Volume 16 Number 22 1988 Nucleic Acids Research

Differential premature termination of transcription as a proposed mechanism for the regulation of coronavirus gene expression

D.A.M.Konings, P.J.Bredenbeek[1], J.F.H.Noten[1], P.Hogeweg[2] and W.J.M.Spaan[1]

European Molecular Biology Laboratory, Postfach 10.2209, 6900 Heidelberg, FRG, [1]Institute of Virology, Department of Infectious Diseases and Immunology, Yalelaan 1 and [2]Bioinformatics Group, Padualaan 8, University of Utrecht, Utrecht, The Netherlands

Received August 2, 1988; Revised September 29, 1988, Accepted October 21, 1988

ABSTRACT

We propose that the different subgenomic mRNA levels of coronaviruses are controlled through differential premature termination of transcription, and are modulated by the relative strength of transcriptional initiation/blockage events. We present the complete set of sequences covering the leader encoding and intergenic regions of the MHV-A59 strain. A computer-assisted analysis of the two now complete sets of these sequences of strain IBV-M42 and MHV-A59 shows that, in contrast to the previous theory, differences amongst stabilities of intermolecular base-pairings between the leader and the intergenic regions are not sufficient to determine the mRNA gradients in both MHV and IBV infected cells. Neither can the accessibility of the interacting regions on the leader and the negative stranded genome, as revealed by secondary structure analysis, explain the mRNA levels. The nested gene organisation itself, on the other hand, could be responsible for observed mRNA levels gradually increasing with gene order. Relatively slow new initiation events at intergenic regions are proposed to block elongation of passing transcripts which, via temporary pausing, can cause premature termination of transcription. This effects longer transcripts more than shorter ones.

1. INTRODUCTION

Coronaviruses form a group of enveloped positive stranded RNA viruses. Mouse hepatitis virus (MHV) and avian infectious bronchitis virus (IBV) are the most thoroughly studied members of the coronavirus family. Their genomes are linear, unsegmented and 27-30 kb in length. The replication of coronaviruses takes place in the cytoplasm of the host cell (1). The genome-length negative stranded RNA serves as a template for the transcription of the genomic RNA and subgenomic mRNAs (2). Depending on whether the virus is IBV or MHV, five or six subgenomic mRNAs can be isolated from infected cells (see Fig. 1 for details on the genome organisation, the nomenclature and sizes of the viral intracellular RNAs). The viral mRNAs are synthesized in non-equimolar but constant amounts throughout infection (3-5). RNase T1 fingerprinting and hybridization studies have shown that the subgenomic mRNAs of IBV and MHV form a 3' coterminal nested set (3, 4, 6-9). U.V. transcription mapping studies have displayed an independent transcription of viral subgenomic RNAs, i.e. they are not produced by processing of larger RNAs (10, 11). However, the subgenomic RNAs share a 5' leader sequence of about 60 nucleotides (IBV) or 72 nucleotides (MHV) (12-14) which is encoded by the 3' end of the negative stranded template (12-16). Several authors proposed a discontinuous transcription process for coronaviral mRNA synthesis as an explanation for these data (12, 15, 17).

Recently, several data compatible with a leader-primed transcription mechanism have been published (18, 19, 20). This transcription mechanism, as postulated by Spaan et al. (12) and Baric et al. (17), is therefore the most favored model to explain the discontinuous transcription mechanism of coronaviruses. In this model the leader RNA is transcribed independently of the mRNA body. After termination of the leader transcription the leader translocates to conserved sequences on the negative strand to serve as a primer for the initation of mRNA body transcription. This transcription mechanism is further supported by sequence data from IBV and MHV demonstrating the sequence complementarity between the intergenic transcription-initiation sites and the 3' end of the leader transcript (12, 14, 16, 21-24). To explain the non-equimolar transcription of coronaviral mRNAs in terms of the leader-primed transcription model it has been suggested - on the basis of a limited (sub)set of MHV-A59 and MHV-JHM intergenic regions - that the degree of complementarity between the free leader and the different intergenic regions is the main factor determining the rate of transcriptional initiation and thereby the level of the different mRNAs in the cytoplasm (12, 22-24). An increase in the degree of complementarity should result in the production of higher levels of the appropriate full-length mRNA.
To evaluate more thoroughly the mechanism responsible for coronaviral gene expression we analysed apart from intermolecular base-pairing potentials, also other structural features of the (sub-)genomic RNA(s) which may play a role in the leader-primed transcription of complete sets of sequence data from two strains. Therefore we determined the sequence of the genome 5' end and the complete set of the intergenic regions of MHV-A59. These data are used to quantify the intermolecular base-pairing and the secondary structure features of the separate leader and intergenic regions of MHV-A59. Since we expect the features determining the transcriptional regulation to be identical for each coronavirus we have carried out the same analysis for IBV, whose sequence of the genomic RNA has been recently described (25 and ref. therein). The comparison of all these structural data points to a different regulation of gene expression of the coronavirus than has been suggested previously. This regulation, reflected by the amount of mRNA, may be related primarily to the linear position of the intergenic regions on the genome. The relative importance of the different structural influences is herein discussed.

2. MATERIAL AND METHODS

(a) cDNA synthesis and cloning

Viral genomic RNA and poly(A)$^+$-selected mRNAs were isolated from purified virions and infected Sac$^-$ cells respectively (26). First and second strand cDNA syntheses were basically carried out according to the method of Gubler and Hoffman (27) using pentanucleotides or oligo-dT as primers. Full details of the procedures used to clone the MHV-A59 genomic RNA will be presented elsewhere (Bredenbeek et al., manuscript in preparation).

(b) Screening and analysis of the recombinants

In situ hybridization of bacterial colonies with kinase-labeled oligonucleotides or nick-translated MHV-A59-specific inserts were performed according to Maniatis et al. (28) and Meinkoth and Wahl (29). Full details on the localisation of the probes used and the mapping of the obtained cDNA clones on the MHV-A59 genome will be described elsewhere

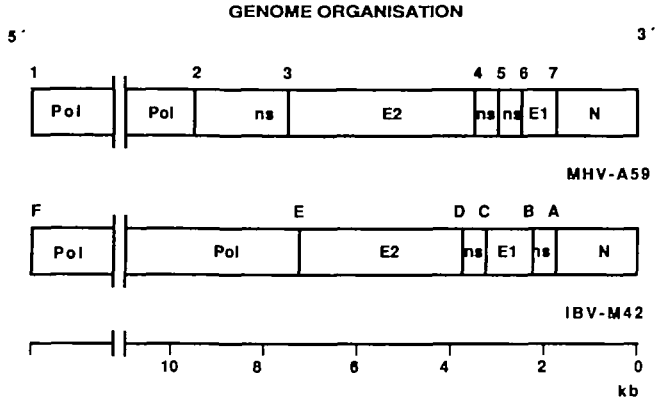

Figure 1

Genomic organisation of the coronaviruses MHV-A59 and IBV-M42. The genomes are presented The horizontal black bars show the location of the conserved intergenic sequences. The numbers (1-7) and letters (A-F) above the intergenic regions indicate the nomenclature of the corresponding mRNA transcripts for MHV-A59 and IBV-M42, respectively. The different proteins are indicated by their nomenclature in the enclosed open bars on the genome. Pol: polymerase protein; E2: spike protein; E1: trans-membrane protein; N: nucleocapsid protein; n.s.: non-structural protein.

(Bredenbeek et al., manuscript in preparation). Plasmid DNA was prepared from positive colonies according to Maniatis (28). Restriction enzyme digestions were performed using standard methods.

(c) DNA sequence analysis

Virus-specific DNA fragments were prepared by digestion of recombinant plasmid DNA with suitable restriction enzymes. After separation by agarose gel electrophoresis, fragments were purified by binding to NA45 paper (Schleicher and Schuell) and subsequently recloned into bacteriophage M13 using standard procedures (28). Single stranded M13 DNA was isolated and used for sequence analysis applying the dideoxynucleotide chain termination procedure of Sanger et al. (30).

(d) Minimal – energy foldings of leader and intergenic regions

Potential secondary structures of the leader transcripts and the intergenic regions were studied using the dynamic programming algorithm described by Williams and Tinoco (31). The algorithm generates all optimal free energy structures as well as defined suboptimal structures; knowledge of these various possible structures is important because the rules of RNA folding, and the (de-)stabilizing effects of other molecules such as proteins, are still poorly understood. The recently improved set of free energy values by Freier et al. (32) were used in the analysis. For the leader transcripts stable secondary structures up to a 20% energy difference to the optimal structure were generated. Since the negative free energy of the leader structures of both strains is relatively low, i.e. about -10 kcal, an energy difference of 20% is equivalent to only a few kcal difference and therefore reasonable to consider. The alternative structures (Fig. 3a) for MHV-A59 are the optimal minimal-

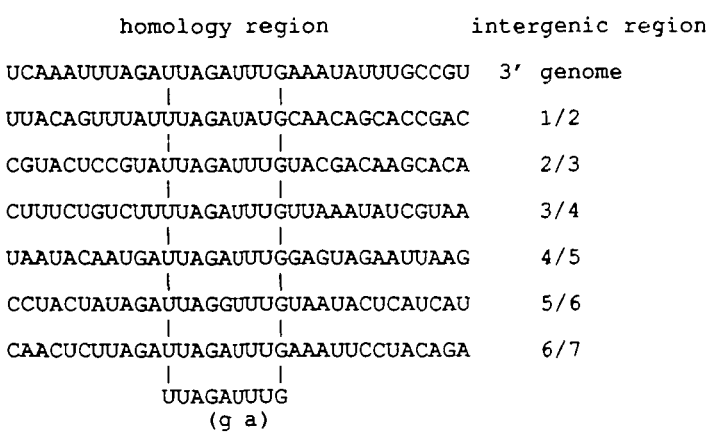

Figure 2
Comparison of all negative stranded fragments covering the MHV-A59 intergenic regions. The nomenclature of the intergenic regions is indicated by the two successive transcriptional units they separate (e.g 1/2).

energy structures. The IBV-M42 structure (Fig. 3b) presented belongs to a couple of most stable and similar structures (the actual length of which has not been determined). The potential secondary structure features of the intergenic regions were studied including flanking sequences up to a 150 nucleotides on both sites of the homology region. The reasons to limit our analysis to short-range interactions (i.e. 150 nucl. on both sides) are the following: (i) in contrast to the findings in structural RNAs, we do not expect very specific and global secondary structure constraints to occur in a protein encoding RNA fragment like the coronavirus genome; (ii) once transcription is occurring anywhere on the RNA template, long-range interactions are disrupted anyway and are thus of no interest; (iii) we have shown that invariance of secondary structure of the core intergenic regions occurs already in smaller parts of the flanking regions (100-200 nucl. in total) than the ones analysed (300 nucl. in total). Since the range of flanking sequences, which still can influence the structure of the homology region, is not defined, we compared the sets of optimal and suboptimal (up to 5% energy difference to the optimal) secondary structures by considering different lengths of flanking sequence. The most stable and frequent appearing structure(s) of each homology region was selected from this comparison. Although not a single defined secondary structure was suggested in all cases by this analysis, almost each set of alternative structures showed a certain tendency, i.e. a more accessible versus a more buried structure of the homology region.

3. DETERMINATION OF SEQUENCES INVOLVED IN THE MHV – A59 MRNA SYNTHESIS

The regions involved in the initiation of the body sequences of mRNA 7, 6 and 5 have been described previously (12, 22). We have determined the remaining three intergenic regions. The comparison of all the MHV-A59 intergenic regions is shown in Figure 2. Except for

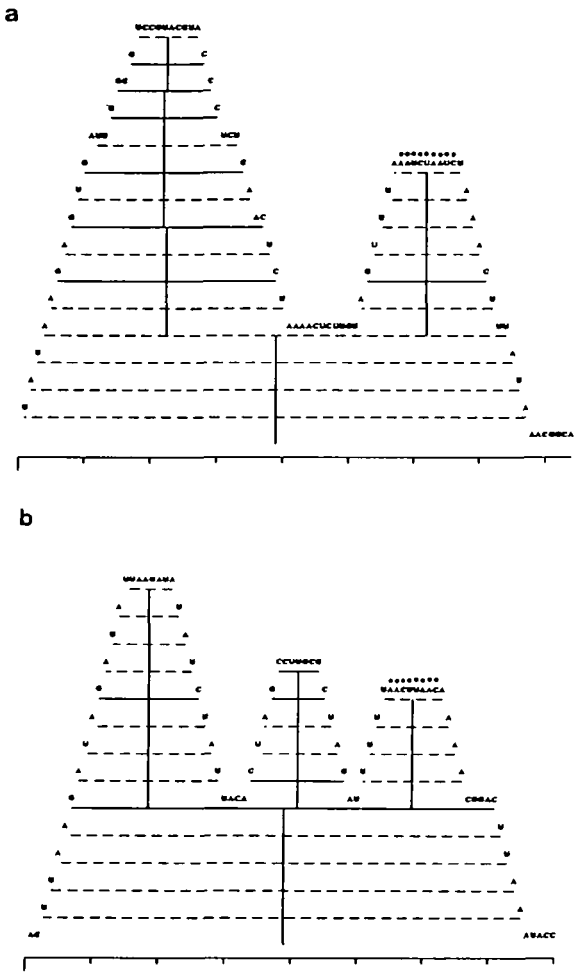

Figure 3

Minimal-energy foldings of the MHV-A59 (a) and IBV-M42 (b) major leader transcripts (see Material and Methods). Structures are represented according to Hogeweg and Hesper (33), in which each base pair is shown by a horizontal line (—: G-C, - - -: A-U and - · -: G-U bonds). • • •: the nucleotides of the homology region on the leader transcript which are supposed to base-pair with the homology region present in the intergenic regions.

a mismatch at the intergenic regions 1/2 and 5/6, the sequence 3' UUAGAUUUG 5' is conserved. This is indicative of these nucleotides importance for the recognition of the RNA polymerase or the polymerase/leader complex. To investigate the presence of this conserved region at the extreme 5' end of the genomic RNA, we have used a leader probe

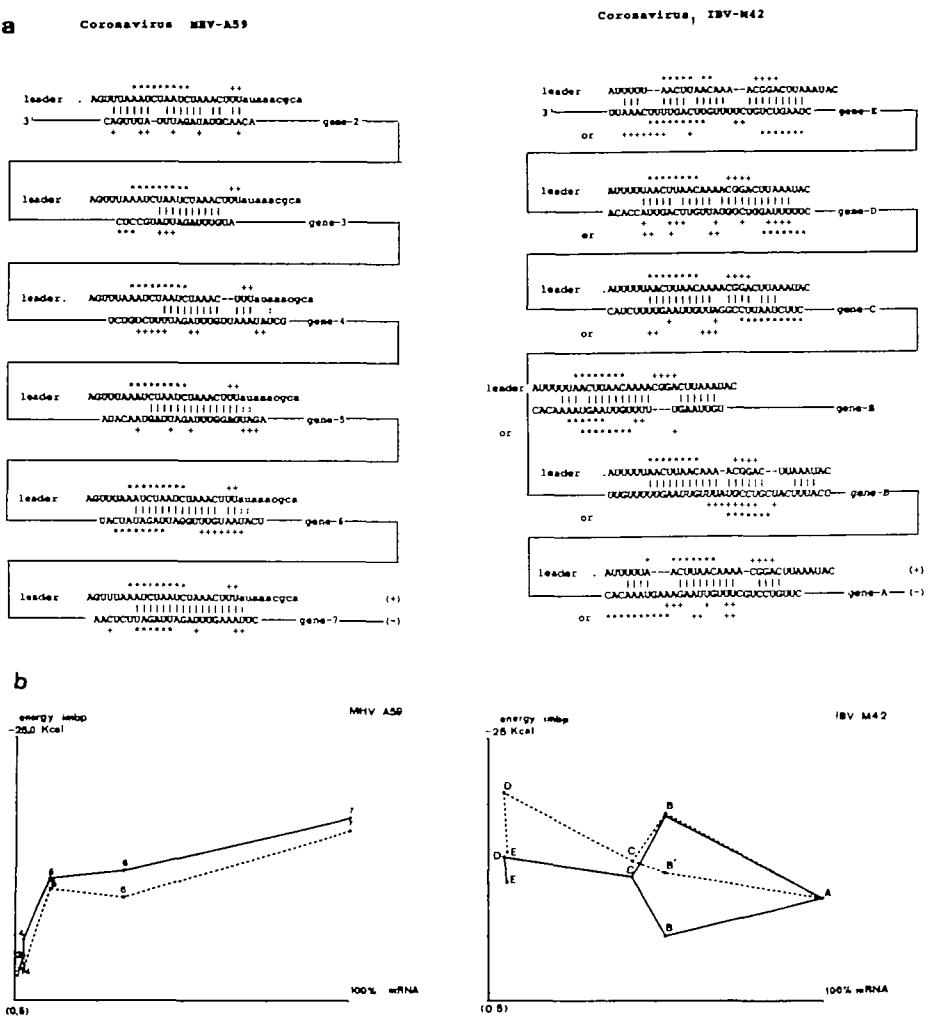

Figure 4

(a) Intra- and inter-molecular base-pairing scheme of the leader transcript and the intergenic promoter regions of coronaviruses MHV-A59 and IBV-M42. The intramolecular structural features are represented above the leader and below the promoter nucleotides in terms of single-stranded nucleotides (* or + symbols) (see Fig 3 for the entire secondary structure of the leaders). The lower case nucleotides in the leader sequence of MHV-A59 represent the 3' part of the minor leader transcript (i.e. larger than 74 nucl.). Alternative "optimal" structures are also presented. The intermolecular base-pairings (| and : symbols) represent stable interactions, they G-U base-pairs which should in principle not be considered as mismatches. The B and B' are alternative pairing schemes and are aligned to each other. For the leader (5' to 3', + strand) only the 3' terminal sequences and for the genome (3' to 5', - strand) only the homology regions and their direct flanking se-

quences are shown. *** = hairin-loop region; +++ = internal-loop or bulge; | = potential base-pairing between leader and intergenic region; : = idem, only pairing downstream of nucleotide 74 in the minor leader transcript of MHV-A59.
(b) Graphic representation of the relation between the free energy of the intermolecular base-pairing and the relative levels of mRNA in the cytoplasm (3, 4). The IBV-M42 molarities (3) have been corrected on recently elucidated mRNA lengths. ——: free energy of the full base-pairing shown; - - - : free energy on the major leader transcript for MHV-A59 (i.e. 74 nucl.) and including additional interactions at the 5' end (i.e. upstream of the UUUU-stretch in the leader) for IBV-M42 (not shown).

to screen the random primed genomic cDNA library. The cDNA inserts of two leader positive clones were recloned in bacteriophage M13 and sequenced (Fig. 2). The obtained leader sequence of the genomic RNA is identical to the leader sequences present at the 5' end of mRNA 5, 6 and 7 of MHV-A59. In addition, the conserved intergenic sequence is also located at the 5' end of the genome. Apart from a extra small repeat (5' AUCUA 3') in the MHV-JHM leader, the leader sequence of MHV-A59 does not differ essentially from the previously published MHV-JHM sequence (23).

4. DIFFERENCES IN POTENTIAL BINDING OF THE LEADER TO THE INTERGENIC REGIONS CAN NOT EXPLAIN MRNA LEVELS

(a) Secondary structure features of the leader transcripts and intergenic regions

The extent of potential base-pairings and the actual frequency of transcriptional initiation will be influenced by the accessibility of the involved nucleotides in both the separate leader and intergenic regions. This means that a strong intramolecular interaction will block, or compete with, a successive intermolecular interaction. On the other hand, the combination of a high base-pairing stability and a coinciding single-strandedness of the interacting regions will be a way to optimise the initiation of leader-primed transcription.
(i) For MHV-A59 the experimentally observed major leader transcripts (74, 77 and 84 nucleotides long) (19) were all considered in the analysis. The foldings of the major MHV transcripts suggest alternative secondary structures, one having an additional weak interaction between the two termini (see Fig. 3a). As also mentioned by Shieh et al. (23) the homology region of MHV (see Fig. 2) resides partly in the loop of the second hairpin and partly in the following 3' flank of the helix (Fig. 3a and 4a). We suppose that the interaction remains restricted to base-pairing with nucleotides present in the weak (A-U-rich) 3' helix of the leader. The leader transcript of IBV can form a structure in which the homology region covers an analogous single-stranded hairpin loop (Fig. 3b and 4a).
(ii) For the intergenic regions optimal and alternative optimal structures are presented for the purpose of comparison (see Fig. 4a, and Material and Methods). A comparison shows clearly that the intergenic regions preceding gene 6 and 7 are more accessible than the ones preceding genes 5, 4 and 3 in MHV (Fig. 4a). This trend is even stronger if one assumes a hairpin loop (* * **) to be more "attractive" than a couple of small internal loops or bulges.

(b) Leader – primed base – paring

The intermolecular base-pairings determined in the present study belong to the most stable interactions, based on the set of recently improved free energy values by Freier et al. (32) (Fig. 4a). In Figure 4b the free energy of the presented base-pairing are graphically

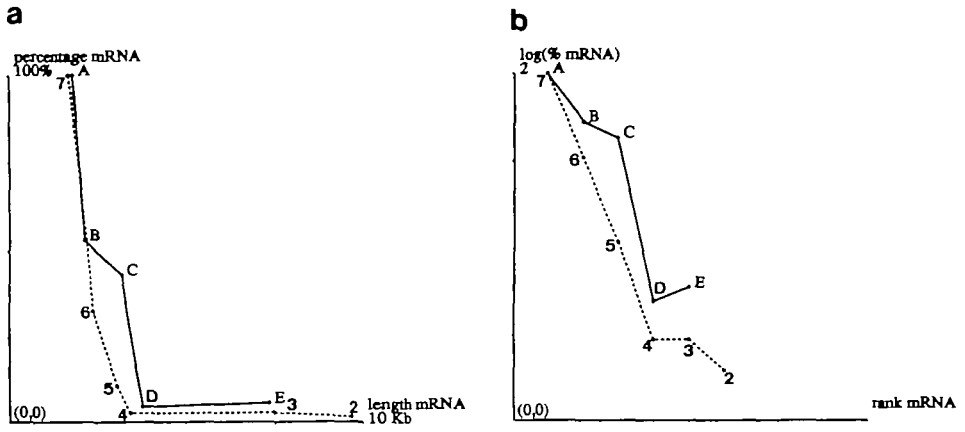

Figure 5

Graphic representation of the relation between the percentage and length of the mRNA transcriptional units by proposing unique single pausing "signals" across or between each gene border (for source of mRNA molarity data see Fig 4.).

quantified. The trend of these free energy values for MHV shows a decrease in stability from intergenic region 7 to region 2. This pattern holds for the set of base-pairings proposed for the longest (84 nucl.) and the shortest (74 nucl.) major leader transcripts alike. It should be noted, however, that the binding potential between the leader and the intergenic regions is not necessarily a linear function of the free energy of the interaction, since an "on-off" effect may be involved. Therefore, the MHV graph alone does not give a clue about the relevance of the relation. For IBV no analogous trend in the free energy values is present (Fig. 4b). Moreover, the stabilities of the different IBV base-pairings are overall higher and slightly less variable. Comparing the interactions of the intergenic regions of IBV reveals for RNA-A versus RNA-B and RNA-C versus RNA-D free energies of interactions which are inversely correlated with their transcription levels (Fig. 4b). The secondary structures of the intergenic regions of RNA-A and RNA-B which are involved in the intermolecular base-pairing are not the same, but this variation can not explain their differences at the mRNA level. Secondly, for RNA-C and RNA-D the structural features of the interaction are very similar whereas the corresponding mRNA levels are not.

From our data (Fig. 4) we conclude that differences between stabilities of intermolecular base-pairings alone are not sufficient to determine the observed levels of the different mRNAs in both coronavirus MHV and IBV. An additional consideration of secondary structure features of the interacting regions on the leader and the negative stranded genome can not explain the observed mRNA gradients, either.

5. NESTED GENE ORDER COULD EXPLAIN MRNA LEVELS

The stable levels of the mature mRNA increase with the gene order in both MHV-A59 and IBV-M42. A mRNA gradient like the one observed could in principle be achieved in several ways by a transcriptional phenomenon related to the genome organisation. Firstly,

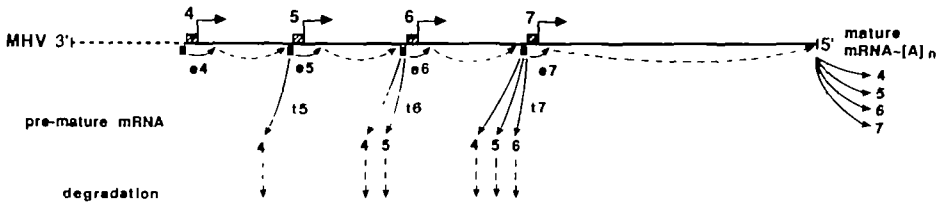

Figure 6
Transcriptional regulation model.
The effect of premature termination of transcription at or close to intergenic regions on the gene regulation of coronaviruses is illustrated for the second half of the MHV genome. The four genes 4, 5, 6 and 7 are marked by their transcription initiation sites (hatched boxes). Elongation between the different initiation sites is shown by a broken line. A certain chance of further elongation (e4, e5, e6 and e7) and premature termination (t5, t6 and t7) of passing transcripts will exist at each of the different intergenic regions, related to their frequency and rate of initiation/blockage (black boxes). Depending on these chances and on the number of initiation sites to be passed, variable proportions of the different mRNA transcripts will become full-length and thus polyadenylated ([A]n). As a result of the absence of a poly-A tail the premature terminated transcripts will have a high chance of rapid degradation (vertical dotted line).

transcription elongation could occur as rate-limiting step in the mRNA synthesis (predicting a linear relation between length and level of mRNA, Fig 5a). Secondly, the presence of elongation pausing sites, due to, e.g., secondary structure, causing a premature termination of a certain proportion of elongation transcripts, will result in unequal levels of full-length mRNA of the different length (predicting a linear relation between length and log[mRNA level]). Thirdly, the occurence of unique single pausing "signals" across or between each gene border which perturb the elongation will cause a rank order in mRNA levels of the successive nested genes (predicting a linear relation between gene rank and log[mRNA level], Fig. 5b). Comparison of these three relations give a weak preference for the last explanation. More important, taking into account that transcription initiation is generally the rate-limiting step in transcription, the "single-site" model is the most likely, though possibly not exclusive, mechanism regulating coronavirus gene expression.
Since the nested, successive genes differ from each other by the number of active initiation sites they have to pass during their transcriptional elongation, site specific transcriptional initiation events at the intergenic regions will affect overlapping transcription units. We suggest blockage of passing transcription complexes related to a relatively slow initiation step to be important and a good candidate for a "single-site" model. The initiation event may cause a temporary pausing which could be followed by premature termination of the elongating transcripts (Fig. 6). As an extension of this regulatory mechanism, we expect that the levels of mRNA to be further modulated by variations in base-pairing abilities between the leader and intergenic regions and in accessibilities of the base-pairing domains in the two separate molecules. The combination of both of these structural features will determine the initiation frequency and possibly also the speed of initiation at a specific intergenic site.

6. DISCUSSION

Our theoretical analysis invalidates the previous working model, i.e. that differential base-pairing potential between the free leader and the intergenic regions is the basic principle regulating gene expression (12, 22-24). We have shown that secondary structure of the intergenic regions, at least locally, can not explain the observed levels, and have argued that an explanation in terms of global secondary structure is invalid because of the extensive amount of evolutionary information required and disruption of global structures during transcription. Based on these conclusions on intra- and inter-molecular base-pairing potentials and the striking relation between the gene order and the steady state mRNA levels (i.e. gradient), we propose differential premature termination of transcription as a minimal-explanation for the observed gene expression. We furthermore expect the different subgenomic mRNA levels of the coronavirus to be additionally modulated by variable transcriptional initiation/blockage events.

However, from the data available, several alternative explanations cannot be excluded. Firstly, differences in mRNA stabilities might also contribute significantly to the observed mRNA levels. From data on gene expression of eukaryotic mRNA degradation we can, however, conclude that parts of 3' untranslated regions of mRNAs are most clearly responsible for their stability and that, so far, there is no general correlation between mRNA length and degree of cytoplasmic stability (34). Since the different mRNAs of coronavirus all have identical 3' terminal sequences it is likely that the different mRNAs show indeed similar stabilities. Secondly, protein specificity could, in principle, act as the main factor determining the initiation event (rate/frequency) at the different intergenic regions and thus regulate gene expression. Apart from the fact that this explanation is too general and would require more evolutionary conserved information, it is, most of all, sharply contrasting the proposed explanation by differential premature termination of transcription in that it does not consider the most striking experimental data: the relation between gene order and mRNA gradient.

It is worth mentioning that, although the trend of mRNA levels is the same in nearly all cases, the measured relative levels of mRNA have been shown to vary to a greater or smaller extent between experiments performed in different laboratoria (3, 4, 10, Bredenbeek, unpublished results). Therefore it is not possible to fit the available data into a quantitative model of the postulated transcriptional regulation we developed.

Finally we discuss how our proposed transcription control mechanism relates to two different phenomena of premature transcriptional pausing/termination in viruses.

(i) The discontinous and processive RNA replication of coronaviruses proposed by Makino et al. (20) describes the transcription by a possible sequence of pausing, falling-off and reinitiation events. This mechanism could account for unusual high levels of RNA-RNA recombination of separate genomes via free RNA intermediates (19, 35, 36). Baric et al. (19) specified several functional RNA intermediates of discrete size. It is not clear, however, if the majority of these intermediates really exist as free premature transcripts or if they occur as classes of halting transcripts on discrete sites on the template. The different pausing sites reported by Baric are unspecifically located in the secondary structure and comparable potential secondary structures exist all over the genome. Thus it is likely that premature pausing sites can in principle be found in all parts of the genes; their num-

ber will correlate with the lengths of the genes. Because of the order rather than length correlation of the observed mRNA levels and because of the relatively long transcription initiation process, we expect that the premature transcription termination proposed by us is more important in the light of gene expression.

(ii) The intracellular mRNAs of the negative stranded rhabdovirus (VSV) do not overlap, they are transcribed from a non-nested set of genes (37). Furthermore, the analogous genes appear in reverse order on the genomes of the two viruses, i.e. the N gene of VSV, in terms of transcriptional direction, is located most upstream whereas it is located most downstream in coronavirus. It is interesting that the reverse gene order of VSV goes along with a different, here sequential transcription mechanism, but results in a similar gradient of the different mRNAs coding for analogous proteins. This means that in both virusses the N gene generates the highest mRNA levels. The sequential transcription of VSV genes implies that the expression of successive genes is dependent on the prior transcription of the upstream gene(s) whereas in coronavirus the transcription of the different genes is independent (10). For VSV it has been shown that termination of transcription occurs upon pausing specifically at or close to gene junctions. This restricted termination (about 30% across gene borders) accounts for the observed VSV mRNA gradient because of the sequential transcription of the genes: transcription initiation rates at the successive genes can only decrease with the direction of transcription. It should be noted, however, that termination of transcription is not inherent to the way transcription initiates (e.g. the sequential transcription of VSV).

We conclude that premature termination of transcription could be of more general importance for viral transcriptional regulation , however via different mechanisms.

The proposed regulation model predicts a group of mRNA transcripts of discrete length, ending at or close to intergenic regions, to be formed in the infected cells. However, the possibility to detect this class of transcripts will rely on their relative stabilities in the cytoplasm, which unfortunately, due to the absence of a poly-A tail, are expected to be low. Furthermore, the unambiguous identification of these transcripts will be hampered since many intermediate RNA species have been detected on the replicative intermediate RNA (19). Our model also predicts that for example introduction of an extra intergenic region in the coronaviral genome will affect the transcription level of upstream transcription units. This would be the most direct approach to test our model and which can be performed as soon as an infectious cDNA clone becomes available.

ACKNOWLEDGMENTS

We would like to thank Dr. B.A.M. van Zeijst and Dr. I. Mattaj for helpful discussions and suggestions and C. Raulfs for help in editing the manuscript. P.B. was sponsored by the Netherlands Foundation for the Advancement of Pure Research (ZWO).

REFERENCES

1. Wilhelmsen, K.C., Leibowitz, J.L., Bond, C.W. and Robb, J.A. (1981). Virology 110, 225-230.
2. Brayton, P.R., Stohlman, S.A and Lai, M.M.C. (1984). Virology 133, 197-201.
3. Stern, D.F. and Kennedy, S.I.T. (1980a). J. Virol. 34, 665-674.

4. Leibowitz, J.L., Wilhelmsen, K.C. and Bond, C.lW. (1981). Virology 114, 39-51.
5. Sawicki, S.G. and Sawicki, D.L. (1986). J. Virol. 57, 328-334.
6. Stern, D.F. and Kennedy, S.I.T. (1980b). J. Virol. 36, 440-449.
7. Lai, M.M.C., Brayton, P.R., Armen, R.C., Pugh, C. and Stohlman, S.A. (1981). J. Virol. 39, 823-834.
8. Spaan, W.J.M., Rottier, P.J.M., Horzinek, M.C. and van der Zeijst, B.A.M. (1982). J. Virol. 42, 432-439.
9. Weiss, S.R. and Leibowitz, J.L. (1983). J. Gen. Virol. 64, 127-133.
10. Jacobs, L., Spaan, W.J.M., Horzinek, M.C. and van der Zeijst, B.A.M. (1981). J. Virol. 39, 401-406.
11. Stern, D.F. and Sefton, B.M. (1982). J. Virol. 42, 755-759.
12. Spaan, W.J.M., Delius, H., Skinner, M., Armstrong, M., Rottier, P.J.M., Smeekens, S., van der Zeijst, B.A.M. and Siddell, S.G. (1983). EMBO J. 2, 1839-1844.
13. Lai, M.M.C., Baric, R.S., Brayton, P.R.and Stohlman, S.A. (1984). Proc. Natl. Acad. Sci. USA 81, 3626-3630.
14. Brown, T.D.K., Boursnell, M.E.G. and Binns, M.M. (1984). J. Gen. Virol. 65, 1437-1442.
15. Lai, M.M.C., Patton, C.D., Baric, R.S. and Stohlman, S.A. (1983). J. Virol. 46, 1027-1033.
16. Brown, T.D.K., Boursnell, M.E.G., Binns, M.M. and Tomley, F.M. (1986). J. Gen. Virol. 67, 221-228.
17. Baric, R.S., Stohlman, S.A. and Lai, M.M.C. (1983). J. Virol. 48, 633-640.
18. Baric, R.S., Stohlman, S.A., Razavi, M.K. and Lai, M.M.C. (1985). Virus Res. 3, 19-33.
19. Baric, R.S., Shieh, C.-K., Stohlman, S.A. and Lai, M.M.C. (1987). Virology 156, 342-354.
20. Makino, S., Stohlman, S.A. and Lai, M.C.M. (1986). Proc. Natl. Acad. Sci. USA 83, 4204-420.
21. Brown, T.D.K. and Boursnell, M.E.G. (1984). Virus Res. 1, 15-24.
22. Budzilowicz, C.J., Wilczynski, S.P. and Weiss, S.R. (1985). J. Virol. 53, 834-840.
23. Shieh, C.-K., Soe, L.H., Makino, S., Chang, M.-F., Stohlman, S.A. and Lai, M.M.C. (1987). Virology 156, 321-330.
24. Bredenbeek. P.J., Charite, J., Noten, J.F.A., Luytjes, W., Horzinek, M.C., van der Zeijst, B.A.M. and Spaan, W.J.M. (1987). in Coronaviruses, M.M.C. Lai and S. Stohlman, eds. Adv. Exp. Med. Biol. 218, pp 165-172.
25. Boursnell, M.E.G., Brown, T.D.K., Foulds, I.J., Green, P.F., Tomley, F.M. and Binns, M.M. (1987). J. Gen. Virol. 68, 57-77.
26. Spaan, W.J.M., Rottier, P.J.M., Horzinek, M.C. and van der Zeijst, B.A.M. (1981). Virology 108, 424-434.
27. Gubler, U. and Hoffman, B.J. (1983). Gene 25, 263-269.
28. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular cloning, a laboratory manual. Cold Spring Harbor laboratory.
29. Meinkoth, J. and Wahl, G. (1984). Anal. Biochem. 138, 267-284.
30. Sanger, F., Couloon, A.R., Barell, B.G., Smith, A-J.,H. and Roe, B.A. (1980). J. Mol. Biol. 143, 161-178.
31. Williams, A.L. and Tinoco, I. (1984). Nucl. Acids Res. 14, 299-315.
32. Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. and Turner, D.H. (1986). Proc. Natl. Acad. Sci. USA 83, 9373-9377.
33. Hogeweg, P. and Hesper, B. (1984). Nucl. Acids Res. 12, 67-74.
34. Shapiro, D.J., Blume, J.E. and Nielsen, D.A. (1987). BioEssays 6, 221-226.
35. Lai, M.M.C., Baric, R.S., Makino, S., Keck, J.G., Egbert, J., Leibowitz, J.L. and Stohlman, S.A. (1985). J. Virol. 56, 449-456.
36. Keck, J.G., Stohlman, S.A., Soe, L.H., Makino, S. and Lai, M.M.C. (1987). Virol. 15, 331-341.
37. Banerjee, A.K. (1987). Microbiol. Rev. 51, 66-87.