The EMBO Journal vol.10 no.10 pp.2773-2782, 1991

Decoding signals for membrane protein assembly using alkaline phosphatase fusions

Karen McGovern[1], Michael Ehrmann and Jon Beckwith

Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA

[1]Present address: Biological Laboratories, Harvard University, Cambridge, MA 02138, USA

Communicated by J.Beckwith

We have used genetic methods to investigate the role of the different domains of a bacterial cytoplasmic membrane protein, MalF, in determining its topology. This was done by analyzing the effects on MalF topology of deleting various domains of the protein using MalF-alkaline phosphatase fusion proteins. Our results show that the cytoplasmic domains of the protein are the pre-eminent topogenic signals. These domains contain information that determines their cytoplasmic location and, thus, the orientation of the membrane spanning segments surrounding them. Periplasmic domains do not appear to have equivalent information specifying their location and membrane spanning segments do not contain information defining their orientation in the membrane. The strength of cytoplasmic domains as topogenic signals varies, correlated with the density of positively charged amino acids within them.

Key words: cytoplasmic domains/MalF/membrane spanning proteins/topology

Introduction

In the last few years, rules governing the topology of membrane proteins have begun to emerge. Topology predictions have been aided by hydropathy analysis of the amino acid sequences of membrane proteins, which can be used to identify the hydrophobic and hydrophilic domains (Jahnig, 1990). Membrane spanning sequences (MSSs) are most often extended hydrophobic domains, usually around 20 amino acids long. The hydrophilic domains of membrane proteins are found protruding from the surface of the membrane either into the cytoplasm or into the aqueous space on the other side of the membrane.

The orientation of the MSSs and the relative positions of the hydrophilic domains can often be predicted from an analysis of the composition of the hydrophilic domains. von Heijne (1986, 1988) has shown that cytoplasmic domains of membrane proteins tend to be enriched for positively charged amino acids. Experimental evidence for the importance of these basic amino acids in determining membrane protein topology is accumulating (reviewed in Boyd and Beckwith, 1990).

The topology of a number of membrane proteins has been studied directly using such techniques as protease susceptibility, location of glycosylation sites, antibody recognition, fusions to reporter proteins, and electron and X-ray diffraction (reviewed in Jennings, 1989).

Despite these advances, a number of questions remain concerning the mechanism of assembly of membrane proteins. For instance, there are contrasting proposals concerning the number of topogenic determinants in membrane proteins. According to one proposal, every MSS which is facing out (with its amino terminus in the cytoplasm) acts as a signal for export of the following hydrophilic domain (Blobel, 1980). Alternatively, it has been suggested that only the amino-terminal cytoplasmic domain and the first MSS constitute a topogenic signal and the rest of the protein inserts in the membrane following the guidance of this signal (Hartmann *et al.*, 1989).

We have used the MalF protein of *Escherichia coli* as a model system for studying membrane protein assembly. MalF is an integral cytoplasmic membrane protein required for maltose transport into the cell. Ordinarily, it is thought to exist in a complex in the membrane with the integral membrane protein, MalG, and the peripheral membrane protein, MalK (Davidson and Nikaido, 1990). A topological model for MalF, which includes eight MSSs, was initially based on hydropathy analysis of the amino acid sequence and on the charge distribution in the hydrophilic domains (Froshauer and Beckwith, 1984). According to this model, the second periplasmic domain is particularly large, being composed of 180 amino acids. Support for this proposed topology has come from studies on alkaline phosphatase (AP) and β-galactosidase fusions to MalF (Boyd *et al.*, 1987; Froshauer *et al.*, 1988). Studies with protein fusions to MalF also revealed that the first MSS, along with the cytoplasmic domain preceding it, can act as a signal for the export of AP and the insertion into the membrane of β-galactosidase (Boyd *et al.*, 1987, 1990). Nevertheless, a MalF protein devoid of the amino-terminal hydrophilic region and the first MSS can still function effectively in maltose transport (Ehrmann and Beckwith, 1991). This last finding suggests that signals other than the first one can function to promote proper insertion of MalF in the membrane. In this paper, we analyze in more detail the features of MalF that are important in determining its topology. By incorporating various deletions into MalF-AP fusions, we can analyze the influence of the different domains of the protein on its assembly. Our studies add further support for a central role of the cytoplasmic domains as topogenic determinants and allow us to distinguish between topogenic determinants of differing strengths. Finally, the results indicate that, in their assembly, membrane proteins do not simply follow the lead of the first topogenic signal, but rather that there can be competition of conflicting signals throughout the protein.

Results

Our approach is to determine the effect of deletions of different MSSs and hydrophilic domains on the topology of

MalF. This is done by incorporating these deletions into several MalF−AP fusions. AP is only enzymatically active when it is translocated across the cytoplasmic membrane to the periplasm (Manoil *et al.*, 1990). Fusions of AP to cytoplasmic domains of MalF result in a low enzymatic activity, while fusions to periplasmic domains give high activity. Thus, if a deletion when introduced into a MalF−AP fusion causes a dramatic change in AP enzymatic activity, it suggests that the topology of the hybrid protein has been altered. In most cases, to further verify protein topology in the various fusions and their deletion derivatives, we determine the pattern of protease susceptibility of the hybrid proteins in spheroplasts. In spheroplasts, those portions of the protein that protrude into the periplasm are exposed to the protease treatment. Often with fusions in which AP is exported to the periplasm, proteolytic treatment yields an AP-sized band, since this protein is highly protease resistant.

The deletions and MalF−AP fusions we have used are shown in Figure 1. The deletions were constructed on plasmids using oligonucleotide-directed mutagenesis (Materials and methods). The isolation of the fusions has been described previously (Boyd *et al.*, 1987). In all cases, the molar amounts of the fusion proteins and their deletion derivatives are far in excess of the MalG and MalK components of the maltose transport system encoded in the chromosomal *mal* genes. Thus, these other proteins, which normally interact with MalF, should not interfere with our topological analysis.

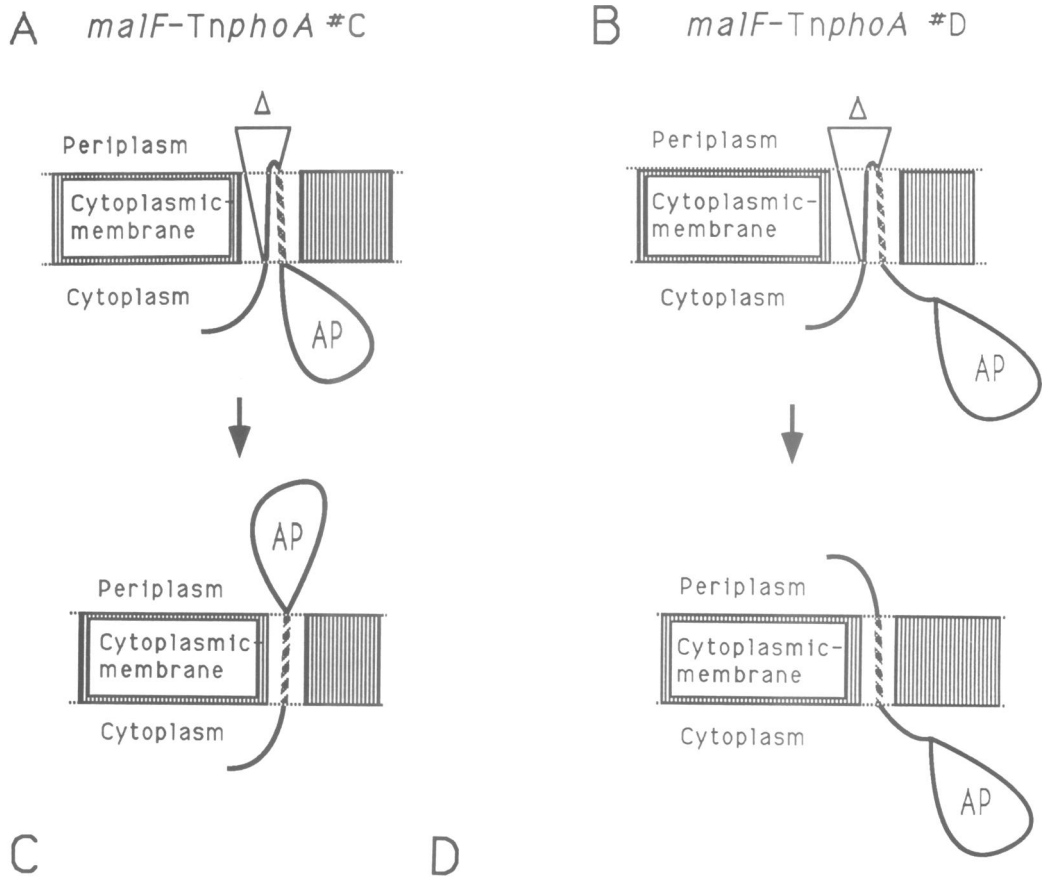
The second MSS of MalF can assume either orientation in the membrane

Several studies have shown that the orientation of a MSS of a membrane protein can be reversed in the membrane by altering the amino acid sequence around it (Audigier *et al.*, 1987; Zerial *et al.*, 1987; Wessels and Spiess, 1988; Laws and Dalbey, 1989; Szczesna-Skorupa and Kemper, 1989). It appears that simply changing the charge distribution in the hydrophilic domains which surround the MSS is sufficient to cause this inversion. We have determined whether this is also the case for MalF. In MalF, the first MSS (MSS1) is facing out while the second MSS (MSS2) is facing in. To determine whether MSS2 can be inverted in its orientation, we have deleted MSS1 and the first periplasmic domain of MalF (deletion I) so that the first positively charged cytoplasmic domain is placed just before MSS2 (Figure 2A). The deletion is introduced into MalF−AP fusion C, which has AP fused to the carboxy-terminal end of MSS2. Since, in fusion C, this end of MSS2 is on the cytoplasmic side of the membrane, the AP is cytoplasmically localized and, therefore, has low enzymatic activity (Boyd *et al.*, 1987). When fusion C carries deletion I, the fusion protein is stable and has the expected mobility on an SDS gel. The CΔI strain now exhibits high AP activity indicating that the AP moiety is localized to the periplasm (Figure 2C). This finding, in turn, means that the orientation of MSS2 has changed so that its carboxy terminus is now on the periplasmic side of the membrane (Figure 2A). Thus, the orientation of MSS2 is not determined by information within its sequence. Rather, its orientation can be altered, presumably by the introduction of basic amino acids in the hydrophilic domain preceding it.

Fig. 1. (**A**) Amino acid sequence of MalF protein. Proposed membrane spanning segments are underlined. The locations of the AP fusion joints of fusion proteins used in these studies are indicated above the sequence. (**B**) Schematic topological representation of MalF−AP fusions. Region of MalF deleted within deletions I, II and III is depicted below.

The second cytoplasmic domain plays a strong role in determining topology

In the C fusion protein, AP is fused directly to MSS2; the hybrid protein is missing the cytoplasmic domain containing basic amino acids that ordinarily follows MSS2. This hydrophilic domain may be important in determining the



C

fusion	AP activity [units] wt	ΔI
C	57	1180
D	40	78

D

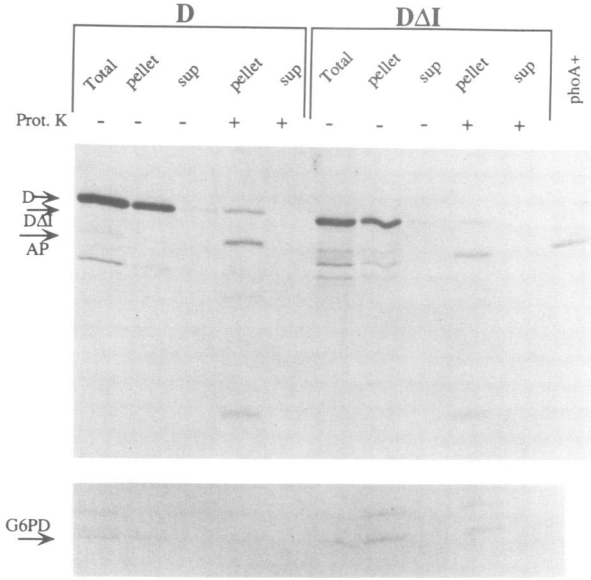
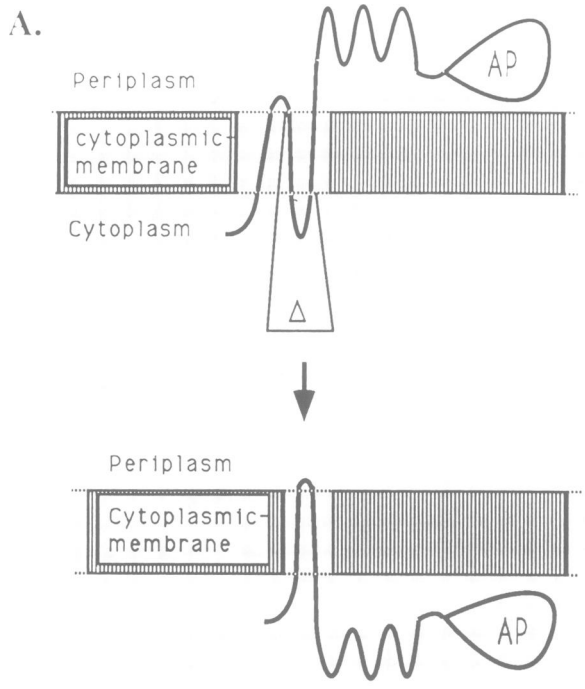


Fig. 2. (**A**) Schematic topological representation of MalF−AP fusions C and CΔI. The region within MalF that is deleted, the first MSS and first periplasmic domain, is outlined. (**B**) Schematic topological representation of MalF−AP fusions D and DΔI. The deletion is depicted as in (A). (**C**) Alkaline phosphatase activities of fusions C and D and their deletion derivatives. (**D**) Proteolysis of fusions D and DΔI. Protease susceptibility was measured by treating spheroplasts with 500 mg/ml proteinase K for 20 min and then separating the pellet (membranes and cytoplasmic fraction) from the sup (proteins exposed to periplasm). Samples were split into two and immunoprecipitated with an anti-AP antibody or anti-G6PD antibody, which is used as a cytoplasmic control for spheroplast integrity. Lanes 1−5, fusion D; lanes 6−10, fusion DΔI. The total cell sample, pellet (pelleted spheroplast) and sup (periplasmic fraction) are labeled. Samples not treated with protease are indicated with a (−), while those treated with protease are indicated by (+). The top panel are samples immunoprecipitated with anti-AP, the bottom panel samples immunoprecipitated with anti-G6PD. Arrows indicate bands representing D, DΔI, AP and G6PD. The proteolytically released fragments of D and DΔI are approximately the same mol. wt as each other and as AP, as determined by running these samples next to each other on another SDS−PAGE gel (data not shown).

normal topology of MalF. To assess the role of the second cytoplasmic domain, we have introduced deletion I into MalF−AP fusion D which has AP fused at the end of this domain (Figure 2B). In contrast to the result with fusion C where the location and activity of AP is altered, deletion I has little effect on the activity of fusion D (Figure 2C). The absence of a substantial increase in AP activity is not due to protein breakdown or reduced synthesis as demonstrated in pulse−chase experiments with labeled protein (data not shown). Thus, simply the placing of the cytoplasmic domain between the MSS and AP results in a dramatic decrease in AP activity.

In the CΔI protein, we proposed that MSS2 is reoriented because of the only flanking cytoplasmic domain, the positively charged amino terminus. In the DΔI protein, MSS2 is flanked by two hydrophilic (and positively charged) sequences, each of which presumably can act as a cytoplasmic anchor in the wild-type protein. The results suggest that the cytoplasmic domain following MSS2 is a strong signal for cytoplasmic localization, which functions even in the presence of the first cytoplasmic domain preceding MSS2.

It is not clear from these results whether the second cytoplasmic domain is a stronger signal for cytoplasmic localization than the amino-terminal cytoplasmic domain, or whether they are equally strong. We can imagine two possible outcomes of competition between these signals. One is that, with this fusion protein, the amino-terminal



B.

malF-TnPhoA fusion #J derivative	AP activity [units]
w.t.	700
ΔII	41

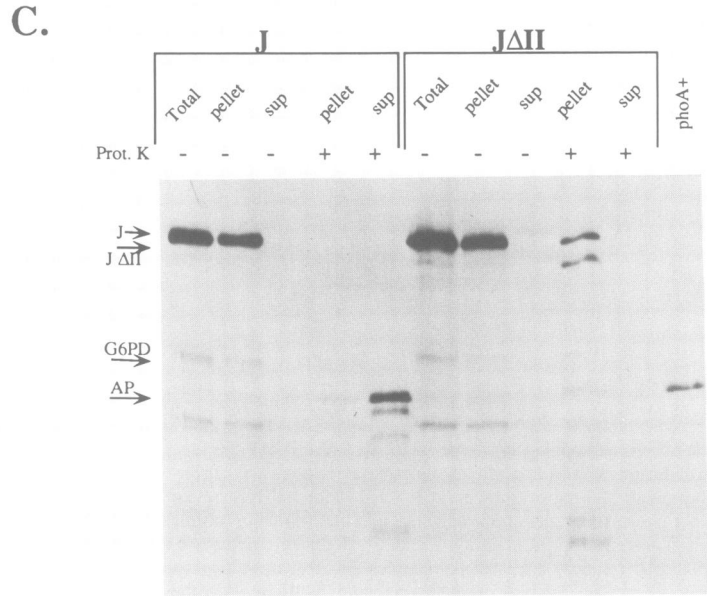


Fig. 3. (**A**) Schematic topological representation of MalF−AP fusions J and JΔII. The region of MalF deleted, the second MSS and the second cytoplasmic domain is outlined. (**B**) Alkaline phosphatase activities of fusion J and JΔII. (**D**) Proteolysis of fusions J and JΔII. Proteolysis was done as described in Figure 2. Lanes 1−5, fusion J; lanes 6−10, fusion JΔII. Arrows indicate bands representing J, JΔII, AP and G6PD.

cytoplasmic domain is exported to the periplasm, with MSS2 retaining its original orientation in the membrane. Alternatively, both hydrophilic domains may be in the cytoplasm and MSS2 is inserted aberrantly in the membrane or not inserted at all.

To distinguish between these two possibilities, we assayed the protease susceptibility of D and DΔI in spheroplasts. If the amino terminus of DΔI has been translocated across the membrane, the protein may be susceptible to protease and yield a slightly smaller fusion protein. Proteolytic cleavage may also occur with the D fusion protein itself, but at the small periplasmic domain between MSS1 and MSS2. In fact, both proteins are susceptible to protease added to spheroplasts (Figure 2D). The protease sensitivity of DΔI suggests that the amino-terminal hydrophilic domain of this protein has been translocated across the membrane. These findings and the AP assay results indicate that the second cytoplasmic domain dominates over the first in this fusion protein.

The proteolytic products seen with DΔI are smaller than might have been expected based on the proposed cleavage sites. This apparent anomaly could be due to entry of protease into the membrane. Alternatively, it might be expected that the proteinase K digestion, which alters the form of the hybrid protein, renders it susceptible to cellular proteases. The glucose-6-phosphate dehydrogenase (G6PD) control shows that spheroplasts are intact during this treatment, suggesting that proteinase K is not entering the cytoplasm. Although the amount of the AP-sized fragment is low compared with the starting level of fusion protein, the sum of the various breakdown products appears to be close to that level.

We report elsewhere (Ehrmann and Beckwith, 1991) that when deletion I is crossed into fusions where AP is fused to several different positions later in MalF, the AP retains its starting location as in fusion D. In all cases (four), the fusion proteins carrying the deletion exhibit the same level of AP activity as the parent fusion. Since all these fusions retain the second cytoplasmic domain, it appears that this sequence presents a strong barrier to the inversion of subsequent portions of the protein and may dominate in the determination of the topology of the rest of the protein. The relationship between this and following signals in the protein is elaborated on in later sections of this paper.

The role of the large periplasmic domain of MalF in determining topology

Cytoplasmic domains of membrane proteins are proposed to have information in them, namely the basic amino acids, that favor a cytoplasmic location. It seemed possible that periplasmic domains also have features to them that determine periplasmic location. Given the size of the large periplasmic domain of MalF, 180 amino acids, it seemed possible that this region might contain such information. To test this possibility, we have introduced into MalF−AP fusion J a deletion (deletion II) of MSS2 and the second cytoplasmic domain which acts as a strong cytoplasmic anchor. In the J fusion protein, AP is fused near the end of the large periplasmic domain (Figure 3A). If this domain contains strong topological signals, deletion II might not alter its periplasmic location. Then, the fusion protein would retain the high activity of the J fusion. However, if the periplasmic domain does not contain strong topological signals, the deletion could result in reorientation of MSS3 causing the periplasmic domain and AP to be localized in the cytoplasm. Such a fusion protein would have low levels of AP enzymatic activity. We find that, while the J fusion protein has high levels of activity, the JΔII fusion protein has low levels (Figure 3B).

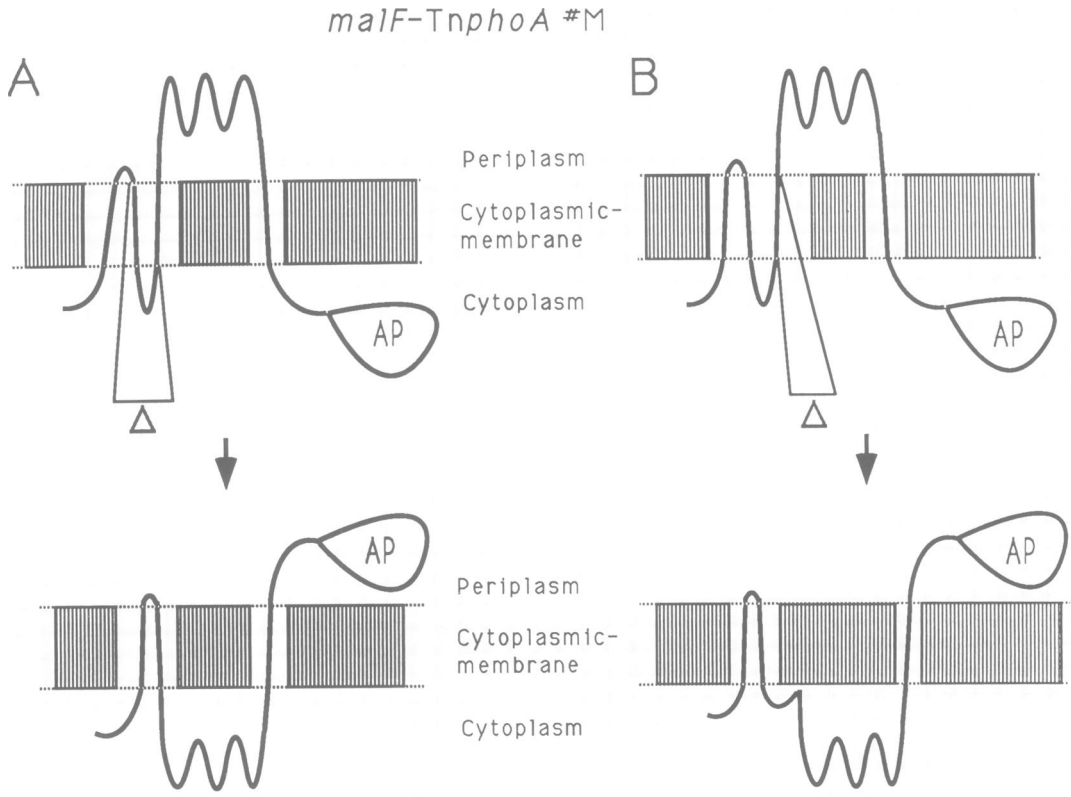
To verify further that the periplasmic domain and AP in the JΔII fusion are cytoplasmically localized, we performed protease susceptibility experiments (Figure 3C). The pattern and location of proteolytic products agrees with the proposed model for the structure of this protein. According to this model, only the small periplasmic domain between MSS1 and MSS2 is exposed on the periplasmic side of the membrane. The major bands seen on this gel would appear to be the fusion protein itself and and breakdown product resulting from cleavage in the small periplasmic domain. (The additional breakdown products of the hybrid protein and their relationship to the starting amount of that protein are explained as above for the DΔI fusion protein.) No AP-sized fragment is released to the supernatant fraction after proteolysis, as is seen with the J fusion protein. The AP assays and the proteolysis experiments indicate that in the JΔII fusion protein, the large periplasmic domain and AP have been altered in location to the cytoplasm. Apparently, no features of the sequence of the periplasmic domain prevent this change.

The third cytoplasmic domain can be translocated across the membrane

The importance of the third cytoplasmic domain in the assembly of MalF has been studied in some detail (Boyd and Beckwith, 1989). The basic amino acids within this domain are required to maintain it in the cytoplasm. To determine whether this portion of the protein is a strong topological signal like the second cytoplasmic domain, we constructed deletions within the MalF−AP fusion protein. The M fusion carries AP joined to the end of the cytoplasmic domain following MSS4 (Figure 4A and B). Both deletion II and a deletion of MSS3 (deletion III) were cloned into the M fusion gene. Whereas M has very low levels of AP activity, both deletions convert M to a high activity fusion (Figure 4C). This change in activity indicates that MSS4 has been inverted in its orientation and that AP and the third cytoplasmic domain have been translocated across the membrane.

To verify this structure for the M fusion protein carrying deletion III, we have carried out protease susceptibility experiments. If our interpretation of the activity change is correct and AP is protruding into the periplasm in the MΔIII protein, we would expect that the AP portion of the fusion protein would be released into the supernatant after protease treatment. In contrast, proteolytic products of the M fusion protein still carrying the AP moiety should not be released and should fractionate with the spheroplasts. Those expectations are realized (Figure 4D). While a very small amount of an AP-sized fragment remains in the pellet fraction after proteolysis of MΔIII, the majority of AP is released to the supernatant fraction.

These results suggest that the removal of an MSS earlier in the protein causes an inversion in orientation of subsequent portions of the protein so that the third cytoplasmic domain, normally a signal for cytoplasmic location, is now in the periplasm. Conversely, the large periplasmic domain is most likely to be internalized in the MΔIII construct, by analogy



C

malF-Tn*phoA* fusion #M derivative	AP activity [units]
wt	77
ΔII	540
ΔIII	416

D

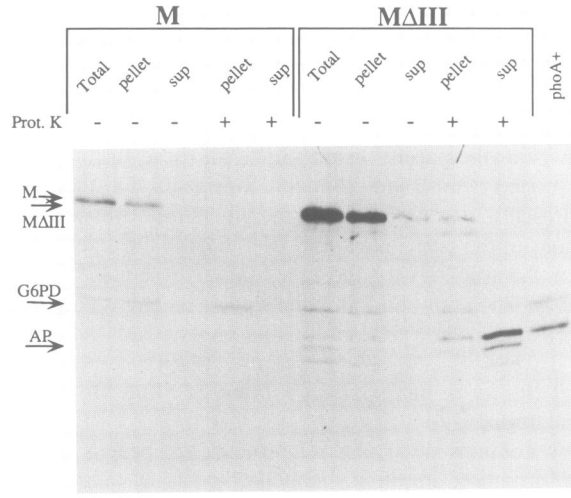
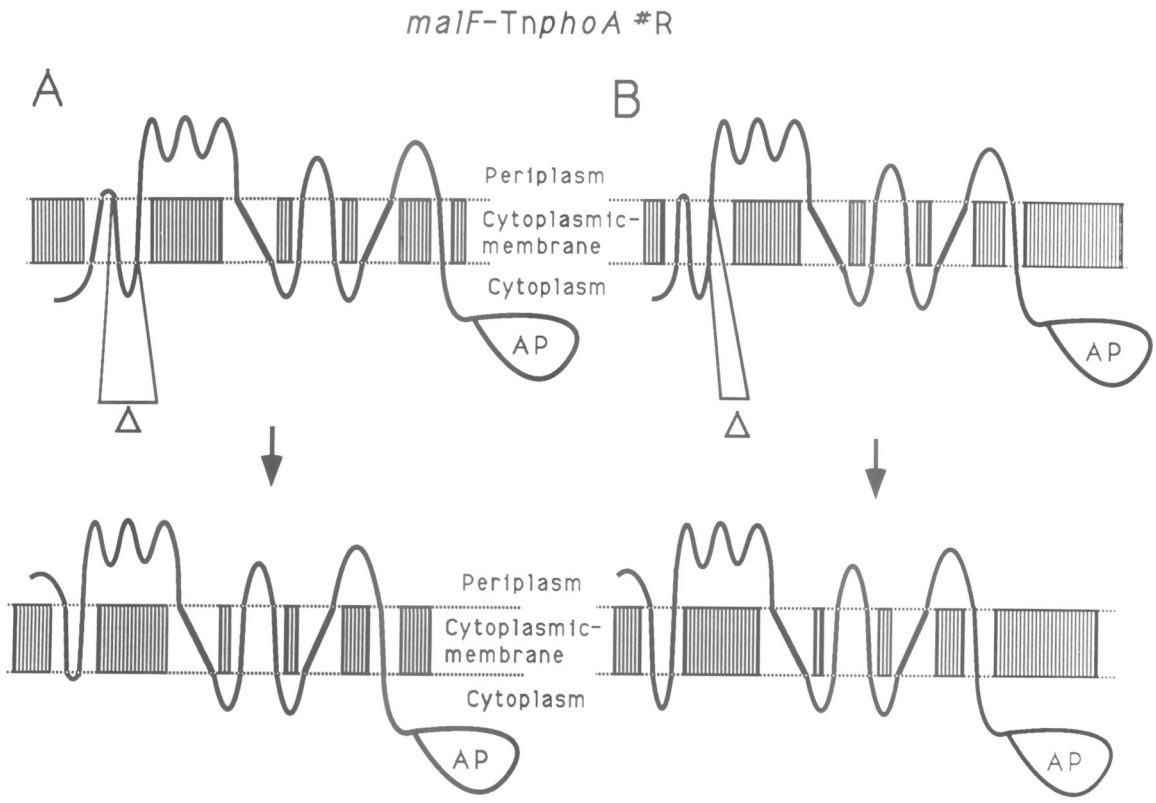


Fig. 4. (**A**) Schematic topological representation of MalF−AP fusions M and MΔII. The region of MalF deleted, MSS2 and the second cytoplasmic domain is outlined. (**B**) Schematic topological representation of MalF−AP fusions M and MΔII. The region of MalF deleted, MSS3, is outlined. (**C**) Alkaline phosphatase activities of fusion M and its deletion derivatives. (**D**) Proteolysis of M and MΔIII. Proteolysis as described except samples were chased with cold Met after labeling for 30 min. Lanes 1−5, fusion M; lanes 6−10, fusion MΔIII. Arrows, which are lined up to the center of the gel since it curves, indicate bands representing M, MΔIII, AP and G6PD.

to the apparent localization of this domain to the cytoplasm in the JΔII derivative.

Topogenic signals late in the MalF protein

In order to analyze topogenic signals late in the MalF protein, we have incorporated deletions II and III into MalF−AP fusion Q (fusion joint in the fourth periplasmic domain) and fusion R (fusion joint at the carboxy terminus of MalF). In the case of fusion Q, the incorporation of the deletions results in low levels of AP activity. The Q fusion itself makes 192 units of AP activity, QΔII 50 units and QΔIII 52 units. These levels are not due to altered amounts or stability of the fusion proteins; they are made in amounts comparable with the parent fusion and are stable (data not shown). Thus, these



C

malF-TnphoA fusion #R derivative	AP activity [units]
wt	16
ΔII	22
ΔIII	20

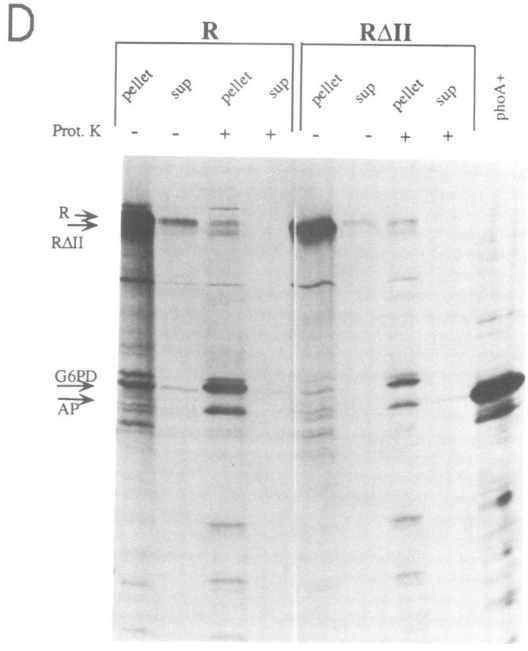


Fig. 5. (**A**) Schematic topological representation of MalF−AP fusions R and RΔII. The region of MalF deleted, MSS2 and the second cytoplasmic domain is outlined. (**B**) Schematic topological representation of MalF−AP fusions R and RΔIII. The region of MalF deleted, the third MSS, is outlined. (**C**) Alkaline phosphatase activities of fusion R and its deletion derivatives. (**D**) Proteolysis of R and RΔII. Proteolysis done as with M fusions. Lanes 1−4, fusion R; lanes 5−18, fusion RΔII. Arrows indicate bands representing R, RΔII, AP and G6PD.

deletions result in a major portion of the hybrid protein being inverted in its orientation. However, the Q fusion has lower levels of AP activity than other AP fusions to periplasmic domains so that the deletions are only reducing the activity to one quarter of the starting level. It is not clear from these activity measurements whether there is complete or only partial reorientation of the hybrid protein. It may be that the population of hybrid proteins represents a mixture with different topologies.

The R fusion protein is of particular interest, because it retains MalF transport activity (Boyd *et al.*, 1987). When deletions II and III are incorporated into this fusion, there is essentially no increase from the low levels of AP enzymatic activity (Figure 5C). In addition, protease susceptibility

experiments with fusion R and RΔII resulted in similar proteolytic patterns, with very little AP-sized protein released into the supernatant (Figure 5D). (The small amounts of AP-sized protein in the supernatant fraction with RΔII may reflect a small percentage of the hybrid protein with a different topology.) These results indicate that there is a topogenic signal(s) late in MalF that acts strongly to prevent inversion of this region. We have presented in Figure 5 models for the RΔII and RΔIII fusions assuming that the carboxy-terminal signals dominate in the process of insertion of the entire protein. Perhaps more likely is a result in which both ends of the molecule insert according to the signal within them. This assembly would then be directed by competing signals that cause the formation of aberrant structures in the regions between them.

Discussion

Our results suggest that the MalF protein contains multiple topogenic determinants which are important in directing its insertion into the membrane in the correct orientation. The simplest interpretation of these studies is that major information for proper assembly is in the cytoplasmic domains. Previous work has identified the positively charged cytoplasmic domain 3 as a topogenic determinant (Boyd and Beckwith, 1989). In this paper, we show that cytoplasmic domain 2 also plays a significant role in directing proper assembly. The incorporation of deletions into MalF−AP fusions in which the AP is fused late in or at the terminus of MalF indicate that there are also signals in this portion of the protein that can dominate in the assembly process.

Altogether, our findings indicate some features of the steps involved in membrane protein assembly. In some cases, when we alter amino-terminal regions so that the export signals would be expected to give an inverted orientation to the protein, the inversion is observed. In these cases, the amino-terminal signal may be initiating the insertion process and subsequent regions simply follow passively its lead. In other cases, the topology of the entire protein is not inverted, indicating that downstream sequences can interfere with the propagation of the initial signal.

Our results suggest that cytoplasmic domains vary in the strength with which they direct their localization to the cytoplasm. In a fusion protein that contains both the first and second cytoplasmic domains (DΔI) separated by only one MSS, it appears that the second signal provides a stronger determinant of cytoplasmic localization than the first which is exported to the periplasm. In derivatives of fusion protein M (MΔII and MΔIII), earlier cytoplasmic signals can dominate the assembly process so that the third cytoplasmic domain is translocated across the membrane. Other cases of export of cytoplasmic domains of membrane proteins as a result of genetic alterations have been reported (Lipp *et al.*, 1989; von Heijne, 1989).

In contrast to the results with fusion M, when the same deletions are introduced into the R fusion protein (RΔII and RΔIII), in which AP is joined to the very end of MalF, the localization of the carboxy terminus to the cytoplasm is not altered. Presumably this last cytoplasmic domain itself or another sequence late in the protein prevents the propagation of the signals that were promoting inversion of the protein as seen with the M fusion. While in this case we have not identified the region of the protein that is preventing inversion of the carboxy terminus, it may well be one of the last two cytoplasmic domains or the combined effects of the two of them. On the other hand, it is possible that other features of the protein are responsible for the stable localization of this portion of the protein.

Table I. Charged residue composition of cytoplasmic domains of MalF

Domain	Basic amino acids		Acidic amino acids		+ Charge	Domain length
	Arg	Lys	Glu	Asp		
1	0	4	0	2	+4	16
2	2	1	0	0	+3	10
3	2	1	1	0	+3	12
4	0	2	1	3	+2	24
5	1	2	0	1	+3	10

The total number of basic and acidic residues in the five cytoplasmic domains of MalF are presented, with the positive charge and length of each of these domains indicated.

The role of cytoplasmic domains in directing proper assembly of membrane proteins has been studied in several systems (reviewed in Boyd and Beckwith, 1990; Yamane *et al.*, 1990). A number of lines of experimentation indicate that the important feature of these domains is their positively charged amino acids. A reasonable hypothesis is that the variation we have observed in the strength of the cytoplasmic domains as topogenic signals depends on the degree or density of charge (Table I; Nilsson and von Heijne, 1990). The domination of the second over the first cytoplasmic domain could be due to differences in net charge. Although the amino-terminal hydrophilic domain contains four basic amino acids, the net charge is +2, while the second cytoplasmic domain has a net charge of +3. Furthermore, the four basic residues of the amino terminus are lysines, whereas the second cytoplasmic domain contains two arginines and one lysine. In some cases, arginines exert a much stronger effect than lysines on protein localization, presumably due to their higher p*K*a (Li *et al.*, 1988; Summers *et al.*, 1989; Zhu and Dalbey, 1989; Akita *et al.*, 1990). In addition, the second cytoplasmic domain (three positive charges in 10 residues) is more compact than the first cytoplasmic domain (four charges in 16 residues).

The finding that cytoplasmic domain 3 can be translocated across the membrane in certain deletion derivatives may be explained by the fact that it has only three positive charges in 12 residues. Similarly, cytoplasmic domain 1, which has a weaker net positive charge than the second cytoplasmic domain, can be localized to the periplasm in certain constructs. The last cytoplasmic domain, which we suspect is a strong determinant of cytoplasmic localization (see above), contains three positive charges in 10 residues similar to the strong determinant, cytoplasmic domain 2.

We have found no evidence for a role of periplasmic domains in the determination of topology. Specifically, the large periplasmic domain of MalF can be readily retained in the cytoplasm, when preceded by the appropriate signals. While these regions may play no strong positive role in assembly, features of their amino acid sequence may be restricted by the need to cross the membrane. First, a high density of positive charges at the amino terminus of a periplasmic domain may well prevent its translocation. Further, as with secreted proteins (Oxender *et al.*, 1980;

Randall and Hardy, 1986), periplasmic domains of membrane proteins may be required to remain in an export-competent state in order to be efficiently translocated. This requirement could put restrictions on the capacity of these domains to fold in the cytoplasm before export. In addition, it is possible that interactions between periplasmic domains contribute to the stabilization of the final structure of membrane proteins.

Our results suggest that the membrane spanning segments of MalF contain little or no information involved in determining proper topology. In many deletion forms of the fusion proteins, the MSSs are inverted from their normal orientation. This point is further supported by the number of systems in which the inversion of such sequences has been readily achieved (Audigier *et al.*, 1987; Zerial *et al.*, 1987; Laws and Dalbey, 1989; Szczesna-Skorupa and Kemper, 1989). However, it is reasonable to expect that interactions between membrane spanning segments will play a role in stabilizing structure and may be particularly important for the proper insertion of those membrane spanning segments that contain significant numbers of hydrophilic amino acids.

The possible role of interactions between domains in membrane protein assembly is discussed elsewhere (Boyd *et al.*, 1990). These interactions are difficult to study with the types of gene fusions described here, since all these hybrid proteins (except for R) are missing varying lengths of their carboxy terminus. However, it may be possible to study these questions using sandwich fusions in which AP is inserted into an otherwise wild-type MalF protein (Ehrmann *et al.*, 1990). Another factor in the membrane assembly of MalF not addressed in this work is its interaction with other components of the maltose transport system. MalF is thought to be part of a complex with MalG and MalK (Davidson and Nikaido, 1990). It may be that the final arrangement of MalF in the membrane is stabilized by these interactions.

We have not related our results on topogenic signals in MalF to its possible interaction with a cellular export machinery. In fact, results described elsewhere suggest that this protein may not require for its insertion into the membrane the secretion machinery as defined by *E.coli sec* genes (McGovern and Beckwith, 1991).

Materials and methods

Bacterial strains and plasmids

The *E.coli* strain used was a *recA*$^-$ derivative of DHB4, which is F'*lacIQ pro/araD139* Δ(*ara*−*leu*)7697 Δ*lacX74* Δ*phoA PvuII phoR* Δ*malF3 galE galK thi rpsL* (Boyd *et al.*, 1987). Plasmids containing the *malF*−*phoA* fusions were derived from plasmid pDHB32 as previously described (Boyd *et al.*, 1987).

Media and enzymes

Media were made according to Miller (1972). Sequenase was obtained from US Biochemicals. All other enzymes were obtained from New England Biolabs and were used according to the manufacturer's recommendations.

Antibodies

Anti-AP antibody was made by C.Gardel in our laboratory. Anti-G6PD antibody was kindly provided by the D.Fraenkel laboratory.

Assay of alkaline phosphatase

Cells were grown overnight in either NZ or M63 media and then diluted 1:100 and grown for several hours at 37°C and assayed as previously described (Michaelis *et al.*, 1983). Assays were done at least three times on duplicate cultures and the average error was ±10%.

Construction of deletions

Oligonucleotide mutagenesis was used to isolate the first deletion of each type, deletions I, II and III, and was done as previously described (Ehrmann *et al.*, 1990). Deletions were checked by restriction analysis and confirmed by DNA sequencing. Subsequent constructions were made by subcloning the deletion from one fusion to another. Subcloning deletions within other fusions was done by exchanging the *malF*−*phoA* fusion within the vector. The promoter and 5'-end of *malF* containing the deletion was left on the vector and a fragment, containing a downstream part of *malF* fused to *phoA*, was ligated in. The fusions each contain different amounts of the *malF* gene. The deletions were subcloned by cutting the plasmid with *Hin*dIII, at a unique site downstram of the *phoA* sequence, and cutting within *malF*, at the unique *Sac*I site. The old fusion was replaced by the new fusion by simply exchanging the *Hin*dIII−*Sac*I fragment. The subclones were verified by restriction analysis.

Proteolysis experiments

Cells were grown in M63 maltose and glycerol (0.2% each) plus 18 amino acids (minus Met and Cys), each at 40 μg/ml, plus 1 μg/ml thiamine and $MgSO_4$ and 50 μg/ml ampicillin. Fusions under Ptac were induced for 15 min with 5 mM isopropyl β-D-thiogalactopyranoside (IPTG). 2 ml of cells were labeled with 20−50 μCi of [^{35}S]Met for 1 min and then either chased with cold Met (0.05%) or stopped on ice. Cells were spun down, washed with spheroplast buffer (40% sucrose, 33 mM Tris, pH 8.0) and then resuspended in 1.2 ml spheroplast buffer. Lysozyme was added to 5 μg/ml and EDTA was added to 1 mM. Samples were divided into 3 × 0.4 ml samples. One represents whole cells, another plus protease, and a third minus protease. After 15 min incubation on ice, proteinase K was added to a final concentration of 500 μg/ml to one sample for 20 min. To all samples, the protease inhibitor phenylmethylsulfonyl fluoride (PMSF) was added to 2 mM. The minus and plus protease samples were spun in the microfuge at 4°C for 10 min to separate the spheroplasts (membranes and cytoplasm) from the proteins exposed to the periplasm (sup). The pelleted spheroplast (pellet) was resuspended in 0.4 ml spheroplast buffer plus 2 mM PMSF. To all samples 30 μl SDS sample buffer (50 mM Tris−HCl, pH 7.4, 1 mM EDTA, pH 8, 1% SDS) and 0.4 ml of 2 × KI buffer (100 mM Tris−HCl, pH 8, 300 mM NaCl, 4% Triton X-100, 2 mM EDTA) was added and cells were lysed with three cycles dry-ice−EtOH freeze−thaw. After spinning out the debris, antibodies were added and samples sat overnight on ice at 4°C. Antibodies to AP and the cytoplasmic protein G6PD were added in all cases except with fusions D and DΔI (Figure 2); due to the similarity in sizes of the D fusion protein and G6PD, the samples were split into two and antibody to AP was added to one sample, antibody to G6PD to the other. Samples were then immunoprecipitated as previously described (Froshauer *et al.*, 1988). They were run on Laemmli 10% SDS−PAGE (Laemmli, 1970).

Acknowledgements

We are grateful to Dan Fraenkel for providing the antibody to G6PD. This research was sponsored in part by a National Research Service award in Genetics to K.M., by NIH grant GM38922 to J.B. and from the gene technology program of DAAD to M.E.

References

Akita,M., Sasaki,S., Matsuyama,S. and Mizushima,S. (1990) *J. Biol. Chem.*, **265**, 8164−8169.

Audigier,Y., Friedlander,M. and Blobel,G. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 5783−5787.

Blobel,G. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 1496−1500.

Boyd,D. and Beckwith,J. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9446−9550.

Boyd,D. and Beckwith,J. (1990) *Cell*, **62**, 1031−1033.

Boyd,D., Manoil,C. and Beckwith,J. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 8525−8529.

Boyd,D., Manoil,C., Froshauer,S., San Millan,J.-L., Green,N., McGovern,K., Lee,C. and Beckwith,J. (1990) In King,J. and Gierasch,L. (eds), *Protein Folding: Deciphering the Second Half of the Genetic Code*. AAAS, Washington DC, pp. 314−322.

Davidson,A.L. and Nikaido,H. (1990) *J. Biol. Chem.*, **265**, 4254−4260.

Ehrmann,M. and Beckwith,J. (1991) *J. Biol. Chem.*, in press.

Ehrmann,M., Boyd,D. and Beckwith,J. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 7574−7578.

Froshauer,S. and Beckwith,J. (1984) *J. Biol. Chem.*, **259**, 10896−10903.

Froshauer,S., Green,N., Boyd,D., McGovern,K. and Beckwith,J. (1988) *J. Mol. Biol.*, **200**, 501−511.

Hartmann,E., Rapoport,T.A. and Lodish,H. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 5786–5790.
Jahnig,F. (1990) *Trends Biochem. Sci.*, **15**, 93–95.
Jennings,M.L. (1989) *Annu. Rev. Biochem.*, **58**, 999–1027.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Laws,J.K. and Dalbey,R.E. (1989) *EMBO J.*, **8**, 2095–2099.
Li,P., Beckwith,J. and Inouye,H. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 7685–7689.
Lipp,J., Flint,N., Haeuptle,M.-T. and Dobberstein,B. (1989) *J. Cell Biol.*, **109**, 2013–2022.
Manoil,C. and Beckwith,J. (1986) *Science*, **233**, 1403–1408.
Manoil,C., Mekalanos,J. and Beckwith,J. (1990) *J. Bacteriol.*, **172**, 515–518.
McGovern,K. and Beckwith,J. (1991) *J. Biol. Chem.*, in press.
Michaelis,S., Inouye,H., Oliver,D. and Beckwith,J. (1983) *J. Bacteriol.*, **254**, 366–374.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Nilsson,I. and von Heijne,G. (1990) *Cell*, **62**, 1135–1141.
Oxender,D.L., Anderson,J.J., Daniels,C.J., Landrick,R., Gunsalus,R.P., Zurawski,G. and Yanofsky,C. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 2005–2009.
Randall,L.L. and Hardy,S.J.S. (1986) *Cell*, **46**, 921–928.
Summers,R.G., Harris,C.R. and Knowles,J.R. (1989) *J. Biol. Chem.*, **264**, 20082–20088.
Szczesna-Skorupa,E. and Kemper,B. (1989) *J. Cell Biol.*, **108**, 1237–1243.
von Heijne,G. (1986) *EMBO J.*, **5**, 3021–3027.
von Heijne,G. (1988) *Eur. J. Biochem.*, **174**, 671–678.
von Heijne,G. (1989) *Nature*, **341**, 456–458.
Wessels,H.P. and Spiess,M. (1988) *Cell*, **55**, 61–70.
Yamane,K., Akiyama,Y., Ito,K. and Mizushima,S.(1990) *J. Biol. Chem.*, **265**, 21166–21171.
Zerial,M., Huylebroeck,D. and Garoff,H. (1987) *Cell*, **48**, 147–155.
Zhu,H.Y. and Dalbey,R.E. (1989) *J. Biol. Chem.*, **264**, 11833–11838.

Received on February 25, 1991; revised on June 25, 1991