Reciprocity between O-GlcNAc and O-Phosphate on the Carboxyl Terminal Domain of RNA Polymerase II†

Frank I. Comer‡,§ and Gerald W. Hart*,‡

Department of Biological Chemistry, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205-2185, and Graduate Program, Department of Biochemistry and Molecular Genetics, The University of Alabama at Birmingham, Birmingham, Alabama 35294

Received December 1, 2000; Revised Manuscript Received May 4, 2001

ABSTRACT: The carboxyl terminal domain of RNA polymerase II has multiple essential roles in transcription initiation, promoter clearance, transcript elongation, and the recruitment of the RNA processing machinery. Specific phosporylation events are associated with the spatial and temporal coordination of these different activities. The CTD is also modified by β-O-linked GlcNAc on a subset of RNA Pol II molecules. Using synthetic CTD substrates, we show here that O-GlcNAc and phosphate modification of the CTD are mutually exclusive at the level of the enzymes responsible for their addition. In addition, we show that O-GlcNAc transferase and CTD kinase have different CTD repeat requirements for enzymatic activity. The K_m values of the two enzymes for CTD substrates are in a similar range, indicating that neither enzyme has a distinct kinetic advantage. Thus, the in vivo regulation of O-GlcNAc and phosphate modification of the CTD may involve the differential association of these two enzymes with the CTD at specific stages during the transcription cycle. Furthermore, direct competition between OGT and CTD kinase in vivo could generate multiple functionally distinct isoforms of RNA Pol II. Taken together, these results suggest that O-GlcNAc may give rise to additional functional states of RNA Pol II and may coordinate with phosphorylation to regulate class II gene transcription.

The production of mature eukaryotic messenger RNA is a complex process that requires the coordinated interactions of RNA Pol II with a multitude of protein factors (*1*−*4*). In addition to specific protein contacts, the process of gene transcription involves reversible phosphrorylation of the unique carboxyl terminal domain (CTD)[1] of the large subunit of RNA Pol II (*5*). A subset of RNA Pol II molecules contains another post-translational modification, O-linked *N*-acetylglucosamine (O-GlcNAc), on the CTD (*6*). O-GlcNAc and phosphate modifications of the RNA Pol II CTD are mutually exclusive, suggesting distinct roles in the transcription cycle.

An immense body of research over the last 15 years has firmly established multiple roles for the CTD of RNA Pol II in eukaryotic transcription and messenger RNA processing (*5*, *7*, *8*). The CTD is a unique domain that consists of multiple tandem repeats of the heptapeptide consensus sequence YSPTSPS. The CTD is conserved throughout the evolutionary ladder, although the number of repeats varies from 26 copies in yeast to 52 copies in mammals. Phosphorylation of the CTD is associated with numerous molecular events, including promoter clearance (*9*), passage through promoter proximal pause sites (*10*, *11*), stabilization of elongation complexes (*12*), and the recruitment of mRNA processing machinery (*7*, *8*, *13*). There are a number of kinases that can phosphorylate the CTD in vitro, and several CTD kinases have demonstrated roles in the transcription cycle. The basal transcription factor TF IIH contains a Cdk7−cyclin H pair that phosphorylates the CTD (*14*−*16*) during the transition from transcription initiation to elongation. This phosphorylation event may enhance promoter clearance by disrupting interactions of the CTD with components of the transcription preinitiation complex, although the exact mechanism of promoter clearance is not yet clear. There is also evidence that CTD phosphorylation is required for the passage of RNA Pol II through 5′ pause sites present in a number of different genes (*10*, *11*, *17*). The HIV-1 protein Tat appears to enhance RNA Pol II processivity by directly stimulating the TF IIH CTD kinase (*18*). In addition to the TF IIH kinase, other CTD kinases influence different aspects of transcription. The positive transcription elongation factor pTEFb is a Cdk9−cyclin T pair that phosphorylates the CTD after initiation and appears to stabilize productive transcription elongation complexes (*19*, *20*). The Cdk8−cyclin C pair can negatively regulate transcription either by phosphorylating and downregulating the TF IIH cyclin H subunit or by directly phosphorylating the CTD (*21*). The cyclin-dependent kinase Cdc2 can also inhibit RNA Pol II transcription by phosphorylating the CTD (*22*). In addition to its effects on transcription, specific

† This work was supported by NIH Grant HD 13563 to G.W.H. and a National Science Foundation Predoctoral Fellowship to F.I.C.

* To whom correspondence should be addressed. E-mail: gwhart@bs.jhmi.edu. Phone: (410) 955-5993. Fax: (410) 614-8804.

‡ The Johns Hopkins University School of Medicine.

§ The University of Alabama at Birmingham.

[1] Abbreviations: UDP-GlcNAc, uridine diphospho-*N*-acetylglucosamine; ATP, adenosine triphosphate; TBS, Tris-buffered saline; CTD, carboxyl terminal domain of RNA polymerase II; CAK, Cdk activating kinase; O-GlcNAc, β-O-Ser/Thr-linked *N*-acetylglucosamine; OGT, O-GlcNAc transferase; TLC, thin-layer chromatography; TPR, tetratricopeptide repeat; CD, circular dichroism.

10.1021/bi0027480 CCC: $20.00
Published on Web 06/09/2001

phosphorylation of the CTD also mediates the association of RNA Pol II with multiple mRNA splicing factors, mRNA capping enzymes, and polyadenylation factors (*8*).

While CTD phosphorylation is associated with the transition from transcription initiation to elongation and with the recruitment of RNA processing machinery, the nonphosphorylated form of RNA Pol II is involved in the formation of transcription competent initiation complexes. Prior phosphorylation of the CTD prevents RNA Pol II from entering into the preinitiation complex (*23*). The nonphosphorylated form of RNA Pol II is modified with O-linked GlcNAc (O-GlcNAc) along the length of the CTD (*6*). O-GlcNAc is a unique form of protein glycosylation that occurs on resident nuclear and cytoplasmic proteins (*24*) of all higher eukaryotes that have been examined to date (*25*−*27*). The O-GlcNAc modification consists of the monosaccharide *N*-acetylglucosamine in β-O-glycosidic linkage to the side chain hydroxyl groups of serine and threonine (*28*). Unlike classical glycosylation, O-GlcNAc is a dynamic modification that turns over rapidly compared to the peptide backbone (*29*, *30*). The modification occurs on a wide variety of proteins involved in many different cellular systems, including chromatin proteins (*31*), protein translation regulatory factors (*32*), RNA processing proteins,[2] and many RNA Pol II transcription factors (*33*). A growing body of evidence points to a regulatory role for O-GlcNAc on many of these proteins (*34*). Indeed, knockout studies have shown that the gene for the O-GlcNAc transferase (OGT) is essential for life at the single-cell level (*35*). O-GlcNAc shares many features in common with phosphorylation, and in some cases, the two modifications are mutually exclusive at a given site (*6*, *36*, *37*). Furthermore, all O-GlcNAc proteins that have been identified to date are also phosphoproteins. The coordinated action of both of these post-translational modifications greatly increases the potential number of functional states of a given protein. In the case of RNA Pol II, which contains multiple phosphate and O-GlcNAc sites, these two modifications could generate an enormous number of functionally distinct isoforms of the protein.

As a first step in examining the interplay between O-GlcNAc modification and phosphorylation of the RNA polymerase CTD, we studied the two modifications at the level of the enzymes responsible for their addition. We found that while the two enzymes have similar kinetics, they show distinct CTD substrate requirements and different modes of substrate binding. Furthermore, we report that O-GlcNAc and phosphate modification of the CTD sequence is reciprocal at the level of OGT and the TF IIH CTD kinase. These results imply that O-GlcNAc and phosphate are reciprocally related on the CTD and that they give rise to different functional variants of RNA Pol II.

MATERIALS AND METHODS

Materials. High Five and Sf9 insect cells were purchased from Invitrogen (San Diego, CA). Grace's insect medium was purchased from GIBCO-BRL (Rockville, MD). UDP-[^{3}H]GlcNAc was purchased from American Radiolabeled Chemicals (St. Louis, MO). UDP-GlcNAc was purchased from Amersham-Phamacia Biotech (Piscataway, NJ). [^{32}P]-ATP was purchased from New England Nuclear (Boston, MA). All peptide synthesis reagents were purchased from NovaBiochem (San Diego, CA). Organic solvents were from J. T. Baker (purchased through Fisher Scientific, Pittsburgh, PA). All other chemicals were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). Fmoc-protected serine-O-GlcNAc and threonine-O-GlcNAc were synthesized in our lab from readily available reagents.

Synthesis of CTD Repeating Peptides. Synthetic peptides bearing 1, 2, 3, 5, or 10 tandem repeats of the RNA Pol II carboxyl terminal domain (CTD) repeat sequence YSPTSPS were synthesized in our lab by standard Fmoc chemistry on a Millipore 9050 automated peptide synthesizer (PE Biosystems, Framingham, MA). All peptides were purified by semipreparative chromatography on a C_{18} reversed phase column using linear gradients of aqueous acetonitrile in 0.1% trifluoroacetic acid. Each peptide was purified by a slightly different gradient according to analytical determination of its chromatographic properties. Synthetic peptides bearing O-linked GlcNAc were prepared by incorporating Fmoc-protected serine-O-GlcNAc or threonine-O-GlcNAc at the desired position as previously described (*38*).

Expression and Purification of OGT. Recombinant OGT was expressed in High Five insect cells cultured in Grace's insect medium as previously described (*39*). Typically, 500 mL of midlog ($\sim$5−7 $\times$ 10^5 cells/mL) cells was infected for 48 h. Cells were harvested by centrifugation and stored at −20 °C until thawing for purification. Recombinant OGT was purified as described previously (*39*) with minor modifications. The linear gradient was replaced with a step elution of 5, 20, and 100 mM imidazole in 20 mM Tris (pH 7.9) and 0.5 M NaCl. OGT eluted in the 100 mM imidazole fraction. The purified enzyme was stored in aliquots at −80 °C. After an aliquot had been thawed, the enzyme was stored at −20 °C. The recombinant OGT Δ5.5 TPR deletion mutant was purified in the same manner.

OGT Assays. The purified enzyme was diluted 1:25 in 20 mM Tris (pH 7.9), 20% glycerol, and 0.02% NaN_3, and the reactions were started with the addition of 25 μL of the diluted enzyme. Depending on the assay, UDP-[^{3}H]GlcNAc was used at different specific activities as indicated in the figure legends. Assays were typically carried out in a final volume of 50 μL in 50 mM sodium cacodylate (pH 6.5) and stopped by the addition of 150 μL of 50 mM formic acid containing 1.0 M NaCl. The reaction mixture was loaded in 1 mL of 50 mM formic acid onto a C_{18} SepPak cartridge (Waters Corp., Milford, MA) equilibrated in 50 mM formic acid. The cartridge was washed successively with 10 mL of 50 mM formic acid, 50 mM formic acid with 1.0 M NaCl, and finally dH_2O. The labeled peptides were eluted directly into 20 mL scintillation vials with 4 mL of 50% methanol, and the amount of incorporated [^{3}H]GlcNAc was determined in a scintillation counter.

Expression and Purification of CTD Kinase. Recombinant baculovirus containing the CDK activating kinase (CAK) and CTD kinase components of the general transcription factor TF IIH (cdk 7, cyclin H, and p36) were a kind gift of A. Levine. The cdk7 and p36 subunits each contain a hemagglutinin epitope tag (HA-cdk 7 and HA-p36, respectively), while the cyclin H subunit contains a His_6 tag (His-cyc H). For a typical enzyme preparation, Sf9 insect cells growing in 500 mL of suspension culture at a density of $\sim$1 $\times$ 10^6

[2] N. Zachara, D. M. Snow, and G. W. Hart, manuscript in preparation.

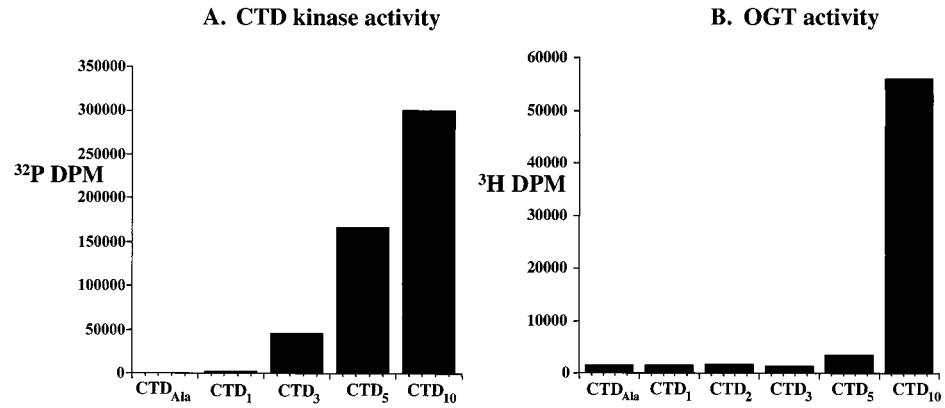

FIGURE 1: CTD kinase and OGT have different CTD repeat requirements. A series of synthetic CTD_n repeat peptides (where n is the number of repeats) were tested for their ability to serve as substrates for CTD kinase (A) and OGT (B). CTD kinase assays were performed at a peptide concentration of 300 μM and contained 1 μCi of [γ-^{32}P]ATP cold diluted to 200 μM. OGT assays contained 1 mM peptide and 1 μCi of UDP-[^{3}H]GlcNAc cold diluted to 0.5 mM.

cells/mL were triply infected with HA-cdk 7, HA-p36, and His-cyc H recombinant baculovirus. The infection was allowed to proceed for 48 h, and the cells were harvested by centrifugation in aliquots containing $\sim 1 \times 10^8$ cells (100 mL culture). Cell pellets were stored at -20 °C until purification. Cell pellets were thawed on ice in lysis buffer [20 mM Tris (pH 7.9), 250 mM NaCl, and 0.5% Nonidet P-40] and then sonicated with three 30 s bursts of a probe sonicator. The lysate was centrifuged at 35000*g* for 30 min. The supernatant fraction was filtered through a 0.2 μm filter and applied batchwise to 0.5 mL of Chelating Sepharose Fast Flow (Amersham-Pharmacia Biotech) charged with $NiSO_4$. The resin was washed with 10 mL of lysis buffer followed by 10 mL of Tris-buffered saline (TBS) [20 mM Tris (pH 7.9) and 150 mM NaCl] containing 5 mM imidazole. The enzyme complex was eluted with 10 mL of TBS containing 100 mM imidazole. The preparation was desalted and concentrated in a centrifugal ultrafiltration device with a molecular weight cutoff of 5000 against three changes of BC-100 buffer [20 mM Tris (pH 7.9), 0.1 mM EDTA, 100 mM KCl, 4 mM $MgCl_2$, and 10% glycerol] containing 1 mM DTT. The preparation was concentrated to $\sim$500 μL, then brought to a final concentration of 40% glycerol, and stored in aliquots at -80 °C.

CTD Kinase Assays. CTD kinase assays were carried out as previously described (*40*). Briefly, assays were carried out for 30 min in a 20 μL reaction volume containing 50 mM Tris (pH 7.9) and 8 mM $MgCl_2$. The amount of [γ-^{32}P]-ATP and the degree of cold dilution varied depending on the experiment, as indicated in the figure legends. Assays were stopped by the addition of 10 μL of 100 mM ATP and 140 μL of 50 mM formic acid containing 1.0 M NaCl. Five microliters was loaded onto a C_{18} reversed phase RPSF TLC plate (Analtech, Newark, DE) prespotted with 10 μL of 100 mM ATP. The plates were developed with an acetonitrile/phosphoric acid/triethylamine mixture (3%:0.1%:0.28% v:v:v in H_2O). The plates were dried, and the products, which remain at the origin, were visualized by autoradiography. The products were scraped from the plate and counted in a scintillation counter.

Alternatively, the products were analyzed on C_{18} SepPak cartridges (Waters Corp.) as follows. The SepPak cartridge was wet with 10 mL of methanol and then equilibrated in 10 mL of 50 mM formic acid. Samples were loaded in 1 mL of 50 mM formic acid containing 1.0 M NaCl. The cartridges were washed successively with 10 mL each of 50 mM formic acid, 50 mM formic acid containing 1.0 M NaCl, and then finally H_2O. The labeled peptides were eluted directly into 20 mL scintillation vials with 4 mL of 50% aqueous methanol and counted in a scintillation counter.

Determination of Michaelis−Menten Constants (K_m). The K_m values for both OGT and CTD kinase were determined using the CTD_{10} peptide over a concentration range of 0.1−5 mM. The K_m of OGT was determined at a UDP-GlcNAc concentration of 0.5 mM at a specific activity of 0.1 Ci/mmol of UDP-[^{3}H]GlcNAc. The CTD kinase K_m was determined at an ATP concentration of 1 mM at a specific activity of 0.1 Ci/mmol of [γ-^{32}P]ATP. The measurements were repeated three times in duplicate for OGT and three times in triplicate for CTD kinase. The K_m values were determined as the negative slope in a Woolf−Augustinsson−Hofstee plot.

CD Spectroscopy Measurements. CTD_n peptides were dissolved in 10 mM Na_2HPO_4 (pH 7.78) at a concentration of 67.75 μM. The peptide concentrations were determined on a fraction of the samples, using the extinction coefficient of tyrosine at 293 nm in 0.1 M NaOH ($\epsilon_{Tyr}^{293nm} = 2.390$ mM^{-1} per Tyr residue). Measurements were taken in triplicate at 15 °C on an Aviv 60DS CD spectropolarimeter.

RESULTS

CTD Kinase and OGT Have Different CTD Substrate Requirements. As a first comparison of OGT and CTD kinase, we determined the minimal number of CTD repeats required for enzymatic activity. CTD kinase exhibits a linear increase in the level of phosphate incorporation as the number of CTD repeats increases (Figure 1A). To our surprise, OGT exhibitshowed very different activity toward the CTD_n series of peptides. OGT showed poor activity toward the CTD peptides until we increased the length to 10 repeats (Figure 1B). This length requirement seemed unusual, as OGT is capable of utilizing many short peptides very efficiently (*39*). To rule out the possibility of a K_m threshold effect, we normalized the peptide concentrations on the basis of the number of CTD repeats and saw the same result (data not shown). These results suggest that OGT may have a different mode of substrate selectivity than CTD kinase.

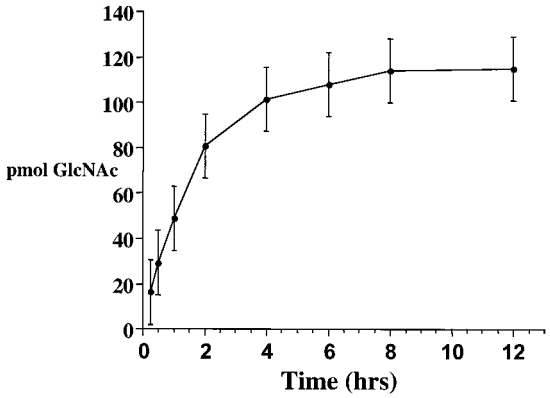

FIGURE 2: Time course of OGT activity toward the CTD_{10} peptide. Reaction mixtures contained 1 mM CTD_{10} peptide and 1 μCi of UDP-[^{3}H]GlcNAc at 0.5 mM. Activity measurements were compared against the CTD alanine-substituted negative control peptide. The experiment was repeated twice in duplicate.

OGT Exhibits Normal Kinetics toward the CTD_{10} Peptide. To further investigate the differences in substrate utilization of CTD kinase and OGT, we performed kinetic analyses of the activities of the two enzymes toward CTD substrates. We first wanted to see if OGT shows a nonstandard time course of incorporation of O-GlcNAc into the CTD_{10} peptide. We hypothesized that the substrate efficiency of the CTD_{10} peptide could be higher than those of the other CTD_n peptides as a result of product activation. If the addition of the first GlcNAc resulted in a more efficient addition of subsequent GlcNAc residues, one would expect a nonlinear increase in the rate of GlcNAc incorporation. Nevertheless, we observed a standard hyperbolic time course of incorporation of GlcNAc into the CTD_{10} peptide (Figure 2), arguing against product activation of the substrate. Likewise, the Michaelis−Menten plot shows a normal substrate saturation curve (Figure 3A), arguing against any form of product activation. A Woolf−Augustinsson−Hofstee reciprocal plot (Figure 3B) shows a K^m value of 313 μM, which is comparable to the known K_m values of OGT for other peptide substrates (*39*).

The Kinetics of CTD Kinase toward the CTD_{10} Peptide Are Similar to Those of OGT. We next determined the kinetics of CTD kinase toward the CTD_{10} peptide. Like OGT, CTD kinase shows a normal hyperbolic Michaelis−Menten substrate saturation curve (Figure 4A). The Woolf−Augustinsson−Hofstee reciprocal plot analysis reveals a K_m value of 898 μM (Figure 4B), which is in a range similar to that of OGT.

The CTD_n Peptide Backbone Exhibits No Major Structural Changes upon Increasing Repeat Length. We next wanted to test whether the differences in CTD substrate utilization between CTD kinase and OGT were due to the acquisition of a stable structural feature upon increasing the number of CTD repeats. Previous work in J. Corden's laboratory demonstrated that the CTD adopts a β-turn configuration in trifluoroethanol solutions (*41*). We wanted to test whether the CTD could adopt a stable structure in aqueous buffers and whether such structures were dependent upon a minimal number of CTD repeats. We thus performed CD spectroscopy experiments on the series of CTD_n peptides (where $n = 1$, 2, 3, 5, and 10). The CD measurements for all of the CTD peptides are consistent with random coil (*42*), rather than any obvious secondary structure (Figure 5). Thus, it is unlikely that the requirement for multiple CTD repeats for both OGT and CTD kinase is based upon the acquisition of a stable structural feature upon increasing the number of CTD repeats. Nevertheless, these results do not rule out the possibility that the catalytic mechanisms of these two enzymes require a proper orientation of the substrate that is only met with a minimum number of CTD repeats.

Glycosylation of the CTD_{10} Peptide Requires Multisite Interaction with OGT. We next tested whether the length requirement of OGT toward CTD peptides could be due to a multisite mode of substrate binding. We reasoned that if OGT requires more than one site of contact with the CTD substrate, then disrupting its subunit structure may impair the ability of the enzyme to utilize the CTD_{10} substrate. Likewise, one might expect shorter nonsubstrate CTD peptides to inhibit glycosylation of the CTD_{10} substrate by interfering with one or more of the sites of binding. Deletion analysis of OGT revealed that truncation of approximately half of the tetratricopeptide repeat motifs of the enzyme results in a loss of its trimeric architecture (*39*). Nevertheless, this deletion mutant, designated Δ5.5 OGT, is equally effective at catalyzing the transfer of O-GlcNAc to several different peptide substrates. We found that this truncated monomeric OGT is unable to glycosylate the CTD_{10} substrate efficiently compared to native trimeric OGT (Figure 6A). This result suggests that there is a requirement for multiple sites of contact with the CTD_{10} substrate, either across subunits or between the active site and one or more additional contact sites in the TPR domain. Since the TPR domain is a known protein interaction domain (*43*), it is possible that it may act as a selectivity filter for a subset of OGT substrates. Our results are reminiscent of previous findings which showed that OGT truncation compromised activity toward protein substrates more than toward a peptide substrate (*44*). As a further test of the multiple-site binding mode of OGT toward CTD substrates, we also performed competition experiments with several CTD peptides. The CTD_1 peptide is a not a substrate for OGT, yet it inhibits glycosylation of the CTD_{10} peptide (Figure 6B). The CTD_1 peptide also inhibits glycosylation of other peptide substrates (data not shown). In contrast, a CTD_1 peptide bearing one O-GlcNAc residue is ineffective at inhibiting glycosylation of the CTD_{10} peptide. Likewise, a CTD_1 peptide that contains alanine substitutions at each of the serine and threonine sites is also incapable of inhibiting OGT. We also performed the converse experiment with CTD kinase and found that the CTD_1 peptide does not inhibit phosphorylation of the CTD_{10} peptide (data not shown). Taken together, these results strongly suggest that OGT employs a unique multisite mode of binding to CTD substrates. This mechanism is distinct from binding to other peptide substrates, as the enzyme is fully capable of glycosylating many short peptides. Therefore, our results may be representative of a novel mechanism for the regulation of OGT activity toward a subset of substrates.

Phosphorylation of the CTD_{10} Peptide Inhibits the Activity of OGT. Previous work in our laboratory showed that O-GlcNAc and phosphate modifications of the CTD of RNA Pol II are mutually exclusive (*6*). One of the goals of this study was to determine whether this reciprocity occurs at the enzymatic level. To address this question directly, we prepared CTD peptides that were either phosphorylated or

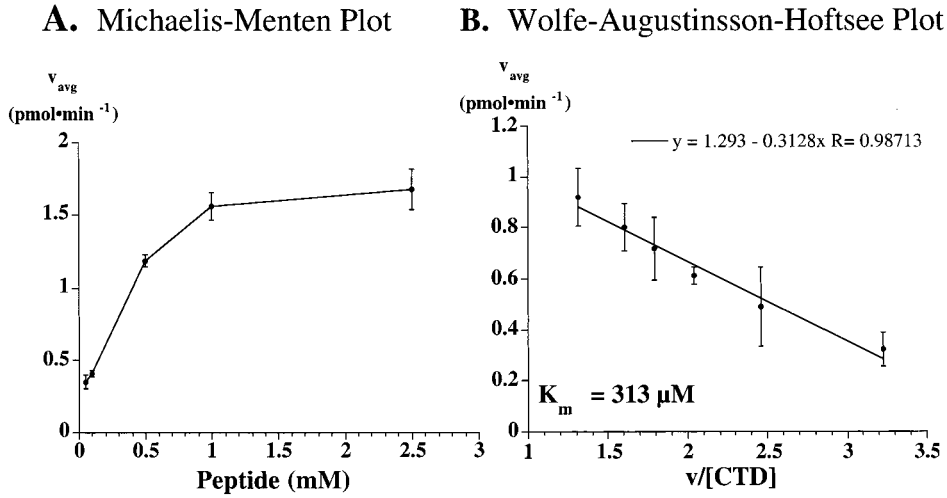

FIGURE 3: Kinetic analysis of OGT activity toward the CTD_{10} peptide. Assays contained 100, 200, 300, 400, 500, and 700 μM CTD_{10} peptide and 2.5 μCi of UDP-[^{3}H]GlcNAc at 0.5 mM. The level of incorporation was measured as ^{3}H counts that bound to a C_{18} SepPak cartridge. Data were assessed with a Michaelis−Menten plot (A) and a Woolf−Augustinsson−Hofstee analysis (B) to obtain K_m values. Data are representative of three experiments, each performed in triplicate.

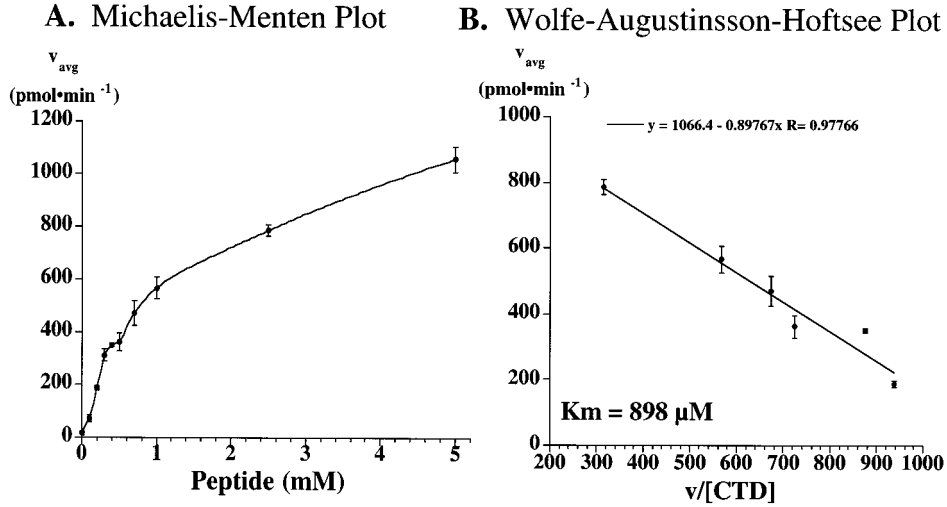

FIGURE 4: Kinetic analysis of CTD kinase activity toward the CTD_{10} peptide. Assays contained 100, 200, 300, 400, 500, 700, and 1000 μM CTD_{10} peptide and 5 μCi of [γ-^{32}P]ATP cold diluted to 1 mM. The level of incorporation was measured as ^{32}P counts that migrated with the peptide on a C_{18} reversed phase TLC plate. Data were assessed with a Michaelis−Menten plot (A) and a Woolf−Augustinsson−Hofstee analysis (B) to obtain K_m values. Data are representative of two experiments, each performed in triplicate.

O-GlcNAc-modified. We prepared a phosphorylated CTD_{10} peptide by in vitro phosphorylation with CTD kinase. Our results show that in vitro phosphorylation completely abolishes the activity of OGT toward the CTD_{10} peptide (Figure 7A). This inhibition is underscored by the fact that the in vitro phosphorylation reaction gives only a substoichiometric yield of three phosphates per CTD_{10} peptide (data not shown). Sequencing of the OGT-labeled CTD_{10} peptide indicates that all of the incorporated GlcNAc is present in the N-terminal end of the peptide (data not shown). The work of Cordes and Krohne suggested that OGT glycosylates the nuclear pore protein gp62 in a sequential manner (*45*). Our results suggest that OGT may glycosylate the full-length CTD in a similar manner, glycosylating N-terminal residues first before moving on to other residues in the CTD sequence. If the CTD kinase also exhibits sequential phosphorylation of the CTD substrate, it would explain why exhaustive phosphorylation is not required for inhibition of glycosylation.

Glycosylation of the CTD Sequence Inhibits the Activity of CTD Kinase. We prepared a synthetic CTD_5 peptide containing Thr-O-GlcNAc at position 4 of each CTD repeat. CTD kinase assays on the CTD_5-Thr-O-GlcNAc peptide demonstrate that the presence of O-GlcNAc effectively blocks the ability of the peptide to become phosphorylated (Figure 7B). Although the major sites of CTD phosphorylation in vivo are the serines at positions 2 and 5 (*5*), the major site of glycosylation is the threonine at position 4 (*6*). Therefore, this result is not merely due to the trivial capping of CTD phosphorylation sites. We also performed experiments to test if substoichiometric glycosylation of the CTD_{10} peptide inhibits its ability to serve as a substrate for the CTD kinase. Using OGT to glycosylate the CTD_{10} sequence, we were only able to achieve a stoichiometry of ~1 mol of GlcNAc per mole of CTD_{10} peptide, despite the presence of 40 potential sites of glycosylation. We found that the CTD kinase was still able to phosphorylate this peptide (data not shown). This result is not surprising, as we found that the

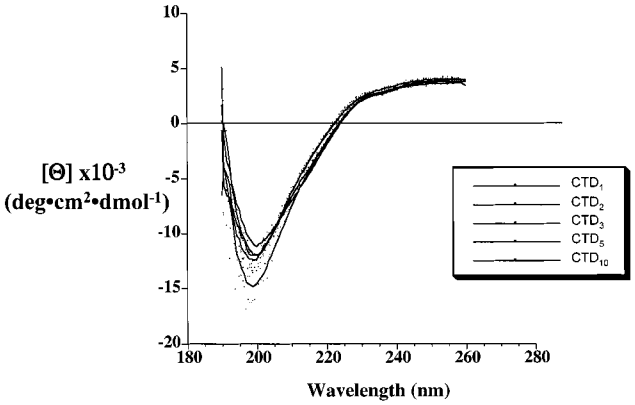

FIGURE 5: CTD peptide backbone showing no major structural change upon increasing the number of CTD repeats. CD spectroscopy measurements were taken on the CTD_n series of peptides in 10 mM Na_2HPO_4 buffer (pH 7.78). Peptide concentrations were 67.75 μM based on the absorbance of tyrosine at 293 nm. All measurements were taken at 15 °C in triplicate.

OGT only glycosylated the extreme N-terminus of the peptide (data not shown), leaving many repeats unmodified and available for the CTD kinase. This result merely suggests that the CTD kinase does not need an unglycosylated N-terminus to phosphorylate downstream sites. However, the presence of a single O-GlcNAc on each repeat of the CTD is sufficient to inhibit phosphorylation. Thus, the two modifications may not always compete for the exact same sites on the CTD, but the presence of O-GlcNAc in the region effectively prevents phosphorylation of adjacent sites, and vice versa.

DISCUSSION

The role of phosphorylation of the unique carboxyl terminal domain (CTD) of RNA Pol II (*5*) has been the subject of a great deal of study over the last 15 years. Phosphorylation of the CTD is associated with the exit of RNA Pol II from the preinitiation complex and the maintenance of transcription elongation complexes. Recent studies clearly demonstrate that CTD phosphorylation is also involved in the recruitment of the RNA processing machinery to nascent transcripts (*5*, *7*, *8*). Kelly et al. (*6*) demonstrated that O-GlcNAc and O-phosphate modifications of the CTD are present on different pools of RNA Pol II. Using a defined in vitro system, we show here that this reciprocity occurs at the level of the enzymes responsible for their addition, the TF IIH CTD kinase and O-GlcNAc transferase. These results suggest that the two modifications may be involved in a concerted mechanism for the regulation of RNA Pol II. A number of studies have suggested that there is remodeling of the CTD phosphorylation pattern during the transcription cycle (*5*). In support of this hypothesis, protein kinase inhibitors diminish the ability of RNA Pol II to produce long transcripts. Furthermore, the transcription elongation factor pTEFb is a CTD kinase (*5*). Nevertheless, it is not clear precisely how the phosphorylation pattern changes during the transcription cycle. The study presented here suffers from the same limitation as many other studies of the CTD. Namely, it does not address the way in which specific patterns of post-translational modification along the entire length of the CTD could affect either the activity of RNA Pol II or its association with transcription factors. Our model for the role of O-GlcNAc on the CTD (Figure 8) and the accepted model for the role of phosphorylation of this domain depict uniform modification along the length of the CTD. It is equally possible that there are distinct subdomains on the CTD that differ in their pattern of post-translational modification. The repetitive nature of the CTD and the vast number of hypothetical CTD isoforms that post-translational modifications might generate make it difficult to address this possibility in a systematic manner.

We found that the CTD kinase and the O-GlcNAc transferase display similar K_m values, suggesting that neither enzyme has a strong kinetic advantage. In contrast, the enzymes have distinct CTD repeat requirements and have different modes of substrate binding. These differences provide potential mechanisms for the differential regulation of the activities of each enzyme. The unusual length-dependent selectivity of OGT for CTD substrates prompted further investigation. This length requirement is different from those of other OGT substrates (*39*), leading us to

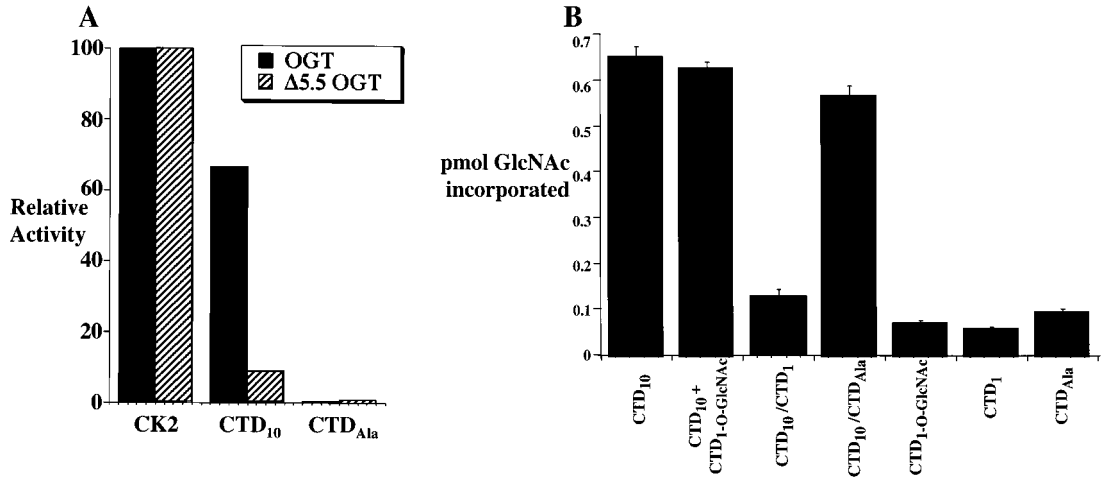

FIGURE 6: Glycosylation of the CTD_{10} peptide requires multisite interactions with OGT. (A) The activity of full-length trimeric OGT was compared to that of the monomeric Δ5.5 TPR OGT. The activity toward the CTD_{10} peptide was normalized relative to the activity toward the CK2 peptide substrate for each enzyme. (B) Peptide mixing experiment. The activity of OGT toward the CTD_{10} peptide was assayed at a CTD_{10} concentration of 1 mM in the presence or absence of a 5-fold molar excess of CTD_1 single-repeat peptides. CTD_1 is YSPTSPS, $CTD_{1-O-GlcNAc}$ YSPT(O-GlcNAc)SPS, and CTD_{Ala} YAPAAPAK.

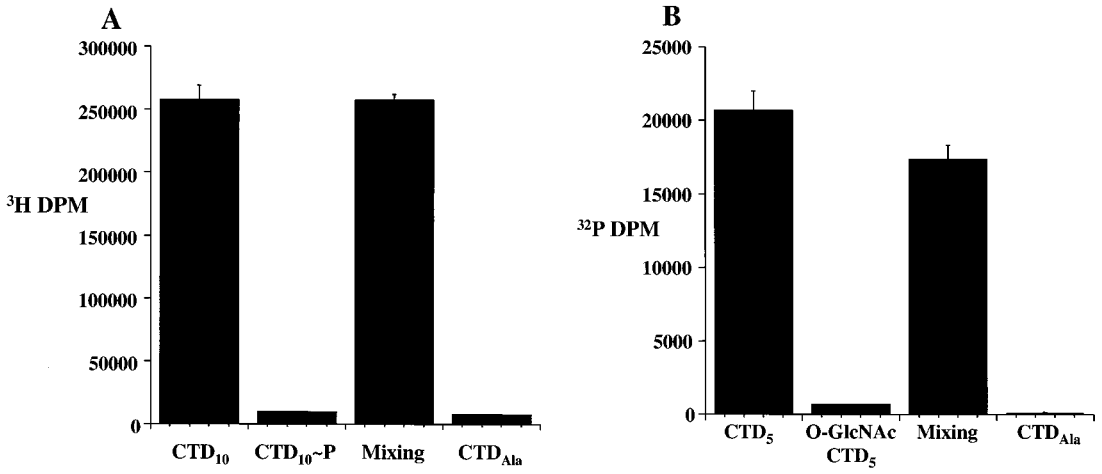

FIGURE 7: Phosphorylation and glycosylation of the CTD sequence are mutually exclusive. (A) Activity of OGT toward the CTD peptide following prior phosphorylation. The CTD_{10} peptide was phosphorylated in vitro with CTD kinase to a stoichiometry of three to four phosphates per molecule. The assays were carried out at a peptide concentration of 500 μM and with 0.5 μCi of UDP-[^{3}H]GlcNAc. For the mixing control, the reaction mixtures contained equal amounts of each peptide (500 μM). The data are representative of three experiments, each performed in triplicate. (B) Activity of CTD kinase toward the O-GlcNAc-modified CTD peptide. A CTD_5 peptide containing O-GlcNAc on all five Thr residues was prepared synthetically. The assays were carried out at peptide concentrations of 100 μM and with 0.1 μCi of [γ-^{32}P]ATP cold diluted to 100 μM. The mixing control contained equal peptide concentrations (100 μM).

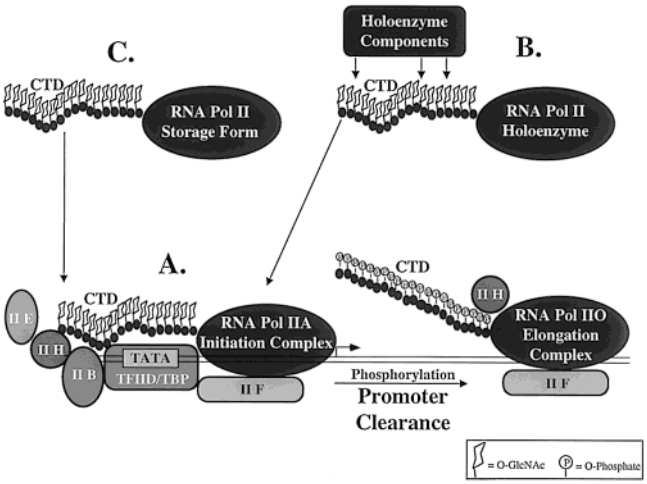

FIGURE 8: Model for the potential roles for O-GlcNAc on the CTD of RNA Pol II. This model outlines three major possibilities for the function of O-GlcNAc on the CTD of RNA Pol II. (A) O-GlcNAc may regulate assembly of RNA Pol II transcription initiation complexes via specific contacts with other components of the transcription machinery. In this model, the O-GlcNAc-modified form of RNA Pol II enters the preinitiation complex. After the initiation complex is complete, an O-GlcNAc specific β-*N*-acetylglucosaminidase would selectively remove the O-GlcNAc residues, thus allowing for CTD phosphorylation and entry into the elongation phase of transcription. (B) O-GlcNAc could mediate interaction of RNA Pol II with components of the RNA Pol II holoenzyme complex. The O-GlcNAc-mediated association of RNA Pol II with holoenzyme components could direct RNA Pol II to active promoters for rapid activation of transcription. (C) O-GlcNAc may protect RNA Pol II from degradation, thus allowing for the accumulation of a readily available storage pool of RNA Pol II. In this model, O-GlcNAc could be removed from the CTD prior to entry into the initiation complex or may remain on the CTD until after formation of the initiation complex. The three possibilities are not mutually exclusive.

hypothesize that OGT may have a unique mode of binding to CTD substrates. We reasoned that OGT may recognize a stable structural feature acquired upon increasing CTD length, but CD spectroscopy measurements argue against such a mechanism (Figure 5). The time course of OGT activity toward the CTD_{10} peptide shows that the reaction follows a simple hyperbolic curve (Figure 2). This observation rules out the possibility of some sort of product activation, as one would expect a nonlinear time course if the initial addition of O-GlcNAc facilitates subsequent glycosylation events. Finally, we investigated whether utilization of CTD substrates requires multisite interactions with OGT. Peptide mixing experiments and OGT truncation mutant studies suggest that OGT requires multiple sites of interaction to utilize CTD substrates (Figure 6). Partial truncation of the tetratricopeptide repeat (TPR) region of OGT abolishes activity toward the CTD_{10} peptide. Since this protein interaction domain (*43*) is also responsible for trimerization of the enzyme (*39*), it is not clear whether this result is due to the loss of the subunit architechture of OGT or the deletion of a CTD binding site in the TPR domain. Further study is necessary to determine whether there is a direct interaction between the CTD and the TPR domain of OGT. If there is a direct interaction, one can envision a mechanism of regulation in which a specific TPR binding protein could specifically block the site of interaction between the CTD and the TPR domain, thus downregulating the activity of OGT toward the CTD.

Since phosphorylation of the CTD is associated with promoter clearance, transcript elongation, and RNA processing, it is conceivable that O-GlcNAc may be involved in early events in the transcription cycle. Figure 8 outlines some of the possible roles of O-GlcNAc on the CTD of RNA Pol II. O-GlcNAc could direct specific protein contacts that mediate the recruitment of RNA Pol II to active promoters (Figure 8A). These associations could also aid in the proper positioning of the enzyme at the start site of transcription. Structural studies from Simanek and co-workers showed that O-GlcNAc modification of the CTD repeat induces a turnlike structure (*46*). It follows that, in addition to preventing CTD phosphorylation, the presence of multiple O-GlcNAc residues along the length of the CTD could dramatically alter the structure of this domain. Such changes in conformation could affect the activity of RNA Pol II, either directly or by

modulating interactions with other components of the transcription machinery. Alternatively, specific factors may interact directly with the O-GlcNAc residues on the CTD. These O-GlcNAc-mediated interactions could regulate the formation of transcription initiation complexes. Another possibility is that O-GlcNAc modification of the CTD mediates the interaction between RNA Pol II and components of the RNA Pol II holoenzyme complex (Figure 8B). Finally, O-GlcNAc modification of the CTD could convert RNA Pol II into a stable storage form that is transcriptionally inactive, but poised for activation through the action of an O-GlcNAc specific β-*N*-acetylglucosaminidase (Figure 8C). The O-GlcNAc modification appears to protect the transcription factor Sp1 from proteasome degradation (*47*). A similar protective mechanism could allow for the accumulation of a readily mobilizable pool of RNA Pol II. In any case, the reciprocity between CTD phosphorylation and glycosylation suggests that the two modifications play distinct roles in the regulation of RNA Pol II. The complexity of coordinating the proper expression of the many eukaryotic genes transcribed by RNA Pol II is likely to require a concerted effort between both phosphorylation and glycosylation of RNA Pol II.

ACKNOWLEDGMENT

We thank Dr. Lisa Kreppel for assistance with the OGT assays and for the preparation of the OGT constructs. We thank Dr. Arnold Levine for the kind gift of the recombinant TF IIH CTD kinase subunits. We also thank Drs. James Wrabl and David Shortle for helpful discussion and for training with and use of the CD spectropolarimeter.

REFERENCES

1. Myer, V. E., and Young, R. A. (1998) *J. Biol. Chem. 273*, 27757−27760.
2. Uptain, S. M., Kane, C. M., and Chamberlin, M. J. (1997) *Annu. Rev. Biochem. 66*, 117−172.
3. Ellwood, K., Chi, T., Huang, W., Mitsouras, K., and Carey, M. (1998) *Cold Spring Harbor Symp. Quant. Biol. 63*, 253−261.
4. Reinberg, D., Orphanides, G., Ebright, R., Akoulitchev, S., Carcamo, J., Cho, H., Cortes, P., Drapkin, R., Flores, O., Ha, I., Inostroza, J. A., Kim, S., Kim, T. K., Kumar, P., Lagrange, T., LeRoy, G., Lu, H., Ma, D. M., Maldonado, E., Merino, A., Mermelstein, F., Olave, I., Sheldon, M., Shiekhattar, R., Zawel, L., et al. (1998) *Cold Spring Harbor Symp. Quant. Biol. 63*, 83−103.
5. Dahmus, M. E. (1996) *J. Biol. Chem. 271*, 19009−19012.
6. Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993) *J. Biol. Chem. 268*, 10416−10424.
7. Corden, J. L., and Patturajan, M. (1997) *Trends Biochem. Sci. 22*, 413−416.
8. Hirose, Y., and Manley, J. L. (2000) *Genes Dev. 14*, 1415−1429.
9. Goodrich, J. A., and Tjian, R. (1994) *Cell 77*, 145−156.
10. O'Brien, T., Hardin, S., Greenleaf, A., and Lis, J. T. (1994) *Nature 370*, 75−77.
11. Kumar, K. P., Akoulitchev, S., and Reinberg, D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9767−9772.
12. Dvir, A., Conaway, R. C., and Conaway, J. W. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9006−9010.
13. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997) *Nature 385*, 357−361.
14. Lu, H., Zawel, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992) *Nature 358*, 641−645.
15. Coin, F., and Egly, J. M. (1998) *Cold Spring Harbor Symp. Quant. Biol. 63*, 105−110.
16. Roy, R., Adamczewski, J. P., Seroz, T., Vermeulen, W., Tassan, J. P., Schaeffer, L., Nigg, E. A., Hoeijmakers, J. H., and Egly, J. M. (1994) *Cell 79*, 1093−1101.
17. Weeks, J. R., Hardin, S. E., Shen, J., Lee, J. M., and Greenleaf, A. L. (1993) *Genes Dev. 7*, 2329−2344.
18. Parada, C. A., and Roeder, R. G. (1996) *Nature 384*, 375−378.
19. Peng, J., Liu, M., Marion, J., Zhu, Y., and Price, D. H. (1998) *Cold Spring Harbor Symp. Quant. Biol. 63*, 365−370.
20. Peng, J., Zhu, Y., Milton, J. T., and Price, D. H. (1998) *Genes Dev. 12*, 755−762.
21. Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000) *Nature 407*, 102−106.
22. Gebara, M. M., Sayre, M. H., and Corden, J. L. (1997) *J. Cell. Biochem. 64*, 390−402.
23. Laybourn, P. J., and Dahmus, M. E. (1990) *J. Biol. Chem. 265*, 13165−13173.
24. Holt, G. D., and Hart, G. W. (1986) *J. Biol. Chem. 261*, 8049−8057.
25. Hart, G. W. (1997) *Annu. Rev. Biochem. 66*, 315−335.
26. Comer, F. I., and Hart, G. W. (1999) *Biochim. Biophys. Acta 1473*, 161−171.
27. Comer, F. I., and Hart, G. W. (2000) *J. Biol. Chem. 275*, 29179−29182.
28. Torres, C.-R., and Hart, G. W. (1984) *J. Biol. Chem. 259*, 3308−3317.
29. Roquemore, E. P., Chevrier, M. R., Cotter, R. J., and Hart, G. W. (1996) *Biochemistry 35*, 3578−3586.
30. Chou, C.-F., Smith, A. J., and Omary, M. B. (1992) *J. Biol. Chem. 267*, 3901−3906.
31. Kelly, W. G., and Hart, G. W. (1989) *Cell 57*, 243−251.
32. Datta, B., Ray, M. K., Chakrabarti, D., Wylie, D. E., and Gupta, N. K. (1989) *J. Biol. Chem. 264*, 20620−20624.
33. Jackson, S. P., and Tjian, R. (1988) *Cell 55*, 125−133.
34. Boehmelt, G., Wakeham, A., Elia, A., Sasaki, T., Plyte, S., Potter, J., Yang, Y., Tsang, E., Ruland, J., Iscove, N. N., Dennis, J. W., and Mak, T. W. (2000) *EMBO J. 19*, 5092−5104.
35. Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 5735−5739.
36. Chou, T.-Y., Hart, G. W., and Dang, C. V. (1995) *J. Biol. Chem. 270*, 18961−18965.
37. Cheng, X., Cole, R. N., Zaia, J., and Hart, G. W. (2000) *Biochemistry 39*, 11609−11620.
38. Greis, K. D., Hayes, B. K., Comer, F. I., Kirk, M., Barnes, S., Lowary, T. L., and Hart, G. W. (1996) *Anal. Biochem. 234*, 38−49.
39. Kreppel, L. K., and Hart, G. W. (1999) *J. Biol. Chem. 274*, 32015−32022.
40. Cisek, L. J., and Corden, J. L. (1991) *Methods Enzymol. 200*, 301−325.
41. Cagas, P. M., and Corden, J. L. (1995) *Proteins 21*, 149−160.
42. Johnson, W. C., Jr. (1990) *Proteins 7*, 205−214.
43. Lamb, J. R., Tugendreich, S., and Hieter, P. (1995) *Trends Biochem. Sci. 20*, 257−259.
44. Lubas, W. A., and Hanover, J. A. (2000) *J. Biol. Chem. 275*, 10983−10988.
45. Cordes, V. C., and Krohne, G. (1993) *Eur. J. Cell Biol. 60*, 185−195.
46. Simanek, E. E., Huang, D. H., Pasternack, L., Machajewski, T. D., Seitz, O., Millar, D. S., Dyson, H. J., and Wong, C. H. (1998) *J. Am. Chem. Soc. 120*, 11567−11575.
47. Han, I., and Kudlow, J. E. (1997) *Mol. Cell. Biol. 17*, 2550−2558.

BI0027480