Hypothesis

Recognition of signal sequences

Alexei V. Finkelstein*, Peter Bendzko and Tom A. Rapoport +

Central Institute of Molecular Biology, GDR Academy of Sciences, 1115 Berlin-Buch, GDR

Received 6 July 1983

The hypothesis assumes that every continuous, entirely hydrophobic sequence of sufficient length, which is not involved in strong intramolecular contacts with other parts of the nascent protein chain, will function as a signal for translocation across the endoplasmic reticulum membrane or across the inner bacterial membrane. The signal peptide is proposed to be deeply immersed into a hydrophobic cleft of the receptor. Accordingly, only the entirely nonpolar peptides can be absorbed and, despite different primary structures, all of them would assume the same conformation dictated by the structure of the receptor pocket.

Signal sequence *Secondary structure stability*

1. INTRODUCTION

It has been firmly established that translocation of nascent secretory or membrane proteins across the endoplasmic reticulum membrane is directed by highly hydrophobic signal sequences [1,2]. There is ample evidence that there exists an ubiquitous recognition system for such sequences; this may be the recently identified signal recognition particle (SRP) [3–6]. Moreover, it has been shown that the signal sequence of an eukaryotic secretory protein can be recognized in a prokaryotic cell [7,8] and vice versa [9]. This suggests a general similarity of all signal sequences. All identified signal sequences occupy 15–35 residues at the N-terminus of the nascent polypeptide. Only one common feature was found in their primary structures: a long continuous hydrophobic region (e.g., [10]). Otherwise, no homology was discovered.

* Permanent address: Institute of Protein Research, Academy of Sciences of the USSR, Poustchino, USSR

+ To whom correspondence should be addressed at: Zentralinstitut für Molekularbiologie der AdW der DDR, Abt, Zelldifferenzierung, 1115 Berlin-Buch, Lindenberger Weg 70, DDR

How are these hydrophobic sequences recognized as signals for translocation given their large differences in lengths and amino acid sequence? One possibility would be their being imbedded into the hydrophobic interior of the membrane [11–13]. However, previous [14–16] and recent [3–6] evidence strongly argue for an initial protein–protein interaction between the signal peptide and a receptor rather than a protein–lipid interaction. Thus, the problem appears to be reduced to the following question: which feature of the signal sequences is recognized by the receptor (SRP)? Only the initial interaction between the signal peptide and the receptor will be considered in this paper and not the following translocation process across the membrane.

It is unlikely that overall hydrophobicity alone is the recognition marker. For example, the hydrophobic C-terminal part of cytochrome b_5 does not compete with a signal sequence in the binding to the receptor [17].

It has been suggested that signal sequences may self-organize to form identical secondary structures and that this preformed structure would be recognized by the receptor [18–21]. This assumption has given rise to the prediction of the secon-

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

dary structure of signal peptides by the rules of Chou and Fasman [22]. However, different authors have reached contradictory conclusions by using this method. They suggested either a common α-helix- [18] or a common β-sheet-formation [19] or even different structures for various signals [23].

In order to clarify this point we have studied the secondary structure stability in signal sequences with the aid of a molecular theory [24–26]. This theory is based on stereochemical considerations and on physical data obtained for polymers of amino acids. It has been successfully applied to predict the stability of secondary structures in polypeptides [27]. This physical aproach can be applied to various environmental conditions, whereas statistical methods [22] deal only with an average environment of amino acids in the X-ray resolved proteins.

It is known that surroundings can alter the secondary structure drastically [28]. So we have studied the structures of signal peptides for all three main types of environment: water, water/hydrophobic boundary, hydrophobic medium (fig. 1).

Firstly, we have considered the signal sequences as surrounded by water only; i.e., without interactions with other macromolecules. Stable secondary structures were observed in all signal sequences, but in some of them α-helices were more stable than β-hairpins (e.g., in pre-lysozyme) while in others the opposite was the case (e.g., in premellitin). The difference in stability between a random coil and the two structures was up to 2 kcal/mol of sequences.

Nearly the same results were obtained if the possibility of an interaction between a signal sequence and the outer surface of the receptor was taken into consideration. This treatment can be performed by assuming that the sequence in question is shallowly immersed in a hydrophobic surface [26]. It corresponds to the usual conditions for a chain in proteins.

These results confirm the conflicting data cited above. Obviously, there is no pre-formed secondary structure common to all signal sequences.

Thus, a common structure necessary for recognition is induced when a signal sequence is immersed deeply in the receptor pocket.

The case of entire insertion of a signal sequence into some non-polar medium seems to be a clue to the understanding of the way of recognition. Unlike the two cases considered above, water molecules cannot compensate here the disruption of intra- and interpolypeptide hydrogen bonds. Thus, the dehydratation of polar groups without some intra- or interchain H-bond formation hinders such deep immersion (the H-bond energy is as high as ~5 kcal/mol of bonds [29]). At the same time the free energy for complete immersion of a hydrophobic side chain is high enough (~2 kcal/mol chains [30]), suggesting a tendency to press any entirely non-polar sequence into a hydrophobic pocket of the receptor as deep as possible.

If immersed completely, the conformation of the protein fragment is determined by the pocket

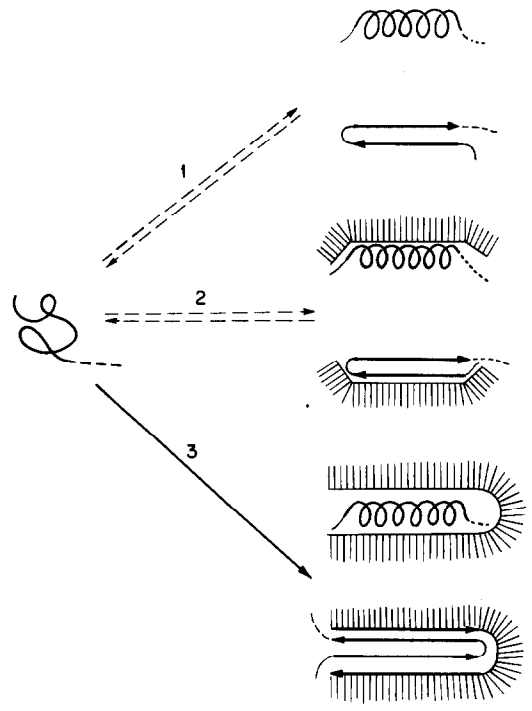

Fig.1. Different possibilities for structure formation in signal peptides: (1) transition from a random coil to either α-helix or β-hairpin in aqueous solution; (2) the signal peptide is shallowly immersed in a hydrophobic surface of the receptor; (3) the hydrophobic segment of the signal peptide is deeply immersed into a hydrophobic cleft of the receptor. If the pocket has no polar groups, helix formation occurs (upper possibility of (3)). An example of a pocket with the possibility of signal-receptor of H-bond formation is given below.

structure as all the H-bond-donors and -acceptors which are screened from water must be saturated. This applies in the first place to all the peptide NH- and CO-groups of the signal peptide and the corresponding groups at the walls of the cleft. If the pocket itself has no polar groups, the H-bond formation must occur within the signal sequence. In this case the signal peptide would form some helical, presumably α-helical, conformation. If the pocket of the receptor contains polar groups, the H-bonds must be formed with the signal peptide. The resulting conformation would be dictated by the interaction with the pocket. If the entire saturation of all H-bond donors and acceptors is impossible (say, if the sequence is not entirely hydrophobic), complete immersion and strong binding cannot occur. For example, changing one of the hydrophobic residues in the hydrophobic core of a signal peptide to a polar, uncharged one, would already reduce the binding constant by a factor of 10^3–10^5. If the change involves the introduction of a charge, the weakening of the binding would be even greater (10^5–10^8). This explains the effect of mutations in the signal sequence which were found to abolish the transport [31].

The absolute value of the binding constant can also be roughly estimated. The hydrophobic free energy, ΔF_{HP}, of maximal immersion of a hydrophobic segment of 10 amino acid residues into an apolar environment can be estimated with the aid of well known parameters (see [30]) to be ~ 20 kcal/mol of sequences. If one assumes (see [32]) that secondary structure formation is just compensated by appropriate H-bond formation, the binding constant, K_B, may be estimated:

$$K_B \approx \exp(-\Delta F_{HP}/RT).\exp(\Delta S_B/R)$$

where:

R = the gas constant;
T = the absolute temperature;
ΔS_B = the entropy for fixation of the sequence in the cleft of the receptor.

The latter can be estimated to be:

$$\exp(\Delta S_B/R \approx (a^3/M^3).(a^2/4\pi L^2).(a/\pi D)$$

where:

$M \simeq$ 12 Å is the distance between particles at 1 mol/l;
L = the length and D = the diameter of the signal peptide fragment ($L \approx D \approx 10$ Å) and $a \approx 1$ Å the dimension of possible vibrations within a Van der Waals' well.

As a result, a value of $K_B \sim 10^7$ l/mol is calculated which is in fair agreement with experimental data [3].

The length of the hydrophobic segment influences greatly the binding constant: the binding constant would increase by a factor of about 10^7 for 15 residues instead of 10, whereas binding would virtually vanish for 5 residues. Considering the uncertainties involved in the estimates given, it appears that a minimum of about 7 contiguous large hydrophobic residues is required.

It should be noted that the binding constant does not significantly depend on the structure of the signal peptide which existed before binding to the receptor which can only be marginally stable ($\leqslant 2$ kcal/mol sequence).

Obviously, any strong interaction of the hydrophobic peptide with other parts of the protein would compete with the receptor binding. A hydrophobic sequence involved in strong internal interactions, such as the strongly structured C-terminus of cytochrome b_5, cannot therefore act as a signal. Thus, effectively recognized signal peptides must be either near the N-terminus of a nascent polypeptide chain, or must follow an unfolded N-terminal part or a completely folded domain with a hydrophilic surface.

The proposed mechanism differs form other known ligand–receptor interactions. Each peptide chain located at the surface of a receptor protein globule can assume several conformations. The choice between them is due to specific interactions of the side chains. In the present case the immersion of the peptide is so deep, and this is the only possibility for entirely apolar sequences, that the conformation is dictated by the requirement of total saturation of all backbone H-bonds screened from water. All other conformations are much less stable. This explains the lack of any sequence homology among different signal sequences which can all be recognized by a single receptor.

The variation among the signal sequences excludes the possibility of a complete complementarity between their surfaces and that of the cleft of

the receptor. How can one visualize strong binding despite this fact? It has been shown recently [33,34] that the 'compact' intermediates in the unfolding of proteins lost their tight Van der Waals' packing but retained secondary structure and compactness (i.e., strong intramolecular binding) due to hydrophobic interactions and hydrogen bonds.

In summary, we propose that by deep immersion of the hydrophobic signal peptides into a hydrophobic pocket of the receptor identical conformations are assumed. Any continuous, highly hydrophobic sequence of sufficient length which is not engaged in interactions with other parts of the polypeptide chain would therefore function as a signal.

REFERENCES

[1] Blobel, G. and Sabatini, D.D. (1971) in: Biomembranes (Manson, L.A. ed) pp. 193–195, Plenum Press, New York.
[2] Blobel, G. and Dobberstein, B. (1975) J. Cell Biol. 67, 835–851.
[3] Walter, P., Ibrahim, I. and Blobel, G. (1981) J. Cell Biol. 91, 545–550.
[4] Walter, P. and Blobel, G. (1981) J. Cell Biol. 91, 551–556.
[5] Walter, P. and Blobel, G. (1981) J. Cell. Biol. 91, 557–561.
[6] Meyer, D.I., Krause, E. and Dobberstein, B. (1982) Nature 297, 647–650.
[7] Talmadge, K., Kaufman, J. and Gilbert, W. (1980) Proc. Natl. Acad. Sci. USA 77, 3988–3992.
[8] Talmadge, K., Stahl, S. and Gilbert, W. (1980) Proc. Natl. Acad. Sci. USA 77, 3369–3373.
[9] Müller, M., Ibrahim, I., Chang, N.Ch., Walter, P. and Blobel, G. (1983) J. Biol. Chem. in press.
[10] Kreil, G. (1981) Annu. Rev. Biochem. 50, 317–348.
[11] Von Heijne, G. and Blomberg, C. (1979) Eur. J. Biochem. 97, 175–181.
[12] Engelman, D.M. and Steitz, T.A. (1981) Cell 23, 411–422.
[13] Wickner, W. (1980) Science 210, 861–868.
[14] Prehn, S., Tsamaloukas, A. and Rapoport, T.A. (1980) Eur. J. Biochem. 107, 185–195.
[15] Majzoub, J.A., Rosenblatt, M., Fennick, B., Maunus, R., Kronenberg, M.H., Potts, J.T. jr and Habener, J.F. (1980) J. Biol. Chem. 265, 11478–11483.
[16] Prehn, S., Nürnberg, P. and Rapoport, T.A. (1981) FEBS Lett. 123, 79–84.
[17] Bendzko, P., Prehn, S., Pfeil, W. and Rapoport, T.A. (1982) Eur. J. Biochem. 123, 121–126.
[18] Von Heijne, G. (1980) Eur. J. Biochem. 103, 431–438.
[19] Steiner, D.F., Quinn, P.S., Chan, S.J., Marsh, J. and Tager, H.S. (1980) Ann. NY Acad. Sci. 343, 1–16.
[20] Garnier, J., Gaye, P., Mercier, J.-C. and Robson, B. (1980) Biochimie 62, 231–239.
[21] Rosenblatt, M., Beaudette, N.V. and Fasman, G.D. (1980) Proc. Natl. Acad. Sci. USA 77, 3983–3987.
[22] Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 211–245.
[23] Austen, B.M. (1979) FEBS Lett. 103, 308–312.
[24] Finkelstein, A.V. and Ptitsyn, O.B. (1976) J. Mol. Biol. 103, 15–24.
[25] Finkelstein, A.V., Ptitsyn, O.B. and Kozitsyn, S.A. (1977) Biopolymers 16, 497–524.
[26] Ptitsyn, O.B. and Finkelstein, A.V. (1983) Biopolymers 21, 15–25.
[27] Ptitsyn, O.B. and Finkelstein, A.V. (1980) Int. J. Quantum Chem. 16, 407–418.
[28] Kushner, V.P. and Frenkel, S.Ya. (1962) Biokhimiya 27, 1111–1115.
[29] Davies, M. and Thompson, D.R. (1956) J. Phys. Chem. 60, 763–773.
[30] Tanford, C. (1980) in: The Hydrophobic Effect, Wiley, New York.
[31] Emr, D.S., Hall, M.N. and Silhavy, T.J. (1980) J. Cell Biol. 86, 701–711.
[32] Ptitsyn, O.B. (1972) J. Pure Appl. Chem. 31, 227–244.
[33] Dolgikh, D., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S.Yu. and Ptitsyn, O.B. (1981) FEBS Lett. 136, 311–315.
[34] Shakhnovitch, E.I. and Finkelstein, A.V. (1982) Dokl. Akad. Nauk SSSR 267, 1247–1250.