Volume 254, number 1,2, 177-183 FEB 07554 August 1989

Intracellular targetting signals of polymeric immunoglobulin receptors are highly conserved between species

George Banting, Brigitte Brake, Paola Braghetta[+], J. Paul Luzio* and Keith K. Stanley°

European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, D-6900 Heidelberg, FRG

Received 6 July 1989

A rat liver cDNA library, constructed in the plasmid expression vector pUEX, was immunoscreened using a rabbit polyclonal antiserum raised against rat liver Golgi membrane proteins. A sub-set of isolated clones were shown to encode the rat polymeric immunoglobulin receptor (pIgR). DNA sequence analysis of these clones provided the complete coding sequence of rat pIgR. Subsequent alignment of rat, rabbit and human predicted amino acid sequences demonstrated that the greatest degree of homology between the three pIgRs lies in their cytoplasmic tails; a region previously shown to be important for correct targetting and trancytosis of rabbit pIgR [(1984) Nature 308, 37–43].

Immunoglobulin receptor, polymeric; Trancytosis: Plasmid pUEX; Expression cloning

1. INTRODUCTION

The polymeric immunoglobulin receptor (plgR) is an integral membrane protein which binds polymeric immunoglobulins (plg) at the basolateral surface of many glandular epithelial cells (sinusoidal surface of hepatocytes). Endocytosis of the receptor-ligand complex is followed by trancytosis to the apical cell surface (bile canalicular front of hepatocytes) and proteolytic cleavage of the receptor to release plg into the apical medium (bile in the case of liver). plg is released in association with the cleaved portion of plgR (termed secretory component, SC) and serves to protect mucosal surfaces from invasion by environmental pathogens. Intramolecular deletion experiments following cloning of the rabbit plgR [1] implied the presence of a lumenally encoded apical targetting sequence and cytoplasmically encoded signal(s) for basolateral targetting and endocytosis [2,3]. Using a polyclonal rabbit serum raised against rat liver Golgi membrane proteins we have isolated cDNA clones encoding rat plgR. The amino acid sequences of the lumenal domains of the rat, rabbit and recently published human [4] receptors are only 36% identical, whilst their cytoplasmic tails display more than 60% identity. This degree of sequence homology in the cytoplasmic tail suggests functional conservation and should help to define signals required for correct sorting/targetting/trancytosis of plgR. It also justifies the strategy of cloning cDNAs from different species in order to gain an insight into functionally conserved areas of proteins.

Correspondence (present) address: G. Banting, Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, England

[+] *Present address:* Istituto di'Istologia ed Embriologia, via Falloppio 50, 1-35100 Padova, Italy

* *Present address:* Department of Clinical Biochemistry, University of Cambridge, Addenbrookes Hospital, Hills Road, Cambridge CB2 2QR, England

° *Present address:* The Heart Research Institute, 145-147 Missenden Road, Camperdown, Sydney, NSW 2050, Australia

Abbreviations: SDS-PAGE, SDS-polyacrylamide gel electrophoresis; SC, secretory component; pIgR, polymeric immunoglobulin receptor

The nucleotide sequence reported here has been submitted to the EMBL/Genbank database under the accession no. X15741

Published by Elsevier Science Publishers B.V. (Biomedical Division)

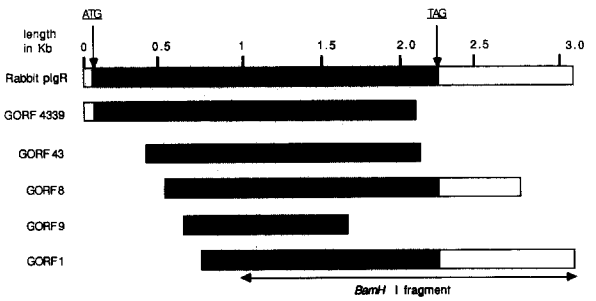


Fig.1. Alignment of clones used to sequence rat pIgR. The largest insert from group B (GORF1) extends into a region with homology to the 3′ non-coding region of rabbit pIgR, but fails to reach the 5′-end of the coding sequence. Whilst some of the other clones in group B encode sequences upstream of the GORF1 start (e.g. GORFs 8,9,43), none reach the 5′-end of the rat pIgR coding region. Rescreening the rat liver cDNA library with a probe made from the 5′-end of GORF43 led to the isolation of clones (e.g. GORF4339) encoding sequences upstream of the initiator methionine. The entire coding region was sequenced in both directions. The DNA sequence has been submitted for inclusion in the EMBL DNA database. The *Bam*HI fragment used as probe in subsequent Northern blot analysis is indicated.

2. EXPERIMENTAL

2.1. *cDNA library construction and immunoscreening*

The rat liver cDNA library was constructed in the plasmid expression vector pUEX [5] using the adaptor cloning strategy [6] and immunoscreened following established procedures [7,8].

2.2. *DNA sequencing*

DNA sequences were produced by a combination of double-strand sequencing in plasmid vectors and single-strand sequencing in M13 vectors using the dideoxy chain-termination method [9] and US Biochem. Corp. (USB) reagents.

2.3. *Electrophoresis and Western blotting*

Samples were prepared for electrophoresis, separated by SDS-PAGE, transferred to nitrocellulose and immunoblotted according to [7–11].

2.4. *Antibody production and immunofluorescence analysis*

Fusion proteins isolated from SDS-PAGE slices were used to immunise rabbits as in [12]. Sections for immunofluorescence were prepared and bound antibody visualised using rhodamine-conjugated sheep anti-rabbit IgG as described [12]. Rabbit anti-rat secretory component (SC) antibodies 552 and 231 have been described elsewhere [14].

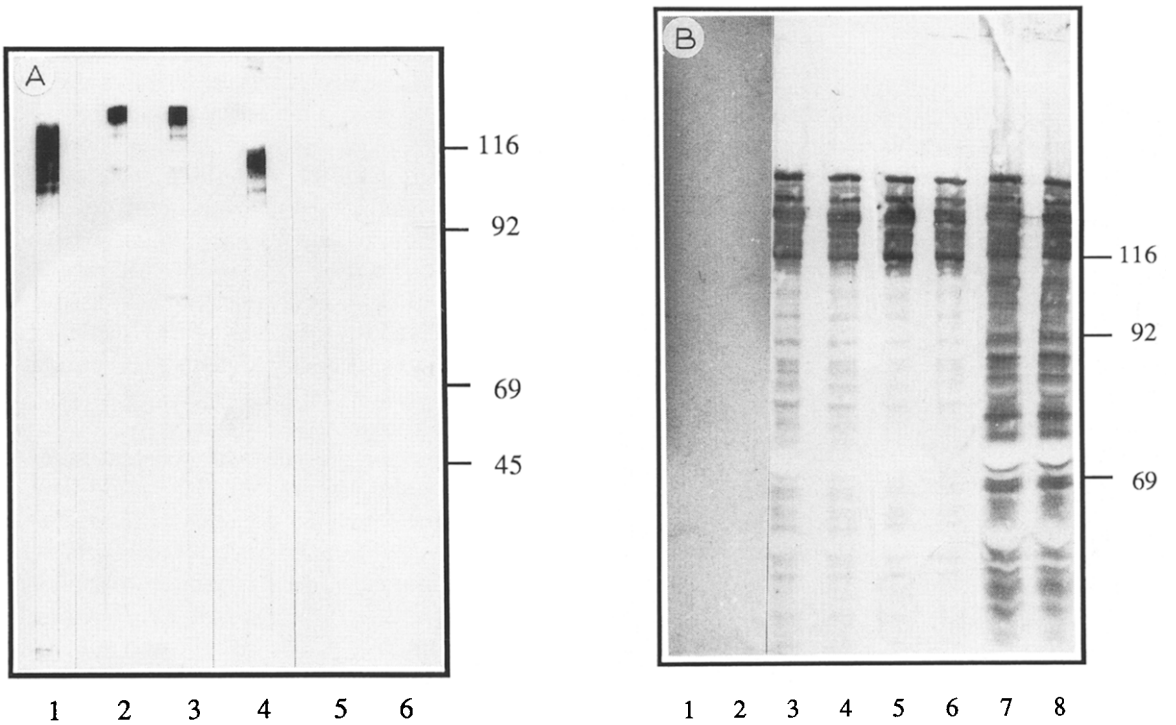
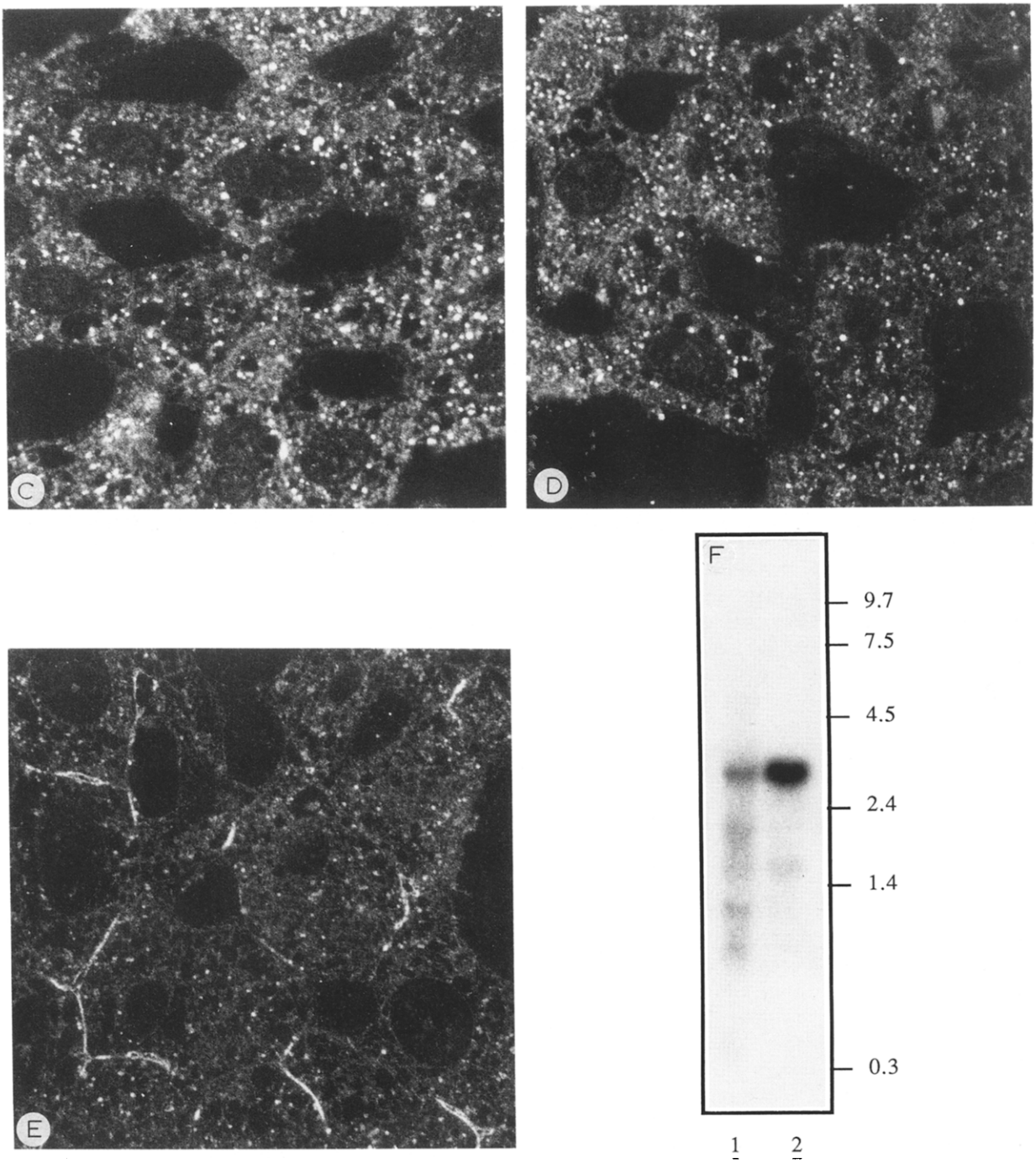


Fig.2. (a) Western blot analysis of TX114 detergent (tracks 1–3) and aqueous (tracks 4–6) phases of rat liver Golgi fractions. Tracks 1,4 were probed with a polyclonal rabbit serum which recognises the heavily glycosylated trans-Golgi network specific integral membrane protein TGN38 [12]. As expected, TGN38 partitions predominantly into the detergent phase. Tracks 2,5 were probed with polyclonal rabbit serum raised against the fusion protein encoded by GORF1, and tracks 3,6 with the anti-rat SC polyclonal rabbit serum 552. Both these antisera detect a molecule of ~120 kDa which partitions almost exclusively into the detergent phase. (b) Western blot analysis of fusion proteins expressed by GORF1. Tracks 1 + 2 were probed with a polyclonal antiserum raised against rat albumin, 3 + 4



with rabbit anti-rat SC antiserum 231, 5 + 6 with rabbit anti-rat SC antiserum 552, and 7 + 8 with rabbit anti-GORF1 antiserum. 231, 552 and anti-GORF1 recognise the GORF1 fusion protein and its degradation products; the anti-albumin antiserum fails to do so. The lower molecular mass bands recognised by the anti-GORF1 polyclonal antiserum correspond to *E. coli* proteins, detection of these proteins being due to the presence of antibodies to *E. coli* proteins raised during immunisation with the crude GORF1 fusion protein. (c–e) Indirect immunofluorescence analysis of rat liver sections using anti-GORF1 antiserum (c), rabbit anti-rat SC antiserum 231 (d) or rabbit anti-rat SC antiserum 552 (e). (f) Northern blot analysis, using a *Bam*HI fragment from GORF1 as probe (see fig.1), of poly(A^+) RNA prepared from rat kidney (track 1) and liver (track 2). 5 μg poly(A^+) RNA were loaded in each track. Numbers in (a,b) denote kDa; in (f) kb.

2.5. *Northern blot analysis*

Previously published procedures were used throughout [12,13].

3. RESULTS AND DISCUSSION

We have recently described a new immunological strategy for the molecular cloning of organelle-specific membrane proteins [12]; a polyclonal rabbit antiserum raised against integral membrane proteins isolated from a rat liver Golgi preparation was used to screen 500 000 clones of a total rat liver cDNA library constructed in the plasmid expression vector pUEX [5]. 68 positive clones were classified into seven groups (A–G) according to cross-hybridization analyses with DNA probes and affinity-purified antibodies. DNA sequences from the ends of the largest insert in each group were used to screen available databases. The 5′ and 3′ sequences of the insert from group B (the group containing most clones, 28/68) showed homology with the rabbit plgR, a molecule previously described as being an abundant and immunogenic component of rat liver Golgi membranes [15]. The complete coding sequence of rat plgR, together with the 5′ and 3′ non-coding sequences (obtained from the clones shown in fig. 1), has been submitted to the EMBL/Genbank database. The predicted amino acid sequence of rat plgR is shown in fig.3.

The fusion protein encoded by GORF1 was used to immunise rabbits, the resulting polyclonal antiserum recognizing a protein of ~120 kDa in Western blot analysis of the TX114 phase of a rat liver Golgi preparation (fig.2a). Another rabbit antiserum (552), raised against rat SC and previously shown to recognise rat plgR [14], detects a molecule with the same mobility on SDS gels (fig.2a), as does a second anti-rat SC antiserum (231[14]: not shown). Antisera 552 and 231 also recognise the GORF1 fusion protein in Western blot analysis (fig.2b). These three antisera (anti-GORF1, 552 and 231) were used in indirect immunofluorescence staining of rat liver sections (fig.2c–e); anti-GORF1 and 231 exhibit similar patterns of punctate intracellular staining, while 552 mainly recognises the bile canalicular front with some of the intracellular punctate staining given by the other antisera suggesting that it may recognise an additional epitope on the plgR. Since the extra epitope recognised by 552 is only detectable at the bile canalicular front it might arise from a conformational change in the plgR at this membrane or be exposed by the cleavage of SC from the intact plgR.

DNA probes from clones encoding the rabbit plgR have previously been shown to recognise two distinct mRNA species in Northern blot analysis of rabbit liver RNA, these having been assumed to arise from differential splicing and to encode the high- and low-molecular forms-mass of the rabbit receptor [1]. Biochemical data have demonstrated that the mature rat plgR exists in a single 120 kDa form [14], suggesting that it is encoded by a single mRNA species. To test this hypothesis, a DNA probe from GORF1 was used to screen a Northern blot of poly(A^+) RNA isolated from rat liver and kidney. A single message of 3.5 kb was detected in both tissues, however it is 20-fold more abundant in liver than in kidney (fig.2f). This differential in mRNA levels is consistent with the high level of expression and trancytosis of plgR in rat hepatocytes and the presence of greater amounts of free SC in rat bile [16] vs rat urine [17].

These biochemical data coupled with the homology to rabbit plgR convinced us that we had cloned the rat plgR and prompted us to compare the rat, rabbit and partial human sequences [4,18] more closely.

Fig.3 shows an alignment of rat, rabbit and human plgR protein sequences, the conserved amino acids being denoted by capital letters. The human sequence is a composite of the SC protein sequence derived by Eiffert et al. [18] (to residue 53 of that sequence) followed by the conceptual translation of the partial cDNA sequence isolated by Krajci et al. [4]. The lumenal domains of the three proteins are only 36% identical, whilst their cytoplasmic tails show greater than 60% identity. Although the lumenal domains demonstrate only limited primary sequence homology, they all share the structural characteristic of encoding five immunoglobulin (Ig) repeat units (fig.3), a feature previously noted in the rabbit plgR [1]. The presence of five potential sites for the addition of *N*-linked sugars in the lumenal domain of rat plgR (fig.3) supports the biochemical data demonstrating substantial *N*-glycosylation of rat plgR [14]. The longest stretch of primary amino acid sequence absolutely conserved between all three

Fig.3. Alignment of rat, rabbit [1]and human [4,18] pIgR sequences using the UWGCG GAP programme [28]; the human sequence is a composite of the SC protein sequence derived by Eiffert et al. [18] (to residue 53 of that sequence) followed by the conceptual translation of the partial cDNA sequence isolated by Krajci et al. [4]. The numbering system defines the residues of the aligned sequences (including gaps), but does not identify individual residues of a particular sequence. The position of predicted signal peptidase cleavage in rat and rabbit pIgRs is indicated by an arrow (based on the weight matrices provided by Von Heijne [29]). Identical residues between all three species are shown in upper case. Sites of potential N-glycosylation [30] are overlined. The lumenal sequence (SVSITCYYP) conserved between human, rat and rabbit pIgRs is boxed, as is the predicted transmembrane region (based on hydropathy plot [28] and sequence motifs [31]). The conserved predicted cleavage site for SC is boxed and arrowed; whilst the second amino acid in this site is not identical between all three species, the change from a glutamate to aspartate residue is highly conservative. The first and last cysteine residues of conserved Ig domains in the lumenal portion of the pIgRs are indicated by closed and open circles respectively. In the cytoplasmic tails conserved tyrosine and serine residues possibly required for correct pIgR sorting are designated by the asterisks.

species (SVSITCYYP) is close to the amino terminus of the mature protein and is the core of a 38-residue region with 29 identities between the three pIgRs. Sequence homology in this region may well reflect functional conservation to maintain the ability to bind IgA dimers non-covalently, a function previously assigned to the amino-terminal domain of rabbit SC [19]. Mostov et al., [1] suggested three possible sites for cleavage of rabbit pIgR to release SC, only one of which is conserved between all three species (fig.3) and thus represents the most likely target for the enzyme which release SC from membrane-bound pIgR.

It has been reported recently that the cytoplasmic domain of rabbit pIgR is important for basolateral targetting, endocytosis and post-endocytotic sorting into one of three possible pathways; trancytosis, recycling or degradation [Mostov, K.E., Casanova, J. and Breitfeld, P., personal communication: ASCB abstr. 2478 (1988) J. Cell. Biol. 107]. The high degree of sequence homology between rat, rabbit and human pIgR cytoplasmic tails presumably reflects conservation of intramolecular signals required for correct targetting/sorting/trancytosis. Deletion of the carboxy-terminal 30 amino acids of rabbit pIgR produces a receptor that is endocytosed much less efficiently than the wild type and, this region of the tail contains a tyrosine residue shared by all three pIgRs (fig.3); since such residues have been implicated in receptor-mediated endocytosis via clathrin-coated pits [20], it is possible that this

Table 1

Amino acid sequence comparison of LAMP proteins

LAMP protein	Ref.	Species	Identity (%) between LAMPs from all four species		
			lumenal	trans-membrane	cyto-plasmic
LAMP-1	24	mouse	36	83	100
LAMP-A	25	human	36	83	100
LEP 100	26	chicken	36	83	100
LGP 120	27	rat	36	83	100

Amino acid sequences of lysosomal associated membrane proteins (LAMPs) were aligned using the University of Wisconsin Genetics Computer Group (UWGCG) GAP programme [28]

tyrosine is instrumental in allowing rapid endocytosis of the plgR. A second cytoplasmic tyrosine residue is shared by all three plgRs, which is closer to the transmembrane region and may also be involved in receptor-mediated endocytosis (the cytoplasmic tail of the low density lipoprotein receptor-related protein [21] contains two tyrosine residues both of which are involved in endocytosis of the protein; Hamann, U., personal communication). Phosphorylation of a serine residue influences post-endocytotic sorting of rabbit plgR [22] [Casanova, J.E. and Mostov, K.E., personal communication: ASCB abstr. 2522 (1988) *J. Cell Biol.* 107], this residue being conserved in rat plgR (fig.3).

Cloning cDNAs encoding the same protein from different species can thus be a powerful means of obtaining insights into functionally conserved areas of the molecule, since conserved regions of amino acid sequence usually represent the cores of domains within the protein structure or amino acids of functional importance [23]. The cytoplasmic tails of other intracellular membrane proteins display a similar degree of conservation, e.g. the lysosomal associated membrane proteins (LAMPs) from various species [24–27] (table 1). Conservation of cysteine residues in the lumenal domains of LAMPS allows sequence alignment and implies the existence of conformational homology in this region in the absence of linear sequence homology, a situation similar to that found with plgR. The identity between LAMP cytoplasmic tails suggests they may also play a role in the correct targetting of such proteins. Recent experiments support this view [Matthews, P.M. and Fambrough, D.M., personal communication: ASCB abstr. 2479 (1988) *J. Cell Biol.* 107] and add validity to the strategy of cloning cDNAs encoding the same protein from more than one species in order to define regions of functionally conserved amino acid sequence.

Acknowledgements: We thank Dr K. Howell for advice, criticism and gifts of antisera 552 and 231, and Dr P. Rice for patient assistance with and supplements to UWGCG programmes. G.B. thanks the Royal Society and EMBO, J.P.L. the von Humboldt organisation and B.B. EMBL for support.

REFERENCES

[1] Mostov, K.E., Friedlander, M. and Blobel, G. (1984) Nature 308, 37–43.
[2] Mostov, K.E., de Bruyn Kops, A. and Deitcher, D.L. (1986) Cell 47, 359–364.
[3] Mostov, K.E., Breitfeld, P. and Harris, J.M. (1987) J. Cell Biol. 105, 2031–2036.
[4] Krajci, P., Solberg, R., Sandberg, M., Oyen, O., Jahnsen, T. and Brandtzaeg, P. (1989) Biochem. Biophys. Res. Commun. 158, 783–789.
[5] Bressan, G.M. and Stanley, K.K. (1987) Nucleic Acids Res. 15, 10056.
[6] Haymerle, H., Herz, J., Bressan, G.M., Frank, R. and Stanley, K.K. (1986) Nucleic Acids Res. 14, 8615–8624.
[7] Stanley, K.K. and Luzio, J.P. (1984) EMBO J. 3, 1429–1434.
[8] Banting, G.S., Pym, B., Darling, S.M. and Goodfellow, P.N. (1989) Mol. Immunol. 26, 181–188.
[9] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[10] Bordier, C. (1981) J. Biol. Chem. 256, 1604–1607.
[11] Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350–4354.
[12] Luzio, J.P., Brake, B., Banting, G., Howell, K., Braghetta, P. and Stanley, K.K. (1989) J. Cell Biol., submitted.
[13] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[14] Sztul, E.S., Howell, K.E. and Palade, G.E. (1985) J. Cell Biol. 100, 1248–1254.
[15] Howell, K.E. and Sztul, E.S. (1982) J. Cell Biol. 95, 242a.
[16] Mullock, B.M., Jones, R.S. and Hinton, R.H. (1980) FEBS Lett. 113, 201–205.
[17] Hinton, R.H., Lau, D.Y.-H., Sanderson, C. and Mullock, B.M. (1985) Protides of the Biological Fluids, vol. 32, pp. 81–84.
[18] Eiffert, H. et al. (1984) Hoppe-Seyler's Z. Physiol. Chem. 365, 1489–1495.
[19] Frutiger, S., Hughes, G.J., Hanley, W.C., Kingzette, M. and Jaton, J.-C. (1986) J. Biol. Chem. 261, 16673–16681.
[20] Lazarovits, J. and Roth, M. (1988) Cell 53, 743–752.
[21] Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H. and Stanley, K.K. (1988) EMBO J. 7, 4119–4127.

[22] Larkin, J.M., Sztul, E. and Palade, G.E. (1986) Proc. Natl. Acad. Sci. USA 83, 4759-4763.
[23] Stanley, K.K. and Herz, J. (1987) EMBO J. 6, 1951-1957.
[24] Chen, J.W., Cha, Y.K., Yuksel, K.U., Gracy, R.W. and August, J.T. (1988) J. Cell Biol. 263, 8754-8758.
[25] Viitala, J., Carlsson, S.R., Siebert, P.D. and Fukuda, M. (1988) Proc. Natl. Acad. Sci. USA 85, 3743-3747.
[26] Fambrough, D.M., Takeyasu, K., Lippincott-Schwarz, J. and Siegel, N.R. (1988) J. Cell Biol. 106, 61-67.
[27] Howe, C.L., Granger, B.L., Hull, M., Green, S.A., Gabel, C.A., Helenius, A. and Mellman, I. (1988) Proc. Natl. Acad. Sci. USA 85, 7577-7581.
[28] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395.
[29] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683-4690.
[30] Marshall, R.D. (1974) Biochem. Soc. Symp. 40, 17-26.
[31] Von Heijne, G. (1985) Curr. Top. Membranes Transp. 24, 151-180.