

EXPRESSION OF PORCINE TRANSMISSIBLE GASTROENTERITIS VIRUS GENES IN E.COLI  
AS  $\beta$ -GALACTOSIDASE CHIMAERIC PROTEINS

Paul Britton, David J. Garwes, Kevin Page and Jean Walmsley  
AFRC Institute for Animal Disease Research, Compton,  
Newbury, Berks, RG16 0NN, U.K.

INTRODUCTION

Transmissible gastroenteritis virus (TGEV) causes gastroenteritis in pigs of all ages but has a high mortality in neonatal piglets resulting in severe economical losses in affected pig farms. In piglets, under two weeks of age, the first clinical sign is usually vomiting 18-24h after infection rapidly followed by a diarrhoea, resulting in loss of weight and dehydration; death usually occurs after 2-5 days (Garwes, 1982). The virus is enveloped, contains a single stranded RNA genome of positive polarity and three structural proteins; a 200 000 dalton surface glycoprotein (the spike or peplomer protein), a 30 000 dalton glycoprotein associated with the viral envelope (the integral membrane protein) and a basic 47 000 dalton phosphorylated protein associated with the viral RNA (the nucleoprotein). During replication of the virus in infected cells four subgenomic species of RNA are produced of which three have been shown to produce the three structural proteins (Millson et al. unpublished results). The fourth RNA species does not appear to produce a polypeptide detectable either in infected cells or by in vitro translation of the RNA though it has the capacity to produce a polypeptide of about 33 000 daltons. The nucleoprotein gene has been copied from the smallest subgenomic RNA species, shown by in vitro translation to produce the nucleoprotein (Millson et al. unpublished results), into cDNA. The complete DNA sequence of the gene has been determined (Britton et al. unpublished results) and a 1.38kb fragment composed of 245 amino acids, corresponding to 68% of the complete gene product, has been fused to the 3' end of the E.coli lacZ gene. The fused genes produced a  $\beta$ -galactosidase-TGEV nucleoprotein fragment chimaeric protein in E.coli. The chimaeric protein has been purified and used to raise antibodies in mice which immune precipitated only the viral nucleoprotein confirming that the DNA sequence assumed to contain the nucleoprotein gene sequence does so. This method provides a useful source of TGEV antigen for virus diagnostic tests and viral protein for use in a research programme aimed at understanding the mechanisms of antigen processing and immune stimulation. The method is useful for confirming the presence of open reading frames (ORFs) and identifying the true product of genes on cDNA either suspected to contain a particular gene or for those previously identified by DNA sequencing.

## MATERIALS AND METHODS

### Virus Production and RNA Isolation

Secondary adult pig thyroid cells (APT/2) were cultured and infected with the British FS772/70 field isolate of TGEV as described previously (Pocock & Garwes, 1975). [<sup>3</sup>H]-labelled viral RNA was produced as described by Millson et al. (unpublished results) using a pig kidney cell line (LLC-PK1 cells) grown in Eagles minimal essential medium supplemented with 50mM N-2-hydroxy-ethylpiperazine-N'-2-ethanesulphonic acid (HEPES) pH6.8, 0.14% sodium bicarbonate and 1μgml<sup>-1</sup> penicillin, streptomycin, mycostatin at 37°C. Confluent cultures were infected with 1 - 10 p.f.u.'s per cell. After 2h the inoculum was removed and replaced with the same medium containing 1μgml actinomycin D to inhibit host cell RNA synthesis. After a further 2h incubation 25 μCi of [<sup>3</sup>H]-uridine (40-60 Ci mmol<sup>-1</sup>, TRK.410, Amersham International) were added per culture bottle. After a further 3h incubation the cells were lysed using guanidinium thiocyanate and the RNA collected as a pellet under 5.7M CsCl as described by Chirgwin et al. (1979). Poly(A)-containing RNA was isolated by chromatography on poly (U) Sepharose.

### Synthesis and Cloning of cDNA

TGEV cDNA was prepared from poly(A)-containing subgenomic RNA, isolated from infected cells, using an oligo(dT)-tailed pUC9 vector-primer produced in our laboratory (Britton et al. unpublished results). Essentially plasmid pUC9 was linearised using PstI, end labelled with 40-50 dT residues using dTTP and terminal transferase. One of the oligo(dT) tails was removed using BamHI and EcoRI to produce the vector-primer. The poly(A)-containing RNA was annealed to the vector-primer and first strand cDNA synthesis carried out using AMV reverse transcriptase. The second strand cDNA synthesis was performed using E.coli RNaseH and DNA polymerase I as described by Okayama and Berg (1982). The DNA was end repaired using T4 DNA polymerase, self ligated and transformed into E.coli cells. Clones containing TGEV cDNA were identified by colony hybridization using [<sup>32</sup>P]-labelled TGEV RNA and the origin of the cDNA identified by hybridization of [<sup>35</sup>S]-labelled recombinant plasmids to glyoxylated TGEV RNA species previously separated by agarose gel electrophoresis and northern blotted onto Biodyne membranes.

### Production of a β-galactosidase-TGEV Nucleoprotein Chimaera

A 1.38kb HindIII fragment, of TGEV cDNA, was purified by electroelution from an agarose gel, ligated into the expression vectors pUR290, pUR291 and pUR292 (Ruther and Muller-Hill, 1983), previously digested with HindIII and treated with alkaline phosphatase, and transformed into E.coli strain JM101 (supE thi Δ(lac-proAB) [F' traD36 proAB lacI<sup>q</sup> ZΔM15]; Messing, 1979). Recombinant plasmids containing the HindIII fragment were identified by colony hybridisation using a [<sup>35</sup>S]-labelled HindIII fragment. Plasmid DNA was isolated from colonies giving a positive signal, the presence of the HindIII fragment was confirmed by digestion with HindIII and the correct orientation of the insert determined using a variety of restriction endonucleases. Several transformants containing recombinant plasmids with the HindIII fragment in the correct orientation and originating from each of the three types of vectors were grown in the absence and presence of 1mM isopropyl-β-D-thiogalactopyranoside (IPTG) and the polypeptides produced were analysed as described by Britton et al. (1982) on 8% SDS-polyacrylamide gels.

### Purification of the Chimaeric Protein

The method used to purify the chimaeric β-galactosidase was a modification of that described by Steers et al. (1971) using affinity chromatography on p-aminophenyl-β-D-thiogalactopyranoside-agarose. Ten

100ml cultures of JM101(pPBU43) were grown in Luria broth containing  $100\mu\text{gml}^{-1}$  ampicillin to an absorbance of 0.5 at 680nm. The *lacZ* gene was induced by the addition of IPTG (1mM) and growth continued for a further 3h. The cells were centrifuged and resuspended in 50ml of buffer B (0.2M Tris-HCl, 0.25M NaCl, 10mM magnesium acetate, 10mM 2-mercaptoethanol and 5% glycerol at pH7.6). Lysozyme was added to a concentration of  $0.4\text{mgml}^{-1}$  and the cells incubated at 4°C for 30min. The protease inhibitors phenyl methyl sulphonyl fluoride (PMSF) and benzamidine were added to final concentrations of 1mM and 5mM respectively and the cells were sonicated. The viscosity of the cell lysate was reduced by the addition of 1mg of RNase A and DNase I followed by incubation at 4°C for 45min. The cell lysate was clarified by centrifugation at 30 000 rpm for 40min at 4°C and dialysed against buffer D (10mM Tris-HCl, 0.25M NaCl, 10mM MgCl<sub>2</sub>, 1mM Na<sub>2</sub>EDTA, 10mM 2-mercaptoethanol, 0.1% Triton X-100, 5mM benzamidine and 1mM PMSF at pH7.6). The sample was then diluted to 100ml with buffer D and applied to a p-aminophenyl-β-D-thiogalacto-pyranoside-agarose column (Product number A-8648, Sigma) previously equilibrated with buffer D. The column was washed with 100ml of buffer D followed by a further 100ml of buffer D without Triton and the chimaeric β-galactosidase eluted with 60ml of 0.1M sodium borate at pH 10. Samples from various fractions were analysed by SDS-PAGE (Britton et al., 1982) and for β-galactosidase activity as described by Miller (1972). Fractions containing the chimaeric protein were pooled, dialysed against buffer D and stored at 4°C.

#### Production of Antisera

Purified chimaeric protein was dialysed against phosphate buffered saline (PBS) pH7.2 and used to immunise Balb/C mice as described by Garwes et al. (1987).

#### Immune precipitation of TGEV Antigens,

TGEV polypeptides were synthesised in LLC-PK1 cells in the presence of [<sup>35</sup>S]-methionine, lysed in RIPA buffer (PBS containing 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1mM methionine and 1mM PMSF), clarified by centrifugation and immune precipitated using mouse serum from controls and animals immunised against either purified TGE virus or the chimaeric protein. Fixed Staph. aureus cells were allowed to adsorb the immune complexes as described by Garwes et al (1987). The bound antigens were analysed by polyacrylamide gel electrophoresis and detected by fluorography.

#### Western Blotting of Antigens

Antigens were separated by polyacrylamide gel electrophoresis, transferred to nitrocellulose in a Trans-Blot apparatus and detected with antisera employing the biotin - <sup>125</sup>I-labelled streptavidin system, as described by Garwes et al (1987).

#### Sequencing of the Junction Between the E.coli lacZ Gene and TGEV cDNA

A restriction fragment from the *Eco*R1 site within the *E.coli lacZ* gene to the *Xba*I site in the TGEV cDNA was subcloned from plasmid pPBU43 into *Eco*R1 and *Xba*I digested sequencing vectors M13mp10 and M13mp11 (Messing, 1983). M13/dideoxynucleotide sequencing (Sanger et al., 1977; Bankier & Barrel, 1983) was carried out using [ $\alpha$ -<sup>35</sup>S]-dATP. The products of the sequencing reactions were analysed on buffer gradient gels (Biggin et al., 1983). A sonic digitiser was used to read data into a BBC model B microcomputer, using a program written by the AFRC Computing Centre, and the data were analysed on a VAX 11/750, using the programs of Staden (1982).

## RESULTS

### Construction and Identification of the Chimaeric Protein

The chimaeric protein was constructed using the pUR290, pUR291 and pUR292 expression plasmids constructed by Ruther and Muller-Hill (1983). The expression plasmids contain the complete *E. coli lacZ* gene, including the *lac* promoter, with the two stop codons replaced by a polylinker insert containing the unique restriction sites *Bam*H1, *Sal*I, *Pst*I, *Hind*III and *Cl*aI followed by new stop codons. The *Pst*I site normally present in the ampicillin resistance gene was removed by NG-mutagenesis. In each plasmid the polylinkers, at the end of the *lacZ* gene, differ slightly in length resulting in nucleotide sequences, encoding different amino acid extensions. The unique restriction sites, present in the polylinkers, are out of phase with each other allowing the insertion of contiguous open reading frames in all three reading frames. The plasmids were also constructed with an *Eco*R1

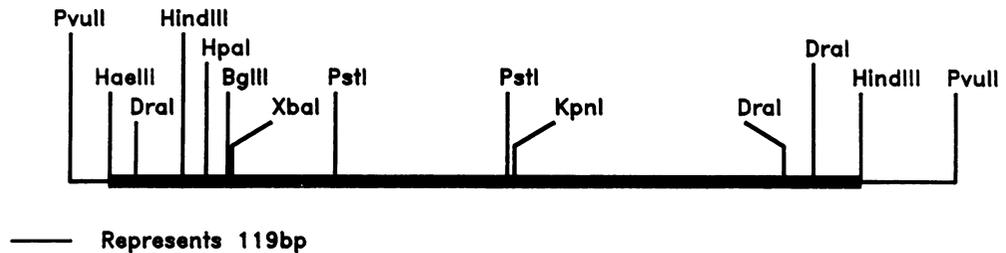


Fig.1. Restriction map of the TGEV cDNA in plasmid pTS15-1. The thick line represents the TGEV cDNA and the thin line part of the pUC9 DNA. The *Hind*III site at the right hand side is within the pUC9 DNA and has the poly (A) tail a few bases to the left. The *Hind*III site at the left hand side is towards, but within, the 5' end (N-terminal end) of the TGEV nucleoprotein gene.

site just after the polylinkers allowing any DNA inserted into the region to be removed as a result of another *Eco*R1 site being present near the end of the *lacZ* gene. The cloning of a DNA insert containing a suspected incomplete ORF into all three plasmids should result in a recombinant plasmid with the ORF contiguous with the *lacZ* gene producing a chimaeric protein consisting of  $\beta$ -galactosidase and the gene product of the inserted DNA fused to the carboxy terminus. The advantage of cloning into these vectors is that the  $\beta$ -galactosidase moiety is still active, allowing purification of the protein by affinity chromatography utilising the  $\beta$ -galactosidase activity. As a result of the fusion being at the 3' end of the *lacZ* gene, corresponding to the carboxy terminus of  $\beta$ -galactosidase, any stop codons within the inserted DNA do not stop the production of  $\beta$ -galactosidase as observed with fusions to the amino terminus of  $\beta$ -galactosidase.

A recombinant plasmid, pTS15-1 (Fig.1), isolated from cDNA copied from the smallest subgenomic RNA was found to contain a cDNA insert of 1.6kb. The plasmid hybridised to all TGEV subgenomic RNA species and to [<sup>32</sup>P]-labelled olig(dT)<sub>12-18</sub> (Britton et al. unpublished results). The 1.38kb HindIII fragment was ligated into all three expression vectors, as described above, and recombinant plasmids isolated with the insert in both orientations in all three types of expression vectors. The correct orientation of the HindIII fragment was determined by the fact that one of the HindIII sites, in the pUC9 DNA of plasmid pTS15-1, is very near to the poly(A) tail of the cDNA. The second HindIII site, present in the TGEV cDNA, therefore, should be within and near the 5' end of the nucleoprotein gene. A plasmid, pPBU43, with the HindIII fragment in the correct orientation in expression vector pUR290, produced a polypeptide of apparent molecular weight 150 000, after induction with 1mM IPTG. The molecular weight of a β-galactosidase monomer is 116 000 daltons implying that the polypeptide produced by plasmid pPBU43 consists of β-galactosidase monomer with a polypeptide extension corresponding to a molecular weight of 34 000. Plasmids containing the HindIII fragment in the opposite orientation in expression vector pUR290 and in the other two expression vectors in either orientation did not direct the synthesis of a larger sized β-galactosidase.

#### Immunological Characterisation of the Chimaeric Protein

Strain JM101(pPBU43) was grown in the presence of 1mM IPTG and [<sup>35</sup>S]-methionine in order to label the induced β-galactosidase chimaera. The cells were incubated with lysozyme, lysed by osmotic shock and the bacterial proteins immune precipitated, as described for TGEV proteins, using anti-TGEV serum and any adsorbed proteins were analysed by SDS-PAGE. Two polypeptides of molecular weights 116 000 and 150 000 were detected but no polypeptides were detected in IPTG induced cultures containing plasmid pUR290, indicating that the β-galactosidase chimaera contained TGEV antigens. The polypeptide of molecular weight 116 000 is the monomeric form of β-galactosidase, of which four identical monomers associate to give active β-galactosidase. The active β-galactosidase precipitated, therefore, must consist of both the 116 000 subunit derived from either the lacZ gene (with the small 5' deletion) carried by the F' plasmid in strain JM101 or from the chimaeric protein after proteolytic cleavage and the chimaeric protein. The chimaeric protein was purified by affinity chromatography utilising the β-galactosidase activity, as described above. Antibodies were raised in mice against the purified product and found to react with TGEV antigens by RIA. The immunological properties of the antibodies were tested by Western blotting and radio immune precipitation assay (RIPA).

(i) Western Blotting. Cell extracts of IPTG-induced JM101(pPBU43), TGEV-infected LLC-PK1 cells, purified β-galactosidase (product number 540-5531LA, BRL) and purified TGEV virus were fractionated by SDS-PAGE and blotted onto nitrocellulose. Following quenching as described above the blots were incubated with mouse control serum, anti-chimaeric protein serum and anti-TGEV serum followed by incubation with biotinylated sheep anti-mouse immunoglobulins and <sup>125</sup>I-labelled streptavidin, molecular weights of the polypeptides were calculated using Sigma molecular weight standards (Product No MW-SDS-200). As can be seen from Fig. 2 the control mouse serum did not detect any polypeptides to any extent. The serum containing antibodies raised against the chimeric protein detected several polypeptides present in the cell extracts. The main polypeptides were of molecular weights 116 000 and 150 000 corresponding to β-galactosidase monomer and the chimaeric protein. Several polypeptides of molecular weights less than 40 000 were detected in the extract from JM101(pPBU43), possibly originating from the proteolytic degradation of the chimaeric protein; one had a molecular weight of 33 000 corresponding to the theoretical size of

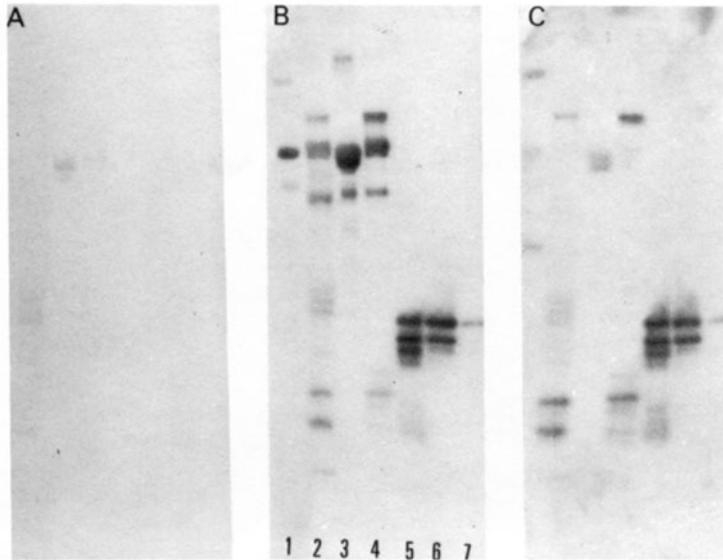


Fig. 2. Autoradiograph of polypeptides bound to nitrocellulose and detected using control mouse serum (A), mouse immune serum raised against either the chimaeric protein (B) or purified TGEV antigens (C). The binding of any mouse proteins were detected with biotinylated sheep anti-mouse immunoglobulins and  $^{125}\text{I}$ -labelled streptavidin. The tracks in each panel were: 1 Molecular weight standards (contained  $\beta$ -galactosidase), 2 Cell extract from IPTG induced JM101(pPB43), 3 Purified  $\beta$ -galactosidase from BRL, 4 Purified chimaeric protein, 5 Cell extract from TGEV infected LLC-PK1 cells, 6 Cell extract from TGEV infected APT/2 cells, 7 Purified TGE virus.

the nucleoprotein fragment as deduced from the DNA sequence. This polypeptide was also detected in the purified chimaeric protein sample though not with coomassie blue staining. A polypeptide of molecular weight 93 000 was detected in the JM101(pPB43) cell extract, the purified chimaeric protein sample and in purified  $\beta$ -galactosidase implying it was a breakdown product of  $\beta$ -galactosidase. Two main polypeptides were detected in virus infected cells of molecular weights 47 000 and 44 000, the former corresponding to the viral nucleoprotein and the second to the nucleoprotein related product previously described (Garwes et al. 1984). The only polypeptide detected in the purified TGEV sample was of molecular weight 47 000 corresponding to the viral nucleoprotein. The anti-TGEV serum detected the polypeptides of molecular weights 150 000 and 33 000 in both the *E.coli* extract and the purified chimaeric protein samples. The antiserum detected  $\beta$ -galactosidase monomer in the *E.coli* extract and purified chimaeric protein samples but only to a slight extent. It also detected the purified  $\beta$ -galactosidase sample but not to the same extent as the anti-chimaeric protein serum, possibly indicating that high concentrations of  $\beta$ -galactosidase may non-specifically bind the immunoglobulins. Polypeptides of molecular weights 47 000 and 44 000 were detected in TGEV infected cells but only the 47 000 polypeptide was detected in the purified virus sample. The peplomer protein was only slightly detected and the integral membrane polypeptide was not detected either due to low levels of these antigens blotted onto the nitrocellulose, coupled with some loss of antigenicity due to denaturation, or low levels of antibodies against these antigens in the serum sample. Detection of the 44 000 polypeptide in the infected cell sample and not in the purified virus sample supports the evidence that this polypeptide is related to the viral nucleoprotein.

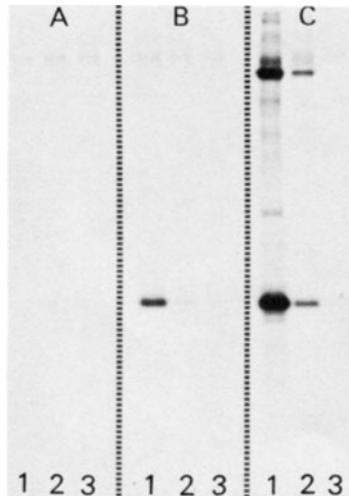


Fig. 3. Fluorograph of [<sup>35</sup>S]-methionine labelled TGEV polypeptides detected by RIPA using control mouse serum (A), anti-chimaeric protein serum (B) and anti-TGEV serum (C). The tracks (1-3) show the polypeptides precipitated from serial dilutions ( $10^{-1}$ - $10^{-3}$ ) of the antigen samples.

(ii) Radio Immune Precipitation Assay. Control mouse, anti-chimaeric protein and anti-TGEV sera were used to immune precipitate either [<sup>35</sup>S]-methionine labelled E.coli proteins from IPTG induced JM101(pPBU43) or cell extracts from TGEV infected LLC-PK1 cells grown in the presence of [<sup>35</sup>S]-methionine as described above. The sera were allowed to react with the cell extracts and formalinised Staph.aureus cells were added to adsorb immune complexes which were then analysed on polyacrylamide gels as described above. Control mouse serum detected only two polypeptides of molecular weights 116 000 and 150 000, in the E.coli extract but to only a very low extent. This suggests that the detection is due to non-specific adsorption or that the mouse serum may contain low levels of antibodies against  $\beta$ -galactosidase arising from a natural E.coli infection. The anti-chimaeric and anti-TGEV sera detected polypeptides of 116 000 and 150 000 from the E.coli extract again indicating that the active  $\beta$ -galactosidase consists of both the monomeric form of  $\beta$ -galactosidase and the chimaeric polypeptide. From Fig. 3 it can be seen that the control mouse serum did not detect any polypeptides in the TGEV infected cell extract. The anti-chimaeric protein serum detected only a single polypeptide of 47 000 in the TGEV infected cell extract. The anti-TGEV serum detected polypeptides of 200 000 and 47 000 in the TGEV infected cell extract.

#### DNA Sequence of the Junction Between the lacZ Gene on pUR290 and the TGEV cDNA

A restriction fragment from the EcoR1 site within the lacZ gene to the Xba1 site in the TGEV cDNA from plasmid pPBU43 was cloned into the DNA sequencing vectors M13mp10 and M13mp11 to determine the DNA sequence of the fragment in both directions. The restriction fragment consisted of the end of the lacZ gene, the polylinker to the HindIII site and 90bp of TGEV cDNA from the HindIII site to the Xba1 site from plasmid pTS15-1. A contiguous reading frame (Fig.4) was identified consisting of the amino acids found in  $\beta$ -galactosidase, corresponding to DNA down stream of the EcoR1 site, the

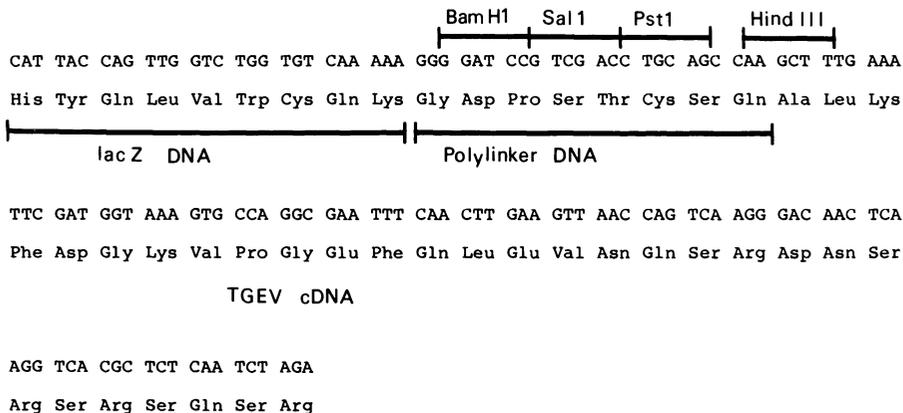


Fig. 4. The DNA sequence from part of the EcoRI-XbaI restriction fragment from plasmid pPBU43. The sequence shows the end of the lacZ gene to the HindIII site within the polylinker DNA from plasmid pUR290. The sequence from the HindIII site is of TGEV cDNA from plasmid pTS15-1. The amino acid sequence deduced is identical to that from the same region of the complete TGEV nucleoprotein gene.

polylinker of pUR290 and a new stretch of amino acids. The new stretch of amino acids were identical to those encoded by the DNA sequence from the HindIII site in the complete DNA sequence, determined from our strain of TGEV, of the TGEV nucleoprotein (Britton et al. unpublished results) and that determined by Kapke and Brian (1986) for the Purdue strain of TGEV.

#### DISCUSSION

The expression vectors used facilitate the construction of fusion proteins by the presence of a number of restriction sites at the 3' end of the lacZ gene. The three vectors have these restriction sites in all three reading frames allowing the insertion of a DNA fragment without prior knowledge of its DNA sequence or the need for the sequential addition of linkers or homopolymer tailing in order to form contiguous reading frames as observed with other expression vectors. Identification and isolation of the fusion proteins are aided by the expression of the lacZ gene, after induction with IPTG. The high molecular weight of the  $\beta$ -galactosidase monomer is advantageous, as there are few cellular proteins of greater size, for identifying the production of a chimaeric protein. As the  $\beta$ -galactosidase moiety of the chimaera is still active, purification by affinity chromatography is relatively simple.

A HindIII restriction fragment from TGEV cDNA believed to contain part of the viral nucleoprotein was cloned into the expression vectors before the DNA sequence was determined. The production of a chimaeric protein was identified by SDS-polyacrylamide gel electrophoresis after induction of the lacZ gene. Western blotting and RIPA data showed that anti-TGEV serum bound to the chimaeric protein indicating that it contained TGEV antigens, although the use of anti-TGEV polyclonal serum did not allow the

identification of the antigen. Production of antibodies against the chimaeric protein produced immunoglobulins against the  $\beta$ -galactosidase moiety and against the polypeptide originating from the TGEV cDNA. The anti-chimaeric protein serum only detected a polypeptide of molecular weight 47 000, corresponding to the viral nucleoprotein, when tested against the TGEV structural proteins. The DNA sequence data showed that the polypeptide, fused to the carboxy terminus of  $\beta$ -galactosidase, was identical to the amino acid sequence subsequently determined from the HindIII site of the DNA sequence, of the smallest TGEV subgenomic RNA species, known to contain TGEV nucleoprotein.

Thus the production of fusion proteins by the method described above provides a useful way of correctly identifying open reading frames. The procedure can be used for the production of polyclonal and monoclonal antibodies to unknown open reading frame products allowing the investigation into the function of unknown gene products and the identification of the whole gene product. Antibodies raised against a fragment of an unknown gene product in a fusion protein may also be used for the purification of the complete gene product using immuno-affinity chromatography. The fusion proteins can also play a useful role in the mapping of monoclonal antibody binding sites leading to the identification of specific epitopes (Mole and Lane, 1985). Interpretation of data from monoclonal antibodies produced against fusion proteins in which the polypeptide would normally be glycosylated may present problems as the conformation of the polypeptide in the fusion protein may be completely different from the natural product giving rise to different epitopes not found on the normal product.

The method is also a convenient source of TGEV antigen for use in a diagnostic test utilising the  $\beta$ -galactosidase activity. Immunoglobulins from the sera of test animals are bound to microtitre plates. The  $\beta$ -galactosidase chimaera is then allowed to bind to any immobilised TGEV antibodies and its presence detected using o-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG) which is converted from a colourless to a yellow compound in the presence of  $\beta$ -galactosidase.

Work is presently being carried out to characterise fusion proteins obtained from cDNA derived from viral RNA known to encode the integral membrane and peplomer proteins.

#### ACKNOWLEDGEMENTS

This research was supported by the Biomolecular Engineering Programme of the Commission of the European Communities Contract No. GB1-2-089-UK(H). We would also like to thank Eli Lilly Research Ltd for providing a one year studentship for Miss J. Walmsley and Dr. T. Harris Cell Tech for providing the pUR expression plasmids.

#### REFERENCES

- Bankier, A. and Barrell, B.G., 1983, Shotgun DNA sequencing, In: "Techniques in the Life Sciences (Nucleic Acid Biochemistry)," B508:1. R.A. Flavell, ed., Elsevier, Ireland.
- Biggin, M.D., Gibson, T.J., and Hong, G.F., 1983, Buffer gradient gels and <sup>35</sup>S label as an aid to rapid DNA sequence determination, Proc. Nat. Acad. Sci., USA, 80:3963.

- Britton, P., Murfitt, D., Parra, F., Jones-Mortimer, M.C. and Kornberg, H.L., 1982, Phosphotransferase-mediated regulation of carbohydrate utilisation in *Escherichia coli* K12: identification of the products of genes on the specialised transducing phages *iex(crr)* and *gsr(tgs)*, *EMBO J.*, 1:907.
- Chirgwin, J.M., Przybila, A.E., MacDonald, R.S., and Rutter, W.J., 1979, Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease, *Biochemistry*, 18:5294.
- Garwes, D.J., 1982, Coronaviruses in Animals, In: "Virus Infection of the Gastrointestinal Tract," D.A.J. Tyrrell and A.Z. Kapikian, eds., Marcel Dekker, Inc., New York.
- Garwes, D. J., Bountiff, L., Millson, G.C., and Elleman, C.J., 1984, Defective replication of porcine transmissible gastroenteritis virus in a continuous cell line, *Adv. Exp. Med. and Biol.*, 173:79.
- Garwes, D.J., Stewart, F. and Elleman, C.J., 1987, Identification of epitopes of immunological importance on the peplomer of porcine transmissible gastroenteritis virus, In: "Coronaviruses," M. Lai, ed., Plenum Press, New York.
- Kapke, P.A. and Brian, D.A., 1986, Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene, *Virology*, 151:41.
- Messing, J., 1979, A multipurpose cloning system based on the single-stranded DNA bacteriophage M13, *Recombinant DNA Technical Bulletin*, NIH 2:43.
- Messing, J., 1983, New M13 vectors for cloning, In: "Methods in Enzymology," 101:20, R. Wu, L. Grossman and K. Moldave, eds., Academic Press, New York.
- Miller, J.H., 1972, "Experiments in Molecular Genetics," Cold Spring Harbor Laboratory, New York.
- Mole, S.E. and Lane, D.P., 1985, Use of simian virus 40 large T-B-galactosidase fusion protein in an immunological analysis of simian virus 40 large T antigen, *J. Virol.*, 54:703.
- Okayama, H., and Berg, P., 1982, High-efficiency cloning of full-length cDNA, *Molec. and Cell. Biol.*, 2:16.
- Pocock, D.H., and Garwes, D.J., 1975, The influence of pH on the growth and stability of transmissible gastroenteritis virus *in vitro*, *Arch. Virol.*, 49:239.
- Ruther, U., and Muller-Hill, B., 1983, Easy identification of cDNA clones, *EMBO J.*, 2:1791.
- Sanger, F., Nicklens, S., and Coulson, A.R., 1977, DNA sequencing with chain terminating inhibitors, *Proc. Nat. Acad. Sci., USA*, 74:5463.
- Staden, R., 1982, Automation of the computer handling of gel reading data produced by the shotgun method of sequencing, *Nucleic acid Res.*, 12:521.
- Steers Jr, E., Cuatrecasas, P., and Pollard, H.B., 1971, The purification of  $\beta$ -galactosidase from *Escherichia coli* by affinity chromatography, *J. Biol. Chem.*, 246:196.