Identification and Characterization of the Woodchuck Hepatitis Virus Origin of DNA Replication

CHRISTOPH SEEGER* AND JOHNNA MARAGOS

Department of Microbiology, Immunology and Parasitology, New York State College of Veterinary Medicine, Cornell University, Ithaca, New York 14853-6401

Received 20 July 1989/Accepted 13 September 1989

Replication of the woodchuck hepatitis virus (WHV) genome requires the synthesis of minus-strand DNA from an RNA template, the pregenome, by reverse transcription. During this reaction, the 5' end of minus-strand DNA becomes covalently linked to a protein. The position of the 5' end of minus-strand DNA was identified previously, but the initiation site for DNA synthesis on pregenomic RNA remained ambiguous because of a sequence repetition at the termini of the RNA template for reverse transcription. Employing a recently designed expression vector for the production of infectious WHV, we localized the origin of minus-strand DNA synthesis to the 3' end of pregenomic RNA. In addition, we identified the nucleotide sequences on pregenomic RNA that provide the signal for the initiation of reverse transcription. Removal of this signal sequence from pregenomic RNA abolished minus-strand DNA synthesis. Insertion of a DNA oligomer bearing this signal sequence at the 3' end of pregenomic RNA restored the production of minus-strand DNA joined to protein. Our results support a model in which protein is the primer for reverse transcription of minus-strand DNA of WHV.

Woodchuck hepatitis virus (WHV), a hepadnavirus, replicates its DNA genome by means of an RNA intermediate, termed pregenomic RNA (for reviews, see references 13 and 5). Pregenomic RNAs with redundant termini are transcribed from covalently closed circular DNAs present in the nuclei of infected cells. Subsequently, the pregenome is packaged into core particles and reverse transcribed to yield relaxed circular DNA, the final product of the viral replication cycle. In its basic concept resembling replication of other retroviral elements, the replication of WHV and other hepadnaviruses is unique in that replication requires the formation of a covalent joint between protein and DNA at the 5' end of the minus-strand DNA (6). This characteristic and the finding that short nascent minus-strand DNA species of duck hepatitis B virus are linked to protein led to the hypothesis that protein rather than RNA is the primer for minus-strand DNA synthesis (15). In support of this model, it has been demonstrated that the terminal protein of duck hepatitis B virus is encoded by the viral polymerase gene, which also encodes the virus-associated reverse transcriptase (1, 23).

Molecular analysis of the major replication intermediates of WHV and other hepadnaviruses revealed the exact positions of the 5' and 3' ends of minus-strand DNA (11, 14, 18, 25). The 5' end of minus-strand DNA was mapped to a short direct repeat, DR1, on the WHV genome. DR1 is located within the terminal redundant region of pregenomic RNA (R region). Therefore, the answer to the question of where minus-strand DNA synthesis actually initiates on pregenomic RNA remained ambiguous. Initiation of DNA synthesis at DR1 close to the 5' end of pregenomic RNA would subsequently require a template switch of a nascent minus-strand DNA to the 3' end of the RNA template, similar to the transfer of nascent minus-strand (strong stop) DNA that occurs during replication of the retroviral genome (24).

To investigate the molecular mechanisms that orchestrate the events leading to the initiation of reverse transcription in

hepadnaviruses, we attempted to identify the signals on WHV pregenomic RNA required for this process. We performed experiments employing a recently developed cytomegalovirus-WHV expression vector for the production of infectious WHV in tissue culture cells (19) and identified the initiation site for minus-strand DNA synthesis on pregenomic RNA. In addition, we have localized the nucleotide sequences on the RNA template that provide the signals for reverse transcription of minus-strand DNA covalently linked to protein.

MATERIALS AND METHODS

Cell culture. Maintenance of HepG2 cells, procedures for DNA transfections, and purification of virion DNA from the culture medium of transfected cells were performed as described previously (19).

Molecular clones. The construction of pCMW82 has been described previously (19). Briefly, WHV clone pWHV81-2 was fused with pBC12/CMV/IL2 (4) to direct the synthesis of WHV pregenomic RNA from the immediate-early cytomegalovirus promoter on pBC12/CMV/IL2. The mutants p82T5 and p82T3 were obtained by oligomer-directed site-specific mutagenesis as described by Taylor et al. (22), using a kit from Amersham Corp. The mutant genotypes were confirmed by nucleotide sequence analysis. The four constructs, 82NS1 through 82NS4, and the five constructs, 82PC1 through 82PC5, were obtained by linearizing pCMW82 with the restriction endonucleases *NsiI* (position 3300) and *SlyI* (position 72), respectively, followed by incubation with the exonuclease *Bal31* (Bethesda Research Laboratories, Inc.) for different times. The ends of the linear DNA were joined with T4 ligase under conditions favoring the formation of circles. The sizes of the deletions were determined by nucleotide sequence analysis. Plasmid pPC2A was obtained by removal of the *NheI*-to-*BstEII* fragment from p82PC2 and plasmid p82d1C was obtained by removal of the *NsiI*-to-*BstEII* fragment from pCMW82. For the construction of p82C, DNA oligomer PCA was inserted into the *NsiI* site of p82d1C, and for the construction of pPC2C1, the *HindIII*-

* Corresponding author.

FIG. 1. Localization of the initiation site for minus-strand DNA synthesis. (A) Two possible models for minus-strand DNA synthesis, depending on the initiation site of DNA synthesis at the 5' (model 1) or 3' (model 2) end of pregenomic RNA. The positions of the direct repeats DR1 and DR2 on the pregenomic RNA are indicated. Hatched boxes mark the DR1 copy which serves as the template for the synthesis of the 5' end of minus-strand DNA. The letters R and r indicate the terminal redundancies of pregenomic RNA and minus-strand DNA, respectively. The solid oblong symbols depict the protein covalently linked to the 5' end of minus-strand DNA. (B) Nucleotide sequences corresponding to the DR1 and r segments on pregenomic RNA and minus-strand DNA, respectively. The position of the 5' end of minus-strand DNA is indicated at position 14 as determined in reference 18. The solid oblong symbol depicts the terminal protein at the 5' end of minus-strand DNA. The position of the dA residue inserted into pCMW82 for the construction of p82T5 and p82T3, respectively, is indicated by the letter A. (C) Primer extension analysis was performed as described in Materials and Methods by using DNA oligomer CS3 as a primer. Lanes 5 to 7 show the DNA products of the primer extension reactions on DNA isolated from virions in the supernatant of HepG2 cells transfected with pCMW82 (lane 5), p82T5 (lane 6), and p82T3 (lane 7). A nucleotide sequence ladder derived from M13 clone Wsh (19), primed with DNA oligomer DR17P, provided size markers (lanes 1 to 4). nt, Nucleotides.

to-*Bst*EII fragment in p82PC2 was replaced with the DNA oligomer PCA.

DNA oligomers. The 42-base-pair long DNA oligomer PCA was prepared by the incubation of equal amounts of DNA oligonucleotides *pca* (5'-AGCTTGGTTACCTATCTTTT-TCACCTGTGCAAGCTT3') and *acp* (5'-GTTACCAAGCTTGCACAGGTGAAAAAGATAGGTAAC-3') in TNE (10 mM Tris [pH 8], 100 mM NaCl, 1 mM EDTA) buffer. The underlined nucleotide sequence corresponds to positions 4 to 21 on the WHV genome. DNA oligonucleotides for primer extension and nucleotide sequencing reactions correspond to positions 3229 to 3248 (CS3), 3125 to 3144 (CS4), and 26 to 3316 (DR17P) on the WHV genome. Primer extension reaction experiments on purified virion DNA were performed as described previously (19).

Nomenclature. Numbering of the WHV type 2 genome was arranged to begin with the first ATG of the precore gene (position 1931 in reference 7).

RESULTS

The 5' end of minus-strand DNA maps to the DR1 repeat close to the 3' end of pregenomic RNA. Determination of the 5' and 3' ends of minus-strand DNA was previously accomplished by primer extension analysis of deproteinized DNA purified from virions of WHV-infected woodchucks (18). The 5' end of minus-strand DNA was mapped to the fourth nucleotide within DR1 (Fig. 1B). Since the 3' end of minus-strand DNA coincides with the 5' end of pregenomic RNA, the complete minus-strand DNA has a short terminally redundant portion, termed *r* (Fig. 1A and B). Since DR1 is contained within the terminally redundant portion of pregenomic RNA, the initiation site for minus-strand DNA synthesis remained ambiguous. The most straightforward expectation is that initiation of reverse transcription occurs at the DR1 close to the 3' end of pregenomic RNA because this would allow for continuous minus-strand DNA synthesis

from the 3' end toward the 5' end of pregenomic RNA (Fig. 1A, model 2). It is also possible, however, that initiation of reverse transcription occurs at the DR1 near the 5' end of pregenomic RNA, requiring a template switch of the 10- to 11-nucleotide-long, nascent minus-strand DNA to the corresponding sequences at the 3' end of pregenomic RNA (Fig. 1A, model 1).

To identify the initiation site for minus-strand DNA synthesis on pregenomic RNA, a deoxyribosyladenine (dA) residue was inserted independently into pCMW82 in the DNA segment corresponding to the 5' or to the 3' copy of the r region (Fig. 1B). pCMW82 contains a cDNA equivalent of WHV pregenomic RNA fused to the cytomegalovirus immediate-early promoter-enhancer region, as described previously (19). The resulting mutants, p82T5 and p82T3, were assayed for virus production after transfection into HepG2 cells. DNA was isolated from virus particles, concentrated from the supernatant of transfected cells, and subjected to primer extension analysis for the determination of the 5' ends of minus-strand DNA. The major DNA products from the extension reaction experiments with virion DNA synthesized with pCMW82 (Fig. 1C, lane 5) and p82T5 (lane 6) were the same size, whereas the DNA product obtained from the reaction with virion DNA from p82T3 was one nucleotide longer compared with the other two products (lane 7). Thus, addition of a dA residue into the 5' copy of the r region (p82T5) did not affect the length of the 5' end of minus-strand DNA (lane 6), whereas addition of the same residue into the 3' copy of r (p82T3) led to the synthesis of minus-strand DNA with an additional nucleotide at its 5' end. Therefore, we concluded that the 5' end of minus-strand DNA maps to the copy of DR1 close to the 3' end of pregenomic RNA and that minus-strand DNA synthesis progresses continuously from this site toward the 5' end of the RNA template (Fig. 1A, model 2). In the following discussion, we will refer to the DR1 at the 3' end of pregenomic RNA as DR1*.

A short sequence motif flanking DR1* is required for initiation of minus-strand DNA synthesis. To identify the nucleotide sequences on pregenomic RNA that provide the signals for the initiation of minus-strand DNA synthesis, we performed a deletion analysis of the region flanking DR1*. Constructs with deletions in the 5' direction of DR1* were made with the exonuclease *Bal31* after linearizing pCMW82 at the *NsiI* site at position 3300 on the WHV genome. Four deletion mutants, spanning nucleotides 3273 to 30, were assayed for the ability to direct the synthesis of minus-strand DNA with correctly initiating 5' ends. The positions of the deleted segments in p82NS1 through p82NS4 are illustrated in Fig. 2A, and the DNA products of the primer extension reactions obtained with DNA purified from virions in the culture medium of transfected HepG2 cells are shown in Fig. 2B. DNA products were obtained with virion DNA expressed from pCMW82 (lane 5), p82NS1 (lane 6), and p82NS2 (lane 7), whereas no detectable reaction products were obtained with virion DNA from p82NS3 (lane 8) and p82NS4 (data not shown). The determination of the lengths of the extended products obtained from reactions with virion DNA from p82NS1 and p82NS2 revealed an exact correlation between the length of the extension product and the size of the deletion in the respective plasmid. This result indicates that the 5' end of minus-strand DNA expressed with p82NS1 and p82NS2 mapped to the same position as the 5' end synthesized with wild-type pCMW82.

These results show that two derivatives of pCMW82 with deletions spanning positions 3273 to 9 remain competent for

minus-strand DNA synthesis, whereas two constructs with deletions extending beyond position 14 have lost the ability to synthesize detectable minus-strand DNA.

Constructs with deletions in the 3' direction of DR1* were created by linearizing pCMW82 at position 72 with the endonuclease *StyI* and by incubation with the exonuclease *Bal31*. Five mutants, designated p82PC1 through p82PC5, spanning positions 18 to 165 were obtained (Fig. 2A) and analyzed for the ability to direct the synthesis of minus-strand DNA as described above. As shown in Fig. 2C, transfection of HepG2 cells with four of the five constructs led to the synthesis of identical DNA products as those obtained with pCMW82, indicating that these plasmids were competent for the expression of correctly initiated minus-strand DNA. The signal of the extension product obtained from reactions with virion DNA from p82PC5 (lane 9) was reduced approximately three- to fivefold compared with the signal obtained from pCMW82 (lane 4). The signal of the DNA product from reactions with virion DNA from p82PC2 was less than 1% of the signal obtained from the other reactions.

The deletion in p82PC2, spanning 132 nucleotides from positions 34 to 165, removes the polyadenylation site on the WHV genome at position 106 and leads to the formation of a pregenomic RNA species with an additional 1,064 nucleotides at its 3' end. This transcript terminates at a second polyadenylation site present on p82PC2 (Fig. 2A). We speculated that the additional 1,064 nucleotides at the 3' end of pregenomic RNA interfered with DNA replication or, alternatively, that the deletion of WHV sequences between positions 34 and 106 removed nucleotide sequences that provided signals necessary for minus-strand DNA synthesis. To distinguish between these possibilities, we reduced the distance between DR1* and the polyadenylation signal by removing the *NheI*-to-*BstEII* fragment of p82PC2. This derivative of p82PC2, pPC2A leads to the transcription of a shorter pregenomic RNA species with only 304 additional nucleotides at its 3' end compared with the natural pregenomic RNA produced by pCMW82. This construct, together with pCMW82 and p82PC2, was transfected into HepG2 cells, and primer extension analysis was performed on virion DNA in the culture medium of the transfected cells as described above. As shown in Fig. 2D, the signal of the DNA products obtained from reactions with virion DNA from pPC2A (lane 9) was comparable to the signal from pCMW82 (lane 7). Hence, removal of the *NheI*-to-*BstEII* fragment from p82PC2 was sufficient to rescue the ability of this construct to express minus-strand DNA. We conclude that the additional 760 nucleotides present on pregenomic RNA transcribed from p82PC2 either prevented RNA packaging or interfered with the synthesis of minus-strand DNA.

The results obtained from the deletion analysis of the 3' terminal region of pregenomic RNA demonstrate that WHV sequences from position 18 to the polyadenylation signal on pregenomic RNA at position 106 are not required for reverse transcription of minus-strand DNA. In addition, our data show that a limited addition of nucleotide sequences to the 3' end of pregenomic RNA does not interfere with the initiation of minus-strand DNA synthesis. In summary, the results from both deletion analyses suggest that a short sequence motif approximately eight nucleotides long, spanning positions 10 to 17, harbors the signals required for the initiation of reverse transcription of minus-strand DNA.

A synthetic DNA oligomer functions as an independent origin of DNA replication. To test whether the nucleotide sequences flanking DR1* provide all the signals necessary

FIG. 2. Identification of the nucleotide sequences required for minus-strand DNA synthesis. (A) The DNA segment on pCMW82 corresponding to the 3' portion of pregenomic RNA with the initiation site for minus-strand DNA synthesis at DR1*. The hatched and solid circles represent the polyadenylation signals on pCMW82 derived from the WHV genome and from pBC12/CMV/IL2, respectively (see text for details). The extent of the deletions in derivatives of pCMW82 is indicated with horizontal arrows. The positions of the restriction sites used for generation of the *Bal*31 deletions and the initiation site for minus-strand (-) DNA are indicated. N, *Nsi*; S, *Syl*. (B) Primer extension analysis was performed as described in Materials and Methods by using DNA oligomer CS4 as a primer. The lengths of the extended DNA products are indicated. A nucleotide sequence ladder derived from the pCMW82 clone primed with DNA oligomer CS3 provided size markers (lanes 1 to 4). Lanes 5 to 8 show the products of the primer extension reactions on virion DNA isolated from the supernatant of HepG2 cells transfected with pCMW82 (lane 5), p82NS1 (lane 6), p82NS2 (lane 7), and p82NS3 (lane 8). The predicted length of the primer-extended DNA product from the reaction with virion DNA from p82NS3 is indicated in parentheses. nt, Nucleotides. (C) Primer extension reaction was performed on virion DNA isolated from the serum of woodchuck W1451 (lane 3) and from the supernatant of HepG2 cells transfected with pCMW82 (lane 4), p82PC1 (lane 5), p82PC2 (lane 6), p82PC3 (lane 7), p82PC4 (lane 8), and p82PC5 (lane 9), employing DNA oligomer CS3 as primer. The nucleotide sequencing ladder shown in lanes 1 and 2 was derived from pCMW82 primed with the same DNA oligomer as used for the primer extension reaction. nt, Nucleotides. (D) Primer extension was performed on virion DNA isolated from the serum of woodchuck W1451 (lane 6) and from the supernatant of HepG2 cells transfected with pCMW82 (lane 7), p82PC2 (lane 8), and pPC2A (lane 9). Lanes 1 to 4 show size markers as indicated in the legend to Fig. 2C. DNA oligomer CS3 served as primer for both the primer extension and the sequencing reactions. nt, Nucleotides.

for the initiation of minus-strand DNA synthesis, we designed a 42-base-pair-long DNA oligomer that contains WHV sequences corresponding to positions 4 to 21 on the WHV genome (Materials and Methods). We created two derivatives of pCMW82 with insertions of DNA oligomer

PCA near the 3' end of pregenomic RNA. Plasmid p82C was generated by replacing the *Nsi*I-to-*Bst*EII fragment of pCMW82 containing the entire R region of pregenomic RNA (including DR1*) with DNA oligomer PCA (Fig. 3A). In plasmid pPC2C1, the *Hind*III-to-*Bst*EII fragment was re-

placed with PCA, leaving DR1* intact and thus rendering a construct with two potential initiation sites for minus-strand DNA synthesis separated by 136 nucleotides. p82d1C is identical to p82C except for the lack of the PCA fragment (Fig. 3A).

These constructs, along with pCMW82 and pPC2A, were transfected into HepG2 cells, and primer extension reactions were performed on DNA purified from virions from the culture medium of the transfected cells. The DNA products of the extension reactions are shown in Fig. 3B. As expected, the extension reaction with DNA expressed from p82d1C did not yield any detectable DNA products (lane 4), since it did not contain an initiation site for minus-strand DNA synthesis. In contrast, DNA products were obtained from reactions with virion DNA expressed from p82C (lane 5). The length of the major reaction product was 11 nucleotides shorter compared with that of the product obtained with virion DNA expressed from pCMW82 (lane 3). This finding implies that the 5' end of minus-strand DNA maps to the position on the PCA sequence that corresponds to the 5' end minus-strand DNA initiating at DR1*.

Primer extension reactions with DNA purified from virions in the supernatant of HepG2 cells transfected with pPC2C1 revealed two DNA products (lane 7). The faster-migrating fragment, 105 nucleotides long, corresponds to the 5' end of minus-strand DNA initiating at DR1*; the slower-migrating fragment, 241 nucleotides long, corresponds to a minus-strand DNA species initiating within the PCA sequence on pPC2C1. Direct comparison of the lengths of the two major extension products with a nucleotide sequencing ladder derived from pPC2C1 demonstrates that both DNA products migrated to the same nucleotide within DR1* and PCA, respectively (lanes 7 to 9). The intensity of the signal of the 241-nucleotide-long DNA product is reduced approximately 10-fold compared with the signal of the faster-migrating species, 105 nucleotides in length. As expected, the 241-nucleotide-long DNA product was not produced in reactions on virion DNA from transfections with pPC2A which lacks the PCA sequence motif (lane 6).

Since minus-strand DNA initiating at DR1* is covalently attached to protein, we asked whether minus-strand DNA originating at the PCA motif was also linked to protein. For this purpose, HepG2 cells were transfected with pPC2C1 and virions were concentrated from the supernatant of the transfected cells. DNA was extracted from virions with and without incubation with proteinase K before phenol extraction. It is known that omission of proteinase K treatment before phenol extraction results in the loss of minus-strand DNA linked to protein from the aqueous phase (6). Primer extension reactions were performed on virion DNA purified by these two procedures. As expected from the previous

results, two major DNA products were obtained from reaction with virion DNA extracted after treatment with proteinase K (Fig. 3C, lane 5). Omission of proteinase K before phenol extraction equally reduced the intensity of the signals from the two reaction products. This result indicates that the 5' end of minus-strand DNA originating at the PCA sequence is covalently linked to protein.

In summary, these results show that the DNA oligomer PCA contains the signals required for the initiation of reverse transcription of minus-strand DNA. The 5' ends of minus-strand DNA initiating at PCA and DR1* map to the same relative positions on PCA and DR1* and are both linked to protein at their 5' ends. Therefore, we conclude that the oligomer PCA contains an origin for minus-strand DNA synthesis of WHV.

DISCUSSION

Replication of the hepadnavirus genome results in the formation of a covalent link between protein and the 5' end of minus-strand DNA. This feature led to the hypothesis that hepadnaviruses, unlike all other known retroviral elements, use protein as a primer for minus-strand DNA synthesis. Precedents for protein-primed DNA synthesis are found in prokaryotes and eukaryotes, as exemplified by bacteriophage ϕ 29 and adenovirus systems (26). The mechanisms require a specific interaction between the primase and nucleotide sequences at the termini of the respective DNA genomes (16, 20). In hepadnaviruses, the situation is different. First, the putative primase must interact with RNA rather than DNA, and second, initiation of DNA synthesis occurs from an internal site on the RNA template and not from its terminus.

To investigate the mechanism underlying the initiation of minus-strand DNA synthesis of hepadnaviruses, we attempted to identify the nucleotide sequences on pregenomic RNA that are required for this process. The experiments described in this report revealed a short sequence motif located at the 3' end of pregenomic RNA that is essential for the initiation of DNA synthesis. We demonstrated that this signal sequence can exert its function from different positions on the 3' end of pregenomic RNA and thus represents the origin of WHV DNA replication.

Minus-strand DNA synthesis is continuous. The results reported in this communication show that the 5' end of minus-strand DNA maps to the DR1 sequence (DR1*) at the 3' end of pregenomic RNA (Fig. 1B). Hence, the synthesis of minus-strand DNA from the pregenomic RNA template can proceed continuously from DR1* to the 5' end of the pregenome. Thus, unlike retroviruses (21, 24) and caulimoviruses (3, 13), the strategy employed by hepadnaviruses for

FIG. 3. Initiation of reverse transcription from a synthetic origin of DNA replication, the DNA oligomer PCA. (A) Maps of the DNA segments corresponding to the 3' ends of pregenomic RNA of derivatives of pCMW82. DR1* and PCA are depicted by hatched and open rectangles, respectively. Hatched and solid circles mark the polyadenylation signals as described in the legend to Fig. 2A. The restriction sites used for the construction of the plasmids shown in lines 2 to 5 are indicated (see text for details). N, *Nsi*I; H, *Hind*III; Nh, *Nhe*I; B, *Bst*EII. WHV sequences are shown as thin lines, and vector sequences derived from pBC12/CMV/IL2 are shown as heavy lines. R, Terminal redundancy on pregenomic RNA. Construction of the derivatives of pCMW82 shown in lines 2 to 7 is described in Materials and Methods and in the Results. (B) Primer extension analysis was performed as described in the legend to Fig. 2C. Nucleotide sequence ladders derived from pCMW82 (lanes 1 and 2) and from pPC2C1 (lanes 8 and 9), primed with oligomer CS3, provided size markers. Lanes 3 to 7 show the extended primer from reactions with virion DNA isolated from the supernatant of HepG2 cells transfected with pCMW82 (lane 3), p82d1C (lane 4), p82C (lane 5), pPC2A (lane 6), and pPC2C1 (lane 7). The lengths of the primer-extended products are indicated. nt, Nucleotides. (C) Primer extension analysis was performed on DNA purified from virions in the supernatant of HepG2 cells transfected with pPC2C1 as described in the legend to Fig. 2C. DNA was extracted with (lane 5) and without (lane 6) treatment of concentrated virus with proteinase K before phenol extraction. The lengths of the primer extended products are indicated. The sequencing ladder shown in lanes 1 to 4 served as size markers. nt, Nucleotides.

minus-strand DNA synthesis does not require the transfer of nascent minus-strand DNA from the 5' end of the RNA template to its 3' end. This conclusion is in agreement with the proposed model for the priming of plus-strand DNA synthesis in which the 5' terminal portion of pregenomic RNA, including DR1, is translocated to DR2 on minus-strand DNA (10, 18). This model implies that, prior to its transfer to DR2, the 5' terminal portion of pregenomic RNA must have already been copied into minus-strand DNA. Initiation of reverse transcription at the 3' end of pregenomic RNA allows for the synthesis of minus- and plus-strand DNA from a single RNA template.

Origin of WHV DNA replication. We have shown that a short sequence motif, less than 8 nucleotides in length and spanning positions 10 to 17, is required for initiation of DNA synthesis. We refer to this sequence as the signal for reverse transcription (SRT). Deletions extending into the SRT region obliterated the production of minus-strand DNA (e.g., p82NS3, Fig. 2A), indicating that initiation of reverse transcription is strictly dependent on the SRT sequence motif on pregenomic RNA.

Since we could not conclude from these experiments that the SRT region functions as an independent origin of DNA replication, we asked whether a synthetic nucleotide sequence corresponding to the SRT sequence motif could direct the synthesis of minus-strand DNA from pregenomic RNA. This question was answered in the affirmative. Insertion of DNA oligomer PCA, comprising the SRT sequence motif, into the 3' portion of pregenomic RNA at two different positions led to the synthesis of minus-strand DNA. This result demonstrates that the PCA sequence can function as an independent origin of DNA replication. However, insertion of the synthetic SRT site in the 3' direction of the natural priming site at DR1* (pPC2C1, Fig. 3B), led to severalfold reduced levels of initiation of minus-strand DNA synthesis at PCA compared with those of initiation at DR1*. This result suggests that additional nucleotide sequences exist on pregenomic RNA that affect the rate of minus-strand DNA synthesis. It is also conceivable that the structural organization of the core particle places the DR1* region on pregenomic RNA in an optimal position for an interaction with the primase-polymerase complex.

An important finding from these experiments is that the nucleotide sequences between DR1* and the poly(A) sequence on pregenomic RNA do not code for functions required for packaging of pregenomic RNA or for the synthesis of minus- and plus-strand DNAs (Fig. 2A). Results obtained from DNA transfection experiments with construct p82PC2 and its derivative, pPC2A, suggest that the length of the 3' portion of pregenomic RNA is critical for DNA synthesis (Fig. 2A and C). The lack of DNA synthesis observed with p82PC2 could be due to the inability of pregenomic RNA exceeding a certain size to be packaged into core particles. Alternatively, it could be due to direct interference of these additional nucleotide sequences with initiation of minus-strand DNA synthesis.

Is protein the primer for reverse transcription? The presence of protein at the 5' end of minus-strand DNA suggests a mechanism in which protein is the primer for DNA synthesis. Although protein priming is an attractive possibility with precedents, the possibility cannot be excluded that priming actually occurs further toward the 3' end of pregenomic RNA, e.g., with the help of an RNA primer hydrogen bonded to pregenomic RNA or alternatively with the help of a 2' hydroxyl group of a nucleotidyl residue of the RNA template, as described for the priming of msDNA synthesis

in bacteria (8, 12). In this model, after the initiation of DNA synthesis protein is added to nascent minus-strand DNA by breaking a DNA internucleotide bond to form a covalent link between protein and minus-strand DNA. Such a mechanism has been proposed for the link between the A protein of bacteriophage ϕ X174 and phage DNA (9) and for the link between the VPg protein and RNA of poliovirus (27).

In contrast to the model in which protein is the primer for DNA synthesis, this alternative model could suggest that the protein at the 5' end of minus-strand DNA has other biochemical functions, for instance those of a topoisomerase or a ligase. Furthermore, according to this model, priming of DNA synthesis could be uncoupled from the reaction that joins protein with DNA. However, the results presented in this report do not indicate that in hepadnaviruses these two events can be separated from each other. For example, removal of DR1* from pregenomic RNA did not lead to the synthesis of any detectable minus-strand DNA species initiating between the positions corresponding to DR1* and the poly(A) sequence at the 3' terminus of pregenomic RNA (p82d1C, Fig. 3B). Furthermore, substitution of WHV sequences located in the 3' direction of the initiation site for minus-strand DNA synthesis with heterologous sequences did not interfere with reverse transcription, suggesting that the 3' terminal portion of pregenomic RNA (R, Fig. 1A) does not contain additional signals important for DNA replication. Finally, the observation that the DNA oligomer PCA contains the signals necessary for both the initiation of minus-strand DNA and the attachment of protein to the 5' end of minus-strand DNA supports the notion that these two events occur concomitantly.

The identification of the WHV origin of DNA replication provides an important basis for a detailed genetic analysis of this region, which may yield important information about the biochemical events that lead to the formation of the covalent link between protein and DNA during replication of the hepadnavirus genome. The results reported in this communication, together with recent reports stating that the reverse transcriptase of hepadnaviruses is synthesized independently from the core gene product (2, 17) and that the terminal protein is encoded by the viral polymerase gene (1), might provide a basis for the development of an *in vitro* system useful for an examination of virus replication under chemically defined conditions.

ACKNOWLEDGMENTS

We thank Bud Tennant for woodchuck serum samples and Volker Vogt and Jim Casey for valuable comments to the manuscript.

This work was supported by Public Health Service grant AI-24972 from the National Institutes of Health.

LITERATURE CITED

1. Bartenschlager, R., and H. Schaller. 1988. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. *EMBO J.* 7:4185-4192.
2. Chang, L.-J., P. Pryciak, D. Ganem, and H. E. Varmus. 1989. Biosynthesis of the reverse transcriptase of hepatitis B viruses involves *de novo* translational initiation not ribosomal frameshifting. *Nature (London)* 337:364-368.
3. Covey, S. N., D. Turner, and G. Mulder. 1983. A small DNA molecule containing covalently-linked ribonucleotides originates from the large intergenic region of the cauliflower mosaic virus genome. *Nucleic Acids Res.* 11:251-264.
4. Cullen, B. R. 1987. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. *Cell* 46:972-982.

5. Ganem, D., and H. E. Varmus. 1987. The molecular biology of the hepatitis B viruses. *Annu. Rev. Biochem.* **56**:651-693.
6. Gerlich, W. H., and W. S. Robinson. 1980. Hepatitis B virus contains protein attached to the 5' terminus of its complete DNA strand. *Cell* **21**:801-809.
7. Kodama, K., N. Ogasawara, H. Yoshikawa, and S. Murakami. 1985. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: evolutionary relationship between hepadnaviruses. *J. Virol.* **56**:978-986.
8. Lampson, B. C., M. Inouye, and S. Inouye. 1989. Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of *Myxococcus xanthus*. *Cell* **56**:701-707.
9. Langeveld, S. A., A. D. M. van Mansfeld, J. M. de Winter, and P. J. Weisbeek. 1979. Cleavage of single-stranded DNA by the A and A* proteins of bacteriophage OX174. *Nucleic Acids Res.* **7**:2177-2188.
10. Lien, J., C. E. Aldrich, and W. S. Mason. 1986. Evidence that a capped oligoribonucleotide is the primer for duck hepatitis B virus plus strand DNA synthesis. *J. Virol.* **57**:229-236.
11. Lien, J., D. J. Petcu, C. E. Aldrich, and W. S. Mason. 1987. Initiation and termination of duck hepatitis B virus DNA synthesis during virus maturation. *J. Virol.* **61**:3832-3840.
12. Lim, D., and W. K. Maas. 1989. Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B. *Cell* **56**:891-904.
13. Mason, W. S., J. M. Taylor, and R. Hull. 1987. Retrovirus genome replication. *Adv. Virus Res.* **32**:35-96.
14. Molnar-Kimber, K. L., J. W. Summers, and W. S. Mason. 1984. Mapping of the cohesive overlap of duck hepatitis B virus DNA and of the site of initiation of reverse transcription. *J. Virol.* **51**:181-191.
15. Molnar-Kimber, K. L., J. W. Summers, J. M. Taylor, and W. S. Mason. 1983. Protein covalently bound to minus strand DNA intermediates of duck hepatitis B virus. *J. Virol.* **45**:165-172.
16. Prieto, I., M. Serrano, J. M. Lazaro, M. Salas, and J. M. Hermose. 1988. Interaction of the bacteriophage 029 protein p6 with double-stranded DNA. *Proc. Natl. Acad. Sci. USA* **85**: 314-318.
17. Schlicht, H. J., G. Radziwill, and H. Schaller. 1989. Synthesis and encapsidation of duck hepatitis B virus reverse transcriptase do not require formation of core-polymerase fusion proteins. *Cell* **56**:85-92.
18. Seeger, C., D. Ganem, and H. E. Varmus. 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. *Science* **232**:477-484.
19. Seeger, C., and J. Maragos. 1989. Molecular analysis of the function of direct repeats and a polypurine tract for plus-strand DNA priming in woodchuck hepatitis virus. *J. Virol.* **63**:1907-1915.
20. Tamanoi, F., and B. W. Stillman. 1982. Function of adenovirus terminal protein in the initiation of DNA replication. *Proc. Natl. Acad. Sci. USA* **79**:2221-2225.
21. Taylor, J. M., and R. Illmensee. 1975. Site on the RNA of an avian sarcoma virus at which primer is bound. *J. Virol.* **16**: 553-558.
22. Taylor, J. W., J. Ott, and F. Eckstein. 1985. The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. *Nucleic Acids Res.* **13**:8765-8784.
23. Toh, H., H. Hayashida, and T. Miyata. 1983. Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. *Nature (London)* **305**:827-829.
24. Varmus, H. E., and R. Swanstrom. 1985. Replication of retroviruses, p. 75-134. *In* R. Weiss, N. Teich, H. E. Varmus, and J. Coffin (ed.), *RNA tumor viruses* 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. Will, H., W. Reiser, T. Weimer, E. Pfaff, M. Buscher, R. Sprengel, R. Cattaneo, and H. Schaller. 1987. Replication strategy of human hepatitis B virus. *J. Virol.* **61**:904-911.
26. Wimmer, E. 1982. Genome-linked proteins of viruses. *Cell* **28**:199-201.
27. Young, D. C., D. M. Tuschall, and J. B. Flanagan. 1985. Poliovirus RNA-dependent RNA polymerase and host cell protein synthesize product RNA twice the size of poliovirus RNA in vitro. *J. Virol.* **54**:256-264.