Mutagenesis of the 3'42 Nucleotide Host Protein Binding Element of the MHV 3'UTR

REED F. JOHNSON AND JULIAN L. LEIBOWITZ
Department of Pathology and Laboratory Medicine, The Texas A&M University System Health Science Center, College Station, Texas 77802-1114

1. INTRODUCTION

The mouse hepatitis virus (MHV) genome is a 32kb message sense single stranded RNA. Studies with defective interfering RNAs have shown that the 3' terminal 447nt are required for replication (Kim et al., 1993; Lin et al., 1993). Previous studies have identified two host protein binding elements within the 3' terminal 166nt of the 3'UTR. The first of these is located at nucleotides 154-129 [assigning position 1 to the first nucleotide 5' of the poly (A)] tail and the second is located within the 3' terminal 42 nucleotides (Liu et al., 1997; Yu and Leibowitz, 1995). Both elements contain the conserved motif UGAARNGAAGUU which is required for host protein binding and for DI RNA replication (Yu and Leibowitz, 1995; Liu et al., 1997). In the present study we further identify nucleotides involved in host protein binding and explore the role of RNA secondary structure in this process.

The Nidoviruses (Coronaviruses and Arteriviruses).
Edited by Ehud Lavi *et al.*, Kluwer Academic/Plenum Publishers, 2001.

2. MATERIALS AND METHODS

2.1 Cells and Lysate preparation

17Cl-1 cells and lysates were prepared as described (Yu and Leibowitz, 1995).

2.2 Mutagenic PCR and Transcription

DE25 served as a template for amplifying and incorporating a T7 promoter into the 3' terminal 42 nucleotide host protein binding element by PCR. The resulting template was used for *in vitro* transcription reactions containing [α^{32}P]-UTP. A portion of each transcription reaction was analyzed by electrophoresis.

2.3 Gel Mobility Shift RNase T1 Protection Assays

Gel mobility shift RNase T_1 protection assays were performed as described, except that 1.4 pmoles of each probe was used (Yu and Leibowitz, 1995). Competitive concentrations of unlabelled wild type probe as well as tRNA were used at 10, 25, 50, 100 fold molar excess to ensure specificity. Assays were quantitated by phosphorimagery. The binding activity of wild type RNA was set at 100%, and after arithmetically correcting for differences in base composition the binding activities of mutants were calculated as a percentage of wild type.

2.4 Computer Modelling

Computer modelling of the 3'42nt host protein binding element was carried out with Mfold, Version 3.0, available at http://www.ibc.wustl.edu/~zuker/rna/node3.html (Zuker et al. 1999).

3. RESULTS

3.1 Gel Mobility Shift RNase T_1 Protection Assays

Mutagenesis of the 11nt motif contained in the last 42nts confirmed this motif's importance for host protein binding and DI replication (Yu and

Leibowitz, 1995). Computer models of the 3' (+)42 protein binding element predicted the formation of a small stem loop (Yu and Leibowitz, 1995). To determine if RNA secondary structure is important for host protein binding to the 3'(+)42 element we performed a series of gel mobility shift RNase T_1 protection assays with RNA probes containing mutations designed to disrupt this predicted stem loop structure. Mutants MT1A, MT2A, MT3C, MT5A, and MD10 were generated by mutagenic PCR and of these mutants, MT3C had the greatest effect on host protein binding. Mutants MT1A, MT2A, MT3C, and MT5A all carry mutations in the 11nt motif (Table 1). The mutations in MT1A and MT2A decreased their respective binding efficiencies to 68.5% and 64% relative to wild type RNA. Mfold, version 3.0, predicted that these mutations would not change the RNA secondary structure (Fig. 1). Mutant MT3C had a binding activity only 6.3% of wild type, while the mutation in MT5A resulted in a binding activity 25.2% of wild type RNA (Table 1). Both of these mutant RNAs were predicted to have dramatic changes in secondary structure from wild type (Fig. 2). For mutant MD10 the last 10 nucleotides were deleted. Deletion of these nucleotides decreased protein binding activity to 36.1% of wild type. Yu and Leibowitz's computer model suggested that the 10 terminal nucleotides were single stranded (Yu and Leibowitz, 1995). Secondary structure modelling with the current Mfold, version 3.0, suggests that the 10 terminal nucleotides are involved in secondary structure formation which may explain the decrease in binding.

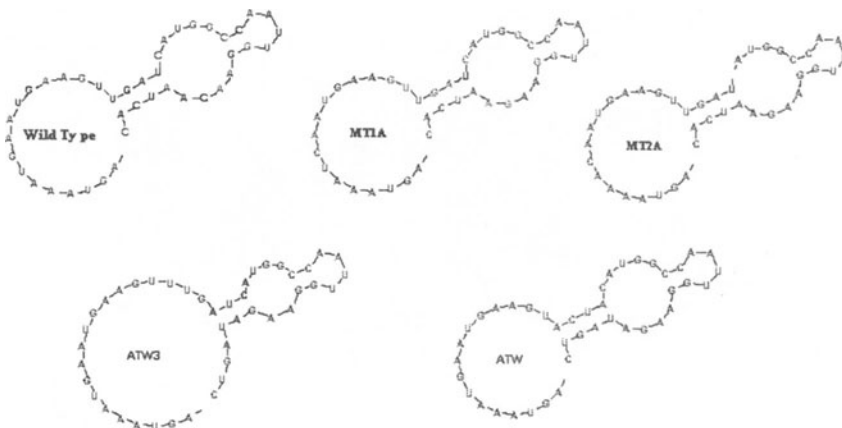

Figure 1. Secondary Structure models for RNA probes with minimal loss in binding activity or increased binding activity. The black dot indicates the 5' terminus of each molecule.

Table 1. Summary of binding activity for all probes.

Probe	Sequence	% Binding	Std Dev	Effect on 2° Structure
Wild type	AGUAAAUGAAUGAAGUUGAUC AUGGCCAAUUGGAAGAAUCAC	100	N/A	N/A
MT1A	AGUAAACUAAUGAAGUUGAUC AUGGCCAAUUGGAAGAAUCAC	68.5	11.3	NONE
MT2A	AGUAAA**AC**AAUGAAGUUGAUC AUGGCCAAUUGGAAGAAUCAC	64	21.7	NONE
MT3C	AGUAAA**ACU**AUGAAGUUGAUC AUGGCCAAUUGGAAGAAUCAC	6.3	6.1	DISRUPTS
MT5A	AGUAAA**ACUUA**GAAGUUGAUC AUGGCCAAUUGGAAGAAUCAC	25.2	5.5	DISRUPTS
MD10	AGUAAAUGAAUGAAGUUGAUC AUGGCCAAUUG	44.1	2.5	DISRUPTS
M24C	AGUAAA**ACU**AUGAAGUUGAUC AUCGCCAAUUGGAAGAAUCAC	74.1	8.6	RESTORES
24C	AGUAAAUGAAUGAAGUUGAUC AUCGCCAAUUGGAAGAAUCAC	73.5	25.4	NONE
ATW5'	AGUAAAUGAAUGAAGU**ACUA**C AUGGCCAAUUGGAAGAAUCAC	7.2	8.7	DISRUPTS
ATW3'	AGUAAAUGAAUGAAGUUGAUC AUGGCCAAUUGGAAGA**UAGU**C	400	84.4	DISRUPTS
ATW	AGUAAAUGAAUGAAGU**ACUA**C AUGGCCAAUUGGAAGA**UAGU**C	185.8	34.7	RESTORES

If primary structure of the 11nt host protein binding motif is all that is required for protein binding, then additional mutations outside of the 11nt motif should have minimal effects on protein binding. If nucleotides outside of the motif also play some role in the RNA:protein interaction, mutating these nucleotides should effect binding. If a specific secondary structure is required for protein binding, then restoration of secondary structure in the presence of a deleterious mutation in the conserved motif should restore host protein binding to near wild type levels. Computer modelling generated compensatory sequence alterations that would restore the wild type secondary structure in the presence of the original deleterious mutations. A compensatory mutation for MT3C was found; changing nucleotide 19 from G to C restored predicted wild type secondary structure (Figs. 1 and 2). This probe, M24C, had a host protein binding activity 74.1% of wild type compared to 6.3% of wild type for MT3C. For probe 24C, a 19 G to C mutation was introduced alone. This mutation lies outside of the host protein binding motif, did not disrupt secondary structure, and had a binding activity 73.5% of wild type (Table 1, Fig. 1).

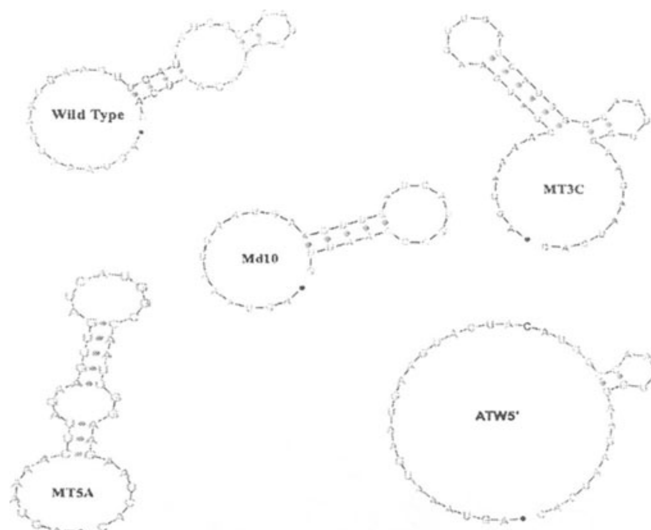

Figure 2. Secondary structure models of RNA probes that did not bind near wild type levels. The black dot indicates the 5' terminus of each molecule.

The computer predicted wild type secondary structure indicated base paring in two regions. Base pairing occurs between nucleotides 2-5 and 23-26, with a second base pairing occurring between nucleotides 10-11 and 16-17. To investigate the role of the base pairings between 2-5 and 23-26 we created the mutants ATW5', ATW3', and ATW (Figs. 1 and 2). When these structures were evaluated for changes in secondary structure, ATW maintained the wild type structure, ATW3' slightly altered its base pairing; ATW5' was almost completely single stranded except for the 10-11 to 16-17 base pairing. Surprisingly, ATW3' host protein binding activity was 4-fold greater than wild type RNA. ATW probe host protein binding ability also increased relative to wild type but not as dramatically, 1.85 fold (Table 1).

Base pairing at positions 10-11 with 16-17 was maintained in wild type, MT1A, MT2A, ATW, ATW3, M24C, and MT3C mutants. Mutations in the 10-11 and 16-17 region were examined with Mfold but compensatory mutations which restored wild type secondary structure could not be identified. Base pairings at positions 10-11 with positions 16-17 in wild type probe, probes with little binding activity, and probes with increased binding activity suggests that this pairing while conserved, is not sufficient for host-protein binding.

4. DISCUSSION

We have shown that secondary structure plays an important role in binding of host proteins to the 3'(+)42 host protein binding element. Eleven probes were analyzed, and four of these probes, MT3C, MT5A, ATW5', and MD10, had a 60% or greater reduction in binding activity and also had significant alterations in their predicted secondary structures. Three probes MT1A, MT2A, and 24C had no predicted changes in secondary structure compared to wild type and had slightly decreased levels of host protein binding activity. M24C, contained the deleterious MT3C mutations along plus compensatory mutation which restored the predicted wild type secondary structure and protein binding activity, thus highlighting the role of secondary structure in host proteins binding to 3'(+)42 RNA.

ATW5' RNA has a predicted secondary structure radically different from the wild type structure and little protein binding activity. The predicted secondary structure of the ATW3' RNA is a variation of the wild type RNA structure (Fig. 1) and its host cell protein binding activity is four-fold greater than wild type. When the ATW3' and ATW5' mutations are placed in the same probe, the secondary structure of the RNA is predicted to match the wild type structure. The protein binding activity of the double mutant RNA is only 1.85-fold greater than wild type, an activity intermediate to that of the single mutations. These data are consistent with the view that secondary structure is a major determinant of protein binding activity of the 3'(+)42 RNA, although primary structure also makes an important contribution.

Non-translated sequences in the MHV genome such as the 5' leader and 3'UTR are essential for viral replication. Deletion studies have shown that *cis*-acting sequences in the last 447 nucleotides of the genome plus the poly (A) tail are essential for replication (Kim et al, 1993; Lin et al., 1993). The 3'(+)42 RNA lies within this region and is contained within the 3' 55 nucleotides required for minus strand RNA synthesis (Lin et al., 1994). Studies in our lab have shown that mutagenesis of the 11nt UGAARNGAAGUU motif within the 3'(+) 42 RNA which disrupt RNA-protein interaction also have a deleterious effect on DI replication(Yu and Leibowitz, 1995). The data presented here indicate that RNA secondary structure plays an essential role in the interaction of the 3'(+) 42 RNA with host proteins. Replication studies with the mutants generated in this work would test the hypothesis that these RNA secondary structures in the 3' UTR are important for replication. Studies of the two mutants with enhanced protein binding activity are likely to be particularly interesting. Such studies are currently under way in our lab.

ACKNOWLEDGEMENTS

This work was supported by National Multiple Sclerosis Society grant RG2203-B-6 and a gift from the Stearman family. The authors would like to thank Santosh Nanda, Qi Liu, and Elena Belyavskaya for all their help.

REFERENCES

Y.-N.Kim, Y.S.Jeong, and S.Makino, Analysis of cis-acting sequences essential for coronavirus defective interfering RNA replication., Virology 197:53 (1993).

Y.-J.Lin and M.M.C.Lai, Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinous sequence for replication., J.Virol. 67:6110 (1993).

Y.Lin, C.Liao, and M.M.C.Lai, Identification of the cis-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription , J.Virol. 68:8131 (1994).

Q.Liu, W.Yu, and J.L.Leibowitz, A specific host cellular protein binding element near the 3' end of mouse hepatitis virus genomic RNA., Virology 232:74 (1997).

W.Yu and J.L.Leibowitz, A conserved motif at the 3' end of mouse hepatitis virus genomic RNA required for host protein binding and viral RNA replication., Virology 214: 128 (1995).

W.Yu and J.L.Leibowitz, Specific binding of host cellular proteins to multiple sites within the 3' end of mouse hepatitis virus genomic RNA., J.Virol. 69:5033 (1995).

M.Zuker, D.H.Mathews, and D.H.Turner, Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide, *in:* "RNA Biochemistry and Biotechnology", J.Barciszewski and B.F.C.Clark, eds., Kluwer Academic Publishers, New York (1999).