Biochimie, 69 (1987) 591 – 600

Enteric coronavirus TGEV : partial sequence of the genomic RNA, its organization and expression

Denis RASSCHAERT, Jacqueline GELFI and Hubert LAUDE*

Institut National de la Recherche Agronomique, Station de Recherches de Virologie et d'Immunologie, F-78850 Thiverval-Grignon, France

(Received 27-3-1987, accepted after revision 24-4-1987)

Summary – The sequence of the 3'-most 8300 nucleotides of the genome RNA of the Purdue-115 strain of the transmissible gastroenteritis virus TGEV, a porcine coronavirus, was determined from cDNA clones. The available sequence corresponds to the part of the genome (total length > 20 kb) expressed through subgenomic mRNAs. The 5 subgenomic and the genomic RNA species detected in TGEV-infected cells form a 3'-coterminal 'nested' structure, a unique feature of *Coronaviridae*.

The transcription initiation site of the TGEV subgenomic RNAs appears to involve the hexameric sequence 5'CTAAAC, which is present upstream from each coding region. In addition to the previously identified genes encoding the three structural proteins, E2, E1 and N, two regions, X1 and X2, corresponding to the non-overlapping portion of mRNAs 4 and 3, may code for so far unidentified non-structural polypeptides. The predicted X1 polypeptide (9.2 kDa) is highly hydrophobic. The sequence of the X2 region allows the translation of two non-overlapping products, *i.e.*, X2a (7.7 kDa) and X2b (18.8 kDa). No RNA species liable to express the extreme 3' open reading frame X3 was found.

coronavirus / transmissible gastroenteritis / TGEV / messenger RNAs / genome structure / gene sequence / non-structural polypeptides-1987)

Résumé – Virus de la gastro-entérite transmissible (TGEV) : séquence partielle, organisation et expression de l'ARN génomique. *La séquence des 8300 nucléotides en région 3' de l'ARN génomique du coronavirus porcin TGEV (souche Purdue-115) a été établie à partir de clones d'ADNc. Par rapport au génome entier (> 20 kb), cela recouvre l'ensemble des séquences exprimées par l'intermédiaire d'ARNs messagers de taille subgénomique. Les 5 espèces d'ARN subgénomiques et l'ARN génomique détectés dans les cellules infectées forment des séquences emboîtées co-terminales en 3', ce qui est caractéristique du mode de réplication des* Coronaviridae. *Une séquence hexamérique, 5'CTAAAC, présente juste en amont de chaque région codante, constituerait le site d'initiation de la transcription des ARN subgénomiques du TGEV. Outre les gènes des 3 protéines structurales E2, E1 et N précédemment identifiés, deux régions X1 et X2, correspondant à la région « unique » des ARNm 4 et 3, pourraient coder pour des polypeptides non-structuraux, actuellement non-identifiés. L'un des polypeptides prédits, X1 (9.2 kDa) est extrêmement hydrophobe.*

* *Author to whom correspondence should be sent.*
Abreviations: **bp**: base pair; **IBV**: infectious bronchitis virus; **kb**: kilobase; **MHV**: murine hepatitis virus; **ORF**: open reading frame; **SSC**: saline sodium citrate; **TGEV**: transmissible gastroenteritis virus.

Deux produits complètement distincts, X2a (7.7 kDa)) et X2b (18.8 kDa), pourraient être traduits à partir du mRNA 3. Aucun ARN susceptible d'exprimer la phase codante située à l'extrémité 3' (X3) n'a été mis en évidence.

coronavirus / gastro-entérite transmissible / TGEV / ARN messagers / structure du génome / séquence des gènes / polypeptides non-structuraux

Introduction

Transmissible gastroenteritis virus (TGEV), an important pathogen of swine neonates, belongs to the *Coronaviridae,* a family of enveloped viruses with a large, positive-stranded RNA as their genome [1]. Earlier studies showed that the TGEV genome consists of a unique RNA molecule, approximately 20 kb in length, which is polyadenylated and infectious similar to that of other coronaviruses [2]. Although the total number of genes encoded has not yet been determined, the TGEV genome codes for at least four polypeptides on the basis of existing protein and nucleotide data. The virions are constructed of three polypeptides, the nucleocapsid (N), the membrane (E1) and the spike or peplomer (E2) polypeptides, the complete sequence of each of which has been recently established [3 – 5]. These three genes account for approximately 6.3 kb of coding information. In addition, at least one non-structural polypeptide is synthesized during virus replication, an RNA dependent-RNA polymerase, which requires Mg^{2+} cations and is probably membrane-bound [11].

Expression of the coronavirus-encoded information proceeds through the synthesis of several distinct mRNA species of subgenomic size. The transcription strategy has been studied in detail on the murine hepatitis virus (MHV) and infectious bronchitis virus (IBV) models. The intracellular RNA species (7 and 6 in number, respectively, including the genome RNA) have been shown to form a nested set with common 3' ends. The translated sequences correspond approximately to the 5' portion which is absent in the next smaller RNA. The subgenomic RNAs contain leader and body sequences joined through a discontinuous transcription. This process relies upon the presence of a short homologous sequence in each intergenic region, most likely acting as a recognition signal for the polymerase – leader complex [6 – 10]. Less information is available concerning TGEV transcription. The number of subgenomic RNA species synthesized in infected cells varies from 4 to 9 in previous literature [11 – 14].

The purpose of this paper is, first, to propose a model of TGEV genome organization and expression based on both sequence analysis of cloned virion RNA and characterization of virus specific intracellular RNAs, and second, to describe the characteristics of additional polypeptides possibly encoded by the genome.

Materials and methods

Virus and cells
The Purdue-115 strain of TGEV was propagated in the PD5-cell line and virions were purified as reported [15].

RNA extraction
Purified virions were treated with proteinase K (200 units/ml; Merck) and 2% SDS for 30 min at 37°C. RNA was extracted once with phenol and twice by phenol – chloroform (1/1) with gentle agitation. After ethanol precipitation with sodium acetate (0.3 M), the RNA pellet was resuspended in sterile bidistilled water and stored at – 80°C. The extraction yield was 40 – 50 μg of RNA for 1 mg of purified virion.

cDNA synthesis
The purified RNA was denatured by methylmercuric hydroxide for 10 min at room temperature [16]. The final concentration of CH_3HgOH in the reverse transcription reaction mix was optimized to 8 mM. The reaction was carried out at 37°C for 2 h in 50 μl containing: 15 μg of extemporaneously denatured RNA, RNasin (100 units; Promega Biotec, Madison), KCl (40 mM), $MgCl_2$ (6 mM), Tris-HCl (40 mM, pH 8.3, at 37°C), 2-mercaptoethanol (56 mM; *i.e.* 7-fold molar excess to CH_3HgOH), dATP, dCTP, dGTP, dTTP (0.5 mM each), [^{3}H]dTTP (100 μCi, 30 Ci/mmol; Amersham); primers pdT 12 – 18 (Pharmacia) or pE2 (sequence specific, 30-mer [5]) (5 μg) and 'super' reverse transcriptase (88 units; Stehelin, Basel). The reaction was stopped with EDTA (20 mM) followed by phenol – chloroform extraction. The RNA – cDNA hybrids were precipitated with ethanol and 2 M ammonium acetate [17]. About 4 μg of cDNA were obtained from 15 μg of RNA.

RNase T2 treatment
The RNA – cDNA hybrid material was subjected to RNase T2 treatment in a volume of 50 μl containing NaCl

(250 mM), sodium acetate (10 mM, pH 4.5) and RNase T2 (17 units; BRL) (S. Van der Werf, Institut Pasteur, personal communication). After a 15 min incubation at 37°C the material was extracted with phenol – chloroform, desalted in a centrifuged Sephadex G – 50 column and ethanol precipitated using 2 M ammonium acetate.

Tailing and cloning of cDNA

Homopolymeric dC tails were added to RNA – cDNA hybrids (500 ng) by incubation (3 min at 37°C) in a 20 μl reaction mixture containing potassium cacodylate (100 mM), Tris-base (25 mM, pH 7.6), $CaCl_2$ (1 mM), DTT (0.2 mM), dCTP (0.2 mM), BSA (0.5 mg/ml; BRL) and terminal deoxynucleotidyl transferase (675 units/ml; Pharmacia P.L.). dC-tailed RNA – cDNA hybrids were annealed to *Pst*I-cut dG-tailed pBR322 (BRL; 1.5 mg/μl, *i.e.*, 2-fold molar excess to RNA – cDNA hybrids) under the following conditions: 20 mM Tris-HCl, pH 7.4; 300 mM NaCl; 1 mM EDTA; at 62°C for 15 min; at 57°C for 2 h then cooled to room temperature. The mixture was used to transform competent *E. coli* RR1 [18] which were plated onto L-agar containing 12 mg/ml of tetracyline. The percentage of ampicillin-sensitive transformants ranged between 60 and 90% in the different experiments.

Screening and mapping

The clones containing an insert exceeding 800 bp were selected [19]. A map of cloned inserts was achieved by means of Northern and Southern blot hybridizations and hexanucleotide restriction enzyme analyses [20]. For Northern blot experiments, total RNA of TGEV-infected PD5 cells was extracted by the guanidium isothiocyanate technique [21] and deposited on a 0.75% denaturing agarose gel containing formaldehyde. RNA transferred onto nitrocellulose was hybridized with nick-translated [^{32}P]dCTP labeled plasmids [20]. Filters were washed in 0.1x SSC + 0.1% SDS at 55°C for 1 h. In Southern blot experiments, identical hybridization and washing conditions were employed.

DNA sequencing

Sonicated plasmid fragments ranging from 500 to 700 bp were subcloned into *Sma*I-cut M13mp18 phage vector [22]. The DNA sequence was determined with the chain termination method [23] using the 17-mer sequencing primer and [^{35}S]dATP (600 Ci/mmol; NEN) as the label. The sequence was determined on polyacrylamide buffer gradient gels [24]. The whole sequence was determined on both strands. Sequencing data were analyzed using the Microgenie sequencing program (March 1985 version, Beckman). The supercoiled plasmid dideoxy-sequencing method [25] was occasionally employed to confirm partial sequence data, using oligonucleotide primers synthesized on a Biosearch 8600 apparatus.

Results

Generation and mapping of cDNA library

RNA extracted from purified TGEV consisting of a large-sized (>20 kb), homogeneous, potentially full-length material, was reverse transcribed after oligodT-priming. Several discrete cDNA species, most likely due to the existence of stable secondary structures in genome RNA, were produced (Fig. 1); a well-defined band of approximately 18 kb, expected to encompass the major structural protein genes, was visible. This material served to generate the pTG2 library. Six recombinant clones (2.15, 2.21, 2.26, 2.27, 2.40, 2.50) were oriented along the genome by means of Northern hybridization with size-fractionated RNAs from TGEV-infected cells (Fig. 2). Clone pTG2.21 (and 2.15, data not shown) contained sequences hybridizing with 6 RNA species, of which the largest one (RNA 1) had the same size as that of virion RNA. Clone pTG2.50 hybridized with all species except RNA 6. Clone

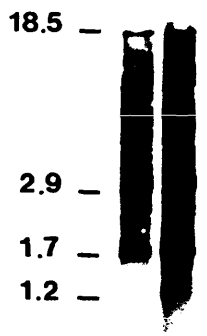


Fig. 1. Electrophoresis of cDNA synthesis products. ^{3}H-labeled cDNA material from two different experiments was analyzed in denaturing 0.75% alkaline agarose gel. The estimated size of the major discrete species is given in kilobases.

pTG2.26 had common sequences exclusively with RNA 1 and 2, whereas clones pTG2.40 (and 2.27, data not shown) possessed sequences only present in RNA 1. This result is consistent with the fact that in coronaviruses, genome RNA and subgenomic RNA species form a nested set with 3′ common sequences. Additional clones were probed against clones 2.26 and 2.15, using Southern blotting. All the selected clones were mapped by restriction enzyme analysis. The overlapping clones were shown to stretch along the 7 kb DNA (Fig. 3). Clones 2.21, 2.15 and 2.26 were sequenced. Subsequently, a second library (pTG6) was produced using a synthetic primer pE2 located 3.8 kb from the 3′ end [5]. Resulting overlapping clones were found to extend the continuum up to 14 500 bases (Fig. 3) of which 8300 bases starting from the 3′ end have been sequenced.

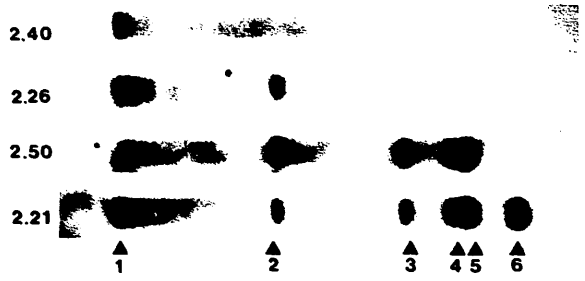


Fig. 2. Northern blot analysis of TGEV intracellular RNAs. Total RNA from TGEV-infected PD5 cells was resolved in formaldehyde 0.75% agarose gel, transferred onto a nitrocellulose filter, then hybridized with 4 different ^{32}P-labeled plasmids (designated at the left). An autoradiograph of each blot is shown. Migration was from left to right. The mRNA species detected are numbered from 1 to 6.

Nucleotide sequence analysis

Seven major open reading frames (ORFs) were identified by stop codon analysis (Fig. 4). As previously reported, the 3 largest ones encode the major structural proteins, E2, E1 and N. In addition, 4 ORFs exceeding 200 bases, designated X2a, X2b, X1 and X3, were detected. The sequence segment extending from the 3′ end of the E2 gene up to the 3′ end of the genome (3920 nucleotides) is displayed in Fig. 5 along with the translation of the main ORFs. During the course of this work, sequences of the E1 and N genes and downstream sequences became available from another group [3, 26]. As seen in Fig. 5, there were only few differences between the two sets of data. The stretch of 111 nucleotides up to the poly(A) is lacking from our data.

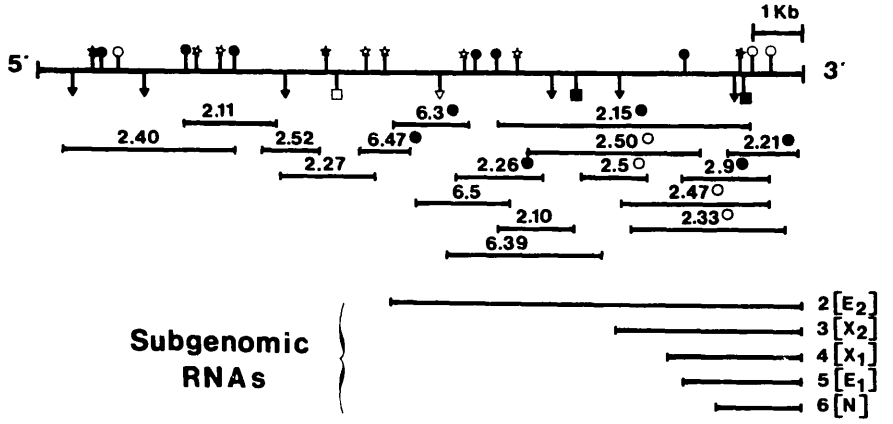


Fig. 3. Restriction endonuclease map of part of TGEV genome (14.5 kb). The length and distribution of cDNA clones selected from the pTG2 and pTG6 libraries are shown. The clones used for sequencing are marked by a solid circle. Open circles indicate clones which have been partially sequenced using plasmid dsDNA as a matrix. Bottom: The 5 subgenomic RNA species identified by Northern hybridization are positioned along the genome map. Restriction enzyme sites: ★: *Bgl*II; □: *Eco*RI; ▼: *Hin*dIII; ☆: *Hpa*I; ○: *Pst*I; ●: *Pvu*II; ■: *Xba*I; ▼: *Xho*I.

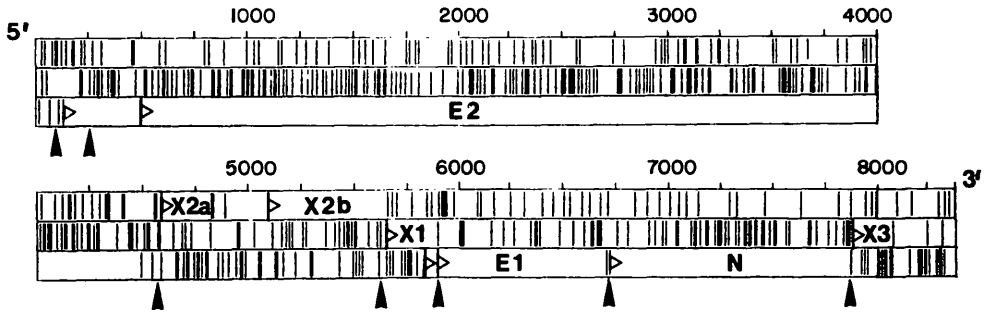


Fig. 4. Stop codon analysis of the virus sense RNA. A computer graphical output of the open reading frames within the first 8300 nucleotides from the 3' end is shown. Stop codons are represented by vertical bars. Bars with an open triangle indicate proximal ATGs in the corresponding frame. Arrowheads beneath the lower frame mark the position of every 5'CTAAAC hexamer found in the sequence.

A remarkable feature of the sequence was the presence of an identical hexamer 5'CTAAAC upstream from the E2, X2a, X1, E1, N and X3 ORFs (Figs. 4 and 5). As suggested for MHV and IBV (see introduction), these homologous sequences are likely to act as initiation sites for the transcription of each mRNA species. Accordingly, it was postulated that the CTAAAC located immediately upstream from the ORFs X2a and X1 ORFs should correspond to the start of the mRNAs 3 and 4, respectively (see Discussion). The non-overlapping region of mRNA 4 appeared to contain a single ORF, X1* (246 bases). The predicted sequence of mRNA 3 might allow translation of two ORFs: X2a, 213 bases long and starting 24 bases downstream from the CTAAAC sequence; and X2b, 495 bases long and starting 570 bases downstream. Three more points were noted regarding X2b: 1) no stop codon occurred up to 267 nucleotides upstream from the potential initiation codon (position 715, Fig. 5); 2) with its 3' end partially overlapping the X1 ORF, X2b is the sole ORF to stretch into the 'unique' sequence of the adjacent smaller RNA; 3) the sequence of the whole X2 region was established on 4 independent clones (see Fig. 3). Surprisingly, 2 of them (pTG2.15 and 2.33) lacked the same 13 base sequence (discontinuous box near position 1000 in Fig. 5); this created an alternative ORF, X2b', only 294 bases long and ending at position 1019 (stop codon overlined).

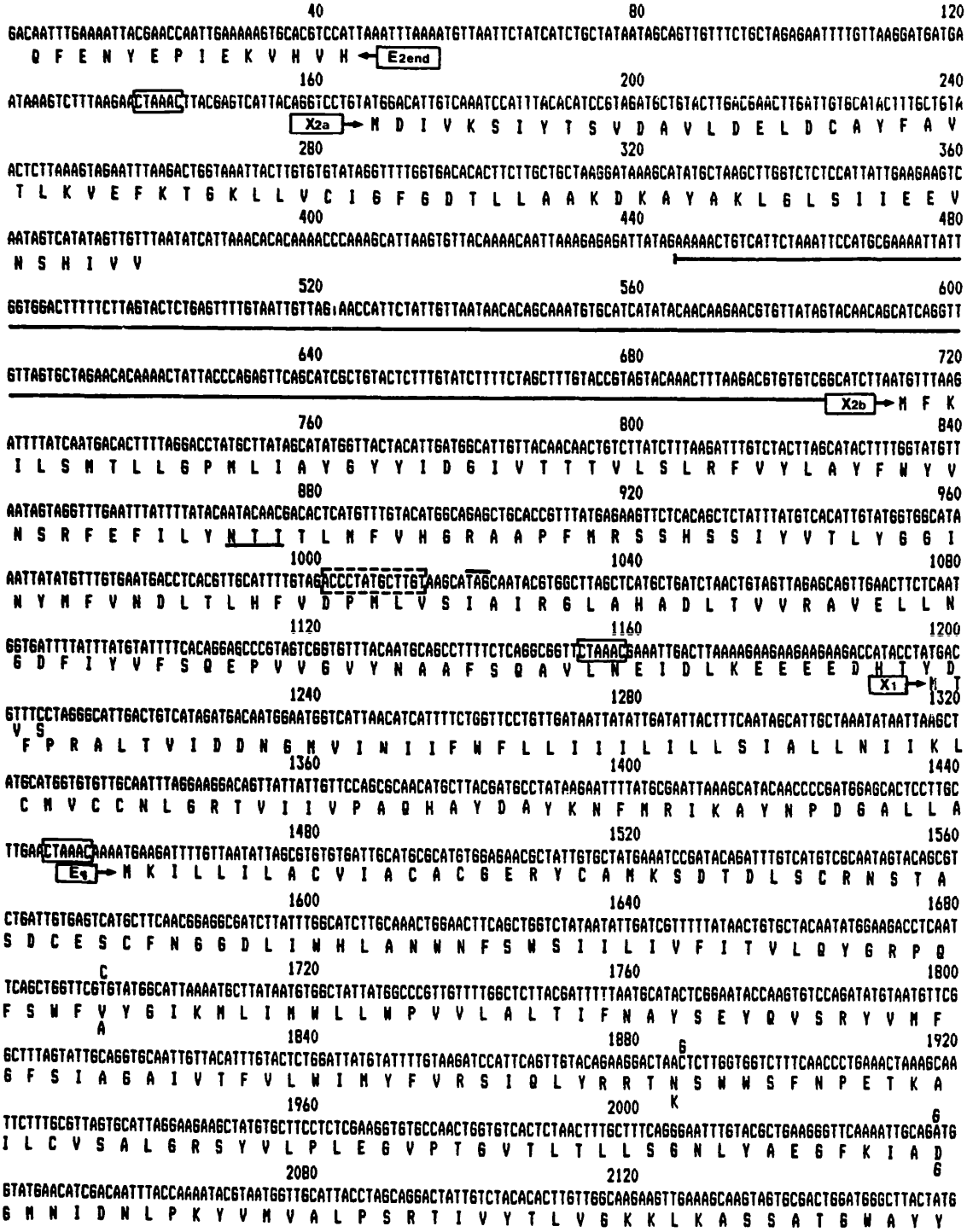
Discussion

Organization and expression of TGEV genome

About 14 500 nucleotides of the 3' end region of TGEV genome were cloned in the pBR322 vector and mapped. All clones used in our study have been derived by direct cloning of a RNA–DNA heteroduplex. According to the size (up to 5 kb) and distribution of the copy fragments, this simple method appeared to be as efficient as that of Gubler & Hoffman (dsDNA synthesis using RNase H) in the case of IBV RNA cloning [27]. Moreover, although we used oligodT, instead of random-priming, clones mapped at more than 14 kb from the 3' end.

The sequence part, 8300 nucleotides at the 3' region, spanned that complete portion of TGEV

*The X1 ORF was observed to contain a 15 bases out-of-frame sequence 5' ATTATATTGATATTA identical to an in-frame sequence found near the 3' end of the E2 gene ([5]; not shown).


E2end
X2a
X2b
X1
E1

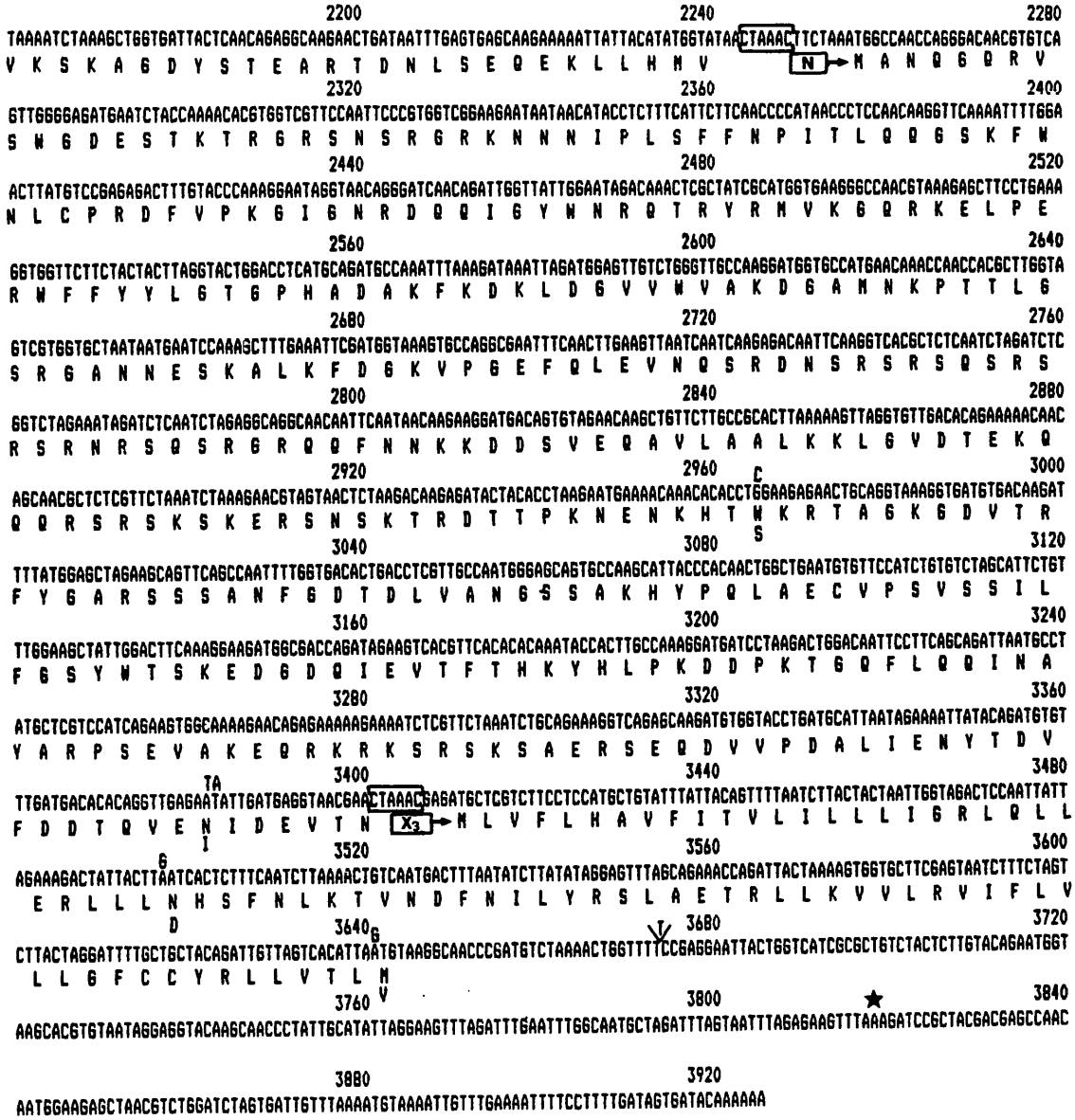


Fig. 5. Sequence of the 3′-most 3920 nucleotides of TGEV genome. The open reading frames are translated in one letter amino acid code. The homologous sequences CTAAAC are boxed. The line upstream from X2b ORF indicates a frame without stop codon. A glycosylation signal present in the X2b product is underlined. Nucleotide and amino acid differences with another published sequence (from position 1400–3820) are indicated. The 111 base long sequence from the star to the poly(A) is taken from [3].

RNA expressed through subgenomic size RNAs, whereas the portion left unsequenced presumably encodes the polymerase. As shown in Fig. 4, the region sequenced comprises the 3 genes encoding the major structural proteins N, E1 and E2, already identified on the basis of their predicted translation products [3–5]. Additionally, three regions, X2a, X2b and X1, might code for non-structural or, less probably, minor structural polypeptides so far unidentified.

As a striking feature, each coding region (except X2b) was preceded by a short consensus sequence 5'$^{AA}_{TT}$CTAAAC, similar to those observed in the genome of MHV (AATC$^{T}_{C}$AAAC, [9]) and IBV (CT$^{T}_{G}$AACAA, [8]). Thus, we believe that these homologous sequences correspond to the site of translation initiation in the TGEV genome. This assumption is strengthened by the finding that the measured size of the non-overlapping region of each intracellular RNA species was in accordance with their respective predicted size (data summarized in Table I). It is worth mentioning that the sequence CTAAAC was never present internally in a TGEV ORF, except in one case, about 150 bases after the start of the E2 gene (Table I; [5]). The CTAAAC sequence located upstream from the X3 ORF, for which no corresponding intracellular RNA species was identified (see below), might also be non-functional for mRNA transcription. If confirmed, this would suggest that additional factors govern the reinitiation of the RNA polymerase–leader complex.

Our results demonstrate that TGEV intracellular RNAs form a 3' co-terminal 'nested' set (Fig. 2), a feature of *Coronaviridae*. In addition, the RNA species pattern is in agreement with that recently published by others [14]. Typically, RNA 5 encoding E1 (2.5 kb) and less abundant RNA 4 (3 kb) appear to be close to each other in size, unlike what was reported by another group [13]. An additional poly(A$^+$)RNA species, 0.7 kb long and rather rare, could have been a candidate for the extreme 3' ORF called X3. However, it was not detected by Northern hybridization using a cDNA probe [14]. A similar result was obtained in our experiments in which total intracellular RNA was analyzed.

The overall view of our data led us to propose the model of the structure of TGEV genome depicted in Fig. 6. Its organization appears to be 'intermediate' between those of MHV and IBV. Like IBV, TGEV possesses 5 subgenomic mRNAs and lacks a subgenomic RNA species larger than the E2 encoding RNA 3, which exists in MHV. On the other hand, the E1 and N genes are adjacent in both MHV and TGEV genomes. The coding regions of TGEV genome are densely packed overall, yet there are almost no overlaps. The intergenic regions consist of 0–15 bases, except the E2–X2a junction, which is 120 bases long (Fig. 5). Every subgenomic RNA species appears to be functionally monocistronic, except RNA 3, which potentially allows the translation of two non-overlapping products, X2a and X2b. It is noteworthy that MHV RNA 5 and IBV RNA D also possess a sequence arrangement which might imply an internal initiation of protein synthesis [28, 29]. This

Table. I. Comparison between the nucleotide position of the homologous regions and the calculated size of the non-overlapping regions of each subgenomic RNA.

TGEV homologous regions			Base distance from the 3' end	Adjacent ORF	Predicted size[a] of the body sequence		RNA species
					Nucleotide data[b]	Experimental data[c]	
5' GTA	CTAAAC	TT 3'	8300	E2	4.5	4.5	2
CTT	CTAAAC	TA	8150	–	4.4	–	Not detected
GAA	CTAAAC	TT	3780	X2a	1	1.1	3
GTT	CTAAAC	GA	2760	X1	0.3	0.4	4
GAA	CTAAAC	AA	2470	E1	0.8	0.7	5
TAA	CTAAAC	TT	1670	N	1.2	1.8	6
GAA	CTAAAC	GA	510	X3	0.5	–	Not detected

[a] In kilobases.
[b] Distance between the two closest homologous sequons.
[c] Difference of size between an RNA species and the next smaller one as measured in a denaturing gel.

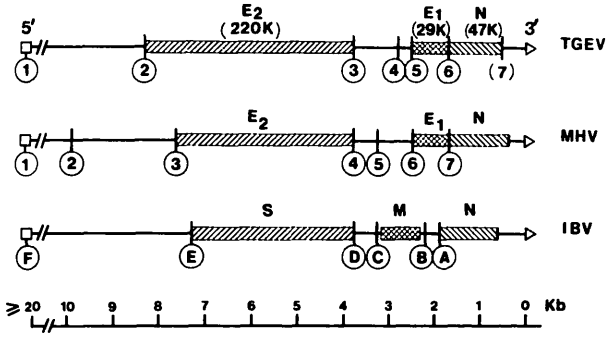


Fig. 6. Compared organization of the genome of three coronaviruses: porcine TGEV, murine MHV and avian IBV. An encircled number or letter placed on the left of a sequence segment indicates the encoding RNA species. The genes coding for the three major structural proteins (peplomer E2, membrane E1 and nucleocapsid N) are represented by hatched boxes. The diagrams of MHV and IBV genomes have been constructed using data from [1, 28, 29].

shared feature might be of biological significance, as for instance a deliberate limitation of the synthesis of the product encoded by the downstream ORF.

Potential primary translation products of mRNAs 4 and 3

The X1 ORF, encoded by the 5′ sequences of mRNA 4, potentially directs the synthesis of an 82 amino acid long polypeptide of 9241 Da, which appears to be extremely hydrophobic (Fig. 7). Its composition is very unusual with 32% leucine + isoleucine residues. The codon usage of X1 does not differ from that of the structural protein genes (data not shown). In particular, codon ATC is unfrequently used for isoleucine (1/14), a bias which would not be expected from a chance ORF. The first available ATG is in an unfavorable context (CxxAUGA) for translation initiation [31].

The mRNA 3 potentially allows the synthesis of two products, X2a and X2b, 71 and 165 amino acids long, respectively. Both ORFs have ATG codon flanking sequences (TxxAUGG, TxxAUGT) which function poorly as initiation signals [31]. Their codon usage suggests that they are not chance ORFs (data not shown). The hydrophilicity profile of X2a (7711 kDa) did not reveal any special feature. The X2b product (18 833 Da) was shown to be hydrophobic overall, with a markedly acidic C-terminus comprising a cluster of 4 glutamic acid residues (position 1180, Fig. 5). As pointed out, the sequence of 2 of the 4 clones spanning this region predicted an alternative product X2b′, 67 amino acids shorter at the C-terminus than X2b (X2b′: 11 413 Da). This finding might reflect a heterogeneity of the virus population, although a cloning artifact cannot be ruled out completely.

It is presently difficult to reconcile the above information with experimental data available for TGEV. *In vitro* translation of mRNA 3 produced a 24 kDa polypeptide which neither comigrated with any intracellular viral protein nor could be immunoprecipitated with anti-virion protein antibodies [14]. A 16–17 kDa non-structural polypeptide, which was unglycosylated and which induced a late antibody response in the host, has been characterized in TGEV-infected cells [32]. A non-structural polypeptide of similar M_r (15 kDa) has been observed in our laboratory, but the latter was shown to incorporate [^{35}S]cysteine (B. Delmas & H. Laude, unpublished results), whereas no cys residue is predicted in X2b. Finally, no smaller polypeptide with an M_r approaching that of X1 or X2a has been identified so far.

Computer investigations revealed no convincing homologies at the DNA or protein level between the TGEV X1 or X2a sequences and the 'non-structural' genes of IBV [29, 33] and MHV [28, 34] (data not shown). However, the TGEV X1 product (Fig. 7) and the highly hydrophobic 7.5 kDa polypeptide predicted by the sequence of IBV mRNA B [33] might have a common (yet unknown)

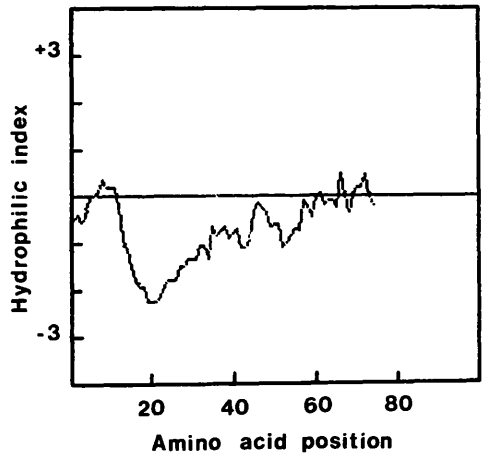


Fig. 7. Hydrophilicity plot of the predicted X1 polypeptide. Running average taken over a heptapeptide using the values of Hopp & Woods [30].

function. In addition, TGEV X2b shows some similarities with IBV 12.4 kDa (mRNA D) and MHV 10.2 kDa (mRNA 5) translation products [29, 28]. They are all produced from a downstream ORF, are hydrophobic overall except for the C-terminus and have an unusually high tyrosine content (7 – 10%). A low sequence homology between these IBV an MHV polypeptides has already been pointed out [29]. In conclusion, the marked resemblance between the structural polypeptides of coronaviruses does not extend to the above-mentioned gene products. Some of them may prove to be key factors in the virus cycle, for instance in transcription – replication switching. One way to achieve their characterization would be to use antisera directed against synthetic peptides derived from the sequence so as to facilitate their identification in infected cell extracts.

Acknowledgments

The authors are grateful to their colleagues J. Cohen, M. Brémont and F. Lefèvre for helpful discussions during this work. We also thank A. Kumar and Kristen Rérat for revising the English manuscript. Parts of these results were presented at the 3rd International Coronavirus Symposium (Asilomar, September 1986).

References

1 Siddell S., Wege H. & Ter Meulen V. (1983) *J. Gen. Virol.* 64, 761–776
2 Brian D.A., Dennis E.D. & Guy J.S. (1980) *J. Virol.* 34, 410–415
3 Kapke P.A., & Brian D.A. (1986) *Virology* 151, 41–49
4 Laude H., Rasschaert D. & Huet J.C. (1987) *J. Gen. Virol.* 68, 1687–1693
5 Rasschaert D. & Laude H. (1987) *J. Gen. Virol.* 68, 1883–1890
6 Spaan W.J.M., Delius H., Skinner M., Armstrong J., Rottier P., Smeekens S., Van der Zeijst B.A.M. & Siddell S. (1983) *EMBO J.* 2, 1839–1844
7 Lai M.M.C., Baric R.S., Brayton P.R. & Stohlman S.A. (1984) *Proc. Natl. Acad. Sci. USA* 81, 3626–3630
8 Brown T.D.K., Boursnell M.E.G., Binns M.M. & Tomley F. (1984) *J. Gen. Virol.* 67, 221–228
9 Budzilowicz C.J., Wilczynski S.P. & Weiss S.R. (1985) *J. Virol.* 53, 834–840
10 Makino S., Stohlman S.A. & Lai M.M.C. (1986) *Proc. Natl. Acad. Sci. USA* 83, 4204–4208
11 Dennis D.E. & Brian D.A. (1982) *J. Virol.* 42, 153–164
12 Garwes D.J., Bountiff L., Millson G.C. & Elleman C.J. (1984) *Adv. Exp. Med. Biol.* 173, 79–93
13 Hu S., Bruszewski J., Boone T. & Souza L. (1984) *in: Modern Approaches to Vaccines* (Chanock R.M. & Lerner R.A., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., pp. 219–223
14 Jacobs L., Van der Zeijst B.A.M. & Horzinek M.C. (1986) *J. Virol.* 57, 1010–1015
15 Laude H., Chapsal J.M., Gelfi J., Labiau S. & Grosclaude J. (1986) *J. Gen. Virol.* 67, 119–130
16 Payvar F. & Schimke R.T. (1979) *J. Biol. Chem.* 254, 7636–7642
17 Okayama H. & Berg P. (1982) *Mol. Cell. Biol.* 2, 161
18 Hanahan D. (1985) *in: DNA cloning* 1 (Glover J.M., ed.), IRL Press, Oxford, pp. 109–135
19 Birnboim H.C. & Doly J. (1979) *Nucleic Acids Res.* 7, 1513–1523
20 Maniatis T., Fritsch E.T. & Sambrouk J. (1982) *in: Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., pp. 545
21 Vaquero C., Sanceau J., Catinot L., Andreu G., Falcoff E. & Falcoff R. (1982) *J. Interferon Res.* 2, 217–228
22 Deininger P.L. (1983) *Anal. Biochem.* 129, 216–223
23 Sanger F., Nicklen S. & Coulson A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467
24 Biggin M.D., Gibson T.J. & Hon G.F. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3963–3965
25 Chen E.Y. & Seeburg P.H. (1985) *DNA* 4, 165–170
26 Kapke P.A., Jung F.Y.C., Brian D.A. & Wesley R. (1987) *in: Biochemistry and Biology of Coronaviruses* (Lai M.M.C. & Stohlmann S., eds.), Plenum Press, New York, (*in press*)
27 Binns M.M., Boursnell M.E.G., Foulds I.J. & Brown T.D.K. (1985) *J. Virol. Methods* 11, 265–269
28 Skinner M.A., Ebner D. & Siddell S.G. (1985) *J. Gen. Virol.* 66, 581–592
29 Boursnell M.E.G., Binns M.M. & Brown T.D.K. (1985) *J. Gen. Virol.* 66, 2253–2258
30 Hopp T.P. & Woods K.R. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3824–3828
31 Kozak M. (1981) *Nucleic Acids Res.* 9, 5233–5252
32 Wesley R.D. & Woods R.D. (1986) *J. Gen. Virol.* 67, 1419–1425
33 Boursnell M.E.G. & Brown T.D.K. (1984) *Gene* 29, 87–92
34 Skinner M.A. & Siddell S.G. (1985) *J. Gen. Virol.* 66, 593–596