Volume 335, number 3, 380–385 FEBS 13360 December 1993

Molecular analysis of *SAR1*-related cDNAs from a mouse pituitary cell line

Katherine A. Shen, Connie M. Hammond, Hsiao-Ping H. Moore*

Department of Molecular and Cell Biology, Division of Cell and Developmental Biology, 571 Life Sciences Addition, University of California, Berkeley, CA 94720, USA

Received 9 October 1993; revised version received 22 October 1993

Vesicular transport between the endoplasmic reticulum (ER) and the Golgi in the yeast *Saccharomyces cerevisiae* requires a Ras-like, small GTP-binding protein, Sar1p [1–3]. Whether a functional homologue operates in export from the ER in mammalian cells is unknown, nor is it clear if transport in other branches of the secretory pathway requires member(s) of a gene family. In this study, we used a PCR approach to examine the complexity of *SAR1*-related sequences expressed in mammalian cells that possess multiple secretory pathways. Amplification of cDNA sequences from rodent pituitary cells with primers corresponding to two conserved GTP binding domains of Sar1p yielded several clones with sequences homologous to Sar1 and/or the closely related ADP-ribosylation factor (ARF) family. Of these, only two showed closer homologies to *S. cerevisiae* Sar1 than members of the ARF family and are designated as *mSARa* and *mSARb*. Northern blot analysis shows that *mSARa* is expressed in most tissues including liver, heart, brain, skeletal muscle and kidney. In contrast, *mSARb* is preferentially expressed in skeletal muscle and liver. The full-length cDNA of *mSARa* isolated from a mouse pituitary AtT-20 cDNA library encodes a protein of 198 amino acids, and is 61.6% identical to Sar1p from *S. cerevisiae*. Thus in contrast to the large rab family of GTP-binding proteins, vesicular transport in mammalian cells appears to be mediated by a relatively small number of Sar1-related proteins.

SAR1; ADP-ribosylation factor; GTPase; Membrane transport; Vesicle budding; Secretion

1. INTRODUCTION

Protein secretion from eucaryotic cells follows a conserved pathway in which proteins are transferred vectorially between successive compartments. Genetic and biochemical analyses of secretion from the yeast *Saccharomyces cerevisiae* have identified a group of genes, *SEC12*, *SEC13*, *SEC16*, *SEC21*, *SEC23* and *SAR1* whose products are essential for the production of transport vesicles between the ER and the Golgi [2,4–8]. The *S. cerevisiae SAR1* gene was first identified as a multi-copy suppressor of a *sec12* temperature sensitive mutation and encodes a 190 amino-acid GTP-binding protein with sequence homology to Ras [1]. The gene is essential for cell growth, and depletion of its product in vitro and in vivo results in the accumulation of secretory proteins in the ER [1–2]. Recent results indicate that Sar1p directly interacts with two other proteins required for ER to Golgi transport. The GTPase activity of Sar1p is stimulated by Sec23p [9], and Sec12p catalyzes the exchange of GDP to GTP on Sar1p [10]. Although a Sar1p homologue in mammalian tissue has not been described, it is likely to exist. Functional homologues of Sar1p and Sec12p have been found in the fission yeast *Schizosaccharomyces pombe* and *Arabidopsis thaliana* [11]. In addition, we have previously described an 85 kDa protein in a variety of mammalian tissues which cross-reacts with antisera against the *S. cerevisiae* Sec23p [12]. A cDNA from mouse fibroblast with significant homology to *S. cerevisiae SEC23* has been recently reported [13].

At present it is unclear whether Sar1p is only required for ER to Golgi transport, or Sar1p and/or other Sar1p-related proteins are involved in transport through successive compartments of the entire secretory pathway. Two other Ras-like GTPases have also been shown to regulate vesicular transport. The Rab proteins are implicated in membrane fusion (reviewed in [14]), and the ADP-ribosylation factors (ARF) in the control of coat protein assembly (reviewed in [15]). In mammals, both are members of large gene families. At least in the case of Rab proteins, the individual members appear to mediate transport between specific sets of membrane compartments. Whether Sar1p also belongs to a family is currently unknown. In this paper, we used a PCR approach to examine the complexity of *SAR1*-related sequences expressed in rodent pituitary cells. Pituitary was chosen because endocrine cells contain several well characterized secretory pathways. In addition to constitutive secretion, the individual pituitary cell secretes a characteristic hormone by a regulated process. The *trans*-Golgi network of these cells thus have the machinery to generate different types of secretory vesicles.

*Corresponding author. Fax: (1) (510) 643 8708.

 Published by Elsevier Science Publishers B.V.

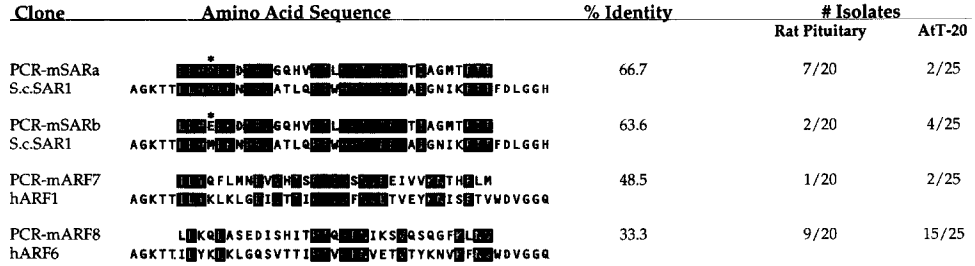

Fig. 1. PCR amplification of *SAR1*-related sequences from rodent pituitary. A PCR reaction was carried out using degenerate oligonucleotide primers encoding the first and third GTP-binding domains of *S. cerevisiae* Sar1p, AGKTT (sense) and FDLGGH (anti-sense) with total RNA from rat pituitary as templates. An identical reaction was performed using a mouse pituitary AtT-20 cell cDNA library. The amplified sequences were cloned into BlueScript vectors and 45 random clones were analyzed by DNA sequencing. Clones with predicted amino acid sequences homologous to known GTPases are tabulated here. These fall into 4 groups: PCR-mSARa and PCR-mSARb (with only one amino acid difference from each other in this region – see asterisks) are most closely related to *S. cerevisiae SAR1* [1]. PCR-mARF7 and PCR-mARF8 are most closely related to human *ARF1* [18] and human *ARF6* [19], respectively. The predicted amino acid sequences of the PCR fragments are aligned to their closest relatives. Identical sequences are shaded in black. Percentage identity shown refers to the percentage of sequence identity in this region. The number of isolates refers to the number of clones representing each sequence in 20 and 25 random clones derived from rat pituitary and AtT-20 cDNA library, respectively.

2. MATERIALS AND METHODS

2.1. *Cloning strategy*

Total RNA was isolated from rat pituitary [16]. First strand cDNA synthesis was carried out using an oligonucleotide primer corresponding to the anti-sense strand of the amino acid sequence FDLGGH. The resulting cDNA was then used as the template for PCR amplification with the following primers. Primer no. 1, 5′-TTTGTCGACG-C(ATGC)GG(ATGC)AA(AG)AC(ATGC)AC-3′, corresponds to the sense-strand of *S.c. SAR1* GTP-binding domain G-1, AGKTT (see Fig. 4). Primer no. 2, 5′-CGGGAATTCTG(ATGC)CC(ATGC)CC-(ATGC)A(AG)(AG)TC(AG)AA-3′, corresponds to the anti-sense strand of *S.c. SAR1* GTP-binding domain G-3, FDLGGH. PCR conditions were: denaturation at 95°C for 30 s, annealing at 60°C for 30 s, extension at 70°C for 1 min, and the reactions were repeated for 35 cycles. The amplified fragments were cloned into pBlueScript II KS$^+$ (Stratagene, La Jolla, CA), and sequenced with the T7 primer (Promega, Madison, WI). A similar PCR reaction was carried out to isolate *SAR1* related sequences from a mouse pituitary cell line AtT-20.

2.2. *Isolation of full-length mSARa cDNA*

The PCR-mSARa fragment from AtT-20 was labeled and used to probe an AtT-20 cDNA library [17]. Hybridization was performed at 37°C in 5 × Denhardt's, 50% formamide, 6 × SSPE (1 × SSPE = 150 mM NaCl/10 mM sodium phosphate, pH 7.4/1 mM EDTA), and 100 μg/ml salmon sperm DNA for 12–16 h. The filters were washed twice in 2 × SSC (1 × SSC = 150 mM NaCl/15 mM Sodium citrate, pH 7.0)/0.1% SDS for 15 min, once in 1 × SSC/0.1% SDS for 15 min and then once in 0.5 × SSC/0.1% SDS for 15 min, all at room temperature. Final wash was performed at 48°C in 0.1 × SSC/0.1% SDS for 2 h. Candidate clones were re-screened, and two positive clones were subjected to restriction mapping and sequencing. DNA sequences were analyzed using the GCG Sequence Analysis Software for UNIX, Version 7.2 (Genetics Computer Group, Inc., Madison, WI).

2.3. *Northern analysis*

Nylon membranes containing poly A$^+$ RNAs from mouse (BALB/c) heart, brain, spleen, liver, lung, skeletal muscle, kidney and testis (2 μg per lane) were obtained from Clontech Laboratories, Inc. (Palo Alto, CA). The blots were probed with the full-length *mSARa* coding region excised from pCDM8-mSAR1 as an *Xho*I/*Not*I fragment, or the PCR fragment of *mSARb* from pBlueScript II KS$^+$-mSARb as an *Eco*RI/*Sal*I fragment. Hybridization was performed at 42°C in 10 × Denhardt's, 50% formamide, 5 × SSPE, 2% SDS and 100 μg/ml

	mSara										
mSara	100	mSarb									
mSarb	97	100	mArf7								
mArf7	18.2	18.2	100	mArf8							
mArf8	18.2	18.2	21.2	100	S.c.Sar1						
S.c.Sar1	66.7	63.6	21.2	18.2	100	S.p.Sar1					
S.p.Sar1	69.7	66.7	21.2	15.2	84.8	100	A.t.Sar1				
A.t.Sar1	66.7	63.6	21.2	21.2	69.7	63.6	100	S.c.Arf1,2			
S.c.Arf1,2	24.2	24.2	45.4	30.3	33.3	27.3	24.2	100	hArf1,3		
hArf1,3	24.2	24.2	48.5	30.3	33.3	27.3	27.3	87.9	100	hArf2,4,5	
hArf2,4,5	24.2	24.2	48.5	30.3	33.3	27.3	27.3	84.8	97	100	hArf6
hArf6	24.2	24.2	39.4	33.3	30.3	27.3	27.3	72.7	78.8	78.8	100

Fig. 2. Comparison of predicted amino acid sequences from PCR amplified fragments and known Sar1p and Arfp. Shown are percentages of amino acid sequence identities between PCR-mSARa, PCR-mSARb, PCR-mARF7, PCR-mARF8 and the corresponding regions of *S. cerevisiae* Sar1p [1], *S. pombe* Sar1p [11], *A. thaliana* Sar1p [11], *S. cerevisiae* Arf1p [24], *S.cerevisiae* Arf2p [24], and human Arf1p [18], Arf2p [20], Arf3p [18], Arf4p [20–21], Arf5p [19], Arf6p [19]. Sequences for *S. c.* Arf1p and *S. c.* Arf2p are identical in this region, as are the sequences between hArf1p and hArf3p, and the sequences between hArf2p, hArf4p and hArf5p.

salmon sperm DNA for 12–16 h. The blot was washed three times with 2 × SSC, 0.05% SDS at room temperature for 20 min, and twice with 0.1 × SSC, 0.1% SDS at 50°C for 20 min.

3. RESULTS

To examine the complexity of *SAR1*-related sequences, we used a PCR strategy to amplify regions encompassing the G-1 and G-3 GTP binding domains (see arrows in Fig. 4) from rat pituitary. Two degenerate oligonucleotide primers encoding the sequences AGKTT and FDLGGH were synthesized for PCR amplification. First-strand cDNA was prepared from rat pituitary total RNA, and used as the template. A similar reaction was carried out using a cDNA library from the mouse anterior pituitary cell line AtT-20 as the template. Of 20 randomly selected clones derived from rat pituitary, 19 show some sequence homologies to small GTP binding proteins. Similarly, 23 of the 25 randomly selected clones from AtT-20 are specific. Clones with identical predicted amino acid sequences were obtained from both rat pituitary and mouse AtT-20 cells. These clones fall into four categories: two sequences are highly homologous to *S. cerevisiae* Sar1p, and the other two show closer homology to the ADP-ribosylation Factor (ARF) members of the Ras-superfamily; all sequences were represented by more than one clone (See Fig. 1). The two sequences that are highly related to Sar1p show 82.4% identity to each other at the nucleotide level, and differ only in one amino acid in this effector domain (see stars in Fig. 1). These two fragments were named PCR-mSARa and PCR-mSARb, and they are 66.7% and 63.6% identical in amino acid sequence to S.c. Sar1p, respectively. This degree of homology is similar to the two known Sar1 homologues from *S. pombe* and *A. thaliana* , which showed 84.9% and 69.7% identity, respectively, to *S. cerevisiae* Sar1 in this region (Fig. 2). The other two sequences are much less homologous to S.c. Sar1p (< 22%); comparison of the predicted amino acid sequences to the SwissProt databank show that they are closest related to the six known human ADP-ribosylation factors [18–21] (Figs. 1 and 2). These sequences are designated as PCR-mARF7 and PCR-mARF8. The predicted amino acid sequence of PCR-mARF7 is more closely related to human Arf1p (48.5% identical) than that of PCR-mARF8 (30.3% identical) (Fig. 2). However, both clones are more distantly related to known ARFs than the known ARFs are among themselves (see Fig. 2 for comparison). For instance, human Arf2, Arf3, Arf4, Arf5, Arf6 are 97%, 100%, 97%, 97%, 79% identical to Arf1 in this region respectively. Of the six human Arfs, Arf6 is most distant from the Arf1–Arf5; PCR-mARF7 is closer to Arf1–Arf5 (48.5% identical) than Arf6 (39.4% identical). PCR-mARF8 is more distant to all Arfs than PCR-mARF7, although it shows slightly higher homology to hArf6 (33.3% identical) than hArf1–hArf5 (30.3% identical). These results suggest that pituitary cells express two closely related *SAR1* genes and a number of ARF-related sequences.

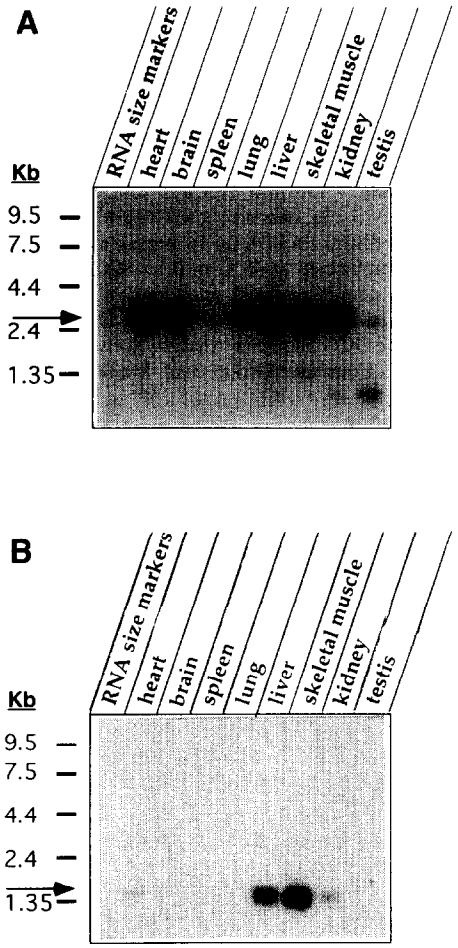

Fig. 3. Northern blot analysis of the expression of *mSARa* and *mSARb*. Two μg of poly-A RNA from various tissues of Balb/c mice was fractionated on formaldehyde gels and transferred to nylon filters. The filter was hybridized with *mSARa* probe corresponding to the entire coding region (Panel A). The filter was stripped and probed again with a PCR fragment encompassing the first and third GTP binding domain of *mSARb* (Panel B). Both probes were gel purified and labeled with [^{32}P]dCTP by random hexamer priming. The autoradiographs shown were exposed for 16 h. Arrows indicate the transcripts specifically labeled by the probes. Numbers indicate the sizes of standard RNA in kilobases (Kb).

The presence of two *mSAR* cDNAs in the pituitary prompted us to examine their tissue distributions. The PCR fragments of *mSARa* and *mSARb* were random-primed labeled for Northern blot analysis with mouse heart, brain, spleen, lung, skeletal muscle, kidney, and testis. Hybridization conditions were chosen such that there is virtually no cross-hybridization between the two clones. As shown in Fig. 3A, *mSARa* hybridizes to a transcript approximately 3 kb in size. This transcript is present in most tissues examined. In contrast, *mSARb* detected a smaller transcript of approximately 1.5 kb (Fig. 3B). This transcript is highly abundant in skeletal muscle and liver, very low in kidney and heart, and not detectable in other tissues.

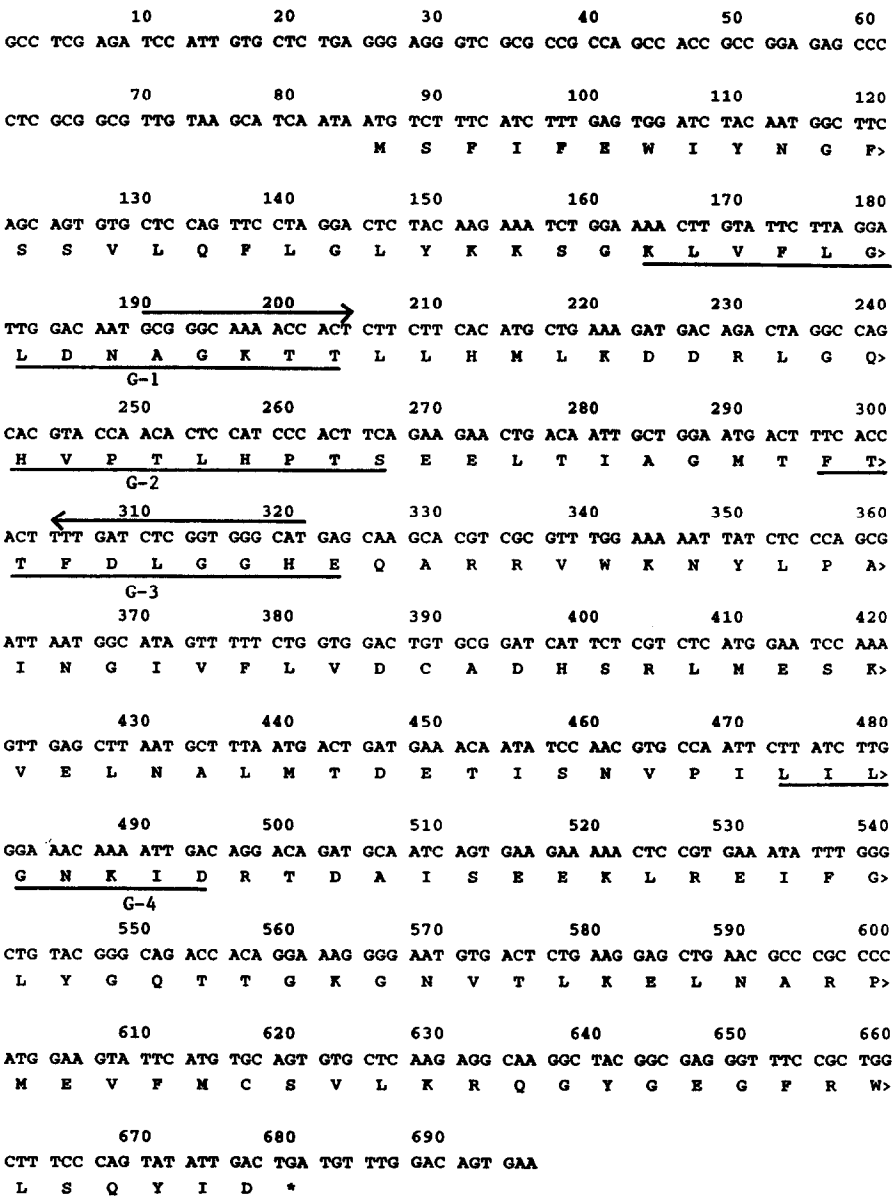

Fig. 4. Nucleotide sequence of mouse *SARa*. Nucleotide sequences and deduced amino acid sequences of mSar-a are shown. Arrows indicate the PCR primers used to amplify *SAR1*-related sequences. Underlined sequences indicate the four GTP-binding domains. This sequence has been submitted to the EMBL Genbank under the accession number L20294.

Constitutive secretion occurs in most cell types. Thus a functional homologue of *S.c. SAR1* is expected to be expressed in most cell types. Of the two PCR-SAR clones, *mSARa* messages are detected in most tissue types and thus this gene is a candidate for the mammalian counterpart of the *S. c. SAR1*. The mSARa PCR fragment was used to isolate the full-length cDNA from an AtT-20 cDNA library. Of 400,000 clones screened, two hybridized strongly to labeled PCR-mSARa and were subjected to sequence analysis. The DNA sequence of the coding region is shown in Fig. 4. The open reading frame is 594 nucleotides in length and encodes a protein with a predicted molecular weight of 21,670 Da. A comparison of *mSARa* with *S.c. SAR1* and the two recently reported *SAR1* from *S. pombe* and *A. thaliana* is shown in Fig. 5. The full length *m SARa* gene shows a high degree of sequence identity to known *SAR1* (61.6%, 60.4%, and 64.6% identical to *S.cerevisiae, S. pombe*, and *A. thaliana*, respectively). The similarity in sequence extends across the entire coding region. Like the other *SAR1*, *mSAR1a* does not have an N-terminal glycine for myristoylation. It also does not contain carboxyl terminal target sites for isoprenylation.

Fig. 5. Predicted amino acid sequence of mouse Sar-a and comparison with Sar1p from other species. The four GTP-binding domains are labeled as G-1, G-2, G-3 and G-4. Identical amino acids are shaded in black.

4. DISCUSSION

Of the three classes of Ras-like, small GTPases involved in vesicular transport, the Rab family constitutes a large family of more than 20 members (reviewed in [14,22]. Molecular cloning of *ARF* from mammals also revealed the existence of multiple genes [18–21]. We show here that in contrast to the *ARF* and *RAB* genes, the complexity of *SAR1*-related cDNAs in rodents appears to be relatively simple. Although the PCR strategy used here may not have pulled out all *SAR1*-related sequences, it probably gave a fairly representative picture of the complexity of these genes. The regions used for PCR primers are typically conserved among members of the sub-families of the Ras GTP-binding proteins; PCR amplification with primers corresponding to these same regions has detected a large number of *RAB* cDNAs [17,23]. Thus it is likely that there are only two *SAR1* genes in rodents, although we cannot exclude the possibility that other genes exist. The two *ARF*-like sequences showed closer homologies to Arfs than Sar1, and thus are probably not functionally homologous to *SAR1*. These two sequences are much more divergent from the known *ARF* sequences, and their exact functions remain to be determined.

The existence of two closely related *SAR* cDNAs in mammals with distinct patterns of tissue distribution raises interesting questions about their functions. The two proteins are not simply isoforms expressed in different cell types, since both are found in a single cell (such as the AtT-20 cell line). One possibility is that they are functionally redundant, as exemplified by Arf1p and Arf2p in *S. cerevisiae* [24]. However, if this were the case, it is not immediately obvious why certain tissues (such as the skeletal muscle) would need higher levels of a specific isoform, i.e. mSar-b, than other tissues. A more likely explanation is that the two forms perform distinct functions. This would be consistent with the finding that only a single functional *SAR1* gene has been found in *S. cerevisiae*, *S. pombe*, or *A. thaliana* that could suppress the ER to Golgi *sec12* mutation. An attractive hypothesis is that one form is necessary for ER to Golgi transport and the other form is the functional counterpart in later stages of constitutive secretion, for instance, Golgi to surface transport. However, the constitutive pathway is present in almost all cells and proteins operating in each step of this pathway are expected to be ubiquitous and present in similar amounts in most tissues, a prediction clearly not consistent with *mSARb*. The high expression of mSar-b in skeletal muscle suggests a function that is highly specialized in this tissue. In this regard, it is interesting to note that there are multiple intracellular calcium storage pools which appear to be in communication with one another, and GTP has been suspected to play a role in this communication event (reviewed in [25]). Perhaps mSar-a is involved in export of proteins via the secretory pathway, while mSar-b regulates traffic between the ER and other calcium storage organelles such as the sarcoplasmic reticulum in muscles.

If mSar-a is indeed operative in the secretory pathway, is it required in every step of the pathway? Biochemical studies show that guanine nucleotide binding by the *S.c.* Sar1p is controlled by a specific GAP protein, Sec23p. If Sar1p is repeatedly used throughout the secretory pathway, both Sar1p and Sec23p should be found in multiple sites. However, localization studies show that immunoreactive Sec23p in pancreatic cells is restricted to the transitional zone between the ER and the Golgi; little or no immunoreactivity could be detected in the Golgi membranes or the *trans* side of the Golgi cisternae [12]. This pattern implies a function unique to ER export. The ER plays an important role to prevent export of mis-folded proteins; perhaps Sar1p (and its associated proteins) function by restricting protein export to a spatially defined area thereby reducing the likelihood of secreting polypeptides that have yet to achieve their final conformation. In this scenario, the

hydrolysis of GTP could somehow be coupled to the recruitment of mature polypeptides to vesicle budding sites, or alternatively, it could directly activate membrane budding. Future experiments will be necessary to address these possibilities.

Acknowledgements: We thank Betty Kim for technical assistance. This work was supported by NIH Grant GM35239, a grant from the Lucille P. Markey Charitable Fund and American Cancer Society Grant CB-89A.

REFERENCES

[1] Nakano, A. and Muramatsu, M. (1989) J. Cell Biol. 109, 2677–2691.
[2] d'Enfert, C. et al. (1991) J. Cell Biol. 114, 663–670.
[3] Barlowe, C., d'Enfert, C. and Schekman, R. (1993) J. Biol. Chem. 268, 874–879.
[4] Hicke, L. and Schekman, R. (1989) EMBO J. 8, 1677–84.
[5] Kaiser, C.A. and Schekman, R. (1990) Cell 61, 723–733.
[6] Rexach, M.F. and Schekman, R. (1991) J. Cell Biol. 114, 219–229.
[7] Hosobuchi, M., Kreis, T. and Schekman, R. (1992) Nature 360, 603–605.
[8] Pryer, N.K. et al. (1993) J. Cell Biol. 120, 865–875.
[9] Yoshihisa, T., Barlowe, C. and Schekman, R. (1993) Science 259, 1466–1468.
[10] Barlowe, C. and Schekman, R. (1993) Nature 365, 347–349.
[11] d'Enfert, C., Gensse, M. and Gaillardin, C. (1992) EMBO J. 11, 4205–4211.
[12] Orci, L. et al. (1991) Proc. Natl. Acad. Sci. USA 88, 8611–8615.
[13] Wadhwa, R. et al. (1993) FEBS Lett. 315, 193–196.
[14] Pfeffer, S.R. (1992) Trends Cell Biol. 2, 41–46.
[15] Rothman, J.E. and Orci, L. (1992) Nature 355, 409–415.
[16] Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning, 2nd Edition.* 1989, New York, Cold Spring Harbor Laboratory Press.
[17] Yu, H., Leaf, D.S. and Moore, H.P.H. (1993) Gene 132, 273–278.
[18] Bobak, D.A. et al. (1989) Proc. Natl. Acad. Sci. USA 86, 6101–6105.
[19] Tsuchiya, M. et al. (1991) J. Biol. Chem. 266, 2772–2777.
[20] Kahn, R.A. et al. (1991) J. Biol. Chem. 266, 2604–2614.
[21] Monaco, L. et al. (1990) Proc. Natl. Acad. Sci. USA 87, 2206–2210.
[22] Balch, W. (1990) Trends Biochem. Sci. 15, 473–477.
[23] Chavrier, P., Simons, K. and Zerial, M. (1992) Gene 112, 261–264.
[24] Stearns, T. et al. (1990) Mol. Cell. Biol. 10, 6690–6699.
[25] Lytton, J. and Nigam, S.K. (1992) Curr. Op. Cell Biol. 4, 220–226.