Antigenic modules in the N-terminal S1 region of the transmissible gastroenteritis virus spike protein

Juan Reguera,1 Desiderio Ordoño,1 César Santiago,1 Luis Enjuanes2 and José M. Casasnovas1

1Department of Macromolecule Structure, Centro Nacional de Biotecnologia, CSIC, Campus Universidad Autónoma, Darwin 3, 28049 Madrid, Spain
2Department of Molecular Biology, Centro Nacional de Biotecnologia, CSIC, Campus Universidad Autónoma, Darwin 3, 28049 Madrid, Spain

The N-terminal S1 region of the transmissible gastroenteritis virus (TGEV) spike (S) glycoprotein contains four antigenic sites (C, B, D and A, from the N- to the C-terminal end) and is engaged in host-cell receptor recognition. The most N-terminal portion of the S1 region, which comprises antigenic sites C and B, is needed for the enteric tropism of TGEV, whereas the major antigenic site A at the C-terminal moiety is required for both respiratory and enteric cell tropism, and is engaged in recognition of the aminopeptidase N (APN) receptor. This study determined the kinetics for binding of a soluble S1 protein to the APN protein. Moreover, the S1 region of the TGEV S protein was dissected, with the aim of identifying discrete modules displaying unique antigenic sites and receptor-binding functions. Following protease treatments and mammalian cell expression methods, four modules or domains (D1–D4) were defined at the S1 region. Papain treatment identified an N-terminal domain (D1) resistant to proteolysis, whereas receptor binding defined a soluble and functional APN receptor-binding domain (D3). This domain was recognized by neutralizing antibodies belonging to the antigenic site A and therefore could be used as an immunogen for the prevention of viral infection. The organization of the four modules in the S1 region of the TGEV S glycoprotein is discussed.

INTRODUCTION

Coronaviruses (CoVs) are enveloped, positive-sense, ssRNA viruses (de Groot et al., 2008; Enjuanes et al., 2008; Masters, 2006) involved in respiratory, enteric, hepatic and neuronal infectious diseases in animals and humans that often lead to important economic losses (Perlman, 1998; Weiss & Navas-Martin, 2005). Since the outbreak of severe acute respiratory syndrome (SARS) in 2002, the interest in CoVs has increased dramatically as potential generators of new, serious zoonotic infectious diseases (Drosten et al., 2003; Holmes, 2005; Lau, 2004; Rota et al., 2003), making evident the necessity for a deeper understanding of the determinants responsible for CoV receptor recognition and tropism.

The CoV spike (S) glycoprotein is anchored to the envelope and is specialized both in receptor recognition and in virus–cell fusion (Bosch et al., 2003; Gallagher & Buchmeier, 2001). Receptor recognition appears to be species specific, and there is significant variability in receptor usage among CoVs. Several cell-surface glycoproteins have been identified as CoV receptors: members of the genus Alphacoronavirus such as transmissible gastroenteritis virus (TGEV) and human (HCoV-229E), canine and feline CoV use the aminopeptidase N (APN) receptor (Delsam et al., 1992; Tresnan et al., 1996; Yeager et al., 1992). Members of the genus Betacoronavirus such as human SARS-CoV use the human angiotensin-converting enzyme 2 (ACE2) (Li et al., 2003), whereas murine hepatitis virus (MHV) uses the cell adhesion molecule CEACAM1a (Yokomori & Lai, 1992). Nevertheless, the existence of alternative receptors that can confer an extended tropism has been revealed for several CoVs. SARS-CoV can use liver/lymph node-specific intercellular adhesion molecule-3-grabbing non-integrin (L-SIGN) for cell attachment and entry (Jeffers et al., 2004), whereas an MHV virus strain isolated from persistently infected cells (MHV/BHK) enters the cells in a heparan sulfate-dependent manner (de Haan et al., 2005; Miura et al., 2008). It has been suggested that binding of the TGEV S protein to sialic acids is responsible for its enteric tropism (Krempl et al., 1997).

The CoV S glycoproteins are type I membrane proteins that form trimers that protrude out of the virus envelope.
and give a crown-like appearance to the CoV particles (Beniac et al., 2006; Delmas & Laude, 1990). The trimeric spikes have a globular region mostly formed by the N-terminal S1 region, and a protein stalk connecting the globular region to the membrane, formed by the C-terminal S2 region (Beniac et al., 2006). The S1 region bears the receptor-binding epitopes (Godet et al., 1994; Wong et al., 2004), whereas S2 is responsible for viral and cell membrane fusion, and adopts a helical conformation characteristic of class I fusion proteins (Bosch et al., 2003; Supek et al., 2004). The CoV S protein is cleaved between the S1 and S2 regions in some CoVs during virus particle maturation (Abraham et al., 1990; Cavanagh et al., 1986; de Haan et al., 2008), but no cleavage was observed in alphacoronaviruses such as TGEV and canine and feline CoVs.

TGEV is a porcine CoV with respiratory and enteric tropism, whereas a naturally derived mutant of this virus, porcine respiratory CoV (PRCV) has only respiratory tropism (Laude et al., 1995; Sánchez et al., 1992). The S1 region of the PRCV S protein has a 224 aa deletion with respect to the TGEV glycoprotein at the N terminus, although both proteins are highly homologous and they bind to the porcine APN (pAPN) receptor. It has been suggested that the N-terminal region of TGEV S protein confers enteric tropism (Ballesteros et al., 1997; Callebaut et al., 1988; Sánchez et al., 1992), apparently related to sialic acid recognition (Krempl & Herrler, 2001; Krempel et al., 1997). The main antigenic profile of the S protein is well characterized for TGEV and other porcine CoVs (Correa et al., 1988; Delmas et al., 1990; Sánchez et al., 1990). The TGEV S protein harbours four antigenic sites, named from the N- to the C-terminal end as C, B, D and A, whereas the PRCV S protein lacks the N-terminal C and B antigenic sites (Callebaut et al., 1988; Sánchez et al., 1990). The position of each antigenic region has been mapped by analysis of mAb-resistant mutant viruses that either overcome antibody neutralization or lack reactivity to specific mAbs (Correa et al., 1990; Gebauer et al., 1991). Receptor-binding domains (RBDs) and neutralizing antigenic sites have been characterized for several CoVs (Breslin et al., 2003; Corapi et al., 1995; Delmas et al., 1990; Sánchez et al., 1990; Tsai et al., 2003; van den Brink et al., 2005; Wong et al., 2004; Wu et al., 2009; Zhang et al., 2004). The RBD of porcine CoVs binding to pAPN has been identified as a 150 aa fragment at the C-terminal moiety of the S1 region and includes the antigenic A site (Godet et al., 1994). These data confirmed the reported inhibition of TGEV binding to ST cells with site A-specific mAbs (Suñé et al., 1990).

The present work shows the dissection of the S1 region of TGEV by expression of soluble protein length variants, which defined four distinct modules or domains (D1–D4) in the N-terminal S1 region of the TGEV glycoprotein. Proper folding of the domains and their antigenic properties were proved with S1-specific mAbs and the pAPN receptor protein. A soluble and isolated domain (D3) displayed receptor-binding activity similar to the full S1 region. The kinetics for binding of the S1 protein to pAPN were determined.

**RESULTS**

**Dissection of the S1 region of the TGEV S glycoprotein**

Aiming to isolate folding modules displaying unique antigenic and functional properties linked to the S1 region of the TGEV S protein, length variants were designed following the antigenic map described with TGEV escape mutants (Correa et al., 1990; Gebauer et al., 1991; Sánchez et al., 1992) (Fig. 1a). We designed two TGEV-derived variants comprising all four antigenic sites: a longer variant comprising the complete S1 region (S1) and a shorter protein (S3) with its C-terminal end after the antigenic A site. Similar variants were also engineered for the same region of the PRCV HOL87 strain (S1H and S3H) (Fig. 1a), which has 96% sequence identity with the TGEV protein (see Supplementary Fig. S1, available in JGV Online). TGEV variants S4 and S5 had sequential deletion of antigenic sites A and D, respectively (Fig. 1a). An S4 protein derived from the PRCV HOL87 strain (S4H) and bearing only antigenic site D was also engineered (Fig. 1a). A short protein fragment including the RBD region and having just antigenic site A (SA) was produced (Fig. 1a).

Proteins were engineered fused to the Fc region of IgG1 (bivalent proteins) and to Flag or haemagglutinin (HA) epitopes (monovalent proteins) at their C-terminal ends, and were produced in 293T and Chinese hamster ovary (CHO) cells (see Methods). We monitored expression for most of the designed protein variants in the supernatant of transfected cells from which they were purified (Fig. 1b). Production of Fc fused to S4 (S4–Fc), S5 (S5–Fc) and SA (SA–Fc) was significantly higher than that of the HA- or Flag-tagged protein variants (not shown), but the longest variants (1 and 3) were not expressed fused to Fc. In addition to the proteins presented in Fig. 1(b), the S3H and S4H proteins derived from PRCV HOL87 strain were also secreted into the supernatant of transfected cells (not shown). Therefore, soluble protein expression allowed us to produce full-length protein and fragments of the S1 moiety of TGEV and PRCV containing unique antigenic sites.

An additional approach for the identification of protein domains was carried out following protease treatments. The S1 variant prepared in CHO Lec cells was treated with trypsin and papain proteases for the identification of fragments resistant to proteolysis. A papain-resistant fragment of about 35 kDa (S35) was seen following electrophoresis of the S1 protein treated with a high papain concentration (Fig. 2a), and a longer intermediate protein (S60, 60 kDa) was observed at a low protease concentration. Mass spectrometry determined that the
mean mass of the papain-resistant fragment was 35 421 Da. The same fragments were observed following treatment of the TGEV S3 variant, but the papain-resistant fragment was not seen with the S1H protein (not shown), which lacks the N-terminal 224 aa of S1 (Fig. 1a); therefore, it

must include the N-terminal region. Sequencing of S1 and the papain-resistant S35 and S60 fragments confirmed that they had the same N terminus (DNFP amino acid sequence). Therefore, they must have been derived from proteolysis of the C-terminal region of the S1 protein. Based on the S60 size, the cleavage site that leads to this intermediate fragment should be on the N-terminal side of antigenic site A. Mass spectrometry of the S35 protein treated overnight at 30 °C with N-glycosidase F (PNGase F; Biolabs), which hydrolyses N-linked glycans (not shown), gave a mass of 27 231 Da, suggesting that, most probably, its C-terminal residue is Ala240. The expected mass for the fragment comprising residues Asp1–Ala240 was 27 228 Da. The mass difference between the PNGase F-treated and untreated protein (8190 Da) was also consistent with the presence of six N-linked glycosylation sites with a mass of about 1343 Da (see Supplementary Fig. S1), attached to glycoproteins produced in CHO Lec cells (Stanley, 1989).

Circular dichroism (CD) of the purified S35 fragment gave an approximate secondary structure composition of 18 % α-helices, 30 % β-sheets and about 20 % β-turns (see Supplementary Fig. S2, available in JGV Online, and Methods). These data, together with its protease resistance properties and its absence in homologous PRCV, indicated

Fig. 1. Design and production of soluble CoV S proteins. (a) Designed S1 length variants derived from the enteric TGEV PUR-MAD and respiratory PRCV HOL87 strains. The open rectangle represents the S1 region, with the approximate location of the C, B, D and A antigenic sites indicated. Antibodies used in this study specific for each site are indicated above. N-Linked glycosylation sites are indicated with an upside-down ‘Y’. The S2 region at the C terminus of S1 is not shown. Solid lines show the soluble length variants designed in this study, with the name on the left and the C-terminal residue of the mature proteins on the right. They are 16 residues shorter than the immature S protein with the signal sequence. (b) 10 % SDS-PAGE under reducing conditions of the partially purified protein variants produced in CHO Lec cells, except for the S4–Fc fusion protein, which was prepared in 293T cells (see Methods). Coomassie blue staining is shown for the S1, S3, S1H, S4 and SA proteins, and an immunoblot (HA mAb) for the HA-tagged S5 variant. Electrophoresis of the S4–Fc protein untreated (−) or treated (+) with thrombin for release of the Fc portion is shown. The fusion protein, S4 and Fc fragments and contaminant immunoglobulins (lg) from the serum are labelled. The broader band of the isolated S4 fragment reflects high glycosylation heterogeneity because of its production in 293T cells. The sizes (kDa) and migration of the molecular mass markers are indicated.

Fig. 2. Protease-resistant fragments in the TGEV S1 protein. (a) 10 % SDS-PAGE of TGEV S1 protein incubated for 16 h at 37 °C in the absence (−) or presence of papain at several enzyme:protein ratios (w/w): 1:500 (1), 1:1000 (2), 1:2000 (3) and 1:10 000 (4). The size (kDa) and migration of the molecular mass markers are indicated. (b) Scheme showing the S1 fragments resulting from papain digestion shown in (a), with approximate sizes of 60 and 35 kDa. The arrows indicate the location of the cleavage sites. The length of the S35 fragment was determined by mass spectrometry to be 240 aa.
that the N-terminal portion must be an independent folding module in the TGEV S protein.

**Folding of S1 length variants proved with conformational antibodies**

We used a panel of mAbs (Sunè et al., 1990) specific for the four antigenic sites defined in the S1 region of TGEV and PRCV to analyse folding of the soluble protein variants (Fig. 1a). ELISAs were performed with purified S proteins and the mAbs 6A.A6 (C site), 1D.B12 (B site), 1D.G3 (D site) and 6A.C3 (A site) (Fig. 3a). All mAbs recognized the TGEV S1 protein, which contains all four antigenic sites (Fig. 1a), whereas mAbs for the antigenic C and B sites mapping to the N-terminal region failed to bind to the PRCV S1H protein. A mAb-binding profile similar to that of the S1 and S1H proteins was observed with the shorter S3 and S3H variants, respectively (not shown). Furthermore, the mAbs recognizing conformational epitopes (1D.B12, 1D.G3 and 6A.C3) (Correa et al., 1990; Gebauer et al., 1991; Sánchez et al., 1992) revealed that the antigenic B, D and A regions were properly folded.

Antibody binding to the S4–Fc and S5–Fc proteins was consistent with the location of the antigenic sites (Fig. 3a and Fig. 1), whereas S4H–Fc was just recognized by the site D mAb 1D.G3 (not shown). Unexpectedly, the S35 and S5 proteins lacking the Fc region were not recognized by mAb 1D.G3 (Fig. 3a) or the site B mAbs 1B.H11 and 8F.B3 (see Supplementary Fig. S3, available in JGV Online), although some binding to the S5 protein was recorded with the latter two mAbs. These results indicated a certain complexity for antigenic site B, as its recovery required the inclusion of an extended polypeptide chain at its C-terminal region, such as has been reported for protein members of the immunoglobulin superfamily with certain mAbs (Casasnovas et al., 1998). Recognition of the SA variant by mAbs 6A.C3 (Fig. 3a) and 1A.F10 (not shown) indicated that this virus protein fragment comprising antigenic site A must be folded correctly.

We observed that mAb 1D.G3 bound to the soluble S1H, but did not bind to PRCV particles (Fig. 3a, b) (Gebauer et al., 1991), indicating that the 1D.G3 epitope must be hidden in the trimeric spike at the PRCV envelope.

**Binding of S1 protein variants to soluble and membrane-bound pAPN**

Binding of the soluble pAPN receptor to plastic-bound TGEV S length variants was determined for a range of protein concentrations (Fig. 4a). The S1, S1H and SA proteins comprising the antigenic A site were recognized by the soluble pAPN protein, whereas the S5 protein was not (Fig. 4a). The binding difference observed between the S1 and S1H proteins at intermediate concentrations could be related to differences in the accessibility of the receptor-binding region in the plastic-bound proteins, but pAPN binding to both proteins at the highest concentration was similar. Binding of soluble pAPN to the S1 protein was blocked specifically by the site A mAb 6A.C3 (Fig. 4b). A similar binding activity of the SA protein and those proteins including the complete S1 region (S1 and S1H) indicated that it should include all the essential epitopes required for receptor recognition.

The involvement of S protein domains in receptor recognition was analysed further using permissive swine testicle (ST) and baby hamster kidney (BHK) cells with (BHK–pAPN cells) or without the pAPN protein on the cell surface. The SA–Fc and S5–Fc proteins were used for cell-binding analysis by flow cytometry (Fig. 4c). We observed specific binding of the SA–Fc protein to the ST and BHK cells expressing pAPN, whereas the S5–Fc protein did not bind to any of the cell lines.

**Kinetics of soluble S1 binding to the pAPN receptor protein**

Surface plasmon resonance was applied to characterize further S protein binding to the pAPN receptor protein (Fig. 5). About 800 resonance units (RU) of purified pAPN with an HA epitope at the N-terminal end were captured first onto a sensor chip with an immobilized HA-specific mAb (see Methods). The TGEV S1 protein was injected onto surfaces with or without pAPN, and the specific increase in the response related to S1 protein binding was monitored for surfaces on which pAPN was captured (Fig. 5).
The soluble S1 protein was injected at concentrations ranging from 1 to 6 μM, which gave binding responses of 50 to ~300 RU (Fig. 5, inset). The sensorgrams did not plateau, indicating that no steady-state binding was reached at the flows used. The binding kinetics were determined from analysis of the association and dissociation phases (Table 1 and Methods). The association kinetic rate ($k_{ass}$) was relatively low for protein–protein interactions, similar to that described for binding of the poliovirus receptor to poliovirus (Xing et al., 2000). The
Table 1. Affinity kinetic rate constants for binding of the TGEV S1 protein to pAPN

Affinity and kinetic rate constants were determined from sensograms recorded at 25 °C during injection of the TGEV S1 protein untreated or treated with endoglycosidase H (eh) through sensor chips having pAPN captured (about 1000 RU) with an HA mAb (see Fig. 5). The S1 protein was produced in CHO Lec cells (see Methods) and therefore contained only high-mannose carbohydrates that can be removed by endoglycosidase H treatment. Means (s) from four experiments carried out at two different flow rates are shown. $K_D$ was determined from the kinetic rates ($k_{on}/k_{off}$).

<table>
<thead>
<tr>
<th>Protein</th>
<th>$k_{on}$ (M$^{-1}$ s$^{-1}$)</th>
<th>$k_{off}$×10$^3$ (s$^{-1}$)</th>
<th>$K_D$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoV S1</td>
<td>17 000 (5140)</td>
<td>2.28 (0.25)</td>
<td>135</td>
</tr>
<tr>
<td>CoV S1eh</td>
<td>17 400 (4580)</td>
<td>2.00 (0.64)</td>
<td>115</td>
</tr>
</tbody>
</table>

Dissociation rate ($k_{off}$) was also low, indicating a strong binding interaction. The TGEV S1 protein was prepared in the lectin-resistant CHO Lec 3.2.8.1 cells and therefore contained only high-mannose carbohydrate (Stanley, 1989), which could be removed by endoglycosidase H treatment. We also determined the affinity of the endoglycosidase H-treated S1 (S1eh) protein, and no significant differences were observed after removal of the glycans (Table 1), showing that they did not contribute to recognition of the pAPN receptor protein.

**DISCUSSION**

In this study, we dissected the S1 region of the TGEV S protein and identified four modules bearing unique antigenic sites, based on expression of soluble S1 length variants and proteolysis experiments (Fig. 6a). Besides its resistance to papain treatment shown here, the natural deletion of the D1 region in certain PRCVs indicated that it must be a unique folding module. D3 could be expressed independently of the remaining S1 fragment and its correct folding was confirmed by both mAb and pAPN receptor binding. We assigned D2 to the protein region between D1 and D3 (Fig. 6a), which carries antigenic site D. This antigenic site was recovered when this region was expressed together with D1 (S4–Fc protein) or isolated (S4H–Fc protein), showing that it is a distinct module in the S1 region of the protein. The C-terminal moiety labelled D4 could be considered a kind of S1 linker to the S2 region of the S protein, although we did not test whether it can fold independently from the remaining S1 fragment.

D3 can fold independently and comprises both antigenic site A and the pAPN receptor-binding epitope. The receptor-binding activity of the SA protein that defined D3 was similar to that of the complete S1 region (Fig. 4a), and the protein could bind to both soluble and membrane-bound pAPN protein. Therefore, this domain should include all the critical residues required for recognition of the pAPN protein. Sequence identity of the TGEV D3 domain with the crystallized receptor-binding domain of SARS-CoV of the genus Betacoronavirus (Li et al., 2005) and NL63-CoV of the genus Alphacoronavirus (Wu et al., 2009) is around 10 and 24 %, respectively. Therefore, we would expect certain structural similarities of the TGEV domain with that of NL63-CoV, whose structure is quite distinct from the SARS-CoV receptor-binding region (Wu et al., 2009), even though both viruses bind to the ACE2 receptor protein.

Antigenic site A displays the major neutralization epitopes in TGEV, and antibodies clustered at this site can efficiently prevent virus binding to the pAPN receptor and infection (Suñé et al., 1990). The soluble SA protein fragment developed here could therefore be used as an immunogenic agent to elicit a neutralizing response. Its IgG fusion variant should have increased stability and therefore could enhance antigen-specific T-cell responses (Nam et al., 2010; Sibéri et al., 2007).

We determined the affinity and kinetics for binding of the TGEV S1 region to its pAPN receptor (Table 1). The determined $K_D$ of around 100 nM for monovalent binding indicated a relatively high-affinity virus–receptor interaction. Previously reported $K_D$ values for receptor binding to viruses range from about 100 nM to 1 µM (Lea et al., 1998; Myszka et al., 2000; Santiago et al., 2002; Xing et al., 2000). The $K_D$ is close to the 50 nM value estimated for binding of a bivalent SARS-CoV S1 fragment to its ACE2 receptor binding domain.
receptor (Wong et al., 2004). The kinetic rates reported here are similar to those described for receptor binding to poliovirus (Xing et al., 2000). The low \( k_{\text{ass}} \) rate (17 000 M\(^{-1}\) s\(^{-1}\)) indicated either a receptor–virus interface with low accessibility or certain conformational changes associated with the interaction. Nevertheless, as seen with poliovirus, TGEV-neutralizing antibodies can block receptor binding, so that the receptor-binding site in the virus must be accessible to antibody neutralization. The \( k_{\text{dis}} \) rate was also low and indicated a relatively strong binding interaction, as has been described for other virus–receptor interactions (Santiago et al., 2002; Xing et al., 2000). The carbohydrates of the heavily glycosylated S protein did not have any effect on receptor-binding affinity (Table 1); therefore, the pAPN-binding region on the S protein must be free of glycans.

Antibody binding to soluble S length variants showed a good correlation with antibody recognition of the S protein in TGEV, but some differences with the recognition of PRCV particles (Fig. 3). Antibodies mapping to the antigenic C, B, D and A sites in TGEV recognized the soluble proteins bearing those antibody epitopes. Nevertheless, mAb 1D.G3 bound efficiently to the soluble S1 and S1H proteins but poorly to the trimeric membrane-bound S protein in the HOL87 PRCV particles (Fig. 3). These data suggested that antigenic site D must be partially buried in the trimer formed by the S protein at the PRCV envelope but exposed in TGEV spikes, indicating that this region (D2) must be differentially displayed in the two related viruses (Fig. 6b). According to our domain arrangement model presented in Fig. 6(b), the N-terminal D2 domain in the PRCV spike is mostly inside the trimer, with the N terminus close to the S2 region. This proximity between the N-terminal region of the S protein and S2 was previously proposed based on MHV mutants (Grosse & Siddell, 1994). Differing from PRCV, TGEV should have D1 inside the trimer, with D2 being more exposed to antibody binding. In good agreement, the N-terminal antigenic C site at D1 was not exposed in freshly purified TGEV particles, although it became accessible to mAb 6A.A6 upon virus binding to plastic or following partial virus denaturation (Gebauer et al., 1991).

The dissection of the S1 region of the TGEV S protein has provided new insights into the modular architecture of the CoV S glycoprotein. We identified four modules or domains in the S1 region, which displayed unique antigenic sites and functions. D3 is specialized in the recognition of the pAPN receptor recognition, but the functional relevance of the D1 domain in recognition of additional receptor molecules is unclear. Binding to an enteric receptor molecule by the N-terminal D1 of TGEV has been suggested to confer its enteric tropism (Krempf & Herrler, 2001). However, the distinct domain arrangement in the S proteins of TGEV and PRCV proposed here could provide differential receptor-binding properties to those viruses in the enteric track. Our model opens up the possibility that binding sites for enteric cellular receptors in TGEV S1 may be allocated to D2, which is not accessible to ligands such as site D mAbs in the PRCV virions. Further structural studies are now required to determine the conformation of the modules defined here and their arrangement in the CoV S glycoprotein.

**METHODS**

**Cells.** Hybridoma, 293T, BHK and ST cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% inactivated FCS. Stable transfected CHO Lec 3.2.8.1 (CHO(Lec) cells (Stanley, 1989) were grown with selective medium containing 30–200 \( \mu \)M \( 1\)-methionine sulfoximine, as described previously (Casasnovas & Springer, 1995; Ordoño et al., 2006). BHK-pAPN cells have been described elsewhere (Delmas et al., 1993).

**Preparation of TGEV S1 length variants and pAPN protein.** The coding sequences for the S1 variants were prepared by PCR from the cDNA clones of virus strains SC11 (GenBank accession no. AJ271965) and HOL87 (Sánchez et al., 1990). The soluble region of the pAPN (residues 36–963 of reference sequence GenBank accession no. P15145) was amplified from the cDNA coding for the receptor protein. The recombinant cDNAs of proteins requiring an exogenous signal peptide for cellular secretion (SA and pAPN) were cloned in frame with the IgK leader sequence and an HA epitope in a vector derived from pDisplay (Invitrogen). All recombinant cDNAs were cloned in a vector derived from the pEF-BOS (Mizushima & Nagata, 1990) in frame with either a Flag (DYKDDDDK) or HA (YPYDVPDYA) epitope, or the human IgG1 Fc region at their C-terminus. A thrombin recognition sequence (LVPRGS) was introduced between the protein and the C-terminal tags. The recombinant vectors were transfected into 293T cells using the calcium phosphate method for transient protein expression (Pear et al., 1993). Transfected 293T cells were maintained in DMEM with 10% FCS or serum-free OPTI-MEM (Invitrogen) for up to 3–4 days post-transfection for protein production. The concentration of the proteins in the cell supernatants was determined by ELISA using purified proteins as reference.

The recombinant cDNAs with the Flag and Fc tags were cloned into the unique Sall and Ncol sites of the pBluescript SK vector and used for preparation of stable CHO(Lec) cells using the glutamine synthetase expression system (Casasnovas & Springer, 1995; Ordoño et al., 2006). Clones secreting the proteins into cell supernatants were selected by ELISA.

**Protein purification and papain digestion.** Proteins were purified from cell supernatants by affinity chromatography. The TGEV S proteins were purified with either mAb 6A.C3 or a Flag M2 mAb (Sigma) coupled to Sepharose, and the pAPN protein with the HA mAb 12AC5 (Roche).Fc fusion variants were purified using a protein A column (GE Healthcare). After loading the cell supernatants, the columns were washed with Tris/saline (pH 8.0) and the proteins eluted with glycine buffer (pH 3.0). The eluted proteins at neutral pH were concentrated and in some cases treated with endoglycosidase H at 30 °C for removal of N-linked glycans. Final size-exclusion chromatography with a Superdex 200 column was usually run with HBS buffer [20 mM HEPES (pH 7.5) and 100 mM NaCl]. Concentration of purified proteins was determined by absorbance at 280 nm, using the extinction coefficient given by its amino acid sequence.

The S1 protein in HBS buffer with 1 mM EDTA and 5 mM cysteine was treated with papain from Carica papaya (Boehringer Mannheim) at different protein:papain ratios (w/w). The reaction was stopped by the addition of E64 (Sigma) at a final concentration of 2 \( \mu \)M, and the
S35 fragment was purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare).

**Antibody- and pAPN-binding assays.** The TGEV S mAbs were purified from hybridoma supernatants using a Hi-Trap protein A column (GE Healthcare). Binding of mAb and pAPN to TGEV S protein variants was carried out in 96-well plates. Proteins (50 μl), either purified or in serum-free 293T cell supernatant, were first bound to wells by overnight incubation at 4 °C. The protein concentration ranged from 10 to 0.1 μg ml⁻¹ for mAb binding or at the concentrations indicated in Fig. 4 for pAPN binding. Wells were blocked with 2 % BSA in PBS, and antibody binding was carried out with the TGEV S mAb at a concentration of 5 μg ml⁻¹ in PBS with 1 % BSA for 1 h at 37 °C. The bound antibody was detected with a secondary biotin-labelled rabbit anti-mouse antibody, HRP-streptavidin and α-phenylenediamine (Zymed). Absorbance at 492 nm was monitored for the wells having S protein and corrected for the absorbance of wells without proteins. Receptor binding to the plastic-bound proteins was carried out with biotin-labelled pAPN protein (10 μg ml⁻¹) for 1 h at 37 °C, and the bound protein was monitored by absorbance determination as carried out for antibody binding.

Antibody binding to the plastic-bound TGEV and PRCV particles was performed as described elsewhere (Sánchez et al., 1990). Partially purified viruses were titrated with mAb 6A.C3 and added in similar amounts to the wells in 96-well plates for the antibody-binding experiments.

Binding of Fc fusion proteins to cell-surface-expressed pAPN was carried out by fluorescence activated cell sorting (FACS). Approximately 2 x 10⁵ cells in 250 μl PBS with 2 % BSA were incubated with 20 μg Fc protein overnight at 4 °C. The cells were washed three times with the binding buffer and incubated for 1 h with an Alexa Fluor 488-labelled goat anti-human Fc secondary antibody (Invitrogen). The cells were washed three times and analysed in a Beckman Coulter EPICS XL-MCL Laser 488 cytometer.

**Kinetics for binding of soluble S1 to pAPN.** Kinetic rate determination by surface plasmon resonance was carried out on a CM5 sensor chip using a BIAcore-X instrument. The HA mAb 12A.C5 was covalently immobilized on the sensor chip. The HA-tagged pAPN protein in HBS buffer was first captured on the sensor chip surface with immobilized mAb and the S1 protein was run at low flow rates of 10 and 20 μl min⁻¹ onto the surface with captured receptor. After each cycle of S1 protein binding to the captured pAPN, the sensor chip was regenerated by two injections of 50 mM citrate buffer (pH 3.0) for 1 min each. Successive binding cycles with a range of S1 protein concentrations (1–6 μM) were carried out in each experiment (Fig. 5). The binding sensorgrams recorded with surfaces lacking pAPN were subtracted from those recorded with captured pAPN receptor. Association and dissociation phases of the corrected sensorgrams were analysed using BIAevaluation software (BIAcore) and the kinetic rates determined by adjusting the sensorgrams to a Langmuir 1:1 binding model.

**CD.** CD spectra (mean of four scans) were recorded in the far-UV region using a J810 spectropolarimeter (Jasco) equipped with a Peltier-type cell holder. Measurements were performed at 20 °C using a scan rate of 20 nm min⁻¹, a response time of 4 s, a bandwidth of 1 nm and a protein concentration of 4.3 μM (1 mm path-length quartz cells). The buffer contribution was subtracted from the protein spectrum and the corrected curve was converted to mean residue ellipticity with a mean molecular mass per residue of 113.45 Da. Secondary structure content was estimated by deconvolution of the normalized spectrum with the GNNN program (Böhme et al., 1992) using a reference dataset with 33 proteins.

**ACKNOWLEDGEMENTS**

This work was supported by grants from MICINN of Spain to J. M. C. (BFU2005-05972 and BFU2008-00971) and L. E. (CIT-010000-2007-8 and BIO2007-00978). L. E. acknowledges support from NIH (ARRA-W00015845) and European Commission FP7 (EC grant no. 245141). J. R. was supported by the Juan de la Cierva program. D. O. was a recipient of a Severo Ochoa fellowship from Fundación Ferrer and 1-year JAE-CSIC fellowship. We acknowledge Nuria Cubells for technical assistance, the CNIO for providing a BIAcore instrument and Margarita Menéndez (IQRF-CSIC) for the CD experiments.

**REFERENCES**


Domains in the TGEV spike protein


