An Analytical Review of Defective Infections of Vesicular Stomatitis Virus

By BENJAMIN M. BLUMBERG* AND DANIEL KOLAKOFSKY

Department of Microbiology, University of Geneva Medical School, 64 Avenue de la Roseraie, CH-1205 Geneva, Switzerland

INTRODUCTION

Following the discovery of interference in influenza infections by incomplete forms of the virus (von Magnus, 1951), Cooper & Bellett (1959) found a similar phenomenon in vesicular stomatitis virus (VSV) infections, and showed that it was due to transmissible components, defective interfering (DI) particles (Hackett, 1964; Huang & Wagner, 1966), which are usually generated when VSV is passaged at a high multiplicity of infection (m.o.i.). DI particles can replicate only during co-infection with standard or non-defective (ND) helper virus, and in such mixed virus infections the replication of the helper virus is greatly repressed. Bellett & Cooper (1959) observed a reciprocal exponential relationship between the m.o.i of DI particles in the inoculum and the yield of infectious virion progeny, and deduced that co-infection of a cell with a single DI particle is sufficient to repress the helper virus in one passage (compare Sekellick & Marcus, 1980). DI particles enjoy a replicative advantage over ND virions, and once generated they quickly become the dominant viral species during high m.o.i. passage. The DI genomes have been shown to be deletion mutants of the ND viral genome (for reviews, see Lazzarini et al., 1981; Perrault, 1981), and most DI genomes that are stable during high m.o.i. passage contain at their 3' ends a sequence of 46 nucleotides which is also found at the 3' end of the ND antigenome ('copyback' Dis). The major exception to this pattern is the DI–LT (Petric & Prevec, 1970) which contains the same organization as the ND genome with most of the L gene deleted. This particle has a limited replicative advantage over the helper virus (Bay & Reichmann, 1982) but is not stable to continued high m.o.i. passage, and is out-competed by Dis which contain the antigenomic 3' end sequences.

The reason for the order of magnitude replicative advantage of the copyback Dis is not known. The rate-limiting step has been linked to genome replication (Huang & Manders, 1972; Palma et al., 1974), but probably is not at the level of genome chain elongation since the relative replicative advantages of Dis which vary greatly in length do not appear to be based on their sizes (Holland et al., 1976; Schnitzlein & Reichmann, 1977a; Kolakofsky, 1981). This advantage might reflect a competition among the template RNAs for a limiting amount of viral replicase (Stampfer et al., 1969; Perrault & Holland, 1972), perhaps due to a higher affinity of the replicase for the DI 3' ends, resulting in an increased rate of chain initiation on DI templates. However, several lines of evidence suggest otherwise. Both ND and DI particles carry into the infected cell a full complement of viral proteins (Kang & Prevec, 1969; Wagner et al., 1969) including a catalytically active, nucleocapsid-specific RNA polymerase (Baltimore et al., 1970; Emerson & Wagner, 1972; Huang & Manders, 1972). The viral replicase may differ from the transcriptase (Clinton et al., 1978; Kingsford & Emerson, 1980); nevertheless, the same viral polymerase that is active in transcription can also replicate the genome in vivo (Perlman & Huang, 1973) and in vitro (Testa et al., 1980). Whatever the nature of the viral replicase, if its binding affinity for the antigenome and the Dis is higher than for the genome, then competition among these templates ought to result in a simple inverse proportionality between DI dose and ND virion progeny that would be described by a rectangular hyperbola. Instead, a reciprocal exponential relationship is observed at low m.o.i. (Bellett & Cooper, 1959; Huang & Wagner, 1966; Schnitzlein & Reichmann, 1977b; Horodyski & Holland, 1980), and is sufficiently
reproducible to be used as an assay for DI particles (Marcus & Sekellick, 1974; Winship & Thacore, 1979). This relationship is not consistent with a simple competition mechanism, nor are the dose-response properties of the heterotypic defective non-interfering particles described by Adachi & Lazzarini (1978). Moreover, the genomic and antigenomic 3' end sequences, the putative polymerase attachment sites, are similar: in VSV (Indiana) 14 out of the first 17 nucleotides are the same, and in VSV (New Jersey) the sequences are identical for the first 20 bases (Keene et al., 1981).

A model for virus replication in mixed infections

One way around these difficulties would be a replication mechanism involving a rate-limiting step subsequent to initiation. This feature is inherent in a model (Kingsbury, 1974; Leppert et al., 1979; Blumberg et al., 1981, 1983; Kolakofsky & Blumberg, 1982) in which the viral polymerase is assumed to function constitutively, being switched between its alternative transcriptase and replicase activities in response to the intracellular concentration of free viral nucleocapsid (N) protein. In this model (Fig. 1), replication not only requires the polymerase to initiate an RNA chain at the 3' end promoter (P-) of the genome template, but also requires the polymerase to read through the leader RNA termination signal (T-). The viral N protein acts as the switching element by its ability to bind specifically to the nascent plus- and minus-strand leader RNA chains, which both contain the sequence for encapsidation (Blumberg & Kolakofsky, 1981; Blumberg et al., 1983). When present in sufficient concentration to encapsidate the nascent leader chain, the N protein causes the viral polymerase to read through the termination signal, leading to synthesis of a full-length antigenome nucleocapsid. Otherwise, as during primary transcription, the polymerase would terminate this chain to give a free leader RNA. Termination of the leader RNA chain then allows re-initiation of RNA synthesis at the promoter for the N protein gene (P_N). Similar leader RNA termination signals (T+) exist on the ND antigenome and the DI templates, but as (P+) is the only promoter on these templates, no further RNA synthesis is possible once the polymerase terminates the leader chain.

If, in the above model, a higher concentration of free N protein were required to cause the polymerase to read through the genomic terminator (T-) than through the antigenomic and DI
Review: Analysis of interference in VSV

terminators \((T+\)), then a low level of free N protein would favour the synthesis of genomes over antigenomes (as is observed, see Soria et al., 1974; Simonsen et al., 1979) and of copyback DIs over ND virion progeny. The model predicts that the N protein controls both its own synthesis and its utilization, so that the concentration of free intracellular N protein is homeostatically regulated. Due to its high stoichiometric requirement in nucleocapsid assembly (approx. 2000 N per genome) the concentration of N protein rises to a level just sufficient to allow replication. In a mixed virus infection, the concentration of N protein would rise enough to allow replication of the DI genomes, but it would remain below that required for normal amplification of the ND genomes.

Below a critical concentration of free intracellular N protein, DI genomes could replicate while ND genomes could only function as templates for transcription. This may be the case where superinfection with too great a multiplicity of DI particles leads to repression of both DI and ND progeny (Bellett & Cooper, 1959; Stampfer et al., 1969; Wild, 1972; Holland & Villareal, 1975; Khan & Lazzarini, 1977). This defective autointerference now might be attributed to a surfeit of easily suppressed DI terminators. When enough DIs are present, newly synthesized N protein goes entirely for DI particle replication, and ND genome amplification is prevented. Thus, only primary mRNA transcripts are made, and these would be seen as a limiting factor in genome replication (Stampfer et al., 1969; Perrault & Holland, 1972; Stamminger & Lazzarini, 1977).

Co-infection of cells with a more modest multiplicity of DI particles may result in the temporal regulation of the virion progeny. For example, the harvest of ND virions is maximal at 6 to 8 h post-infection, while DI progeny steadily increase up to 17 h post-infection (Khan & Lazzarini, 1977). The cytopathic effects which set in at about 6 h post-infection may in part prevent further ND amplification, because the high rate of N protein synthesis required for ND genome replication in competition with DI genome replication cannot be maintained. However, DI amplification can continue due to the lower concentration of free N protein needed to suppress the copyback DI terminators. Perhaps the most striking example of temporal regulation is the observation of Huang & Wagner (1966) that when superinfection with DI particles is progressively delayed in time following infection with standard virus, the yield of ND progeny increases, until at a delay of 2.5 h the full normal yield of ND virus is regained, even though the DI progeny are simultaneously amplified (Stampfer et al., 1969; Khan & Lazzarini, 1977). Thus, ND replication is literally suppressed by co-infection with DI particles: for if the ND genomes are given the time to amplify, so that secondary transcription becomes important, they can replicate normally in the presence of DIs. Such temporal regulation would necessarily be lacking in a mechanism involving competition at the initiation step, since that implies a proportionate amount of initiation even at the disfavoured promoter, and ND genome amplification early in infection could not be avoided.

Mathematical analysis of the model

The model may be formulated and evaluated by reference to the experimental data. The viral polymerase is assumed to function constitutively, so that the rate of synthesis of its RNA products will be proportional to the product of three factors: the promoter strength of the template RNA, the concentration of this RNA, and a switching function which expresses the fact that a given polymerase must either transcribe or replicate its template. A simple switching function for genome replication, which will take values between 0 and 1.0 like a probability function because transcription and replication are mutually exclusive events, is that for the S-shaped curve (Fig. 2):

\[
A = (1 + Te^{-R})^{-1}
\]  

The form of this equation and its possible physical significance will be discussed later. Each parameter of this function can be related to an element of the model. The value of A (the ratio of polymerase read-through to termination events at the terminator, i.e. the probability of replication) approaches its maximum value of 1.0 when T (defined as the ‘strength’ of the terminator) is small, and for a given value of T, the value of A increases as R (defined as the
concentration of free N proteins available to encapsidate a given concentration of nascent RNA chains) increases. With this definition, R will be a dimensionless ratio of concentrations. A dimensionless value for the terminator strength may also be given in terms of the ratio of read-through events to polymerase starts. The observation of 8 to 10% read-through \textit{in vitro} of the leader gene terminator (Herman & Lazzarini, 1981), and of 1 to 3% read-through \textit{in vivo} of the L→5' terminal extracistronic region (Schubert & Lazzarini, 1981), suggest that this ratio may be of the order of 0.01 to 0.1. The terminator strengths would be the inverse of this ratio; thus, T may take values of the order of 10 to 100. At the 50% read-through level in this range of terminator strengths (A = 0.5; shaded area of Fig. 2), R takes values of 2 to 4. This suggests that a low concentration of free intracellular genome and antigenome nucleocapsids may be sufficient to suppress the terminators and support viral replication.

The rate of synthesis of a plus-strand ND antigenome from its minus-strand genome template can now be written:

$$\frac{d}{dt} (ND^+) = P_\Theta (ND^-) A_\Theta$$

(2)

where $P_\Theta$ represents the promoter strength of the genome in units of polymerase starts per template per unit time, $(ND^-)$ is the genome template concentration, and $A_\Theta$ expresses its probability of being replicated. Likewise, the rate of ND genome synthesis from its antigenome template would be:

$$\frac{d}{dt} (ND^-) = P_\Theta (ND^+) A_\Theta$$

(3)

To test these equations with some experimental data, the rates of ND antigenome and genome synthesis can be compared as a ratio:

$$\frac{d(ND^+)}{d(ND^-)} = \frac{P_\Theta (ND^-) (1 + T_\Theta e^{-R})}{P_\Theta (ND^+) (1 + T_\Theta e^{-R})}$$

(4)

Soria \textit{et al.} (1974) and Simonsen \textit{et al.} (1979) have reported that late in a standard virus infection, about 80% of newly synthesized full-length genomic RNA is minus-strand. However, due to the selective maturation of only minus-strand genomes, the ratio of total intracellular ND genomes to antigenomes was found to remain at about 60/40. If it is assumed that $P_\Theta = P_\Theta$ because of the similarity of the putative 3' end genomic and antigenomic promoter sequences, then
Review: Analysis of interference in VSV

substitution of these values into equation (4) gives:

$$20/80 = 1.0 \times 60/40 \times (1 + T_{\Theta}e^{-R})/(1 + T_{\Theta}e^{-R}).$$

This equation reduces to $T_{\Theta} = P_{\Theta}T_{\Theta^*} + 5e^R$, a result consistent with the central assumption of this model, that the minus-strand genomic terminator is stronger than its plus-strand antigenomic counterpart. These assumptions do not exclude other possibilities, however. The promoter need not be confined to the exact 3’ end of the genome or antigenome, as pointed out by Isaac & Keene (1982). Furthermore, on a per template basis, 30 times as much minus-strand leader RNA as plus-strand leader RNA was found in mixed virus-infected cells (Leppert & Kolakofsky, 1980), suggesting that the antigenomic (and DI) promoter may be considerably stronger than its genomic counterpart. In line with this possibility, equation (4) also leads to $P_{\Theta} = 6P_{\Theta}$ if it is assumed instead that $T_{\Theta} = T_{\Theta^*}$. Thus, the concepts embodied in equation (4), which pertain only to ND replication, cannot decide between these mechanistic alternatives. The implications of the possible dependence on $R$ of the terminator strengths will be discussed later.

Another test of the model would be the derivation of a reciprocal exponential relationship like that of Bellett & Cooper (1959) by including terms for DI replication in the analysis. The rate equation for DI genome synthesis can be written:

$$\frac{d}{dt} (DI) = 2P_{\Theta}(DI)A_{\Theta}$$

The statistical factor of 2 reflects the fact that the plus-strand DIs are the templates for the minus-strand DIs and vice versa, and only plus-signs appear because all copyback DIs share the plus-strand antigenomic 3’ end sequence. Equation (5) may be expanded and rearranged to give:

$$\frac{d(DI)}{(DI)} = 2P_{\Theta}(1 + T_{\Theta}e^{-R})^{-1}dt$$

As long as $R$ is reasonably constant, an assumption which is justified by the homeostatic regulation of N protein synthesis, equation (6) may be directly integrated:

$$(DI) = (DI)_0 e^{Kt}, \text{ where } K = 2P_{\Theta}(1 + T_{\Theta}e^{-R})^{-1}$$

Comparing the rate of ND genome synthesis (equation 3) with that of DI genome synthesis (equation 6) as a ratio, gives:

$$\frac{d(ND^{-})}{d(DI)} = \frac{P_{\Theta}(ND^+)A_{\Theta}}{2P_{\Theta}(DI)A_{\Theta}} = \frac{(ND^+)}{2(DI)}$$

Recalling that $(ND^+)/((ND^{-}) = 40/60$ over the course of infection, substituting for $(DI)$ according to equation (7) and rearranging, gives a differential equation relating DI dose to ND virions:

$$\frac{d(ND^{-})}{(ND^{-})} = \frac{e^{-Kt}}{3(DI)_0}d(DI)$$

Assuming as above that $R$ is reasonably constant, integration gives:

$$\ln \frac{(ND^{-})}{(ND^{-})_0} = \frac{(DI)e^{-Kt}}{3(DI)_0}, \text{ where again } K = 2P_{\Theta}(1 + T_{\Theta}e^{-R})^{-1}$$

Under conditions of mixed virus infections where the input ratio of DIs to ND virus is varied, and samples are taken at a fixed time when the harvest is maximal, as in the study of Bellett & Cooper (1959), then equation (10) becomes the reciprocal exponential relationship described in Fig. 1 of these authors:

$$\ln \frac{(ND^{-})}{(ND^{-})_0} = \frac{\text{Constant}}{(DI)_0}$$

(11)
Inclusion in the mathematical analysis of a rate equation for DI replication thus leads to an analytic equation (11) consistent with experiment. This suggests that the replicative advantage of DIs over ND virions is due to control of the viral polymerase at a terminator, rather than to competition for this enzyme at the initiation of RNA synthesis.

**DISCUSSION**

The derivation from the model of a reciprocal exponential relationship between DI dose and ND virion progeny adds weight to the concept of the terminator as a functional element of VSV replication, even though the mathematics are greatly oversimplified. No explicit account is taken of the common observation that leader RNA synthesis proceeds at rates many times greater than either viral mRNA transcription or genome replication. This is acceptable provided that only polymerase starts which lead to transcription or replication are scored as such; the actual rate of polymerase starts will be many times the figure given below. Constant terms to account for the initial infecting virus particles have been omitted, for mathematical clarity. Since the key comparisons (equations 4 and 8) are formulated as ratios, the resulting errors will be minimal provided that the m.o.i. of ND and DI particles are modest (≤10) and near equal. Interestingly, both the mathematics and the experimental observations of Bellett & Cooper (1959) break down when the m.o.i. of ND or DI particles is too high. In so far as the differential rate equations (2), (3) and (5) reflect the properties of viral RNA synthesis, the reciprocal exponential relationship that arises naturally by integration reflects the properties of virus replication. The rate equations, in turn, depend on the switching function (1), which identifies the mutually exclusive pathways of transcription and replication as the all-or-none event deduced by Bellett & Cooper (1959).

One key operational element in the VSV replication model is the behaviour of the viral RNA polymerase. It freely engages its genome template at the single 3' entry promoter (Emerson, 1982) and is thus always 'on', like a pulsating electrical current, but can be 'switched' between two circuits at the terminator. This abrupt sort of switching is well described by the Hill (1910) equation, which was in fact derived from differential equations for the break of a steady electric current. The sigmoid form of equation (1), that of the Hill equation, was later found to describe many biological phenomena such as the binding of oxygen to haemoglobin. The analogies are striking between the rectangular hyperbola expected for polymerase competition at the initiation step of VSV replication and oxygen binding to the single haem of myoglobin, on the one hand, and the reciprocal exponential relationship derived from a Hill equation, reflecting regulation at a terminator, and oxygen binding to the haemoglobin tetramer, on the other (see inter alia, Fruton & Simmons, 1961). The sigmoid shape of this latter curve was interpreted to reflect cooperativity between the four haem groups (Pauling, 1935; for review, see Perutz, 1979), and similar curves reflecting cooperative interaction are often seen where stringent control is required. A good example is the phage T4 gene 32 protein, which functions in T4 DNA replication by binding cooperatively to ssDNA, and similarly controls its own synthesis by binding to its mRNA at an unstructured region that also functions as the ribosome binding site (von Hippel et al., 1982). Another is the 'hair trigger' control of λ phage lysogeny by λ repressor (Johnson et al., 1981), which enables the phage to switch to lytic growth in response to a transient signal, in contrast to the non-cooperative binding of the lac repressor to its operator. Evidence for cooperativity in the binding of N protein to VSV leader RNAs has already been presented (Blumberg et al., 1983). In general, whenever a constitutive synthetic process can be modulated between two alternative pathways, the products of the favoured pathway will become amplified. When, in addition, a reciprocal exponential relationship like that of Bellett & Cooper is observed, this modulation is likely to reflect the cooperative interaction of regulatory proteins.

Equation (10) also suggests why the concentration of free N protein may have a profound effect on the proportions of ND and DI progeny: the value of R (which is proportional to the concentration of N protein) is magnified twice, by its appearance as an exponent in the function K, which is itself an exponent. Huang & Wagner (1966) reported that the yield of infectious progeny virus in the culture supernatant of a standard infection was 1.6 × 10^7 p.f.u./ml at 4 h
post-infection, and $2.0 \times 10^6$ p.f.u./ml at 8 h post-infection. Assuming exponential replication of the virus during this period, the calculated first-order rate constant for replication is $1.8 \text{ h}^{-1}$, corresponding to a yield of 10000 progeny virus after 5.1 h of replication from a single cell infected with one standard virus. This number represents the product of the frequency of polymerase starts on the antigenome template (i.e. $P_\Theta$) and the probability $A_{\Theta} = (1 + T_{\Theta} e^{-R})^{-1}$ that a start will lead to a full-length genome. Thus, $P_\Theta = 1.8(1 + T_{\Theta} e^{-R}) \text{ h}^{-1}$. Substituting $T_{\Theta} = 10$, the presumed low end value for the strength of the easily suppressed antigenomic terminator, gives $P_\Theta$ a maximum value of 19 h$^{-1}$ when R is small. If only one polymerase at a time can engage the 46 nucleotide-long leader gene, and the average rate of chain elongation is 3.7 nucleotides/s (Iverson & Rose, 1981), then the polymerase start rate is 283 h$^{-1}$. The probability that a start will lead to replication in vivo is thus $A_{\Theta} = 19/283 = 0.067$, which corresponds to $R \leq 2$ (Fig. 2, cross-hatched area). The critical linkage between these numbers, inherent in the double exponential relationship between R and replication, predicts that R will be small at all levels of replication. Limited availability of N protein may be a means of maximizing the specificity of leader chain encapsidation in vivo. Once specifically nucleated, nucleocapsids will elongate in the presence of a low concentration of free N protein because assembly is highly cooperative (Blumberg et al., 1983). In this context, the observations of Grubman & Shafritz (1977) and of Rosen et al. (1982) may indicate that in vivo, N protein may sometimes encapsidate any RNA that contains an accessible nucleation site, such as the poly(A) tail of mRNA.

The recognition of the operative importance of the self-regulation of N protein synthesis was another key to the construction of the present model. A specific effect of the homeostatic control of free intracellular N protein concentration is to impose a limit on R in equation (1), and hence to limit the effects of variation of $T_{\Theta}$ with respect to $T_{\Theta}$. The dependence on R of the terminator strengths, which may reflect the cooperativity of N in encapsidation, clearly cannot be without limit. Mathematically, this effect also imposes closure on the integrals arising from equations (6) and (9), the values of which would otherwise be indeterminate and the derivation of equation (10) impossible. However, the model itself is also clearly oversimplified. It does not specify what sequences may be involved in the termination signal, nor does it take explicit account of the possible roles of host cell factors or of viral proteins other than the N protein. In the context of the terminator, notions of the viral polymerase may take on new aspects. The PolR1 strains of VSV described by Perrault et al. (1981), in which the polymerase reads through the terminator at an abnormally high frequency, is thought to owe this property to a mutated N protein.

CONCLUDING REMARKS

In this model, we have attempted to assemble into a coherent framework some salient features of defective infections of VSV, such as temporal regulation and the reciprocal exponential relationship of Bellett & Cooper (1959). One key element in the construction of the model was the homeostatic VSV N protein, which serves as the feedback link in VSV amplification. Other properties of such a self-regulatory protein can then be linked to the coordinate control of other genes of the virus. Another key element was the switching of the constitutively functioning viral RNA polymerase at a terminator. The identity of the terminator is not specified in the model. However, the work of Giorgi et al. (1983) suggests that minus-strand leader RNAs may contain at positions 22 to 33 from their 5' end the sequence $\text{AAAAAAAYAAAA}$ which is not present in plus-strand leader RNAs. Interestingly, the genomic complement of this sequence is the region of contact with the viral NS protein (Isaac & Keene, 1982). Since the N protein nucleation signal appears to involve an A residue in every third position (Blumberg et al., 1983), this sequence may act as a second nucleation site in nascent minus-strand RNA chains. It is tempting to speculate that enhanced cooperativity of encapsidation at this interior site may help 'push' the polymerase through the terminator (T+) of the template, making it more easily 'suppressible'.

Much praise is due Rosette Bandelier and Colette Pasquier for technical excellence in a series of experiments (unpublished) which could only be interpreted through this analysis. We thank Bob Hipskind and Laurent Roux, and D. W. Kingsbury, R. A. Lazzarini and M. E. Reichmann for helpful commentary. This work was supported by grants from the Swiss National Science Foundation.
REFERENCES


