Analysis of Nuclear Pore Protein p62 Glycosylation

W. A. Lubas, M. Smith, C. M. Starr, and J. A. Hanover*

Laboratory of Biochemistry and Metabolism, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

Received September 12, 1994[⊗]

ABSTRACT: Glycoprotein components of the nuclear pore are essential for nuclear transport and are modified by both glycosylation and phosphorylation. The function and control of these post-translational modifications are poorly understood. Glycosylation of the major rat nuclear pore glycoprotein, p62, was examined *in vitro* using recombinant p62 as a substrate. Rat p62 was expressed in *Escherichia coli* and purified to near homogeneity. Kinetic analysis using a partially purified mammalian transferase suggests that the recombinant protein is an excellent substrate (K_m = 0.30 μM) for the transfer of GlcNAc from UDP-GlcNAc (K_m = 1.8 μM). Localization of the sites of *O*-linked GlcNAc glycosylation of rat p62 was performed by a combination of deletion analysis of *in vitro* translation products and by immunoprecipitation of [^{14}C]GlcNAc-labeled proteolytic fragments. The amino terminus of rat p62 is poorly glycosylated with no *O*-linked GlcNAc sites between Lys^{22} and Lys^{97}; the carboxyl terminus has one known glycosylation site at Ser^{471}. The majority of the glycosylation sites in rat p62 are likely to occur on the six clustered Ser residues in the central Ser/Thr-rich region from Ser^{270} to Thr^{294}. A synthetic peptide derived from this region is a good substrate for *O*-GlcNAc addition (K_m = 30 μM) and a potent competitive inhibitor of p62 glycosylation (K_i = 15 μM). It is proposed that this Ser/Thr-rich domain functions as a linker region between the amino-terminal β-pleated sheet and the carboxyl terminal α-helical domains. *O*-Glycosylation and phosphorylation of this linker region could provide a dynamic means of altering the conformation of p62 during nuclear pore assembly and disassembly.

The nuclear pore complex provides an aqueous channel for active transport of proteins and RNA between the nucleus and cytoplasm. It is a supramolecular structure with an estimated molecular mass of 124 MDa and is composed of an estimated 100−200 different proteins (Miller et al., 1991). A subset of these proteins, ranging in molecular mass from 45 to 210 kDa, was initially identified by monoclonal antiserum raised against nuclear pore complex−lamina extracts of intact nuclei (Davis & Blobel, 1986; Snow et al., 1987; Park et al., 1987). These proteins were shown to contain a novel carbohydrate linkage, *O*-linked GlcNAc, which was part of the epitope recognized by the monoclonal antibody (Park et al., 1987; Holt et al., 1987b; Hanover et al., 1987).

The mammalian nuclear pore protein, p62, initially cloned in rat (D'Onofrio et al., 1988; Starr et al., 1990) and more recently in human (Carmo-Fonseca et al., 1991), mouse, and *Xenopus laevis* (Cordes et al, 1991) is a member of the family of *O*-linked GlcNAc containing nuclear pore proteins. A second member of this family, p180, has also been recently cloned from rat (Sukegawa & Blobel, 1993). Both of these proteins have been shown to be glycosylated with 10−12 *O*-GlcNAc residues per molecule (Holt et al., 1987b). We previously mapped one site of glycosylation in rat p62 to Ser^{471} in the carboxyl terminus (D'Onofrio et al., 1988). More recent reports have helped to localize glycosylation sites in human p62 to the amino terminal 301 amino acids (Carmo-Fonseca et al., 1991) and to the central part of the murine protein between amino acids 222 and 330 and between amino acids 153 and 435 in *X. laevis* (Cordes et al., 1991). Consistent with these known glycosylation sites, histidine-tagged fusion peptides derived from mouse p62 have been shown to be glycosylated by rabbit reticulocyte lysate (Cordes & Krohne, 1993). Yet only two of these peptides, 1−341 and 248−341, were capable of inhibiting glycosylation of the intact protein at the concentrations tested.

Utilizing two complementary approaches, we show that 60% of the glycosylation of p62 occurs in the central Ser/Thr-rich region from Ser^{270} to Thr^{294}, most likely on the six Ser residues in this region. We describe the expression and purification of unglycosylated p62 from *Escherichia coli* and show that the recombinant molecule serves as a substrate for *in vitro* glycosylation by a partially purified mammalian *O*-GlcNAc transferase.[1] Kinetic analysis of the glycosylation of recombinant p62 is consistent with a single class of acceptor sites (K_m = 0.30 μM). We quantitatively examined the glycosylation of a synthetic peptide derived from the Ser/Thr-rich region corresponding to Gly^{257} to Ser^{299} and show it to be both a good substrate (K_m = 30 μM) and a pure competitive inhibitor of p62 glycosylation by the mammalian O-GlcNAc transferase (K_i = 15 μM).

EXPERIMENTAL PROCEDURES

Purification of Rat p62 Expressed in E. coli BL21 (pLys-S). Cultures of BL21 (pLys-S) cells containing the pET-3a expression vector (Novagen) encoding p62 were grown at 37 °C and 200 rpm in LB media (Digene) supplemented with

* Address correspondence to Bldg 10, Rm 9B07, LCBB, NIDDK, National Institutes of Health, Bethesda, MD 20892; Telephone: (301) 496-0943; Fax: (301) 496-0839.

[⊗] Abstract published in *Advance ACS Abstracts,* January 1, 1995.

[1] Abbreviations: IPTG, isopropyl β-thiogalactopyranoside; PMSF, phenylmethanesulfonyl fluoride; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; *O*-GlcNAc transferase, uridine diphospho-*N*-acetylglucosamine:polypeptide β-*N*-acetylglucosaminyltransferase.

ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL). Purification of p62 from these cultures was performed by a modification of a method we have previously reported (Buss et al., 1994). After reaching an OD_{650} of 0.6−0.8, cultures were induced with IPTG (0.4 mM) for 90 min and harvested by centrifugation at 6000*g* for 15 min (Sorvall RC-5B). The cells were washed with phosphate-buffered saline, pH 7.2, and the pellet from 1 L of culture was resuspended in 12.5 mL of 20% sucrose, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 1 mM PMSF. The cells were centrifuged, and the pellet was resuspended in 10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1 mM PMSF, and 5 mg/mL lysozyme; stored on ice for 30 min; centrifuged; and finally frozen at −70 °C. After thawing, the pellet from 1 L of culture was sonicated on ice (6× 30 s, setting 4, Heat Systems Ultrasonic Inc.) in 20 mL of 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 1 mM PMSF (20 mL/L of culture); frozen; thawed; and centrifuged at 30000*g* for 20 min to isolate the inclusion bodies. The inclusion bodies were washed once with 0.2% deoxycholate, 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 1 mM PMSF, washed twice with the same buffer without deoxycholate, and then solubilized by sonication (3× 30 s) in 6 M urea, 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 2 mM dithiothreitol. A partially purified p62 preparation (30−50% of the total protein) could be obtained by dialysis of this suspension against three changes with 20 mM Tris-HCl, pH 7.4, and 1 mM dithiothreitol, or the protein could be further purified by anion-exchange chromatography on a Mono Q FPLC column (Pharmacia).

Mono Q Purification of Inclusion Bodies. To remove poorly soluble inclusion body proteins, the inclusion body preparation was centrifuged at 120000g_{av} (Beckman 60Ti rotor) for 30 min, and the resulting supernatant was passed through a 0.45-μm Millex-HA filter unit (Millipore). The p62-enriched preparation was chromatographed on a Mono Q HR 10/10 column equilibrated in 6 M urea, 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, and 2 mM dithiothreitol (equilibration buffer) at 4 °C using a Pharmacia FPLC system. During a typical run, the column was washed with 20 mL of equilibration buffer and eluted with a linear gradient (50 mL) of 0−0.3 M NaCl in equilibration buffer at a flow rate of 1 mL/min. Column fractions were analyzed using SDS−PAGE, and fractions containing p62 were pooled prior to dialysis against 20 mM Tris-HCl, pH 7.4, with or without 1 mM dithiothreitol at 4 °C. The circular dichroism spectrum from 190 to 240 nm was determined on recombinant Mono Q-purified p62 by using a constrained statistical regularization procedure (Provencher & Glockner, 1981).

Purification of O-GlcNAc Transferase. The GlcNAc transferase used in these experiments was partially purified from rabbit reticulocyte lysate (Promega) according to the method of Haltiwanger et al. (1992a) except that the 30% ammonium sulfate-saturated pellet was loaded on a 10-mL phenyl-Sepharose 4B column washed with 10 mM Tris-HCl, pH 8.0, 0.25 M ammonium sulfate, and 1 mM dithiothreitol (50 mL), followed by 10 mM Tris-HCl, pH 8.0, 0.1 M ammonium sulfate, and 1 mM dithiothreitol (50 mL), and eluted with 60% ethylene glycol, 10 mM Tris-HCl, pH 8.0, and 1 mM dithiothreitol (100 mL). Fractions (2 mL) collected throughout the separation were assayed for protein (absorbance at 280 nm) and *O*-GlcNAc transferase activity, and those containing activity were pooled. The enzyme was dialyzed against 50 mM Tris-HCl, pH 7.4, and 1 mM $MgCl_2$ and concentrated using a Centricon 30 microconcentrator (Amicon). Pierce BCA protein assay reagent was used to estimate protein concentrations.

O-GlcNAc Transferase Assay. Purified p62 (5 μg) was bound to nitrocellulose membranes (Schleicher & Schuell, BA85, 0.5 cm in diameter) in a total volume of 40 μL in 1.5-mL Eppendorf microcentrifuge tubes. The protein binding capacity of these membranes was 30 μg, and protein binding was found to be linear up to at least 14 μg using radiolabeled p62 (data not shown). Similar results were obtained when p62 was bound to the nitrocellulose membranes following dialysis against 2 M urea, 50 mM Tris-HCl, pH 7.4, and 1 mM $MgCl_2$ or 50 mM Tris-HCl, pH 7.4, and 1 mM $MgCl_2$. The buffer containing 2 M urea was preferred to prevent excessive aggregation and precipitation of p62 during storage at 4 °C. Bovine serum albumin (5 mg/mL, fraction V) was added in excess to block the rest of the protein binding sites on the nitrocellulose membranes. The membranes were rinsed with water to remove excess unbound bovine serum albumin prior to assays. The reaction mixture contained 50 mM Tris-HCl, pH 7.4, 12.5 mM $MgCl_2$, and 4−6 μCi (12 μM) UDP-[^{3}H]GlcNAc (Dupont) in a final volume of 40 μL. Partially purified enzyme (either the 30% NH_4SO_4 precipitate of rabbit reticulocyte lysate or the phenyl-Sepharose purified enzyme) was added at the start of the incubation (37 °C, 220 rpm for 60 min), and assays were terminated by aspiration of the reaction mixture. Membranes were washed with four changes of water, and radioactivity incorporated into the p62 bound to the nitrocellulose membranes was determined by liquid scintillation spectrometry. Under these standard conditions, the assay was found to be linear with respect to enzyme concentration (50−230 μg of 30% NH_4SO_4 precipitate or 4−40 μg of phenyl-Sepharose purified enzyme) and time (0−90 min, data not shown).

Deletion Mutants. The various deletion mutants of rat p62 were generated by polymerase chain reaction or by restriction digestion. Full-length rat p62 was cloned into the transcription vector pGEM-3Zf$^+$ (Promega, pGEMp62). Transcription was carried out *in vitro* following linearization with *Bam*HI.

Carboxyl-Terminal Deletions. Deletions removing various amounts of the carboxyl terminus are designated p62−6 (405 amino acids), p62−8 (447 amino acids), and p62−9 (475 amino acids). The construct p62−6 was generated by subcloning p62 cDNA into the pGEM7f vector downstream of the T7 RNA polymerase promoter region to generate pGEM7Zp62. This was cut with restriction enzymes *Pfl*M1 and *Bam*HI, and the 5′ overhang was filled in with Klenow fragment. Blunt-end ligation was then performed at room temperature. To generate p62−8, pGEM7Zp62 was cut with *Ppu*MI and *Bam*HI, and the appropriate band was isolated with low melting point agarose. The vector was then directly religated with DNA ligase. The construct p62−9 was formed by cutting pGEM7Zf+p62 with *Bgl*II and *Bam*HI, filling in the overhangs, and religating.

For *in vitro* transcription, p62−6, p62−8, and p62−9 were linearized with *Nsi*I and incubated with T7 RNA polymerase. All transcripts were capped using a commercially available mRNA capping kit (mCAP, Stratagene).

Other Deletions. The It deletion removes the Ser/Thr-rich central domain of the p62 coding sequence (Ser^{270}-Thr^{294}). The It construct was generated by polymerase chain

reaction from the vector pGEM3Zf+p62. The amino terminus of p62 was amplified using the SP6 promoter primer and the primer 5′MI(AGGCAAAGCCGGCACCAGGAGCTGCTC). The carboxyl terminus was amplified using the T7 promoter primer and 3′MI(TCCTGGTGCCGGCTTTGCCTTAAGCCTG). The primers 5′MI and 3′MI are complementary so they could be annealed and the whole construct amplified using the T7 and M13 reverse primers to generate It. The polymerase chain reaction product was cut with *Eco*RI and ligated into the vector pGEM3Zf− downstream of the Sp6 polymerase promoter.

The CO deletion lacks all of the Ser/Thr-rich region and amino terminus of p62. It was generated by fusing the 5′ flanking region of p62 with the carboxyl-terminal region. pGEMp62 was amplified by polymerase chain reaction using the Sp6 promoter and the primer 3′CO5-1(CGCGTAGGTCATGGCTGTGGAATTTCCAGG) for the 5′ flanking region (including the Met start site). The T7 promoter primer and 5′CO3-1(ACAGCCATGACCTACGCGCAGCTGGAA) were used to amplify the carboxy-terminal region. 3′CO5-1 and 5′CO3-1 contain complimentary ends, and the two polymerase chain reaction products were annealed and amplified with the T7 and Sp6 promoter primers. The product was cut with *Eco*RI and ligated into pGEM3Zf+ downstream of the Sp6 promoter.

The construct ItNH was prepared in a fashion identical to that of It. This polymerase chain reaction product was found to have a single base pair insertion, resulting in an aberrant translation of and premature termination of protein just following the Ser/Thr-rich region. The protein produced was expected to be 310 amino acids in length. All of the mutants were sequenced to confirm the coding sequence.

In Vitro Translation of Recombinant p62. The p62 transcript was translated in rabbit reticulocyte lysate (Bethesda Research Laboratories) or wheat germ extract (Promega) containing [^{35}S]methionine (1000 Ci/mmol) (Amersham Corp.) as previously described (Starr et al., 1990).

Protease Digestion of p62. p62 (0.27 mg of protein) was incubated in rabbit reticulocyte lysate (Promega, 50% by volume) with 3 μCi UDP-[14C]GlcNAc (New England Nuclear) at 37 °C for 16 h. The glycosylated p62 was precipitated with 90% ice-cold acetone and resuspended in sample buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% 2-mercaptoethanol, 0.0025% bromophenol blue) for SDS-PAGE on 10% Tris−glycine precast gels (Novex). Gels were run at constant current 25 mA/gel; stained with 0.1% Coomassie Brilliant Blue R-250, 40% methanol, and 10% acetic acid for 10 min; and destained with 10% methanol and 10% acetic acid for 20−30 min. Bands corresponding to glycosylated p62 were cut out and electroeluted using and Isco Gel Eluter set at 1 W for 16 h. Glycosylated p62 (0.04 mg of protein) was incubated with 15 μg of endoproteinase Glu-C or endoproteinase Lys-C (0.2 mg/mL in resuspension buffer provided by Promega) for 9 h at 37 °C to allow for proteolysis to go to completion; under these conditions, proteolysis was >90% complete after a 60-min incubation. Samples were heated to 80 °C for 5 min and immunoprecipitated using antipeptide antisera 474, 475, and 476 recognizing various regions of rat p62. The regions of rat p62 corresponding to these peptides is indicated in Figures 5A,B. Preparation of the antisera 474 has been previously described (Starr et al., 1990). Antisera 476 was prepared by immunizing a rabbit with a fusion protein containing β-galactosidase linked to the carboxyl-terminal 154 amino acids of p62 (D'Onofrio et al., 1988). Antisera 475 was prepared against the peptide AQLESLINKWSLELEDQERH. Immunoprecipitation was carried out as described below. Protease digests, diluted 1/30 with RIPA buffer (20 mM Tris-HCl, pH 7.2, 0.15 M NaCl, 1% aprotinin (Sigma), 1% deoxycholate, 1% NP40, 0.1% SDS), were incubated with p62 antisera (474, 475, 476) diluted 1:50 in 20 mM Tris-HCl, pH 7.2, 0.15 M NaCl, 0.5% aprotinin, 0.5% NP40, and 0.02% SDS for 30 min at 4 °C in 10 × 75 polypropylene tubes and immunoprecipitated with 5% protein A−agarose (Boehringer Mannheim). The pellet was washed once with high salt buffer (1 M NaCl, 20 mM Tris-HCl, pH 7.2, 0.5% aprotinin, 0.5% NP40, 0.02% SDS) followed by two washes with low salt buffer (0.15 M NaCl, 20 mM Tris-HCl, pH 7.2, 0.5% aprotinin, 0.5% NP40, 0.02% SDS). Following SDS−PAGE, the immunoprecipitated proteolytic fragments were visualized by autoradiography on X-Omat film (Kodak).

Synthetic Peptide Substrate for O-Linked GlcNAc Transferase. A 43 amino acid peptide corresponding to the central Ser/Thr-rich region of p62 was produced by solid-phase synthesis on an Applied Biosystems Model 418 peptide synthesizer. This sequence is GFSLKAPGAAPGASTSTTTTTTTTTTTTASTSSSTTTTGFLAS. The peptide was subsequently purified by reverse-phase HPLC (C-4, data not shown).

RESULTS

Expression and Purification of Recombinant p62 from E. coli. Rat p62 was cloned into the *Nde*I and *Bam*HI sites in the pET-3a vector, as shown in Figure 1A. This expression vector was introduced into *E. coli* BL-21 (pLys-S) cells which have an inducible T7 polymerase. After growth and induction, the cells were harvested, and inclusion bodies containing most of the newly synthesized p62 were isolated. The recombinant rat p62 was solubilized from these inclusion bodies with 6 M urea and further purified to near homogeneity by anion-exchange chromatography on Mono Q FPLC (Figure 1B). The purified protein remained soluble during dialysis against low ionic strength buffers.

The observed secondary structure in the recombinant purified p62 was found to be consistent with proper refolding of the protein following dialysis. The circular dichroism spectrum suggests an α-helical content of 24% and 42% β-pleated sheet structure (data not shown).

In Vitro Glycosylation of Recombinant p62. Rabbit reticulocyte lysate had been previously shown to glycosylate p62 during cell-free translation of *in vitro* transcripts (Starr & Hanover, 1990). Here we examined the ability of the recombinant p62 produced in *E. coli* to be glycosylated by reticulocyte lysate. Recombinant p62 was incubated for various intervals with UDP-[^{14}C]GlcNAc and rabbit reticulocyte lysate. The extent of glycosylation was monitored by SDS−PAGE and was dependent upon the concentration of added p62 (Figure 2).

A more convenient assay was developed for detailed kinetic analysis of p62 glycosylation. Renatured recombinant p62 was immobilized on filters allowing simple removal of unincorporated UDP-[^{14}C]GlcNAc. This assay was demonstrated to be linear with respect to time and *O*-GlcNAc transferase concentration (see Experimental Procedures).

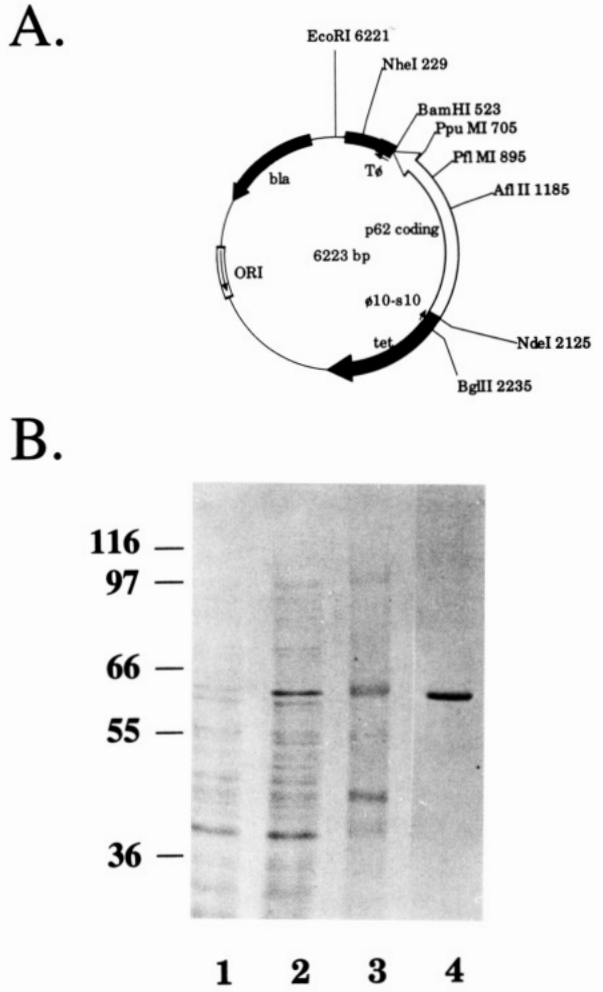

FIGURE 1: Expression and purification of rat nuclear p62 from BL-21 (pLys) cells transformed with the p62 Pet3a expression vector. (A) For protein expression, the p62 coding region was inserted into the Pet3a expression vector at the *Nde*I and *Bam*HI restriction sites, between the T7 promoter, ø10-s10, and T7 terminator, Tø. Abbreviations: (bla) ampicillin resistance marker and (tet) tetracycline resistance marker. (B) SDS−PAGE analysis of proteins from Pet3a transformed cells. Cells were grown in log phase (lane 1) and induced for 90 min with IPTG (lane 2); inclusion bodies were isolated (lane 3) from which p62 was purified by Mono Q FPLC (lane 4). Proteins were detected with Coomassie Brilliant Blue. The migration profile of molecular mass standards is shown to the left.

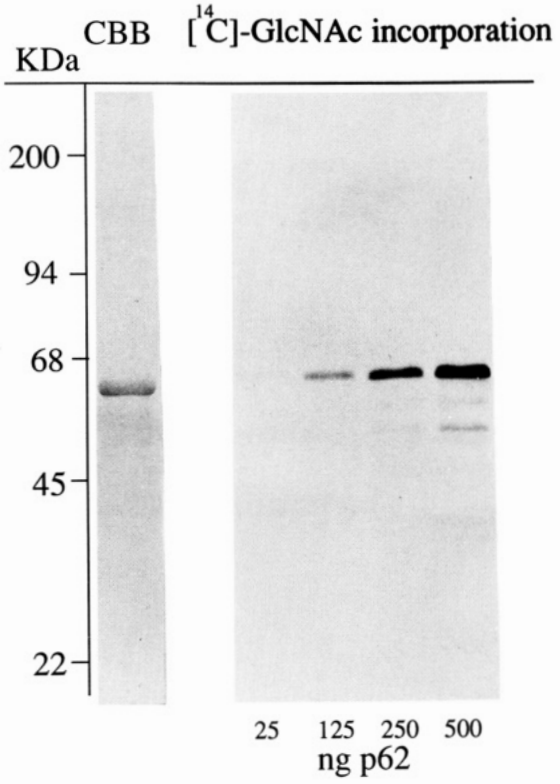

FIGURE 2: *In vitro* glycosylation of recombinant p62. Purified recombinant p62 (25−500 ηg) was incubated with rabbit reticulocyte lysate (25 μL/assay) and UDP-[^{14}C]GlcNAc (0.5 μCi) for 1 h at 37 °C in a final volume of 50 μL. Reactions were stopped by acetone precipitation; proteins were separated by SDS−PAGE; and the incorporation of [^{14}C]GlcNAc was visualized by autoradiography. The panel on the left shows Coomassie Brilliant Blue R 250 (CBB) staining of purified recombinant p62. The panel on the right is an autoradiograph of [^{14}C]GlcNAc incorporation. The migration profile of molecular mass standards is shown to the left.

Analysis of reaction rate as a function of substrate concentration is consistent with a single set of modified sites with a $K_m = 0.30$ μM with respect to p62 and a $K_m = 1.8$ μM with respect to UDP-GlcNAc (Figure 3).

Localization of Regions of Glycosylation in p62. Using galactosyltransferase to label all terminal GlcNAc residues on proteins isolated from rat liver nuclei (Holt et al., 1987b), it was shown that there are 10−12 *O*-linked GlcNAc moieties on p62. Only the site at Ser^{471} has so far been identified (D'Onofrio et al., 1988). In order to identify the other sites of *O*-linked glycosylation, we constructed recombinant p62 molecules with deletions in the amino terminal, central Ser/Thr-rich, and carboxy terminal regions of the protein as shown in Figure 4B. We took advantage of the observation that rabbit reticulocyte lysate is capable of *O*-linked GlcNAc glycosylation, whereas wheat germ extract is devoid of this activity. By comparing the size difference of the *in vitro* translation products produced in the two systems, we estimated what fraction of the total amount of glycosylation could be attributed to the deleted part of the protein. As shown in Figure 4A, the original p62 translation product undergoes a 3200-Da shift due to glycosylation. This is somewhat more than the 2000−2500-Da shift that would be expected from the addition of 10−12 GlcNAc residues and suggests that glycosylation probably alters protein conformation and/or SDS binding. The CO construct, which has the first 328 amino acids deleted leaving only the carboxy terminus, showed no significant shift following glycosylation, thereby suggesting that the carboxyl terminus is very poorly glycosylated. This, however, does not rule out glycosylation at Ser^{471} in this construct since the molecular mass shift due to a single GlcNAc would probably be too small to be detected as a gel shift in this experiment. The ItNH construct, which has the final 215 amino acids of the carboxy terminus deleted, showed a 3070-Da shift following glycosylation. This again is consistent with nearly all of the glycosylation sites in p62 being located in the amino terminal and central serine−threonine region. Additional experiments with 50, 75, and 125 amino acid carboxy terminal deletions yielded results consistent with those obtained with ItNH (data not shown). The It construct, which has the central 25 amino acids containing the Ser/Thr-rich region deleted, only undergoes a 1200-Da shift after glycosylation, that is, only 38% (1200/3200) of the glycosylation of p62 occurs when the Ser/Thr-rich region is deleted. These data argue that the bulk of the glycosylation, 62%, must be in the Ser/Thr-rich region.

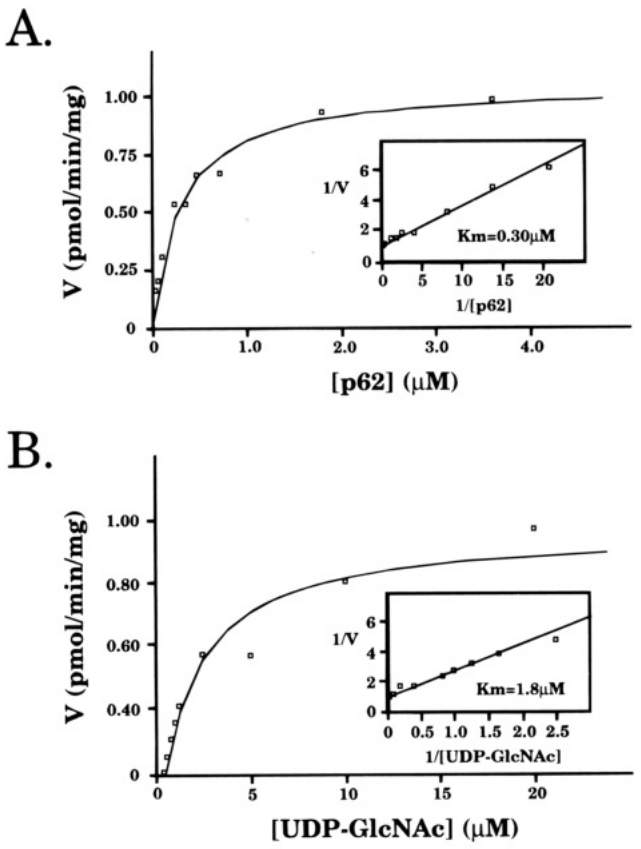

FIGURE 3: Steady-state kinetics of the mammalian *O*-linked GlcNAc transferase. (A) Effect of p62 substrate concentration on *O*-linked GlcNAc transferase acticity. Assays were done under standard conditions as described under Experimental Procedures, except using 0.05−3.6 μM p62 and saturating amounts of UDP-[^{14}C]GlcNAc (12 μM). (B) Effect of UDP-GlcNAc concentration on *O*-linked GlcNAc transferase activity. Assays were done under standard conditions as described under Experimental Procedures, except using 0.4−20 μM (0.6−30 μCi) of UDP-[^{14}C]GlcNAc and saturating amounts of p62 (2.4 μM).

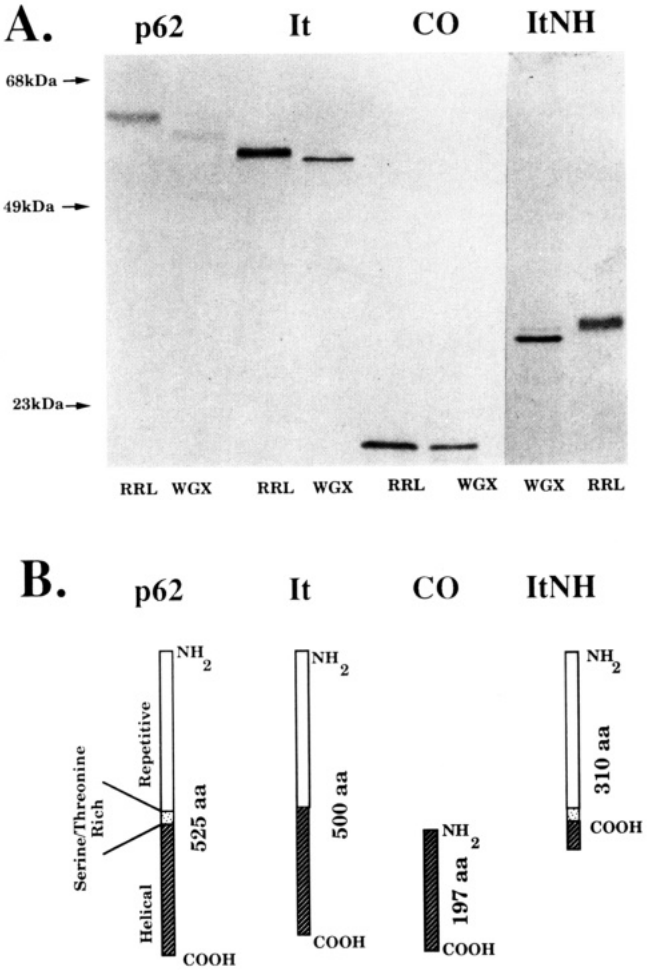

FIGURE 4: Deletion analysis of *in vitro* translation products. (A) p62 or deletion mutants It, CO, and ItNH were translated in rabbit reticulocyte lysate (RRL) or wheat germ extract (WGX) as described under Experimental Procedures. The *in vitro* translation products were separated by SDS−PAGE and visualized by autoradiography. Migration of molecular mass standards is shown at the left. (B) Diagrammatic representation of the domain structure of p62. Amino terminal repetitive (white), central Ser/Thr-rich (stippled), and carboxylic helical (hatched) domains of p62 and the deletion mutants It, CO, and ItNH.

Because of the uncertainty involved in determining the number of glycosylation sites based on molecular mass changes, an alternative approach was developed (Figure 5). In these experiments, recombinant p62 previously radiolabeled with UDP-[^{14}C]GlcNAc was proteolyzed with either endoproteinase Lys C, which cleaves at lysine residues, or endoproteinase Glu C, which cleaves at glutamic acid residues. The radiolabeled proteolytic fragments resulting from these digests were then identified by immunoprecipitation with specific p62 antiserum. The predicted proteolytic cleavage sites as well as the regions of p62 recognized by antisera 474, 475, and 476 are shown in Figure 5. Initial control experiments demonstrated that all of these antisera were capable of immunoprecipitating the intact p62 molecule. Therefore, if the labeled glycosylation sites were physically separated from the antipeptide epitope by proteolysis, the antisera would not immunoprecipitate the radiochemically labeled fragments.

The amino terminal antiserum 474 immunoprecipitated the main proteolytic fragment from endoproteinase Glu C cleavage of glycosylated p62; however, no similar labeled fragment was recovered following endoproteinase Lys C cleavage of p62. This suggests an absence of glycosylation in the 7200-Da fragment generated by proteolysis at Lys^{22} and Lys^{97} and containing the amino terminal epitope recognized by antiserum 474 (Figure 5). Antiserum 475, on the other hand, precipitates a much larger than predicted fragment, suggesting that glycosylation of this peptide may be masking potential cleavage sites at Lys^{374} and Lys^{404} on both sides of this epitope.

Glycosylation of a Ser/Thr-Rich Peptide Derived from p62. Synthetic peptides have been previously utilized to study the substrate specificity of the *O*-linked GlcNAc transferase (Haltiwanger et al., 1990). While YSDSPSTST (K_m = 9.9 mM) (Haltiwanger et al., 1992a) was shown to be a substrate for the enzyme, YSDSGSTST did not serve as a substrate, indicating that a stretch of serines and threonines alone is not sufficient for activity. We also did not find any activity for several commercially available poly-L-Ser polypeptides (data not shown). However, since the major glycosylated region of the purified recombinant p62 molecule appeared to be at the central Ser/Thr-rich domain, we chemically synthesized and HPLC-purified a 43 amino acid peptide from Gly^{257} to Ser^{299} to see if it would be recognized as a substrate by the mammalian *O*-linked GlcNAc transferase. As shown in Figure 6, this peptide was found to be a much better substrate for the GlcNAc transferase (K_m = 30 μM) than had been previously described for the peptide

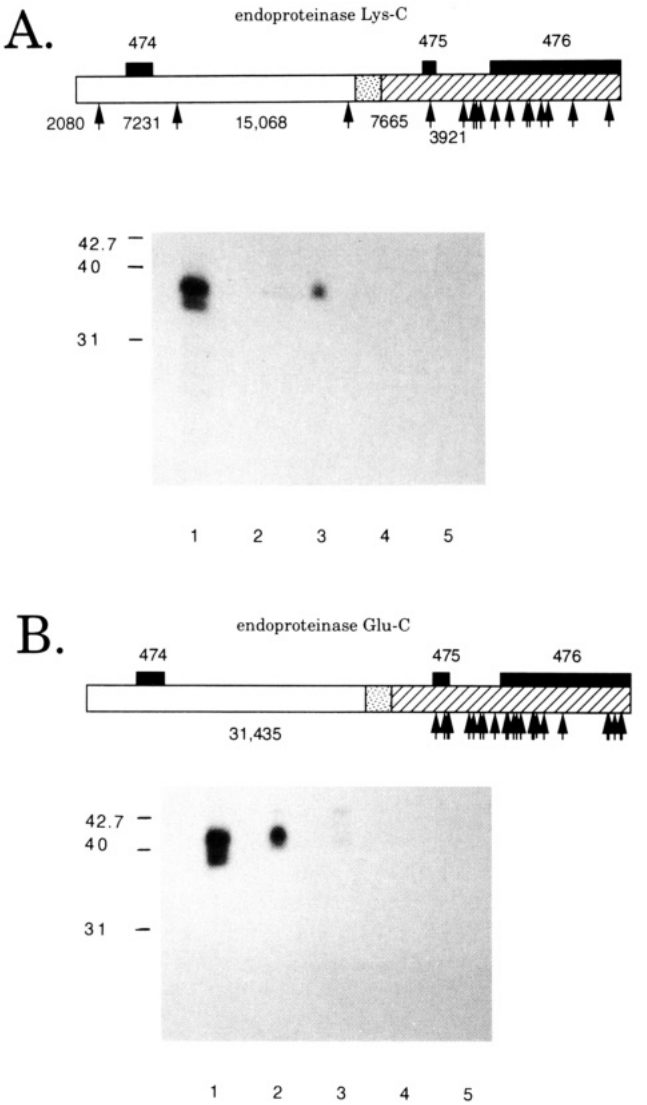

FIGURE 5: Immunoprecipitation of [^{14}C]GlcNAc labeled proteolytic fragments, following *in vitro* glycosylation of p62. Proteolytic fragments of [^{14}C]GlcNAc labeled p62 digested with endoproteinase Lys-C (A) or endoproteinase Glu-C (B) were immunoprecipitated with antisera 474, 475, and 476, which recognize the amino terminal, central, and carboxyl terminal portions of the protein, respectively. The immunoprecipitated products were separated by SDS−PAGE and visualized by autoradiography. The migration profile of molecular weight standards is shown to the left. The radiolabeled proteolytic fragments prior to immunoprecipitation (lane 1) or after immunoprecipitation with antisera 474 (lane 2), 475 (lane 3), and 476 (lane 4) and control rabbit preimmune serum (lane 5) are shown. A diagrammatic representation of the structural domains of p62, including antisera binding sites and potential protease cleavage sites, is shown at the top of each figure. The intact p62 molecule contains amino terminal repetitive (open box), central Ser/Thr-rich (stippled box), and carboxyl terminal helical (hatched box) domains. The location of the p62 polyclonal antibody binding sites are depicted by the closed boxes above each figure. The potential protease cleavage sites are represented by the arrows beneath each figure. The predicted molecular masses of the larger fragments expected from complete cleavage of all proteolytic sites is given below each fragment.

substrate YSDSPSTST (K_m = 9.9 mM). This Ser/Thr-rich peptide was also found to be a good inhibitor of p62 glycosylation (K_i = 15 μM, Figure 7), and the linear K_m/V_{max} versus inhibitor [peptide] plot confirmed that the peptide was acting as a pure competitive inhibitor of p62 glycosylation.

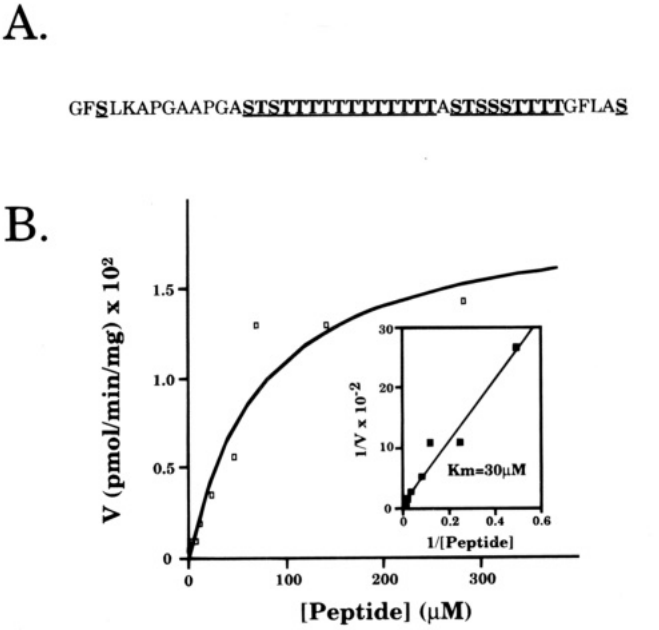

FIGURE 6: *In vitro* glycosylation of a Ser/Thr-rich peptide derived from p62 by the mammalian *O*-linked GlcNAc transferase. (A) Protein sequence of the chemically synthesized peptide derived from p62 Gly^{257}−Ser^{299}. Hydroxyamino acids are identified in bold type and underlined as potential glycosylation sites. (B) Glycosylation of the Ser/Thr-rich peptide by the *O*-GlcNAc transferase. Assays were done under standard conditions as described under Experimental Procedures, with 2−285 μM Ser/Thr-rich peptide and saturating amounts of UDP-GlcNAc (12 μM). Similar results were obtained when glycosylation assays were done in solution and analyzed by SDS−PAGE, and autoradiographs were quantified by densitometry (data not shown).

DISCUSSION

In this paper, we expressed the major rat nuclear pore protein, p62, in *E. coli* and purified the unglycosylated protein to near homogeneity. The recombinant p62 was an excellent substrate for the mammalian *O*-linked GlcNAc transferase; its apparent K_m was 30 000-fold lower than any previously reported substrate (Haltiwanger et al., 1992a). Consistent with its cytoplasmic localization, the *O*-GlcNAc transferase showed a high affinity for UDP-GlcNAc, (K_m = 1.8 μM). This K_m value is quite similar to other sugar nucleotide-utilizing proteins exposed to the cytoplasm such as the UDP-GlcNAc and UDP-Glc transporters present in the microsomal vesicles derived from rat liver endoplasmic reticulum (K_m = 3−4 μM) (Hirschberg & Snider, 1987).

Rat p62 is modified with 10−12 *O*-linked GlcNAc moieties almost exclusively on Ser residues (Holt et al., 1987b). We previously mapped on site of glycosylation at Ser^{471} in the carboxyl terminus (D'Onofrio et al., 1988). Previously, it has been suggested that the glycosylation sites in human p62 are present in the amino terminal 301 amino acids (Carmo-Fonseca et al., 1991). Deletion analysis also provided indirect evidence that the region between amino acids 153 and 435 in *X. laevis* p62 and between amino acids 222 and 330 in mouse p62 are glycosylated (Cordes et al., 1991). Based on the observation of a time lag in glycosylation for several histidine-tagged fusion peptides from different regions of mouse p62, the existence of high and low affinity acceptor sites in these peptides was proposed (Cordes & Krohne, 1993). However, our kinetic analysis using intact recombinant p62 or synthetic peptides over a

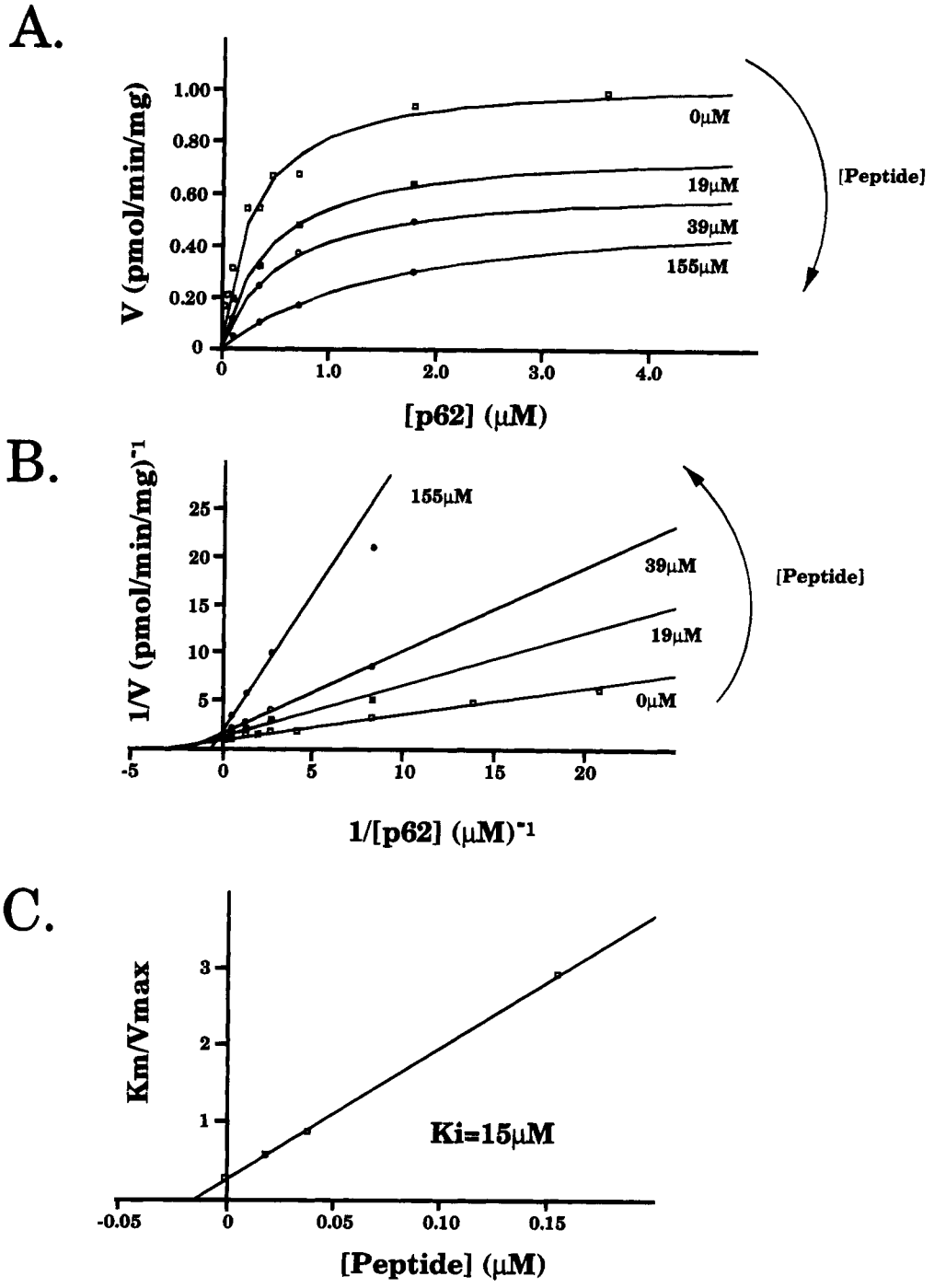

FIGURE 7: Inhibition of p62 glycosylation by the Ser/Thr-rich peptide. (A) The *v* (velocity) versus p62 concentration plot for the mammalian GlcNAc transferase in the presence of increasing concentrations of inhibitor peptide (0−155 μM). Assays were performed under standard conditions as described under Experimental Procedures, with 0.05−3.6 μM p62 and saturating UDP-GlcNAc (12 μM). (B) Effect of the Ser/Thr-rich peptide on the Lineweaver−Burk plot for the mammalian GlcNAc transferase. (C) Plot of K_m/V_{max} versus Ser/Thr-rich peptide concentration for the mammalian GlcNAc transferase.

wide range of concentrations are all consistent with a single set of acceptor sites. The observed lag in glycosylation between different p62 fusion peptides may instead be due to differential aggregation between peptides, since p62 has been shown to aggregate into dimers using specific regions at both the carboxylic and amino terminal ends of the molecule (Buss et al., 1994). Alternatively, the N-terminal extensions, which made up as much as one-sixth of these fusion peptides, might alter binding to nearby glycosylation sites.

In this paper, using two complementary approaches, we are able to localize 6−8 of the potential 10−12 glycosylation sites to the central Ser/Thr-rich region from amino acid Ser^{270} to Thr^{294}. Since this region of the molecule contains only six Ser residues and it has already been shown that glycosylation of p62 occurs almost exclusively on Ser residues, it is likely that all of these Ser residues are normally glycosylated. Because the sequence requirements for *O*-linked GlcNAc addition are not yet established, it is of interest to examine the sequence of these likely glycosylation sites. The Ser residues at these sites are located within stretches also rich in Thr. Only one of these Ser residues is located within three residues of a Pro in contrast to the original observation that Pro residues are typically found one

to three amino acid residues from the attachment site (Haltiwanger et al., 1992b), similar to *O*-linked GalNAc glycosylation (Wilson et al., 1991). More recently, however, several glycosylation sites on the serum response transcription factor (Reason et al., 1992) and on talin (Hagmann et al., 1992) have been identified that lack Pro residues nearby, suggesting that Pro residues are not an absolute requirement for glycosylation and that accessibility of the site to the enzyme instead of primary amino acid sequence may determine which hydroxyamino acids are glycosylated. However, the fact that the enzyme seems to specifically glycosylate Ser residues in regions rich in Thr suggests a strict substrate specificity for the enzyme. It is interesting in this regard that the Ser residues found between Ser^{270} and Thr^{294}, ASTSTT and ASTSSSTT, are located in regions similar to known glycosylation sites in the serum response transcription factor (Reason et al., 1992) i.e., TTST and TQTSSSGT.

We chemically synthesized a 43 amino acid peptide corresponding to Gly^{257} to Ser^{299} from the central Ser/Thr region of the molecule and free of regions previously shown to be involved in aggregating p62. Here we show that such a peptide can act as a substrate for the *O*-GlcNAc transferase *in vitro* (K_m = 30 μM) and that it can compete with glycosylation of p62 as a pure competitive inhibitor (K_i = 15 μM).

Since the initial identification of *O*-linked GlcNAc on the surface of lymphocytes (Torres & Hart, 1984) and the nuclear pore proteins (Holt et al., 1987b), it has been found on many other cytoplasmic and nuclear proteins including RNA polymerase II and its associated transcription factors (Kelly et al., 1993; Reason et al., 1992; Jackson & Tjian, 1988; Lichtsteiner & Schibler, 1989; Datta et al., 1989), cytoskeletal proteins (Dong et al., 1993; Roquemore et al., 1992; Hagmann et al., 1992; Holt et al., 1987a; Luthi et al., 1991; Chou et al., 1992; King & Hounsell, 1989), a component of clathrin-coated pits (Murphy et al., 1994), viral proteins (Privalsky, 1990; Whitford & Faulkner, 1992; Mullis et al., 1990; Caillet-Boudin et al., 1989; Benko et al., 1988; Gonzalez & Burrone, 1991), schistosomes (Nyame et al., 1987), Drosophila chromatin (Kelly & Hart, 1989), and a phosphoprotein phosphatase (Meikrantz et al., 1991). Most of these proteins are poorly glycosylated, typically having less than three potential sites per molecule and often with less than 50% of these sites occupied. The few exceptions are the nucleoporins p62, POM121, and p180 and the transcription factor Sp1, which are suggested to be more heavily glycosylated (Holt et al., 1987b; Jackson & Tjian, 1988; Hallberg et al., 1993). It is interesting to note that all of these highly glycosylated proteins have Ser/Thr-rich regions. As we demonstrated here, roughly 60% of the total glycosylation sites of rodent p62 occur in the Ser/Thr-rich region; one may expect that similar regions in p180, POM121, and Sp1 may also be the major sites of glycosylation in these proteins. Hydroxyamino acid-rich domains have also recently been implicated in receptor internalization and signal transduction (Moro et al., 1993; Benya et al., 1993; Merida et al., 1993). While it is tempting to attribute these activities to phosphorylation at these sites, none of these proteins have yet been examined for glycosylation with *O*-linked GlcNAc.

The function of the highly glycosylated Ser/Thr-rich domain in p62 is not clear, but its location as a "hinge region" in the central part of the protein, between the amino terminal β-pleated sheet and the carboxyl terminal α-helical domains, suggests a potential function for *O*-glycosylation of p62. Since clustering of *O*-linked sugars has been shown to be important in providing conformational rigidity to the protein backbone, *O*-glycosylation of p62 may be important for extending the backbone of the protein and producing a stiffer elongated molecule (Butenhof & Gerken, 1993; Jentoft, 1990). It is possible that dynamic glycosylation of the hinge region could serve to trigger supramolecular assembly of the nuclear pore proteins following mitosis.

REFERENCES

Benko, D. M., Haltiwanger, R. S., Hart, G. W., & Gibson, W. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2573−2577.

Benya, R. V., Fathi, Z., Battey, J. F., & Jensen, R. T. (1993) *J. Biol. Chem. 268*, 20285−20290.

Buss, F., Kent, H., Stewart, M., Bailer, S. M., & Hanover, J. A. (1994) *J. Cell Sci. 107* (Pt 2), 631−638.

Butenhof, K. J., & Gerken, T. A. (1993) *Biochemistry 32*, 2650−2663.

Caillet-Boudin, M., Strecker, G., & Michalski, J. (1989) *Eur. J. Biochem. 184*, 205−211.

Carmo-Fonseca, M., Kern, H., & Hurt, E. C. (1991) *Eur. J. Cell Biol. 55*, 17−30.

Chou, C. F., Smith, A. J., & Omary, M. B. (1992) *J. Biol. Chem. 267*, 3901−3906.

Cordes, V. C., & Krohne, G. (1993) *Eur. J. Cell. Biol. 60*, 185−195.

Cordes, V., Waizenegger, I., & Krohne, G. (1991) *Eur. J. Cell Biol. 55*, 31−47.

Datta, B., Ray, M. K., Chakrabarti, D., Wylie, D. E., & Gupta, N. K. (1989) *J. Biol. Chem. 264*, 20620−20624.

Davis, L. I., & Blobel, G. (1986) *Cell 45*, 699−709.

Dong, D. L., Xu, Z. S., Chevrier, M. R., Cotter, R. J., Cleveland, D. W., & Hart, G. W. (1993) *J. Biol. Chem. 268*, 16679−16687.

D'Onofrio, M., Starr, C. M., Park, M. K., Holt, G. D., Haltiwanger, R. S., Hart, G. W., & Hanover, J. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9595−9599.

Gonzalez, S. A., & Burrone, O. R. (1991) *Virology 182*, 8−16.

Hagmann, J., Grob, M., & Burger, M. M. (1992) *J. Biol. Chem. 267*, 14424−14428.

Hallberg, E., Wozniak, R. W., & Blobel, G. (1993) *J. Cell Biol. 122*, 513−521.

Haltiwanger, R. S., Holt, G. D., & Hart, G. W. (1990) *J. Biol. Chem. 265*, 2563−2568.

Haltiwanger, R. S., Blomberg, M. A., & Hart, G. W. (1992a) *J. Biol. Chem. 267*, 9005−9013.

Haltiwanger, R. S., Kelly, W. G., Roquemore, E. P., Blomberg, M. A., Dong, L. Y., Kreppel, L., Chou, T. Y., & Hart, G. W. (1992b) *Biochem. Soc. Trans. 20*, 264−269.

Hanover, J. A., Cohen, C. K., Willingham, M. C., & Park, M. K. (1987) *J. Biol. Chem. 262*, 9887−9894.

Hirschberg, C. B., & Snider, M. D. (1987) *Annu. Rev. Biochem. 56*, 63−87.

Holt, G. D., Haltiwanger, R. S., Torres, C. R., & Hart, G. W. (1987a) *J. Biol. Chem. 262*, 14847−14850.

Holt, G. D., Snow, C. M., Senior, A., Haltiwanger, R. S., Gerace, L., & Hart, G. W. (1987b) *J. Cell Biol. 104*, 1157−1164.

Jackson, S. P., & Tjian, R. (1988) *Cell 55*, 125−133.

Jentoft, N. (1990) *Trends Biochem. Sci. 15*, 291−294.

Kelly, W. G., & Hart, G. W. (1989) *Cell 57*, 243−251.

Kelly, W. G., Dahmus, M. E., & Hart, G. W. (1993) *J. Biol. Chem. 268*, 10416−10424.

King, I. A., & Hounsell, E. F. (1989) *J. Biol. Chem. 264*, 14022−14028.

Lichtsteiner, S., & Schibler, U. (1989) *Cell 57*, 1179−1187.
Luthi, T., Haltiwanger, R. S., Greengard, P., & Bahler, M. (1991) *J. Neurochem. 56*, 1493−1498.
Meikrantz, W., Smith, D. M., Sladicka, M. M., & Schlegel, R. A. (1991) *J. Cell Sci. 98* (Pt 3), 303−307.
Merida, I., Williamson, P., Kuziel, W. A., Greene, W. C., & Gaulton, G. N. (1993) *J. Biol. Chem. 268*, 6765−6770.
Miller, M., Park, M. K., & Hanover, J. A. (1991) *Physiol. Rev. 71*, 909−949.
Moro, O., Lameh, J., & Sadee, W. (1993) *J. Biol. Chem. 268*, 6862−6865.
Mullis, K. G., Haltiwanger, R. S., Hart, G. W., Marchase, R. B., & Engler, J. A. (1990) *J. Virol. 64*, 5317−5323.
Murphy, J., Hanover, J. A., Froehlich, M., Dubois, G., & Keen, J. H. (1994) *J. Biol. Chem. 269*, 21346−21352.
Nyame, K., Cummings, R. D., & Damian, R. T. (1987) *J. Biol. Chem. 262*, 7990−7995.
Park, M. K., D'Onofrio, M., Willingham, M. C., & Hanover, J. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6462−6466.
Privalsky, M. L. (1990) *J. Virol. 64*, 463−466.
Provencher, S. W., & Glockner, J. (1981) *Biochemistry 20*, 33−37.
Reason, A. J., Morris, H. R., Panico, M., Marais, R., Treisman, R. H., Haltiwanger, R. S., Hart, G. W., Kelly, W. G., & Dell, A. (1992) *J. Biol. Chem. 267*, 16911−16921.
Roquemore, E. P., Dell, A., Morris, H. R., Panico, M., Reason, A. J., Savoy, L. A., Wistow, G. J., Zigler, J. S. J., Earles, B. J., & Hart, G. W. (1992) *J. Biol. Chem. 267*, 555−563.
Snow, C. M., Senior, A., & Gerace, L. (1987) *J. Cell Biol. 104*, 1143−1156.
Starr, C. M., & Hanover, J. A. (1990) *J. Biol. Chem. 265*, 6868−6873.
Starr, C. M., D'Onofrio, M., Park, M. K., & Hanover, J. A. (1990) *J. Cell Biol. 10*, 1861−1871.
Sukegawa, J., & Blobel, G. (1993) *Cell 72*, 29−38.
Torres, C. R., & Hart, G. W. (1984) *J. Biol. Chem. 259*, 3308−3317.
Whitford, M., & Faulkner, P. (1992) *J. Virol. 66*, 3324−3329.
Wilson, I. B., Gavel, Y., & von Heijne, G. (1991) *Biochem. J. 275* (Pt 2), 529−534.

BI9421533