BIOCHEMISTRY OF CORONAVIRUSES 1980

Brian W.J. Mahy

Division of Virology, Department of Pathology,

University of Cambridge, England

INTRODUCTION

The coronavirus is something of a wolf in sheep's clothing. Originally identified as a large pleomorphic enveloped RNA virus with unique external morphological features and a helical internal nucleocapsid (Tyrrell et al., 1975), the close parallels with negative strand RNA viruses prompted many investigators to search for a coronavirion transcriptase. It soon became clear however, that, despite their morphological appearance, the coronaviruses possess an infectious genome RNA which must, by definition, be of positive polarity.

Considerable advances in our knowledge of coronavirus biochemistry have been made during the past two years, most of which are described in this volume, but still there are enormous gaps, particularly in understanding the mechanism by which these viruses multiply in the infected host cells.

THE GENOME

Several laboratories have determined the size of coronavirus genome RNA by gel electrophoresis in comparison with marker RNAs of known size. Table 1 summarises some of the values reported in this volume, and indicates an average genome molecular weight for a variety of different coronaviruses of 6 to 7 x 10^6. Thus

V. ter Meulen et al. (eds.), *Biochemistry and Biology of Coronaviruses*

coronavirus genome RNA is larger than that of any other known single-stranded RNA virus, and might be expected to code for as many as twelve average sized (50K) polypeptides.

A major question regarding coronavirus RNA is whether or not it consists of non-covalently bound subunits. Despite reports that the RNA of porcine (Garwes et al., 1975) or human (Hierholzer et al., this volume) coronaviruses dissociates on heating or in the presence of chaotropic agents, the weight of evidence suggests that this is a technical artifact. There is clear published data that avian (Lomniczi and Kennedy, 1977; Schochetman et al., 1977), murine (Lai and Stohlman, 1978; Wege et al., 1978) as well as porcine (Brian et al., 1980) and human (MacNaughton and Madge, 1978) RNAs do not dissociate into a subunit form as does oncornavirus RNA. Furthermore, complexity measurements of the avian coronavirus genome indicate that it is haploid (Lomniczi and Kennedy, 1977).

The presence of poly (A) within the genome RNA has been reported by several laboratories, though curiously most found that only 30 % of the genome RNA molecules were polyadenylated. The possibility that some genome RNA molecules do not contain poly (A), as with certain picornavirus RNAs (Frisby et al., 1976) cannot be ruled out. But since all the evidence derives from selection of the coronavirus genome RNA by binding to oligo-dT columns, the most likely explanation for only 30 % binding is that the poly (A) stretch is quite short so that some molecules do not bind sufficiently strongly.

Indeed, by binding to poly(U) sepharose after ribonuclease A digestion, a 2s poly (A) fraction was obtained from human coronavirus OC43, which is consistent with a poly(A) length of only 19 adenylate residues (Hierholzer et al., this volume). So far it has not been unequivocally established that the poly (A) is 3'-terminal, though this is presumed to be the case (MacNaughton and Madge, 1978).

Both Lai and Stohlman (this volume), for murine hepatitis virus, and Stern et al. (this volume) for avian IBV have found that the 5'-terminus of the genome RNA contains a methylated cap structure, which in the latter case is apparently of the form m^7GpppA (Kennedy, personal communication). This further emphasizes the positive strand nature of the coronavirus genome, which

Table 1. Comparative Sizes of Coronavirus-Specific RNAs Molecular Weights (X 10^{-6}) from Gel Electrophoresis

VIRUS	GENOME RNA	SUBGENOMIC RNAs					SOURCE
	1	2	3	4	5	6	7	(THIS VOLUME)
AVIAN IBV	6.9	-	2.6	1.5	1.3	0.9	0.8	STERN ET AL.
HUMAN 229 E	6.5	NOT REPORTED						HIERHOLZER ET AL.
HUMAN 229 E	6.0	2.8	2.1	1.3	1.0	0.8	0.6	WEISS AND LEIBOWITZ
HUMAN OC43	6.1	NOT REPORTED						HIERHOLZER ET AL.
MURINE A59	6.1	3.4	2.6	1.2	1.1	0.9	0.6	LEIBOWITZ AND WEISS
MURINE A59	5.6	4.0	3.0	1.4	1.2	0.9	0.6	VAN DER ZEIJST ET AL.
MURINE JHM	6.7	3.4	2.8	1.4	1.2	0.9	0.6	WEGE ET AL.
MURINE JHM	6.0	3.6	3.0	1.3	1.2	0.9	0.6	WEISS AND LEIBOWITZ
PORCINE TGEV	6.8	3.2	(1.6)*	1.4	(1.1)*	0.9	0.7	DENNIS AND BRIAN

*MINOR RNA SPECIES FOUND

must function directly as mRNA in the infected cell. Evidence that the isolated genome RNA is infectious has been obtained for murine (Wege et al., 1978), avian (Schochetman et al., 1977; Lomniczi, 1977) and porcine (Norman et al., 1968; Brian et al., 1980) coronaviruses.

INTRACELLULAR VIRUS-SPECIFIC RNAs

When coronavirus-infected cells are radioactively labelled with ^{3}H-uridine in the presence of sufficient actinomycin D to inhibit host cell mRNA synthesis, seven virus-specific polyadenylated RNAs can normally be detected. In this respect coronavirus genome expression differs from that of picornaviruses (where only genome-sized RNA functions as mRNA) or caliciviruses and togaviruses (where two mRNA species are found) but has some similarity to that of the other major group of positive strand viruses, the oncornaviruses.

It is clear from Table 1 that there is good agreement between various investigators on the sizes of the intracellular RNAs. In each case, RNA of genome size can be detected in infected cells in addition to six subgenomic RNA species, with the exception of IBV, where only five subgenomic RNAs are reported (Stern and Kennedy, 1980).

The sum of the molecular weights of the six subgenomic RNAs is considerably greater than that of the genome RNA; this reflects the fact that the intracellular RNAs share common sequences. Both Weiss and Leibowitz, and Stern et al. (this volume), have shown that all the virus-specific RNAs share a common 3'-terminal sequence. T_1 ribonuclease mapping studies indicate that all RNAs can be arranged as a 'nested set', with the sequence of each RNA being contained within the sequences of all larger RNAs, extending inward from the 3'-terminus of genome RNA (Leibowitz and Weiss, Stern et al., this volume).

In vitro translation of the subgenomic RNAs in micrococcal nuclease-treated rabbit reticulocyte or L cell lysates (Siddell et al this volume; Leibowitz and Weiss this volume) or in Xenopus oocytes (van der Zeijst et al. this volume) indicates that the size of any particular translation product corresponds to the difference in coding capacity between any one RNA and the next smallest RNA. For example, RNA 7 encodes the 60 K nucleocapsid protein, and RNA 6 a 23K precursor to the

envelope protein gp25. Similarly, the largest molecular weight difference is found between RNAs 3 and 4 (about 1.2×10^6) and this allows RNA 3 to encode the major high molecular weight envelope protein. Thus only the 5'-terminal portion of each subgenomic RNA is translated, and presumably there must exist a termination signal to prevent further read-through. It is still not clear how the genome RNA normally acts as a mRNA; the in vitro translation products which have been observed are of high molecular weight and do not correspond to any intracellular virus-specific proteins so far observed (Leibowitz and Weiss this volume). However, since the virion RNA is infectious, it is likely to encode one or more non-structural proteins required for its own replication.

The mechanism by which the virus-specific RNAs are synthesised is as yet unknown. Since they have the same polarity as genome RNA they must be transcribed from a complementary (negative strand) RNA template. No such template, or double-stranded replicative intermediate molecule, has yet been identified. Assuming that it exists, there are two likely mechanisms for subgenomic RNA synthesis: transcription of a full length plus strand copy of the template followed by processing into six smaller mRNAs, or independent initiation of the synthesis of each mRNA on the template molecule. Evidence that the latter scheme may be correct has been obtained by van der Zeijst et al (this volume), who determined the UV target sizes of the subgenomic RNA species induced in murine A59 virus-infected cells. Complete agreement was found between the physical sizes of RNAs 1 to 5 as determined by agarose gel electrophoresis and the calculated UV dose required to block their synthesis. These data appear to rule out the possibility that coronavirus RNAs are processed from a large precursor molecule. The existence of a short 5'-terminal sequence common to all the RNAs and derived by a splicing or polymerase jumping mechanism cannot, however, be excluded.

The nature of the virus-specific RNA polymerase complex is unknown, although Dennis and Brian (this volume) have found a cytoplasmic membrane-associated RNA-dependent RNA polymerase activity in cells infected with porcine coronavirus. On the other hand, Evans and Simpson (1980) have presented evidence that avian coronavirus replication requires participation of the host cell nucleus and is sensitive to inhibition of host cell DNA-dependent RNA polymerase form II by a-amanitin.

Since avian coronavirus multiplication is quite insensitive to 1 ug/ml actinomycin D in the growth medium (Kennedy, personal communication) and enucleated cells will apparently support the growth of murine coronavirus (Bond, personal communication) it is difficult to reconcile these findings until further experimental work has been carried out.

VIRUS-SPECIFIC PROTEINS

Many investigators have studied the number and sizes of the virion structural proteins, and a summary of results reported in this volume for a variety of coronaviruses is presented in Table 2. In all viruses studied so far, three major virion proteins can be detected. The most prominent of these is a 50 - 60 K protein which in association with RNA forms the helical nucleocapsid. As might be predicted from studies of negative strand virus nucleocapsid proteins, the coronavirus protein is found to be phosphorylated, presumably to facilitate interaction with the genome RNA. Siddell et al. (1981) have described a virion protein kinase activity but it is not clear whether this is a virus-specific enzyme or in which virion-associated proteins the kinase activity resides. Phosphorylation occurs on serine, but not threonine residues (Stohlman and Lai, 1979; Siddell et al., 1981).

The other two well-defined virion proteins are envelope glycoproteins. The largest forms the petal-like structure (peplomers) of the 'crown' and apparently has a conventional mode of synthesis and glycosylation similar to the glycoprotein of other enveloped viruses. This glycoprotein, which some investigators refer to as envelope 2 protein, is found in some viruses in related forms, with molecular weights of 90 K and about 180 K. The processing events which give rise to these two forms are at present unclear. The tryptic peptide patterns of the 90 K species and 180 K species are apparently identical (Sturman and Holmes, 1977). The high molecular weight form is trypsin sensitive, as are the spike glycoproteins of myxoviruses (Klenk and Rott, 1980) and is converted by trypsin treatment into the 90 K form(s) (Sturman and Holmes, 1977). As a consequence the relative amount of gp 90 to gp 180 may vary depending on the proteolytic activity in the host cell in which the virus is grown. This in turn reflects virus infectivity, since cleavage of gp 180 -gp 90 is probably required for infectivity (Storz et al., this volume). An interesting model for the structure of this large envelope glyco-

Table 2. Comparison of Coronavirus Structural Proteins

VIRUS	NUCLEOCAPSID	MEMBRANE	ENVELOPE	OTHER	SOURCE (THIS VOLUME)
AVIAN IBV	PP 51	gp 31	gp 90	gp 84 P 14	STERN ET AL.
BOVINE BECV	P 50	gp 23	gp 90/180	gp 65	STORZ ET AL.
BOVINE BECV	50	28	125	65 50 45 36	LAPORTE AND BOBULESCO
CANINE CCV	P 50	gp 32	gp 204	gp 22	GARWES
HUMAN 229 E	P 50	gp 24	gp 105	gp 160 P 22	MACNAUGHTON
HUMAN 229 E	P 47	gp 17	gp 105/gp 196	gp 165 gp 66 gp 31	HIERHOLZER ET AL
HUMAN OC43	P 47	gp 15	gp 104/gp 191	gp 165 gp 60 P 30	HIERHOLZER ET AL
MURINE A59	P 50	gp 23	gp 90/180		STURMAN
MURINE A59	60	25	170		BOND ET AL.
MURINE JHM	63	25	170		BOND ET AL.
MURINE JHM	PP 60	gp 25	gp 98/170	gp 65 P 23	SIDDELL ET AL.
PORCINE TGEV	P 50	gp 30	gp 200		GARWES

protein, containing a single trypsin-sensitive site and disulphide bonds necessary for the conformation of the molecule is presented by Sturman (this volume), and based on studies with murine coronavirus A 59.

There is general agreement that the smaller envelope glycoprotein (gp 25) fulfils a similar role in virion structure to that of the matrix protein of negative strand viruses, though in contrast the coronavirion protein is glycosylated. Some investigators refer to it as envelope I, but 'membrane protein' would probably be a more easily recognized designation. This protein is embedded in the lipid bilayer, and is largely unaltered by incubation of the virion with bromelain, which removes the peplomers, although the carbohydrate moiety is lost by such treatment (Sturman, 1977).

Glycosylation of this membrane protein is unusual in that it is insensitive to tunicamycin, in contrast to all other virus glycoproteins studied so far (including the 90 K coronavirus glycoprotein), (Holmes et al., Niemann and Klenk, this volume). This indicates that dolichol-linked N-acetylglucosamine plays no part in its synthesis, and it is probably that the carbohydrate linkages are of the o-glycosidic type as found in mucin (Niemann and Klenk, this volume). These results strongly suggest a novel mode of synthesis for the 25 K glycoprotein, involving a different cellular compartment to that involved in 90 K glycoprotein synthesis. In agreement, with this idea, it is found that neither mannose nor fucose can be incorporated into this protein (Storz et al, this volume). When coronavirus infects cells in the presence of tunicamycin, synthesis of the 90 K peplomer protein is inhibited, but 'bald' particles still bud out from the cells, and contain only the nucleocapsid and 25 K membrane protein: the particles are non-infectious, however (Holmes et al; Sturman, this volume).

Table 2 shows that in addition to the three well-defined virion proteins, several investigators report an additional glycoprotein of about 65 K and an additional 30 K protein has been reported for human coronavirus. Definition of the role and location of these additional proteins in the virion must await further investigation.

Currently there is no concensus of opinion regarding the virus-induced non-structural proteins found in coronavirus-infected cells. Comparison of

results from different laboratories lead to no useful conclusions, but now that well defined virus-cell systems are under investigation, particularly with the murine and porcine viruses, a great deal will be learnt in this area over the next few years. Such studies are necessary to provide a firm basis for investigation of the wide range of persistently-infected cells now available. Unravelling the molecular basis of this persistence both in vitro and ultimately in vivo, is a major goal for future research.

REFERENCES

Brian, D.A., Dennis, D.E. and Guy, J.S. (1980). Genome of porcine transmissible gastroenteritits virus. J. Virol. 34: 410-415.

Evans, M.R. and Simpson, R.W. (1980). The coronavirus avian infectious bronchitis virus requires the cell nucleus and host transcritpional factors.
Frisby, D., Smith, J., Jeffers, V. and Porter, A. (1976). Size and location of poly (A) in encephalomyocarditis virus RNA. Nucleic Acids Res. 3, 2789-2809.

Garwes, D.J., Pocock, D.H. and Wijaszka, T.M. (1975). Identification of heat-dissociable RNA complexes in two porcine coronaviruses. Nature, London 257, 508-510.

Klenk, H.D., and Rott, R. (1980). Cotranslational and post-translational processing of viral glycoproteins. Current Topics in Microbiol. and Immunol., 90, 19-48.

Lai, M.M.C. and Stohlman, S.A. (1978). RNA of mouse hepatitis virus. J. Virol. 26, 236-242.

Lomniczi, B. (1977). Biological properties of avian coronavirus RNA. J. gen. Virol. 36, 531-533.

Lomniczi, B. and Kennedy, I. (1977). Genome of infectious bronchitis virus. J. Virol. 24, 99-107.

Macnaughton, M.R. and Madge, M.H., (1978). The genome of human coronavirus strain 229E. J. gen. Virol. 39, 497-504.

Norman, J.O., and McClurkin, A.W. (1968). Infectious nucleic acid from a transmissible agent causing gastroenteritis in pigs. J. Comp. Pathol. 78, 227-235.

Schochetman, G., Stevens, R.H. and Simpson, R.W. (1977). Presence of infectious polyadenylated RNA in the coronavirus avian bronchitis virus. Virology 77, 772-782.

Siddell, S.G., Barthel, A. and ter Meulen, V. (1981). Coronavirus JHM. A virion-associated protein kinase. J. gen. Virol. 52, in press.

Stern, D.F. and Kennedy, S.I.T., (1980). Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. J. Virol. 36, 440-449.

Stohlman, S.A. and Lai, M.M.C. (1979). Phosphoproteins of murine hepatitis viruses. J. Virol. 32, 672-675.

Sturman, L.S. (1977). Characterization of a coronavirus I. Structural proteins: effects of preparative conditions on the migration of protein in polyacrylamide gels. Virology 77, 637-649.

Sturman, L.S. and Holmes, K.V. (1977). Characterization of a coronavirus. II. Glycoproteins of the viral envelope: tryptic peptide analysis. Virology 77, 650-660.

Tyrrell, D.A.J., Almeida, J.D., Cunningham, C.H., Dowdle, W.R., Hofstad, M.S., McIntosh, K., Tajima, M., Zakstelskaya, L.Y., Easterday, B.C., Kapikian, A. and Bingham, R.W. (1975). Coronaviridae. Intervirology 5, 76-82.

Wege, H., Müller, A., and ter Meulen, V. (1978). Genomic RNA of the murine coronavirus JHM. J. gen. Virol. 41, 217-227.