Transcytosis-associated protein (TAP)/p115 is a general fusion factor required for binding of vesicles to acceptor membranes

MARGARIDA BARROSO, DAVID S. NELSON, AND ELIZABETH SZTUL*

Department of Molecular Biology, Princeton University, Princeton, NJ 08544

Communicated by James E. Rothman, Memorial Sloan-Kettering Cancer Center, New York, NY, September 8, 1994 (received for review August 25, 1994)

ABSTRACT Transcytosis-associated protein (TAP) is found on transcytotic vesicles (TCVs) and is required for their fusion with the target membrane. We developed a cell-free assay capable of differentiating targeting/binding of TCVs to membrane from later fusion events. We found that TAP mediates stable association of TCVs with the target membrane. The sequence of rat liver TAP (959-amino acid open reading frame) encodes a protein that contains (i) an Nterminal region (amino acids 1-649), (ii) an internal region with several coiled-coil stretches (amino acids 650-930), and (iii) a C-terminal acidic region (amino acids 931-959). Comparisons between TAP and other sequences indicate that TAP is identical to p115, a protein involved in cis to medial Golgi transport, and homologous to Uso1p, a yeast protein involved in endoplasmic reticulum to Golgi transport. Our findings suggest that TAP/p115/Usop1 is a general factor acting within the secretory and endocytic pathways to bind transport vesicles prior to membrane fusion.

Vesicular transport (transcytosis) of proteins from the basolateral to apical plasma membrane (PM) of polarized epithelial cells has been extensively studied (1, 2). We have designed a cell-free assay reconstituting the last step of transcytosis, the fusion of transcytotic vesicles (TCVs) with the apical PM (3-6), and have identified two proteins, *N*-ethylmaleimidesensitive fusion protein (NSF) and transcytosis-associated protein (TAP), which are required for fusion (3, 7). We now further analyze TAP's role by establishing a modified *in vitro* assay that differentiates between the targeting/binding stage and subsequent steps of the fusion reaction. We show that TAP is required for the binding of TCVs to the PM.

We have cloned and sequenced TAP.[†] TAP's sequence is identical to that of p115 (8), a protein originally shown to be required for *cis* to *medial* Golgi transport (9) and subsequently for an uncoupled reaction measuring exclusively intra-Golgi fusion (10). TAP also shares highly conserved regions with a yeast protein Uso1p, which among other pleiotropic effects has been implicated in endoplasmic reticulum (ER) to Golgi traffic (11).

MATERIALS AND METHODS

SDS/PAGE and Immunoblotting. Samples were processed for SDS/PAGE and immunoblotting as described (3, 12). Immunoblots were processed by chemiluminescence and filters were exposed to x-ray film.

Cell-Free Assay. In vivo radiolabeled donor fractions and unlabeled target fractions and cytosol were prepared from rat livers as described (3). Fusion assays and analysis of polymeric IgA receptor (pIgA-R) were performed as described (3). In some experiments, reaction mixtures were centrifuged, and supernatant and pellets were separated. For the immunodepletion experiments, anti-TAP antibodies were purified from culture supernatants using protein G-Sepharose (Pharmacia) and coupled to Affi-Gel 10 (Bio-Rad) support matrix. TAP-immunodepleted cytosol was prepared by incubating cytosol with 4G2 monoclonal antibodies (mAbs) covalently bound to the Affi-Gel support at 4°C.

TAP Purification. Rat livers were homogenized in 25 mM Tris·HCl, pH 7.4/250 mM sucrose/150 mM KCl/1 mM dithiothreitol. The homogenate was centrifuged, the supernatant was collected, and ammonium sulfate was added to a concentration of 40%. The pellets were recovered, resuspended in TD (25 mM Tris-HCl/1 mM dithiothreitol) and diluted to a conductivity equal to TD containing 150 mM KCl (150KTD). This material was loaded onto a DEAE-Sepharose (Pharmacia) column equilibrated with 150KTD. The column was eluted with a gradient of 150-600KTD. Fractions were analyzed by immunoblotting with polyclonal anti-TAP antibody. A concentrated pool was loaded onto a Superose 6 (Pharmacia) column equilibrated with 150KTD. TAP-containing fractions were pooled and loaded onto a Mono Q (Pharmacia) column equilibrated with 150KTD. The column was eluted with a gradient of 150-600KTD.

mAbs. Mono Q fractions were used as antigen to generate antibodies in mice (13). Screening of hybridomas was performed by immunoblotting. Two cell lines secreting anti-TAP antibodies were used in this study.

Cloning and Sequencing. We used degenerate oligonucleotide primers corresponding to amino acid residues QHD-NIVTH (amino acids 647–654) and TQQASQIQ (amino acids 688–695) to amplify a cDNA fragment from reversetranscribed total liver RNA (SuperScript kit; GIBCO/BRL). The resulting PCR fragment was cloned and sequenced and was 84% homologous to a human partial cDNA sequence (GenBank accession no. Z24991). The fragment was used to screen an oligo(dT)-primed λ ZAP II cDNA library from a rat liver cell line. A 5'-terminal restriction fragment of a partial cDNA clone was used to screen a random-primed rat liver cDNA library in pUEX. A single clone (\approx 1.6 kb) was isolated that contained the 5' end of the TAP coding sequence. Standard procedures were used for cloning and sequencing (14).

RESULTS AND DISCUSSION

Anti-TAP mAbs. Previously, we demonstrated that cytosol immunodepleted of TAP failed to support fusion of TCVs with apical PM (3), but these experiments used a polyspecific polyclonal antibody raised against purified TCVs (15). To

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PM, plasma membrane; TCV, transcytotic vesicle; NSF, *N*-ethylmaleimide-sensitive fusion protein; TAP, transcytosisassociated protein; ER, endoplasmic reticulum; pIgA-R, polymeric IgA receptor; mAb, monoclonal antibody; ORF, open reading frame; SNAP, soluble NSF attachment protein; SNARE, SNAP receptor. *To whom reprint requests should be addressed.

[†]The sequence reported in this paper has been deposited in the GenBank data base (accession no. U15589).

unequivocally show that TAP is required for transcytotic fusion, we used mAb immunodepletion to prepare TAP-free cytosol. Such cytosol was then used in the cell-free assay to determine whether it could support fusion and, if not, whether addition of purified TAP could reconstitute fusion.

To test the mAbs, cytosolic proteins were immunoblotted with either the mAbs or the polyclonal anti-TAP antibody. As shown in Fig. 1*A*, for the 5D6 mAb, only TAP was detected (lane 2), while the polyclonal anti-TAP antibody reacted with TAP but also detected other minor proteins (lane 3). To ensure that the mAbs recognize the same protein as the polyclonal anti-TAP antibody, cytosol was subjected to immunoprecipitation with 4G2 and 5D6 mAbs or without antibodies and then immunoblotted with the polyclonal anti-TAP antibody. As shown in Fig. 1*B*, TAP was present in the +mAb immunoprecipitates (lanes 1 and 4) but was absent from the -mAb immunoprecipitate (lane 3). These results indicate that the mAbs immunoprecipitate TAP. The 4G2 mAb was used to immunopurify TAP from cytosol (Fig. 1*C*).

TAP Is Required for Transcytotic Fusion. 4G2 mAb was used to immunodeplete cytosol (removing >95% of TAP), which was then tested in the transcytotic cell-free assay. The donor fraction was prepared from in vivo radiolabeled rats and contains TCVs that carry radiolabeled 120-kDa pIgA-R. The target fraction was prepared from unlabeled rats and contains an inside-out apical PM that has an exoprotease capable of cleaving pIgA-R to an ≈90-kDa fragment. When donor and target fractions were mixed with cytosol and an ATPregenerating system and incubated at 37°C, fusion occurred as measured by the cleavage of 120-kDa pIgA-R to the ≈90-kDa fragment (Fig. 2, lane 1). Under control conditions, 50-80% fusion is routinely observed. A similar result (lane 4) was observed when mock-depleted cytosol was used. However, no fusion was observed (lane 7) when TAP-depleted cytosol was tested. Addition of purified TAP (lane 8) reversed the inhibition to the same extent as addition of untreated cytosol (lane 9). Addition of purified TAP or untreated cytosol to reaction mixtures containing untreated cytosol (lanes 2 and 3) or to reaction mixtures containing mock-depleted cytosol (lanes 5 and 6) did not increase fusion efficiency.

Dissection of the Transcytotic Assay. To address the role of TAP at a more refined level, the cell-free assay was dissected into distinct stages by changing the end-point readout; instead of measuring fusion, we assayed the association of TCVs with the target PM. We developed a velocity centrifugation protocol (based on the large difference in size between the donor TCVs and the target PM), which can separate free TCVs from

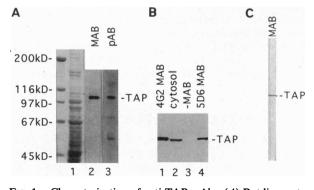


FIG. 1. Characterization of anti-TAP mAbs. (A) Rat liver cytosol was processed by SDS/PAGE and stained with Coomassie blue (lane 1) or immunoblotted with 5D6 mAb (lane 2) or with polyclonal anti-TAP antibody (lane 3). (B) Cytosol was immunoprecipitated with 4G2 and 5D6 mAbs or without antibody. Immunoprecipitated proteins were immunoblotted with the polyclonal anti-TAP antibody. Cytosol was included to indicate position of TAP. (C) 4G2 mAb was used to immunopurify TAP from cytosol. A Coomassie blue-stained SDS/ polyacrylamide gel of purified material is shown.

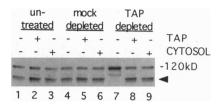


FIG. 2. TAP is required for transcytotic fusion. Fusion reaction mixture containing untreated cytosol, mock-depleted cytosol, or TAP-depleted cytosol were incubated at 37°C for 60 min (lanes 1, 4, and 7) or were supplemented with purified TAP (lanes 2, 5, and 8) or with untreated cytosol (lanes 3, 6, and 9) prior to incubation at 37°C. pIgA-R was immunoprecipitated and analyzed by SDS/PAGE and fluorography.

the target PM. The donor or the target fraction was incubated in a complete reaction mixture at 37°C for 60 min and then centrifuged. The extent of sedimentation of the donor fraction in the absence of the fusing partner was analyzed by examining the resulting pellet and supernatant for the content of the 120-kDa form of pIgA-R. Pelletability of the target fraction in the absence of donor was assayed by examining the pellet and supernatant fractions for content of the apical PM protein dipeptidyl-peptidase IV. As shown in Fig 3A (untreated bars), the centrifugation results in minimal (~10%) pelleting of TCVs, but as shown in Fig. 3B (untreated bars), it results in extensive (~65%) sedimentation of target PM.

Donor or target fractions were supplemented with reaction mixtures containing TAP-depleted cytosol or an ATPdepleting system at 37°C for 60 min and centrifuged. As shown

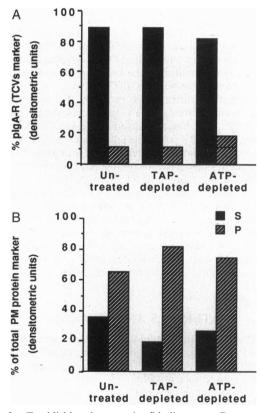


FIG. 3. Establishing the targeting/binding assay. Donor or target fraction was incubated in a reaction mixture containing untreated cytosol (untreated bars), TAP-depleted cytosol (TAP-depleted bars), or untreated cytosol and ATP-depleting system (ATP-depleted bars). Reaction mixtures were incubated at 37° C for 60 min and then separated by centrifugation. (A) Amount of pelleted and nonpelleted donor TCVs was assayed by immunoblotting with anti-pIgA-R antibodies. (B) Pelleting characteristics of the target PM was determined by immunoblots with anti-dipeptidyl-peptidase IV antibodies.

in Fig. 3A (TAP-depleted and ATP-depleted bars), there was no change in the sedimentation characteristics of TCVs. Similarly, as shown in Fig. 3B (TAP-depleted and ATPdepleted bars), the pelletability of the target apical PM under such conditions is similar to that seen in the control reaction.

TAP Is Required for Binding of TCVs to Target Membrane. To examine the stage at which TAP is required for transcytotic fusion, we analyzed the centrifugation behavior of three fusion reaction mixtures: (i) control reaction mixture (containing donor and acceptor, untreated cytosol, and an ATPregenerating system); (ii) reaction mixture containing TAPdepleted cytosol; and (iii) reaction mixture containing an ATP-depleting system. As shown in Fig. 4A, fusion occurs in the control (lane 1) reaction mixture but is inhibited in the TAP-depleted (lane 2) and ATP-depleted (lane 3) reaction mixtures. An aliquot of each reaction mixture was subjected to centrifugation and the resulting supernatant and pellet fractions were analyzed by SDS/PAGE and fluorography to determine their content of radiolabeled pIgA-R. As shown in Fig. 4B (lanes 1 and 2), \approx 82% of pIgA-R present in the control reaction mixture was recovered in the pellet, with the remainder found in the supernatant. The pIgA-R recovered in the pellet consists predominantly of the 90-kDa fragment [which is released into the sealed lumen of the target PM (3) and is therefore pelletable], and low levels of the uncleaved 120-kDa form. The $\approx 20\%$ of 120-kDa pIgA-R recovered in the pellet is higher than the $\approx 10\%$ that pellets in the absence of target PM (Fig. 3A, untreated bars) and might be due to pIgA-R present in TCVs bound to the PM but not yet fused.

Different results were obtained when a reaction mixture containing TAP-depleted cytosol was incubated at 37°C for 60 min and then subjected to centrifugation. As shown in Fig. 4B (lanes 3 and 4), the majority of pIgA-R remained in the supernatant (the $\approx 10\%$ of pIgA-R found in the pellet is analogous to the background level seen in Fig. 3A, untreated bars). All of the pIgA-R was in the 120-kDa form. These results suggest that TAP is required for binding of TCVs to the target PM or, alternatively, that the centrifugation conditions are too stringent and dissociate bound TCVs from the target PM if fusion is prevented. To distinguish between these possibilities, we examined the distribution of pIgA-R when an ATPdepleted reaction mixture was centrifuged (16). As shown in Fig. 4B (lanes 5 and 6), the distribution of pIgA-R was $\approx 50\%$ in the pellet and $\approx 50\%$ in the supernatant, indicating that TCVs associate with the target PM in the absence of fusion. The pIgA-R recovered in the pellet is in the 120-kDa form. The fact that only $\approx 40\%$ ($\approx 50\%$ minus $\approx 10\%$ for background) of pIgA-R was associated with the target PM in the absence of ATP (lane 6) but $\approx 60\%$ of TCVs fuse with the target PM in the presence of ATP (lane 2) might be due to a limiting number

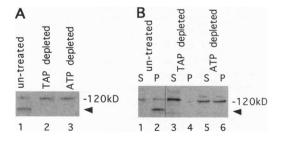


FIG. 4. TAP is required for targeting/binding of TCVs to the target apical PM. (A) Reaction mixtures containing untreated cytosol (lane 1), TAP-depleted cytosol (lane 2), or ATP-depleting system (lane 3) were incubated at 37° C for 60 min. Reactions were terminated and pIgA-R was immunoprecipitated and analyzed by SDS/PAGE and fluorography. (B) Analogous reaction mixtures were centrifuged and the pellet (P) and supernatant (S) fractions were immunoprecipitated with anti-pIgA-R antibodies. Immunoprecipitates were analyzed by SDS/PAGE and fluorography.

of binding sites on the target PM, suggesting that in the course of a normal reaction, the same PM sites are used for fusion of multiple TCVs.

Molecular Analysis of TAP. The sequence of full-length TAP cDNA (3860 nucleotides) reveals an open reading frame (ORF) of 959 amino acids, 96 nucleotides of 5' noncoding sequence, and a 3' untranslated sequence of 884 nucleotides, including a termination codon (nucleotides 2878–80, triple asterisks), a polyadenylylation signal (nucleotides 3721–3726, boldface), and the beginning of the poly(A) tail (nucleotide 2748) (Fig. 5). The sequences of five TAP tryptic peptides obtained by microsequencing are present within the cDNA sequence (boxed residues), indicating that the ORF encodes rat liver TAP.

There are two in-frame ATGs (nucleotides 1–3 and 22–24, marked with single asterisks). The context of the first ATG does not conform to the consensus sequence for eukaryotic translation initiation (17) except for the conserved purine at position -3. The second ATG conforms perfectly to the translation start consensus sequence.

The predicted secondary structure of TAP includes (*i*) an N-terminal region (amino acids 1–650); (*ii*) an internal α -helical region (amino acids 650–930) containing several coiledcoil domains; and (*iii*) a C-terminal acidic region (amino acids 931–959). The N-terminal region contains two proline-rich regions at positions 357–363 and 597–604 (double asterisks). The N-terminal 70-amino acid region is basic, with a calculated pI of 10.22; the C terminus, with a calculated pI of 3.07, is acidic, suggesting that TAP might form a dipole.

The α -helical region contains three domains with strong coiled-coil-forming potential (I, amino acids 650–709; III, amino acids 783–828; IV, amino acids 843–930; shaded letters) as defined by the COILS2 program (18). Another region ($\approx 80\%$ probability of forming a coiled-coil structure) was detected between amino acids 728 and 766 (region II). Helix-breaking prolines and glycines are found flanking each of the putative coiled-coil regions. Electron microscopic analysis (data not shown) indicates that TAP is superficially similar to myosin in that two TAP polypeptides, each composed of a globular (≈ 9 nm) head and an elongated (≈ 45 nm) tail, form a homodimer by parallel association of the tails. The coiled-coil region of TAP shows $\approx 40\%$ similarity and $\approx 20\%$ identity with other coiled-coil-containing proteins such as myosins, tropomyosins, kinesins, CLIP-170, golgin p160, giantin, and tpr gene product (19–21).

Using the BLAST program we found the strongest homology between TAP and a human cDNA encoding a 99-amino acid ORF with a sequence 92% identical to the TAP sequence (amino acids 605-703 of TAP). A weaker homology was found to Uso1p, a yeast protein involved in ER to Golgi traffic. Uso1p is a hydrophilic protein of 1790 amino acids with a 1010-amino acid coiled-coil region (11). Using BESTFIT (of Genetics Computer Group package) and BLAST programs, we found 42% identity (65% similarity) between amino acids 21-85 of TAP and 18-84 of Uso1p (Fig. 5B, boxed residues) and 33% identity (53% similarity) between amino acids 98-259 of TAP and 118-277 of Uso1p. Within this last region there is a stretch of high homology (61% identity, 77%) similarity) between amino acids 201-252 of TAP and 220-271 of Uso1p (Fig. 5B, boxed amino acids). The C-terminal homology includes the acidic region, which has 85% similarity between amino acids 932-954 of TAP and 1767-1788 of Uso1p (Fig. 5C, acidic amino acids in boldface) and the region that immediately precedes it with 43% identity over 37 amino acids (amino acids 898-934) (Fig. 5C, boxed amino acids). Interestingly, the sequence immediately preceding the start of the acidic domain, SKLKDLG, is 100% identical in both proteins (double underlined amino acids). The overall homology (28.5% identity between amino acids 1–644 of TAP and 1–760 of Uso1p and 28.65% identity between the coiled-coil domains) of both proteins is relatively low.

Α	aaaa	gcto	rtcgt	gttt	tctt								ctto			-60 -1
1	atg Met	aat Asn	ttc Phe	ctc Leu	cgc Arg	ggg Gly	gtg Val	atg Met	ggg Gly	ggt Gly	cag Gln	agt Ser	gcc Ala	gga Gly	ccc Pro	45
16	cag Gln	cac His	aca Thr	gaa Glu	gct Ala	gag Glu	acg Thr	att Ile	cag Gln	aag Lys	ctc Leu	tgt Cys	gac Asp	cgc Arg	gta Val	90
31	gct Ala	tca Ser	tct Ser	act Thr	tta Leu	ctg Leu	gac Asp	gac Asp	cga Arg	aga Arg	aat Asn	gct Ala	gtc Val	cgt Arg	gcc Ala	135
46	ctt Leu	aag Lys	tca Ser	ctg Leu	tct Ser	aag Lys	aaa Lys	tac Tyr	cgc Arg	ttg Leu	gaa Glu	gtg Val	gga Gly	atc Ile	caa Gln	180
61	gcc Ala	atg Met	gag Glu	cat His	ctt Leu	atc Ile	cac His	gtc Val	tta Leu	cag Gln	aca Thr	gat Asp	cgt Arg	tcg Ser	gat Asp	225
76	tct Ser	gaa Glu	ata Ile	ata Ile	gcg Ala	tat Tyr	gct Ala	ttg Leu	gac Asp	aca Thr	ctc Leu	tat Tyr	aat Asn	ata Ile	ata Ile	270
91	tcc Ser	aat Asn	gat Asp	gaa Glu	gag Glu	gaa Glu	gaa Glu	gta Val	gaa Glu	gaa Glu	aat Asn	tct Ser	aca Thr	aga Arg	cag Gln	315
106	agt	gag	gat	ttg	gga	agc	caa	ttt	aca	gaa	att	ttc	att Ile	aag	cag	360
121	cca Pro	gaa Glu	aat Asn	gtc Val	act Thr	ctc Leu	ctg Leu	tta Leu	tct Ser	ttg Leu	ttg Leu	gag Glu	gag Glu	ttt Phe	gat Asp	405
136	ttc Phe	cac His	gtc Val	cga Arg	tgg Trp	cct Pro	ggt Gly	gtg Val	aga Arg	ctt Leu	ctg Leu	act Thr	tct Ser	ctt Leu	tta Leu	450
151	aaa Lys	cag Gln	cta Leu	ggg Gly	cct Pro	cca Pro	gtg Val	cag Gln	caa Gln	atc Ile	atc Ile	cta Leu	gtc Val	agc Ser	ccc Pro	495
166	atg Met	ggt Gly	gtt Val	tca Ser	aaa Lys	tta Leu	atg Met	gac Asp	ttg Leu	ctg Leu	gca Ala	gat Asp	tcc Ser	agg Arg	gaa Glu	540
181	att Ile	ata Ile	cgt Arg	aat Asn	gat Asp	ggc Gly	gtc Val	tta Leu	cta Leu	ctg Leu	cag Gln	gca Ala	tta Leu	aca Thr	agg Arg	585
196	agc Ser	aac Asn	gga Gly	gcg Ala	atc Ile	cag Gln	aaa Lys	att Ile	gtt Val	gct Ala	ttt Phe	gaa Glu	aat Asn	gct Ala	ttc Phe	630
211	gag Glu	aga Arg	cta Leu	ctg Leu	gac Asp	att Ile	att Ile	aca Thr	gag Glu	gag Glu	ggg Gly	aac Asn	agc Ser	gat Asp	gga Gly	675
226	ggt Gly	ata Ile	gta Val	gtg Val	gaa Glu	gat Asp	tgt Cys	ttg Leu	att Ile	ttg Leu	ctc Leu	caa Gln	aat Asn	ttg Leu	tta Leu	720
241	aaa Lys	aac Asn	aac Asn	aat Asn	tcc Ser	aat Asn	caa Gln	aat Asn	ttt Phe	ttt Phe	aaa Lys	gaa Glu	ggc Gly	tct Ser	tat Tyr	765
256	att Ile	caa Gln	cgt Arg	atg Met	aaa Lys	gct Ala	tgg Trp	ttt Phe	gaa Glu	gtt Val	gga Gly	gat Asp	gaa Glu	aat Asn	cct Pro	810
271	ggt Gly	tgg Trp	tca Ser	gca Ala	cag Gln	aaa Lys	gtg Val	acc Thr	aat Asn	ctt Leu	cat His	tta Leu	atg Met	ctc Leu	cag Gln	855
286	ctt Leu	gtg Val	cgg Arg	gta Val	ctg Leu	gtg Val	tct Ser	ccc Pro	acc Thr	aac Asn	cct Pro	ccc Pro	ggt Gly	gct Ala	acc Thr	900
301	agt Ser	agc Ser	tgc Cys	cag Gln	aag Lys	gcc Ala	atg Met	ttc Phe	cag Gln	tgt Cys	ggg Gly	tta Leu	cta Leu	caa Gln	cag Gln	945
316	ctt Leu	tgc Cys	act Thr	atc Ile	ctg Leu	atg Met	gct Ala	acc Thr	gga Gly	att Ile	cct Pro	gct Ala	gat Asp	atc Ile	ctg Leu	990
331	åct Thr	gag Glu	acc Thr	ata Ile	aat Asn	act Thr	gta Val	tca Ser	gaa Glu	gtt Val	att Ile	cga Arg	ggc Gly	tgc Cys	caa Gln	1035
346	gta Val	aat Asn	caa Gln	gac Asp	tac Tyr	ttt Phe	gct Ala	tct Ser	gtg Val	aat Asn	gcg Ala	cct Pro	tca Ser	aat Asn	ccc Pro	1080
361	cca Pro **	cga Arg	ccg Pro **	gca Ala	atc Ile	gtt Val	gtg Val	ctg Leu	ctc Leu	atg Met	tcc Ser	atg Met	gtc Val	aac Asn	gag Glu	1125
376	agg Arg	cag Gln	cca Pro	ttt Phe	gta Val	ctg Leu	cgc Arg	tgc Cys	gcc Ala	gtg Val	ctc Leu	tac Tyr	tgt Cys	ttc Phe	cag Gln	1170
391	tgc Cys	ttc Phe	ctc Leu	tat Tyr	aaa Lys	aac Asn	gag Glu	aaa Lys	gga Gly	caa Gln	gga Gly	gag Glu	att Ile	gtg Val	gcc Ala	1215
406	act Thr	ctg Leu	cta Leu	cct Pro	tcc Ser	act Thr	att Ile	gac Asp	gca Ala	aca Thr	ggt Gly	aac Asn	tcg Ser	gtc Val	tca Ser	1260
42 1	gct Ala	ggc Gly	cag Gln	ctg Leu	ctc Leu	tgc Cys	gga Gly	ggt Gly	ttg Leu	ttt Phe	tcc Ser	aca Thr	gac Asp	tcc Ser	ctc Leu	1305
436	tca Ser	aac Asn	tgg Trp	tgt Cys	gct Ala	gcc Ala	gtg Val	gcc Ala	ctc Leu	gcc Ala	cat His	gca Ala	ctg Leu	caa Gln	ggg Gly	1350
451	aac Asn	gcc Ala	acc Thr	cag Gln	aag Lys	gag Glu	cag Gln	ctg Leu	ctc Leu	agg Arg	gtt Val	cag Gln	ctg Leu	gcc Ala	acc Thr	1395
466	agc Ser	atc Ile	ggc Gly	aac Asn	ccc Pro	cca Pro	gtg Val	tcg Ser	ctc Leu	ttg Leu	cag Gln	cag Gln	tgc Cys	acc Thr	aac Asn	1440
481	att Ile	ctg Leu	tcc Ser	cag Gln	gga Gly	agc Ser	aag Lys	ata Ile	cag Gln	aca Thr	aga Arg	gtc Val	GJÀ aàa	ttg Leu	ttg Leu	1485
496	atg Met	ttg Leu	ctt Leu	tgt Cys	acc Thr	tgg Trp	tta Leu	agc Ser	aac Asn	tgt Cys	ccg Pro	att Ile	gct Ala	gta Val	aca Thr	1530
511	cac His	ttt Phe	ctt Leu	cac His	aac Asn	tca Ser	gcc Ala	aat Asn	gtt Val	cca Pro	ttt Phe	ctt Leu	aca Thr	gga Gly	cag Gln	1575
526	att Ile	gca Ala	gaa Glu	aat Asn	ctc Leu	gga Gly	gaa Glu	gaa Glu	gag Glu	cag Gln	ttg Leu	gtc Val	caa Gln	ggc Gly	tta Leu	1620
541	tgt Cys	gcc Ala	ctt Leu	ctt Leu	ttg Leu	ggc Gly	att Ile	tca Ser	att Ile	tat Tyr	ttc Phe	aac Asn	gac Asp	aac Asn	tca Ser	1665
556	cta Leu	gaa Glu	aac Asn	tac Tyr	acg Thr	aaa Lys	gag Glu	aaa Lys	cta Leu	aag Lys	caa Gln	cta Leu	ata Ile	gag Glu	aag Lys	1710
571	agg Arg	att Ile	ggc Gly	aaa Lys	gag Glu	aat Asn	tac Tyr	ata Ile	gag Glu	aaa Lys	ctt Leu	gga Gly	ttt Phe	att Ile	agc Ser	1755
586	Lys	His	Glu	Leu	Tyr	Pro	Arg	Ala	Ser	Gln	Lys	Pro **	cag Gln	Pro **	Asn	1800
601	ttt Phe	ccg Pro **	agt Ser	cca Pro **	gaa Glu	tac Tyr	atg Met	ata Ile	ttt Phe	gat Asp	cat His	gag Glu	ttt Phe	aca Thr	aaa Lys	1845

616															aag Lys	1890
631	tcc Ser	ago Ser	gaa Glu	gaa Glu	gat Asp	aag Lys	aaa Lys	gag Glu	gaa Glu	gag Glu	gta Val	aag Lys	aaa Lys	acc Thr	Leu	1935
646	gaa Glu	Gln	cat His	gat Asp	aat Asi	att 11e	gtg Val	act Thr	cac H1s	tat Tyr	aag Lys	aat Asn	gtg Val	atc Ile	cgt Arg	1980
661	gag Glu	caa Gln	gac Amp	cta Leu	caa Gln	ctg Leu	gag Glu	gaa Glu	cta Leu	aaa Lys	cag Gln	caa Gin	gtg Val	tcc Ser	aca Thr	2025
676	ctg Leu	r aaa Lys	tgt Cys	cag Gln	aac Asn	gag Glu	cag Gln	ctg Leu	caa Gln	aca Thr	gca Ala	gtc Val	acc Thr	cag Gln	cag Gln	2070
691	gct Ala	tct Ser	cag Gln	att Ile	cag Gln	cag Gln	cac His	aag Lys	gat Asp	cag Gln	tac Tyr	aac Asn	ctc Leu	cto Leu	aaa Lys	2115
706	gtt Val	cag Gln	ctc Leu	ggg Gly	aaa Lys	gac Asp	aat Asn	cac His	cat His	caa Gln	ggt Gly	tct Ser	cac His	agt Ser	gac Asp	2160
721	ggg Gly	gct Ala	cag Gln	gta Val	aat Asn	ggc Gly	att Ile (cag Cln	cca Pro	gag Glu	gaa Glu	atc Ile	agt Ser	cgg Arg	ctg Leu	2205
736	agg Arg	gaa Glu	gag Glu	ata Ile	gaa Glu	gaa Glu	tta Leu	aga Arg	agt Ser	cat His	cag Gln	gtg Val	ctc Leu	tta Leu	cag Gln	2250
751	agc Ser	cag Glm	cta Leu	gct Ala	gaa Glu	aag Lys	gac Map	act Thr	gtg Val	att Ile	gaa Glu	aat Ann	ttg Leu	aga Arg	tct Ser	2295
766	tca Ser	caa Gln	gtg Val	tct Ser	ggc Gly	atg Met	agt Ser	gaa Glu	cag Gln	gct Ala	tta Leu	gca Ala	acg Thr	tgt Cys	tca Ser	2340
781	ccc Pro	aga Arg	gat Mip	gcc Ala	gag Glu	caa Gin	gtt Val	gca Ala	gag Glu	tta Leu	aaa Lys	cag Gln	gaa Glu	ttg Leu	tca Ser	2385
795	gca Ale	tta Leu	aag Lys	tcc Ser	cag Gln	tta Leu	tgt Cys	tca Ser	cag Cln	tct Ser	ctg Leu	gag Glu	atc Ile	act Thr	aga Arg	2430
811	ctc Leu	cag Gln	aca Thr	gag Glu	aac Asn	cgt Arg	gag Glu	ctg Leu	caa Gln	cag Gln	aga Arg	gca Ala	gag Glu	acc Thr	ttg Leu	2475
826	gca Ala	aag Lys	tca Ser	gtt Val	cct Pro	gta Val	gaa Glu	gga Gly	gag Glu	agt Ser	gag Glu	ctt Leu	gtg Val	aca Thr	gcc Ala	2520
841	gca Ala	aaa Lys	act Thr	acg Thr	gat MP	gta Val	gag Glu	gga Gly	agg Arg	ctg Leu	tct Ser	gcc Ala	ctg Leu	ctg Leu	cag Gin	2565
855	gag Glu	acc Thr	aaa Lys	gag Glu	tta Leu	aag Lys	aat Asn	gag Glu	att Ile	aaa Lys	gca Ala	ttg Leu	tct Ser	gag Glu	gag Glu	2610
871	agg Arg	acc Thr	tcc Ser	att Ile	caa Gln	aag Lys	cag Gln	ctg Leu	gac Marp	tcc Ser	tcc Ser	aac Asn	agc Ser	acc Thr	atc Ile	2655
886	gcc Ala	atc Ile	cta Leu	caa Gin	acg Thr	gag Glu	aaa Lys	gac Asp	aag Lys	cta Leu	tac Tyr	ttg Leu	gaa Glu	gtt Val	acg Thr	2700
901	gat Mip	tct Ser	aag Lys	aaa L y s	gaa Glu	caa Gin	gat Asp	gat Mip	ctt Leu	ttg Leu	gtg Val	ctg Leu	ttg Leu	gca Ale	gat Asp	2745
016	caa Gin	gat Asp	cag Gin	aaa Lys	atc Ile	ctg Leu	tca Ser	ctg Leu	aag Lva	agt	aaa LNS	ctc	aag L ve	gat Asp	ctt Leu	2790
916		-														
931	ggt	cat	cca Pro	gtt	gaa	gaa	gaa	gat	gaa	tct	gga	gac	caa	gaa Glu	gat	2835
	ggt Gly gat <u>Asp</u>	cat His gat Asp	cca Pro gat Asp	gtt Val gaa Glu	gaa <u>Glu</u> ctt Leu	gaa Glu gac Asp	gaa Glu gat Asp	gat Asp ggt Glv	gaa Glu gac Asp	tct Ser agg Arg	gga Glv gac Asp	gac Asp cag Gln	caa Gln gat Asp	Glu atc Ile	gat <u>Asp</u> tag	2880
931	ggt Gly gat ctt: gct gaa agaa agaa agaa agaa agaa agaa a	cat His gat Asp ttat gttt acacc gcggg aagg cttc gcagg gagc cttc ttct	cca Pro gat	gtt Val gaa <u>Glu</u> aggaa actac ccaga actaf	gaa <u>Glu</u> ctt <u>Leu</u> acaat cacaat dytttt ttctt aaggo cacaga tttct gcctc cgcct cgcct cagtt caata cagat	gaa <u>Glu</u> gac Asp agag acaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaa caaaaaa	gaa Glu gatt Asp gattaa aatto gggg cagg gggg cagg gggg cagg gggg cagg gggg cag ca	gat Asp ggt Glv atgtc agcca cgtc agcca cgtt agcca gggaa gcgg ggaa gctgg ctgg	gaa Glu gac Asp catto agaad cagaad cagaad cagaad cagaad cagaad ytttt ytttt gatta aaaaa	tct <u>Ser</u> agg Arg tgtgti cgtgg cgtgg cgtgg cgtgg cgtgg cgtgg ccccet accac cccct ccct cccct ccc ccc ccct ccctt ccct cccct ccc ccct ccct ccct ccct ccc ccct ccct ccct ccct ccc ccct ccct ccct ccc ccc ccct ccct ccc ccc ccct ccct ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc ccc cccc	gga <u>Glv</u> gac Asp cgaag gggaa gggaa gcaga cctta acatt cggct gggt ccgga cctga ccgga cctga ccgga cctga ccgga cctga gggt ccgga ccga gga ccta acatt ccgga ga ccta acatt ccgga ga ccta acatt ccgga ga ccta acatt ccgga ccga cc	gac Asp cag <u>Gln</u> jaagt aagto ctgg ctgat cctgg ccgg dtgat tcaa ttgat ttaa ttaa ttatt	caa <u>Gln</u> gat Asp cetta cate cate cate cate cate cate cat	Glu atc Ile aagaa cact. ttca aggt cgtg ccttg agcaa aaagg cacaa aaaga cacaa attt cgtc	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	
931 946	ggt Gly gat ctt: gct gaa agaa agaa agaa agaa agaa agaa a	cat His gat Asp ttgg tttat acacc gctt gcg gg acttc gcag gagc cagt ttct atag gagc	ccca Pro gat Asp cccta ttgatc tatta actg gccta ttgatc tagtat cccaa actg tcgata tagtt ttgg ttgg	gtt Val ggaa Glu ggaa cctco gtatg ccaga cc	gaa <u>Glu</u> ctt Leu acaat cacaa gtttt cacaa ggctc cacaa ggctc cacaa cagat acagat acagat	gaa Glu gac Asp aaaaa ctaat cctg acctg cacgt agtct caata cacct caata aacco caata aacco	gaa Glu gatta Asp gatta catto cggto cggo cago cago cago cago cago cago cag	gat Asp ggt Glv ggt ggcad gccad gccad gggad gggg	gaa Glu gac Asp catto agaad caaaa cagta cagta cagta cagta cagta tgat gatt gat	tett ser agg stgtgt caggg cgtgg cgtgg cgtcg caetaa caetaa caetaa caetaa caetaa caetaa caetaa caetaa caetaa caetaa caetaa caetaa	gga Glv gac Asp gaag gaag ggaag ggaag ccts acatt cggct ccggg ccgg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccgg ccgg ccgg ccgg ccgg ccgg ccg ccggg cc cc	gac Asp cag Gln aago aagt cccc gggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg ccggg cccc cc	caa <u>Gln</u> gat <u>Asp</u> cctta atto gacta ctaato ctaato ggggaa aggga agggaa aggaa agggaa agggaa agggaa agggaa agggaa agggaa agggaa agggaa agggaa aggaa aggaa aggaa aggga aggaa agggaa agggaa agggaa agggaa a agggaa aggga	Glu atc Ile aagaa cact. ttca aggt cgtg ccttg agcaa aaagg cacaa aaaga cacaa attt cgtc	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931	ggt Gly gat ctt: gct gaa agaa agaa agaa agaa agaa agaa a	cat His gat Asp ttat gttti gttti ggtti gggg cagc gcagg gagc cagt. ttat gat gagg gagc 1	cca Pro gat Asp cccta tgat tatt agcat gat actgf ccca actgf ccca actgf ccca tggaa tagtt tatt t	gtt Val gaa Glu aggaa cctco gtatg cctco gtatg cctco gtatg cctco cctco gtatg cctco cctco gtatg cctco cctco gtatg cctco cctco cctco gtatg cctco cctco gtatg cctco cctco gtatg cctco cctco gtatg cctco cctco cctco gtatg cctco cc	gaa <u>Glu</u> ctt <u>Leu</u> acaat cacaa gyttt ttttt aaggy acaga gyttt cacaa gyttt cacaa gyttt cacaa aggy acaga gyttt cacaa aggy acaga gyttt cacaa aggy acaga gyttt cacaa aggy acaga gyttt cacaa aggy acaga gyttt cacaa aggy acaga gyttt cacaa aggy acaga gyttt cacaa aggy cacaa a aggy cacaa aggy cacaa aggy cacaa a a aggy cacaa aggy cacaa aggy cacaa aggy cacaa aggy cacaa aggy cacaa a c a a a a a a a a a a a a a c a a a a a a a a a a a a a a	gaa Glu gac Aspo aaaaa ctga cc	gaa <u>Glu</u> gatt Asp gattaa aatto gattaa aatto gggg gggg ggg	gat Asp ggt Glv agaca ggca ggca ggca ggca ggtt agcct agcct agca ggaa gga	gaa Glu gac Asp catto cagaa cccat cagaa c cagaa cccat cagaa c cagaa c cagaa c cagaa c cagaa c c cagaa c cagaa c cagaa c cagaa c cagaa c cagaa c cagaa c c cagaa c c c c	tet ser agg fgtft cgtgg cgtgg agtt actaa gggad cgtct cctet catcaa gggad cact	gga Glv gac Asp ggaa ggaa ggaa ggaa ggaa ggaa ggaa gg	gac Asp cag Gln gaagt ctg ctg ctg ctg ctg ctg ctg ctg ctg c	caa Gln gat Asp cctta catto cacca ctato cacca ctato cacca gat catto cacca catto cato	Glu atc Ile aagaa cact. ttca aggt cgtg ccttg agcaa aaagg cacaa aaaga cacaa attt cgtc	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat Aso ctt: aaaa gaaa agac ccca gag aaaa atg tgaa tat	cat His gat Asp ttat gctt: acacc gctg cagg: cagt: ttct ttct 1 1 51	CCA Pro gat Asp cccta cctgac tgac gcta accg cca accg cca accg cca accg cca accg tattt tattt holli C KKYRI	gtt Val gaa Glu aggaa cctcc gtatg ccage ccag cca	gaa <u>Glu</u> ctt Leu acaat cacaat	gaa Glu gac Asp aaaaa aaco aaaaa aaco cata aaco aaaa aaco aaaa aaco aaaa aaco aaaa aaco aaaa aaco aaaa aaco aaaaaaaa	gaa Glu gat Asp gatta ascc ggtg gggg gggg gggg gggg gggg gggg g	gat Asp Glv atgtc agaca gccac ggcac ggcac ggcac ggcac ggcac ggcac gctg gcac gctg gcac gcac	gaa Glu gac Asp catto ca	tct Ser agg Arg tgtft ccagg cdt9g agttc actaa tcctct actaa cact TLLD TTLIS DT.I	gga Glv gac Asp ggac ggac ggac ggac ggac ggac coggac ggac	gac Asp cag GIn taago taago totot cro	Caa Gln gat Asp cctta catco gact catco gacta ctaat cggact catco ggaca sgtgc agtgc agtgc agtgc sgtgc sgts catco ggaca .: Asp 	Glu atc Ile agga act; tta aggt act; gtg gtg gtg acttg acttg acttg acttg acttg acttg acttg acttg acttg acttg acttg acttg acttg acttg actt actt	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat agaa agaa agaa agaa agaa agaa agaa	cat His gat Asp ttaaa ttggttt gtttt ggttt cag cttc cgggg aagg cttc cgggg aagg cag ttaaa gag cttc ttat ttat	cca Pro gat Asp cctsg tgatg tatt aatag tggatg tagtt tggatgga	gtt Val gaa Glu aggaa cctco catag ccag ccagg ccag cc	gaa <u>Glu</u> ett Leu acaat zacaat zacaat zacaga tttet caaac geete zagat z zagat z zagat z zagat z zagat z z z z z z z z z z z z z z z z z z z	gaa Glu gac Asp cagac cactor c	gaa Glu gatt Asp gattaas catco ggtg gggg caago caago caago caago caagatt agatt 	gat Asp ggt GlV agaca gccac gc	gaa Glu gac Asp catto agaac cccat cagaa grigcc agact gatt gatt gatt gatt gatt gatt	tct Ser agg gtgtl ccagg gtgtg atttc agtat cctct agaat agcaa cctct TLLN TLLS DT.I.::I. ETIL	gga Glv. gac Gly. ggac ggaa ggaa ggaa ggaa ggaa ggaa gga	gac Asp cag gln taagg cag cag cag cag cag cag cag cag cag	Caa Gln gat Asp cotta aato gact gacta aato ggggg gggg aa saggcta taat sIS SIS SIS SIS E : DLT	Glu atc Ile agat cattica aggt catti gtt gtt gtt gtt gtt agaa attt agaa agaa attt gtt agaa a a a a a a a a a a a a a a a a a a a	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat gctt gctt gaa agaa agaa agaa agaa a	cat His gat Asp ttagg ttagg ttg ttg ttg ttg ttg ttg ttg	CCA Pro gat Asp cccts tattt agcat ccca gccca tagtt tattt agcat tagtt tag	gtt Val gaa Glu aggaa cctco catg ccaga cco	gaa <u>Glu</u> ctt Leu acaat acaaa gcttct caaaa gcttct caaaa gcctc caaaa acaagat ac	gaa Glu gac Asp aaaa cctga acctga cc	gaa Glu gat Asp gatta aacto gggg ccago ggggc cago gggc cago gggc cago gggc cago gggc cago gggc cago gggc cago gggc cago gggc cago ggg cago ggg cago ggg cago cago	gat Asp ggt atgtc agaca cette agcca cette agcca ggcaa gcca ggcaa gcca gcca ggcaa gcca gcca ggcaa gcca ggcaa gcca ggcaa gcca agcca agcca cette ggcaa gcca agcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca agcca cette ggcaa gcca ggcaa gcca ggcaa gcca ggcaa gcca gcca gcca gcca gcca gcca gcca gcca gcca gcca gcca gcca gcca ggcaa gcca ggcaa gcca ggcaa gcca ggcaa gcca ggcaa gcca ggcaa gcca ggcaa gcca gca g	gaa Glu gac Aspo catto agaac cocat cagta cagta cagta cagta cagta cagta tgttt aaaaa tgttt RVASS II] IAYAL IYAAIL IYAAIL EI	tct Ser agg gtgtt ccagg gtgt attcc acta actaa gggaa cacta tctcc agaat cacta tctcc agaat cacta tctcc agaat cacta tctc agaat cacta tctc agaat cacta tctc agaat cacta tctc agaat cacta tctc agaat cacta tctc agaat cacta tctc agaat cacta tctc tctc	gga Glv gac Gaag ggac ggac ggac ggac ggac cast coggac cogg	gac Asp cag gin taago tgat tcat tcat tcat tcat tcat tcat tcat	Caa Gln gat Asp cotta gacta gacta trato gggaa gggaa gggaa gggaa gggaa sub gggaa sub gggaa sub gggaa sub gggaa trato ggaa trato trato ggaa trato ggaa trato	Glu atc Ile agaa ttcaatgt actgg cttg actgg ccttg acaca aaggc cacaa attti gtca acaca aaga ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acaca ccttg gtca acacaca ccttg gtca acacacaca ccttg gtca acacacacaca ccttg gtca acacacacacacacaca gtc ccttg gtca acacacacacacacacacacacacacacacacacac	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat <u>Asp</u> ctttaaaa gaaa agaa agaa agaa agaa atga ttat	cat His gat Asaa ttgg ttat gat ccg gag ct ccg gag cag gag ct cag ttat tat gat tat tag ttat tag gat ttat tag gat taa a ttgg gat taa a t gat taa a t gat taa a t gat taa a t gat taa a t gat taa a tag ttat c gag ttat tag gat taa a a gat taa a a gat taa a a gat taa a a gat taa a gat taa a gat taa a gat taa a gat taa a gat ta a gat ta a gat ta a gag gag	CCCA Pro gat Asp ccctfg tgatct agcat gctat tagtt tggt tagtt tggt tagtt satcg actag gcca tagtt satcg tagtt satcg sa	gtt Val ggaa Glu sctco sctor s	gaa Glu ctt Leu acaat gcctc gcctc gcctc gcctc gcctc gcctc aaag ataa ataa	gaa Glu gac Asp Asp caaaa cctga cctga cctga cctga cctga cata aacco aacco aaco aacco a	gaa Glu gat Aso gattaaa catco cago gggc cago cago gggc cago cago ca	gat Asp ggt Glv ggt ggaa ggaa ggaa ggaa ggaa ggaa ggaa	gaa Glu gac Asp cattg agaaa ccattg agaac ccattg gatt gatt	tct Ser Arg gtgtt ccagg cgtgg ggga ttctc actaa cacta cacta ggga cctct agaata cacta lii: FILDD II: FILD FI ii WIADP	gga Glv gac Asp gaag gaag gaag gaag gaag gaag gaag caga catt aatt a	gac Asp cag gad dig dig dig dig dig dig dig dig dig di	Caa Gln Asp cctta	Glu atc Ile aga catta aggt catta ggg catta aggt gc cata agg catta aggt gc cata agg catta aggt gc cata aggt gc catta aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggt gc cata aggc gc cata agc agc agc agc agc agc agc agc agc ag	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat <u>Asp</u> ctt: aaa gaga aaa gccc gag gaga aaa tga tta t	cat His gat taaaat ttggtt cggggg aagg cagt ttct ttc	CCCa Pro gat ASD CCCCta Ctggacg tagtot tagto	gtt Val gaa Glu aggaa actai acti actai act	gaa Glu ctt Leu acaat cacat cacaat cacaat cacaat cacaat cacaat cacaat cacaat cacat	gaa Glu gac Asp ctga ctga ctga cctga cata acctg acctg ac	gaa Glu gat Asp gat acto ggg ggg gggg caag ggg caag ggg caag ggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag gggg caag ggg caag g caag g caag g caag g caag g caag g caag g caag caa s c s c	gat Asp GIV GIV GIV GIV Asp Control GIV Asp Control GIV Asp Control GIV Asp Control GIV Asp Control	gaa Glu gac Asp catto agaad ccata agaad ccata agaad ggac gatt gatt gatt gatt aaaaa tgtt RVASS II RVASS II EI SVAIL	tct Sar agg ftgtt ccagg gtgt agtt actac actac agaat cact TLLD TLLS TLLS TLLS TLLS TLLS TLLS TLLS	gga Gly gac gac gaag gaag gaag gaag gaag gaag	gac Asp Cag Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln	Caa Gin Asp active catto	Glu atc Ile aggatizatti string	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat Asp cgt: aaa agaa agaa ccca gaga aaa agtc tgaa tgaa	cat His gat his gat ttgtttat gtttat gttat ggag cagg ggag cagg ggag cagg ggag cagg gag cagg gag cagg gag cagg gag cagg a a tt g gag cagg gag cagg gag cagg gag cagg gag cagg a a a a	CCCA Pro gat ASD CCCCL2 CCCC2 GCCA GCCA GCCA GCCA GCCA GCCA GC	gtt Val ggaa Glu aggaa ccag ccag ccag ccag ccag ccag cc	gaa Glu ctt Leu acaat cacaa gtttt cacaa gcttc gcctc acaga gcct acaga gcct acagcc gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga gcct acaga acaga acaga acaga gcct acaga gcct acaga ac ac ac	gaa Glu gac Asp aaaa taat cctg aact aact	gaa Glu gat asp gat asp gat cag ggt ggt ggt ggt ggt ggt ggt ggt ggt g	gat Asp ggt aggc gga gga gga gga gga gga gga gga gg	gaa Glu gac Asp cattg agac cattg agac cattg agac cattg agac cattg agac cattg agac cattg agac cattg agac cattg agac agac	tet Ser agg Arg ftgtt cagg cgtg agttc agtact actact agg agg agg agg agg agg agg agg agg ag	gga gly gac Asp gac ygac ygac cast satt action ggat cogga co	gac Asp cag gac dag gac dag dag dag dag dag dag dag dag dag dag	Caa Gln gat Asp ctta catto gact gact catoo gaca gact catoo gaca gaca catoo gaca satoo gaca gaca gaca gaca satoo gaca gaca gaca gaca gaca gaca gac gaca gaca gaca gac gac	Glu atc Ile agaa actggt gtgg gtgg gtgg gtgg gtgg gt	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat Asp ctt: gct: gca aaaa agaa agaa agaa agaa agaa agaa	cat His gat his gat gtttat gtttat gctgg gccgg gagct ttcg gagc ttcc gagc gttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagc ttcg gagct ttcg gagc g ttcg gagc g ttcg gagc gagc	CCCA Pro gat ASD CCCCCA gatct gatct gatct gatct gatca gata gata	gtt Val gaa Glu aggaa cctac ccag ccag ccag ccag ccag cc	gaa Glu ctt Leu acaat acaat acaa acaat acaa gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gcc gcc gc	gaa Glu gac Asp aaat aag acg ctg aacg ctaa aacg cacg cacg cac	gaa Glu gat ASD gatta ASD gatta acto gggg cago gggg cago gggg cago gggg cago cago	gat Asp ggt Glv aggcc ggct ggcca ggg	Gaa Glu gac Asp catto agaac catto catto agaac catto gatt gatt aagaac gatc secat secat secat catto agaac secat seco	tct Ser agg dtgtt ccagg ggtt acta ggtt acta ggtt acta ggg ggc cct acta acta	gga gly gac Asp gac gaag gaag gaag gaag gaag gaag gaag cast satt atgaa satt atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta	gac Asp cag Gln aagt iaagt ictot ictot ictot ictot gggg ictog ictog ictog ictog ictot gggg ictot i i i i i i i i i i i i i i i i i i	Caa Gin gat Asp cotta catto ca	Glu atc atcatta aga actatta agg actatta	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat Asp ctt: gct: gca aaaa agaa agaa agaa agaa agaa agaa	cat His gat his gat gtttat gtttat gctgg gccgg gagct ttcg gagc ttcc gagc gttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagct ttcg gagc ttcg gagct ttcg gagc g ttcg gagc g ttcg gagc gagc	CCA Pro gat Asp ccctg ccttg act gat act gat act gat act gat act sat ccca act gat sat ccca act gat sat ccca ccca ccca ccca ccca ccca ccc	gtt Val gaa Glu aggaa cctac ccag ccag ccag ccag ccag cc	gaa Glu ctt Leu acaat acaat acaa acaat acaa gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gct ccc gcc gcc gc	gaa Glu gac Asp aaat aag acg ctg aacg ctaa aacg cacg cacg cac	gaa Glu gat ASD gatta ASD gatta acto gggg cago gggg cago gggg cago gggg cago cago	gat Asp ggt Glv aggcc ggct ggcca ggg	Gaa Glu gac Asp catto agaac catto catto agaac catto gatt gatt aagaac gatc secat secat secat catto agaac secat seco	tct Ser agg dtgtt ccagg ggtt acta ggtt acta ggtt acta ggg ggc cct acta acta	gga gly gac Asp gac gaag gaag gaag gaag gaag gaag gaag cast satt atgaa satt atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta atgaa satta	gac Asp cag Gln aagt iaagt ictot ictot ictot ictot gggg ictog ictog ictog ictog ictot gggg ictot i i i i i i i i i i i i i i i i i i	Caa Gin gat Asp cotta catto ca	Glu atc Ile agga cacta aggt cacta aggt cacta aggt cacta aga cacta aga cacta aga cacta aga cacta aga cacta aga cacta aga cacta aggt cacta aga cacta aggt cacta aga cacta cacta cacta aga cacta	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B	ggt Gly gat Asp gca gaaa aggac agag aaa aggac cca gaga aaa tat 1 1 1 1 1 1 2 2	cat s His gat Ltaggittic cactor gcaggic gcaggic cactor gcaggic cactor gcaggic cactor gcaggic cactor gcaggic cactor gcaggic atata a 1 1 51 48 98 98 29 48 77 96 27 46	CCCa Pro ggat Asp ccctz gccttg gcctt actor gcctt actor gcctt actor gcctt actor gcctt actor actor gcctt gcctt actor gcctt gcctt actor gcctt gcctt actor gcctt gcctt actor gcctt gcctt gcctt actor gcctt	gtt Val gaa Glu aggaa Colu sctco sctco sctco sctco sctco sctage sctata s	gaa Glu ctt Leu acaat acaat acaat caage caage tte cte caage	gaa Glu gac Asp casa casa gaca casa casa casa casa casa	gaa Glu gatt Aso jatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor gatta actor ac	gat aspo ggt ggt ggc ggc ggc ggc ggc ggc	gaa Glu gac Asp catto ca	tct Ser agg Cagge Cagge Cagge Cagge Cacta	gga Gly gac Asp gac gac gac gac ggac ggac ggac ggac gg	gac Asp cag Gin aago ictgat ictga i ictga i ictga i ictga i i i i i i i i i i i i i i i i i i i	Caa Gln gat Asp acta gact gact gact gact gact gact gact	Glu atc IIa again actg gcd actg actg gcd actg gcd actg gcd actg gcd actg gcd actg gcd actg actg gcd actg	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706
931 946 B TAP USO1	ggt Gly gat Asp gca gaaa aggac agag aaa aggac cca gaga aaa tat 1 1 1 1 1 1 2 2	cat his cat hi	CCCA Pro gat ASD CCCCCA gatct gatct gatct gatct gatca gata gata	gtt Val gaa Glu aggaa cag ccago ccag	gaa Glu ctt Leu acaat acaag gcatc acaag ac	gaa Glu gac Asp aaaa ctaat cctg2 acct	gaa Glu gattaa acto agat acto acto agat acto acto agat acto acto acto acto acto acto acto ac	gat Asp ggt Glv adgdc ggda ggda ggda ggda ggda ggda ggd	gaa Glu gac Asp catto ca	tct Ser aggg Arc ccagg ccagg ccagg ccagg ccact rtling rtli	gga Gly gac Asp gac gaag gaaa ggac ggac ggac ggac ggac	gac Asp cag Gin raagt cot gggc ttgat tcac tgat tcac tgat tcac tcac	caa Gln gat Asp cctta ccccc cgact ccccc ccccc ccccc ccccc ccccc ccccc cccc	Glu atc atcatta aga actatta agg actatta	gat Asp tag *** cagt ata tgtg tcta tcta tgta tcca tcca	2880 2939 3057 3116 3234 3293 3352 3411 3470 3529 3588 3647 3706

FIG. 5. (Legend appears at the bottom of the opposite page.)

The amino acid sequence of TAP was found to be identical to the unpublished sequence of p115 (8), a protein involved in intra-Golgi fusion (9, 10), indicating that TAP/p115 is a general factor operational at multiple membrane fusions.

TAP/p115 and Uso1p are present in the cytoplasm and associated with intracellular membranes (9, 11, 15, 22). TAP was initially identified as a component of TCVs (15), but subsequently we found that TAP is also present on secretory vesicles derived from the TGN and within the Golgi complex (unpublished results). p115 is detected predominantly within the Golgi complex in bovine fibroblasts (9) but its ultrastructural localization is currently unknown. Likewise, the subcellular localization of Uso1p has not been defined but it is likely that Uso1p is present on vesicles operational between the ER and the Golgi complex.

Vesicle targeting and fusion are dependent on reversible and ordered interactions between membrane receptors and soluble cytosolic factors (23, 24). Specifically, proteins present in the vesicles [v-SNAREs; e.g., synaptobrevin (SNARE, SNAP receptor)] would interact with a complementary protein (a t-SNARE; e.g., syntaxin) on the target membrane and this association would form the site for recruitment of the "fusion machinery," a complex that includes NSF and SNAPs (7, 25, 26). Our data showing that ATP-depletion allows the docking of TCVs to the target apical PM supports the hypothesis that the NSF-SNAP complex assembles onto docked vesicles.

How do we superimpose the requirement for TAP/p115 onto the proposed targeting/fusion scheme? Targeting specificities might result from the interactions of v-SNAREs and t-SNAREs. However, vesicle docking prior to targeting might demand the creation of a network of stable interactions between the vesicular and target membranes, facilitating the specific binding of v-SNAREs to t-SNAREs. We propose that TAP/p115 acts as a vesicular "anchor" by interacting with the target membrane and holding the vesicular and target membranes in proximity. This suggestion is based on data from the neuronal field; synaptic vesicles appear normally docked at the target membrane even when v-SNAREs are cleaved by clostridial toxins (and thus unable to form a stable complex), suggesting that SNAREs are not required for association of vesicles with the target membrane and other molecules must perform this function (27). We propose that TAP/p115 may provide such an activity.

We thank Dr. J. Rothman for advice and helpful insights. We are grateful to Drs. G. Banting and T. Dawson for their generous gifts of cDNA libraries. We acknowledge Dr. G. Waters for providing results prior to publication. We thank Drs. J. Schwarzbauer, D. Winkelman, and A. Lupas for scientific input. We acknowledge the assistance of the Princeton Sequencing Facility and Princeton Monoclonal Facility. This work was supported by grants from the National Institutes of Health and the National Science Foundation.

- 1. Barroso, M. & Sztul, E. S. (1994) J. Cell Biol. 124, 83-100.
- Apodaca, G., Katz, L. A. & Mostov, K. E. (1994) J. Cell Biol. 125, 67–86.
- Sztul, E., Colombo, M., Stahl, P. & Samanta, R. (1993) J. Biol. Chem. 268, 1876–1885.
- Sztul, E. S., Howell, K. E. & Palade, G. E. (1985) J. Cell Biol. 100, 1248–1254.
- 5. Musil, L. S. & Baenziger, J. U. (1987) J. Cell Biol. 104, 1725-1733.
- Musil, L. S. & Baenzinger, J. U. (1988) J. Biol. Chem. 263, 15799-15808.
- 7. Clary, D. O., Griff, I. C. & Rothman, J. E. (1990) Cell 61, 709-721.
- Sapperstein, S. K., Walter, D. M., Grosvenor, A. R., Heuser, J. E. & Waters, M. G. (1995) Proc. Natl. Acad. Sci. USA 92, 522–526.
- Waters, M. G., Clary, D. O. & Rothman, J. E. (1992) J. Cell Biol. 118, 1015–1026.
- Elazar, Z., Orci, L., Ostermann, J., Amherdt, M., Tanigawa, G. & Rothman, J. E. (1994) J. Cell Biol. 124, 415-424.
- 11. Nakajima, H., Hirata, A., Ogawa, Y., Yonehara, T., Yoda, K. & Yamasaki, M. (1991) J. Cell Biol. 113, 245-260.
- 12. Sztul, E. S., Howell, K. E. & Palade, G. E. (1983) J. Cell Biol. 97, 1582–1591.
- 13. Harlow, E. & Lane, D. (1988) Antibodies: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY).
- Sambrook, J., Maniatis, T. & Fritsch, E. F. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY).
- 15. Sztul, E., Kaplin, A., Saucan, L. & Palade, G. (1991) Cell 64, 81-89.
- Wattenberg, B. W., Balch, W. E. & Rothman, J. E. (1986) J. Biol. Chem. 261, 2202–2207.
- 17. Kozak, M. (1989) J. Cell Biol. 108, 229-241.
- 18. Lupas, A., Van Dyke, M. & Stock, J. (1991) Science 252, 1162–1164.
- Espreafico, E. M., Cheney, R. E., Matteoli, M., Nascimento, A. A. C., De Camilli, P. V., Larson, R. E. & Mooseker, M. S. (1992) J. Cell Biol. 119, 1541-1557.
- Warrick, H. M., De Lozanne, A., Leinwand, L. A. & Spudich, J. A. (1986) Proc. Natl. Acad. Sci. USA 83, 9433–9437.
- 21. Pierre, P., Scheel, J., Rickard, J. E. & Kreis, T. E. (1992) Cell 70, 887–900.
- Gow, A., Nelson, D., Rahim, S. & Sztul, E. (1993) *Eur. J. Cell Biol.* 61, 184–188.
- 23. Rothman, J. E. & Warren, G. (1994) Curr. Biol. 4, 220-232.
- Ferro-Novick, S. & Jahn, R. (1994) Nature (London) 370, 191– 193.
- Söllner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P. & Rothman, J. E. (1993) *Nature (London)* 362, 318-324.
- Whiteheart, S. W., Rossnagel, K., Buhrow, S. A., Brunner, M., Jaenicke, R. & Rothman, J. E. (1994) J. Cell Biol. 126, 945-954.
- 27. Südhof, T. C., De Camilli, P., Niemann, H. & Jahn, R. (1993) Cell 75, 1-4.

FIG. 5 (On opposite page). Sequence of TAP and homology alignments. (A) cDNA and protein sequence of rat liver TAP. Numbers on right correspond to nucleotide position and those on left correspond to amino acid position. Amino acid sequences of five peptides obtained from tryptic digestion of purified TAP are boxed. The two putative initiator methionines are marked with single asterisks. Stop codon is indicated by triple asterisks. Putative coiled-coil domains are in shaded letters. Proline-rich sequences are marked with double asterisks. Acidic domain is underlined. (B) Alignment of TAP and Uso1p N-terminal sequences. Homologous regions are boxed. (C) Alignment of TAP and Uso1p C-terminal sequences. Acidic domain is indicated by residues in boldface letters. Homologous region is boxed and a stretch of identity is double underlined.