Nucleic Acids Research, 1995, Vol. 23, No. 3 377–382

# Localization of binding site for encephalomyocarditis virus RNA polymerase in the 3′-noncoding region of the viral RNA

**Taian Cui and Alan G. Porter***

Institute of Molecular and Cell Biology, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511, Republic of Singapore

Received November 4, 1994; Revised and Accepted January 4, 1995

## ABSTRACT

**We previously showed that encephalomyocarditis (EMC) virus RNA-dependent RNA polymerase ($3D^{pol}$) binds specifically to 3′-terminal segments of EMC virus RNA. This binding, which depends on both the 3′-noncoding region (3′-NCR) and 3′-poly (A) tail [together denoted 3′-NCR(A)], may be an important step in the initiation of virus replication. In this paper, the 3′-NCR and 3′-poly(A) were separately transcribed then mixed, but no complex with $3D^{pol}$ was obtained, showing that covalent atttachment of the 3′-poly(A) to the 3′-NCR is essential for complex formation. Mutational and deletion analyses localized a critical determinant of $3D^{pol}$ binding to a U-rich sequence located 38–49 nucleotides upstream of the 3′-poly(A). Similar analyses led to the identification of a sequence of A residues between positions +10 and +15 of the 3′-poly(A) which are also critical for $3D^{pol}$ binding. As U-rich and A-rich regions are important for $3D^{pol}$ binding, a speculative model is proposed in which $3D^{pol}$ induces and stabilizes the base-pairing of the 3′-poly(A) with the adjacent U-rich sequence to form an unusual pseudoknot structure to which $3D^{pol}$ binds with high affinity.**

## INTRODUCTION

All picornaviruses including polioviruses, encephalomyocarditis (EMC) virus and foot-and-mouth disease (FMD) virus, encode an RNA-dependent RNA polymerase ($3D^{pol}$) which copies both the plus and minus strand viral RNAs during virus replication (1). Besides its function in replication, poliovirus $3D^{pol}$ also modifies the specificity of the viral protease $3C^{pro}$ (2,3); and FMDV $3D^{pol}$, which is believed to have nuclease activity is packed into virus particles with viral RNA (4). Prior to the initiation of replication, $3D^{pol}$ must presumably also recognize and form a pre-initiation complex with the 3′-terminal region of the viral RNA.

The 3′-terminal regions of bacterial, plant and animal viral RNAs are known to be important in virus replication and specifically in template selection by the cognate viral polymerase (5–11). In some cases, the 3′-poly(A) has been found to be important for viral RNA synthesis (6,7). Genetic studies with poliovirus have established that the 3′-noncoding region (3′-NCR) of poliovirus RNA is required for *in vivo* synthesis of minus strand RNAs (12) and that the genetically coded 3′-poly(A) is essential for virus replication (13,14). Moreover, partially purified poliovirus $3D^{pol}$ bound equally well to full-length poliovirus RNA and 3′-terminal fragments of poliovirus RNA (15). Purified $3D^{pol}$ of EMC virus was able to form a specific complex *in vitro* with 106 nucleotides of the heteropolymeric 3′-NCR of EMC virus RNA provided that a 3′-poly(A) tail of 30 nucleotides was also present, but poly(A) alone could not form a complex with $3D^{pol}$ (16). These results suggested that the binding site for $3D^{pol}$ in the 3′-NCR(A) region must include the 3′-poly(A) tail.

In this paper we localize the sequences in the 3′-NCR and 3′-poly(A) regions of EMC virus RNA which are critical for efficient binding to EMC virus $3D^{pol}$ and we propose a model for the binding reaction based on a novel interaction between the 3′-poly(A) and a U-rich sequence in the heteropolymeric 3′-NCR.

## MATERIALS AND METHODS

### Purification and assay of recombinant $3D^{pol}$

Recombinant EMC virus $3D^{pol}$ was expressed, purified and assayed as described previously (17,18). To obtain a concentrated protein preparation, the eluant from the glutathione–Sepharose 4B column containing $3D^{pol}$ was precipitated using $(NH_4)_2SO_4$ (16). After centrifugation and removal of the supernatant, the pellet was resuspended in 50 mM Tris–HCl pH 8.8, 150 mM NaCl, 1 mM EDTA and 1 mM phenylmethylsulfonyl fluoride. The concentration of $3D^{pol}$ was measured by Bio-Rad protein assay.

### Plasmid construction

The construction of plasmid pJC2 has been described (16). pTAO is identical to pJC3 (16). To construct plasmids pTA25, pTA20,

* To whom correspondence should be addressed

pTA15, pTA10 and pTA5, the DNA of pJC2 was amplified using the following primers:

Foreward primer 1: 5′-ATTTAGGTGACACTATAGAATACTC AAGCT-3′
Reverse primers 2–6: 5′-TTTAAGC$(T)_n$CTCTATCTATTTATT TTACTACT-3′

(where n = 26, 21, 16, 11 and 6, equivalent to pTA25, pTA20, pTA15, pTA10 and pTA5, respectively). DNA fragments obtained from the PCR amplification were double-digested with *Mlu*I and *Hin*dIII and cloned into pGEM-7Zf(+) digested with the same two restriction enzymes. The plasmids pTC2, pEC2 and pTC6 were constructed by amplifying pJC2 DNA using the following primers:

primer 7: 5′-AGGATCCATTTAGGTGACACTATAGAATACT GCAGACAGGGTTCTTCTACTTTG-3′
primer 8: 5′-AGGATCCTAATACGACTCACTATAGGGCGA AAATAAATAGATAGAGAAAAA-3′
primer 9: 5′-TTTAAGCTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTCTCTATC-3′
primer 10: 5′-TTTAAGCTTTTTTTTTTTTTTTTAAATTTTT TTTTTTTCTCTATC-3′

The PCR products obtained from primers 7 and 9 (for plasmid pTC2), 8 and 9 (for pEC2) and 7 and 10 (for pTC6) were digested with *Bam*HI and *Hin*dIII and cloned into plasmid pUC19 digested with the same pair of restriction enzymes. To construct plasmids pTC3, pTC4 and pTC5, primer 9 together with the following primers was used to generate the PCR fragment (pJC2 DNA as template):

primer 11: 5′-TACTGCAGACAGGGTTCTTCTACAAAGCA AGATAGTCTAG-3′
primer 12: 5′-TACTGCAGACAGGGAACTTCTACTTTGCA AGATAGTCTAG-3′
primer 13: 5′-TACTGCAGACAGGGTTCAACTACTTTGCA AGATAGTCTAG-3′

After digesting with *Pst*I and *Hin*dIII, the fragments obtained from primers 9 plus 11, 9 plus 12 and 9 plus 13 were cloned into pTC2 digested with *Pst*I and *Hin*dIII, giving plasmids pTC3, pTC4 and pTC5, respectively.

Plasmids pTC7 and pTC8 were obtained by amplifying pJC2 DNA using primer 9 and the following two primers, respectively:

Primer 14: 5′-AAAGGATCCATTTAGGTGACACTATAGTTC TTCTACTTTGCAAGATAGTCTAG-3′
Primer 15: 5′-AAAGGATCCATTTAGGTGACACTATAGCA AGATAGTCTAGAGTAGTAAAATA-3′

The DNA fragments obtained from the PCR amplification were cloned directly into pGEM-T and then digested with *Bam*HI and *Hin*dIII. The small inserts obtained were then subcloned into pUC19, giving rise to plasmids pTC7 and pTC8. All cloned DNAs were sequenced using the dsDNA Cycle Sequencing System (Life Technologies, Inc.).

### *In vitro* transcription of RNAs

All plasmids except pTAO were linearized with *Hin*dIII prior to transcription, which was based on the protocol of Promega Corp. (pTAO was linearized with *Eco*RI). The reaction mixture of 20 µl containing 1.0 µg plasmid DNA, 10 mM DTT, 20 U of ribonuclease inhibitor, 50 µCi of [α-$^{33}$P] UTP (NEN DuPont), 12 µM UTP, 2.5 mM each of ATP, CTP and GTP and 20 U of SP6 RNA polymerase (Promega) (except pEC2 for which T7 RNA polymerase was used) was incubated at 40°C for 60 min. The DNA template was removed with 40 U RNase-free DNase I (Promega Corp.) for 10 min at 37°C. Purification was achieved either by phenol/chloroform extraction followed by ammonium acetate/ethanol precipitation or by using Chroma Spin plus TE-30 columns (Clontech). The quantitation of RNA was performed as previously described (16).

### RNA band shift assay

The standard binding reactions were performed in 12 µl containing 50 mM Tris–HCl pH 7.2, 10 mM DTT, 60 mM NaCl, 60 mM KCl, 2 mM $MgCl_2$, 1% glycerol, ~100 ng of $^{33}$P-labeled viral RNA transcript and 20 µg EMC virus 3D$^{pol}$ at 30°C for 15 min. After adding gel loading buffer, the samples were electrophoresed in a 6% non-denaturing polyacrylamide gel and autoradiographed (Fig. 1A) as previously described (16). For the small transcripts of pTC7 and pTC8, a 12% polyacrylamide gel was used. All bandshift experiments were done at least three times.

### RNA secondary structure prediction

RNA secondary structures were predicted (Figs 2 and 3A) using the RNAFOLD programme (PC/Gene, Intelligenetics, Inc.).

### Quantitation of bands on autoradiographs

The quantitation of images of labeled RNA bands was performed using a Visage 2000 Image system (BioImage Products, Ann Arbor, USA). The percentage of the RNA transcript bound to 3D$^{pol}$ was calculated from the optical density of the bands corresponding to the bound and free RNA.

## RESULTS

### Minimum length of 3′-poly(A) required for 3D$^{pol}$ binding

Using pairs of PCR primers, deletions were made progressively from the 3′-end of the 3′-poly(A) tail of EMC virus RNA beginning with the transcript JC2 (16), which consists of 106 nucleotides of the 3′-NCR and 30 adenylate residues of the 3′-poly(A). When the 3′-poly(A) was shortened in steps of five nucleotides from 25 to 15 adenylates, there was a gradual decrease from 40 to 24% in the proportion of $^{33}$P-labeled transcripts bound to 3D$^{pol}$ (Fig. 1A and B). However, a further reduction in the size of the 3′-poly(A) from 15 to 10 adenylates resulted in a drastic drop in RNA bound from 24 to 1.8% (Fig. 1B). Removal of 25 adenylates or complete deletion of the 3′-poly(A) abolished the binding of 3D$^{pol}$ (Fig. 1A and B). These results show that at least 10–15 A residues are necessary for the efficient binding of 3D$^{pol}$ to the 3′-NCR region.

To provide further evidence that adenylates between the tenth and fifteenth A residue of the 3′-poly(A) tail (A116–A121; Fig. 3A) are essential for 3D$^{pol}$ binding, three consecutive A residues (A119–A121) were mutated to UUU, resulting in almost total loss of 3D$^{pol}$ binding (Fig. 3C, lane 2; compare with Fig. 3B, lane 2).

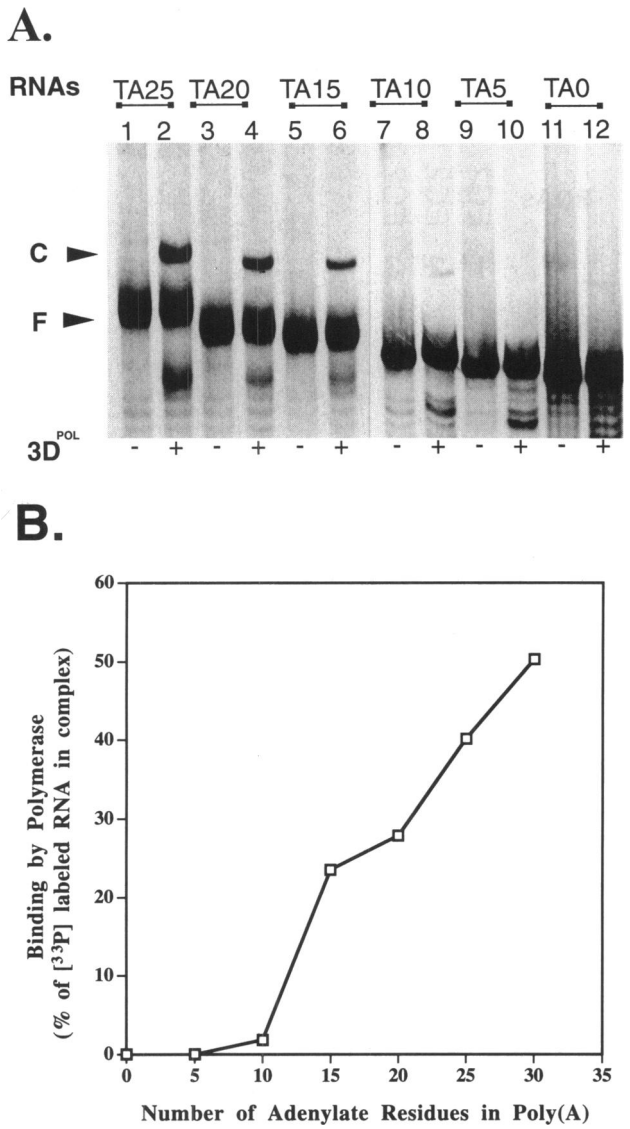


**Figure 1.** Analysis of 3′-NCR(A) constructs with shorter 3′-poly(A) tracts. (**A**) Non-denaturing PAGE bandshift analysis of $3D^{pol}$:3′-NCR(A) complex formation. TAO to TA25 refers to the length of the 3′-poly(A) in nucleotides. Each RNA was subjected to bandshift analysis with (+) and without (–) $3D^{pol}$. C = complex, F = free RNA. (**B**) Quantitation of 3′-NCR(A) bound to $3D^{pol}$ in constructs with progressively shorter 3′-poly(A). The percent binding was calculated from dividing the amount of bound RNA by the bound plus free RNA as determined by scanning the autoradiograph and computing the optical density (Materials and Methods).

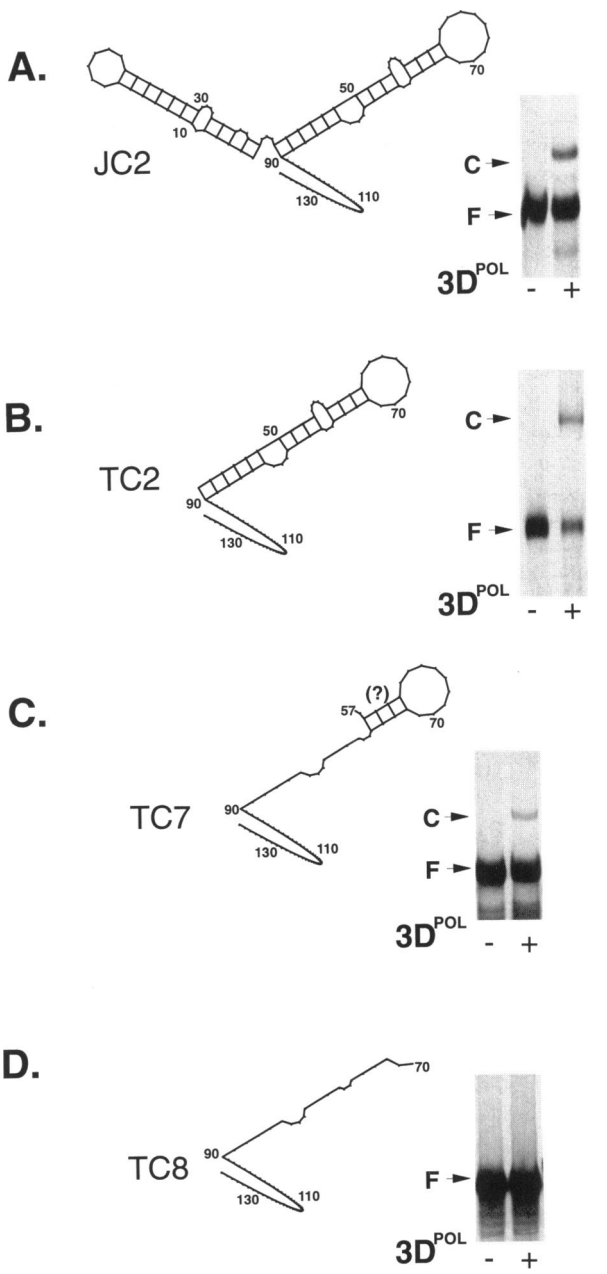


**Figure 2.** The predicted secondary structures of the 3′-NCR(A) region and 5′-shortened derivatives together with their binding to $3D^{pol}$ (PAGE bandshift analysis). (**A**) JC2 is the 3′-NCR(A) control. (**B**) TC2 is 43 5′-terminal nucleotides deleted from the 3′-NCR(A). (**C**) TC7 is 56 5′-terminal nucleotides deleted from the 3′-NCR(A). (**D**) TC8 is 69 5′-terminal nucleotides deleted from the 3′-NCR(A). C = complex, F = free RNA.

## Minimum length of the heteropolymeric 3′-NCR required for $3D^{pol}$ binding

The nucleotide sequence of the 3′-NCR(A) region transcribed from plasmid pJC2 (Fig. 1 of ref. 16) comprises 106 nucleotides of the 126 nucleotide 3′-NCR located upstream of the 30 nucleotide 3′-poly(A) tail. The secondary structure predicted for this region suggests the potential of the free 3′-NCR to form two almost adjacent stem and loop structures (Fig. 2A) separated from the 3′-poly(A) by a non-base-paired sequence of 16 nucleotides (Fig. 3A). Significant binding of $3D^{pol}$ to JC2 RNA occurred (Fig. 2A) as was previously observed (16). Deletions were made progressively from the 5′-end of JC2 RNA to determine the minimum length of 3′-NCR required for $3D^{pol}$ binding. Deletion mutants with the loss of nucleotides 1–43 or 1–56 still bound to $3D^{pol}$ (Fig. 2B and C), but removal of nucleotides 1–69 (Fig. 2D) or 1–89 (generating EC2 RNA; data not shown) completely abolished binding. Thus, the 3′-terminal ~49 nucleotides of the heteropolymeric 3′-NCR are essential for $3D^{pol}$ binding and

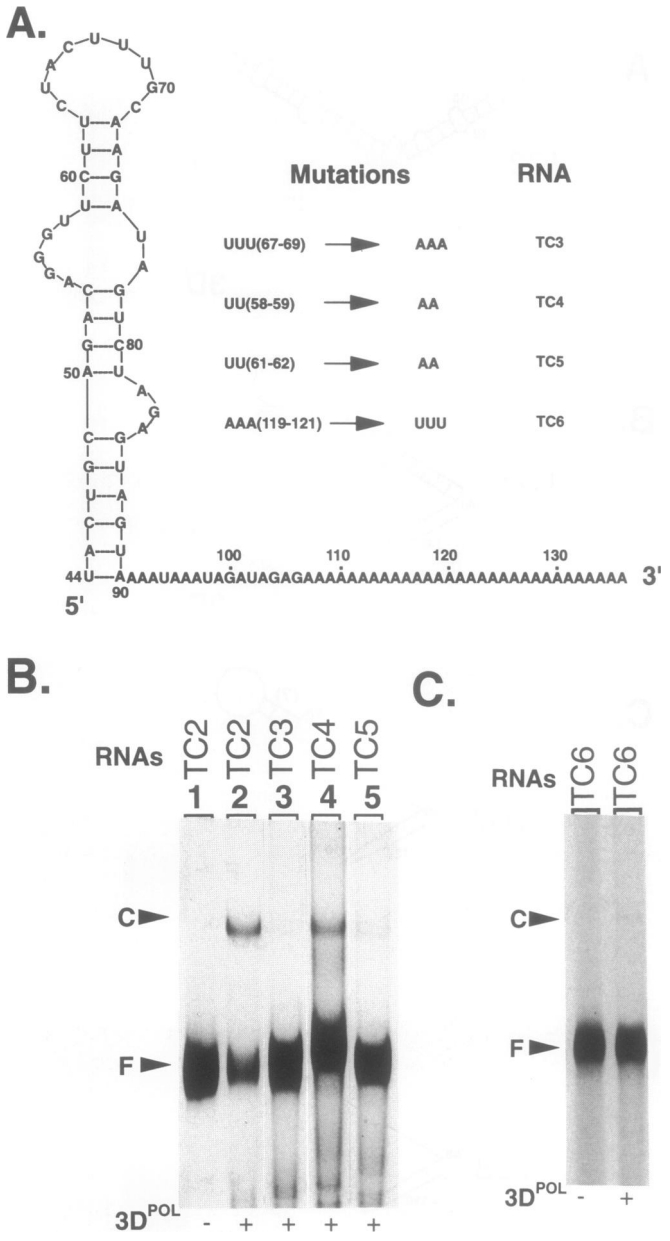


**Figure 3.** Analysis of binding of mutated 3′-NCR(A) regions to $3D^{pol}$. (**A**) Positions of double or triple mutations in the predicted secondary structure of TC2 (Fig. 2B). (**B** and **C**) PAGE bandshift analysis of mutants shown in panel A (TC3 to TC6). C = complex, F = free RNA.

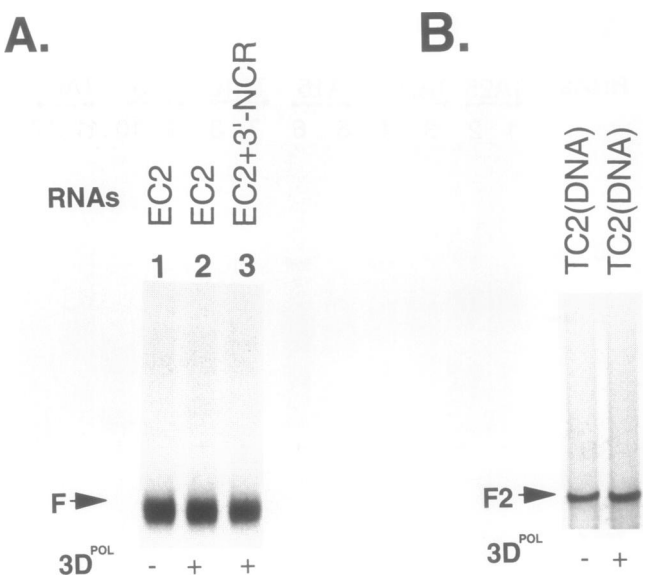


**Figure 4.** The 3′-NCR and 3′-poly(A) must be covalently joined and in the form of RNA. (**A**) Failure of separately transcribed 3′-NCR and 3′-poly(A) to bind to $3D^{pol}$ (PAGE bandshift analysis). Lane 1, $^{33}$P-labelled EC2 [17 nucleotides of 3′-NCR and 30 nucleotides of 3′-poly(A)]; lane 2, as lane 1 plus $3D^{pol}$; lane 3, as lane 2 plus unlabelled 3′-NCR (JC3, ref. 16). F = free EC2 RNA. (**B**) Failure of single-stranded DNA equivalent of 3′-NCR(A) region to bind to $3D^{pol}$. The 3′-poly(dA) consisted of 30 deoxyribonucleotides and is equivalent to JC2 (Fig. 2A). F2 = free DNA.

critical determinants of this binding lie in a U-rich sequence of 12 nucleotides between residues 58 and 69 (Fig. 3A).

### The importance of a U-rich region in the stem–loop adjacent to the 3′-poly(A)

As both the 3′-NCR and 3′-poly(A) are required for $3D^{pol}$ binding (16), the 3′-NCR may interact with the 3′-poly(A) even though base-pairing was not predicted by the RNAFOLD programme (Fig. 2A). We tested the idea that an interaction between the 3′-poly(A) and the important U-rich sequence (5′-UUCUUCUA-CUUU-3′) located between nucleotides 58 and 69 (Fig. 3A) might be induced and stabilized by $3D^{pol}$. Starting with transcript TC2 (Fig. 2B), various U residues in this region were converted to A by *in vitro* mutagenesis in order to test their importance in $3D^{pol}$ binding. Conversion of UU [58–59] to AA giving TC4 was without effect (Fig. 3B, lane 4), but conversion of either UUU [67–69] to AAA giving TC3, or UU [61–62] to AA giving TC5 significantly impaired the ability of the RNA to bind to $3D^{pol}$ (Fig. 3B, lanes 3 and 5). In two other experiments, binding of the TC3 and TC5 mutant RNAs to $3D^{pol}$ was undetectable (data not shown). This further shows the importance of part of the U-rich region in $3D^{pol}$ binding and raises the possibility that it may interact with the important oligoadenylate sequence in the 3′-poly(A) between A116 and A121 (Fig. 3A).

### Complex formation depends on ribonucleic acid and covalent linkage of 3′-NCR to 3′-poly(A)

To determine whether the 30 nucleotide 3′-poly(A) must be covalently attached to the 3′-NCR for binding to $3D^{pol}$, the 3′-NCR without poly(A) (JC3; 16) was transcribed and mixed with a transcript comprising the 3′-terminal 17 nucleotides of the 3′-NCR and the complete 3′-poly(A) (EC2). $3D^{pol}$ repeatedly failed to bind to the unlinked RNAs, regardless of whether it was the JC3 RNA or the EC2 RNA that was labeled with $^{33}$P-UTP (Fig. 4A). This indicates that the 3′-NCR must be covalently joined to the 3′-poly(A) for $3D^{pol}$ binding to occur.

Finally, it was of interest to know whether DNA would substitute for RNA in complex formation with $3D^{pol}$. A synthetic single-stranded DNA fragment exactly equivalent to the 3′-NCR(A) was unable to bind to $3D^{pol}$ at 30°C (Fig. 4B) or at 25°C (data not shown). Thus, a 2′-hydroxyl in the ribose moieties is essential for $3D^{pol}$ binding.

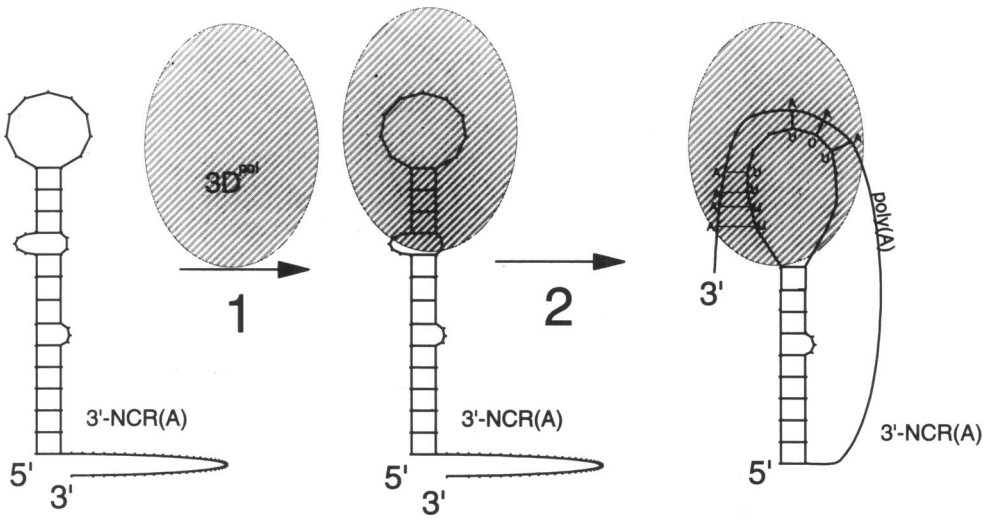


**Figure 5.** Model proposed for formation of complex between 3′-NCR(A) and 3D$^{pol}$. The shaded area represents EMC virus 3D$^{pol}$. In step 1, 3D$^{pol}$ binds with low affinity to the loop and upper stem. In step 2, the upper stem is unwound and the U-rich sequence is now able to base-pair with the 3′-poly(A). This creates the high affinity binding site for 3D$^{pol}$.

## DISCUSSION

The core binding site for EMC virus 3D$^{pol}$ has been mapped to a region of not more than ~64 nucleotides, comprising the 3′-terminal ~49 nucleotides of the heteropolymeric 3′-NCR and 10–15 A residues of the 3′-poly(A). An important question is what role does the 3′-poly(A) tail play in 3D$^{pol}$ recognition and pre-initiation complex formation? It is possible that the 3′-NCR and 3′-poly(A) do not interact with each other but interact with distinct binding sites on 3D$^{pol}$. Alternatively, the 3′-poly(A) base-pairs with U-rich sequences in the 3′-NCR, forming a unique binding site for 3D$^{pol}$. Several considerations argue against the former possibility. First, synthetic poly(A), the EMC virus 3′-NCR and polyadenylated globin mRNA do not compete with the EMC virus 3′-NCR(A) for 3D$^{pol}$ binding (16). Secondly, the 3′-NCR must be covalently linked to the 3′-poly(A), implying a specific interaction between them.

A stretch of eight uridylates in a 12-nucleotide sequence that could potentially base-pair with the 3′-poly(A) tail lies at or near the tip of a predicted stem–loop structure (Fig. 3A). The deletion analyses showed that these 12 nucleotides are essential for 3D$^{pol}$ binding. Moreover, mutagenesis of certain of these U residues (UU[61–62] or UUU[67–69]) resulted in almost total loss of 3D$^{pol}$ binding. As mutagenesis of UU[58–59] was without effect, these results suggest that the U-rich sequence between positions 61 and 69 must be an important part of the 3D$^{pol}$ binding site. The transcript TC7 beginning at nucleotide 58 formed a complex with 3D$^{pol}$ (Fig. 2C) indicating that the predicted two lower stems (Fig. 3A) are not critical for 3D$^{pol}$ binding. In TC7, the top stem is probably energetically too weak to exist by itself (Fig. 2C). Therefore, the predicted secondary structure illustrated in Figure 3A may not be particularly relevant for complex formation with 3D$^{pol}$. Systematic deletion analysis of the 3′-poly(A) clearly showed the importance of A residues located 10–15 nucleotides from the heteropolymeric 3′-NCR. Consistent with this interpretation, mutagenesis of three consecutive A residues within the +10 to +15 region of the 3′-poly(A) abolished 3D$^{pol}$ binding to the 3′-NCR(A). It had previously been found that poliovirus RNAs with 3′-poly(A) tracts of less than 20 nucleotides were much less infectious (13) and that EMC virus cDNA clones were non-infectious if the 3′-poly(A) tail is less than 15 nucleotides in length (A. Palmenberg, personal communication). We suggest that the 3D$^{pol}$:3′-NCR(A) pre-initiation complex for virus RNA replication is unable to assemble in the EMC cDNA constructs with less than 15 adenylates.

A speculative model for the RNA structure in the preinitiation complex which takes into account our data and theoretical considerations is presented in Figure 5. In this model, 3D$^{pol}$ binds with low affinity to the 3′-NCR around the U-rich region between nucleotides 58 and 69 (Fig. 3A) without involvement of the 3′-poly(A). This induces and stabilizes U:A base-pairing between the 3′-NCR and the +10 to +15 region of the 3′-poly(A) generating at least part of the high affinity site for 3D$^{pol}$ (a pseudoknot). This might explain why the RNAFOLD programme did not predict base-pairing of the 3′-poly(A) in the free 3′-NCR(A) (Figs 2A and 3A). Evidence for base-pairing involving 3′-poly(A) tails of viral RNAs has been obtained for Cowpea Mosaic Virus (6) and poliovirus (19), a member of the picornavirus family. However, these structures are not classified as pseudoknots and it is not known whether they are important in viral RNA polymerase recognition and binding.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contribution from Dr P. A. Walker throughout this work. We would also like to thank Drs Y. Lam and Y. Wang for critically reading the manuscript, Mr Y. Jiang for valuable discussions and Dr K. Pleij (Leiden University, The Netherlands) for communicating unpublished results on the RNA secondary structure predictions. We would also like to thank Francis Leong for photography and Miss Pearly Aw for typing the manuscript.

## REFERENCES


1 Porter, A. G. (1993) *J. Virol.*, **67**:6917–6921.
2 Jore, J., de Geus, B., Jackson, R.J., Pouwels, P.H. and Enger-Valk, B.E. (1988) *J. Gen. Virol.* **69**:1627–1636.
3 Ypma-Wong, M.F., Dewalt, P.G., Johnson, V.H., Lamb, J.G. and Semler, B. (1988) *Virology* **166**:265–270.
4 Newman, J.F.E., Piatti, P.G., Gorman, B.M., Burrage, T.G., Ryan, M.D., Flint, M. and Brown, F. (1994) *Proc. Natl. Acad. Sci. USA* **91**:733–737.
5 Miller, W.A., Bujarski, J.J., Dreher, T.W. and Hall, T.C. (1986) *J. Mol. Biol.* **187**:537–546.
6 Eggen, R., Verver, J., Wellink, J., Pleij, K., van Kammen, A.B. and Goldbach, R. (1989) *Virology* **173**:456–464.
7 Jupin, I., Bouzoubaa, S., Richards, K., Jonard, G. and Guilley, H. (1990) *Virology* **178**:281–284.
8 Kuhn, R.J., Hong, Z. and Strauss, J.H. (1990) *J. Virol.* **64**:1465–1476.
9 Takamatsu, N., Watanabe, Y., Meshi, T. and Okada, Y. (1990) *J. Virol.* **64**:3686–3693.
10 Gargouri-Bouzid, R., David, C. and Haenni, A.-L. (1991) *FEBS Lett.* **294**:56–58.
11 Barrera, I., Schuppli, D., Sogo, J.M. and Weber, H. (1993) *J. Mol. Biol.* **232**:512–521.
12 Sarnow, P., Bernstein, J.D. and Baltimore, D. (1986) *Proc. Natl. Acad. Sci. USA* **83**:571–575.
13 Spector, D.H. and Baltimore, D. (1974) *Proc. Natl. Acad. Sci. USA* **71**:2983–2987.
14 Sarnow, P. (1989) *J. Virol.* **63**:467–470.
15 Oberste, M.S. and Flanegan, J.B. (1988) *Nucleic Acids Res.* **16**:10339–10352.
16 Cui, T., Sankar, S. and Porter, A.G. (1993) *J. Biol. Chem.* **268**:26093–26098.
17 Sankar, S. and Porter, A.G. (1991) *J. Virol.* **65**:2993–3000.
18 Sankar, S. and Porter, A.G. (1992) *J. Biol. Chem.* **267**:10168–10176.
19 Jacobson, S.J., Konings, D.A.M. and Sarnow, P. (1993) *J. Virol.* **67**:2961–2971.