:551–556.
17. Walter, P. and Blobel, G. (1981) *J. Cell Biol.* **91**:557–561.
18. Rothman, J. E. and Lodish, H. E. (1977) *Nature,* **269**:575–580.
19. Andrews, D. W., Walter, P. and Ottensmeyer, F. P. (1985) *Proc. Natl. Acad. Sci. USA,* **82**:785–789.
20. Singer, S. J., Maher, P. A. and Yaffe, M. P. (1987) *Proc. Natl. Acad. Sci. USA,* **84**:1015–1019.
21. Rapoport, T. A., Heinrich, R., Walter, P. and Schulmeister, T. (1987) *J. Mol. Biol.* **195**:621–636.
22. Ainger, K. J. and Meyer, D. I. (1986) *EMBO J.* **5**:951–955.
23. Roitsch, T. and Lehle, L. (1988) *Eur. J. Biochem.* **174**:699–705.
24. Kaiser, E. T. (1986) in: *"Protein Engineering Applications in Science, Medicine and Industry",* (Inouye, M. and Sarma, R. eds.) Academic Press Inc. (London) Ltd. pp. 71–79.
25. Waxman, L and Goldberg, L. A. (1986) *Science,* **232**:500–503.
26. Rosenblatt, M., Beaudette, N. V. and Fasman, G. D. (1980) *Proc. Natl. Acad. Sci. USA,* **77**:3983–3987.
27. Briggs, M., Cornell, D., Dluhy, R. and Gierasch, L. (1986) *Science* **233**:206–208.