THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 265, No. 33, Issue of November 25, pp. 20117-20122, 1990
Printed in U.S.A.
© 1990 by The American Society for Biochemistry and Molecular Biology, Inc.

# Sequence Specificity of mRNA $N^6$-Adenosine Methyltransferase*

(Received for publication, May 23, 1990)

**Tunde Csepany‡, Alison Lin§, Carl J. Baldick, Jr., and Karen Beemon¶**
*From the Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218*

**The sequence specificity of chicken mRNA $N^6$-adenosine methyltransferase has been investigated *in vivo*. Localization of six new $N^6$-methyladenosine sites on Rous sarcoma virus (RSV) virion RNA has confirmed our extended consensus sequence for methylation: RGACU, where R is usually a G (7/12). We have also observed A (2/12) and U (3/12) at the −2 position (relative to m$^6$A at +1) but never a C. At the +3 position, the U was observed 10/12 times; an A and a C were observed once each in weakly methylated sequences. The extent of methylation varied between the different sites up to a maximum of about 90%. To test the significance of this consensus sequence, it was altered by site-specific mutagenesis, and methylation was assayed after transfection of mutated RSV DNA into chicken embryo fibroblasts. We found that changing the G at −1 or the U at +3 to any other residue inhibited methylation. However, inhibition of methylation at all four of the major sites in the RSV *src* gene did not detectably alter the steady-state levels of the three viral RNA species or viral infectivity. Additional mutants that inactivated the *src* protein kinase activity produced less virus and exhibited relatively less *src* mRNA in infected cells.**

$N^6$-Methyladenosine (m$^6$A)[1] is the major methylated residue found at internal sites in mRNAs of higher eukaryotes (reviewed in Banerjee (1980)). Pre-mRNAs are methylated in the nucleus prior to polyadenylation (Perry *et al.*, 1975), and the bulk of the methylated bases appears to be conserved during processing to mature cytoplasmic mRNA (Chen-Kiang *et al.*, 1979). Studies in a number of different cell types have demonstrated that internal m$^6$A methylation occurs exclusively at the sequences GAC or AAC, with GAC being more frequently used (Canaani *et al.*, 1979; Dimock and Stoltzfus, 1977; Schibler *et al.*, 1977; Wei and Moss, 1977). GAC and AAC should appear every 32 nucleotides in a random sequence; however, only one m$^6$A residue is seen per approximately 1000 nucleotides in total mammalian cell mRNA (Adams and Cory, 1975; Desrosiers *et al.*, 1975; Lavi *et al.*, 1977; Perry *et al.*, 1975; Wei *et al.*, 1976). This suggests either that methylation occurs inefficiently at many sites or that there is additional enzyme specificity for methylation, perhaps at the level of RNA secondary structure or of more extensive sequence specificity. Some effort has been made to identify conserved nucleotides flanking the RAC consensus sequence. At the −2 position (2 bases upstream of the m$^6$A at +1), a purine was observed 90% of the time in mouse mRNA and hnRNA, and A, U, or C was observed at the +3 position (Schibler *et al.*, 1977). Similarly, in HeLa cell mRNA, the base following m$^6$AC is never G (Wei and Moss, 1977). RGACU was observed to be a preferred methylation sequence in Rous sarcoma virus (RSV) RNA on the basis of comparison of seven m$^6$A sites; however, one of these sites had a U at −2 (Kane and Beemon, 1985).

If m$^6$A residues were distributed uniformly between mRNAs, an average mammalian cell mRNA would have approximately three m$^6$A residues (Adams and Cory, 1975; Desrosiers *et al.*, 1975; Lavi *et al.*, 1977; Perry *et al.*, 1975; Wei *et al.*, 1976). The distribution of m$^6$A residues in specific mRNAs has been studied in only a few cases; however, it is definitely nonrandom. Both the number and positions of m$^6$A residues vary with different mRNAs. Most of the m$^6$A in adenovirus is found within the 5′ two-thirds of the viral mRNAs (Sommer *et al.*, 1976), and simian virus 40 late mRNAs are methylated near their 5′-ends (Canaani *et al.*, 1979). In contrast, bovine prolactin mRNA is methylated at a single site which is within its 3′-noncoding region (Horowitz *et al.*, 1984; Narayan and Rottman, 1988). RSV has from 12 to 15 m$^6$A residues per genomic RNA subunit, on average (Beemon and Keith, 1977; Furuichi *et al.*, 1975; Stoltzfus and Dimock, 1976), predominantly in the 3′-half of the virion RNA (Beemon and Keith, 1977; Kane and Beemon, 1985). Precise localization on the RSV genome has identified 3 m$^6$A residues in *src* and 4 in *env* coding sequences (Kane and Beemon, 1985). The extent of methylation varies between different sites in RSV RNA, with some being nearly stoichiometric while others are only rarely used (Dimock and Stoltzfus, 1977; Kane and Beemon, 1985). There are numerous RGACU sequences in the 5′-half and some in the 3′-half of RSV RNA which are not detectably methylated (Beemon and Keith, 1977; Kane and Beemon, 1985).

In this paper, we have extended our previous localization of methylated bases in RSV virion RNA at the nucleotide level and have identified six additional m$^6$A sites in the *src* coding region of RSV. Comparison of sequences containing these sites has confirmed our extended consensus sequence for methylation: RGACU. Site-specific mutagenesis of the G at −1 or of the U at +3 in this consensus sequence inhibited methylation, as did mutation of the invariant C at +2 in an earlier study (Kane and Beemon, 1987). However, mutation of the four major m$^6$A sites in the *src* gene of RSV did not have an appreciable effect on steady-state levels of viral RNA nor on viral infectivity.

## EXPERIMENTAL PROCEDURES

*Materials*—CEF primary cultures were prepared from gs$^-$ chf$^-$ eggs supplied by SPAFAS, Inc., Norwich, CT. A full-length clone of PR-

* This work was supported by National Institutes of Health Grant CA33199 from the National Cancer Institute. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Dept. of Neurology and Psychiatry, Medical University of Debrecen, Debrecen H-4012, Hungary.
§ Present address: Baylor College of Medicine, Houston, TX 77030.
¶ To whom correspondence should be sent.

[1] The abbreviations used are: m$^6$A, $N^6$-methyladenosine; RSV, Rous sarcoma virus; CEF, chicken embryo fibroblast; PR-C, Prague strain, subgroup C, RSV.

C RSV DNA in pBR322 (pATV8K) (Cobrinik *et al.*, 1987) was a gift from R. Katz and A. Skalka, Fox Chase Cancer Center. Synthetic oligonucleotides were either obtained from Operon, Alameda, CA, or were synthesized on an Applied Biosystems model 380A synthesizer. [$^{32}$P]Orthophosphate was obtained from ICN. [$\gamma$-$^{32}$P]ATP (3000 Ci/mmol) and [$\alpha$-$^{32}$P]CTP (400–800 Ci/mmol) were from Amersham Corp. RNases A, $T_1$, $T_2$ were obtained from Calbiochem; P1 nuclease was from Bethesda Research Laboratories; T4 polynucleotide kinase was from U. S. Biochemical Corp.; T7 RNA polymerase was from Pharmacia LKB Biotechnology Inc. Marker pm$^6$A was from P-L Biochemicals. RNAzol was obtained from Cinna/Biotecx. Plasmid pGEM-2 was from Promega Biotec.

*Site-specific Mutagenesis*—Site-specific mutagenesis was carried out to alter sequences at m$^6$A sites by the method of Zoller and Smith (1984) as previously described (Kane and Beemon, 1987). Mutated DNA fragments were sequenced by the method of Sanger *et al.* (1977), as adapted for plasmid DNA (Wallace *et al.*, 1981). The mutated *SacI* fragment was then inserted into the complete RSV genomic clone, pATV8K (Cobrinik *et al.*, 1987). Mutations in the consensus sequence flanking the m$^6$A at site 8014 (numbering is according to the sequence of Schwartz *et al.* (1983)) were made in a plasmid bearing previous mutations (7415U/7425U) that prevented adenine methylation at nucleotides 7414 and 7424 (Kane and Beemon, 1987); thus three methylation sites were mutated in all. Mutation 8016G changed the U at +3 to a G but did not alter the amino acid sequence of pp60$^{src}$. In contrast, mutation 8013U/8016A was a double mutation, which altered amino acid 295 of the *src* protein from Lys to Asn. The resulting mutant failed to transform CEFs, as has been observed with other mutations which alter this amino acid at the ATP-binding site of *src* (Kamps and Sefton, 1986). Fourth site mutations (at m$^6$A site 8485) were originally constructed in a plasmid bearing the 8013U/8016A and 7415U/7425U mutations; all of the viruses generated were nontransforming. RNA from these viruses was analyzed to determine the effects of the mutations on methylation (Table II). However, for functional characterization of methylation-deficient viruses, mutants 8484U and 8487G were reconstructed on an 8016G, 7415U/7425U background by replacing the *StuI* fragment (nucleotide 3316–8042) of these viral constructs with the corresponding fragment from 8016G. This yielded transforming viruses with four mutations at m$^6$A sites. These transforming viruses were used in studies regarding the possible functional significance of m$^6$A (Table III, Fig. 2).

*Transfection*—Plasmids containing viral DNA were permuted around a unique *KpnI* site (Cobrinik *et al.*, 1987). Before transfection, the plasmids were cut with *KpnI* and religated to generate concatamers. One $\mu$g of DNA was transfected onto chick embryo fibroblast cells with DEAE-dextran, followed by a shock with 10% dimethyl sulfoxide as described (Arrigo and Beemon, 1988).

*Preparation and Analysis of Viral RNA*—Viral RNA was prepared 7–14 days after infection of CEFs. Radiolabeling of infected cells with $^{32}P_i$ and isolation of genomic RNA from virus particles were as previously described (Kane and Beemon, 1985). For isolation of 35 S virion RNA subunits, the 70 S RNA was denatured by boiling for 45 s and was sedimented through a second 10–25% sucrose gradient. The total m$^6$A content of RNA was determined by digestion with RNases A, $T_1$, and $T_2$, followed by two-dimensional electrophoretic separation of the mononucleotides as described previously (Beemon and Keith, 1977). Fragments of genomic RNA were isolated by hybridization to nitrocellulose filters to which viral DNA restriction fragments had been bound (Kane and Beemon, 1985). RNAs were digested with RNase $T_1$ and fingerprinted by the method of Barrell (1971) using electrophoresis on DEAE-cellulose paper in the second dimension. To localize m$^6$A sites specifically, selected oligonucleotides were eluted from the fingerprint and digested to mononucleotides with RNases A, $T_1$, and $T_2$. The resulting mononucleotides were separated by electrophoresis on Whatman No. 3MM paper in pH 3.5 buffer (pyridine/acetic acid). The Ap was eluted from the paper and separated from m$^6$Ap by thin-layer chromatography on cellulose plates as described earlier (Kane and Beemon, 1985).

This same procedure was used to analyze RNA fragments containing mutated m$^6$A sites. Isolation of a 149-nucleotide RNA fragment, by hybridization to a *MluI-PstI* RSV DNA fragment (nucleotide 7901–8050), allowed analysis of site 8014; and a 246-nucleotide fragment, obtained by hybridization to a *BglI-PstI* fragment (nucleotide 8416–8662), was used to analyze site 8485. To confirm analyses of fragments containing mutated m$^6$A sites, unlabeled viral RNA fragments were digested to 3′-monophosphates with RNases A, $T_1$, and $T_2$ and 5′-end-labeled with T4 polynucleotide kinase and 1 $\mu$Ci of [$\gamma$-$^{32}$P]ATP. After removal of the 3′-phosphate with P1 nuclease, the 5′-monophosphates were resolved by two-dimensional thin-layer chromatography as described (Horowitz *et al.*, 1984). The presence of pm$^6$A was detected by comigration with a cold marker nucleotide.

*RNase Protection Assays*—The levels of *src* and *env* mRNAs and unspliced RNAs were detected by RNase protection assays. Total cellular RNA was prepared after cell lysis in RNAzol as described (Chomczynski and Sacchi, 1987). RNA species were detected by RNase protection assays (Melton *et al.*, 1984). The riboprobe template for assay of *env* mRNA was constructed by insertion of a *BamHI* (nucleotide 4716) to *XhoI* (nucleotide 5258) fragment of PR-C RSV into pGEM-2. After template digestion with *DdeI* (nucleotide 5045), the antisense riboprobe was transcribed with T7 RNA polymerase. 50 $\mu$Ci of [$\alpha$-$^{32}$P]CTP were used in this reaction at a final CTP concentration of 15 $\mu$M. Hybridization with this probe yielded a 213-nucleotide unspliced RNA and a 180-nucleotide spliced *env* mRNA. Splicing at the *src* splice acceptor site was assayed with a riboprobe made from a template spanning RSV sites *XhoI* (nucleotide 6983) to *Tth111-I* (nucleotide 7330). Hybridization of this probe to RNAs from infected cells yielded a 347-nucleotide RNA which included unspliced RNA and *env* mRNA, and a 276-nucleotide *src*-specific mRNA. Hybridization was carried out for approximately 18 h at 55 °C. RNase digestion was in 10 $\mu$g/ml RNase A and 10 units/ml RNase $T_1$ at 23 °C for 1 h.

*Reverse Transcriptase Assays*—The clarified cell culture media was assayed for viral reverse transcriptase activity as a measure of virus production as described (Kane and Beemon, 1987).

## RESULTS

*Sequence Specificity of Methylation in RSV RNA and Chicken mRNA*—The average m$^6$Ap content of Prague C RSV 35 S virion RNA was determined by nucleotide analysis to be 0.14%, using *in vivo* $^{32}$P-labeled RNA (data not shown). This suggested an average of 12.6 m$^6$Ap residues/9.5-kilobase RNA, in good agreement with the level previously observed for Prague B RSV RNA (Beemon and Keith, 1977). Individual RNase $T_1$ oligonucleotides from a homochromatography fingerprint of total 35 S RSV RNA were also analyzed for their m$^6$Ap content. This study detected 18 different methylated oligonucleotides, which varied extensively in their extent of methylation. Since most of these oligonucleotides are not unique in the viral genome, more than 18 potential m$^6$A sites are possible. As methylation was previously observed to be incomplete at most sites (Kane and Beemon, 1985), it was expected that more than 12 methylated sites would be found in RSV RNA to generate an average of 12.6 m$^6$A residues.

The oligonucleotide containing the highest fraction of m$^6$Ap (relative to Ap) in RSV 35 S RNA was (G)ACUG (data not shown). This sequence contained 4–5 times as much m$^6$Ap as the next highest set of methylated sequences: (G)ACUUUN and (G)ACUUG. A lower level of m$^6$Ap content was observed in a variety of oligonucleotides beginning (G)ACUC/A. A barely detectable level of m$^6$Ap was observed in several oligonucleotides beginning (G)ACA/C. Finally, no m$^6$A was observed in (G)ACG, (G)ACCG, or (G)ACCCG. The differences observed in the average extent of methylation of these different oligonucleotides may reflect different affinities of the methyltransferase for the sequences contained in the oligonucleotide. Alternatively, a single oligonucleotide may be derived from a mixture of extended sequences, each having different affinities for the methyltransferase.

A similar m$^6$Ap analysis was carried out on total polyadenylated RNA from chicken embryo fibroblasts. We were surprised to find that the overall level of m$^6$Ap in total cellular polyadenylated RNA was 0.055%, which is amost 3 times lower than that of the virion RNA. Analysis of intracellular viral RNAs also showed a lower level of m$^6$Ap than did the corresponding regions in virion RNA.[2] As in virion RNA, the cellular mRNA oligonucleotide found to contain the most

[2] T. Csepany, unpublished results.

$m^6Ap$ was (G)ACUG, followed by (G)ACUUUN. Lower methylation was detected in (G)ACUUG, (G)ACUAAG, and (G)ACCUG, and no detectable $m^6Ap$ was observed in (G)ACG, (G)ACCG, (G)ACCCG, (G)ACCCAG, or (G)ACACCG (data not shown). We did not detect $m^6A$ in AAC sequences unambiguously by this method; however, there were several potential AAC substrates (in oligonucleotides that also contained GAC), which were methylated at a low level. In conclusion, the sequence specificity of the methyltransferase appeared to be similar for viral and cellular mRNAs, but the extent of methylation was significantly higher for RSV virion RNA than for cellular mRNA.

*Precise Localization of $m^6A$ Sites in RSV RNA*—To study the sequence specificity of the $m^6A$ methyltransferase within a larger context than a T1 oligonucleotide, we have extended our previous studies on localization of $m^6A$ in RSV RNA (Kane and Beemon, 1985). The region of RSV RNA from nucleotide 4715 to 9312 (the 3′ terminus) has been found to contain the bulk of the methylated bases, and a detailed analysis of RSV RNA regions between nucleotide 4715–5258 and nucleotide 6185–8050 was carried out previously (Beemon and Keith, 1977; Kane and Beemon, 1985). In the present study we analyzed RSV RNA regions between nucleotides 5564 and 6185 and between 8051 and 8662. The region from nucleotide 5564 to 6185 is within the *env* coding region and contains several potential $m^6A$ sites (GAC or AAC), including one at nucleotide 5849 that fits our extended consensus sequence RGACU. However, analysis of these potential sites failed to identify any methylated sites in this region.

In contrast, we identified five new $m^6A$ sites in the *src* gene region of the genomic RNA between nucleotide 8051 and 8662 (Table I). In addition to the four sites localized precisely, methylation was observed at nucleotide 8366 and/or 8369 in the sequence AGGACAACG. We have not determined if one of the A residues shown in bold type is the preferred methylation site or if both the GAC and the AAC sites in this oligonucleotide can serve as substrates for the methyltransferase. In addition, we were able to resolve the ambiguity at $m^6A$ site 7890/7981 (Kane and Beemon, 1985) by analyzing a smaller RNA fragment. A low level of methylation was detected in the RNA fragment extending from nucleotide 8662 to 9180, but methylation sites were not definitively identified within it. The fragment containing nucleotides 9181–9312 and nucleotides 1–255 had no detectable $m^6A$ residues. We observed extensive variation in the extent of methylation between these different sites, as summarized in Table I.

By comparing the 12 different methylated sequences precisely localized in RSV RNA (Table I), we noticed that the GAC sequence was invariant. Further, a U was present at position +3 in 10 out of 12 sites. The exceptions at nucleotide 8413 and 8633 had an A or a C at +3, respectively, and both of these sites were weakly methylated. At position −2, we observed a G 7 times, an A 2 times, and a U 3 times. We have never observed a C at position −2. At positions −3 and +4, the sequence appeared to be random. In conclusion, there seems to be a strong preferred sequence of methylation: RGACU, although the R at −2 can be replaced by a U. Analysis of the most highly methylated RSV sequences shows three GACUG sequences and one each of GACUUG and GACUCUG. This is consistent with the analysis of total viral RNA described above, which found that GACUG was the most highly methylated sequence, followed by GACUUUN and GACUUG.

*Mutagenesis of the $m^6A$ Consensus Sequence*—To assess the significance of various nucleotides within the consensus sequence to recognition by an $m^6A$ methyltransferase, we performed site-specific mutagenesis on the RSV genome. We hoped these mutations would also help to address the functional significance of $m^6A$ methylation, so we chose to mutate various parts of the consensus methylation sequences of the four strong $m^6A$ sites localized in the *src* gene: nucleotides 7414, 7424, 8014, and 8485. A double viral mutant had been constructed previously in which $m^6A$ sites 7414 and 7424 were simultaneously made methylation-minus by mutating the two Cs at the +2 positions to U (7415U/7425U) (Kane and Beemon, 1987). This mutant was additionally mutated at $m^6A$ site 8014 to construct mutants 8016G (+3 U to G) and 8013U/8016A. These three-site mutants were additionally mutated at site 8485 to generate a series of six different four-site mutants. The G at −1 (nucleotide 8484) was changed to A, U, or C; alternatively, the U at +3 (nucleotide 8487) was changed to A, G, or C. All of these mutations are described in Table II.

Mutated viral DNAs were used to transfect CEFs, and virion RNA was harvested. We initially determined the presence or absence of $m^6A$ in short RNA fragments known to contain a single $m^6A$ site or a mutated version of this site by analyzing *in vivo* labeled RNA (Kane and Beemon, 1985). The major drawback of the *in vivo* labeling method was the difficulty in obtaining sufficient radiolabeled RNA to carry out the multiple steps required for $m^6A$ analysis. This was particularly true for the nontransforming mutant viruses, which were expressed at a lower level than the wild type. Therefore, these mutants were also analyzed by an *in vitro* labeling method (Horowitz *et al.*, 1984). A representative two-dimensional separation of *in vitro* labeled RNA fragments is shown in Fig. 1. The wild-type 246-nucleotide RNA fragment, which contains a single $m^6A$ site at nucleotide 8485 (*panel A*) was compared with the corresponding fragment derived from mutant 8484A in which GAC has been converted to AAC (*panel B*). The mutant clearly has a greatly reduced level of $m^6A$ in this fragment.

Mutating the bases at positions −1 (G to A, C or U), +2 (C to U), or +3 (U to A, C or G) all resulted in substantial inhibition of methylation, so that no $m^6A$ was detected at the mutated site (Table II). We have not analyzed effects of other

TABLE I
*Sequence specificity of methylation at $m^6A$ sites*

| $m^6A$ site | Sequence | Methylation extent[a] |
|---|---|---|
| *nt*[b] | | |
| 6394 | GAGACUAG[c,d] | Low |
| 6447 | GGGACUUAUUG[d] | Low |
| 6507 | UUGACUUCUUG[d] | Low |
| 6718 | AGGACUG[d] | High |
| 7414 | UGGACUG[d] | High |
| 7424 | CGGACUUG[d] | High |
| 7890 | GGGACUCG | Low |
| 8014 | AAGACUCUG[d] | High |
| 8339 | CUGACUUCG | Low |
| 8366 | AGGACAACG[e] | Low |
| 8413 | UGGACAG | Low |
| 8485 | CUGACUG | High |
| 8633 | AGGACCCUG | Low |
| Consensus sequence | RGACU | |

[a] High methylation denotes 50–90% methylation at a specific site; low methylation means approximately 20–30% of all RSV genomic RNAs are methylated at this nucleotide.

[b] Nucleotides.

[c] The methylated nucleotide is shown in bold type.

[d] These sites were previously identified in Kane and Beemon (1985).

[e] We have not determined which A is methylated in this oligonucleotide.

TABLE II

*Mutations at m⁶A sites in RSV RNA*

Mutations at the 8014 site were made in constructs previously mutated at both nucleotides 7414 and 7424. Similarly, the 8485 site mutations were made in constructs previously mutated at nucleotides 7414, 7424, and 8014.

| Mutation | m⁶A site | Sequence | Mutation position | Methylation |
|---|---|---|---|---|
| None | 7414 | GGACU | None | + |
| 7415U[a] | | GGAU**U**[b] | +2 C→U | − |
| None | 7424 | GGACU | None | + |
| 7425U[a] | | GGAU**U**[b] | +2 C→U | − |
| None | 8014 | AGACU | None | + |
| 8016G | | AGAC**G** | +3 U→G | − |
| 8013U/8016A | | A**U**AC**A** | −1 G→U/+3 U→A | − |
| None | 8485 | UGACU | None | + |
| 8484A | | U**A**ACU | −1 G→A | − |
| 8484U | | U**U**ACU | −1 G→U | − |
| 8484C | | U**C**ACU | −1 G→C | − |
| 8487G | | UGAC**G** | +3 U→G | − |
| 8487A | | UGAC**A** | +3 U→A | − |
| 8487C | | UGAC**C** | +3 U→C | − |

[a] This mutant was described in Kane and Beemon (1987).
[b] Mutated nucleotides are shown in bold type.

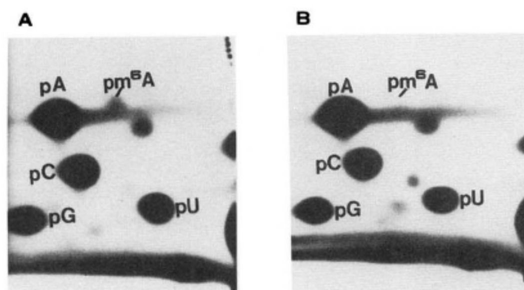


FIG. 1. **Analysis of m⁶Ap at site 8485 in wild-type RSV and mutant 8484A.** The presence of m⁶Ap in a 246-nucleotide fragment containing a single m⁶A site at nucleotide 8485 was investigated by the *in vitro* labeling method described under "Experimental Procedures." This involved two-dimensional thin-layer chromatography. The first dimension is from bottom to top, and the second is from left to right. *A*, wild-type virus; *B*, mutant 8484A.

mutations at these sites nor of mutations at positions +1 (the methylated A is presumed to be invariant) or at −2. It was surprising that all of the mutations at site +3 were inhibitory, since GACAG sequences (as found in mutant 8487A) have been observed to be methylated in total viral and cellular RNA, albeit weakly. In contrast, GACG (as in mutant 8487G) has never been observed to be methylated in any system. We have not observed methylation of GACCG sequences (as in mutant 8487C), although other GACC sequences are methylated. These particular mutations were made at site 8485 (wild-type sequence, UGACUG), which has a nonpreferred U nucleotide at position −2 and a G at +4. It is possible that the +3 mutations would have had a less inhibitory effect if combined with a G at position −2 or with a different residue at +4. We were also surprised that conversion of GAC to AAC (mutant 8484A) inhibited methylation (Fig. 1) since AAC is observed to be methylated in total cell RNA, although less frequently than GAC (Canaani *et al.*, 1979; Dimock and Stoltzfus, 1977; Schibler *et al.*, 1977; Wei and Moss, 1977). Perhaps this mutation would have had less effect in a different context than that at 8485. Alternatively, the sequence context required for methylation at AAC sites may be quite different than that for GAC sites.

*Functional Analyses of Methylation Mutants*—We have begun to assess the effect of these methylation-minus mutations on RSV RNA metabolism. Our earliest mutagenesis studies involved mutations at two sites: 7414 and 7424. This double mutation did not have any observable effect on RSV infectivity, focus formation, intracellular RNA levels, or their subcellular location (Kane and Beemon, 1987). From this original double mutation, additional mutations have been added at m⁶A sites 8014 and 8485, generating viruses with three and four mutations in the *src* gene. These mutations were on a background which preserved transforming activity of the virus (8016G). Steady-state levels of unspliced, *env*, and *src* mRNAs in CEFs transfected with the viral mutants were quantified by an RNase protection assay, using riboprobes that span the *env* and the *src* splice acceptor sites. Fig. 2 shows the results of an experiment using two of the four-site mutants and one of the three-site mutants. The levels of *env* and *src* mRNAs relative to unspliced RNAs were determined by densitometric scanning of the autoradiograms. Cells infected by these transforming viruses lacking three or even all four of the major m⁶A sites in the *src* gene generated steady-state levels of viral *src*, *env*, and unspliced RNAs that were indistinguishable from the wild-type levels.

The infectivity of the various mutant-transforming viruses was also studied by means of reverse transcriptase assays. Cells infected with equivalent reverse transcriptase units of virus from methylation-defective mutants were found to produce levels of virus comparable to those of wild-type virus (Table III). Thus, no effect of the methylation inhibition was observed at the level of viral replication or infectivity.

In contrast, analysis of virion reverse transcriptase and

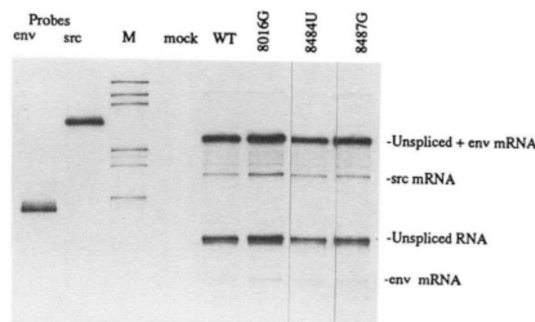


FIG. 2. **RSV RNAs in chicken cells infected with wild-type (*WT*) or mutant viruses.** The levels of unspliced, *env*, and *src* mRNAs were determined by an RNase protection assay as described under "Experimental Procedures." Probes spanning the *env* and *src* splice acceptor sites were used simultaneously in this assay. *M* represents RNA markers. The four-site mutants 8487G and 8484U also contained mutations 8016G and 7415U/7425U.

TABLE III

*Reverse transcriptase assays*

Duplicate plates of CEFs were infected with equivalent reverse transcriptase units of wild-type or mutant virus obtained by transfection. After the cells became transformed, 50 μl of culture supernatants from each plate were assayed for reverse transcriptase activity in duplicate. Reverse transcriptase activity shown in each experiment is the average of duplicate transfected plates, normalized with respect to wild-type virus.

| Virus | Reverse transcriptase activity | | | Mean |
|---|---|---|---|---|
| | Exp. 1 | Exp. 2 | Exp. 3 | |
| Wild type | 100 | 100 | 100 | 100 |
| 8016G | 115 | 130 | 79 | 108 |
| 8484U[a] | 117 | 109 | 86 | 104 |
| 8487G[a] | 148 | 146 | 108 | 134 |

[a] The four-site mutants 8487G and 8484U also contained mutations 8016G and 7415U/7425U.

viral RNA levels revealed severalfold less virus produced by the transformation-defective mutant 8013U/8016A and by all of the four-site mutants derived from it (data not shown). Interestingly, these mutants all inactivated the *src* protein tyrosine kinase activity by mutating the ATP-binding site (amino acid 295) from Lys to Asn. Similar mutations at this site have been shown to inactivate the transforming ability of the virus (Kamps and Sefton, 1986). When we analyzed the steady-state levels of viral RNA in cells transfected with this transformation-defective class of mutants, we discovered that the *src* mRNA levels were decreased relative to those of the unspliced and *env* mRNAs (Csepany *et al.*, 1990). Further, the *src* protein levels were decreased relative to the *gag* protein levels in immunoprecipitates from [$^{35}$S]methionine-labeled cells.[2] These results suggest that cell transformation rather than methylation may be facilitating expression of the *src* mRNA, since analogous mutants which inhibited methylation but did not affect the *src* protein sequence had wild-type levels of *src* mRNAs (Fig. 2).

## DISCUSSION

Localization of 12 m$^6$A residues in the RSV genome has led to the derivation of an extended consensus sequence for m$^6$A, RGACU. While the central GAC sequence was invariant, we have observed some variation from this consensus sequence at positions −2 and +3 (U rather than R at position −2 at three sites, and one site each with a C and an A at position +3). The R at −2 was predominantly G in RSV. The m$^6$A site localized in bovine prolactin RNA was at an AGACU sequence (Narayan and Rottman, 1988), which also fits this consensus. The expected incidence of RGACU would be every 512 nucleotides in a random sequence; thus, it is a considerably better fit than the limited RAC sequence to the actual average distribution of m$^6$A observed in mRNAs (about 1 m$^6$A/1000 nucleotides) (Horowitz *et al.*, 1984; Lavi *et al.*, 1977; Perry *et al.*, 1975; Wei *et al.*, 1976). We have also obtained evidence of variation in the levels of methylation between different sites. In RSV RNA, the RGACU sequence is definitely preferred, but not every such sequence is methylated. The sites that are methylated in viral RNA vary in extent from approximately 20% up to 90%. The m$^6$A sites in RSV RNA appear frequently to be tightly clustered together; for example, four sites were identified between nucleotides 8339 and 8485. It also appears that whole regions of an RNA may not be available for methylation. There are several RGACU sequences in the 5′-half of RSV RNA, but they do not appear to be methylated. Thus, the average number of m$^6$A sites in a specific RNA appears to reflect both incomplete methylation at most sites and an extended consensus sequence. The fact that certain regions of an RNA are highly methylated and others are not suggests the possibility of favorable secondary structures or other factors involved in determining the specificity of methylation.

We have begun to assess the relevance of the RGACU consensus sequence for methylase specificity by performing oligonucleotide-directed mutagenesis. These studies have revealed the importance of the G (−1), C (+2), and U (+3) residues, since methylation was severely inhibited when any one of these bases was mutated. We assume that the methylated A is also obligatory but have not mutated it. As expected based on the consensus sequence, alteration of the C at +2 or changing the G at −1 to a pyrimidine prevented methylation. However, it was surprising that the G to A mutation at nucleotide 8483, which generated the sequence UAACU, also inhibited methylation. Although AAC, as well as GAC, sequences are reportedly methylated in a variety of mRNAs (Dimock and Stoltzfus, 1977; Schibler *et al.*, 1977; Wei and Moss, 1977), we have not identified any unambiguously methylated AAC sites in RSV. It is possible that methylation of an AAC requires a different context than that for GAC. It has not been established whether a single m$^6$A methylase is active at both GAC and AAC sites or whether there are two or more independent methylases with different sequence specificities. An m$^6$A methylase has not been purified; however, methylase activity has been detected in HeLa cell nuclear extracts (Narayan and Rottman, 1988). In RSV RNA methylated GAC sites appear to be much more abundant than AAC sites (Dimock and Stoltzfus, 1977; Kane and Beemon, 1985; this paper). It is possible that the AAC sites are used less efficiently, making them difficult to detect in a specific RNA. The finding that virion RNA was more highly methylated than chicken cell mRNA or viral mRNA raises the possibility that there may also be m$^6$A demethylases in the cell.

Previous studies using methylation inhibitors have suggested a possible role for m$^6$A in RNA processing or transport of mRNA from the nucleus to the cytoplasm (Camper *et al.*, 1984; Finkel and Groner, 1983; Stoltzfus and Dane, 1982). We have generated mutants of RSV that are undermethylated at specific sites so as to assess the functional significance of the methylation. The RSV *src* region has four strong m$^6$A sites as well as numerous weaker methylation sites. We have inhibited methylation at all four of the strong sites but have observed no noticeable effect on steady-state levels of viral RNA. Viral replication and infectivity, measured by reverse transcriptase assays, were also unaffected by undermethylation in the *src* gene.

Only those methylation-minus mutants which inactivated the *src* kinase activity manifested a phenotype different from the wild-type virus. These mutants generated a lower level of reverse transcriptase activity than wild type, possibly due to a lack of *src*-mediated activation of transcription from the RSV long terminal repeat as has recently been seen in serum-starved rat cells (Dutta *et al.*, 1990). In addition, the nontransforming mutants had a lower relative level of *src* mRNA and protein (Csepany *et al.*, 1990). Both of these phenotypes correlated with transformation rather than with methylation. The *src* protein has recently been reported to induce stabilization of cellular 9E3 mRNA (Stoeckle and Hanafusa, 1989); however, enhancement of v-*src* mRNA levels in RSV-transformed cells has not been reported previously. The alterations in *src* mRNA levels in the nontransforming mutants could be at the level of mRNA splicing, transport, or stability.

There are several possible explanations for the lack of apparent effect of the *src* methylation mutations on RNA metabolism. It is possible that methylation facilitates RNA splicing or transport but that the effects are too subtle to be seen at the level of steady-state RNA. Alternatively, m$^6$A sites with a function may represent a minority of the total m$^6$A sites. It is possible that the multiple m$^6$A sites in RSV are somewhat degenerate so that not all need to be methylated for function to be maintained. Since the RSV *src* gene is a transduced cellular gene that has lost its introns (Takeya and Hanafusa, 1983), it is possible that the m$^6$A sites in *src* played a role in processing of the c-*src* progenitor but that they are not involved in v-*src* mRNA metabolism. In conclusion, in this study we were able to learn more about the sequence specificity of the m$^6$A methylase; however, the function of this modification in mRNA remains elusive.

*Acknowledgments*—We thank Susan Kane and Kristen Carlberg for helpful discussions, Susan Kane and Stacie Dalrymple for review-

ing the manuscript, and Sandra Buchman for excellent technical assistance.

## REFERENCES

Adams, J. M., and Cory, S. (1975) *Nature* **255,** 28–33

Arrigo, S., and Beemon. K. (1988) *Mol. Cell. Biol.* **8,** 4858–4867

Banerjee, A. K. (1980) *Microbiol. Rev.* **44,** 175–205

Barrell, B. G. (1971) in *Procedures in Nucleic Acid Research* (Cantoni, G. L., and Davies, D. R., eds) Vol. 2, pp. 751–795, Harper & Row Publishers, Inc., New York

Beemon, K., and Keith, J. (1977) *J. Mol. Biol.* **113,** 165–179

Camper, S. A., Albers, R. J., Coward, J. K., and Rottman, F. M. (1984) *Mol. Cell. Biol.* **4,** 538–543

Canaani, D., Kahana, C., Lavi, S., and Groner, Y. (1979) *Nucleic Acids Res.* **6,** 2879–2899

Chen-Kiang, S., Nevins, J. R., and Darnell, J. E., Jr. (1979) *J. Mol. Biol.* **135,** 733–752

Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159

Cobrinik, D., Katz, R., Terry, R., Skalka, A. M., and Leis, J. (1987) *J. Virol.* **61,** 1999–2008

Csepany, T., Kane, S. E., Lin, A., Baldick, C. J., and Beemon, K. (1990) in *Nucleic Acid Methylation* (Clawson, G. A., Willis, D. B., Weissbach, A., and Jones, P. A., eds) pp. 83–91, Alan R. Liss, Inc., New York

Desrosiers, R. C., Friderici, K. H., and Rottman, F. M. (1975) *Biochemistry* **14,** 4367–4374

Dimock, K., and Stoltzfus, C. M. (1977) *Biochemistry* **16,** 471–478

Dutta, A., Stoeckle, M. Y., and Hanafusa, H. (1990) *Genes & Dev.* **4,** 243–254

Finkel, D., and Groner, Y. (1983) *Virology* **131,** 409–425

Furuichi, Y., Shatkin, A. J., Stavnezer, E., and Bishop, J. M. (1975) *Nature* **257,** 618–620

Horowitz, S., Horowitz, A., Nilsen, T. W., Munns, T. W., and Rottman, F. M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 5667–5671

Kamps, M. P., and Sefton, B. M. (1986) *Mol. Cell. Biol.* **6,** 751–757

Kane, S. E., and Beemon, K. (1985) *Mol. Cell. Biol.* **5,** 2298–2306

Kane, S. E., and Beemon, K. (1987) *J. Biol. Chem.* **262,** 3422–3428

Lavi, U., Fernandez-Munoz, R., and Darnell, J. E. (1977) *Nucleic Acids Res.* **4,** 63–69

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984) *Nucleic Acids Res.* **2,** 7035–7056

Narayan, P., and Rottman, F. M. (1988) *Science* **242,** 1159–1162

Perry, R. P., Kelley, D. E., Friderici, K., and Rottman, F. M. (1975) *Cell* **6,** 13–16

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467

Schibler, U., Kelley, D. E., and Perry, R. P. (1977) *J. Mol. Biol.* **115,** 695–714

Schwartz, D. E., Tizard, R., and Gilbert, W. (1983) *Cell* **38,** 317–331

Sommer, S., Salditt-Georgieff, M., Bachenheimer, S., Darnell, J. E., Furuichi, Y., Morgan, M., and Shatkin, A. J. (1976) *Nucleic Acids Res.* **3,** 749–765

Stoeckle, M. Y., and Hanafusa, H. (1989) *Mol. Cell. Biol.* **9,** 4738–4745

Stoltzfus, C. M., and Dane, R. W. (1982) *J. Virol.* **42,** 918–931

Stoltzfus, C. M., and Dimock, K. (1976) *J. Virol.* **18,** 586–595

Takeya, T., and Hanafusa, H. (1983) *Cell* **32,** 881–890

Wallace, R. B., Johnson, M. J., Suggs, S. V., Miyoshi, K., Bhatt, R., and Hakura, K. (1981) *Gene (Amst.)* **16,** 21–26

Wei, C.-M., and Moss, B. (1977) *Biochemistry* **16,** 1672–1676

Wei, C.-M., Gershowitz, A., and Moss, B. (1976) *Biochemistry* **15,** 397–401

Zoller, M. J., and Smith, M. (1984) *DNA (NY)* **3,** 479–488