

Sequence of the nucleoprotein gene from a virulent British field isolate of transmissible gastroenteritis virus and its expression in *Saccharomyces cerevisiae*

P. Britton*, R. S. Cármenes,[†] K. W. Page, D. J. Garwes and F. Parra[†]

Division of Microbiology, A.F.R.C. Institute for Animal Disease Research, Compton Laboratory, Compton, nr. Newbury, Berkshire, RG16 0NN, UK.

Summary

Subgenomic mRNA from a virulent isolate of porcine transmissible gastroenteritis virus (TGEV) was used to produce cDNA which was sequenced. Two non-overlapping open reading frames (ORFs) were identified. The largest, encoding a polypeptide of 382 amino acids (relative molecular mass (M_r) 43 483), was shown to be the viral nucleoprotein gene. The second ORF, found 3' to the larger ORF, encodes a polypeptide of 78 amino acids (M_r 9068) which has yet to be assigned to a viral product. The nucleoprotein gene was expressed in yeast cells under the control of two types of yeast promoters: the constitutive PGK promoter, and the inducible *GAL1* promoter. Yeast cells containing recombinant plasmids, with the nucleoprotein gene in the correct orientation, produced a polypeptide of M_r 47 000, identical to the viral product, that reacted with a specific monoclonal antibody.

Introduction

The coronaviruses comprise a large group of enveloped positive-stranded RNA viruses and cause a variety of infections in a wide range of animal hosts (Garwes, 1982; Siddell *et al.*, 1983). Transmissible gastroenteritis virus (TGEV) is a coronavirus that causes gastroenteritis in pigs, resulting in a high mortality among neonates. TGEV is composed of a capped polyadenylated single-stranded RNA genome of M_r 6.8×10^6 or 20 kb in length (Garwes *et al.*, 1975; Dennis and Brian, 1981) and three major structural polypeptides: a surface glycoprotein (spike or peplomer protein) with a monomeric M_r 200 000, a glycosylated integral membrane protein observed as a series of polypeptides of M_r 28 000–31 000, and a basic phos-

phorylated protein (the nucleoprotein) of M_r 47 000 associated with the viral RNA (Garwes and Pocock, 1975). Infection of susceptible cells with TGEV results in the synthesis of a 'nested' set of subgenomic mRNA species with common 3' termini but different extensions in the 5' direction, a phenomenon also observed with other coronaviruses such as infectious bronchitis virus (IBV, Stern and Kennedy, 1980), though the size and number of the mRNA species vary with the virus. TGEV appears to have, in addition to the genomic sized RNA species, four major (11.2, 3.9, 2.6 and 1.7 kb) and two low abundant (3.0 and 0.7 kb) mRNA species (Britton *et al.*, 1986; Jacobs *et al.*, 1986).

Each coronavirus subgenomic mRNA species is believed to direct the synthesis of a protein from the 5' proximal region (Sturman and Holmes, 1983; Stern and Sefton, 1984). *In vitro* translation studies have shown that the mRNA species of 1.7 kb directs the synthesis of the viral nucleoprotein (Britton *et al.*, 1986; Jacobs *et al.*, 1986). Kapke and Brian (1986) have sequenced the 3' end of genomic RNA from the avirulent Purdue strain of TGEV and identified two complete open reading frames within the first 2022 bp of the genome. One of the open reading frames encodes a polypeptide of M_r 43 426 and has properties similar to those found in the IBV and mouse hepatitis virus (MHV) nucleoproteins. Here we report the production of cDNA from subgenomic mRNA species isolated from TGEV-infected cells using a virulent British isolate of TGEV.

E. coli cells are frequently unable to process eukaryotic proteins correctly, often producing the protein in the wrong conformational structure as a result of incorrect folding, leading to the degradation of foreign gene products. Several heterologous proteins have been successfully expressed in *Saccharomyces cerevisiae*, including eukaryotic proteins (calf chymosin, Mellor *et al.*, 1983) and viral antigens (hepatitis B surface antigen (HBsAg), Valenzuela *et al.*, 1982; Miyanojara *et al.*, 1983). *S. cerevisiae* cells are easy to grow and unlike *E. coli* cells have the advantage that they lack any pathogenicity, making them particularly useful as a source of large quantities of hormones and antigens of interest in human and veterinary medicine. We have expressed the nucleoprotein gene in yeast cells and the product was shown to be identical to the TGEV nucleoprotein by immunoblotting using a specific monoclonal antibody.

Received 25 June, 1987; revised 20 September, 1987. *For correspondence. [†]Present address: Departamento de Bioquímica, Universidad Oviedo, Oviedo, Spain.

Results

Cloning from TGEV mRNA species

TGEV poly(A)-containing mRNA species were fractionated on non-denaturing sucrose gradients and TGEV cDNA, derived from the 1.7 kb mRNA species, was synthesized by the vector-primer system, and isolated as described in *Experimental procedures*. Out of 417 colonies transformed with recombinant DNA, 36 contained DNA that hybridized to ^{32}P -labelled TGEV poly(A)-containing RNA. Three recombinant plasmids, shown to contain cDNA inserts of 1.5 kbp and designated pTS13-2, pTS15-1 and pTS15-2, were labelled with [^{35}S]-dATP by nick translation and hybridized to glyoxylated TGEV mRNA species separated on 1% agarose gels as described in *Experimental procedures*. All three recombinant plasmids were found to hybridize to all TGEV RNA species. In contrast, other recombinant plasmids derived from different mRNA species only reacted to one or more TGEV RNA species. This observation indicated that pTS13-2, pTS15-1 and pTS15-2 contained regions of cDNA complementary to regions present on all the TGEV 'nested' mRNA species. This phenomenon can only occur if the cDNA was derived from the 3' end of the RNA and since the sizes of the cDNA inserts were about 1.5 kbp, then most of the 1.7 kb mRNA species was copied. Plasmid pTS15-1 was purified and a restriction map constructed (Fig. 1A). The end containing the poly(A) tail was confirmed by probing Southern blots with ^{32}P -labelled oligo(dT)₁₂₋₁₈ oligonucleotides.

TGEV poly(A)-containing mRNA species isolated from infected cells were primed using a 90 bp *Hind*III-XbaI fragment purified from the 5' end of the TGEV cDNA cloned

in plasmid pTS15-1. The cDNA generated was end-repaired and cloned into plasmid pUC9 following the addition of *Eco*RI linkers. A recombinant plasmid, pF4F-36 (Fig. 1B), was identified and found to contain a cDNA insert of 1.15 kbp with restriction sites at one end identical to those at the 5' end of plasmid pTS15-1 (Fig. 1A), indicating that the two cDNA fragments were contiguous.

Sequencing of TGEV cDNA

The cDNA from plasmids pTS15-1 and pF4F-36 was subcloned into M13mp vectors using either the shotgun method (Bankier and Barrell, 1983) or by using specific restriction fragments and sequencing entirely in both directions. The corresponding DNA sequence 37 bp from the 5' end of the first open reading frame to the poly(A) tail is illustrated in Fig. 2. The cDNA was translated in all six reading frames, of which only the translation in the virus-sense strand revealed any open reading frames. The longest open reading frame, initiating from the ATG at position 38, of 1149 bp overlapped both cDNA inserts and encoded a polypeptide of 382 amino acids, with an M_r of 43 483. The basic nature and size of the polypeptide indicated that it corresponded to the viral nucleoprotein which is observed to have M_r 42 000 and 47 000 in TGEV-infected cells and M_r 47 000 in virions (Garwes *et al.*, 1984). The difference between the theoretical and predicted molecular weights of the protein may result from the degree of phosphorylation. Garwes *et al.* (1984) for TGEV and Stohlman and Lai (1979) for MHV have shown that the nucleoproteins are phosphorylated on serine residues, which represent 41 (11%) of the residues in strain FS772/70 and 39 (10%) in the Purdue strain of TGEV (Kapke and Brian,

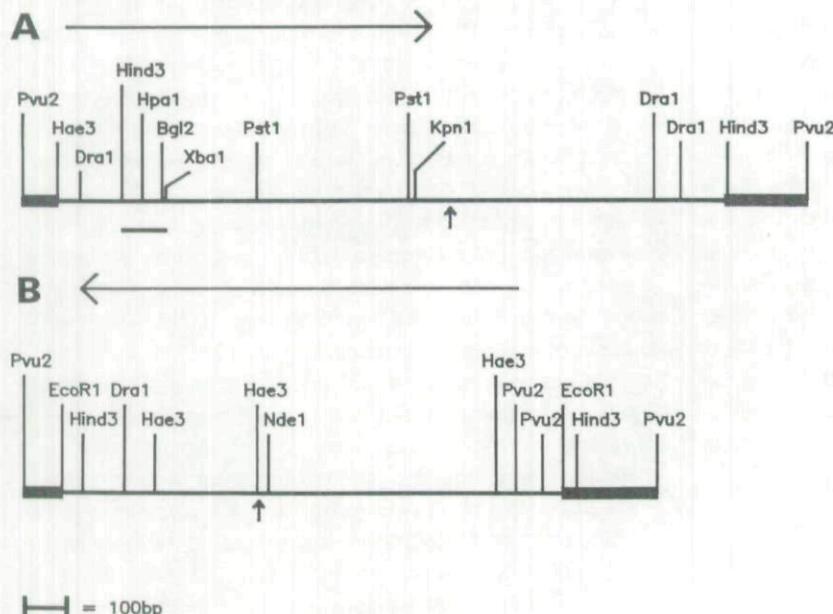


Fig. 1. Restriction endonuclease maps of (A) pTS15-1 and (B) pF4F-36. The thick black lines represent DNA from the vector pUC9. The large arrows show the direction of the TGEV cDNA inserts, as compared to the viral genome, when mapped in pUC9. The small arrows show on (B) the start of the nucleoprotein gene and on (A) the start of the small open reading frame postulated to produce the polypeptide of M_r 9068. The bar under pTS15-1 indicates the 90 bp fragment used to synthesize pF4F-36.

AGAAAAATTATACATATGGTATAACTAACTCTAAATGGCCAACCAGGGACAACGTGTTAGTGGGGGATGAATCCACAAAATCGTGGTCGCTCCATTCCGTGGTCGGPAGAG 120
 E K L L H M V * M A N Q G Q R V S W G D E S T K I R G R S N S R G R K S * *
 Purdue strain T N
 TAATAACATACCTCTTCATTCTCAACCCCATAACCTCCAGCAAGGTCAAAAATTGGAACTTATGTCGAGAGACTTGTACCCAAGGAATAGGTAAACAGGATCAACAGATTGG 240
 N N I P L S F F N P I T L Q Q G S K F W N L C P R D F V P K G I G N R D Q Q I G
 TTATGGAATAGACAAACTCGCTATCGCATGGTGAAGGGCAACGTAAAGAGCTTCTGAAAGGTGGTCTCTACTACTTAGGCACTGGACCTCATGCAGATGCCAATTIAAGATAA 360
 Y W N R Q T R Y R M V K G Q R K E L P E R W F F Y Y L G T G P H A D A K F K D K
 ATAGATGGAGTTGCTGGGTGCAAGGATGGTGCATGAAACAAACCAACACGCTGGTAGTCGTTGCTAATAATGAATCCAAGCTTGAAATCGATGGTAAAGTCCAGGG 480
 L D G V V W V A K D G A M N K P T T L G S R G A N N E S K A L K F D G K V P G E
 ATTCAACTTGAAGTAAACCACTGCAAGGGACAACCTCAGGTACCGCTCAATCTAGATCTGGCTAGAAATAGATCTAGAGGCAAGCAATCCAATAACAAGAAGGTGA 600
 F Q L E V N Q S R D N S R S Q S R S R S R N R S Q S R G R Q Q S N N K R D D F
 CAGTGAGAACAGCTGTTCTGCGCACTTAAAAGTTAGGTGACACAGAAAACAGCAACGCTCTGTTCTAAATCTAAAGAACGTAGTAACTCTAAACAAGAGATACTAC 720
 S V E Q A V L A A L K K L G V D T E K Q Q Q R S C S K S K E R S N S K T R D T T R
 GCTTAAGATGAAAACAAACACACTGGAAGAGAACTGAGGTAAAGGTGATGACAAGAATTATGGAGCTAGAAGCAGTCAGCCAATTGGTGACAGTGACCTCGTIGCCAATGG 840
 P K N E N K H T W K R T A G K G D V T R F Y G A R S S S S A N F G D S D L V A N G T
 GAGCAGTGCAAGCATTACCCACAATGGCTGAATGTTCTAGCATCTGTTAGCATTTGGAAAGCTATTGGACTCAAAAGGAAGATGGCAGCAGATAAGTCAGTTACACA 960
 S S A K H Y P Q L A E C V P S V S S I L F G S Y W T S K E D G D Q I E V T F T H
 CAAATACCACTTGCAAAGGATGATCTAAACTGAACTTCTAGCATCTGCTATGCTCTAGCATTTGGAAAGCTATTGGACTCAAAAGGAAGATGGCAGCAGATAAGTCAGTTACACA 1080
 K Y H L P K D D P K T E Q F L Q Q I N A Y A R P S E V A K E Q R K R K S R S K S G
 TGAGAAAGGTAGAGCAAGAGGGTGTACCTGATGCTTAATAGAAAATTACAGATGTTGATGACACACAGGTGAGATGATGAGGTAAAGCAACTAACGAGATGCTCGTC 1200
 A E R S E Q E V V P D A L I E N Y T D V F D D T Q V E M I D E V T N * M L V I
 CTCTCCATGCTGTTTACAGTTAACCTTACTACTTGTAGACTTCATTAGAAAGATTACTTAATCATTCATTCTCAATCTAAACTGTTGCTGATTAAATTC 1320
 L L H A V F I T V L T L L L G R L Q L L E R L L L N H S F N L K T V A D F N I
 TTATATAGGAGTTAGCAGAAACAGATTACTAAAGTGGCTCGATTAATCTTTAGTCATTAGGATTGCTATAGATGTTAGTCATATAATGTAAGGCAACCCATTGCAATTAGGAAGTTAGAATTG 1440
 L Y R S L A E T R L L K V V L R L I F L V L L G F C C Y R L L V I L M * T V
 TCTAAAATGGTTTCCGAGGAACACTGGCATCGCGTGTACTCTGTACAGAATGTAAGCACGTGTAATAGGAGGTACAAGCAACCCATTGCAATTAGGAAGTTAGAATTG 1560
 GATTGGCAATGCTAGTTAGTAATTAGAGAAAGTTAACGATCCCTACGACGAGCAACATGGAAGAGCTAACGTCGGTACTAGTGAATTGTTAAATGAAAATTGTTGAAAAT 1680
 TTTCTTTGATAGTGTATACAAAAAAAAAAAAAAAAAAAAAA
 1740

Fig. 2. The nucleotide sequence derived from part of the cDNA insert in pF4F-36 and the cDNA insert from pTS15-1. The amino acid sequences are part of the integral membrane protein (bases 1-25) as postulated by Kapke and Brian (1986), the nucleoprotein (bases 39-1185) and a hypothetical hydrophobic protein (bases 1192-1428). The ACTAAC and TGGAAGAGCT sequences postulated to be involved in RNA transcription are explained in the text. Amino acids below the major sequence are substitutions found in the Purdue strain (Kapke and Brian, 1986). The boxed amino acids are identical residues found between IBV (Boursnell *et al.*, 1985) and MHV (Skinner and Siddell, 1983). The broken lines refer to the two major and weaker areas of homology between the three nucleoproteins derived by dot-matrix analysis (Fig. 3). These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00542.

1986). The polypeptide sequence has three areas of homology with the nucleoproteins from IBV (Boursnell *et al.*, 1985) and MHV (Skinner and Siddell, 1983) as analysed by dot matrix (Fig. 3). Previous work (Britton *et al.*, 1987) showed that antibodies raised against a purified chimaeric protein, produced from the pTS15-1 *Hind*III fragment (1.38 kbp) fused to the 3' end of the *lacZ* gene, immunoprecipitated TGEV nucleoprotein. This confirmed that the cDNA carried on pTS15-1 contained most of the nucleoprotein gene.

A second open reading frame, present only in pTS15-1, is found 6 bp 3' to the nucleoprotein gene. This second

open reading frame, initiating from the ATG at position 1192, is composed of 237 bp and encodes a polypeptide of 78 amino acids of M_r 9068. Initial analysis using a universal codon usage table (Staden and McLachlan, 1982), or a codon usage table constructed from the TGEV nucleoprotein gene, suggested that the polypeptide might not be produced because of its poor codon usage. The smaller open reading frame should be present on an mRNA species of 0.7 kb; the smallest mRNA species detected in TGEV-infected cells is the 0.7 kb low-abundance species (Britton *et al.*, 1986), showing that a polypeptide of M_r 9068 could be produced by this mRNA species.

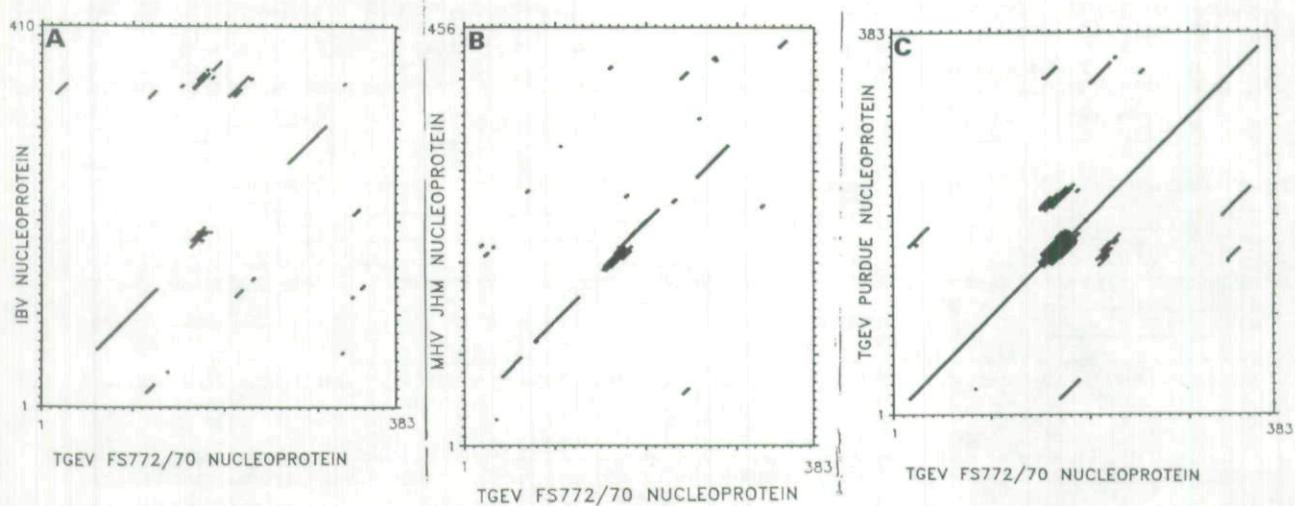


Fig. 3. Comparison of the sequences of coronavirus nucleoproteins from three antigenically distinct subgroups using the UWGCG programs COMPARE and DOTPLOT (Devereux *et al.*, 1984).

A. IBV Beaudette nucleoprotein (Boursnell *et al.*, 1985) and TGEV FS772/70 nucleoprotein.
 B. MHV JHM nucleoprotein (Skinner and Siddell, 1983) and TGEV FS772/70 nucleoprotein.
 C. TGEV Purdue nucleoprotein (Kapke and Brian, 1986) and TGEV FS772/70 nucleoprotein.
 The comparisons used a window of 30 residues with a stringency of 8.

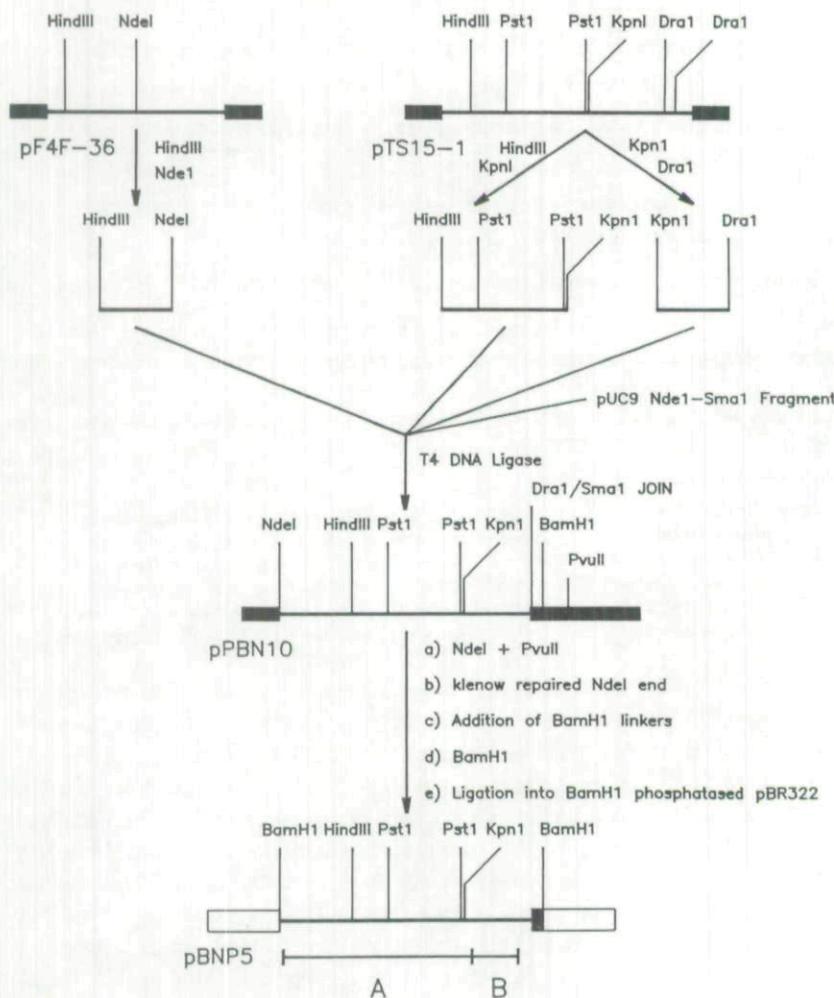


Fig. 4. Schematic representation of the construction of the TGEV nucleoprotein gene BamHI cassette. The thin line represents the TGEV cDNA, the thick line represents part of pUC9 DNA, and the open box represents part of pBR322 DNA. (A) shows the position of the nucleoprotein gene and (B) the position of the open reading frame capable of producing a polypeptide of M, 9000.

Construction of the TGEV nucleoprotein gene

Plasmids pTS15-1 and pF4F-36 were digested with restriction endonucleases to produce fragments that did not overlap with pUC9 fragments and the nucleoprotein gene was constructed by cloning the fragments into the purified 2.5 kbp *Nde*I-*Sma*I fragment from plasmid pUC9 (Fig. 4). A recombinant plasmid, pPBN10, was found to consist of the truncated pUC9 vector with an insert consisting of the fragments from plasmids pTS15-1 and pF4F-36 in the correct order. Plasmid pPBN10 contained the TGEV nucleoprotein gene initiating 19 bp 3' from the *Nde*I site. The plasmid also contained the smaller 3' ORF which initiates, in a different reading frame, 6 bp 3' from the end of the nucleoprotein gene and terminates 166 bp 5' to the *Dra*I/*Sma*I combined site between the TGEV cDNA insert and the pUC9 DNA.

The 1.76 kbp *Nde*I-*Pvu*II fragment from pPBN10, containing the complete nucleoprotein gene, was purified, repaired using the Klenow fragment of *E. coli* DNA polymerase I, and *Bam*HI linkers were added (Fig. 4). The resulting 1.58 kbp DNA fragment was gel-purified and cloned into the *Bam*HI site of plasmid pBR322. A recombinant plasmid, pBNP5 (Fig. 4), consisting of the 1.58 kbp *Bam*HI fragment in pBR322, was used as a source of the nucleoprotein gene in a *Bam*HI cassette.

Construction of yeast recombinant plasmids for expression

Under a constitutive promoter. The 1.58 kbp *Bam*HI fragment from plasmid pBNP5 was ligated into the *Bgl*II site of the yeast expression plasmid, pMA91, to position it downstream of the constitutive yeast PGK promoter. Two recombinant plasmids were identified as having the 1.58 kbp *Bam*HI fragment in opposite orientations in pMA91. Plasmid pYNP1 consisted of the TGEV nucleoprotein in the correct orientation for expression from the yeast PGK, and pYNP2 had the *Bam*HI fragment in the opposite orientation so that expression from the PGK promoter was prohibited.

Under an inducible promoter. The 1.58 kbp *Bam*HI fragment from plasmid pBNP5 was ligated into the *Bam*HI site of the yeast expression plasmid pB620 to position it downstream of the galactose inducible yeast *GAL1* promoter. A recombinant plasmid, pYNG1, was found to contain the TGEV nucleoprotein gene in the correct orientation for expression under the control of the yeast *GAL1* promoter.

The expression plasmid pB620 does not contain a yeast terminator sequence, unlike plasmid pMA91, which is believed to improve the efficiency of expression. The 0.72 kbp *Bgl*II-*Bam*HI fragment, containing the PGK terminator

sequence, was purified from plasmid pMA91 and ligated onto the 1.58 kbp *Bam*HI fragment from pBNP5. The resulting mixture was electrophoresed into an agarose gel and a DNA fragment corresponding to 2.3 kbp was purified and ligated into the *Bam*HI site of plasmid pBR322. A recombinant plasmid, pBRNP2, was isolated which contained the 0.72 kbp fragment in the correct orientation and fused onto the correct end of the TGEV nucleoprotein gene-containing fragment. The 2.3 kbp *Bam*HI fragment was cloned into the *Bam*HI site of plasmid pB620 as described for the 1.58 kbp *Bam*HI fragment. Two recombinant plasmids were isolated which had the 2.3 kbp fragment in opposite orientations. Plasmid pYNG2 (Fig. 5) consisted of the fragment in the correct orientation for expression from the yeast *GAL1* promoter but differed from plasmid pYNG1 in that it now had the PGK termination sequence spliced to the other end. Plasmid pYNG3 was similar to pYNG2 except that the fragment was in the opposite orientation prohibiting expression of the TGEV nucleoprotein gene from the *GAL1* promoter.

Transformation and stability of recombinant plasmids in yeast cells

The recombinant plasmids pYNP1, pYNP2, pYNG1, pYNG2 and pYNG3 and expression vectors pMA91 and pB620 were transformed into the *S. cerevisiae* strain BWG1-7A by the spheroplast method. Yeast transformants were selected at 30°C for their ability to grow in the absence of leucine (pMA91, pYNP1 and pYNP2) or uracil (pB620, pYNG1, pYNG2 and pYNG3). The generation time of the yeast strain before or after transformation with any of the plasmids was the same (3.5 h in synthetic media). The stability of the plasmids in the yeast strain was determined by growing the cells in complex non-selective media (YPD) and plating the cells on either synthetic selective media (SD) or on YPD. After seven generations in YPD, 90% of cells conserved pMA91 and its derivatives pYNP1 and pYNP2 and 70% of cells conserved pB620 and its derivatives pYNG1, pYNG2 and pYNG3. In order to check that the yeast cells had not modified the recombinant plasmids in any way, plasmid DNA was rescued from the yeast cells, amplified through DH1 *E. coli* cells and analysed by restriction endonuclease digestion. None of the plasmids rescued from the yeast cells appeared to be modified.

Detection of TGEV nucleoprotein in yeast cells

Yeast cells transformed with the above plasmids were grown in the presence of 2% glucose or 2% galactose to show the production of any polypeptides under the control of the yeast promoters. To determine whether any TGEV-specific polypeptides could be detected in transformed

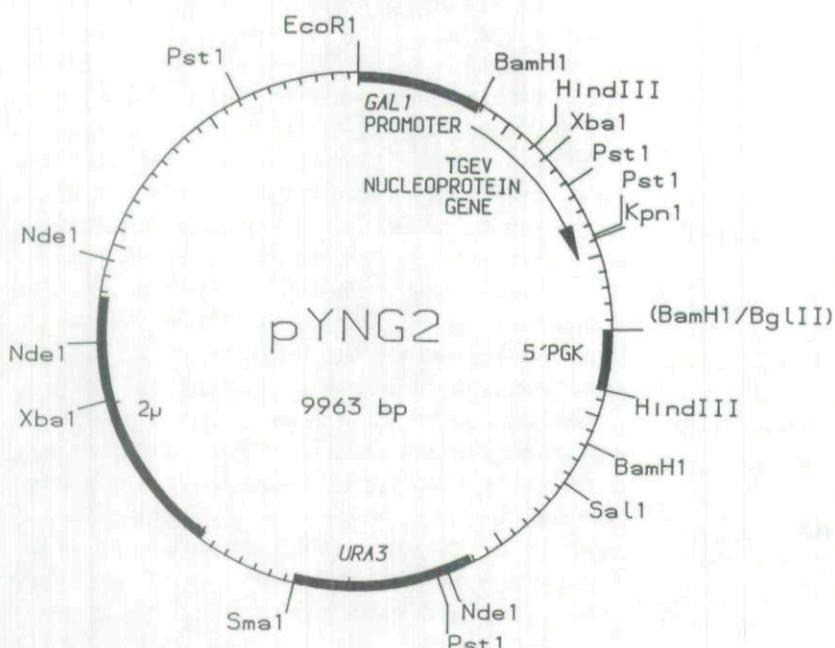


Fig. 5. Restriction map of the recombinant yeast expression plasmid, pYNG2. The TGEV nucleoprotein gene is represented by an arrow, sequences derived from yeast DNA are shown as thick lines, and sequences derived from pBR322 are shown as thin lines. pYNG1 was similar, except that it does not contain the sequences between the second BamHI site and the BamHI/BglII junction. pYNG3 contained the TGEV BamHI insert in the opposite orientation.

yeast the cells were disrupted with glass beads and the cell lysates were dot blotted onto nitrocellulose membranes and probed with a specific mouse monoclonal antibody as described in *Experimental procedures*. Yeast cells transformed with pYNP1 grown in the presence of 2% glucose or 2% galactose and cells transformed with pYNG1 and pYNG3 grown in the presence of 2% galactose but not 2% glucose reacted with the monoclonal antibody. None of the other cell lysates reacted with the antibody. The nucleoprotein was also detected in yeast spheroplasts using this monoclonal antibody by immunofluorescence, and in cell extracts by enzyme-linked immunoabsorbance assay (ELISA) (data not shown).

Cell lysates were prepared from cells transformed with pYNP1 and pYNP2 grown in the presence of 2% glucose, and pYNG1, pYNG2 and pYNG3 grown in the presence of 2% glucose or 2% galactose and analysed on 12% SDS-polyacrylamide gels. The fractionated polypeptides were electrophoretically transferred onto nitrocellulose membranes which were probed as described. From Fig. 6A it can be seen that the TGEV nucleoprotein gene cloned in the correct orientation in plasmid pYNP1 and in plasmids pYNG1 and pYNG2, in the presence of 2% galactose, produced a polypeptide of M_r 47000 detected by the nucleoprotein monoclonal antibody DA3. The nucleoprotein gene cassette cloned in the opposite orientation did not produce the polypeptide of M_r 47000 or any other polypeptide detectable with DA3 under any conditions. Expression from the GAL1 promoter appeared to be more efficient than from the PGK promoter. The presence of the PGK termination sequence following the nucleoprotein

gene under the control of the GAL1 promoter did not seem to increase the efficiency of expression. There was no distinguishable difference in the sizes between the TGEV nucleoprotein produced in the yeast cells and that present in purified viral particles (Fig. 6B). Nucleoprotein degradation products, of which the main one had an M_r of 37000 (Fig. 6A), also reacted with the nucleoprotein monoclonal antibody.

Discussion

The complete nucleoprotein gene, from a British field isolate of TGEV, was cloned over two cDNA fragments. Both cDNA clones were sequenced and two potential open reading frames were identified in the viral sense strand. The initiation codons of both open reading frames were preceded by the heptameric sequence ACTAAC, which is similar to the sequences preceding the nucleoprotein, matrix, spike genes and other open reading frames (whose products have not yet been identified) in IBV and MHV genomes. The heptameric sequence contains the hexameric sequence, CTAAAC, found by Kapke and Brian (1986) for the Purdue strain of TGEV. It has been suggested that this sequence is involved in the initiation of mRNA synthesis for other coronaviruses (Shieh *et al.*, 1987).

The sequence context of the initiation codon in the nucleoprotein gene, TAAATGG, often occurs among functional eukaryotic initiator sequences, and the sequence at the initiation codon for the second open reading frame, GAGATGC, is also favoured for eukaryotic initiation

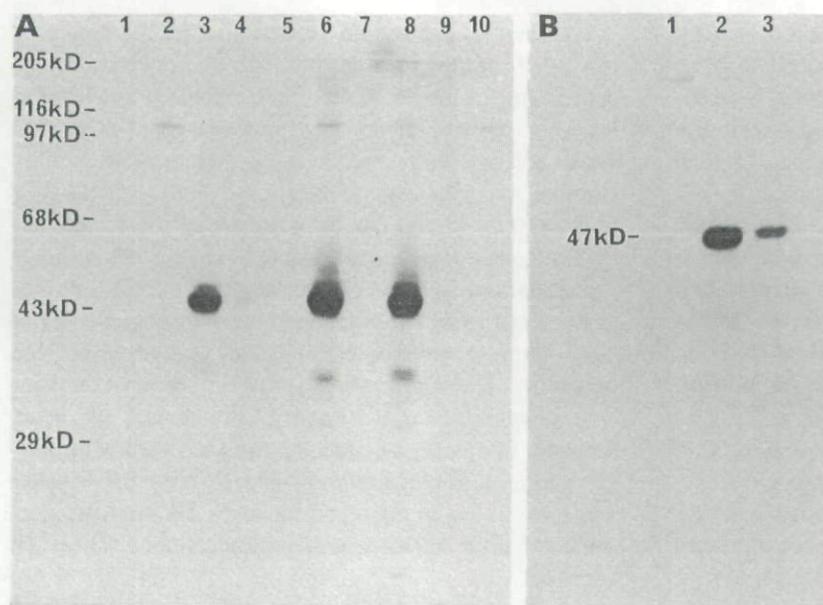


Fig. 6. Detection by immunoblotting of TGEV nucleoprotein in yeast cells and comparison with the TGEV nucleoprotein present in purified virions. Samples were probed with a specific monoclonal antibody as described in the text.

A. The cell-free extracts were as follows: lanes 1 and 2 are from the yeast recipient cell; lanes 3 and 4 cells transformed with pYNP1 and pYNP2; lanes 5 and 6 cells transformed with pYNG1; lanes 7 and 8 cells transformed with pYNG2; and lanes 9 and 10 cells transformed with pYNG3. The cells used for lanes 1, 3, 4, 5, 7 and 9 were grown in the presence of glucose and cells used for lanes 2, 6, 8 and 10 were grown in the presence of galactose. B. Lanes 1 and 2 are extracts from cells transformed with pYNG3 and pYNG2 grown in the presence of galactose. Lane 3 contained fractionated proteins from purified TGEV virions.

sequences (Kozak, 1983). Both TGEV genes terminate with the same single stop codon, TAA. The 10-base sequence, TGGAAGAGCT, observed by Kapke and Brian (1986), found 69 bp from the poly(A) tail, is similar to that found in both IBV (81 bp from the poly(A) tail) and MHV (82 bp from the poly(A) tail) and may be a recognition site or binding site involved in the synthesis of the negative RNA strand (Boursnell *et al.*, 1985; Kapke and Brian, 1986).

Knowledge of the amino acid sequences of gene products from different viral strains can help to locate domains in the proteins conserved for their structure and function. Comparison of all the genes between virulent and avirulent strains may be useful for the identification of sequences involved in the pathogenicity of the virus. The amino acid sequence of the nucleoprotein from FS772/70 and Purdue strains of TGEV differ by 2.1% and most of the changes are relatively conservative. However, the change at base 636 results in an amino acid substitution of cysteine for strain FS772/70 to arginine for Purdue. There are three areas of homology between the nucleoproteins of TGEV (strains FS772/70 or Purdue), IBV (Beaudette) and MHV (JHM) (Figs. 2 and 3), indicating that all three coronavirus subgroups share common ancestry. The areas of homology, shown in Fig. 2, between TGEV amino acid residues 65-123 and 257-291 appear to be present to a similar extent in the three viruses. However, the homology in the region between amino acid residues 155-190 is much less between IBV than MHV, indicating that TGEV more closely resembles MHV. Dot matrix analysis shows that this region between TGEV strains FS772/70 and Purdue has multiple repetitions (Fig. 3), probably resulting from a

series of alignments of serine-arginine repeats, which could be a site for interaction with genomic RNA. There is no detectable homology between the nucleotide sequences of the nucleoprotein genes of the three viruses. This suggests that the three regions of protein homology may be involved in either RNA interaction or the formation of the correct conformation for interaction. There is no antigenic homology between the nucleoproteins of TGEV, IBV and MHV (Pedersen *et al.*, 1978), suggesting that the antigenic sites are present in the non-homologous regions. Motz *et al.* (1986) have shown that major antigenic sites on the Epstein-Barr virus 138 kD early protein occurred at hydrophilic β -turns. Determination of the secondary structure of TGEV nucleoprotein using the PREDICT program and the hydrophilicity of the primary structure showed that the areas of non-homology with IBV and MHV contained potential hydrophilic β -turns, indicating that the areas responsible for antigenicity may fall within these regions. None of the amino acid substitutions between TGEV FS772/70 and Purdue strains fall within the two main homology regions, except for a conservative serine/threonine substitution at amino acid position 262, which indicates that the differences between the nucleoproteins of the two strains will probably not be responsible for the difference in the pathogenicity of the two strains, though there may be antigenic variations.

The 276-bp open reading frames of the two TGEV strains differed in amino acid sequence by 10.3%. Most of the changes observed are in the first base of the triplet, although they also resulted in conservative amino acid substitutions, with the exception of the alanine (FS772/70) to arginine (Purdue) change at position 39 that results

from a double substitution at positions 1 and 2 of the triplet. Substitution of asparagine for aspartate at position 30 of the gene results in the introduction of a potential *N*-glycosylation site in the FS772/70 strain as compared to the Purdue strain. To date, no product has been identified in either infected cells or virions that corresponds to this polypeptide. However, Garwes (unpublished results) has detected a polypeptide of M_r 12000–14000, produced in infected cells, which does not appear to be incorporated into virions. The polypeptide is probably similar to the polypeptide of M_r 17000 identified by Wesley and Woods (1986) and later sized as M_r 14000 (Wesley and Woods, 1987) produced by the Purdue strain of TGEV. The presence of a potential glycosylation site at amino acid 30 of FS772/70, which may add up to 2000 to the molecular weight of the polypeptide, and the hydrophobicity of the molecule may cause the difference in the molecular weight between the theoretical and observed values. Three oligopeptides have been synthesized for the production of antibodies in order to probe infected cells and virions for the presence of the polypeptide of M_r 9068.

The TGEV nucleoprotein gene was constructed from the two cDNA clones and sandwiched between *Bam*H I linkers to produce a gene cassette for transfer to suitable expression vectors. The gene cassette, inserted in the correct orientation in the yeast expression plasmids, produced a polypeptide under the control of the yeast promoters, identical in size to the nucleoprotein found in TGEV virions and which reacted with a TGEV nucleoprotein monoclonal antibody. This observation demonstrated that the nucleoprotein gene construction was correct. To our knowledge, this is the first coronavirus protein expressed in yeast cells. The TGEV nucleoprotein gene or its product did not appear to be toxic to the yeast cells. The nucleoprotein gene inserted in the wrong orientation in the yeast expression vectors did not produce the TGEV nucleoprotein, indicating that the viral cDNA did not contain sequences capable of acting as a promoter in the yeast cells.

The PGK (phosphoglycerate kinase) promoter has been shown to be very efficient (Mellor *et al.*, 1985), which makes it useful for the expression of foreign proteins. Expression of the TGEV nucleoprotein gene under the control of the PGK promoter was constitutive, as observed for other foreign genes (Mellor *et al.*, 1983) and to overcome any potential toxicity of the TGEV nucleoprotein the gene cassette was also subcloned into a yeast expression vector under the control of the *GAL1* inducible promoter. This allowed selection of yeast cells containing the recombinant plasmids with the nucleoprotein gene but without expression from the *GAL1* promoter. The TGEV nucleoprotein was expressed in the presence of galactose but not glucose and the product was the same size and reacted with the same monoclonal antibody as the product obtained under the control of the PGK promoter.

The level of expression observed under the control of the *GAL1* promoter was higher than that observed under the control of the PGK promoter, although the latter is known to be one of the most efficient yeast promoters when acting with its natural gene (Mellor *et al.*, 1985). Comparison of the yeast (Sharp *et al.*, 1986) and the TGEV nucleoprotein codon usage showed that there was little difference in codon usage (data not shown), which should not cause any reduction in efficiency of expression. The presence of yeast transcription termination sequences has been found to enhance the efficiency of expression. The expression plasmid containing the PGK promoter contains a termination site downstream of the cloning site used, whereas the plasmid containing the *GAL1* promoter contains no yeast termination sequence. No apparent increase in expression was observed between the products from the *GAL1* promoter in the presence or absence of the PGK terminator.

In addition to the TGEV nucleoprotein gene, the *Bam*H I cassette also encoded the second potential open reading frame. While several genes can be translated from polycistronic mRNA in *E. coli*, no polycistronic mRNA has been found in yeast. Expression occurs only from the first open reading frame downstream of a promoter when an artificial or foreign gene is introduced into yeast cells as observed with other eukaryotic cells. Thus it is unlikely that any expression will occur from the second open reading frame. The antibodies raised against one of the three synthetic oligopeptides have been used in an ELISA test on yeast extracts for the presence of the potential product of this open reading frame but no product has been detected so far (data not shown).

Experimental procedures

Strains, plasmids and media

E. coli strains DH1, C600 and JM83 were used for routine plasmid construction. *E. coli* transformants were selected on LB plates containing ampicillin ($100 \mu\text{g ml}^{-1}$). *Saccharomyces cerevisiae* strain BWG1-7A (*a*, *leu2-3*, *leu2-112*, *his4-519*, *ade1-100*, *ura3-52*, *GAL⁺*; Guarente *et al.*, 1982) was used as a host strain for the expression of recombinant plasmids. The constitutive PGK expression vector, pMA91, has been described by Mellor *et al.* (1983). The inducible *GAL1-GAL10* expression vector, pB620, was derived from pCGE329 by the addition of a *Bam*H I site to the *GAL1* side of the 0.74-kbp *GAL1-GAL10* divergent promoter (Boeke *et al.*, 1985). Yeast cells were grown in synthetic media containing 2% glucose or galactose and 0.67% yeast nitrogen base (Difco) supplemented with the required amino acids.

Preparation of viral RNA

TGEV strain FS772/70 was grown in LLC-PK1 cells in the presence of [^3H]-uracil and actinomycin D (Garwes *et al.*, 1984).

Messenger RNA was isolated from the cells and purified from other RNA species on poly(U) Sepharose as described previously (Britton *et al.*, 1987). The TGEV mRNA species were fractionated on 15–30% (w/v) linear sucrose gradients in 50 mM Tris-HCl (pH 7.4), 100 mM LiCl, 1 mM EDTA, and 0.1% SDS, for 5 h at 110000×g_{av} and 10°C using an MSE 65 preparative ultracentrifuge.

Digestion and analysis of plasmid DNA

Standard recombinant DNA methods were used (Maniatis *et al.*, 1982) with enzymes purchased from New England Biolabs (CP Laboratories, Bishop's Stortford, England) unless otherwise stated in the text. DNA fragments were isolated from agarose gels by electroelution. Ligation reactions were carried out as described by Britton *et al.* (1984). *E. coli* cells were routinely transformed using the RbCl method developed by V. Simanis (Hanahan, 1985). Vector DNA was routinely treated with alkaline phosphatase prior to ligation. Yeast cells were transformed using the spheroplast method described by Rothstein (1985).

Preparation of primer-vector

Preparation of the oligo(dT) tailed vector primer. Purified pUC9 was digested using *Pst*I, and homopolymer tails of about 50 deoxythymidylate (dT) residues were added to the DNA in a solution containing 0.14 M potassium cacodylate, 30 mM Tris, 0.1 mM dithiothreitol (DTT), 1 mM CoCl₂, 4.4 mM dTTP, 25 µCi of [methyl,1',2'-³H]-dTTP (TRK.576, 97 Ci mmol⁻¹, Amersham International) and 250 units of terminal transferase (Pharmacia Ltd., Milton Keynes, England) at pH 7.0 and 37°C for 2 h (Okayama and Berg, 1982). The extent of tailing was analysed by digestion of the vector-primer with *Hae*II. An increase in the size of the two smallest fragments, compared to *Pst*I- and *Hae*II-digested pUC9, indicated the addition of oligo(dT) tails. One of the dT homopolymer tails was removed by digestion with *Bam*HI and *Eco*RI and chromatography on CL-6B Sepharose (Pharmacia Ltd.).

Cloning of the nucleoprotein mRNA species

Synthesis of first-strand cDNA. A sample of the oligo(dT)-tailed vector-primer was added to 2 µg of RNA solution, containing the 1.7 kb TGEV mRNA species (prepared as described above), previously denatured at 60°C for 5 min. Synthesis of first-strand cDNA was carried out in a solution containing 50 mM Tris-HCl, pH 8.3, 1 mM dNTPs, 8 mM MgCl₂, 50 mM NaCl, 10 mM DTT, 70 units RNasin, 80 mg ml⁻¹ actinomycin D, 42 µCi [α -³⁵S]-dATP (SJ.304, 650 Ci mmol⁻¹, Amersham International), 72 units AMV reverse transcriptase (AH20183, Anglian Biotechnology Ltd., Colchester, England) at 42°C for 2 h. Untreated deoxynucleotides were removed by chromatography through Sephadex G-50.

Synthesis of second-strand cDNA synthesis. The second strand of cDNA was produced by a modification of the RNA/DNA replacement method of Okayama and Berg (1982). The vector-primer containing the mRNA/cDNA hybrid was redissolved in 50 µl of 20 mM Tris-HCl, pH 7.5, 4 mM MgCl₂, 10 mM (NH₄)₂SO₄, 125 mM KCl, 0.15 mM NAD⁺, 25 µg nuclease-free BSA, 25 µM dNTPs, 20 µCi [³²P]-dCTP (1996-231, 3200 Ci mmol⁻¹, New

England Nuclear) 15 units DNA polymerase I (Boehringer), 2 units *E. coli* DNA ligase (Pharmacia Ltd.), 2 units RNase H (BRL, Gibco Ltd., Paisley) and incubated for 1 h at 15°C and 30 min at 22°C.

Cloning using a specific restriction fragment for priming cDNA synthesis. The 90 bp *Hind*III-XbaI fragment (Fig. 1) from plasmid pTS15-1 was purified. A sample of the fragment was boiled for 5 min and added to poly(A) mRNA, isolated from TGEV-infected LLC-PK1 cells, at 60°C. The mixture was incubated at 60°C for 5 min and cooled to 42°C. First- and second-strand cDNA syntheses were carried out as described above.

Transformation of cDNA clones into *E. coli*

cDNA was end-repaired using T4 DNA polymerase as described by Maniatis *et al.* (1982). The cDNA-vector-primer was self-ligated using T4 DNA ligase at 15°C for 24 h followed by treatment with RNA ligase (Pharmacia Ltd.) for a further 15 h. The cDNA derived from the restriction fragment primer was ligated to phosphorylated *Eco*RI linkers (Pharmacia Ltd.). cDNA of 1 kbp was purified and ligated into pUC9 previously cut with *Eco*RI. The ligation mixtures were routinely diluted five-fold and used to transform *E. coli* cells. cDNA derived using the vector-primer method was transformed into strain JM83 and cDNA derived from the restriction fragment primer method was transformed into strain C600 using the method of Hanahan (1983).

Detection of TGEV cDNA-containing transformants

All ampicillin-resistant colonies were replica plated onto Biodyne membranes (P/N BNNG225 1.2 µm, PAL, Portsmouth, England) and the plasmid copy number was amplified by transferring the membranes onto LB plates containing chloramphenicol (25 µg ml⁻¹; Hanahan and Meselson, 1980). The cells were lysed according to the method of Grunstein and Hogness (1975), but including the 10% SDS step described by Maniatis *et al.* (1982). The membranes were pre-hybridized in buffer containing formamide and poly(A) (Boehringer) as described by Maniatis *et al.* (1982). ³²P-labelled TGEV poly(A)-containing RNA (6×10⁶ c.p.m.) was used to probe the cDNA for 42 h at 42°C. The membranes were then washed four times at room temperature in X2 SSC containing 0.1% SDS (X1 SSC = 0.15 M NaCl, 0.015 M trisodium citrate pH 7.0), once in X1 SSC containing 0.1% SDS at 68°C for 1.5 h and autoradiographed. Plasmid DNA was prepared and purified from positive clones as described by Britton *et al.* (1984).

Northern blotting of TGEV subgenomic mRNA species

TGEV poly(A)-containing RNA was glyoxylated as described by Maniatis *et al.* (1982) and separated on a 1% agarose gel. The RNA was immediately transferred onto Biodyne membranes in X20 SSC for 18 h and baked at 80°C for 2 h.

Nick translation of recombinant plasmid DNA and hybridization to TGEV RNA

Recombinant plasmid DNA was nick-translated using DNA polymerase I and DNase I as described by Maniatis *et al.* (1982) except that [α -³⁵S]-dATP was used for labelling the DNA and the reaction was carried out for 2 h at 16°C. Unreacted nucleotides

were removed by chromatography through Sephadex G-50. Labelled recombinant plasmid DNA was used to probe the TGEV mRNA Northern blots.

Subcloning for M13 sequencing

Random subclones of the cDNA inserts from pTS15-1 and pF4F-36 were generated by purifying HindIII, PvuII or the HaeIII-PvuII fragments from pTS15-1 and the EcoRI fragment from pF4F-36 and cloning sonicated (Deininger, 1983) fragments into *Sma*I-cut, phosphatase-treated M13mp10 (Amersham International). In addition, various restriction endonuclease fragments, deduced from the restriction maps (Fig. 1), were cloned into M13mp-vectors to cover areas of sequence partially covered by random clones.

DNA sequencing and sequence analysis

M13/di-deoxynucleotide sequencing (Sanger *et al.*, 1977; Bankier and Barrell, 1983) was carried out using [α -³⁵S]-dATP (Amersham International) and sequencing reactions were analysed on buffer gradient gels (Biggin *et al.*, 1983). A sonic digitizer (Graf/Bar, Science Accessories Corporation) was used to read the data, which were analysed on a VAX 11/750, using the programs of Staden (1982, 1984) and UWGCG (Devereux *et al.*, 1984). Protein secondary structure was determined from its primary structure using PREDICT, a program that compares eight different prediction algorithms (Eliopoulos *et al.*, 1982).

Expression of TGEV nucleoprotein in *S. cerevisiae*

Yeast cells were isolated at mid-exponential phase ($A_{600} = 1-2$), pelleted and resuspended in 1 ml of phosphate-buffered saline (PBS, 10 mM potassium phosphate, 150 mM NaCl, pH 7.2) containing 1 mM PMSF (phenyl methyl sulphonyl fluoride) and 5 mM benzamidine. Cell-free extracts were obtained by disruption of the cells with glass beads (diam. 0.45 mm) for 3 min at 4°C and clarified by centrifugation. Samples containing 100 µg of protein were fractionated on 12% SDS-polyacrylamide gels as described by Britton *et al.* (1982). Following electrophoresis, the separated proteins were transferred onto nitrocellulose membranes (Schleicher & Schuell, BA85, pore size 0.45 µm) using an electroblotting apparatus (Biorad Laboratories Ltd.) at 100 V for 2 h in electrophoresis running buffer with 0.025% SDS and 20% methanol.

Detection of TGEV nucleoprotein

Nitrocellulose membranes containing the fractionated yeast proteins were washed several times with PBS containing 1% gelatine and 0.1% Tween 20. The presence of TGEV nucleoprotein was determined using a monoclonal antibody (DA3: previously raised against TGEV nucleoprotein in mice; Garwes *et al.*, 1987), rabbit anti-mouse antibodies, and 5–10 µCi ¹²⁵I-protein A (Amersham International plc, 32 mCi mg⁻¹ Protein A, code No. IM112) per blot.

Acknowledgements

This research was supported by the Biomolecular Engineering Programme of the Commission of the European Communities, Contract No. GB1-2-089-UK. R. S. Cármenes (from Departamento de Bioquímica, Universidad de Oviedo, Spain) was supported by an EEC training grant. Dr F. Parra was supported by a NATO Fellowship. We would like to thank Dr R. Serrano (EMBL, Heidelberg, RFA) for supplying plasmid pB620 and yeast strain BWG1-7A, Dr G. Yarranton (CellTech Ltd.) for plasmid pMA91, and Miss Fiona Stewart of this Institute for the monoclonal antibody DA3.

mento de Bioquímica, Universidad de Oviedo, Spain) was supported by an EEC training grant. Dr F. Parra was supported by a NATO Fellowship. We would like to thank Dr R. Serrano (EMBL, Heidelberg, RFA) for supplying plasmid pB620 and yeast strain BWG1-7A, Dr G. Yarranton (CellTech Ltd.) for plasmid pMA91, and Miss Fiona Stewart of this Institute for the monoclonal antibody DA3.

References

- Bankier, A., and Barrell, B.G. (1983) Shotgun DNA sequencing. In *Techniques in the Life Sciences (Biochemistry)*, Vol. B5: *Techniques in Nucleic Acid Biochemistry*, B508. R.A. Flavell (ed.) Ireland: Elsevier, pp. 1–34.
- Biggin, M.D., Gibson, T.J., and Hong, G.F. (1983) Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proc Nat Acad Sci USA* **80**: 3963–3965.
- Boeke, J.D., Garfinkel, D.J., Styles, C.A., and Fink, G.R. (1985) Ty elements transpose through an RNA intermediate. *Cell* **40**: 491–500.
- Boursnell, M.E.G., Binns, M.M., Foulds, I.J., and Brown, T.D.K. (1985) Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. *J Gen Virol* **66**: 573–580.
- Britton, P., Garwes, D.J., Millson, G.C., Page, K., Bountif, L., Stewart, F., and Walmsley, J. (1986) Towards a genetically-engineered vaccine against porcine transmissible gastroenteritis virus. In *Biomolecular Engineering in the European Community. Final Report*. Magnien, E. (ed.) The Netherlands: Martinus Nijhoff, pp. 301–313.
- Britton, P., Garwes, D.J., Page, K., and Walmsley, J. (1987) Expression of porcine transmissible gastroenteritis virus genes in *E. coli* as β-galactosidase chimaeric proteins. In *Coronaviruses*. Lai, M., and Stohlmeyer, S. (eds). New York: Plenum Press, pp. 55–64.
- Britton, P., Lee, L.G., Murfitt, D., Boronat, A., Jones-Mortimer, M.C., and Kornberg, H.L. (1984) Location and direction of transcription of the *ptsH* and *ptsI* genes on the *Escherichia coli* K12 genome. *J Gen Microbiol* **130**: 861–868.
- 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 λ*ie*(*crr*) and λ*gsr*(*tgs*). *EMBO J* **1**: 907–911.
- Deininger, P.L. (1983) Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. *Anal Biochem* **129**: 216–223.
- Dennis, D.E., and Brian, D.A. (1981) Coronavirus cell-associated RNA-dependent RNA polymerase. *Adv Exp Med Biol* **142**: 155–170.
- Devereux, J., Haeblerli, P., and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucl Acids Res* **12**: 387–395.
- Eliopoulos, E.E., Geddes, A.J., Brett, M., Pappin, D.J.C., and Findlay, J.B.C. (1982) A structural model for the chromophore-binding domain of ovine rhodopsin. *Int J Biol Macromol* **4**: 263–268.
- Garwes, D.J. (1982) Coronaviruses in animals. In *Virus Infections of the Gastrointestinal Tract*. Tyrell, D.A.J., and Kapikian, A.Z. (eds). New York: Marcel Dekker Inc., pp. 315–359.
- Garwes, D.J., and Pocock, D.H. (1975) The polypeptide structure of transmissible gastroenteritis virus. *J Gen Virol* **29**: 25–34.
- Garwes, D.J., Bountif, L., Millson, G.C., and Elleman, C.J. (1984) Defective replication of porcine transmissible gastroenteritis virus in a continuous cell line. *Adv Exp Med Biol* **178**: 79–93.

- Garwes, D.J., Pocock, D.H., and Wijaszka, T.M. (1975) Identification of heat-dissociable RNA complexes in two porcine coronaviruses. *Nature* **257**: 508–510.
- 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*. Lai, M., and Stohlman, S. (eds). New York: Plenum Press, pp. 509–516.
- Grunstein, M., and Hogness, D. (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc Natl Acad Sci USA* **72**: 3961–3965.
- Guarente, L., Yocom, R.R., and Gifford, P. (1982) A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. *Proc Natl Acad Sci USA* **79**: 7410–7414.
- Hanahan, D., and Meselson, M. (1980) Plasmid screening at high colony density. *Gene* **10**: 63–67.
- Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**: 557–580.
- Hanahan, D. (1985) Techniques for transformation of *Escherichia coli*. In *DNA Cloning, A Practical Approach*, Vol. II, Glover, D.M. (ed.). Oxford: IRL Press, pp. 109–135.
- Jacobs, L., van der Zeijst, B.A.M., and Horzinek, M.C. (1986) Characterization and translation of transmissible gastroenteritis virus mRNAs. *J Virol* **57**: 1010–1015.
- Kapke, P.A., and Brian, D.A. (1986) Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. *Virology* **151**: 41–49.
- Kozak, M. (1983) Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbiol Rev* **47**: 1–45.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.
- Mellor, J., Dobson, M.J., Roberts, N.A., Kingsman, A.J., and Kingsman, S.M. (1985) Factors affecting heterologous gene expression in *Saccharomyces cerevisiae*. *Gene* **33**: 215–226.
- Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J., and Kingsman, S.M. (1983) Efficient synthesis of enzymatically active calf chymosin in *Saccharomyces cerevisiae*. *Gene* **24**: 1–14.
- Miyanohara, A., Toh-e, A., Nozaki, C., Hamadan, F., Ohtomo, N., and Matsubara, K. (1983) Expression of hepatitis B surface antigen gene in yeast. *Proc Natl Acad Sci USA* **80**: 1–5.
- Motz, M., Fan, J., Seibl, R., Jilg, W., and Wolf, H. (1986) Expression of the Epstein-Barr virus 138 kDa early protein in *Escherichia coli* for use as antigen in diagnostic tests. *Gene* **42**: 303–312.
- Okayama, H., and Berg, P. (1982) High-efficiency cloning of full-length cDNA. *Mol Cell Biol* **2**: 161–170.
- Pedersen, N.C., Ward, J., and Mengeling, W.L. (1978) Antigenic relationship of the feline infectious peritonitis virus to coronaviruses of other species. *Arch Virol* **58**: 45–53.
- Rothstein, R. (1985) Cloning in yeast. In *DNA Cloning, A Practical Approach*, Vol. II, Glover, D.M. (ed.). Oxford: IRL Press, pp. 45–66.
- Sanger, F., Nicklins, S., and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463–5467.
- Sharp, P.M., Tuohy, T.M.F., and Mosurski, K.R. (1986) Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucl Acids Res* **14**: 5125–5143.
- Shieh, C-K., Soe, L.H., Makino, S., Chang, M-F., Stohlman, S.A., and Lai, M.M.C. (1987) The 5'-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. *Virology* **156**: 321–330.
- Siddell, S., Wege, H., and Ter Meulen, V. (1983) The biology of coronaviruses. *J Gen Virol* **64**: 761–776.
- Skinner, M.A., and Siddell, S.G. (1983) Coronavirus JHM: nucleotide sequence of the mRNA that encodes nucleocapsid protein. *Nucl Acids Res* **11**: 5045–5054.
- Staden, R., and McLachlan, A.D. (1982) Codon preference and its use in identifying protein coding regions in long DNA sequences. *Nucl Acids Res* **10**: 141–156.
- Staden, R. (1982) Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucl Acids Res* **10**: 4731–4751.
- Staden, R. (1984) Graphic methods to determine the function of nucleic acid sequences. *Nucl Acids Res* **12**: 521–538.
- Stern, D.F., and Kennedy, S.I.T. (1980) Coronavirus multiplication strategy. I. Identification and characterisation of virus-specified RNA. *J Virol* **34**: 665–674.
- Stern, D.F., and Sefton, B.M. (1984) Coronavirus multiplication: the locations of genes for the virion proteins on the avian infectious bronchitis virus genome. *J Virol* **50**: 22–29.
- Stohlman, S.A., and Lai, M.M.C. (1979) Phosphoproteins of murine hepatitis virus. *J Virol* **32**: 672–675.
- Sturman, L.S., and Holmes, K.V. (1983) The molecular biology of coronaviruses. *Adv Virus Res* **28**: 35–112.
- Valenzuela, P., Medina, A., Rutter, W.J., Ammerer, G., and Hall, B.D. (1982) Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. *Nature* **298**: 347–350.
- Wesley, R.D., and Woods, R.D. (1986) Identification of a 17,000 molecular weight antigenic polypeptide in transmissible gastroenteritis virus-infected cells. *J Gen Virol* **67**: 1419–1425.
- Wesley, R.D., Woods, R.D., and Kapke, P. Antibody response in swine to individual transmissible gastroenteritis virus (TGEV) proteins. In *Coronaviruses*. Lai, M., and Stohlman, S. (eds). New York: Plenum Press, pp. 475–482.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.