Reverse transcription-associated dephosphorylation of hepadnavirus nucleocapsids

David H. Perlman*1, Eric A. Berg†‡, Peter B. O’Connor‡, Catherine E. Costello‡, and Jianming Hu*‡§

*Department of Microbiology and ‡Mass Spectrometry Resource, Boston University School of Medicine, Boston, MA 02118; and §Department of Microbiology and Immunology, Pennsylvania State University College of Medicine, Hershey, PA 17033

Edited by Francis V. Chisari, The Scripps Research Institute, La Jolla, CA, and approved May 9, 2005 (received for review March 24, 2005)

Hepatitis B viruses (HBVs), or hepadnaviruses, are pararetroviruses that replicate their DNA genome via an RNA intermediate (pregenomic RNA, pgRNA) by reverse transcription (1, 2). Similar to classical retroviruses, hepadnaviruses begin assembly with the formation of intracellular nucleocapsid (NC) particles that specifically package viral pgRNA. However, they undergo reverse transcription before membrane envelopment of NCs and secretion as virions, rather than subsequent to entry into new target cells, like classical retroviruses (2, 3). This replication strategy is manifested in the hepadnaviral maturation phenomenon: secreted hepadnaviral virions contain only the mature, double-stranded relaxed circular (or the mature, double-stranded linear) DNA species, whereas intracellular pools of NCs contain a mix of nucleic acid species, including the pgRNA and DNA species from all of the intermediate stages of reverse transcription (2).

The maturation phenomenon implies that mature DNA-containing NCs may be selectively enveloped and secreted. It was proposed that DNA synthesis and NC envelopment may be coupled through a maturation signal (2). Strong support for this maturation signal model has been obtained (4–6), particularly through the use of a synchronized duck HBV (DHBV) replication system (7, 8) that demonstrated that mature NCs alone are endowed with intrinsic secretion competence, thereby effectively excluding other alternative models (9). Yet, the nature of this maturation signal, or of any maturation-associated dynamic changes in the NCs, other than the differences in the NCs’ nucleic acid content per se, remains unknown.

The viral core protein, which forms the viral NC shell that packages the viral genome, has been implicated as a mediator of the hepadnaviral maturation signal by accumulating evidence. Its N-terminal domain is sufficient for assembly of the capsid shell (10–13), whereas its extremely basic C-terminal domain (CTD) can bind nucleic acid and is essential for packaging pgRNA into NCs and for progression of reverse transcription within these NCs (6, 12, 14–17). Mutations in the HBV core protein can, indeed, affect NC envelopment and secretion (18–22). Also, the CTD contains three S-P motifs (plus one T-P motif for DHBV) that have so far been identified as major phosphorylation sites (23, 24), mutations to which affect various stages of the viral lifecycle, including pgRNA packaging and DNA synthesis (16, 17, 23, 25–28). Furthermore, it has been shown that, whereas the CTD phosphosites within intracellular NCs are heavily and heterogeneously phosphorylated, those within secreted virions are hypophosphorylated or nonphosphorylated (29). Still, few details are known of the relationship between CTD dephosphorylation events and the intracellular NC maturation process.

Using the DHBV model, we have developed a method to obtain separated NCs from the three extremes of hepadnaviral maturation, i.e., immature (RNA-containing), mature (dsDNA-containing), and secreted virion-derived NCs. We have undertaken comprehensive MS analyses of the core protein from these NCs, using recently developed techniques and instrumentation to study labile posttranslational modifications (30). Our analyses have uncovered additional core modifications and revealed how these core modifications change with NC maturation.

Methods

Preparation of Virions and Virion-Derived NCs. DHBV virions were purified from the culture medium of D2 cells (31) as described (9). Virion-positive fractions were desalted and exchanged into TNE P/P/β [10 mM Tris, pH 8.0/1 mM EDTA, pH 8.0/100 mM NaCl/a mixture of phosphatase inhibitors (10 mM NaF, 10 mM sodium pyrophosphate, 2 mM Na3VO4, and 50 mM β-glycero-phosphate)/Complete protease inhibitors (Roche Molecular Biochemicals)/0.05% β-mercaptoethanol] by using Centricron YM-100 microconcentrators (Millipore). Subsequently, they were exchanged into a modified HCB2 (32) P/P/β-T [20 mM Tris-HCl, pH 7.5/50 mM NaCl/I 1 mM EDTA/0.11% (vol/vol) Triton X-100/0.1% (vol/vol) Nonidet P-40/the phosphatase inhibitor mixture/Complete protease inhibitors/0.05% β-mer-

1This paper was submitted directly (Track II) to the PNAS office.
2Abbreviations: CAD, collisionally activated dissociation; CTD, C-terminal domain; ACN, acetonitrile; HBV, hepatitis B virus; DHBV, duck HBV; FT-ICR, Fourier transform ion cyclotron resonance; NC, nucleocapsid; pgRNA, pregenomic RNA; Q-oTOF, quadrupole orthogonal TOF; SORI, sustained off-resonance irradiation; VC, vibrational cooling.
3D.H.P. and E.A.B. contributed equally to this work.
4To whom correspondence should be addressed. E-mail: juh13@psu.edu.

© 2005 by The National Academy of Sciences of the USA
Separation and Purification of Intracellular NCs of Different Maturity.

Intracellular NCs were isolated from D2 cells, or Dstet-5 cells, which express a surface (S) gene deletion mutant of DHBV (33), as described (7), with the following modifications. The mixture of phosphatase and protease inhibitors was added to all buffers, and HCB2 P/P/β [the same as HCB2 P/P/β-T, except with 0.01% (vol/vol) Triton X-100] was used to make up the sucrose gradient solutions. One milliliter of NC sample was layered over a 15–30% linear sucrose gradient (32.5 ml) and spun in a Sorvall Surespin 630 rotor for 4 h at 27,000 rpm. NCs with mixed maturity (mixed NCs) were harvested from the peak core fractions. To further purify the mature NCs, fractions containing mature NCs were pooled and subjected to repeated rounds of the same gradient centrifugation. Immature NCs were harvested from the top fractions containing abundant amounts of core protein and viral pgRNA, but little or no viral DNA. Approximately 100 μg of immature NCs and 1 μg of mature NCs could be purified from 10 × 100-mm dishes of cells.

Southern and Western Analyses. NC DNA extraction (7, 8) and Southern blotting (34, 35) were performed as described. SDS/PAGE and Western blotting were conducted by using polyclonal anti-DHBV core protein antibodies (36) as described (37). Purified NCs, like histones (38), required the addition of 5 μg of proteamine sulfate per lane to prevent extensive smearing.

Digestion of the Core Protein and Preparation of Peptides. Protein bands were visualized by zinc staining (Zink Stain Kit, Bio-Rad) after SDS/PAGE and digested in-gel (39), using a 1:2 or a 1:20 weight ratio of endoproteinase Lys-C or trypsin to core protein. Peptides were eluted with 20 mM ammonium bicarbonate, pH 8.8, then with 50% acetonitrile (ACN)/1% trifluoroacetic acid (TFA), and finally with 100% ACN, using mild sonication. Eluates were dried down, then resuspended in 0.1% TFA or a step gradient of 10%, 20%, 30%, 40%, and 50% ACN in 0.1% TFA or a step gradient of 10%, 20%, 30%, 40%, and 50% ACN in 0.1% TFA. Hydrophilic core phosphopeptides were typically eluted in the 20% ACN fraction.

MALDI-TOF MS. Core peptides were analyzed by using a Reflex IV (Bruker, Billerica, MA) MALDI (40) TOF MS in positive ion. Peptides were spotted onto an Anchor Chip (Bruker) MALDI target by using the dried droplet technique with a 2.5-dihydroxybenzoic acid matrix. Both external (Bruker) and internal (enzyme autolysis and sequenced core peptides) standards were used for spectral calibration. Acquired MS data were analyzed by using XTOF (Bruker) and other MS/MS software (freeware edition, Genomic Solutions, Ann Arbor, MI, http://65.219.84.5/moverzhtml), and experimental mass values were compared to theoretical values generated by the online program MS-DIGEST (http://prospector.ucsf.edu) or the sequence analysis software GPMAW (Lighthouse Data, Odense, Denmark).

Vibrational Cooling (VC) MALDI-Fourier Transform Ion Cyclotron Resonance (FT-ICR) MS. Core peptides were analyzed with a modified IonSpec FT-ICR mass spectrometer with a Cryomagnetics 7T actively shielded superconducting electromagnet by using a capacitively coupled closed cylindrical ICR cell with a home-built VC MALDI source (30, 41). MALDI was performed by using a nitrogen laser (337 nm). Peptide solutions were spotted (1:1 ratio) with 10 mg/ml α-cyano-4-hydroxycinnamic acid, desorbed with a 10-shot laser pulse sequence, and accumulated for 0.5 s into ~1–10 mbar nitrogen cooling gas. During accumulation, the pulse valve, backed by 15 mbar nitrogen, was pulsed for 4 ms, and ions were allowed to cool for 3 s. For MS/MS tandem mass spectrometry, ions of interest were isolated by using stored waveform inverse Fourier transform axial excitation/ejection (42) and fragmented by using sustained off-resonance irradiation (SORI)-collisionally activated dissociation (CAD) (43). Briefly, ions were irradiated with an off-resonance (1.5%) frequency at 10 V pp during a 4-ms gas pulse (44). For MS/MS/MS experiments, ions were allowed to cool, and another SORI-CAD experiment was applied, using an off-resonant frequency with the resultant peptide fragment (m/z 1,402,760 minus H2PO4) as the basis for calculation. Data were analyzed by using Boston University Data Analysis (BUDA) software developed at the Boston University School of Medicine Mass Spectrometry Resource. Spectra were calibrated either externally by using peptide standards or internally with definitively sequenced peptides derived from the core protein. Mass accuracy was >3 ppm, with typical broadband mass resolution of ~1:50,000.

Electrospray Ionization Quadrupole Orthogonal TOF (Q-oTOF) MS. Hydrophilic core peptides were analyzed by using an MDS Sciex/Applied Biosystems QStar Pulsar i triple Q-oTOF mass spectrometer (Applied Biosystems) with a Protona nanospray source (Protona, Odense, Denmark) (45). Desalted peptide samples were dried in a centrifugal evaporator and resuspended in 50% ACN in 1% formic acid. Nanospray tips were pulled in-house to 1 μm diameter by using a Sutter Instruments (Novato, CA) P-97 micropipette puller. Spectra were acquired at a spray voltage of 1,200–1,400 V. Tandem MS (CAD) were acquired by using user-controlled ramped collision energies (18–50 V) and helium as the collision gas. Data were analyzed with ANALYST software (Applied Biosystems) by manually reconstituting resultant spectra and comparing experimental data with theoretical fragment masses.

Fig. 1. Purification of immature and mature DHBV NC species. (Left) Intracellular NCs were harvested from Dstet-5 cells (33) and separated into immature NCs and mature NCs by successive rounds of velocity gradient centrifugation. NC DNA was isolated from gradient fractions and analyzed by Southern blotting. (Inset Left) Further separation of immature and mature NC species by subsequent rounds of velocity gradient ultracentrifugation. NCs from pooled mature fractions (as marked by brackets) showed a DNA size marker (M). (Right) Further separation of immature and mature NC species by successive rounds of velocity gradient ultracentrifugation. NC DNA was isolated from gradient fractions and analyzed by Southern blotting. (Inset Right) The detection of immature, pgRNA-containing NCs (RNA NC) in one of the top fractions by resolving the NCs on a native agarose gel, followed by probing with a riboprobe specific for the pgRNA (37).
Results

Separation and Purification of NC Species of Different Maturity. To ascertain whether hepadnaviral NC maturation is associated with any NC biochemical or biophysical changes other than the genome maturity itself, we sought to obtain NCs from the three extremes of maturation (i.e., immature, pgRNA-containing NCs; mature, dsDNA containing NCs; and secreted virion-derived NCs) for analysis. DHBV virions were purified from the culture supernatant of D2 cells, which replicate and secrete WT DHBV NCs) for analysis. DHBV virions were purified from the culture supernatant of D2 cells, which replicate and secrete WT DHBV (31), with a stringent, two-step procedure as described (9). Virion-derived NCs were then purified from lysed virions by velocity gradient centrifugation. Intracellular immature and mature NCs were isolated from both the D2 and surface-negative Dstet5 cells (33).

As evidenced by Southern analysis of viral DNA isolated from gradient fractions (Fig. 1), we noted that it was possible to separate the immature and mature NC species by velocity gradient ultracentrifugation, likely because of different biophysical characteristics (e.g., mass) resulting from their difference in nucleic acid content. The NC species containing the more mature (i.e., longer) DNA genomes traveled to the bottom of the velocity gradient, whereas the NCs containing the less mature (i.e., shorter) DNA and viral pgRNA remained retarded near the top to various degrees depending on their maturity (Fig. 1 Left and data not shown). Pooling fractions containing the most mature NCs and subjecting the pooled NCs to second rounds of gradient ultracentrifugation led to further separation of immature and mature species (Fig. 1 Right). By repeating this cycle four times, we were able to achieve complete separation of the mature dsDNA-containing NCs from any less mature DNA-containing NCs or completely immature NCs. Southern blotting showed that the purified mature NCs contained only the mature dsDNA, like virions (Fig. 2 A). Furthermore, there was a one-to-one stoichiometry of NC particle to genome equivalent in the purified mature NCs, an estimate obtained by quantifying the amount of core protein by Western blotting and the number of genome equivalents by Southern blotting (Fig. 2 and data not shown). Immature NCs were harvested from fractions near the top of each gradient, which contained core proteins and pgRNA but no detectable viral DNA (Fig. 1 and data not shown). Therefore, this procedure enabled us to obtain pure mature NCs, as well as immature NCs, in quantities sufficient for extensive biochemical and structural analysis. In addition, large amounts of NCs of mixed maturity were harvested from the peak core fractions (Fig. 1). Dephosphorylation of Mature NCs as Detected by SDS PAGE. Unfractionated intracellular DHBV core protein has been shown to

Fig. 2. Maturation-associated dephosphorylation of hepadnaviral NCs revealed by SDS PAGE. Intracellular NCs were purified as described in Fig. 1. Virions were isolated from the culture supernatant of D2 cells (9) and lysed with detergent, and virion-derived NCs were isolated by velocity gradient centrifugation. (A) Southern blot analysis of DNA from mature NCs after four rounds of enrichment by velocity gradient ultracentrifugation (lane 1). Virion DNA served as a marker for maturity (lane 2). (B) Western blot analysis of mixed maturity NCs harvested from D2 cells (lane 1), immature NCs from D2 cells (lane 2), or Dstet5 cells (lane 3), mature NCs from D2 (lane 4) or Dstet5 cells (lane 5), and virion-derived NCs (lane 6).

Fig. 3. Summary of the core peptides detected by MS analyses. Coverage of detected peptide ions is shown as underlined intervals underneath the core protein sequence (the highly basic CTD is in bold). Detected Lys-C and tryptic peptides are indicated by solid and dashed lines, respectively. Vertical bars on underlined intervals represent the termini of internal or overlapping peptides. Observed posttranslational modifications are as indicated: Ac, acetylation, s, phosphoserine, and t, phosphothreonine. Previously identified sites of phosphorylation are indicated by arrows; sites identified in this study are indicated by arrowheads.

Fig. 4. Detection of a core phosphorylation site at S230 by using VC MALDI FT-ICR MS. NCs of mixed maturity, purified, and digested in-gel by Lys-C were analyzed by using modified IonSpec FT-ICR MS, equipped with a home-built VC MALDI source (30, 41, 46, 47). (A) Mass spectrum (m/z 800–1,600) of core Lys-C peptides, showing the peptides 32–39 and 221–231, with and without phosphorylation, using the matrix a-cyano-4-hydroxycinnamic acid. The sequence from 221 to 231 is shown. (B) SORI-CAD tandem mass spectrum of the isolated ion (m/z 1,402.76). The detected fragment ion with phosphoric acid loss is labeled. (C) M5/M5/MS mass spectrum following a second SORI-CAD experiment using the resonance frequency of m/z 1,304.788 as the target for the calculated off-resonance frequency used. Detected fragment ions, as well as predicted side-chain losses, are labeled with standard b/z/y nomenclature. Modification by phosphate (HPO4) is designated by a P within a droplet. a, dehydroalanine.

Fig. 1. Summary of the core peptides detected by MS analyses. Coverage of detected peptide ions is shown as underlined intervals underneath the core protein sequence (the highly basic CTD is in bold). Detected Lys-C and tryptic peptides are indicated by solid and dashed lines, respectively. Vertical bars on underlined intervals represent the termini of internal or overlapping peptides. Observed posttranslational modifications are as indicated: Ac, acetylation, s, phosphoserine, and t, phosphothreonine. Previously identified sites of phosphorylation are indicated by arrows; sites identified in this study are indicated by arrowheads.
run as a series of phosphorylation-dependent bands by SDS/PAGE, whereas the virion-derived core runs as a single band (29). As an indirect assessment of any potential changes in core phosphorylation as a function of NC maturation, the purified mature NCs were compared with immature NCs and virion-derived NCs by SDS-PAGE followed by Western blotting (Fig. 2B) or silver staining (data not shown). Both analyses clearly showed that the core protein from the mature NCs resembled that of the virions, in that it migrated as a single band, in sharp contrast to that of immature NCs (and mixed NCs), which migrated as a heterogeneous series of bands, a reflection of their heterogeneous phosphorylation state.

Identification of Two Additional Sites of Core Phosphorylation, S230 and S232. To confirm and extend these initial indications of maturation-associated core dephosphorylation and to detect any additional core modifications that may be associated with maturation, we pursued extensive MS analyses of the DHBV core protein, starting with the mixed NCs. Through modification of the standard techniques for sample preparation we were able to obtain >96% sequence coverage of the core protein (Fig. 3) and identified additional posttranslational modifications.

A recently developed VC MALDI ionization source, which minimizes the loss of labile modification species (30, 41, 46, 47), was used in conjunction with an FT-ICR mass analyzer for the analysis of core Lys-C peptides. Ions coincident with expected proteolytic core peptides were detected throughout the region m/z 800–4,500 (data not shown). Moreover, ions consistent with core phosphopeptides were readily detected, particularly one at m/z 1,402.762, corresponding to a potential core phosphopeptide, representing core amino acids 221–231 (221TTVYGYRKRSSR231) (Fig. 4A). To localize the phosphosite, the phosphopeptide was sequenced by SORI-CAD fragmentation in the FT-ICR cell. Isolation of the peptide [M+H]+ ion at m/z 1,402.762, followed by a SORI-CAD experiment (tandem MS), resulted in the primary loss of phosphoric acid (H3PO4), with little other fragmentation detected (Fig. 4B). Thus, a SORI-CAD MS/MS experiment was necessitated to localize the dephosphorylated and dehydrated residue, which yielded fragmentation around the peptide C terminus, producing abundant b8, b9, and b10 fragment ions (Fig. 4C). The phosphorylation site was localized to S230 by the detection of the dehydroalanine (phospho-S230 that had undergone phosphoric acid loss) residue contained within the y4 and b10 fragment ions. Side-chain cleavage and other neutral losses typical of such high-energy CAD were also detected.

Because of their hydrophilicity, phosphopeptides can be expected to bind only weakly to reversed-phase chromatography medium, and basic phosphopeptides, like the Lys-C peptides of the highly basic core protein, even more weakly. We exploited this fact to elute core phosphopeptides selectively from the C18 ZipTip with low ratios of organic solvent. Under these conditions, most of the other peptides were retained on the chromatography medium. Using this technique, followed by MALDI-TOF MS (40), we detected peptide ions corresponding to the presence of six different phosphorylation sites on Lys-C C-terminal peptides, including the four previously described sites on the 232–262 peptide, the S230 site described above, and an additional site between 232 and 262. Treatment of these peptides with shrimp alkaline phosphatase before MS analysis removed...
all six phosphates, with only completely dephosphorylated peptide ions then being detected (Fig. 7, which is published as supporting information on the PNAS web site).

Using the same hydrophilic core Lys-C peptides, we conducted nanospray Q-oTOF MS (45) and readily detected multiple charge states (3+ to 6+) of the 232–262 peptide that appeared to contain two, three, four, and five phosphate modifications (Fig. 8 and Table 1, which are published as supporting information on the PNAS web site). We specifically targeted for CAD the pentaphosphorylated peptide 5+ ion at m/z 773.35 because it could contain at least one previously unreported site of phosphorylation. Fragmentation results from the CAD experiment (Fig. 5 and Table 2, which is published as supporting information on the PNAS web site), while exhibiting a substantial amount of phosphoric acid loss, also produced a significant amount of peptide backbone fragmentation that generated ions retaining some or all of their labile phosphate moieties. These ions matched theoretical fragments from the core C-terminal 232–262 peptide and permitted the specific localization of all five of the phosphorylation sites, resulting in the detection of a phosphosite at S232 and the confirmation of the four previously described phosphosites at T239, S245, S257, and S259 (Fig. 5). Notably, most of the peptide backbone fragmentation (y- and b-series) occurred between phosphosites and adjacent downstream proline residues.

Complete Dephosphorylation of the Core Protein from Mature and Virion-Derived NCs at the Four Known and the Two Additional Phosphosites. To assess how the core phosphosites changed during the course of NC maturation, we compared the MALDI-TOF mass spectra of hydrophilic Lys-C peptides derived from the purified core protein species from the three extremes of NC maturation (Fig. 6 and Table 3, which is published as supporting information on the PNAS web site). Ions corresponding to unmodified core peptides were detectable in each sample of NCs of different maturity (for example, ions corresponding to peptides spanning amino acids 23–31, 32–38, 178–201, 209–217, 221–230, and 232–262). In striking contrast, ions corresponding to the singly phosphorylated 221–231 peptide, and the singly, doubly, triply, quadruply, and quintuply phosphorylated 232–262 peptide, were detectable only in the immature NC samples (Fig. 6A) and were absent in both the mature intracellular NCs (Fig. 6B) and the virion-derived NCs (Fig. 6C). This complete change in the core phosphorylation state occurred in direct contrast to another posttranslational core modification, N-terminal methionine acetylation, also identified by our MS analyses, which did not change over the course of NC maturation (Fig. 3 and data not shown).

Discussion

By developing techniques capable of separating the DHBV NCs from the three extremes of maturation (immature, mature, and virion-derived) and subjecting them to comprehensive MS analyses, we have been able to carry out a detailed in vitro biochemical analysis of the dynamics of NC maturation. Our results demonstrate that the core protein from the mature NCs, like that from virion-derived NCs, was dephosphorylated at all observed sites of phosphorylation, including four previously identified and two additional sites identified in this study. Importantly, the same results were obtained with the intracellular NCs isolated from cells that replicated a surface protein-minus variant that precluded virion formation, as well as from cells that replicated the WT DHBV (Figs. 2 and 6). Thus, we effectively eliminated the possibility that we might have copurified “intracellular virions” (i.e., virions in the secretion pathway en route to the cell surface) as a contaminant along with the mature NCs. These results clearly indicated that NC maturation was associated with core dephosphorylation, independent of, and most likely before envelopment. This maturation-associated core dephosphorylation indicates a biochemical change of the NCs that can be correlated to maturation.

A unique MS instrument, recently developed in-house by the Boston University Mass Spectrometry Resource, was designed to perform MALDI under collisional cooling conditions, thereby minimizing losses of labile analyte functional groups and facilitating the detection of singly and multiply modified species, including phosphopeptides (30, 41, 46, 47). Configuration of this VC MALDI ion source with an FT-ICR mass analyzer provides a very high degree of mass accuracy, high sensitivity (46), and the capacity to perform multiple sequential fragmentation experiments. This instrument first allowed us to detect additional core phosphosites, and then to successfully localize one of these sites to S230 by MS/MS/MS, by conducting sequential SORI-CAD fragmentation reactions.

The second additional core phosphosite was localized by sequencing a pentaphosphorylated peptide containing this site, using electrospray ionization Q-oTOF MS (45). With this in-
indicate that the maturing NCs undergo active dephosphorylation stages of viral pgRNA packaging and DNA synthesis. Our results activated protein kinases (49, 50).

quently because of the activity of the abundant S residues. These residue juxtapositions appeared to provide low-energy fragmentation channels along the peptide backbone (N-terminal to proline) that effectively competed with cleavage of the phosphodiester bond of the pentaphosphopeptide, which contains four prolines pre-

sitions appeared to provide low-energy fragmentation channels along the peptide backbone (N-terminal to proline) that effectively competed with cleavage of the phosphodiester bond of the pentaphosphopeptide, which contains four prolines pre-

This remarkable sequencing of a pentaphosphopeptide was facilitated by phosphorylated S residues, of the pentaphosphopeptide, which contains four prolines pre-

of the pentaphosphopeptide, which contains four prolines pre-

phosites, including the second phosphosite at S232, as well as some or all of their phosphate moieties, despite the usually labile nature of the phosphate modification. Preferential backbone fragmentation N-terminal to proline, particularly when proline is preceded by acidic residues, has been documented (48). Importantly, we observed a similar effect in the fragmentation of the pentaphosphopeptide, which contains four prolines preceded by phosphorylated S/T residues. These residue juxtapositions appeared to provide low-energy fragmentation channels along the peptide backbone (N-terminal to each proline) that effectively competed with cleavage of the phosphodiester bond in the phosphorylated residue adjacent to the prolines. The success of this complex phospholocalization suggests that this methodology may be routinely applicable to the sequencing of other important proline-rich phosphopeptides, which occur frequently because of the activity of the abundant S/T-P-directed kinases, such as the cyclin-dependent kinases or mitogen-activated protein kinases (49, 50).

Core phosphorylation is known to be important for the early stages of viral pgRNA packaging and DNA synthesis. Our results indicate that the maturing NCs undergo active dephosphorylation as they carry out reverse transcription. A tantalizing hypothesis suggested by these results is that the hepadnaviral maturation signal, originating with the production of mature dsDNA within the NC, is transmitted through the NC shell in the form of the complete dephosphorylation of the core CTD, and is ultimately relayed from the dephosphorylated NC to the viral surface proteins. Also consistent with this notion are the reports that changes in the core phosphorylation state can affect its nucleic acid binding properties (51, 52), which may act to sense and transmit information on the nature of the nucleic acid within the maturing NC. However, evidence is still lacking in support of a direct role for mature NC dephosphorylation in triggering envelopment. In light of the discovery of two additional core CTD phosphosites and the complete core dephosphorylation associated with NC maturation shown here, it will be critical now to carry out further studies of all core phosphosites to carefully ascertain the role of core dephosphorylation in DHBV NC envelopment. The availability of large amounts of staged NCs should now facilitate time-resolved molecular imaging studies (53) of hepadnavirus NCs across the maturation process. These studies will be extremely valuable to visualize how the dynamics of core phosphorylation and conformational change are involved in the coupling of viral DNA synthesis to virion secretion, helping to make HBV one of the most infectious particles known.

We thank William Mason (Fox Chase Cancer Center, Philadelphia) for the anti-DHBV core antibodies and D2 cells. Christoph Seeger (Fox Chase Cancer Center) for the Dstet 5 cells, and William Mason and Jesse Summers for a critical reading of the manuscript. This work was supported by National Institutes of Health Public Health Service Grant R01 AI43453 (to J.H.) and National Institutes of Health Grants P41 RR10888 and S10 RR15942 (to C.E.C.).


