JOURNAL OF VIROLOGY, Oct. 1996, p. 7085–7091 Vol. 70, No. 10
0022-538X/96/$04.00+0

A Hydrophobic Heptad Repeat of the Core Protein of Woodchuck Hepatitis Virus Is Required for Capsid Assembly

MINSHU YU,* ROGER H. MILLER,† SUZANNE EMERSON, AND ROBERT H. PURCELL

Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892-0470

Received 14 February 1996/Accepted 18 July 1996

The capsid particle of hepadnaviruses is assembled from its dimer precursors. However, the mechanism of the protein-protein interaction is still poorly understood. A small region in the capsid protein of woodchuck hepatitis virus (WHV) contains four hydrophobic residues, including leucine 101, leucine 108, valine 115, and phenylalanine 122, that are conserved and spaced every seventh residue in the primary sequence to form a hydrophobic heptad repeat (*hhr*). A hydrophobic force often plays an important role in the interaction of proteins. Therefore, to investigate the role of this region in capsid assembly, we individually changed the codons specifying these four hydrophobic amino acids to codons specifying alanine or proline. In addition, we examined the in vivo infectivity of a WHV genome bearing a naturally occurring single amino acid change (histidine 104→proline) in the *hhr* region. The phenotype of each altered genome was determined in both eukaryotic and prokaryotic systems by a capsid protein assay and electron microscopic examination. We show that replacement of any one of the four hydrophobic residues with alanine did not prevent capsid assembly. However, assembled capsid particles were not detected if combinations of any two of the four residues were substituted with alanines or if the spacing of these four hydrophobic residues was changed. An individual introduction of a proline (which dramatically changes the secondary structure of proteins) into different positions of this small region also abolished capsid assembly in vitro or viral replication in vivo. These results suggested that the *hhr* region of the core protein of WHV was critical for capsid assembly.

Woodchuck hepatitis virus (WHV) is a small DNA virus that belongs to the hepadnavirus family (34, 40). The WHV nucleocapsid consists of a partially duplex, open circular DNA genome and a virus-encoded polymerase enclosed within a protein shell of core protein. The nucleocapsid is surrounded by a lipid bilayer containing envelope surface proteins. The core protein performs a central function in viral replication. During initiation of infection, once the virus penetrates permissive cells and the envelope is removed, the viral genome is immediately delivered by the nucleocapsid to the nucleus. Here the genome is released from the capsid shell to begin replication. In later steps of infection, the core protein is assembled into a capsid shell in which a viral RNA pregenome transcribed from the viral DNA is encapsidated. Within this nucleocapsid, the core protein regulates viral DNA synthesis, including initiation of synthesis of minus-strand DNA and elongation of plus-strand DNA (43, 45). The nucleocapsid further participates in amplification of covalently closed circular DNA in the nucleus (33, 37) and interacts with envelope proteins to form enveloped virus particles.

The capsid particles of hepadnavirus are assembled from dimers of the core protein into a highly ordered icosahedral conformation (7, 46). The primary sequence of core protein contains two major functional regions. The carboxyl-terminal region of the core protein is required for general nucleic acid binding (18, 27), phosphorylation (25, 44, 45), and nuclear localization (42). The amino-terminal region from amino acid positions 10 to 140 provides the lateral contact that is required for self-assembly of 180 or 240 subunits into an icosahedral protein shell in the cytoplasm of infected cells (4, 14, 31). The N terminus of the core protein is thought not to be exposed on the surface of the capsid particle and probably is buried within the structure to form a bond that promotes the protein-protein interaction (4, 32). However, the mechanism of the core-core subunit interaction is not well understood at the level of either dimer formation or assembly. Several smaller regions and specific amino acid residues in this N-terminal region were found to be involved in capsid assembly (4, 7, 28, 41). Insertions of two or four amino acids in different regions of the N terminus of the core protein of hepatitis B virus (HBV) showed that some mutations resulted in defective capsid assembly but some did not (2). These results suggested that the N-terminal region may contain several domains that are required for capsid assembly at different levels.

The core protein of both mammalian and avian hepadnaviruses possesses an N-terminal region containing four conserved hydrophobic amino acid residues (i.e., leucine 101, leucine 108, valine 115, and phenylalanine 122 in WHV) that are spaced every seventh residue in the primary sequence to form a leucine zipper-like motif (22, 29) that we have termed the hydrophobic heptad repeat (*hhr*). Since a hydrophobic force often plays an important role in protein-protein interactions, it was of interest to determine whether the *hhr* of the WHV core protein was involved in capsid assembly. Thus, we changed the codons for these four hydrophobic amino acid residues to codons specifying alanine or proline and determined the phenotype of each mutant in both eukaryotic and prokaryotic systems. In this paper, we report that the *hhr* domain of the core protein of WHV is critical for capsid assembly.

* Corresponding author. Mailing address: Hepatitis Viruses Section, Laboratory of Infectious Diseases, 7 Center Dr. 0470, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892-0470. Phone: (301) 496-6227. Fax: (301) 402-0524. Electronic mail address: myu@helix.nih.gov.

† Present address: Targeted Interventions Branch, Basic Sciences Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892.

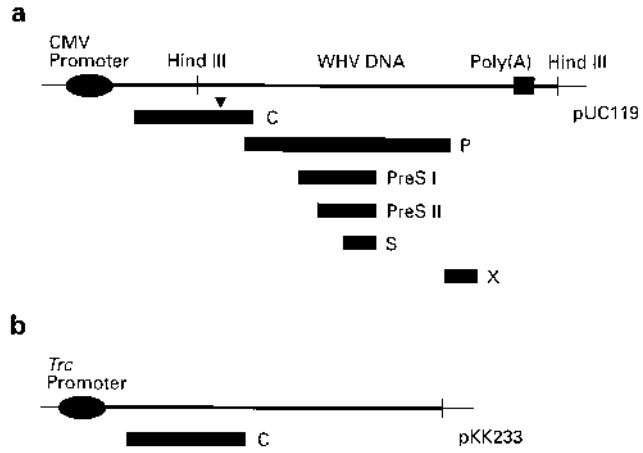


FIG. 1. Diagrams of the pUC119/CMV/WHV eukaryotic expression plasmid and pKK233/WHV/CORE prokaryotic expression plasmid. (a) The pregenome-encoding fragment of WHV8 was ligated downstream of a cytomegalovirus promoter region cloned in the vector pUC119. The construct pUC119/CMV/WHV is shown relative to open reading frames for the core (C), polymerase (P), X, and envelope (pre-S I, pre-S II, and S) genes of WHV. The location of the *hhr* domain is symbolized by an inverted triangle. (b) The entire WHV8 core gene was ligated to a *trc* promoter cloned in the vector pKK233-2. The ATG of the core gene is located directly downstream of the *trc* promoter.

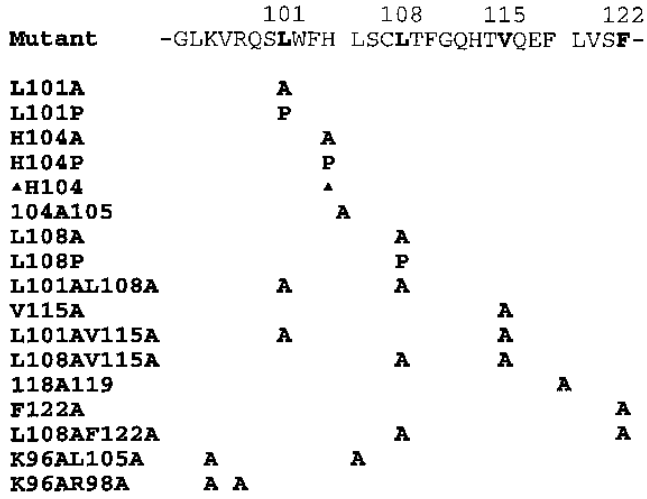


FIG. 2. Map of WHV core mutants used in the analysis of capsid protein function. The sequence from amino acids 94 to 122 of the wild-type core protein is displayed. Amino acids are numbered from the AUG of the core protein at nucleotide 2021 according to the work of Girones et al. (16). For each of the mutant core proteins depicted, the new amino acids are indicated at the relative positions of the wild-type sequence.

MATERIALS AND METHODS

Plasmids. An eukaryotic expression plasmid, pUC119 CMV/WHV, was constructed, a plasmid in which 4.7 kb of WHV strain 8 (WHV8) (16) pregenome-encoding sequence, beginning with nucleotide 1916 (according to the numbering scheme of Girones et al. [16]) and ending with the second *Hin*dIII site at nucleotide 2190, was cloned directly downstream of an immediate early cytomegalovirus promoter (*Stu*I-*Sac*I fragment) isolated from pBC12/CMV/IL2 (12). The WHV pregenome-encoding sequence driven by the cytomegalovirus promoter was subsequently cloned between the *Hin*cII and *Eco*RI sites in the polylinker of the vector pUC119 (Fig. 1a). The plasmid served as a wild-type control since expression of the plasmid in permissive cells produced virus particles which were infectious when inoculated into a woodchuck (unpublished data).

A prokaryotic expression plasmid, pKK233/WHV/CORE, was kindly supplied by F. Schoedel. The plasmid consisted of the entire core gene of WHV and a *trc* promoter cloned in pKK233-2. We replaced the cloned core gene originally existing in the pKK233/WHV/CORE plasmid with the WHV8 core gene (Fig. 1b).

Construction of mutant genomes. In order to study the *hhr* region spanning amino acids 95 to 122 of the core protein, we used site-directed mutagenesis to introduce a series of point mutations resulting in individual or double replacement of relevant amino acids by alanine or proline (Fig. 2). Since the open reading frame of the core protein does not overlap the open reading frames of other viral proteins in the WHV genome, fully sequenced restriction fragments that contained the mutation of interest were substituted for the corresponding fragment of the infectious pUC119 CMV/WHV expression plasmid. The mutation of interest was also introduced into the core gene of pKK233/WHV/CORE expression plasmid. All clones were further characterized by sequencing and restriction enzyme analysis.

Cell culture and transfection. The human hepatoma cell line Huh7 was used for transient expression of WHV pregenome-encoding DNA. The Huh7 cells were maintained in Dulbecco's modified Eagle's medium–F-12 medium supplemented with 10% fetal bovine serum. The calcium phosphate coprecipitation procedure was carried out to transfect DNA into the Huh7 cells as previously described (43) with a modification. Briefly, 20 μg of DNA was added to a 60-mm dish containing 1.5×10^6 cells. Four hours later, cells were washed once with Dulbecco's modified Eagle's medium–F-12 medium and 1 ml of 15% glycerol solution was added to the cells with occasional rocking for 4 min. Cells were washed once again with the same medium, and then 4 ml of Dulbecco's modified Eagle's medium–F-12 medium with 10% fetal bovine serum was added. The cells were incubated at 37°C with 3% CO_2 for 4 days until transfected cells were harvested.

Induction of expression of WHV core mutants in bacterial cells. The pKK233/WHV/CORE mutants were expressed in DH5α cells as previously reported (13, 24) with minor modifications. Briefly, the wild-type core plasmid or a mutant core plasmid was transformed into *Escherichia coli* DH5α. The transformed cells were grown on a Luria-Bertani (LB) agar plate containing carbenicillin (100 μg/ml) and 2% glucose at 30°C overnight. A single colony of plasmid-containing cells was used to inoculate 2 ml of LB medium containing carbenicillin (100 μg/ml) and 2% glucose. The cells were allowed to grow at 30°C until the optical density at 600 nm reached approximately 0.4. Then, the cells were briefly washed once with LB medium containing only carbenicillin (100 μg/ml) and then cultured in LB medium containing carbenicillin (100 μg/ml) and isopropyl-β-D-thiogalactopyranoside (IPTG; 250 μM) in the absence of glucose at 30°C for 3 h.

Extraction and assay of intact viral nucleocapsids from transfected cells. Cells transfected with capsid mutants were frozen at −70°C for 30 min and lysed by adding 0.5 ml of lysis buffer (50 mM Tris-HCl [pH 8.0], 1 mM EDTA, 1% Nonidet P-40) and incubating at 37°C for 10 min. The intact capsid particles were extracted and subjected to electrophoresis in a 1% agarose gel. Subsequently, the capsids were transferred to a nylon or a nitrocellulose membrane that was probed for core antigens with a rabbit antibody to native core protein and ^{125}I-protein A. In order to detect viral nucleic acids within intact blotted capsid particles, the membrane was soaked in 0.2 N NaOH and 1.5 mM NaCl for 15 min, neutralized in 0.2 N Tris-HCl (pH 4.3) and 1.5 mM NaCl for 10 min, and then probed for viral RNA and plus-strand DNA or minus-strand DNA with ^{32}P-labelled strand-specific probes as previously described (45). In some cases, the blot was first probed with minus-strand riboprobe, stripped by washing with 0.2 N NaOH and 1.5 M NaCl for 30 min, then neutralized in 0.2 N Tris-HCl (pH 4.3) and 1.5 M NaCl for 15 min, and finally rehybridized with plus-strand riboprobe. When the blot was rehybridized, the radioautography exposure times were lengthened to compensate for any DNA losses during washing, and this resulted in a higher background.

SDS-polyacrylamide gel electrophoresis (PAGE) and Western blot (immunoblot) analysis. An equal volume of 2× sample loading buffer (0.125% M Tris-HCl [pH 6.8], 0.4% sodium dodecyl sulfate [SDS], 20% glycerol, 10% β-mercaptoethanol) was added to the cell lysate or bacterial suspension. Samples were heated in a boiling water bath for 5 min and applied to an SDS–12% polyacrylamide gel. Gel electrophoresis was performed at a constant current of 9 mA for 16 h. The proteins in the gel were transferred to a nitrocellulose membrane that was probed for core antigen with a rabbit anti-WHV core antibody and ^{125}I-protein A.

In vivo transfection of woodchucks with viral genomes and assay of serum. Transfection of woodchucks with monomeric WHV DNA with or without mutations was carried out as previously described (8). Briefly, 50 μg of the purified monomeric WHV DNA was inoculated into the liver of adult or neonatal woodchucks. For adult woodchucks, assays for WHV surface antigen, antibody against WHV surface antigen, and antibody against WHV core antigen were performed 1 month postinoculation and then twice monthly for 9 months. For neonatal woodchucks, the assays were performed at 3 months and then monthly thereafter for 1 year.

Cell fixation and transmission electron microscopy (EM). The DH5α cells transformed with pKK233/WHV/CORE mutant expression plasmids were pelleted by centrifugation at 900 × *g* for 10 min. The pellet was washed once with 10 ml of 1× phosphate-buffered saline and suspended in 5 ml of fixative solution (1% glutaraldehyde and 2% formaldehyde in 0.1 M sodium cacodylate buffer, pH 7.4) at room temperature for 2 h and stored at 4°C. The fixed samples were examined as previously described (17, 41). Briefly, the cell pellet was embedded

in 2% agar (45°C), which was allowed to solidify on ice, and then trimmed into small cubes for EM sample processing. The sample was rinsed with 0.1 M cacodylate buffer (pH 7.4) and postfixed by 1% osmium in the same buffer for 1 h at room temperature. The sample was briefly washed in 4.5% sucrose solution, dehydrated in graded ethanol (35, 50, 75, 95, and 100%) followed by propylene oxide, and infiltrated overnight in a 1:1 mixture of propylene oxide and LX-112 epoxy resin (Ladd). The sample was embedded in pure epoxy resin in Beem capsules and allowed to cure for 48 h in a 60°C oven. Thin sections, 60 to 70 nm, were made with an Ultra Cut E microtome (Leica) and mounted on 200-mesh copper grids. The thin sections were double stained in an Ultrostainer (Leica) with uranyl acetate and lead citrate and then carbon coated in a vacuum evaporator. The samples were photographed with a Hitachi H-7000 electron microscope operated at 100 kV.

RESULTS

In vivo assay of infectivity of a WHV genome with a His→Pro change in the *hhr*. Initially, we cloned supercoiled WHV DNA (WHV7-11 and WHV7-9) from the hepatocytes of a chronically infected woodchuck. Each recircularized monomeric genome was transfected into the liver of noninfected, neonatal woodchucks (11 for WHV7-11, 7 for WHV7-9) to determine infectivity. By 3 months posttransfection, surface and core antigen of WHV appeared in 10 of 11 woodchucks transfected with WHV7-11 DNA (9), indicating that the WHV7-11 genome was infectious. However, the WHV7-9 genome failed to initiate an infection in the transfected woodchucks, indicating that the WHV7-9 genome was defective in viral replication. DNA sequence analysis of the complete genomes showed that there was only a single nucleotide difference between the two genomes. This single nucleotide change, from an A in WHV7-11 to a C in WHV7-9 at nucleotide position 2331, resulted in the replacement of histidine 104 (in WHV7-11) with a proline (in WHV7-9) in the core protein. The amino acid sequence of other viral proteins encoded by WHV7-9 was not affected by this naturally occurring point mutation. Thus, this experiment demonstrated that a substitution of histidine 104 of core protein with a proline completely inhibited viral replication in vivo.

To determine the stage at which this histidine-to-proline mutation inhibited viral replication, we introduced the same mutation into the core gene of the eukaryotic expression plasmid, pUC119 CMV/WHV, and transfected the mutant genome into Huh7 cells. The phenotype of the mutant was determined at several different stages of viral replication, including synthesis of core protein, RNA packaging, DNA synthesis, and capsid assembly. The result showed that the replacement of histidine 104 of the core protein with a proline did not prevent core protein expression but did abolish capsid assembly (see below). Since histidine 104 is located in the *hhr* region, we speculated that the *hhr* region may be required for viral capsid assembly.

Molecular analysis of *hhr* mutants in eukaryotic cells. Since one of the major functions of the core protein in viral replication is to provide a protein shell for viral nucleic acids, it is not surprising that the *hhr* region of core protein may be involved with the assembly of core proteins into a capsid particle. The *hhr* domain contains four conserved hydrophobic residues: leucine 101, leucine 108, valine 115, and phenylalanine 122. We presumed that these specific hydrophobic amino acid residues might act coordinately in capsid assembly and that substitution with different hydrophobic residues might disrupt the function of this domain. In order to determine the importance of these specific hydrophobic residues in capsid assembly without dramatically changing the primary sequence or hydrophobicity of the core polypeptide, we first substituted alanine for specific hydrophobic amino acids at individual sites. All of the core mutants expressed core proteins in transfected Huh7 cells as assayed by SDS-PAGE and Western blot analysis (Fig. 3a).

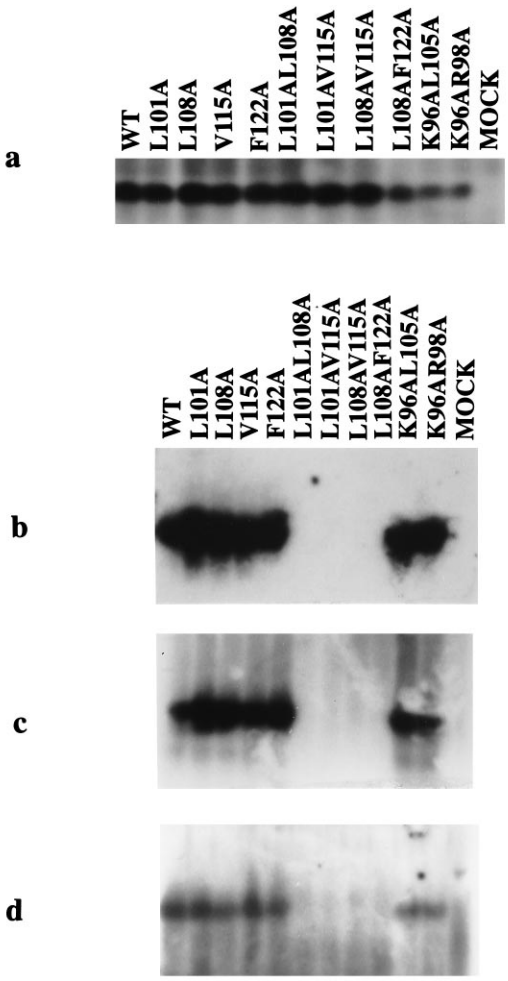


FIG. 3. Effects of alanine substitution mutations in the *hhr* domain on core protein synthesis, capsid formation, RNA packaging, and viral DNA synthesis in eukaryotic cells. Huh7 cells transfected with plasmids expressing the indicated capsid-defective genome were lysed, and the core proteins were denatured and analyzed by SDS-PAGE and then transferred to a filter which was probed for the core antigens with an antibody against the denatured core protein of WHV (a). The assembled capsids from the cytoplasm were analyzed by agarose gel electrophoresis and transferred to a filter for immunostaining with an antibody to native core protein of WHV (b) or detection of total plus-strand viral nucleic acids with minus-strand RNA probe (c) and examination of minus-strand and double-stranded DNA within intact capsid particles (d). The blot shown in panel d was originally hybridized with minus-strand RNA probe, stripped, and rehybridized with a plus-strand probe. WT, wild type. MOCK, nontransfected Huh7 cells.

The ability of each mutant protein to assemble into capsid was determined by electrophoresis of particles through a nondenaturing agarose gel followed by Western blot analysis. The results showed that mutants L101A, L108A, V115A, and F122A, which substituted alanine for the respective hydrophobic residue, produced an electrophoretic pattern similar to that produced by the wild-type capsids (Fig. 3b). Therefore, a single substitution of an alanine residue for any one of the four hydrophobic residues does not significantly disrupt the structure of an assembled capsid. To determine whether these mutant capsids were functional in other steps of replication besides assembly, we measured the viral nucleic acids contained within the capsids of each mutant compared with those of wild type. The results showed that the assembled capsids produced by the individual alanine substitution mutants contained normal levels of plus-strand nucleic acids, including viral RNA and

DNA, and normal levels of full-length minus-strand nucleic acids (Fig. 3c and d). Therefore, these altered capsid proteins were all able to encapsidate viral RNA pregenome and to synthesize viral DNA, as well as to assemble into capsids.

Next, the effect of double mutations in the *hhr* region was determined by substituting alanine for the hydrophobic residues in sets of two to yield mutants L101AL108A, L101AV115A, L108AV115A, and L108AF122A. All of the mutants produced core proteins in transfected Huh7 cells (Fig. 3a). However, assembled capsid particles of the double mutants were not detected by agarose gel analysis even though the mutants produced a normal level of core protein (Fig. 3b). These results indicated that double substitution of these four hydrophobic residues with another hydrophobic residue, alanine, completely abolished capsid assembly and suggested that these four hydrophobic residues may coordinately maintain the structure of an intact capsid particle. To exclude the possibility that a random double mutation in this region could result in the same defect in capsid assembly, we made two double mutations within this small region as controls. Either substitution of lysine 96 and arginine 98 with alanine residues (K96AR98A) or replacement of lysine 96 and leucine 105 with alanine residues (K96AL105A) resulted in a normal level of assembled capsid particles containing viral nucleic acids (Fig. 3b, c, and d), suggesting that the particular four hydrophobic amino acid residues studied in the *hhr* domain are required for capsid assembly.

Since we showed that the identity of the amino acid at positions 101, 108, 115, and 122 was critical for capsid assembly, it was of interest to determine whether the spacing of the four individual hydrophobic residues was important as well. To address this question, we made three types of mutants. In the first, an alanine residue was inserted after histidine 104 (104A105) or after phenylalanine 118 (118A119). In the second type of mutant, histidine 104 was deleted. In the third type, histidine 104 was replaced by an alanine. Each type of mutant produced core proteins in transfected Huh7 cells (Fig. 4a). Both the first type and the second type of mutation completely inhibited capsid assembly (Fig. 4b). However, the third type of mutation did not prevent capsid assembly, RNA encapsidation, or DNA synthesis (Fig. 4b and c). These results suggested that the correct spacing of the four hydrophobic residues was necessary for capsid assembly.

Since these four hydrophobic residues in the *hhr* region appeared to be sensitive to any double substitution rather than sensitive to a single residue substitution with alanine, the *hhr* region may contain a critical structure that is essential for core-core interaction. Because proline functions as a helix breaker and is able to dramatically change the secondary structure of proteins, we introduced proline mutations into this region to determine whether an alteration in the secondary structure of the *hhr* region was detrimental to capsid assembly. Since the individual substitution of leucine 101, histidine 104, or leucine 108 with alanine did not inhibit capsid assembly, we chose these three residues as targets to examine the effects of a proline substitution. Therefore, leucine 101, histidine 104, or leucine 108 was individually replaced by proline to yield three proline substitution mutants, L101P, H104P, and L108P. All three mutants produced a normal level of core proteins as assayed by SDS-PAGE and Western blot analysis (Fig. 4a) but failed to show any signal for capsid particles in the native agarose gel assay compared with the signal from a wild-type capsid (Fig. 4b). These results supported the argument that the *hhr* region functions as a domain that is required for capsid assembly.

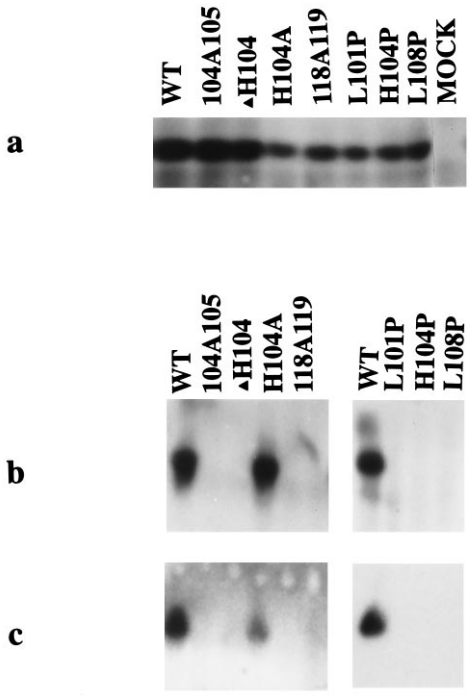
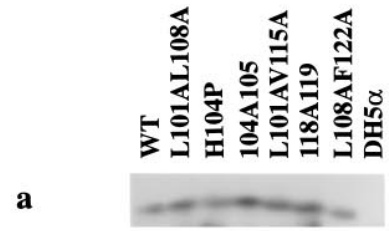


FIG. 4. Effects of spacing changes and proline mutations in the *hhr* domain on capsid assembly and viral RNA packaging. Huh7 cells were transfected with a plasmid expressing the indicated capsid-defective genome. The patterns of denatured core proteins, intact capsid particles, and total plus-strand nucleic acids within capsids of the mutants were analyzed by SDS-PAGE and Western blot (a) by agarose gel electrophoresis and Western blot (b), or by hybridization of plus-strand nucleic acids (c). WT, wild type. MOCK, nontransfected Huh7 cells.

EM examination of capsid assembly-defective mutants in prokaryotic cells. In order to directly observe the assembly situation of all assembly-defective mutants in intact cells, we reexamined by EM all of the mutants in bacterial cells that were defective in capsid assembly as determined by a native agarose gel assay. Core proteins of human HBV and duck HBV (DHBV) assemble efficiently in bacteria into capsid particles (4, 11, 14, 31) that are clearly visible by EM (41). Therefore, we introduced each of the above double mutations or proline mutations into the WHV core gene of a prokaryotic expression plasmid. All of the mutants produced core proteins in bacteria as assayed by SDS-PAGE and Western blot analysis (Fig. 5a). The assembled capsid particles of wild type were observed within sectioned cells by EM (Fig. 5b-1). The assembled capsid particles were found in approximately 5% of the cells (50 cells positive in 1,000 examined) and ranged in size from 20 to 25 nm. In contrast, the altered core proteins with double mutations or proline mutations failed to form standard capsid particles. Capsid particles of the mutants were not found in the cells or in the inclusion bodies (Fig. 5b-2 to -7). However, the mutants containing a single substitution of one of the four hydrophobic residues with alanine assembled capsid particles at a normal level in bacterial cells (data not shown). These results provided further evidence that the native agarose gel technique was able to reliably determine the phenotype of deficient assembly.

DISCUSSION

The results of our experiments provided evidence that a small region of the core protein of WHV that contained four hydrophobic residues (i.e., leucine 101, leucine 108, valine



a

b

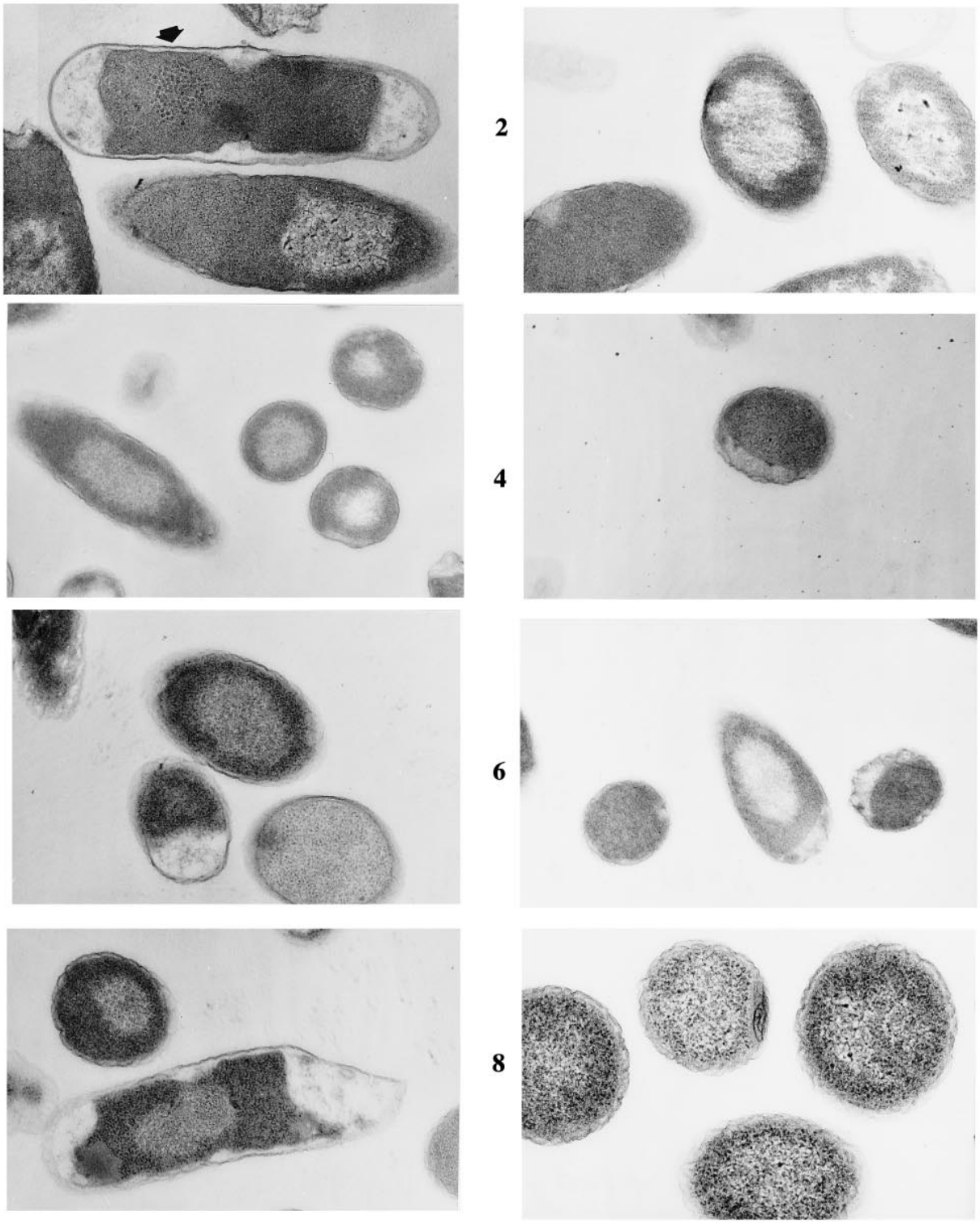


FIG. 5. Expression and EM examination of core mutants in prokaryotic cells. DH5α cells transformed with plasmids expressing a mutant core protein were induced by IPTG and lysed. (a) The denatured mutant core proteins were analyzed by SDS-PAGE and Western blot. WT, wild type. (b-1) The intact capsid particles were examined by EM in fixed DH5α cells. The samples were examined with a Hitachi H-7000 electron microscope at 100 kV. The wild-type capsid particles, ranging in size from 20 to 25 nm, are indicated by a bold arrow. (b-2 to -7) EM photographs of bacterial cells that contained the capsid assembly-deficient mutants are shown as follows: 2, L101AL108A; 3, H104P; 4, 104A105; 5, L101AV115A; 6, 118A119; 7, L108AF122A. (b-8) An EM photograph of non-transformed DH5α cells is shown as a control.

115, and phenylalanine 122) was required for capsid assembly. Replacement of one of the four individual hydrophobic residues with alanine did not disrupt capsid assembly in the cytoplasm. However, combinations of substitutions of any two hydrophobic residues with alanine residues resulted in a failure to detect formation of capsids in the cytoplasm. In contrast, a double mutation of two amino acids other than these four (K96AR98A and K96AL105A) within this small region did not perceptibly influence capsid assembly. These results suggested that core proteins with mutations in two of the four critical positions were not able to assemble in the cytoplasm and that a coordination of hydrophobic residues in this region was important for capsid assembly. Since capsid formation cannot tolerate randomly paired substitutions in the *hhr* region, the particular identity of each residue at the seventh position seems important for the mechanism of capsid assembly for WHV.

Clearly, the *hhr* is a critical region for capsid function. The spacing within the *hhr* is important for the maintenance of the interactions of the core proteins of WHV in the cytoplasm. Mutants containing either deletion of a single histidine residue at position 104 between leucine 101 and leucine 108 or insertion of an alanine residue after histidine 104 or after phenylalanine 118 were defective in capsid assembly. However, when histidine 104 was substituted with alanine and the distance between these two leucines was unchanged, the capsid proteins assembled in the cytoplasm with an efficiency comparable to that of wild-type proteins. These results suggested that the distance of six amino acids between two specific hydrophobic residues was essential for the maintenance of the conformation of assembled capsid proteins. On the other hand, an alteration of secondary structure in the *hhr* by introduction of a single proline residue into different positions of this region resulted in defective capsid assembly. Furthermore, a WHV genome bearing a single proline substitution (H104→P) in the *hhr* of the core protein failed to initiate viral replication in vivo. Presumably, the *hhr* domain could tolerate a minor change, such as a single substitution of one of the four hydrophobic residues with another hydrophobic residue, but could not tolerate a major change, such as an insertion or a deletion. These structural characteristics were also found in a typical leucine zipper domain in which every seventh residue in the heptad is a leucine. A single substitution of the leucine at every seventh position with another hydrophobic residue or even with a charged residue in a leucine zipper domain did not affect protein-protein interaction while a double or triple substitution did (15, 22, 23, 29, 36). The leucine zipper function was also destroyed by an insertion or deletion of one or two residues (29). Therefore, it seems reasonable to postulate that these four hydrophobic residues in the *hhr* may have a structural role that contributes to protein-protein interaction.

The *hhr* domain is structurally and functionally conserved in the hepadnavirus family. Leucine 101, leucine 108, and valine 115 in the core protein of WHV are invariant with respect to HBV, DHBV, ground squirrel hepatitis virus, and heron hepatitis virus core proteins. Phenylalanine 122 in the core protein of WHV is highly conserved in all mammalian, but not avian, hepadnaviruses. The amino acid at the position corresponding to phenylalanine 122 in the core protein of WHV is a leucine in DHBV and heron HBV (Fig. 6). We found that substitution of the first hydrophobic residue, leucine, with proline in the *hhr* of DHBV resulted in a defect in the capsid assembly of DHBV (data not shown). Weizsacker et al. reported that a region encompassing 196 amino acids of the DHBV core protein can be functionally replaced by a corresponding region containing 120 amino acids of the HBV core protein (39). These two functionally exchangeable regions both contain the *hhr* domain. Chang et al. (7) also found that truncation of 44 amino acids from the C terminus of HBV core did not affect the interaction between core monomers. However, deletion of 68 amino acids from the C terminus prevented such interaction (7). We suggest that this C-terminal truncation partially destroyed the *hhr* domain. Moreover, Akarca and Lok recently examined the entire core gene of HBV from 69 patients with chronic HBV infection and found that nucleotide mutations rarely occurred in a region encoding amino acids 101 to 124 of the core protein, indicating that this region is probably important for viral replication (1). Interestingly, this domain encompasses the *hhr* region. These findings suggested that the *hhr* domain of the core protein was important in capsid assembly of all hepadnaviruses. Although the *hhr* region is important for viral replication, we still do not understand why four hydrophobic residues are conserved and yet apparently three of the four conserved residues are sufficient for the function of capsid assembly. We can only speculate at this time whether the fourth residue provides an extra degree of stability to the virus particle or perhaps provides a margin of error in case of a minor change at any one of the four particular positions.

```
WHV    -GLKVRQSLWFHLSCLTFGQHTVQEFLVSFGVWIRT-
GSHV   -GLKVRQTLWFHLSCLTFGQHTVQEFLVSFGVEIRT-
HBV    -GLKFRQLLWFHISCLTFGRETVLEYLVSFGVWIRT-
HERON  -LDKARRLLWWHYNCLLWGEATVTNYISRLRTWLST-
DHBV   -LDRARRLLWWHYNCLLWGEAQVTNYISRLRTWLST-
```

FIG. 6. Comparison of the amino acid sequences in the *hhr* domain of the core protein in the hepadnavirus family including WHV, ground squirrel hepatitis virus, HBV, heron hepatitis virus, and DHBV. The conserved hydrophobic amino acid residue located at every seventh position in this region is indicated in boldface.

The core proteins are assembled into a highly organized icosahedral structure that contains the viral genome and polymerase. This is a complicated and highly ordered process in which various mechanisms may be involved. A cysteine-cysteine cross-linking of capsid particles (28) and amino acids 37 to 67 of the core protein of HBV (7) and DHBV (39, 41) appeared to be involved in this process. Therefore, although we provided evidence to demonstrate that the *hhr* was associated with capsid assembly, we cannot exclude the possibility that some of the amino acid residues in this region may participate in other assembly pathways of the capsid particle through a domain that overlaps the *hhr* region.

Since the *hhr* contains a four-heptad repeat consisting of a hydrophobic residue at every seventh position, the sequence appears to resemble a leucine zipper motif. Recently, similar structures have been reported in other proteins (3, 6, 20, 21, 26, 35, 38). These structures were different from that of a typical leucine zipper, and some did not even exhibit a coiled-coil structure by computer analysis (10). However, all these structures were characterized by a heptad repeat of hydrophobic residues and involvement in protein-protein interaction. It was established that the structure of the leucine zipper may not be strictly limited to a leucine residue at every seventh position, and although there is a distinct preference for leucine (30), the seventh amino acid may be another hydrophobic residue (5, 15, 19, 36). These analogous structures also possess a dimerization capacity for protein-protein interaction (3, 20). The hepadnaviral core monomers also prefer to form dimers in *cis* (7, 46). Therefore, we speculate that the *hhr* of the core protein may play a role in the dimerization of the core monomers at the early stage of hepadnaviral capsid assembly.

ACKNOWLEDGMENTS

We thank Florian Schoedel for kindly providing plasmid pKK233/WHV/CORE, Paul Cote for kindly providing WHV core antibody, Kunio Nagashima for assistance with EM, Tatiana Tsareva for synthesis and purification of oligonucleotides, Ying Huang and Marianne Lewis for technical assistance, and Theo Heller for helpful discussion.

REFERENCES

1. **Akarca, U. S., and A. S. F. Lok.** 1995. Naturally occurring hepatitis B virus core gene mutations. Hepatology **22:**50–60.
2. **Beames, B., and R. E. Lanford.** 1995. Insertions within the hepatitis B virus capsid protein influence capsid formation and RNA encapsidation. J. Virol. **69:**6833–6838.
3. **Bernstein, H. B., S. P. Tucker, S. R. Kar, S. A. Mcpherson, D. T. Mcpherson, J. W. Dubay, J. Lebowitz, R. W. Compans, and E. Hunter.** 1995. Oligomerization of the hydrophobic heptad repeat of gp41. J. Virol. **69:**2745–2750.
4. **Birnbaum, F., and M. Nassal.** 1990. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. J. Virol. **64:**3319–3330.
5. **Busch, S. J., and P. Sassone-Corsi.** 1990. Dimers, leucine zippers and DNA-binding domains. Trends Genet. **6:**36–40.
6. **Chan, Y.-L., J. Olvera, A. Gluck, and I. G. Wool.** 1994. A leucine zipper-like motif and a basic region-leucine zipper-like element in rat ribosomal protein L13a. J. Biol. Chem. **269:**5589–5594.
7. **Chang, C., S. Zhou, D. Ganem, and D. N. Standring.** 1994. Phenotypic mixing between different hepadnavirus nucleocapsid proteins reveals C protein dimerization to be *cis* preferential. J. Virol. **68:**5225–5231.
8. **Chen, H.-S., S. Kaneko, R. Girones, R. W. Anderson, W. E. Hornbuckle, B. C. Tennant, P. J. Cote, J. L. Gerin, R. H. Purcell, and R. H. Miller.** 1993. The woodchuck hepatitis virus X gene is important for establishment of virus infection in woodchucks. J. Virol. **67:**1218–1226.
9. **Cohen, J. I., R. H. Miller, B. Rosenblum, K. Denniston, J. L. Gerin, and R. H. Purcell.** 1988. Sequence comparison of woodchuck hepatitis virus replicative forms shows conservation of the genome. Virology **162:**12–20.
10. **Coll, J. M.** 1995. Heptad-repeat sequences in the glycoprotein of rhabdoviruses. Virus Genes **10:**107–114.
11. **Crowther, R. A., N. A. Kiselev, B. Bottcher, J. A. Berriman, G. P. Borisova, V. Ose, and P. Pumpens.** 1994. Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. Cell **77:**943–950.
12. **Cullen, B. R.** 1987. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell **46:**972–982.
13. **Edman, J. C., R. A. Hallewell, P. Valenzuela, H. M. Goodman, and W. J. Rutter.** 1981. Synthesis of hepatitis B surface and core antigens in *E. coli*. Nature (London) **291:**503–506.
14. **Gallina, A., F. Bonelli, L. Zentilin, G. Rindi, M. Muttini, and G. Milanesi.** 1989. A recombinant hepatitis B core antigen polypeptide with the protamine-like domain deleted self-assembles into capsid particles but fails to bind nucleic acids. J. Virol. **63:**4645–4652.
15. **Gentz, R., F. J. Rauscher, C. Abate, and T. Curran.** 1989. Parallel association of Fos and Jun leucine zipper juxtaposes DNA-binding domains. Science **243:**1695–1699.
16. **Girones, R., P. J. Cote, W. E. Hornbuckle, B. C. Tennant, J. L. Gerin, R. H. Purcell, and R. H. Miller.** 1989. Complete nucleotide sequence of a molecular clone of woodchuck hepatitis virus that is infectious in the natural host. Proc. Natl. Acad. Sci. USA **86:**1846–1849.
17. **Gonda, M. A., D. L. Fine, and M. Gregg.** 1978. Squirrel monkey retrovirus: electron microscopy of a virus from new world monkeys and comparison with Mason-Pfizer monkey virus. Arch. Virol. **56:**297–307.
18. **Hatton, T., S. Zhou, and D. N. Standring.** 1992. RNA- and DNA-binding activities in hepatitis B virus capsid protein: a model for their roles in viral replication. J. Virol. **66:**5232–5241.
19. **Hemmerich, P., A. von Mikecz, F. Neumann, O. Sozeri, G. Wolff-Vorbeck, R. Zoebelein, and U. Krawinkel.** 1993. Structural and functional properties of ribosomal protein L7 from humans and rodents. Nucleic Acids Res. **21:**223–231.
20. **Ho, C.-Y., J. G. Adamson, R. S. Hodges, and M. Smith.** 1994. Heterodimerization of the yeast MATa1 and MATα2 proteins is mediated by two leucine zipper-like coiled-coil motifs. EMBO J. **13:**1403–1413.
21. **Ing, Y. L., I. W. L. Leung, H. H. Q. Heng, L.-C. Tsui, and N. J. Lassam.** 1994. MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain and a leucine zipper-basic region domain. Oncogene **9:**1745–1750.
22. **Kouzarides, T., and E. Ziff.** 1988. The role of the leucine zipper in the fos-jun interaction. Nature (London) **336:**646–651.
23. **Landschulz, W. H., P. E. Johnson, and S. L. McKnight.** 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science **243:**1681–1688.
24. **Leadbetter, G., and K. Murray.** 1979. Hepatitis B virus genes and their expression in E. coli. Nature (London) **282:**575–579.
25. **Liao, W., and J.-H. Ou.** 1995. Phosphorylation and nuclear localization of the hepatitis B virus core protein: significance of serine in the three repeated SPRRR motifs. J. Virol. **69:**1025–1029.
26. **Mukai, H., and Y. Ono.** 1994. A novel protein kinase with leucine zipper-like sequences: its catalytic domain is highly homologous to that of protein kinase C. Biochem. Biophys. Res. Commun. **199:**897–904.
27. **Nassal, M.** 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. J. Virol. **66:**4107–4116.
28. **Nassal, M., A. Rieger, and O. Steinau.** 1992. Topological analysis of the hepatitis B virus core particle by cysteine-cysteine cross-linking. J. Mol. Biol. **225:**1013–1025.
29. **Neuberg, M., M. Schuermann, J. B. Hunter, and R. Muller.** 1989. Two functionally different regions in Fos are required for the sequence-specific DNA interaction of the Fos/Jun protein complex. Nature (London) **338:**589–590.
30. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1989. Evidence that the leucine zipper is a coiled coil. Science **243:**538–542.
31. **Salfed, J., E. Pfaff, M. Noah, and H. Schaller.** 1989. Antigenic determinants and functional domains in core antigen and e antigen from hepatitis B virus. J. Virol. **63:**798–808.
32. **Seifer, M., and D. N. Standring.** 1994. A protease-sensitive hinge linking the two domains of the hepatitis B virus core protein is exposed on the viral capsid surface. J. Virol. **68:**5548–5555.
33. **Summers, J., P. Smith, and A. Horwich.** 1990. Hepadnaviral envelope proteins regulate amplification of covalently closed circular DNA. J. Virol. **64:**2819–2824.
34. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc. Natl. Acad. Sci. USA **75:**4533–4537.
35. **Tsurugi, K., and K. Mitsui.** 1991. Bilateral hydrophobic zipper as a hypothetical structure which binds acidic ribosomal protein family together on ribosomes in yeast *Saccharomyces cerevisiae*. Biochem. Biophys. Res. Commun. **174:**1318–1323.
36. **Turner, R., and R. Tjian.** 1989. Leucine repeats and an adjacent DNA-binding domain mediate the formation of functional c-fos-c-jun heterodimers. Science **243:**1689–1694.
37. **Tuttleman, J., C. Pourcel, and J. Summers.** 1986. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. Cell **47:**451–460.
38. **Wang, S.-Z., and R. Adler.** 1994. A developmental regulated basic-leucine zipper-like gene and its expression in embryonic retina and lens. Proc. Natl. Acad. Sci. USA **91:**1351–1355.
39. **Weizsacker, F., S. Wieland, and H. E. Blum.** 1995. Identification of two separable modules in the duck hepatitis B virus core protein. J. Virol. **69:**2704–2707.
40. **Werner, B. G., J. M. Smolec, R. Snyder, and J. Summers.** 1980. Serological relationship of woodchuck hepatitis virus to human hepatitis B virus. J. Virol. **37:**239–243.
41. **Yang, W., J. Guo, Z. Ying, S. Hua, W. Dong, and H. Chen.** 1994. Capsid assembly and involved function analysis of twelve core protein mutants of duck hepatitis B virus. J. Virol. **68:**338–345.
42. **Yeh, C.-T., Y.-F. Liaw, and J. H. Ou.** 1990. The arginine-rich domain of hepatitis B virus precore and core protein contains a signal for nuclear transport. J. Virol. **64:**6141–6147.
43. **Yu, M., and J. Summers.** 1991. A domain of the hepadnavirus capsid protein is specifically required for DNA maturation and virus assembly. J. Virol. **65:**2511–2517.
44. **Yu, M., and J. Summers.** 1994. Phosphorylation of the duck hepatitis B virus capsid protein associated with conformational changes in the C terminus. J. Virol. **68:**2965–2969.
45. **Yu, M., and J. Summers.** 1994. Multiple functions of capsid protein phosphorylation in duck hepatitis B virus replication. J. Virol. **68:**4341–4348.
46. **Zhou, S., and D. N. Standring.** 1992. Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. Proc. Natl. Acad. Sci. USA **89:**10046–10050.