RNA (1996), 2:1046–1057. Cambridge University Press. Printed in the USA.

Promoter elements in the influenza vRNA terminal structure*

RAMON FLICK, GABRIELE NEUMANN,[1] ERICH HOFFMANN, ELISABETH NEUMEIER, and GERD HOBOM

Institut für Mikrobiologie und Molekularbiologie, Frankfurter Str. 107, 35392 Giessen, Germany

ABSTRACT

The role of the partially double-stranded influenza vRNA terminal structure and its constitutive elements as a promoter signal was studied in vivo in a series of nucleotide substitution and insertion derivatives. A series of single and complementary double exchanges restoring intrastrand base pairing shows that a distal promoter element consists of a six-base pair double-stranded RNA rod in long-range complementary interaction. Within the distal element, all base pair positions are freely exchangeable, and hence no nucleotide-specific recognition could be identified. The proximal promoter element consists of nine partially complementary nucleotides at the vRNA 5′ and 3′ end. The nine plus six base pair panhandle rod of protein-free vRNA is interrupted by a central third element, a single unpaired nucleotide: adenosine 10 or various substitute residues, which appears to cause a bulged conformation in the overall structure. Mutagenization studies in the promoter proximal element indicate that, upon binding to polymerase, nucleotides at positions 2 and 3 interact with positions 9 and 8 within each branch (5′ or 3′) in short-range base pairing. In this conformation, the intermediate positions 4–7 are exposed as a single-stranded tetra-loop, which includes invariant guanosine residue 5 in the top conformational position of the 5′ segment loop. Altogether, the three base paired segments in angular conjunction to each other adopt a conformation that is described in a ''corkscrew model'' for an activated stage of vRNA/polymerase interaction.

Keywords: conformation; corkscrew model; panhandle structure; RNA polymerase I; tetra-loop; viral polymerase

INTRODUCTION

Double-stranded RNA panhandle structures have been proposed to involve both ends of the otherwise single-stranded eight viral RNA (vRNA) molecules that collectively constitute the genome of influenza viruses. This structural interpretation relates to 16 and 15 terminal nucleotides, predominantly complementary to each other at both ends of such molecules, of which 13 and 12 residues are almost always identical among all eight vRNA segments (Hsu et al., 1987; Stoeckle et al., 1987; Yamanaka et al., 1991). Viral RNA polymerase has been observed to be associated with that terminal structure (Honda et al., 1988; Fodor et al., 1993), whereas the major internal, single-stranded region of each vRNA molecule is known to be covered by NP subunits in cooperative interaction (Martin & Helenius, 1991). A different kind of panhandle structure has also been observed in vitro for internally shortened vRNA molecules in the absence of proteins (Baudin et al., 1994), and various such structures have been discussed with regard to their potential significance for transcription and replication of orthomyxoviral RNAs.

Although initially the vRNA promoter signal had been assumed to reside exclusively in the 3′ end of each segment (Li & Palese, 1992; Seong & Brownlee, 1992), a strong association of viral polymerase, also with the 5′ end of each segment, changed that pattern (Hagen et al., 1994). Accordingly, both ends in complementary, partially double-stranded conformation have been concluded to be required for initiation of transcription

Reprint requests to: Gerd Hobom, Institut für Mikrobiologie und Molekularbiologie, Frankfurter Str. 107, 35392 Giessen, Germany; e-mail:gerd.hobom@mikro.bio.uni-giessen.de.

* We dedicate this article to Professor R. Rott on the occasion of his seventieth birthday.

[1] Present address: Institute of Virology and Molecular Biology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA.

as a result of mutagenization studies in vitro and in vivo (Fodor et al., 1994, 1995; Tiley et al., 1994; Neumann & Hobom, 1995). Various structural elements or sequence positions have been proposed to be involved directly in RNA–protein interaction with viral polymerase. Among these are positions $\bar{9}$–$\bar{11}$ (numbers marked in this way are used in referring to positions counted from the 3′ end) as concluded by Fodor et al. (1993, 1995), or 5′ position 5, as demonstrated in vivo (Neumann & Hobom, 1995). However, several experimental results still remain in contradiction to each other.

The vRNA promoter structure serves in initiation of two different reactions of RNA synthesis. Early in infection, mRNA is synthesized in a specific cap-snatching mechanism (Krug et al., 1989) and, later in infection, full-length complementary RNA (cRNA) is transcribed, probably dependent on the absence or presence of newly synthesized NP protein. The plus-strand cRNA molecule with the potential to form a double-stranded structure of its own, similar but not identical to the corresponding vRNA, is then used as a template for progeny minus-strand vRNA synthesis. Inherent asymmetries between both vRNA and cRNA promoter structures may result in a different recognition by RNA polymerase and in asymmetric pool sizes for cRNA versus vRNA molecules (around 1:10; Mukaigawa et al., 1991). The viral RNA polymerase complex consisting of proteins PB1, PB2, and PA catalyzes all three modes of influenza RNA synthesis following its specific binding to the terminal structures of all influenza vRNAs and cRNAs. Its PB2 subunit also recruits capped cellular primer RNAs via cellular pre-mRNA cleavage for initiation of viral mRNA synthesis (Cianci et al., 1995; Shi et al., 1995).

RNA polymerase I transcription of appropriately constructed influenza cDNA templates was developed in this laboratory for in vivo transcription of vRNA-like molecules. DNA transfection and infection with influenza virus resulted in viral progeny carrying the foreign vRNA segment, which could be passaged into next generations of influenza viruses (Zobel et al., 1993; Neumann et al., 1994). For the analysis of the promoter function, the influenza gene reading frame was replaced by the chloramphenicol acetyltransferase (CAT) gene. Among a first series of single- and multiple-point mutations, several promoter-up mutations were described, resulting in up to 20-fold enhancement of CAT expression from such RNAs, and in their accumulation during passaging (Neumann & Hobom, 1995). This result prompted us to initiate a more extended series of promoter variants in order to investigate the requirement for various elements of that partially double-stranded structure. Our results suggest a new model for the polymerase-interactive vRNA conformation, and point out invariable nucleotide residues that are likely to represent positions of direct RNA–protein interactions.

RESULTS

Nucleotide A10 constitutes a flexible joint within an angular structure of two rigid elements

For the in vivo analysis of influenza vRNA promoter elements, the method of RNA polymerase I transcription was used for synthesis of vRNA molecules, which contained the CAT coding sequence inserted exactly between complete influenza terminal untranslated sequences. In these cDNA constructs, CAT reading frames are transcribed in antisense, i.e., in vRNA orientation, covalently linked to serial variants of the 34 and 29 nt of flanking promoter-containing sequences, as present in segment 4. Two strong promoter-up mutation derivatives, pHL1104 and pHL1102, have been employed as references throughout these studies rather than wild-type vRNA. pHL1104 cDNA carries three point mutations in the vRNA 3′-terminal region, resulting in G$\bar{3}$ A, U$\bar{5}$ C, and C$\bar{8}$ U transitions, which were designed to create a sequence exactly complementary to the 5′ vRNA sequence over nine 5′- and 3′-terminal positions (Fig. 4A). In the related double mutant pHL1102, position U$\bar{5}$ is retained in wild-type configuration. A more than 20- or more than 15-fold increase of CAT activities over wild-type level, respectively, allows for a more sensitive differentiation of intermediate levels of activity, whereas both mutant structures are close enough so that functional data may be extrapolated back to wild-type influenza vRNA.

A working model for the influenza vRNA promoter structure consists of three distinct elements. The proximal and the distal rigid RNA segments include up to nine base pairs (positions 1–9 and $\bar{1}$–$\bar{9}$ from the 5′ and 3′ ends) and six base pairs (on average; positions 11–16 and $\bar{10}$–$\bar{15}$) that are separated by a single residue, adenosine 10, in an unpaired position without any complementary nucleotide in the 3′ segment. A10 may be expected to cause an angle or kink between the two rigid RNA segments, and may be present either in an intrahelical position due to stacking interactions (with A11 or G9) or may be exposed externally above a continuous rod extending over both of these segments (Wimberly et al., 1993). Accordingly, A10 might either be recognized indirectly by causing a bulge or an overall angular shape of a two-rod structure. Alternatively, it might be recognized in an exposed position directly and specifically by viral RNA polymerase in its vRNA-specific binding (as is true for several other protein/RNA recognition reactions; Wu & Uhlenbeck, 1987; Riordan et al., 1992).

Single nucleotide substitutions at position A10 left CAT expression rates unaffected when A10 was exchanged into guanosine 10, and caused only moderate reductions (to around 50%) when A10 was converted into a C10 or U10 residue, both in DNA-transfected plus infected B82 cells, or after viral passage in MDCK

cells (see Fig. 1). Due to a situation of A10 and A11 both facing a single uridine residue, $\overline{10}$, in the 3′ segment, it is not entirely certain whether U$\overline{10}$ is involved in base pairing with A11, and A10 stays in the unpaired, angular position, or vice versa. Therefore, also A11 has been exchanged in a series of nucleotide substitutions, but all of those resulted in a loss of activity or, at most, in spurious expression rates. Activity was regained, however, upon introduction of secondary substitutions in position U$\overline{10}$, if designed to restore the potential for base pair formation (e.g., C11:G$\overline{10}$ in pHL1282, not shown). From these results, we conclude that it is A10 that is located in the unpaired, angular position, as the panhandle model is normally drawn, whereas A11 and U$\overline{10}$ constitute the first base pair of the distal dsRNA segment. The latter argument is in agreement with further experimental data (see below).

The data shown in Figure 1 support the conclusion that the angular structural element in the promoter structure, because of its variability, is recognized indirectly. Complete inactivity, however, resulted upon deletion of that position (pHL1152; see Neumann & Hobom, 1995).

Some degree of promoter activity was observed also for several constructs that deviate further (in one way or another) from the preferred situation of a single unpaired residue in position 10. These include: insertion of a second unpaired adenosine (but not of a guanosine) residue next to A10, or insertion of a uridine residue opposite A10, in the 3′ segment. An inserted U residue opposite A10, with base pairing potential for that angular residue, still retained a small degree of CAT expression or polymerase binding activity, as was shown in vivo (pHL1140: 10% relative activity; Neumann & Hobom, 1995) and in vitro (compare Fig. 7 in Tiley et al., 1994). Also, upon insertion of two uridine residues next to each other and opposite A10, a very low rate of CAT expression was retained (pHL1354, not shown). Insertion of an extra cytosine residue opposite G10 (substitute for A10), with potential to form a G-C base pair in this irregular position, destroyed the promoter activity entirely, however, and several other related constructs also remained inactive.

In summary, our studies in the A10 region indicate that a flexible joint has to be present in this location. In interacting with viral RNA polymerase, this flexible joint may facilitate an angular conformation within an otherwise rigid RNA double-helical structure (for pHL1140 and similar constructs via forced bending and squeezing out the extra uridine residue), so that a dual interaction of its two constitutive elements with the enzyme becomes possible. The role of the A10 unpaired residue, however, is indirect, and can be fulfilled by several other related structures, at least within a pHL1104 vRNA framework, which potentially may be more rigid than wild-type vRNA.

The distal element of the vRNA promoter structure is dsRNA with no sequence specificity

All influenza A vRNA molecules contain a perfect series of mostly six complementary sequence positions:

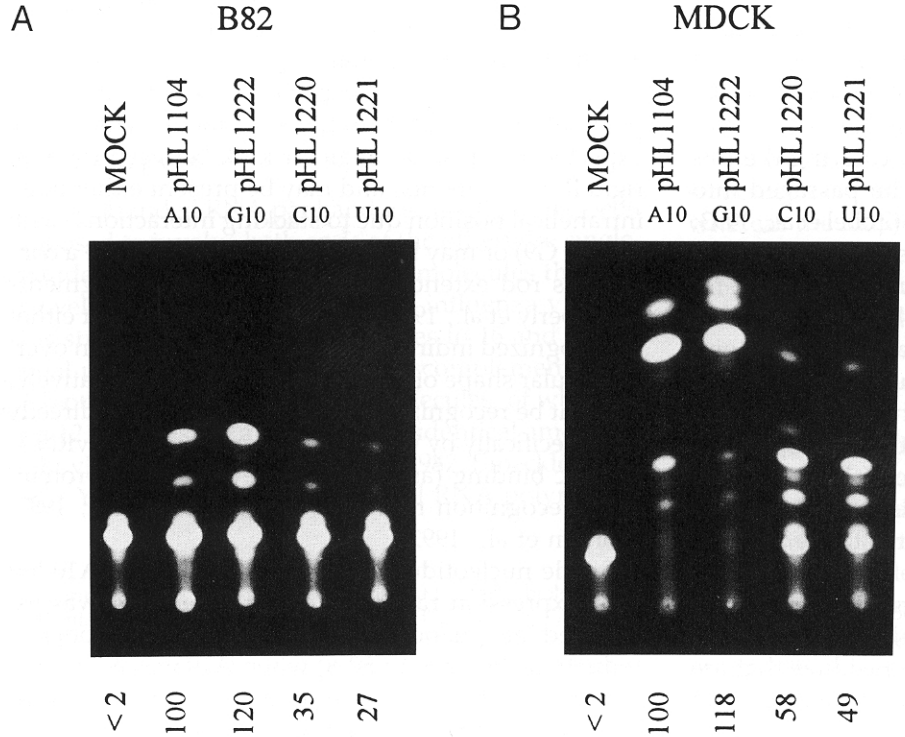

FIGURE 1. Nucleotide substitutions in angular position A10 of vRNA. CAT activities of cell lysates were analyzed as described in the Materials and methods. Acetylated reaction products were separated by chromatography. Nucleotide substitutions in position 10 from the 5′ end are indicated above each lane, relative to parental pHL1104 vRNA. CAT activities relative to standard pHL1104 are indicated below each lane and represent the average of three or more experiments. **A:** B82 cells were transfected with recombinant DNA and infected with influenza virus. Cell lysates were prepared 8 h post infection and analyzed for CAT activity as described above. **B:** MDCK cells were infected with recombinant influenza virus, i.e., supernatants obtained at 8 h post infection from B82 cells, treated as described above. CAT activity of cell lysates was determined at 8 h post infection.

11–16 opposite $\overline{10}$–$\overline{15}$. The same is true for other orthomyxoviruses (Stoeckle et al., 1987; Lee & Seong, 1996). Sequence diversity is directly observed for base pairs 14–16 versus $\overline{13}$–$\overline{15}$, because these are different among the various influenza segments, even if always adhering to the base pairing principle. In a systematic approach, the three proximal base pairs of this segment were studied in vivo one by one in single- and two-step complementary double exchanges. In every case, a compensating double exchange that restores base pair formation according to a double-strand model resulted in full promoter activity, even though a different base pair than originally present was finally inserted in such two-step transition *or* transversion approaches (Fig. 2B). Single-step intermediates with a single mismatch position, however, were either entirely inactive or instead only marginally reduced in activity (Fig. 2C). The latter result was observed for all G:U-carrying vRNA mismatch constructs in this series, such as pHL1276, and is consistent with the known potential for forming a low-energy base pair, in particular in a context of preserved base pairing to its left and right; it was not observed for any of the A:C mismatch vRNA constructs. Although that series of experiments gives direct proof for the absence of nucleotide-specific interactions with RNA polymerase in the distal structural segment, it also proves the absolute requirement for a double-stranded structure in this region of the vRNA promoter. Equivalently, insertion of an extra base pair (A:U) proximal to A11:U$\overline{10}$ resulted in full promoter activity, although that insertion shifts and rotates all consecutive base pairs of the distal element (data not shown). Similar proof for base pairing in this region was obtained in in vitro experiments (Fodor et al., 1994; Tiley et al., 1994).

Patterns of single-nucleotide substitutions in the proximal vRNA promoter element

The vRNA promoter was shown earlier to support a substantially increased level of reporter gene expression rates upon introduction of three nucleotide substitutions, which were designed to result in an uninterrupted, perfectly complementary structure in the proximal element. In pHL1104, the proximal RNA segment therefore potentially consists of nine instead of the regular six base pairs plus three mismatch positions (two of them being G:U pairs, however) as present in wild-type vRNA; in the related double mutant pHL1102, one of the wild-type vRNA G:U pairs is retained (Neumann & Hobom, 1995). The results achieved appear to argue for a role of the double-stranded RNA structure also in the proximal promoter segment, at least during one step of interaction with polymerase. However, unlike before, in the distal RNA segment, construction of an initial similar set of single and complementary double exchanges did not lead to unequivocal results in this region.

We therefore decided to work out a complete pattern of, altogether, 53 single-nucleotide substitutions in the proximal element of the pHL1104 influenza promoter structure covering both the terminal 5′ and 3′ regions, i.e., residues A1–C9 and U$\overline{1}$–G$\overline{9}$ (and in another incomplete series also for pHL1102), to determine their influence on CAT reporter gene expression rates in vivo. Data for positions 4, 5, and 6 of the 5′ end are shown as an example in Figure 3. The entire pattern of the pHL1104 results is tabulated in Figure 4A. The inherent structure of these data is most evident from the 3′ branch in the lower part of the figure: central positions $\overline{4}$–$\overline{6}$ are freely exchangeable, whereas external positions $\overline{2}$, $\overline{3}$, and $\overline{7}$, $\overline{8}$, $\overline{9}$ do not tolerate any *single* nucleotide deviation from the original pHL1104 sequence without complete or at least severe loss of activity. The same pattern can also be discerned from the 5′ branch in the upper part, even if complicated by several additional features. This concerns an absolute requirement for G5 (Neumann & Hobom, 1995), which is unparalleled in the 3′ branch and, therefore, appears to reflect a nucleotide-specific interaction with viral polymerase in a locally single-stranded conformation. In part, this argument may also hold for positions A4 and A7, because several single-nucleotide exchanges lead to moderate or severe reduction in promoter activity. In the external positions, no transversions, but several transitions, such as A8G or C9U—indicative of the underlying secondary structure—appear to be tolerated here, even if accompanied by a partial loss of activity (see below).

Patterns of complementary dual substitutions in the proximal element are explained by a "corkscrew model"

With regard to influenza wild-type sequence, both constructs pHL1102 and pHL1104 constitute *double* mutants exactly in those positions $\overline{3}$ and $\overline{8}$, which will tolerate no single substitutions. pHL1148 constitutes a quadruple mutant in sensitive positions $\overline{3}$, $\overline{8}$, 3, and 8. All three constructs are highly active, however, far above wild-type level (Neumann & Hobom, 1995). This paradox can be resolved on the basis of a new structural interpretation for the vRNA structure in its mode of binding to viral RNA polymerase, the corkscrew model. This model, as shown in Figure 4B, depicts the double and quadruple mutations mentioned above as *complementary* double exchanges able to reconstitute short-range intrastrand base pairs $\overline{3}$:$\overline{8}$ and 3:8. Similarly, the single-nucleotide transitions mentioned above, which may be tolerated and lead to partially reduced, but certainly not abolished promoter activity, are reinterpreted with respect to this model to constitute G:U (instead of G:C or A:U) base pairs in

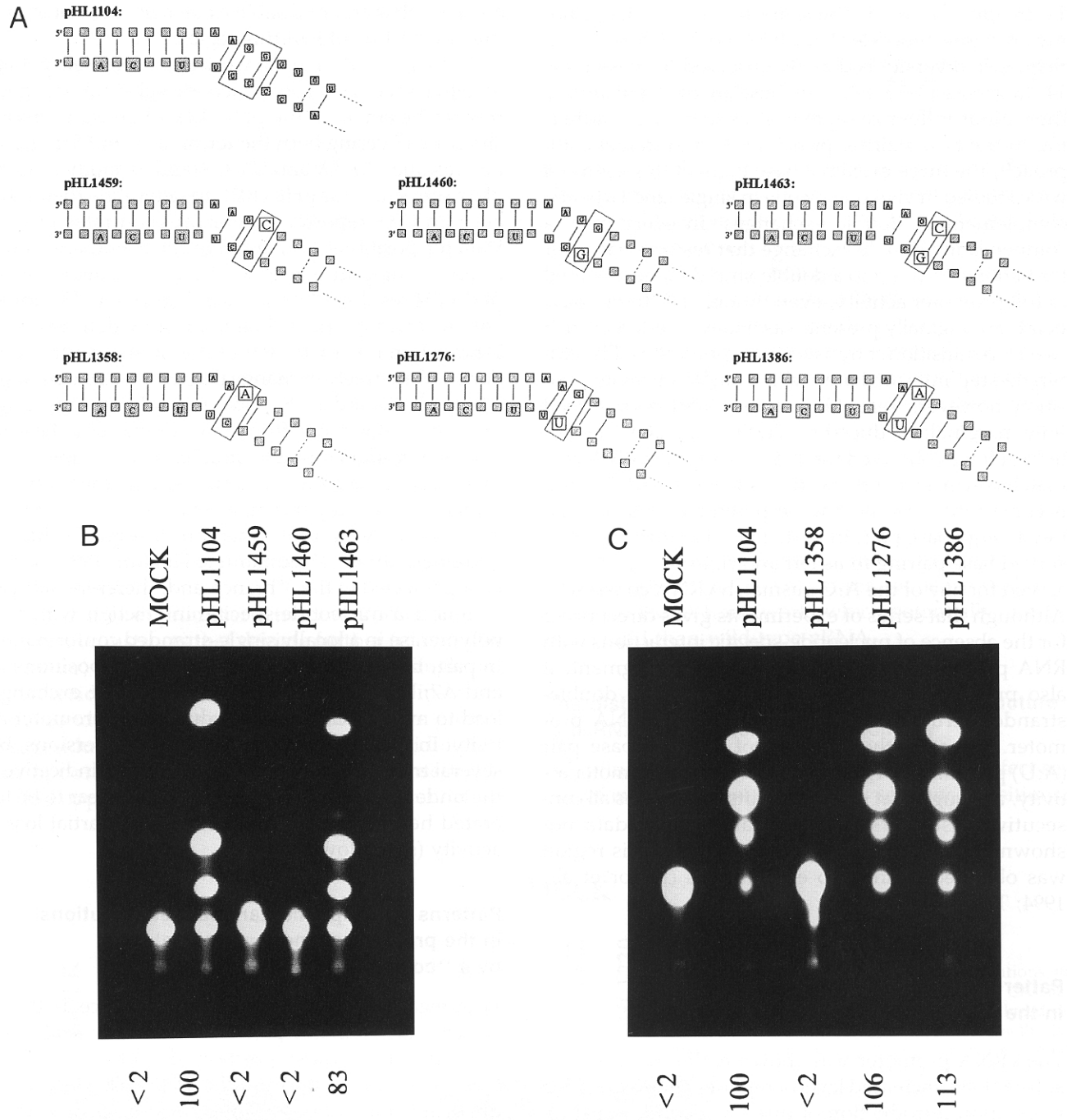

FIGURE 2. Single and complementary double substitutions at the second and third base pair of the distal promoter element. **A:** Molecular constitutions of distal element base pair 3 and base pair 2 variants in comparison to pHL1104, which is used as the reference construct throughout. pHL1104 deviates at three 3′ positions (larger grey boxes) from the vRNA wild-type sequence, such that a maximum number of panhandle base pairs is realized in the proximal promoter element. Specific mutations introduced for the analysis of the distal promoter element are indicated in larger white boxes. **B:** CAT activities of MDCK cell lysates infected with base pair 3 variants. Scaling of CAT expression relates to pHL1104 and mock reference data. **C:** CAT activities of MDCK cell lysates obtained after infection with base pair 2 variants.

corkscrew paired positions 3:8 or 2:9, whereas A:C mismatch constellations in the same positions result in minimal or no activity.

The corkscrew model in Figure 4B suggested additional experiments of complementary double exchanges with regard to the predicted short-range intrastrand base pairing positions in both branches of the proximal element. Several of these are shown in Figures 5 and 6. The results of these and other experiments for both pHL1104 and pHL1102 basic variants can be summa-

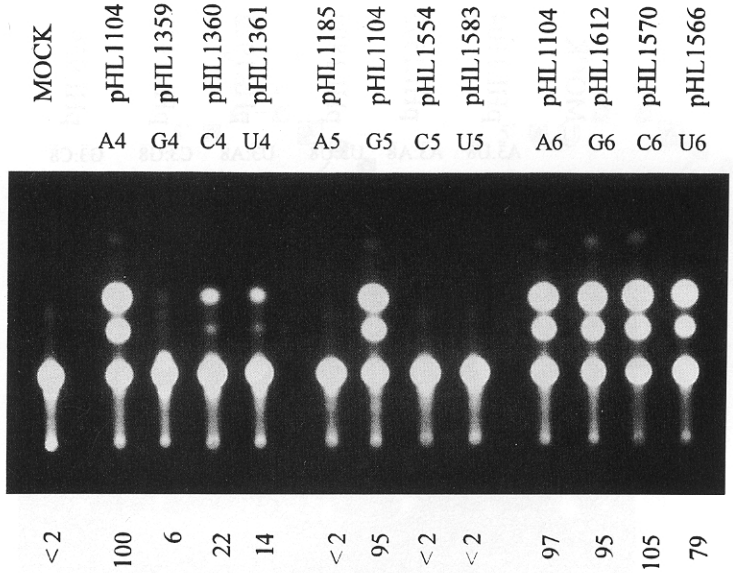

FIGURE 3. Comparative analysis of single-nucleotide substitution derivates for positions 4, 5, and 6 in the 5′ branch of the proximal element. Positions and type of nucleotide substitutions as compared to pHL1104 are indicated above each lane. For the structure of pHL1104, compare Figure 4A. CAT activities of MDCK cell lysates were analyzed at 8 h post infection, i.e., after passage of supernatants from DNA-transfected and influenza virus-infected B82 cells. CAT expression data relative to pHL1104 are indicated below each lane, and these data are also used for tabulating expression rates in Figure 4A, accordingly.

rized as follows: short-range base pairing is crucial for both, but is more important for the 3′ branch than for the 5′ branch, which appears to be stabilized in addition by single-strand-specific interactions with polymerase, particularly evident for G5, but also valid for A7 (or A7/C7) (Fig. 4A). In the 5′ branch, single G:U mismatches in either base pair (2:9 or 3:8) do yield functional promoters, even if partially reduced in their activity. In the 3′ branch, surprisingly, any other base pairing constellation in positions $\bar{3}$ and $\bar{8}$ was superior

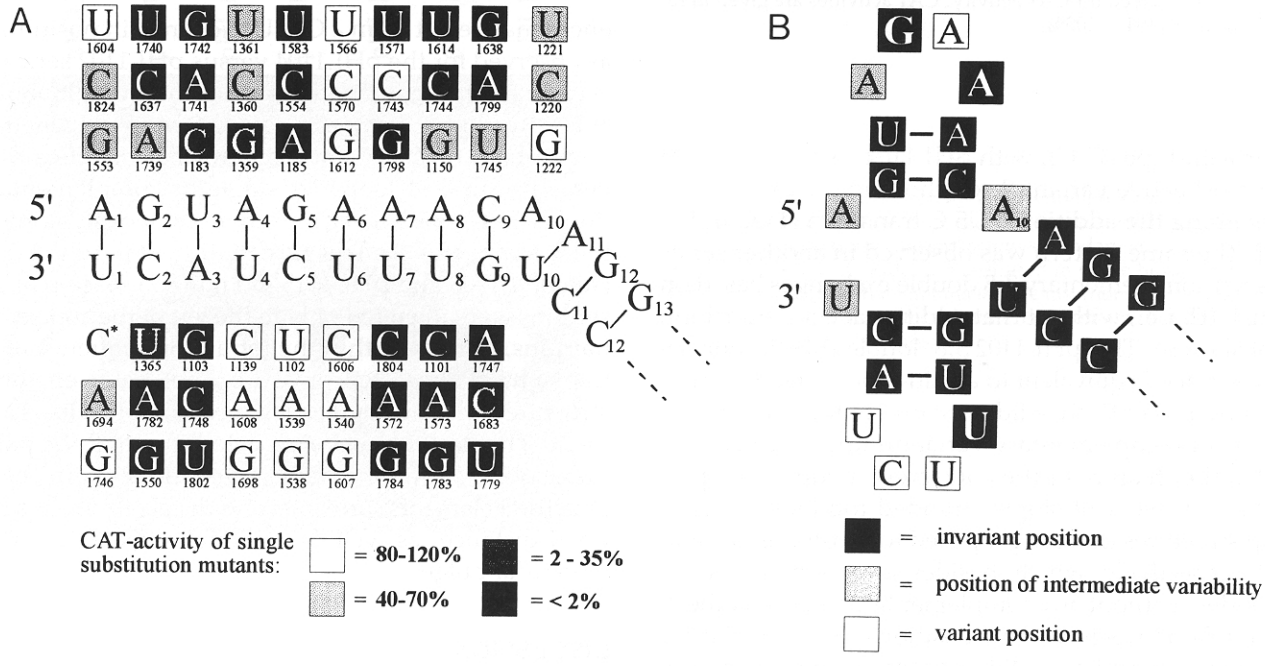

FIGURE 4. Summary of CAT expression rates for single-nucleotide substitution derivatives in the proximal promoter element of pHL1104 vRNA. **A:** Nucleotide sequence of pHL1104 vRNA is shown in the central two lanes in complementary panhandle conformation for its 5′ and 3′ ends. CAT expression rates (average of three or more experiments each) have been tabulated above and below that sequence for a complete set of 56 single-substitution mutants in the 5′ and 3′ branch, respectively. Tabulation of CAT activities reflects a scaling relative to pHL1104 CAT expression rates as explained below the figure; for representative experimental data, compare Figure 1 (A10) and Figure 3 (A4, G5, A6). *Due to the mechanism of RNA polymerase I transcription termination, the substitution U$\bar{1}$C cannot be analyzed (Zobel et al., 1993). **B:** Proposed secondary structure of pHL1104 vRNA according to the corkscrew model, as suggested by the patterns of single-substitution data in A, and supported by complementary double-substitution results as shown in Figures 5 and 6.

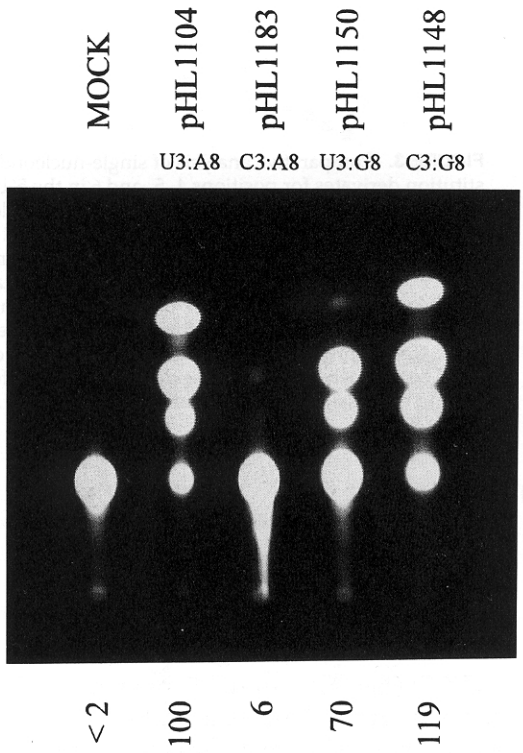

FIGURE 5. Single and complementary double substitutions for proposed corkscrew base pair 3:8. Nucleotide compositions in positions 3 and 8 in the 5′ branch of the proximal promoter element are indicated above each lane. Cell lysates were prepared 8 h post infection from MDCK cells that had been infected with recombinant influenza A virus and assayed for CAT activity. CAT activities are given in relation to pHL1104 = 100%.

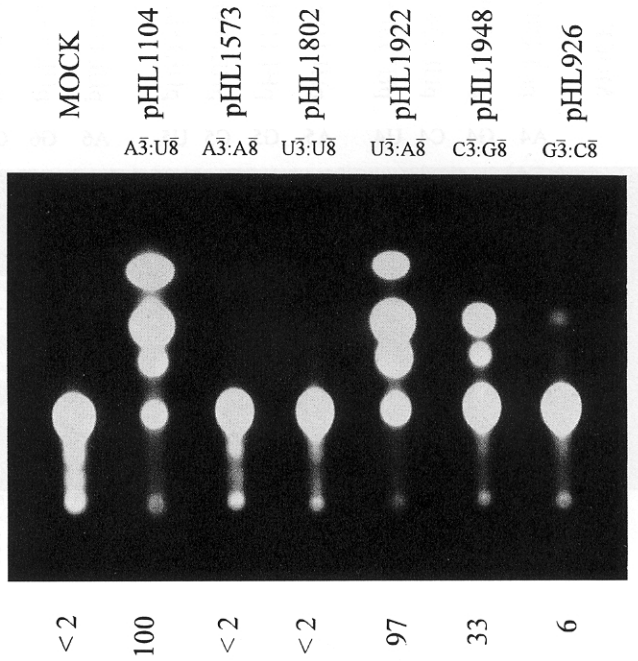

FIGURE 6. Single and complementary double substitutions for proposed corkscrew base pair 3̄:8̄. Nucleotide composition in positions 3̄ and 8̄ in the 3′ branch of the proximal promoter element are indicated above each lane. Cell lysates were prepared 8 h post infection from MDCK cells that had been infected with recombinant influenza A virus and assayed for CAT activity.

over wild-type (G:C), with pHL1104 (A:U) remaining the most active variant throughout in the main series containing the additional U5̄ C transition (Fig. 6). Exactly the same pattern was observed in another series of such complementary 3̄:8̄ double exchanges based on pHL1102, i.e., without that additional U5̄ C transition (not shown). This pHL1102 3̄:8̄ double substitution series is indeed equivalent to a wild-type variation series, because it is entirely wild-type except for 3̄:8̄, i.e., the points of complementary variation under study.

Another feature of the corkscrew model is the prediction of two 4-nt single-stranded top-loop or tetra-loop structures that are proposed for positions 4–7 and 4̄–7̄, respectively, which, besides several invariant nucleotides (without any complementarity to each other), also include freely variant positions, in particular for residues A6 and U6̄. With regard to the results obtained from single-nucleotide exchanges as tabulated in Figure 4A, major specific nucleotide/protein interactions besides G5 might be expected for positions 4 and 7, as well as 7̄ in the 3′ branch, if not resulting from inherent tetra-loop constraints. These nucleotides are supposed to be exposed in single-stranded tetra-loop conformation according to that model. Minor influences may result from a C ↔ U exchange in position 5̄, as observed for the pHL1104 versus pHL1102 series of secondary mutagenizations (Neumann & Hobom, 1995). In another approach, we wanted to test whether one or both of these 4-nt loops would tolerate a conversion into 5-nt loops by single or complementary double-nucleotide insertions, simultaneously done in the 5′ and 3′ branch. This turned out not to be the case (see pHL1708 and pHL1615 in Figure 7). Base pair insertions were tolerated only to the left of the corkscrew hairpins, but not within the tetra-loop sequence or to the right of these hairpins, i.e., not between these structures and the angular A10 position plus distal element. This result may be expected if viral RNA polymerase is simultaneously interacting with both structural elements, proximal and distal of A10, at a defined distance, as well as torsional angle across that structural bulge.

DISCUSSION

A vRNA panhandle structure has been introduced earlier as a general model for influenza viral RNA molecules. It proposes a partially double-stranded conformation involving both ends of an otherwise single-stranded RNA segment. This structure was suggested after terminally double-stranded RNP structures had been observed in an electron microscopy study on in-

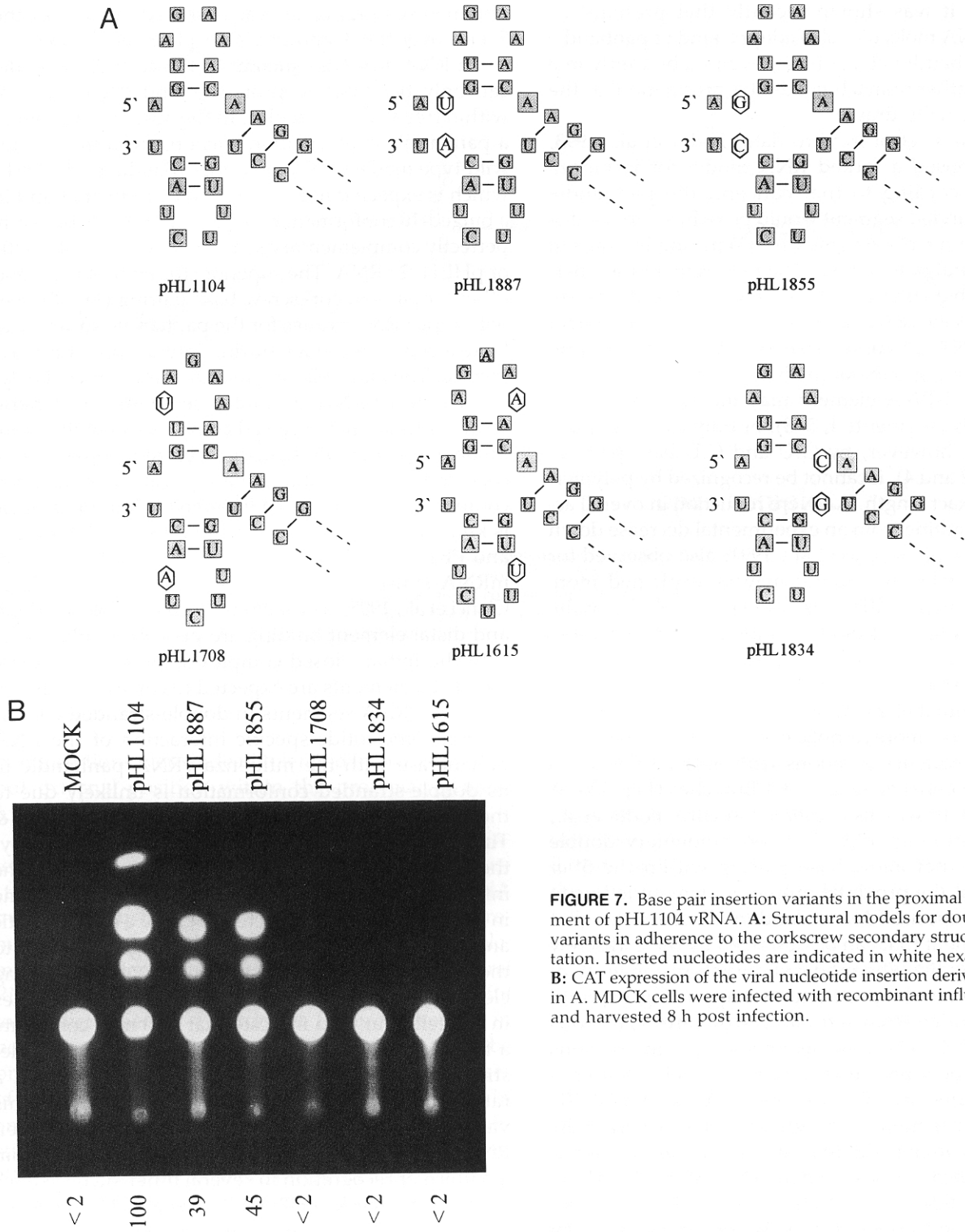

FIGURE 7. Base pair insertion variants in the proximal promoter element of pHL1104 vRNA. **A:** Structural models for double insertion variants in adherence to the corkscrew secondary structure interpretation. Inserted nucleotides are indicated in white hexagonal boxes. **B:** CAT expression of the viral nucleotide insertion derivatives shown in A. MDCK cells were infected with recombinant influenza viruses and harvested 8 h post infection.

fluenza RNA segments in virions (Hsu et al., 1987). The detailed conformation of the influenza vRNA terminal structure was analyzed in several in vitro studies for their function as a viral promoter signal (Li & Palese, 1992; Seong & Brownlee, 1992; Piccone et al., 1993; Fodor et al., 1994, 1995; Tiley et al. 1994), with partially contradictory results. Following the development of techniques for site-directed mutagenesis of influenza viruses (Luytjes et al., 1989; Zobel et al., 1993; Neumann et al., 1994), such questions now can be studied more appropriately in vivo. The proteins of the viral RNP complex (PB1, PB2, PA, and NP) undoubtedly contribute to the stability of the vRNA conformation both in vivo and upon reconstitution in vitro.

However, it was shown recently that protein-free model vRNA molecules also adopt a kind of panhandle structure (Baudin et al., 1994), even if aberrantly in a bulged-4 rather than a bulged-10 conformation, as the model is usually drawn.

Based on a set of in vitro data, Fodor et al. (1993, 1994) proposed a forked RNA model for influenza vRNA. According to that concept, the panhandle double-stranded segment would be reduced to the distal element (nt 11–16 versus $\bar{10}$–$\bar{15}$) during interaction with the viral polymerase. The requirement for a perfectly double-stranded RNA region without any sequence specificity was confirmed for the distal element of the vRNA promoter structure via compensating double exchanges in our in vivo experiments. Because that distal dsRNA element may include seven (segment 3), six (segments 1, 5–8), or only five base pairs (extended, however, by G:U and C:U base pairing; segments 2 and 4), it cannot be recognized by polymerase in its exact length. A severe reduction in overall activity as resulting from an experimental decrease down to four base pairs (Luo et al., 1991; also observed for pHL1349, data not shown) may be explained more readily by the resulting loss in overall stability of the remaining 4-bp RNA double strand rather than by specific recognition of the length of that helix.

An extensive series of single-nucleotide substitutions in the proximal element of the vRNA promoter structure revealed a more complicated pattern of variant and invariant sequence positions with only partial correspondence between its 5′ and 3′ branches (Fig. 4A). A similar pattern was also observed in vitro (Fodor et al., 1995). Short-range (2:9, 3:8) complementary double exchanges that allow base pairing within the 5′ *or* 3′ branch of the proximal promoter element led us to propose a special RNA secondary structure for that vRNA element that can be described by the corkscrew model (Fig. 4B). The model shows short-range, two-base pair stem-loop structures supporting exposed single-stranded RNA tetra-loop structures (positions 4–7, and $\bar{7}$–$\bar{4}$) where nucleotide-specific interactions with viral polymerase may occur, possibly within a protein cavity. By exposing an AGAA and a UUCU (or UUUU) sequence, the vRNA 5′ and 3′ corkscrew tetra-loops do not adhere to any of the standard tetra-loop sequences, as observed in rRNA (Wyatt & Tinoco, 1993), MS2 RNA so far (Valegård et al., 1994), or ribozyme RNA so far (Pley et al., 1994). In general structural agreement with such tetra-loop elements, however, the vRNA 5′ branch contains a major invariant and probably exposed nucleotide residue: G5 at position 2 within the tetra-loop conformation, whereas positions 1 and 4 (A4 and A7), which are supposed to sterically interact with each other in specific hydrogen bonding constellations, show limited degrees of variability. The 3′ branch contains only one major invariant tetra-loop position, U$\bar{7}$. This difference may relate to a tighter polymerase binding, as was observed in vitro for the 5′ end over the 3′ end of vRNA (Tiley et al., 1994).

An RNA structure such as in the corkscrew model can only be stable in protein-bound RNA, but not within free RNA molecules. In the absence of protein, a panhandle conformation was observed in vitro for wild-type model vRNA molecules (Baudin et al., 1994), which is expected to be stabilized even further and in a bulged-10 conformation, in case of perfectly or nearly perfectly complementary sequences, such as pHL1104 or pHL1102 vRNA. The superior efficiency of these and other variants in corkscrew base pairing (Fig. 6) over wild-type vRNA argues for the panhandle structure to be realized at one stage during interaction with polymerase. The initial binding reaction between viral polymerase and vRNA promoter segment—in general terms, formation of a closed complex—may therefore involve a panhandle structure before it converts to a corkscrew conformation, the equivalent of an open complex structure (Fig. 8). Influenza viral polymerase is known to consist of three subunits in a 1:1:1 ratio, and PB2 was observed to be involved in initiation of mRNA synthesis (and possibly in 3′ end binding; Cianci et al., 1995). The subunit(s) responsible for 5′ end and distal element binding are presently unknown.

In the initial, closed complex state, both proximal and distal elements are expected to consist of complementary vRNA segments in double-stranded conformation. Nucleotide-specific interaction of the RNA polymerase with the influenza vRNA panhandle in its double-stranded conformation is unlikely due to the narrow major groove in A-form RNAs (Wyatt & Tinoco, 1993). Specific recognition at this stage may, therefore, depend on the third structural element of the influenza promoter-structure, the unpaired adenosine in 5′ position 10. It is expected to cause a characteristic angle (and torsion) between the two rigid elements to the left and right, which may serve as a discriminating element against a background of other RNA molecules in the cell. Our data indicate that A10 may constitute a flexible joint between the other two rigid, double-stranded elements, and can be exchanged within a rather broad spectrum of structural alternatives. This variability excludes any direct interaction between RNA polymerase and angular A10 itself, different from a pattern of recognition in several other such systems (Wu & Uhlenbeck, 1987; Wolters et al., 1992). A flexible joint of this kind may be required for a simultaneous interaction of a spherical enzyme protein with both of the rigid elements.

Interaction with both the proximal and the distal element is expected to continue even after vRNA is converted into an open complex of vRNA/polymerase binding, i.e., extending across the angle caused by A10. In accordance, no base pair insertion was tolerated between the two elements of recognition in the immediate neighborhood of that angle, whereas base

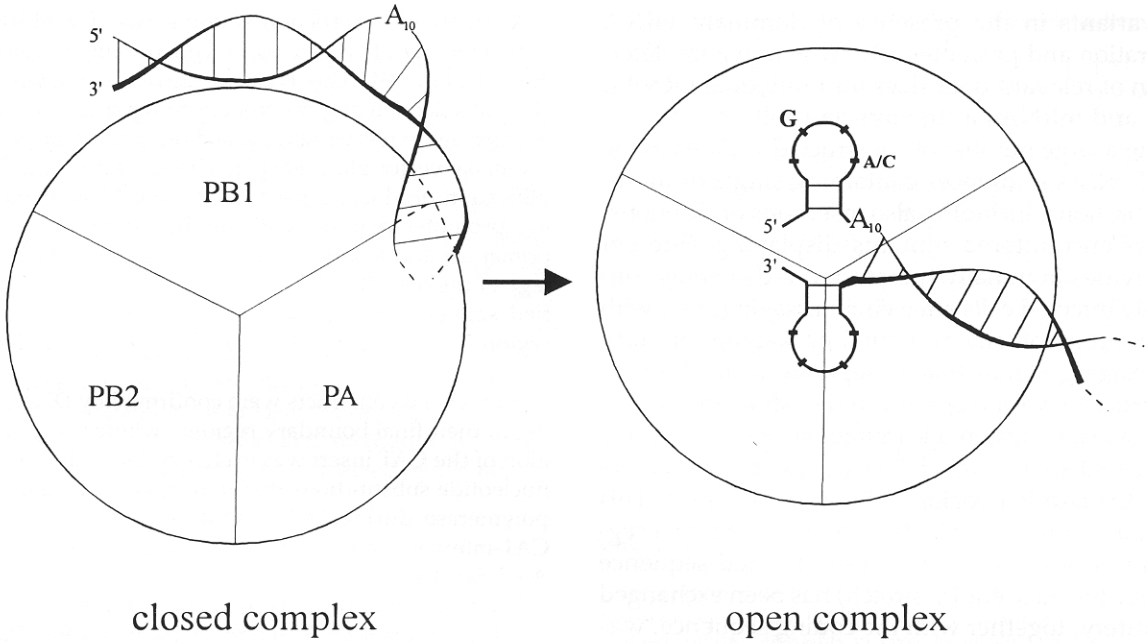

FIGURE 8. Model for two-step interaction of viral RNA polymerase with the vRNA terminal promoter sequence. Initial interaction of viral polymerase with the bulged-10 panhandle vRNA structure (wild-type or perfectly double-stranded, as in pHL1104) is shown as an equivalent of a closed complex formation. The protein-bound vRNA is proposed to convert into an open complex interaction with viral polymerase according to the corkscrew model. The three subunits of the viral polymerase (PB1, PB2, and PA) are drawn arbitrarily with regard to their interaction with different elements of the promoter structure.

pair insertions were indeed tolerated in an equivalent, but external position within the proximal element (compare structures in Fig. 7A).

All considerations of asymmetric double-strand interactions face yet another problem. Not only minus-strand vRNA molecules, but also plus-strand cRNA molecules may be expected to form terminal RNA promoter structures, again because of their largely complementary sequence and interaction with the same viral RNA polymerase. Although at first it may be doubtful to predict whether a particular feature observed in polymerase/RNA interactions in vivo might have to be projected upon vRNA *or* cRNA structures, a number of results argues in favor of a vRNA interpretation. In analyzing the distal element, all single-nucleotide substitutions leading to a G:U mismatch in vRNA (C:A in cRNA) are reduced only slightly in their promoter activities, whereas all C:A mismatch constructs in relation to vRNA (G:U in cRNA) are as inactive as all other mismatch constructs resulting from transversions: e.g., A:A, G:G (Fig. 2C). Similarly, all G:U mismatch constellations within the corkscrew base pairs and no C:A pairs (Fig. 5) are also tolerated at much higher promoter levels, and therefore, again consistently relate to the vRNA structure. If projected onto cRNA instead, it would have to be the C:A mismatch constructs that end up in the 80–100% activity range, whereas the G:U pairs would be all entirely or almost entirely inactive, throughout all the dsRNA base pairs, e.g., including base pair 13:$\overline{12}$, i.e., the central position in an undisputed double-stranded RNA segment. And finally, all of the known tetra-loop conformations have position 2 from the 5′ end (G5 in vRNA) exposed at the top of the structure, i.e., ready for specific interaction with viral polymerase. Because of the opposite polarity, in the cRNA corkscrew structure this exposed position is occupied by cRNA U$\bar{6}$, i.e., the complementary nucleotide to vRNA A6, which is freely exchangeable and therefore unlikely to bind specifically to polymerase.

All these data argue in favor of a vRNA promoter interpretation, which is present in an infected cell in about 10-fold excess over cRNA (Mukaigawa et al., 1991) and via its initiation of mRNA synthesis also gives rise to CAT protein synthesis and the actual measurement of CAT activity. A final decision, whether the specific requirement for G5 reflects an interaction of vRNA-G5 with RNA polymerase or instead an interaction of cRNA-C$\bar{5}$ (in the 3′ branch) with the enzyme, cannot be made, however, at this stage. It also remains an open question how the cRNA promoter may be functioning in the presence of a U$\overline{10}$ angular residue in the 3′ end (not tolerated in vRNA, see pHL1164; Neumann & Hobom, 1995) instead of an A10 angular residue as in vRNA, and also without any invariant residue in particular in position $\bar{6}$ (to be exposed as position 2 in a 5′ cRNA tetra-loop). Analysis of cRNA pro-

moter variants in the presence of dominant vRNA concentration and promoter activity will require determination of relevant pool sizes for individual vRNAs, cRNAs, and mRNAs in the infected cell.

Among a large number of constructed variants in the terminal vRNA sequences containing single or multiple substitutions, including also insertions or deletions, we never encountered mutants displaying different CAT activities in initially transfected cells versus consecutively infected cells after viral passaging, i.e., with and without involvement of viral packaging. Because any packaging signal that is separate from the promoter structure would be expected to show up in such an extensive and systematic mutagenization study, we conclude that such a signal does not exist as a separate entity in the terminal region, or indeed elsewhere. This conclusion is supported by the observation that a segment in which the remaining untranslated sequence (except for the adjacent U_6-stretch) has been exchanged in its entirety, together with the coding sequence, was not excluded from viral passaging (E. Hoffmann & G. Hobom, unpubl. results). These data support a statistical mode of influenza RNP encapsidation during virion formation at the plasma membrane.

MATERIALS AND METHODS

Plasmid constructions

Most of the influenza promoter variants were constructed as derivatives of plasmid pHL1104, another less extensive series was related to pHL1102. These plasmids, containing an ampicillin-resistance gene for selection in *Escherichia coli* and the components of a murine RNA polymerase I system for transcription in mouse cells, serve as a cDNA template for synthesis of influenza vRNA-type molecules after DNA transfection in vivo. In both plasmids, the viral segment 4 cDNA was inserted in antisense orientation between the rDNA promoter and rDNA terminator signals, with the CAT reporter gene exactly replacing the HA coding sequence. In more detail, that CAT gene was flanked by the noncoding sequences of segment 4 (34 and 29 nt), including, however, three or two nucleotide substitutions, respectively, in the 3′-terminal region of the transcribed vRNA (G$\bar{3}$ A and C$\bar{8}$ U in both plasmids, U$\bar{5}$ C in pHL1104 only; Neumann & Hobom, 1995).

Variations in either vRNA promoter region were constructed by using PCR techniques. For mutagenization reactions in the 5′ region, a general primer, complementary to a section of the rDNA promoter, and a specific primer, containing the desired nucleotide substitution(s), were used for amplification. Mutations in the 3′ vRNA promoter region were introduced similarly using a common primer hybridizing to a position in the CAT-gene, plus a specific primer carrying the nucleotide substitution(s) to be introduced into the 3′-terminal vRNA segment. Amplification products of the 5′ vRNA segment were cloned as 56-bp *Bgl* II/*Spe* I fragments, and PCR products of the 3′-segment of viral cDNA (573 bp) were cloned via *Sca* I/*Nco* I restriction into pHL1104 DNA correspondingly deleted for the pre-existing promoter sequence.

Alternatively, restriction endonuclease *Bsm*BI (or its isoschizomer *Esp* 3I) was used for constructing promoter variants of pHL1104. This type IIs restriction enzyme cleaves DNA at a defined distance from its recognition site and is able to delete its own externally located recognition site, both from a suitable vector plasmid (pHL1261) and from the PCR amplification products consisting of the CAT gene plus flanking promoter regions, with modifications introduced via primer oligonucleotides. This cloning procedure takes advantage of the *Bsm* BI properties so that no additional or modified sequence will remain in the manipulated promoter region, in spite of the absence of any local "regular" cleavage site.

All plasmid constructs were confirmed by DNA sequencing in their final boundary regions, whereas the central region of the CAT insert was exchanged in order to eliminate nucleotide substitutions that may have been caused by *Tfl*-polymerase during PCR amplification. For this purpose, CAT-internal restriction sites *Nco* I/*Bsp*M II, *Spe* I/*Bam*H I, or *Ase* I/*Sac* II were used.

For a combination of constructs carrying various 5′-end *and* 3′-end variations, 3′ mutant fragments were obtained by *Spe* I and *Bam*H I restriction and inserted into the appropriate 5′-terminal variation plasmids now used as vector DNAs.

Cells and viruses

Influenza A/FPV/Bratislava virus was propagated in MDCK cells. For transfection and passaging experiments, B82 cells (a mouse L cell line) and MDCK cells were used, respectively.

Lipofectamin DNA transfection and influenza virus helper infection

For DNA transfection, approximately 10^7 B82 cells were used. Five micrograms of plasmid DNA and 30 μL Lipofectamin (Lipofectamine™, GIBCO/BRL, 1 μg/μL) were mixed in serum-free Dulbecco minimal essential medium (DMEM) and incubated for 15 min at room temperature. Cells were washed twice with serum-free medium, and incubated with the DNA/Lipofectamin mixture for 1–6 h. After further incubation with DMEM containing 10% serum (FCS) for 1 h, the transfected cells were washed with PBS+ (2.5 mM $MgCl_2$, 3.4 mM $CaCl_2$ added), and superinfected with influenza helper viruses (Influenza A/FPV/Bratislava) at a multiplicity of infection of 0.1–1. After 30–45 min, the cells were washed again and incubated with DMEM (+ 10% FCS) for 8 h. A complete replication cycle takes place during this period.

Passaging of virus-containing supernatants

At 8 h post influenza infection, DNA-transfected and infected cells were harvested for CAT assays, and the corresponding supernatants for passaging of viral progeny. Cell debris was removed by centrifugation (10,000 rpm, 5 min), and 1 mL of the B82 supernatants was transferred to approximately 10^7 MDCK cells and incubated for 30–45 min. After 8 h of infection, the cytopathic effect was observed, and cells and supernatants were treated in another round as described before.

CAT assay

Cell extracts were prepared as described by Gorman et al. (1982). Fifty-five microliters of cell lysate were mixed with 10 μL acetyl-CoA (4 mM), and 15 μL of fluorescent-labeled chloramphenicol (borondipyrromethane difluoride fluorophore: BODIPY™ CAM substrate, Flash Cat kit, Stratagene) and incubated at 37 °C for 3–15 h. For extraction of reaction products, 1 mL ethylacetate was added and, after centrifugation for 1 min at 13,000 rpm, the upper phase containing the acetylated chloramphenicol was isolated and the solvent evaporated (either 24 h at room temperature, or 45 min in a vacuum centrifuge). The resulting pellet was resuspended in 20 μL ethylacetate and the acetylchloramphenicol reaction products were separated by thin-layer chromatography using a solvent mixture of chloroform and methanol (87:13%). Finally, the reaction products were visualized by UV illumination and photography.

ACKNOWLEDGMENTS

We acknowledge G. Kochs and F. Weber, Freiburg, for inspiring discussions. We thank C. Beckort, N. Pohl, and U. Ruppert for expert technical assistance. Influenza strain FPV Bratislava (mouse-adapted) was kindly provided by P. Staeheli. This work was supported by the Deutsche Forschungsgemeinschaft. R.F. was supported by the Boehringer Ingelheim Fonds. This work is in partial fulfillment of the requirements for the Ph.D. degree of R.F. (University of Giessen).

Received May 10, 1996; returned for revision June 10, 1996; revised manuscript received July 23, 1996

REFERENCES

Baudin F, Bach C, Cusack S, Ruigrok RWH. 1994. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J 13*:3158–3165.

Cianci C, Tiley L, Krystal M. 1995. Differential activation of the influenza virus polymerase via template RNA binding. *J Virol 69*:3995–3999.

Fodor E, Pritlove DC, Brownlee GG. 1994. The influenza virus panhandle is involved in the initiation of transcription. *J Virol 68*:4092–4096.

Fodor E, Pritlove DC, Brownlee GG. 1995. Characterization of the RNA fork model of virion RNA in the initiation of transcription in influenza A virus. *J Virol 69*:4012–4019.

Fodor E, Seong BL, Brownlee GG. 1993. Photochemical cross-linking of influenza A polymerase to its virion RNA promoter defines a polymerase binding site at residues 9 to 12 of the promoter. *J Gen Virol 74*:1327–1333.

Gorman M, Moffat L, Howard B. 1982. Recombinant genomes which express chloramphenicol acetyl-transferase in mammalian cells. *Mol Cell Biol 2*:1044–1057.

Hagen M, Chung TDY, Butcher JA, Krystal M. 1994. Recombinant influenza virus polymerase: Requirement of both 5′ and 3′ viral ends for endonuclease activity. *J Virol 68*:1509–1515.

Honda A, Ueda K, Nagata K, Ishihama A. 1988. RNA polymerase of influenza virus. Role of NP on RNA chain elongation. *J Biochem 104*:1021–1026.

Hsu M, Parvin JD, Gupta S, Krystal M, Palese P. 1987. Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc Natl Acad Sci USA 84*:8140–8144.

Krug RM, Alonso-Caplan FV, Julkunen I, Katz MG. 1989. Expression and replication of the influenza virus genome. In: Krug RM, ed. *The influenza viruses*. New York: Plenum Press. p 89–152.

Lee YS, Seong BL. 1996. Mutational analysis of influenza B virus RNA transcription in vitro. *J Virol 70*:1231–1236.

Li X, Palese P. 1992. Mutational analysis of the promoter required for influenza virus virion RNA synthesis. *J Virol 66*:4331–4338.

Luo G, Luytjes W, Enami M, Palese P. 1991. The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure. *J Virol 65*:2861–2867.

Luytjes W, Krystal M, Enami M, Parvin JD, Palese P. 1989. Amplification, expression and packaging of a foreign gene by influenza virus. *Cell 59*:1107–1113.

Martin K, Helenius A. 1991. Transport of incoming influenza nucleocapsids into the nucleus. *J Virol 65*:232–244.

Mukaigawa J, Hatada E, Fukuda R, Shimizu K. 1991. Involvement of the influenza A virus PB2 protein in the regulation of viral gene expression. *J Gen Virol 72*:2661–2670.

Neumann G, Hobom G. 1995. Mutational analysis of influenza virus promoter elements in vivo. *J Gen Virol 76*:1709–1717.

Neumann G, Zobel A, Hobom G. 1994. RNA polymerase I-mediated expression of influenza viral RNA molecules. *Virology 202*:477–479.

Piccone ME, Fernandez-Sesma A, Palese P. 1993. Mutational analysis of the influenza virus vRNA promoter. *Virus Res 28*:99–112.

Pley HW, Flaherty KM, McKay DB. 1994. Three-dimensional structure of a hammerhead ribozyme. *Nature 372*:68–74.

Riordan FA, Bhattacharyya A, McAteer S, Lilley DMJ. 1992. Kinking of RNA helices by bulged bases, and the structure of the human immunodeficiency virus transactivator response element. *J Mol Biol 266*:305–310.

Seong BL, Brownlee GG. 1992. Nucleotides 9 to 11 of the influenza A virion RNA promoter are crucial for activity in vitro. *J Gen Virol 73*:3115–3124.

Shi L, Summers DF, Pen Q, Galarza JM. 1995. Influenza A virus RNA polymerase subunit PB2 is the endonuclease which cleaves host cell mRNA and functions only as the trimeric enzyme. *Virology 208*:38–47.

Stoeckle MY, Shaw MW, Choppin PW. 1987. Segment-specific and common nucleotide sequences in the noncoding regions of influenza B virus genome RNAs. *Proc Natl Acad Sci USA 84*:2703–2707.

Tiley LS, Hagen M, Matthews JT, Krystal M. 1994. Sequence-specific binding of influenza virus RNA polymerase to sequences located to the 5′ ends of the viral RNAs. *J Virol 68*:5108–5116.

Valegård K, Murray JB, Stockley PG, Stonehouse NJ, Liljas L. 1994. Crystal structure of an RNA bacteriophage coat protein–operator complex. *Nature 371*:623–626.

Wimberly B, Varani G, Tinoco I Jr. 1993. Conformation of loop E from eukaryotic 5 S RNA. *Biochemistry 32*:1078–1087.

Wolters J. 1992. The nature of preferred hairpin structures in 16S-like rRNA variable regions. *Nucleic Acids Res 20*:1843–1850.

Wu HN, Uhlenbeck OC. 1987. Role of a bulged A residue in a specific RNA-protein interaction. *Biochemistry 26*:8221–8229.

Wyatt JR, Tinoco Jr I. 1993. RNA structural elements and RNA function. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 465–496.

Yamanaka K, Ogasawara N, Yoshikawa H, Ishihama A, Nagata K. 1991. In vivo analysis of the promoter structure of the influenza virus RNA genome using a transfection system with an engineered RNA. *Proc Natl Acad Sci USA 88*:5369–5373.

Zobel A, Neumann G, Hobom G. 1993. RNA polymerase I catalysed transcription of insert viral cDNA. *Nucleic Acids Res 21*:3607–3614.