

Identification, sequencing and expression of an integral membrane protein of the *trans*-Golgi network (TGN38)

J. Paul LUZIO,* Brigitte BRAKE,† George BANTING,‡ Kathryn E. HOWELL,§ Paola BRAGHETTA|| and Keith K. STANLEY¶

European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, D-6900 Heidelberg, Federal Republic of Germany

Organelle-specific integral membrane proteins were identified by a novel strategy which gives rise to monospecific antibodies to these proteins as well as to the cDNA clones encoding them. A cDNA expression library was screened with a polyclonal antiserum raised against Triton X-114-extracted organelle proteins and clones were then grouped using antibodies affinity-purified on individual fusion proteins. The identification, molecular cloning and sequencing are described of a type 1 membrane protein (TGN38) which is located specifically in the *trans*-Golgi network.

INTRODUCTION

An important approach to the identification of target and retention signals of resident membrane proteins of organelles is cDNA cloning followed by sequence comparisons and molecular 'cut-and-paste' experiments. Since very few cDNA clones coding for resident membrane proteins of intracellular organelles have been described, there is clearly a need to develop a general methodology for their isolation and characterization. We describe here a technique which generates both cDNA clones and specific antisera against membrane proteins of defined organelles. In essence, this involves the isolation of a large number of different cDNA clones using a polyclonal serum raised against a mixture of proteins, followed by classification of the clones using antibodies which are affinity-purified on individual expressed fusion proteins.

In the present study we have isolated a number of cDNA clones and specific antibodies using a polyclonal serum raised against a mixture of rat liver Golgi membrane proteins. We report the deduced amino acid sequence of one of these, a type 1 membrane protein (TGN38), which is located specifically in the *trans*-Golgi network, a structure in which the sorting of newly synthesized proteins into separate carrier vesicles for delivery to post-Golgi destinations occurs [1–4]. Transient transfection experiments suggested that the predicted cytoplasmic tail contains information necessary for intracellular targeting and localization.

EXPERIMENTAL

Preparation of Golgi membrane fraction antigen

Male Sprague–Dawley rats (100–120 g; Süddeutsche Tierzucht, Tuttlingen, Germany) were starved overnight and killed

(without ethanol pre-treatment). The livers were removed and homogenized in ice-cold 0.25 M-sucrose containing 2 µg of leupeptin/ml and 20 µg of pepstatin/ml. All solutions until the final step contained these proteinase inhibitors. Golgi fractions were isolated on a sucrose gradient as described [5]. The GF₃ fraction isolated at the 0.86/1.15 M-sucrose interface was removed and immediately frozen at –20 °C. For antigen preparation the GF₃ fractions from several isolations were thawed, pooled and prepared in four sequential steps. (i) A membrane subfraction was prepared at high pH (100 mM-Na₂CO₃, pH 11.3) as described [6]. (ii) The membrane pellet was resuspended in a high-salt solution [500 mM-KCl/50 mM-Tris/HCl (pH 7.4)/5 mM-MgCl₂] and repelleted. (iii) The resulting pellet was resuspended in 0.25 M-sucrose to a protein concentration of 2 mg/ml, loaded on top of a continuous sucrose gradient (density 1.05–1.14 g/cm³) and centrifuged in an SW 40 rotor at 200 000 g_{av} for 18 h. The membrane band at a density of 1.098 g/cm³ was recovered, pelleted to remove the sucrose and proteinase inhibitors, and resuspended in phosphate buffered saline (PBS; 0.15 M-NaCl/2 mM-NaH₂PO₄/16 mM-Na₂HPO₄) (without proteinase inhibitors). (iv) The membrane preparation was extracted once with Triton X-114 according to the procedure of Bordier [7]. The total detergent phase (3.2 mg of protein derived from 30 rat livers) was divided into appropriate portions for immunization, frozen in liquid N₂ and stored at –80 °C.

Antisera

A polyclonal antiserum (coded 239) against a detergent-extracted Golgi membrane fraction was raised in specific pathogen-free rabbits, which received a total of 210 µg of protein during the immunization schedule. The protocol was that of Louvard *et al.* [8], with the first injection into the popliteal lymph node. Antisera to whole Golgi fractions and antibodies to expression vector pUEX fusion proteins were raised using the

Abbreviations used: NRK cells, normal rat kidney cells; VSV, vesicular stomatitis virus.

* Present address: Department of Clinical Biochemistry, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, U.K.

† Present address: Institut für Immunologie und Genetik, Abt. 710, DKFZ, Im Neuenheimer Feld 280, D-6900 Heidelberg, Federal Republic of Germany.

‡ Present address: Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, U.K.

§ Present address: University of Colorado Medical School, Department of Cellular and Structural Biology, B111, 4200 E 9th Ave., Denver, CO 80262, U.S.A.

|| Present address: Instituto di Istologia ed Embriologia, via Falloppio 50, I-35100 Padova, Italy.

¶ Present address: The Heart Research Institute, 145–147 Missenden Road, Camperdown, Sydney, New South Wales 2050, Australia.

* To whom correspondence should be addressed.

The nucleotide sequence data reported will appear in the EMBL, Gen Bank and DDBJ Nucleotide Sequence Databases under the accession number X53565.

same immunization schedule. In the latter case, the immunogen was purified by SDS/PAGE and extracted from gel slices by homogenization in PBS. A total of 200 µg of protein was used for the entire immunization schedule.

cDNA cloning

Poly(A)⁺ RNA from rat liver was prepared by acid guanidinium thiocyanate/phenol/chloroform extraction [9] and affinity chromatography on oligo(dT)-cellulose. cDNA was synthesized [10] using a random hexanucleotide (Pharmacia) as primer and a library of more than 6×10^5 independent clones was constructed in the expression vector pUEX1 [11] using adaptors [12]. Less than 10% of colonies in the rat liver library were without inserts and the average insert size was greater than 1 kb, as judged by *Bam*HI digestion of small-scale plasmid preparations. Expression of clones, antibody screening, affinity purification of antibodies on expressed fusion proteins and DNA hybridization studies were performed as previously described [13,14].

Immunofluorescence microscopy

Normal rat kidney (NRK) cells were grown on glass coverslips in Dulbecco's modified Eagle's medium/fetal calf serum to approx. 50% confluence before use. For immunofluorescence analysis the cells were washed $\times 3$ with PBS and then fixed and permeabilized using cold (-20°C) methanol for 5 min. After further washing with PBS, the cells were stained with rabbit antibodies and rhodamine-labelled goat anti-rabbit antibodies (Medac, Hamburg, Germany) essentially as previously described [15]. Surface localization of immunofluorescence was confirmed with paraformaldehyde-fixed cells in the absence of detergent permeabilization [15].

Immunoelectron microscopy

Immunoelectron microscopy of frozen sections from NRK cells and vesicular stomatitis virus (VSV)-infected 20°C -blocked NRK cells was carried out essentially as previously described [16].

Surface labelling and immunoprecipitation

NRK cells were labelled at 0°C with ^{125}I using Iodogen as a catalyst [17] as described by Herz *et al.* [18], or with Bolton-Hunter reagent (du Pont-NEN) as described by the manufacturer after resuspending 10^7 cells in 200 µl of 0.1 M-borate-buffered saline at pH 8.5. Cells were solubilized and immunoprecipitated using 2 µl of anti-(fusion protein) antiserum as described [18].

Gel electrophoresis and Western blots

SDS/PAGE and Western blot analysis were as previously described [14]; approx. 20 µg of membrane protein was loaded per gel track.

DNA sequence analysis

cDNA sequence was obtained using the dideoxy chain-termination method [19] and the Sequenase enzyme (United States Biochemical Corporation) by a combination of random fragment sequencing in the M13 derivative mp8 and double-stranded DNA sequencing from synthetic oligonucleotide primers [20].

Transient transfections

CV-1 cells were transiently transfected with DNA cloned into the late replacement vector pSV51 [21]. Cells were used 24 h after transfection. The tail-deleted mutant of TGN38 cDNA was prepared using the oligonucleotide 3'-GGTGTGTTTCT-AGAATCTAACGAAAACGG-5', by established techniques (Amersham mutagenesis kit), so that a stop codon was introduced

C-terminal to the membrane-spanning region, with the C-terminus being Ser rather than Lys at position 311.

RESULTS

A polyclonal antiserum (coded 239) was raised against membrane proteins of a rat liver Golgi fraction and used to screen 500,000 clones of a rat liver cDNA library prepared in the prokaryotic expression vector pUEX [11]. After colony purification, 71 positive clones were obtained with the mixed polyclonal serum. Seven epitope classes were identified by purification of the polyclonal serum on individual expressed fusion proteins as previously described for epitope mapping [22,23]. The use of pUEX, in which large amounts of fusion protein may be generated in a controlled fashion, facilitated

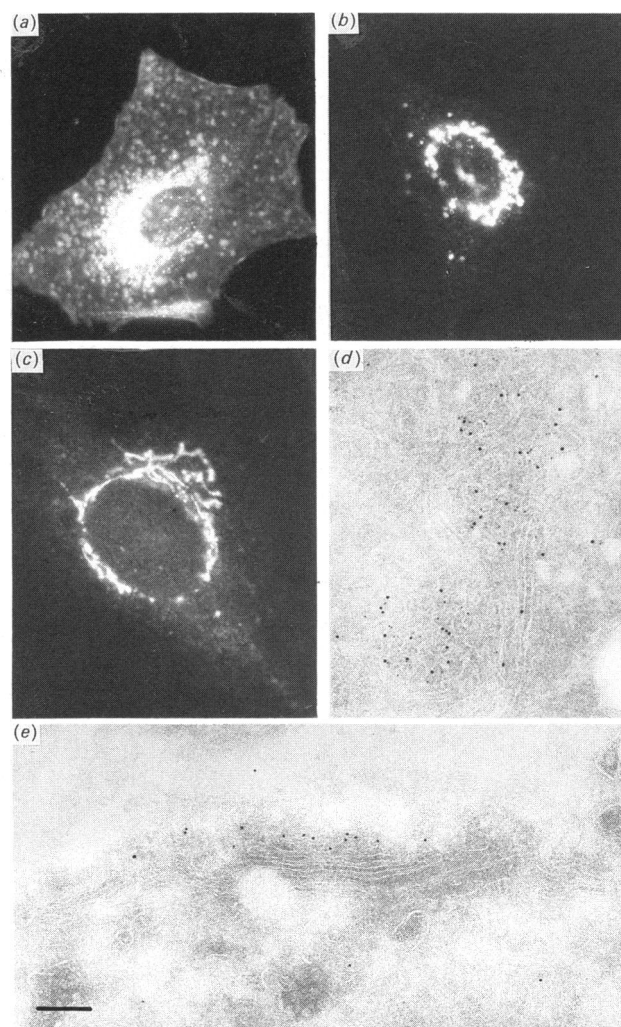


Fig. 1. Immunolocalization of TGN38 in NRK cells

(a-c) Indirect immunofluorescence using (a) anti-(Golgi membrane protein) serum; (b) affinity-purified antibodies eluted from a TGN38 (group A) fusion protein; and (c) antiserum against TGN38 (group A) fusion protein. (d, e) Immunoelectronmicroscopic localization of (d) thin frozen section of NRK cells labelled with antiserum against TGN38 (group A) fusion protein and 12 nm gold coupled to sheep anti-rabbit IgG, and (e) thin frozen section of VSV-infected NRK cells, maintained at 20°C , labelled with rabbit antiserum against TGN38 (group A) fusion protein and monoclonal anti-(VSV G protein) antibody followed by 12 nm gold coupled sheep anti-(rabbit IgG) and 5 nm gold coupled to sheep anti-(mouse IgG). Bar = 0.2 µm.

further characterization of the groups of clones. Western blotting provided the apparent M_r values of the products of cloned cDNAs and immunofluorescence studies determined their intracellular localization. The two most abundant groups of cDNA clones (accounting for 28 and 13 clones respectively) were shown by nucleotide sequence analysis to encode known antigenic membrane proteins of purified rat liver Golgi fractions, the polymeric Ig receptor [24] and 'microsomal' glutathione *S*-transferase [25]. This paper concerns a group of eight overlapping clones (group A) which we show to encode a membrane protein of the *trans*-Golgi network.

In Fig. 1 the indirect immunofluorescence staining of NRK cells with the anti-(Golgi membrane protein) antiserum is compared before and after affinity purification on expressed fusion protein from group A (Figs. 1a and 1b). An identical pattern of immunofluorescence to that seen with the affinity-purified antibody was obtained with a polyclonal antiserum raised against

the fusion protein used for affinity purification (Fig. 1c). This antiserum was also used for immunoelectron microscopy in NRK cells. Gold particles were only observed on the *trans* face of the Golgi complex in thin frozen sections (Fig. 1d). In cells infected with VSV and incubated at 20 °C before fixation to enlarge and define the *trans*-Golgi network [2], both the viral envelope VSV-G protein and the group A antigen co-localized to the *trans*-Golgi network (Fig. 1e).

The antiserum raised against the group A fusion protein was also used to immunoprecipitate biosynthetically labelled protein from NRK cells. When cells were labelled with [³⁵S]methionine and [³⁵S]cysteine, no immunoprecipitated protein was detected. This result was consistent with the absence of cysteine from the sequence of the protein (see below) and the presence of only a single methionine residue in the mature protein. We also tried ³⁵SO₄²⁻ labelling, since the *trans*-Golgi network is the site of tyrosine sulphation of some secreted and membrane proteins

```

1  CAGACTACAGGATGCAGTTCCTGGTTGCGTGTCTGCTGAGTGTGCGGTGGCGAGAGCTCTGCCATCCGCTTCTAAGCCAAACAACACATCAAGTGA
    M Q F L V A L L L L S V A V A R A L P S A S K P N N T S S E
    -17                                     1 * * 10
101 AAATAATCTCCGATTCAACCTTCCACGCCCTGCCTCTGGGGTGTATTTCTCAACAGGTAAAGACTAACAGACCGACTGATCAGAGGTGGAGTCG
    N N P P I Q P S T P L P P G V D I S Q Q V K T N R P T D Q R L E S
    20                                     30                                     40
201 GATAAAGAAGGCCAGGACAAGACGGTGGCTAGGACTTCGGCATCAGTTTCCAGTGGGGTGAATCTGCGACCAATCTCAATTTGGATGATTCAAAAAGC
    D K E G Q D K T V A R T S A S V S S G V E S A T N L N L D D S K K H
    50                                     60                                     70                                     80
301 ACCCCGAAACAGCAGATGCCAAATTGAAGGAAACTTTGCAGCAGCTTTTGCTGTGGACCCCAACAGGAAAGTCAGGTCAGAAGTTTACAAAAGATTC
    P E T A D A K L K E T L Q Q L L P V D P K Q E K S G Q K F T K D S
    90                                     100                                     110
401 AGGCAGCCCCACAGGAGGTGATTCCGACAATACCACAGGAGGTGATTCTAACAAGACCACAGGAGTTGATTCTGACAAGACATCAGGAGGTGATTCTAAC
    G S P T G G D S D N T T G G D S N K T T G V D S D K T S G G D S N
    120 *                                     *                                     140
501 AAGCCACAGGAAGTGATAATGACAAGCCACAGGAGGTGATTCTAACAAGCCACCTCCAAAGTCCCTCCAATACAGAGACTCCAAAAATTGACAAG
    K P T G S D N D K P T G G D S N K P T S K V P S N T E T P K I D K V
    150                                     160                                     170                                     180
601 TCCAGCTAAGTGAAGGACAGAAACCAACTTATTTCCAAACTGAATCTGGGGAAGAACTGGCAGGGGACTCTGACTTCTCTTTAAGCCAGAGAA
    Q L T E K G Q K P T L I S K T E S G E K L A G D S D F S L K P E K
    190                                     200                                     210
701 AGGAGATAAGTCTTCAGAGCCTACTGAAGATGTGGAACCAAGGAGATTGAAGAGGTGATACAGAGCCCAAGAGGTTCCGCACTTGAAGAAGAGAAT
    G D K S S E P T E D V E T K E I E E G D T E P E E G S P L E E E N
    220                                     230                                     240
801 GAAAAGGTGTGGGCCCTTCCTCCAGTGAGAACCAAGAGGGGACACTTACAGATTCTATGAAGGATGAGAAGGATGACCATATAAGGACAATTCTGGAA
    E K V L G P S S S E N Q E G T L T D S M K D E K D D H Y K D N S G N
    250                                     260                                     270 *
901 ACACCAGTGCAGAAAGCAGCCACTTCTTTGCTTATCTGGTGACCGTGTCTTCTTGTGCTGTCTCTATATTGCTTACCACAACAAACGAAAGATTAT
    T S A E S S H F F A Y L V T A A V L V A V L Y I A Y H N K R K I I
    290                                     300                                     310
1001 TGCTTTTGCCCTGGAAGGAAAAAGATCCAAAGTCTACTCGGAGGCCAAAGGCCAGTGACTACCAACGTTTGAACCTAAAGCTTTGATTTTGTCTG
    A F A L E G K R S K V T R R P K A S D Y Q R L N L K L *
    320                                     330                                     340
1101 CAAGAACCTTGTCTCCCTGCTGATTGTTTCCGAATCAAGAGAAATGAAAGACTGACCTGTGGGGTTTGGTGGCAAGTCTGGGCTGGCAGAGATGGGA
1201 TTGCTTCAGGTTTTCGACCTGCACCTTGATGACAGTCCCTGTGGTGGTGTATTTTACTTTTGTGTTCCCTTATGTGATGTGTCCATCTTCCTCCATT
1301 CTTGGAGTACAAGCAGAGGGAGTGTGTGGAACATGGCTGTGACCAAGGAAGACCCACCTTCCTTAGGCAGGCCACACTATGGACACCCCTCCATAA
1401 CCCCCTTATCTTTGCACTCTAGGAGCAAACTACAGAGGTGCTTTCTAGACAAGATGTAGAAGATGATTCTGTGGGATGCTGATGGGCAGAGACAGGGTT
1501 GAAATCTTCTCCACAGAGGCCTCAGCTGCAGACCAGACTAGCAAGTTAAAGAGTGTCAAGGTCAGTCCGCTGCAGCTCTGTATATCTCTGAAATTT
1601 GCACTCGAGTTGTCTAGCACAAGAGGACTCAGTCTCTGAAAGTAAATCACTGTCGACATGGGATTTTTTTTTTTTAAATATGGTGAATGATTATTTG
1701 AACTGAACCAAGTGTGTTCTTGCATACAGTCAACCAATAGGAAGTCTCTGGAAGAAGCAAGGTTTATGCCTTGTGGAAGAAAACACATAGGCGCTTGGT
1801 ATACAGAAAGAGTTTCTCAGTCACTGTTTGTCTTCATGGTCTGCAGTAGTAAAGATAGACATTAGAAACTCCGTGCGTTCTCCACATTCTCCACAC
1901 GTTTTGAATGCCAGGGAATGAGTAATATCTTCTCTATCTGTCTAGTACTCATGAATTTGAAGAAACCAAGTAGGTTCCCTGAGATAACCAAGAT
2001 GTCTTGAATGTCTGTCTTAAAGTTTATTCAGGTTCTTCTGAAAGTAAATCACTGTCGACATGGGATTTTTTTTTTTTAAATATGGTGAATGATTATTTG
2101 AAATCTGAGTCTGTGCTTGTGAGTGTGAGAGTTGAAGAAGACCTGGTCTTGGAGCTCTTCTCCATGCTGACTGTAGCATGTGCCTGAAGTTTGTATGCTTGG
2201 CAGGTGCACTCTAGAAATGAGGTGCTGCTGCTTCCGACTGTGCATAAGTCTGCTGTTACTTTTCTAGCTGAGAAGCCGTTTTCAGTCTCCC
2301 TCAGCTTTCGTAACCTTCTTCTCAGTTCCCACTCTTCCATCTCTCTCTCTCTTCTAGCCCTGTTGTACACCTCTGATGAAACCTTCTCTTCTAGCC
2401 TAGTGTGATGCTCCCTACCCGTTCCAGTTGGTGTCTTCTTATTCACCTTCCCTCTCTCCCTGACATTGTCAGGTTTGGCTTGGTCTGTTCTCTCT
2501 GTTTCCTAAGGTTACTCAGTTAACAGATACCAATAAAACAGCCCATTTTACAGCTTATAACACTTGGTAATGATGGATGCTGGGGTCTGTGCATGTT
2601 GACATCTAGGTGCTCTTGCCTTGTCTTCTTATGCGTGGGGGAGGGTGCCTCAGTCTCTTGTCTTGTACTGTCGATCCATTATT
2701 TATCCTTCTGTGACCTCACATTCTTGTGCTCTTCCCAAGAACCA 2749

```

Fig. 2. Nucleotide and deduced amino acid sequence of TGN38

Underlined groups of amino acids show, from the *N*-terminus, the putative signal peptide, six consecutive repeats of 8 amino acids, and the putative membrane-spanning segment. Residues marked with an asterisk are potential sites for attachment of *N*-linked oligosaccharides. The numbering of amino acids is based on the most likely cleavage point of signal peptidase [27]

[26], but observed no immunoprecipitable material. However, after labelling with a mixture of ^{14}C -labelled amino acids, immunoprecipitated protein ran on SDS/PAGE as a broad band of M_r 85000–95000 (results not shown), consistent with the apparent mobility of protein on Western blots (see Fig. 3). Immunoprecipitation from a total membrane fraction labelled with ^{125}I (using either Iodogen or Bolton–Hunter reagent) gave a similar broad band on SDS/PAGE of M_r 85000–95000, though when only the cell surface of intact NRK cells was labelled with these reagents, immunoprecipitated protein was not detected (results not shown).

The sequence of the 2749 bp insert in one of the group A clones was determined by dideoxy sequencing of both strands of the cDNA (Fig. 2). None of the other original seven clones, or clones obtained by hybridization screening of the original library, extended this sequence. Within the nucleotide sequence an open reading frame of 357 amino acids was found, commencing with an initiating methionine codon 12 bases downstream of the cloning adaptors. A hydrophobicity plot of the derived amino acid sequence showed only two hydrophobic regions corresponding to a predicted [27] *N*-terminal signal peptide of 17 amino acids and a putative *trans*-membrane region towards the C-terminus. The short length of the hydrophobic sequence at residues –17 to –1 and the absence of a stop transfer sequence indicate that this is a leader peptide. The predicted *trans*-membrane region commences with a histidine and is terminated by an asparagine and three basic amino acids, which are usual features of the *trans*-membrane region of integral membrane proteins [28], indicating that the protein encoded is a type 1 *trans*-membrane protein [29], with a C-terminal cytoplasmic tail of 33 amino acids.

Analysis of the derived primary sequence using sequence comparison programs revealed the existence of a unit of amino acids repeated consecutively six times, commencing with the proline at position 116 (Fig. 2). At the DNA level the repeat is even more highly conserved, as most of the amino acid changes are caused by single base changes. Only 13 out of 144 bases are changed when the second repeat sequence is compared with the others. Secondary structure predictions [30] using the derived amino acid sequence were only able to find one region in the luminal domain which could form an α -helix with high probability (amino acids 82–96), and no regions of β -pleated sheet. A noticeable feature of the primary sequence is a relatively proline-rich region (8 out of 39 residues) at the *N*-terminus of the predicted mature polypeptide. The relatively high content of proline and glycine residues in the luminal portion of the protein suggests that rather little of the protein is likely to be folded in α -helix and β -sheet structures.

The mature polypeptide has a predicted M_r of 38307, in contrast to the apparent M_r on SDS/PAGE of 85000–95000 (Fig. 3), suggesting that the protein is highly glycosylated and/or anomalously folded. Since there are only five consensus sites for *N*-glycosylation *N*-terminal to the predicted *trans*-membrane region (Fig. 2), the protein could also be post-translationally modified with *O*-linked carbohydrates. The clusters of serine-rich sequence and poor staining with protein stains on gels would be consistent with this. Together with an acid-rich region with a net negative charge of –18 in 57 residues (amino acids 219–276) adjacent to the predicted *trans*-membrane segment, the high level of predicted glycosylation must lead to a very polar structure. From the size of the predicted mature polypeptide (M_r 38307), and its intracellular location, we propose the name TGN38.

The difficulties encountered in metabolically labelling TGN38 meant that classical enzyme analysis of its glycosylation was impractical. Thus, in order to confirm (i) the localization of TGN38 and (ii) that the isolated cDNA was capable of giving

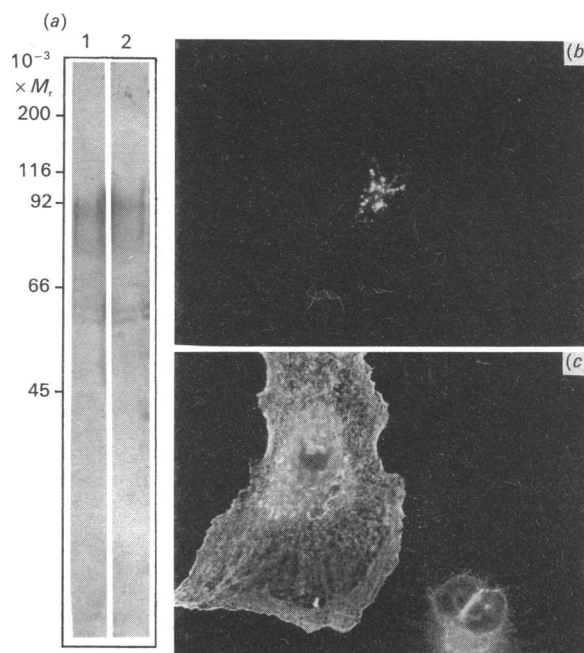


Fig. 3. Western blots and immunofluorescence of transfected CV-1 cells

Western blots using antiserum raised against group A (TGN38) fusion protein. Lane 1, rat liver membrane pellets (100000 g, 1 h, [14]); lane 2, lysate from CV-1 cells transiently transfected with TGN38 full-length cDNA in pSV51. (b, c) Indirect immunofluorescence, using antiserum raised against TGN38 fusion protein as primary antibody, of CV-1 cells transiently transfected with pSV51 containing TGN38 full-length cDNA (b) or TGN38 cDNA in which the putative cytoplasmic tail was deleted (c). Representative cells remaining on the coverslip are shown in each case. Mock-transfected CV-1 cells showed no staining by Western blotting or immunofluorescence (results not shown).

rise to M_r 85000–95000 molecules, transient transfections of CV-1 cells were performed. CV-1 cells were selected since the antiserum raised against TGN38 fusion protein did not react with these primate cells. Western blotting of CV-1 cells transiently transfected with the full-length TGN38 cDNA, encoding a predicted M_r 38307 open reading frame, in a simian virus 40 late-replacement vector showed a band of M_r 85000–95000 (Fig. 3), demonstrating that the product of the expressed cDNA could be post-translationally modified to produce a mature protein of the expected size *in vivo* (Fig. 3). Immunofluorescence experiments localized TGN38 to the Golgi region of all the transfected cells at 24 h post-transfection (Fig. 3). At 48 h post-transfection, when expression reaches a peak level [21], most fluorescent cells had fallen off the coverslips and died. These cells left behind a footprint which was strongly immunofluorescent with TGN38 antisera, suggesting a surface location of TGN38. Control transfected cells expressing surface influenza virus haemagglutinin did not fall off the coverslips even after 48 h of expression, indicating that the change in adherence was caused by over-expression of TGN38. By counting immunofluorescent cells plus footprints at 24 h post-transfection, 50% of transfected cells showed only Golgi region localization of stain, 30% also showed cell surface localization and 20% of cells had fallen off the coverslips. The data show that TGN38 retention in the Golgi complex is conserved across species, even though TGN38 epitopes are not, and that over-expression of TGN38 can overwhelm the retention system. In contrast, when cells were transfected with a construct in which the predicted cytoplasmic tail was deleted, most of the transfected cells (85%) had fallen off the glass coverslips by 24 h after transfection, and of the remaining

adherent immunofluorescent cells, > 90 % showed surface expression and additional intracellular patterns of staining (Fig. 3), providing evidence that Golgi retention signals are contained within the cytoplasmic tail.

DISCUSSION

The present work provides a strategy for the molecular cloning of membrane proteins from intracellular organelles when neither antibodies to specific proteins nor oligonucleotide probes based on known amino acid sequences are available. The strategy can be used to obtain specific probes for individual components of any mixture of proteins of biological interest and is a development of a method used to confirm antibody-positive clones in expression cDNA libraries [31], which has also been used for epitope mapping [22]. By using the bacterial expression vector pUEX [11], in which large amounts of fusion protein may be generated in a controlled fashion, the affinity purification step is simplified. The affinity-purified antibodies may be used for determining the size and location of the protein in the cell. In addition, monospecific antisera may be produced against defined regions of the protein, without recourse to protein purification or monoclonal antibody production, by immunizing rabbits with fusion proteins [32]. By virtue of the clonal nature of bacterial fusion proteins the antibodies raised against the cloned epitopes (ACEs) are particularly useful for immunofluorescence and immunoelectron microscopy because of the lack of cross-reaction with other eucaryotic proteins [32].

Using the strategy described we have isolated cDNA clones coding for the pIg receptor [24] and 'microsomal' glutathione *S*-transferase [23], two known integral membrane proteins of Golgi fractions, and in addition have characterized and cloned a resident membrane protein of the *trans*-Golgi network, TGN38. TGN38 is the first integral membrane protein restricted to the *trans*-Golgi network to be defined and the first type 1 membrane protein of the Golgi complex to be cloned. It has properties similar to a protein characterized with a monoclonal antibody coded 18B11 [33] which was found to react with a *trans*-Golgi integral membrane protein (GIMP₁). Both proteins have an *M_r* of 85000–100000, are probably heavily *O*-glycosylated with a polypeptide chain *M_r* of approx. 40000, and are difficult to label with [³⁵S]cysteine or [³⁵S]methionine. The sequence of TGN38 is consistent with a structural rather than enzymic function, having little predicted secondary structure and being unusually polar with a high level of acidic and glycosylated amino acids.

The address signals on TGN38 which direct it to the *trans*-Golgi network and cause its retention there despite traffic pathways to other subcellular locations are clearly of great importance. For endocytosed receptors [34–36], a resident lysosomal membrane protein [37] and endoplasmic reticulum transmembrane proteins [38], the cytoplasmic tail has been shown to contain the signals determining intracellular location. Intriguingly, although it is not normally present on the cell surface, TGN38 contains in residues 333–337 a structural feature that has been suggested by Vega & Strominger [39] to constitute a specific signal for endocytosis, though it does not have the NPXY sequence thought to be important by Chen *et al.* [40]. The transient transfection experiments conducted in the present study provide preliminary evidence that the cytoplasmic tail of TGN38 contains important information for correct intracellular targeting and retention of this membrane protein, though the variable level of expression in the transfected cells indicates the need to construct permanently transfected cloned cell lines to study further its targeting function. It is of interest to note that the 33-amino-acid cytoplasmic tail of TGN38 is very basic in character

with a net positive charge of +9. This is similar to the cytoplasmic tail of the G1 membrane glycoprotein of Uukuniemi virus which has been shown to bud into the Golgi complex [41]. An interesting feature of this region of positive charge in the viral protein (residues 444–453 [41] similar to residues 318–327 of TGN38) is that it is predicted to form an amphipathic α -helix.

We thank John Dickson for technical assistance, and Gareth Griffiths and Heinz Horstmann for performing electron microscopy. G. B. was supported by the Royal Society and J.P.L. was funded in part from grants from the Alexander von Humboldt Foundation and the Ciba-Geigy Trust.

REFERENCES

- Farquhar, M. G. (1985) *Annu. Rev. Cell Biol.* **1**, 447–488
- Griffiths, G. & Simons, K. (1986) *Science* **234**, 438–443
- Orci, L., Ravazzola, M., Amherdt, M., Perrelet, A., Powell, S. K., Quinn, D. L. & Moore, H.-P. H. (1987) *Cell* **51**, 1039–1051
- Cooper, M. S., Cornell-Bell, A. H., Chernjavsky, A., Dani, J. W. & Smith, S. J. (1990) *Cell* **61**, 135–145
- Howell, K. E., Ito, A. & Palade, G. E. (1978) *J. Cell Biol.* **79**, 581–589
- Howell, K. E. & Palade, G. E. (1982) *J. Cell Biol.* **92**, 822–832
- Bordier, C. (1981) *J. Biol. Chem.* **256**, 1604–1607
- Louvard, D., Reggio, H. & Warren, G. (1982) *J. Cell Biol.* **92**, 92–107
- Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159
- Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269
- Bressan, G. M. & Stanley, K. K. (1987) *Nucleic Acids Res.* **15**, 10056
- Haymerle, H., Herz, J., Bressan, G. M., Frank, R. & Stanley, K. K. (1986) *Nucleic Acids Res.* **14**, 8615–8624
- Stanley, K. K. & Luzio, J. P. (1984) *EMBO J.* **3**, 1429–1434
- Brake, B., Braghetta, P., Banting, G., Luzio, J. P. & Stanley, K. K. (1990) *Biochem. J.* **267**, 631–637
- Stanley, K. K., Burke, B., Pitt, T., Siddle, K. & Luzio, J. P. (1983) *Exp. Cell Res.* **144**, 39–46
- Griffiths, G., Pfeiffer, S., Simons, K. & Matlin, K. (1985) *J. Cell Biol.* **101**, 949–964
- Fraker, P. J. & Speck, J. C. (1978) *Biochem. Biophys. Res. Commun.* **80**, 849–857
- Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H. & Stanley, K. K. (1988) *EMBO J.* **7**, 4119–4127
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Chen, E. Y. & Seeburg, P. H. (1985) *DNA* **4**, 165–170
- Huylebroeck, D., Maertens, G., Verhoeven, M., Lopez, C., Raeymakers, A., Jou, W. M. & Fiers, W. (1988) *Gene* **66**, 163–181
- Stanley, K. K. & Herz, J. (1987) *EMBO J.* **6**, 1951–1957
- Luzio, J. P., Brake, B., Banting, G., Howell, K., Bressan, G., Braghetta, P. & Stanley, K. K. (1990) *Biochem. Soc. Trans.* **18**, 148–149
- Banting, G., Brake, B., Braghetta, P., Luzio, J. P. & Stanley, K. K. (1989) *FEBS Lett.* **254**, 177–188
- DeJong, J. L., Morgenstern, R., Jornvall, H., DePierre, J. W. & Tu, C.-P. D. (1988) *J. Biol. Chem.* **263**, 8430–8436
- Huttner, W. B. (1988) *Annu. Rev. Physiol.* **50**, 363–376
- von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690
- Rao, J. K. M. & Argos, P. (1986) *Biochim. Biophys. Acta* **869**, 197–214
- Wickner, W. T. & Lodish, H. F. (1985) *Science* **230**, 400–407
- Gilbrat, J. F., Garnier, J. & Robson, B. (1987) *J. Mol. Biol.* **198**, 425–443
- Weinberger, C., Hollenberg, S. M., Ong, E. S., Harmon, J. M., Brower, S. T., Cidlowski, J., Thompson, E. B., Roesenfeld, M. G. & Evans, R. M. (1985) *Science* **228**, 740–742
- Tooze, S. A. & Stanley, K. K. (1986) *J. Virol.* **60**, 928–934
- Yuan, L., Barriocanal, J. G., Bonifacio, J. S. & Sandoval, I. V. (1987) *J. Cell Biol.* **105**, 215–227
- Breitfeld, P. P., Casanova, J. E., Simister, N. E., Ross, S. A., McKinnon, W. C. & Mostov, K. E. (1989) *Curr. Opin. Cell Biol.* **1**, 617–623
- Davis, C. G., van Driel, I. R., Russell, D. W., Brown, M. S. & Goldstein, J. L. (1987) *J. Biol. Chem.* **262**, 4075–4082

36. Lazarovits, J. & Roth, M. (1988) *Cell* **53**, 743–752
37. Matthews, P. M. & Fambrough, D. M. (1988) *J. Cell Biol.* **107**, 2479a
38. Nilsson, T., Jackson, M. & Peterson, P. A. (1989) *Cell* **58**, 707–718
39. Vega, M. A. & Strominger, J. L. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2688–2692
40. Chen, W.-J., Goldstein, J. L. & Brown, M. S. (1990) *J. Biol. Chem.* **265**, 3116–3123
41. Rönnholm, R. & Pettersson, R. F. (1987) *Virology* **160**, 191–202

Received 1 May 1990; accepted 24 May 1990