2005

Characterization of NCOAT, a bifunctional enzyme with O-GlcNAcase and histone acetyltransferase activity.

Clifford A. Toleman
University of Alabama at Birmingham

CHARACTERIZATION OF NCOAT, A BIFUNCTIONAL ENZYME WITH *O*-GlcNAcase AND HISTONE ACETYLTRANSFERASE ACTIVITY

by

CLIFFORD A. TOLEMAN

A DISSERTATION

Submitted to the graduate school of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2005

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Ph.D. Program Cell Biology

Name of Candidate Clifford A. Toleman

Committee Chair Jeffrey E. Kudlow

Title Characterization of NCOAT, a Bifunctional Enzyme with *O*-GlcNAcase and Histone Acetyltransferase Activity

O-GlcNAcylation is a ubiquitous post-translational modification whereby the monosaccharide *N*-acetylglucosamine (GlcNAc) is affixed via a beta linkage to specific serine or threonine hydroxyls on target proteins. These modifications can be selectively removed by the enzyme *O*-GlcNAcase, which hydrolyzes the glycosidic bond and liberates the free protein and GlcNAc. *O*-GlcNAcase recognizes the same common core substrate, GlcNAc, that chitinases, hexosaminidases, and hyaluronidases recognize and is therefore likely to exhibit a mode of catalysis and substrate binding remarkably similar to that of these enzymes, which have a highly conserved mechanism for these functions as well as conserved amino acids which carry out these events. This information, along with the use of naturally occurring splice variant isoforms, was used to find the active site of *O*-GlcNAcase and to elucidate amino acids necessary for efficient catalysis. Once demonstrated, this information was utilized to determine the mechanism by which the diabetogenic drug and selective *O*-GlcNAcase inhibitor streptozotocin can perform its inhibitory effects on the enzyme. Since streptozotocin is a GlcNAc analog, the enzyme will attempt to catalyze it as it would a normal substrate, converting streptozotocin to a transition-state analog that is more stable than the natural ligand transition-state is and therefore competes for the active site more expediently. Over the course of these studies, *O*-

GlcNAcase was realized to have a domain in its C terminus with secondary structure similar to those of several different acetyltransferase active sites. *O*-GlcNAcase was found to possess acetyltransferase activity for specific target lysines in nucleosomal and free histone substrates, prompting us to rename the enzyme NCOAT, for nuclear and cytosolic *O*-GlcNAcase and acetyltransferase. The active site for this second activity was found to have secondary structure nearly identical to those previously characterized and had critical catalytic residues in positions corresponding to the counterpart enzymes.

DEDICATION

I would like to dedicate the enclosed work to my parents, Clifford Toleman, Sr., and Sandra Toleman, for raising me to be the man who is capable of publishing this dissertation. I thank them for always fostering my interests, giving me every assistance within their power, and readily accepting all of my choices. It is this dissertation that puts not only my work but all of their hard work into a tangible form- just one of many for which they should be grateful and of which they should be proud.

ACKNOWLEDGMENTS

There are a number of people I thank for their contributions during my dissertation research. Principally, I must express my full indebtedness to my graduate mentor, Dr. Jeffrey Kudlow, who has given me the opportunity and invaluable advice and teaching. He has always enthusiastically supported my ideas and has provided an ideal working environment for all of those in it, without which my interaction with my colleagues would not have been so fruitful. Accordingly, I thank the members of Dr. Kudlow's laboratory with whom I have worked for all of their assistance and companionship: Xiaoyang Yang, Fengxue Zhang, Lacinda Reisland, Damon Bowe, Thomas Whisenhunt, Yong Hu, Ping Huang, and Kai Wang. I extend my particular thanks to Dr. Andrew Paterson for all of his supreme kindness, patience, and help in its multitude of forums. He is a role model for his work ethic.

I give thanks to my graduate committee members, Drs. Xinbin Chen, Jeffrey Engler, Rakesh Patel, and Bradley Yoder, for their support and guidance in overseeing of my research. I also extend my gratitude to Ray Moore and Landon Wilson at the mass spectroscopy center and Dr. Ronald Shin at the nuclear magnetic resonance facility. Their experimentation and expertise in their respective fields were indispensable to my research.

Finally, my accomplishments to date could not have been realized without the encouragement and goodwill of my family. To put into words all of the reasons I have to be thankful to them would result in a document much larger than this dissertation itself.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure *Page*

INTRODUCTION

CHARACTERIZATION OF THE HISTONE ACETYLTRANSFERASE (HAT) DOMAIN OF A BIFUNCTIONAL PROTEIN WITH ACTIVABLE *O*-GlcNAcase AND HAT ACTIVITIES

THE LOCATION AND CHARACTERIZATION OF THE *O*-GlcNAcase ACTIVE SITE

LIST OF FIGURES (Continued)

Figure *Page*

LIST OF ABBREVIATIONS

Asn	aspartate
Asp	aspartic acid
ATP	adenosine 5'-triphosphate
bNCOAT	bacterial nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase
CBP	cyclic 3',5' adenosine monophosphate-responsive element-binding protein binding protein
CREB	cyclic 3',5' adenosine monophosphate-responsive element-binding protein
Cys	cysteine
GlcNAc	*N*-acetylglucosamine
GLUT2	glucose transporter type 2
GST	glutathione *S*-transferase
kDa	kiloDalton
mNCOAT	mammalian nuclear and cytosolic *O*-GlcNAcase and acetyltransferase
NCOAT	nuclear and cytosolic *O*-GlcNAcase and acetyltransferase
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
O-GlcNAc	*O*-linked *N*-acetylglucosamine
O-GlcNAcase	β-*N*-acetylglucosaminidase
OGT	β-*N*-acetylglucosaminyltransferase

*p*NP-GlcNAc	*para*-nitrophenyl *N*-acetylglucosamine
SCOP	Structural Classification of Protein
Ser	serine
SMART	simple modular architecture research tool
STZ	streptozotocin
TPR	tetratricopeptide
UDP	uridine diphosphate

INTRODUCTION

Hexosamine Pathway and O-GlcNAcylation as a Post-translational Modification

The ability to adapt to nutrient flux is essential to the survival of cells and organisms on all levels. Glucose, the primary fuel for the generation of ATP energy, is one such nutrient that must be tightly monitored. In vertebrates, the pancreatic β-cell is the primary regulator of glucose metabolism on the organism level, and evidence suggests that the hexosamine biosynthesis pathway acts as a glucose sensing mechanism and can function to regulate glucose homeostasis within a single cell (1). Glucose can enter a β-cell via GLUT2 glucose transporters. Upon entering the cell, the enzyme hexokinase utilizes ATP to phosphorylate glucose on the hydroxyl group of the C6 carbon, converting it to glucose-6-phosphate. Glucose-6-phosphate is then converted by phosphoglucose isomerase to fructose-6-phosphate, where it awaits one of several pathways. The two principal fates are the breakdown of this substrate for energy in the glycolytic pathway and its conversion to glycogen for storage. Approximately 1 to 5% of fructose-6-phosphate enters the hexosamine pathway, where it is converted to glucosamine-6-phosphate by glutamine: fructose-6-phosphate amide transferase (2). Ultimately, the substrate UDP-*N*-acetylglucosamine (UDP-GlcNAc) is generated in this pathway and is thereafter used in the synthesis of glycoproteins (3). These sugar residues are commonly linked to proteins in two different ways. *N*-linked glycosylation occurs when attachment is made to the amide nitrogen of asparagine side chains. *N*-linked glycosylations contain a minimum of four sugars in addition to the terminal GlcNAc, and this modification is typically

implicated in the directing of proteins through the endoplasmic reticulum-Golgi-plasmalemma pathway.

First described on cell surface proteins in 1984, *O*-linked GlcNAcylation is a form of post-translational modification where a single sugar, the monosaccharide *N*-acetylglucosamine, is affixed via a beta linkage to the hydroxyl group of specific serine or threonine residues on target proteins (4, 5). In rare cases, *O*-linked *N*-acetyl-glucosamine (*O*-GlcNAc) linkage can occur on lysine side chains. While this modification has not been witnessed in bacteria, it is otherwise ubiquitous in all eukaryotic cells that have been studied. *O*-GlcNAcylation can occur on both extracellular and intracellular proteins and can be found on both cytosolic and nuclear members of the latter (4, 5). *O*-GlcNAc additions are reversible and highly dynamic; as suggested previously, the protein-bound and free GlcNAc levels are determined by metabolic changes (glucose availability) in the cell.

The enzyme that catalyzes the reaction between the target protein and UDP-GlcNAc, with UDP as the leaving group, is β-*N*-acetylglucosaminyltransferase (OGT) (6). OGT is a highly conserved protein throughout evolution (6, 7) and murine knock-out models have proven to be embryonic lethal, indicating that the ability to glycosylate proteins is essential for the completion of embryogenesis (8). OGT exists as a heterotrimer, with two large (100 kDa) peptides and one smaller (78 kDa) peptide (9). The relative levels of protein as well as that of its mRNA vary within different tissues. The highest concentrations of transcripts are detected in the pancreas, with a strikingly high concentration localized in the β-cells of the islets of Langerhans (10-12). Interestingly, the diabetogenic drug alloxan, a uridine analog, is an OGT inhibitor; the administration of this drug corresponds to a cellular increase in free *O*-GlcNAc (13).

The enzyme then responsible for the converse reaction, removing the *O*-linked GlcNAc from protein, is β-*N*-acetylglucosaminidase (*O*-GlcNAcase). While *O*-GlcNAcase can complex with other proteins, the active enzyme functions as a monomer, contrary to most characterized hexosaminidases, which are homomultimeric (14).

O-GlcNAcase is ubiquitously present at comparable levels in all tissues but has a range of activity that is tissue specific. "Strong" *O*-GlcNAcase activity can be detected in the spleen, lung, kidney, and liver; "weaker" activity is present in the heart, uterus, and spinal cord (4). Particularly enriched activity occurs in the skeletal muscle, brain, and pancreas, presumably to compensate for the high amounts of the corresponding OGT localized in these tissues.

Hundreds of proteins to date have been shown to be *O*-GlcNAc modified, with a great many more as yet untested that possess potential *O*-GlcNAcylation sites (15). The target proteins which are regulated via this modification function in a range of cellular processes, most of which are of fundamental importance to the viability of the cell. These include cell cycle control and cell division (16-19), protein import-export through the nuclear pores (16, 17), translation (20), transcription (20-24), proteosomal degradation (25), and others. Cellular protein-*O*-GlcNAc levels can be modulated by mitogens, intracellular signals, growth factors, or cell-cycle and developmental stages (26). The functional significance of this modification is not always fully understood in all cases; however, growing evidence suggests it can play a role in regulating protein stability (22, 27), regulating protein-protein interactions (16, 28, 29), governing subcellular localization (26), and activating or inactivating a target enzyme's catalysis (30, 31). In addition to its specific functions on proteins, *O*-glycosylation on serine and threonine residues can also compete directly with *O*-phosphorylation for the protein free-hydroxyl pool, and this es-

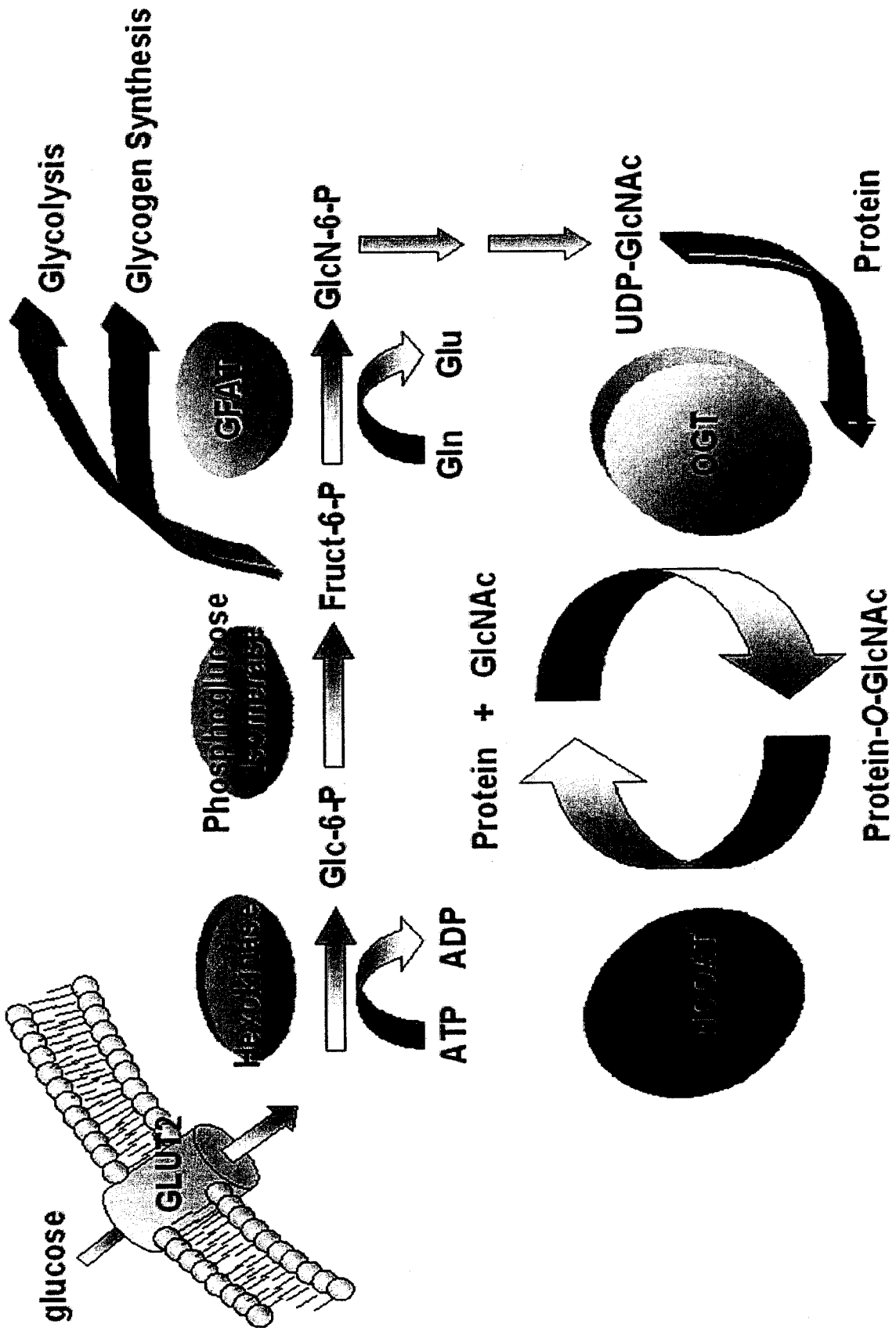
tablishes an additional level of regulatory control for these cellular functions (32). An outline of the process, from glucose entry into a cell to its incorporation and removal from proteins, is highlighted in Fig. 1.

The Current Understanding of O-GlcNAcase

Lysosomal hexosaminidases have been studied for decades in attempts to explicate the mechanisms involved in inheritable metabolic diseases such as Tay-Sachs and Sandhoff diseases (33). These hexosaminidases, naturally, have acidic pH optima, and they possess the ability to cleave both GlcNAc and *N*-acetylgalactosamine linkages. Outside the lysosome, there is only a single enzyme that can remove *O*-GlcNAc modifications on proteins. This enzyme is *O*-GlcNAcase. In humans, the single *O*-GlcNAcase gene is found at chromosomal locus 10q24 (34) near the insulin degrading enzyme; in rats, this gene is found at a genetically comparable region on chromosome 1 (34). Both loci have been mapped as regions of genetic susceptibility for diabetes (34, 35).

In contrast to the modulation of phosphate removal, which can be accounted for by approximately 150 different phosphatases, *O*-GlcNAcase, as stated above, is the only cellular protein capable of cleaving *O*-GlcNAc linkages on proteins in the nucleus and cytoplasm, a finding that implies its great importance in cell viability. *O*-GlcNAcase, like OGT, is highly conserved through evolution from yeast to humans and, as mentioned previously, is expressed in all tissues. As of 2000, the protein had no known sequence homologies, even to other glycosidases; neither did it have any of the reported protein motifs found in databases (36). PROSITE scan results have identified several possible modification sites, including phosphorylation sites and both *O*- and *N*-linked glycosylation sites. Potential myristylation and sumoylation sites also exist on the protein (37).

FIG. 1. **Hexosamine pathway and *O*-GlcNAcylation as a post-translational modification.** Glucose can enter a cell via one of four glucose transporters. After entering the cell it is rapidly converted to glucose-6-phosphate (*Glc-6-P*) by the enzyme hexokinase. Glucose-6-phosphate is then isomerized into fructose-6-phosphate (*Fruct-6-P*), where it will go on to be utilized for energy or stored for future use as glycogen through the process of glycogen synthesis. Approximately 5% of the fructose-6-phosphate enters the hexosamine pathway, where the enzyme glutamine: fructose-6-phosphate amide transferase (*GFAT*) converts it to glucosamine-6-phosphate (*GlcN-6-P*), in a step that also converts glutamine (*Gln*) to glutamate (*Glu*). Eventually, glucosamine-6-phosphate is converted to UDP-GlcNAc, which becomes the substrate for *O*-GlcNAc transferase (*OGT*), which will catalyze the addition of the GlcNAc to a target protein. Protein-*O*-GlcNAcylation is removed by the enzyme NCOAT, with its *O*-GlcNAcase activity.



Currently no evidence exists for the finding of any of these modifications. Other potential modification sites may still be present, but these have likewise remained unexplored.

In addition to its characterization as an *O*-GlcNAcase, the protein was also independently identified as a hyaluronidase using two separate assay systems (38). This observation is not surprising, since hyaluronic acid, the substrate for hyaluronidases, is a long molecule of repetitive GlcNAc β (1→4) glucuronic acid disaccharides. It is probable that *O*-GlcNAcase can function as a hyaluronidase more or less by default. The group responsible for these findings originally named the protein meningioma expressed antigen 5, due to its discovery as an immunogenic antigen found in serum libraries from patients with meningiomas (38). They also reported a splice variant of this protein (lacking the C-terminal 240 amino acids); this isoform retained its ability to cleave hyaluronic acid, therefore providing the first evidence that the active site for the enzyme lies in the N-terminal two-thirds of the protein (39). Independently, *O*-GlcNAcase was found to be a substrate for caspase-3 cleavage, resulting in a protein of two nearly equal halves. This cleavage was reported to have no effect on *O*-GlcNAcase activity (14) and narrows down the potential active site to the N-terminal half of the protein.

O-GlcNAcase as a Potential Histone Acetyltransferase

Histone acetylation is a tightly regulated and dynamic process. At the heart of this process are histone acetyltransferases, which can utilize acetyl coenzyme A to catalyze the addition of acetyl groups to target lysine residues on the tails of one or more of the core histones. The activity of these histone acetyltransferases as well as those for a large and growing number of transcription factors is regulated by post-translational modification, including *O*-GlcNAcylation. It should not be surprising, then, to find the enzymes

responsible for the addition and removal of these groups (OGT and *O*-GlcNAcase) to be in close association with these factors, either transiently or in complex. Indeed, OGT has recently been described to exist in histone deacetylase-containing corepressor complexes (21). Independently, OGT has also been found to associate very tightly with *O*-GlcNAcase in pull-down assays (unpublished results from our laboratory). The possibility therefore exists that *O*-GlcNAcase can localize to transcriptional complexes on histones.

Interestingly, when subjected to SMART computer analysis, *O*-GlcNAcase has been found to possess, in its C terminus (amino acids 772-900), a sequence with domain composition similar to that of several different acetyltransferases (ATs). Upon closer examination, secondary structure predictions reveal that *O*-GlcNAcases contain the signature pattern of secondary structures that appear at the active site of these ATs (eight of the ATs with similar composition have been crystallized and their secondary structural motifs determined) (40-47). The pattern consists of four helices (referred to as motifs A-D) that form the catalytic pocket. While the primary sequences within each motif vary greatly and while the lengths and sequences of the intervening loops do as well, the overall *secondary* structural pattern is conserved (41, 48). Such variations cause many potential ATs to be missed in database searches, since even in the catalytic core, where sequences are commonly similar among family members, closely related ATs will typically only retain only as much as 10% sequence homology and no more than 50% sequence similarity (41).

Finally, the reactions of those ATs with structural similarity to *O*-GlcNAcase have a conserved acid-base catalytic mechanism; however, an exception exists for Esa1, which has recently been proposed to proceed via a "ping-pong" mechanism involving a

self-acetylated intermediate (49). It has been demonstrated in these enzymes that a general base (aspartate or glutamate) must deprotonate the lysine amino group (which due to its high pK_a enters the enzyme in its protonated state) before acetylation can occur. The acetyl group from the reactive thioester of acetyl-coenzyme A can then be transferred to the ε-amino group of the substrate lysine (50). A general acid (with one exception, a tyrosine, among the nine examined) then makes a direct hydrogen bond to the coenzyme A sulfur atom, first making the carbonyl more electrophilic and then donating the proton to facilitate the departure of the thiolate leaving group (40). The locations of these two critical residues vary within individual sequences but superimpose to the same locations within the aforementioned tertiary and secondary structural motifs (41, 48). When various *O*-GlcNAcases are aligned, among the few invariant residues in the C terminus are two aspartates and a tyrosine, located in comparable positions when secondary structure prediction methods are used. Other bifunctional ATs have been characterized recently (51), and these findings make the AT potential of *O*-GlcNAcase an even more appealing possibility.

The O-GlcNAcase Active Site and Glycosidase Catalysis

In order to hydrolyze the particularly stable glycosidic bond, glycosidases have evolved a highly proficient and generally conserved mechanism for catalysis. There exist over a hundred different families of glycoside hydrolases (classified by their sequence and core sugar substrate), within which lie incredible structural diversity. However, many active site features are retained ubiquitously among subfamilies (52). Therefore, these enzymes share a mode of action; however, some may "retain" or "invert" the stereochemical conformation of the departing sugar in the process (52, 53).

Through the use of inhibitors as well as mutagenesis and crystal structure techniques, the mechanisms of these glycosidases have been elucidated. The mechanism involves two carbonyl groups suitably placed in the active site to allow both the substrate and a water molecule to bind between them. The first functions as the general acid catalyst, donating a proton to the oxygen atom on the anomeric carbon, C1. The second acts as a nucleophile, allowing the formation of a covalent glycosyl-enzyme intermediate. The first carbonyl, now acting as a general base, deprotonates the incoming water molecule; the resulting hydroxyl, a strong nucleophile, then attacks the C1 carbon, causing the release of the sugar (52, 53).

A variation of this general mechanism has been demonstrated in families 18, 20, and 56 glycosidases. These families include the chitinases, hyaluronidases, and hexosaminidases, which share *N*-acetylglucosamine as their common core substrate. Due to the character of the substrate, the 2-acetamido group itself can act as the reaction nucleophile, generating a bicyclic, high-energy oxazolinium ion intermediate as its transition state. Therefore, the enzymatic reactions of these glycosidases proceed via a substrate-assisted mechanism (54, 55).

An assumption can be made that such a mechanism holds true for *O*-GlcNAcase since, like these other enzyme families, it recognizes GlcNAc as its core substrate. This information can be utilized to investigate the potential location of the active site for *O*-GlcNAcase activity. Evidence is being accumulated, again through mutagenesis and crystallographic techniques, suggesting that the active site of chitinase and hyaluronidase family members contains a "catalytic doublet" consisting of the amino acid motif D-ψ-D/E, where ψ is any hydrophobic residue (56-58). The second carbonyl in the doublet, either an aspartic acid or a glutamic acid, is the general acid residue and both protonates

the oxygen at the reaction center and deprotonates the incoming water molecule. The first carbonyl in the series (always an aspartic acid) functions to keep the general acid protonated. The first carbonyl is also necessary to position the *N*-acetyl side chain of the GlcNAc to allow this *N*-acetyl group to act as a nucleophile. The second carbonyl also stabilizes the transition state through hydrogen bonding to the nitrogen atom of the bicyclic intermediate. Together, these two carbonyls lie in close enough proximity (being separated by a single amino acid, they reside on the same plane of the two-dimensional structure) to also force the resulting hydroxyl ion to attack the reaction center to complete the reaction (56-58).

Examination of the other families 18, 20, and 56 members reveals that they commonly possess this "catalytic doublet"; the roles of these residues have been experimentally proven in a large number of these enzymes. A look into the sequences of several species' *O*-GlcNAcase likewise reveals the presence of the shared D-ψ-D/E catalytic motif. These residues in *O*-GlcNAcase (residues 175-177) are the most likely candidates for investigation into the amino acids necessary for *O*-GlcNAcase enzymatic activity.

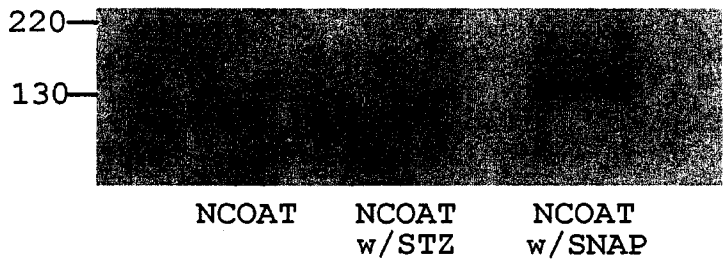
Mechanism of O-GlcNAcase Inhibition by Streptozotocin

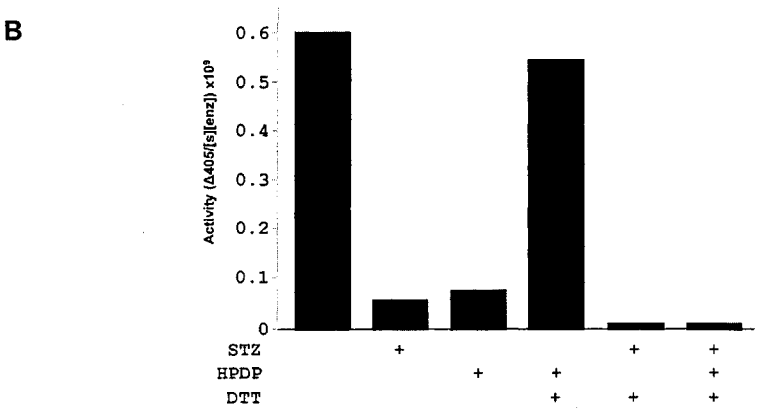
Streptozotocin (STZ) is an *N*-acetylglucosamine analog with an *N*-methyl-*N*-nitrosourea group linked to the C-2 carbon of the sugar ring. This drug, which can selectively enter cells such as a β-cell that express the GLUT2 glucose transporter, is routinely used to acutely induce pancreatic β-cell apoptosis when models of diabetes are desired in laboratory animals. It has also been shown to be an effective anticancer agent for the treatment of pancreatic islet cell tumors (59).

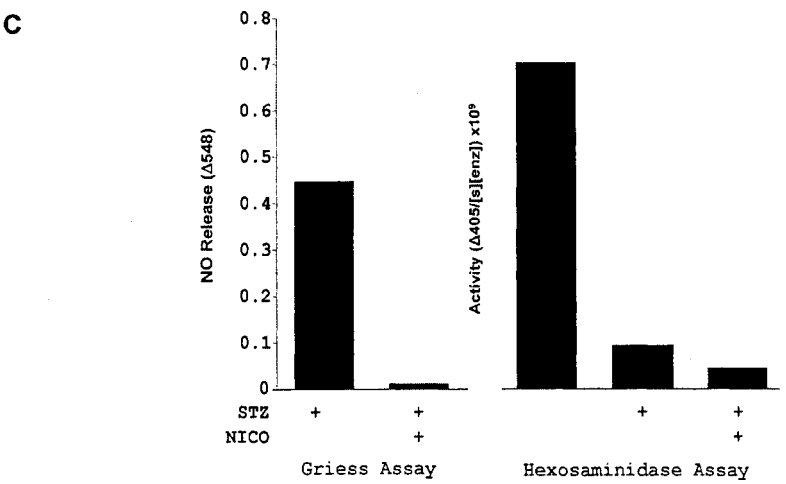
STZ has been reported to be a genotoxic methylating reagent (60) (*i.e.* a member of a group of neoplastic drugs named alkylnitrosoureas). STZ has the ability to methylate DNA, producing significant levels of N^7-methylguanine, N^7-methyladenine, N^3-methyladenine and O^6-methylguanine. Such alkylation results in double-stranded DNA breaks, the accumulation and inefficient repair of which can result in cell death (61). In addition, STZ is known to be an NO donor, and this has also been proposed to be responsible for toxicity to β-cells through oxidative damage to numerous cell processes (62). Independently, it has been demonstrated that STZ can specifically inhibit *O*-GlcNAcase *in vitro*. Furthermore, STZ inhibits *O*-GlcNAcase activity in both a dose- and time-dependent manner. This effect would seemingly be greatly detrimental to a β-cell, which, as mentioned, is heavily reliant on a sensitive *O*-GlcNAcase to counteract the high OGT levels present in these cells. In support of this idea, *in vivo* data have shown a rise in the level of *O*-glycosylated proteins in pancreatic β-cells as well as β-cell death following STZ treatment, an observation consistent with an impairment of the enzyme responsible for their removal from proteins (3).

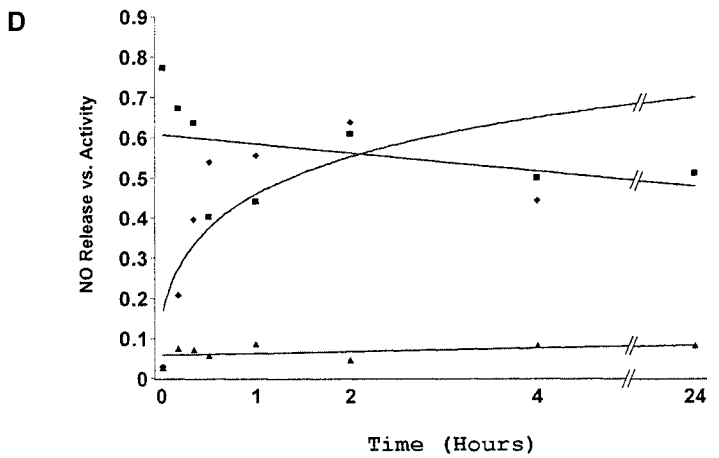
The drug *N*-ethylmaleimide, a sulfhydril-specific alkylating reagent, has also been demonstrated to inhibit the enzymatic activity of *O*-GlcNAcase, a finding that is indicative of a free cysteine in or around the active site of the enzyme (4). Since STZ is an NO donor and since select NO donors have been shown to target and modify free cysteine thiols on proteins, in many cases inhibiting the activity of the enzyme (63), one rationale as to why STZ can impair *O*-GlcNAcase is through this *S*-nitrosation mechanism. However, when an *S*-nitrosation assay (63) is performed, the promiscuous *S*-nitrosating reagent *S*-nitrosoacetylpenicillamine can label free cysteines in the protein; on the other hand, STZ-treated samples resemble those of untreated *O*-GlcNAcase (Fig. 2*A*). Second,

FIG. 2. **The inhibition of *O*-GlcNAcase activity by streptozotocin is NO independent.** *A*, *S*-nitrosation assay. *O*-GlcNAcase (*NCOAT*) was left untreated, treated with STZ, or treated with the *S*-nitrosating reagent *S*-nitrosoacetylpenicillamine (*SNAP*). *B*, NCOAT was treated with and inhibited by STZ or biotinylated *N*-(hexyl)-3'-(2'-pyridyldithio) propionamide (*HPDP*); (*lanes 2* and *3*). The HPDP treatment is reversible with 50 mM dithiothreitol (*DTT*); (*lane 4*), while STZ treatment was not (*lane 5*). Blocking the free cysteine with HPDP and then treating with STZ, followed by removal of the bound HPDP, still resulted in *O*-GlcNAcase inhibition (*lane 6*). *C*, NO release when STZ is dissolved in water, as measured by a change in spectrophotometric absorbance when Griess Reagent binds to liberated NO (*lane 1*). Free NO is quenched by inclusion of nicotinamide (*NICO*) in copper sulfate buffer (*lane 2*). Even after nicotinamide quenching of NO, STZ can still inhibit *O*-GlcNAcase activity (*lanes 4* and *5*). *D*, NO release over time, as measured by Griess assay, is indicated by ◆. *O*-GlcNAcase activity over time when left at room temperature is indicated by ■. *O*-GlcNAcase activity after treatment with STZ that has been left at room temperature for the indicated times is shown with ▲.









S-nitrosyl-cysteine modifications are readily reversible with high concentrations of dithiothreitol. When *O*-GlcNAcase is treated with STZ, the inhibition is not reversible with dithiothreitol (Fig. 2*B*). In addition, the free cysteine in the active site could not be protected with a reversible, biotinylated sulfhydril reagent, *N*-(hexyl)-3'-(2'-pyridyldithio) propionamide which, like *N*-ethylmaleimide, can inhibit activity. Furthermore, the inhibition of *O*-GlcNAcase activity by STZ is completely NO independent. When STZ is dissolved in solution, as mentioned above, it rapidly releases its NO, an observation that can be monitored by Griess assay (63) (Fig. 2*C*). The release of NO can be scavenged by the addition of nicotinamide and copper sulfate (Fig. 2*C*). Even after the NO has been removed in such a fashion, STZ can still inhibit *O*-GlcNAcase activity (Fig. 2*C*). Finally, after monitoring the NO release by STZ in this same manner, one can still observe that, even after all the NO has been released, the resulting compound can still fully inhibit *O*-GlcNAcase activity (Fig. 2*D*). Given these data, the means by which STZ can inhibit the *O*-GlcNAcase activity must be achieved by some mechanism unrelated to NO.

CHARACTERIZATION OF THE HISTONE ACETYLTRANSFERASE (HAT) DOMAIN OF A BIFUNCTIONAL PROTEIN WITH ACTIVABLE *O*-GlcNAcase AND HAT ACTIVITIES

by

CLIFFORD TOLEMAN, ANDREW J. PATERSON, THOMAS R. WHISENHUNT, AND JEFFREY E. KUDLOW

Journal of Biological Chemistry 279 (51): 53665-53673, 2004

Format adapted and errata corrected for dissertation

Abstract

Histones and transcription factors are regulated by a number of post-translational modifications that in turn regulate the transcriptional activity of genes. These modifications occur in large, multisubunit complexes. We have reported previously that mSin3A can recruit *O*-GlcNAc transferase (OGT) along with histone deacetylase into such a corepressor complex. This physical association allows OGT to act cooperatively with histone deacetylation in gene repression by catalyzing the *O*-GlcNAc modification on specific transcription factors to inhibit their activity. For rapid, reversible gene regulation, the enzymes responsible for the converse reactions must be present. Here, we report that *O*-GlcNAcase, which is responsible for the removal of *O*-GlcNAc additions on nuclear and cytosolic proteins, possesses intrinsic histone acetyltransferase (HAT) activity *in vitro*. Free as well as reconstituted nucleosomal histones are substrates of this bifunctional enzyme. This protein, now termed NCOAT (nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase) has a typical HAT domain that has both active and inactive states. This finding demonstrates that NCOAT may be regulated to reduce the state of glycosylation of transcriptional activators while increasing the acetylation of histones to allow for the concerted activation of eukaryotic gene transcription.

Introduction

The genomes of eukaryotes are assembled into the highly condensed structure of chromatin. Chromatin is composed of repeating units of nucleosomes that are comprised

of DNA coiled around an octameric particle consisting of two molecules of each core histone, H2A, H2B, H3, and H4 (1, 2). The acetylation state of these histone protein tails has for some time been known to be correlated with gene expression, where transcriptionally competent loci are hyperacetylated and silenced loci hypoacetylated (3-5). Histone acetyltransferases (HATs) compose a superfamily of enzymes broken into subfamilies based on sequence similarities, active site size, and the presence or absence of other protein domains (6). In each HAT protein there exists a structurally homologous region that composes the active site and includes four universally present motifs, designated A-D, that form the scaffold of the catalytic core (6). Recently published three-dimensional structures of various acetyltransferases (ATs) has allowed for a comprehensive view of the roles played by each motif as well as the roles of many amino acids structurally conserved therein (6-8). HATs and histone deacetylases (HDACs) act competitively within large multiprotein complexes that recruit them to their nucleosomal substrates on DNA and give them the ability to contribute to the activation or repression of gene expression, respectively (9). The covalent addition of the monosaccharide, *N*-acetylglucosamine (GlcNAc) to serine or threonine residues of proteins is catalyzed by the enzyme, *O*-GlcNAc transferase (OGT) that is encoded by a single gene (10). Recently, we reported that the corepressor, mSin3A, known to recruit HDAC (9, 11), also recruits OGT via its TPR domains to specific genes (12). OGT can thereby contribute along with HDAC to the repression of gene expression through addition of *O*-GlcNAc modifications on transcriptional activators, inhibiting their activity. Such inhibitory effects have been witnessed on the transactivation domain of Sp1 (13), the C-terminal tail of RNA polymerase II (14) and the $TAF_{II}130$ recruitment domain of (15) among others. It has also been documented

that the repression of genes is associated with the hyperglycosylation of the proteins bound to their promoters (12).

Many genes, including those responsive to variable hormone levels, can be activated or repressed to maintain homeostasis. For the repressed gene state to be reversible, the complex responsible for the repressed state must be exchanged with a complex that allows gene activation (16, 17). As part of the activation process, the action of the enzymes residing in the repression complex would need to be removed, including the inhibitory modification of many transcriptional activators by *O*-GlcNAc. The enzyme *O*-GlcNAcase, which is the only enzyme capable of catalyzing the removal of these regulatory *O*-GlcNAc modifications on proteins in the nucleus and cytosol, must necessarily then play a role in cooperation with HATs in the activation of genes, just as OGT plays its role along with HDAC in repression (12). Of interest in this regard, it has recently been reported that the C terminus of *O*-GlcNAcase contains a domain with similar composition to eleven different AT active sites, as predicted by SMART computer analysis (18). While it has been argued that the *O*-GlcNAcase may have evolved from an AT (19), this protein could potentially play a dual role in the reversibility of corepression by removing *O*-GlcNAc modification from activators while also adding acetyl groups to histones, allowing a target gene to be expressed. Here we report that *O*-GlcNAcase does in fact possess acetyltransferase activity *in vitro* for a synthetic histone substrate tail as well as for free core histones and reconstituted oligonucleosome substrates. However, the HAT activity is regulated and can only be observed when the enzyme is expressed in mammalian cells. The active site for this domain lies in the C terminus of the protein, where it resembles other acetyltransferases both structurally and in catalytic mechanism and has complete functional distinction from the N-terminal *O*-GlcNAcase domain. Because the en-

zyme is bifunctional with two important enzymatic domains, we have renamed it nuclear cytoplasmic *O*-GlcNAcase and acetyl transferase (NCOAT).

Experimental Procedures

Plasmids and Recombinant NCOAT—A pUC118-pTM hybrid expression vector was designed by removing a 2-kb fragment of pTM containing T7 and glutathione-*S*-transferase (GST) elements and inserting it into the pUC118 MCS. This construct was used for GST peptide expression. A PCR product containing full-length mouse NCOAT was then cloned into the pUC118-pTM hybrid for N-terminal GST fusions or into pcDNA 3.1 (Invitrogen) for GAL4 fusions. Mouse NCOAT was also inserted into a pGEX vector (Amersham Biosciences) for expression in bacteria (bNCOAT). All recombinant DNA manipulations were performed by standard procedures (20).

Splice Variant Identification and Cloning—NCOATs were cloned by reverse transcription-PCR from Goto Kakazaki or Sprague-Dawley rat brains. The full-length and variant isoforms were cloned into pBluescript and sequenced. The full-length rat NCOAT was very similar to the mouse NCOAT at the nucleotide level, and therefore the 3'-end of the mouse cDNA was replaced with the 3'-end of the respective Goto Kakazaki or Sprague-Dawley rat cDNA to create chimeric NCOAT molecules with the missing exons. These constructs were cloned into the pUC-pTM vector.

Truncations and Site-directed Mutagenesis—Full-length pcDNA-NCOAT was digested with EcoRI to excise a fragment that consisted of nucleotides 1-1747 of NCOAT. This digested plasmid was religated onto itself for use as an expression vector

for a GAL4 fusion to the NCOAT acetyltransferase domain (NCOAT nucleotides 1748-2771). The AT domain fusion was also placed into the pGEX vector for bacterial expression. For the expression of NCOAT with a deletion of its AT domain, the pcDNA-GAL4-NCOAT vector was transformed into DM-1 cells (Invitrogen) then the plasmid was digested with PmeI and PflMI. The resulting vector, missing NCOAT nucleotides 2086-2771 was blunt-end-ligated. For point mutations, the pUC118-pTM-NCOAT construct was modified by site-directed mutagenesis using a four-primer cassette strategy to introduce the substitution (mutated residues underlined) and sites for selection (bold). Standard PCR amplifications were performed with the following oligonucleotides: for D853N, 5'-**GATAT***C*CATAAAAAAGTGACTGACCCGAGTGTTGCC-3'; for D884N, 5'-CTGTGAAGTAAGACCAGATAATAAAA**GGATCC**CTGG-3'; for Y891F, 5'-GGATTCTGGAATTTT*T*CAGC**AAGCTT**GGCTG-3'. Flanking oligonucleotides were as follows: 5'-CATGGATCCAGTCAACCTGACCTTATTGG-3' (upstream) and 5'-ACGCCAAGCTCGAAATTAACC-3' (downstream). The products were separated by electrophoresis, extracted, and purified by QIAquick extraction kit (Qiagen). The purified products were digested with Bsu36I and KpnI and ligated into pUC118-pTM. Each ligation mixture was transformed into DH5α cells (Invitrogen) and screened for the selection sites.

Cell Culture and Bacterial Expression—BSC-40 cells (ATCC) were maintained in Dulbecco's modified Eagle's medium (Invitrogen) with 10% fetal bovine serum (Atlanta Biologicals) with 0.2 mg/ml penicillin and 0.4 mg/ml gentamicin according to standard procedures. 20 μg of plasmid DNA were transfected by electroporation, incubated overnight, and then infected with VTF-7 vaccinia virus. Cells were harvested after over-

night incubation in 0.5% Nonidet P-40 lysis buffer containing 50 mM Tris-HCl, 0.5 M NaCl, 20% glycerol (v/w), 5 mM $MgCl_2$, 0.2 mM EDTA, 5 mM dithiothreitol, pH 8.0. Bacterial protein expression was achieved by transforming a 500-ml culture of BL-21 *Escherichia coli* in Terrific Broth (BD Biosciences) with pGEX-mNCOAT and expression induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside (final). After 1 h, 1 mM benzamidine (final) was added, and incubation was continued another 2 h. Cultured cells were spun down and resuspended in Nonidet P-40 lysis buffer with lysozyme. Lysed cells were sonicated and spun down, and supernatant was collected in 500-µl aliquots and stored at -80 °C until use.

Affinity Purification and Immunoprecipitation—GST or GST fusion proteins were purified from lysates using glutathione-Sepharose (Amersham Biosciences) and stringently washed three times with 5 volumes of radioimmune precipitation assay buffer. Proteins were eluted for filter binding assays and in assays with oligonucleosomes in gel slices. The fusion proteins were left resin-bound for pretreatment with streptozotocin (STZ) or when the proteins were resolved on SDS-PAGE. Elutions were carried out according to manufacturer's protocol. Immunoprecipitation of GAL4 and GAL4 fusion proteins were performed by preclearing lysates with 20 µl of a 50:50 protein A:protein G-Sepharose bead mixture (Amersham Biosciences). Precleared lysates were then incubated with an N-19 GAL4 DNA binding domain antibody (Santa Cruz Biotechnology) for 1 h at 4 °C and then with a 100-µl mixture of 50:50 bead mixture for an additional 2 h. Immunoprecipitates were collected by centrifugation and washed. An aliquot of each purified enzyme was quantitated by Bio-Rad protein assay to ensure equivalent amounts of protein were used in each assay and were comparable in concentration to the positive

controls. Bacterial *O*-GlcNAcase proteins were also run on SDS-PAGE and Western blotted with α-GST according to standard protocols. Resin-bound incubations with mammalian whole cell lysates were carried out for 1 h at 4 °C and then washed.

Oligonucleosome Reconstitution—Histone octamers were assembled as described by Dyer *et al.* (21) using unfractionated type IIA calf thymus whole histones (Sigma). Briefly, these histones were dissolved in 4 ml of unfolding buffer (6 M guanidine HCl) at a concentration of 2 mg/ml for 2 h at room temperature then dialyzed against 3 changes of refolding buffer (2 M NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM dithiothreitol) at 4 °C for 6 h, overnight, and 6 h. Precipitated protein was removed, and the histone mixture was concentrated in a YM-10 protein concentrator (Amicon) down to 1 ml. Octamer reconstitution was confirmed by non-denaturing gel electrophoresis. Nucleosomal DNA was prepared using 186-bp HeLa cell α-satellite DNA as described by Tanaka *et al.* (22). Finally, oligonucleosomes were reconstituted using the standard salt dialysis method as described by Dyer *et al.* (21) using equal concentrations of the histone mixture and DNA. Reconstitution was confirmed by mobility shift assay using 15 μg of nucleosome particles in both 0.7% agarose gel and 5% polyacrylamide gel as described elsewhere (21, 22). Shifted bands were excised from these gels to ensure only nucleosomal histones were used. These gel fragments were minced and diluted in HAT buffer for 1 h. These solutions were used in the oligonucleosomal HAT assays.

HAT and O-GlcNAcase Assays—Filter binding assays were performed as described elsewhere (23, 24). 60 μl of purified enzyme were added to 20 μg of synthetic histone H4 peptide (Upstate Biotechnology), 100 μl of 5 mg/ml bovine serum albumin, 29

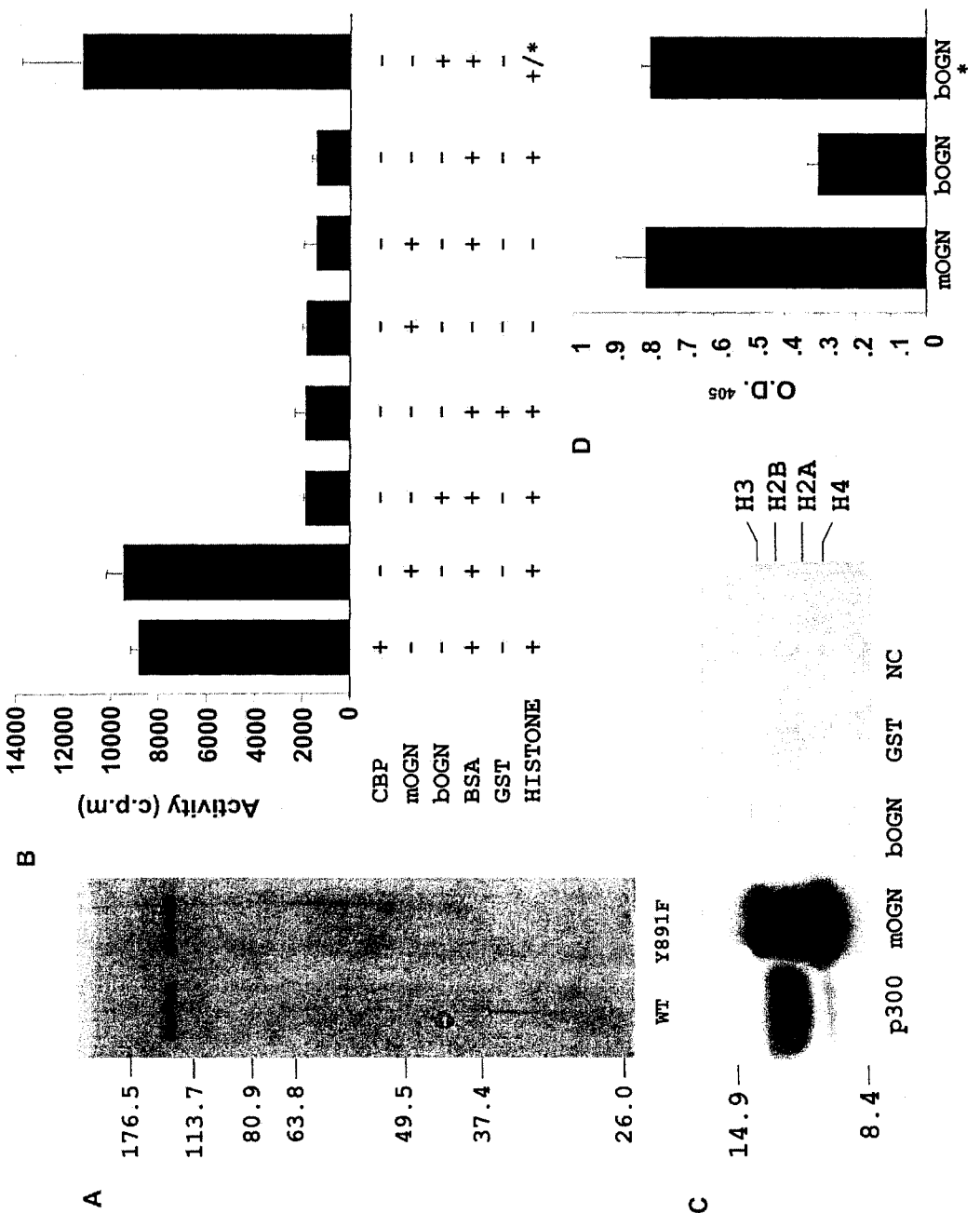
μl of 10x HAT buffer (50 mM Tris-HCl, pH 8.0, 10% glycerol, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 10 mM sodium butyrate), and 3 μl of [^{3}H]acetyl co-enzyme A (4.7 mCi/mmol). Reactions were carried out for 1 h at room temperature and then spotted onto 0.45-μm nitrocellulose membranes (Whatman), dried, and unincorporated [^{3}H] removed with three 15-min washes using 50 mM Na_2CO_3, pH 9.2. Dried filters were then suspended in scintillation fluid, and acetyl incorporation was measured by scintillation counter. HAT assays to be resolved on SDS-PAGE were performed as above except purified proteins were left resin-bound in 120 μl of 1x HAT buffer, and 2 μl of [^{14}C]acetyl-CoA (50 mCi/mmol) were used. Calf thymus histones were reconstituted in 1x HAT buffer and 20 μg used per reaction. 40 μl of each reaction were run on a 15% SDS-PAGE and radiolabeling monitored by autoradiography. Where noted, CBP (Pierce) or p300 HAT domain (Upstate Biotechnology) were used as positive controls. Streptozotocin (Sigma) was prepared as a 1 M fresh solution in 1x HAT buffer. 6 μl were added to each 60-μl solution of resin-bound protein and incubated for 1 h at 37 °C. STZ was removed with three washes using excess 1x HAT buffer before addition of HAT assay components. STZ treatment before hexosaminidase assays was performed as described elsewhere (25). HAT assays on oligonucleosome substrates were performed as described by Tse *et al.* (4) with the following exceptions: minced gel fragments (15 μg) were incubated with eluted *O*-GlcNAcase in the presence of either 10 μM acetyl-CoA or 2 μl of [^{14}C]acetyl-CoA (50 mCi/mmol) in assays to be resolved by Western blot or autoradiograph, respectively. Final assay volumes were brought up to 70 μl in 1x HAT buffer. *O*-GlcNAcase assays were carried out as described elsewhere (26). Caspase 3 (R&D Systems) reactions were carried out as described by Hart and co-workers (27).

Secondary Structure Prediction—Structural Classification of Protein analysis was used to reveal the presence of the mixed beta sheet motif associated with the acyl-CoA *N*-acyltransferase (Nat) superfamily of proteins within each NCOAT sequence indicated (19). This region was then subjected to secondary structure prediction by Jpred (28). The presence of each α-helix and β-sheet was confirmed individually using CHOFAS secondary structure prediction (version 2.0u61) (29).

Results

O-GlcNAcase Possesses Intrinsic Histone Acetyltransferase Activity in Vitro—To determine whether *O*-GlcNAcase possessed acetyltransferase activity, recombinant GST-tagged *O*-GlcNAcase was expressed in both mammalian BSC-40 cells and *E. coli* BL-21 cells and affinity-purified. The purity of the protein samples were determined by Coomassie-plus staining (Fig. 1*A*). GST-*O*-GlcNAcase was visible as a dominant band at 130 kDa. There were other slightly detectable products copurified, most of which were determined by Western blot to be truncated GST-*O*-GlcNAcase products. Eluted protein was incubated with a synthetic histone H4 tail and [^{3}H]acetyl-CoA for a filter binding assay or with core histones and [^{14}C]acetyl-CoA for resolution on 15% SDS-PAGE and detection by autoradiography. These results were directly compared with those for the global coactivators CBP or p300. The HAT activity, as measured by scintillation counting, of the *O*-GlcNAcase expressed in mammalian cells, but not that expressed in *E. coli*, was comparable with that obtained when using CBP. The activity of equimolar amounts of CBP and *O*-GlcNAcase were comparable on the histone H4 tail substrate (Fig. 1*B*) under dose and time conditions determined to be within the linear range for enzyme activity (data not shown). The counts obtained in this assay were determined to be only those for

FIG. 1. **Mammalian-expressed *O*-GlcNAcase has intrinsic HAT activity.** *A*, Coomassie staining confirming expression and purity of *O*-GlcNAcase samples used in subsequent assays. *Lane 2* shows that the same slightly detectable copurified products also coprecipitate with a catalytically dead mutant enzyme. *B*, filter binding assays were performed on *O*-GlcNAcase expressed in either mammalian (*mOGN*) or bacterial (*bOGN*) cells with a synthetic histone H4 tail and [^{3}H]acetyl-CoA and compared with CBP. Assays were performed in triplicate. The *asterisk* indicates preincubation with mammalian whole cell lysate. *C*, autoradiograph using core histones and [^{14}C]acetyl-CoA. Molecular mass markers in kDa are to the *left* of the gel. p300 was used as a positive control. Reactions in the absence of enzyme were used as a negative control (*NC*). In these reactions, *O*-GlcNAcase and bovine serum albumin showed levels of labeling similar to background (data not shown). *D*, *O*-GlcNAcase activity associated with mammalian *O*-GlcNAcase (*mOGN*), bacterial *O*-GlcNAcase (*bOGN*), and bacterial *O*-GlcNAcase treated with mammalian whole cell lysate (indicated by the *asterisk*). Units on the *y* axis, as measured by spectrophotometric absorbance at 405 nm, represent the amount of *para*-nitrophenol cleavage when 1 mmol of *para*-nitrophenyl-GlcNAc is used in the presence of 1 μmol of enzyme.



^{3}H incorporation into the histone substrate, since *O*-GlcNAcase could neither acetylate itself (Fig. 1*B, lane 5*) nor the bovine serum albumin in the reaction mixture (*lane 6*), both of which would adhere as equally as the histone substrate to the nitrocellulose membrane. The contingency that a separate, distinct HAT, which may copurify with *O*-GlcNAcase, may be responsible for this activity seems unlikely, given the relative purity determined in Fig. 1*A*. However, to fully ensure that the trace amount of copurified products observable in Fig. 1*A* were not responsible for HAT activity, we expressed a protein containing a single point mutation that abolishes HAT activity and ran this sample side by side with wild type active *O*-GlcNAcase. As shown in Fig. 1*A*, the copurified bands were identical in each case, however the $Tyr^{891} \rightarrow$Phe mutation, a well characterized residue involved in catalysis in other HATs (see below), resulted in an absence of HAT activity. In no known cases has this residue been shown to be involved in protein-protein binding. Because we were in the linear range for enzyme activity, the specific activity can be determined, which would necessitate that the observed small quantities of impurity would have a specific activity much higher, by orders of magnitude, than CBP to accumulate acetates on the histone substrate. Conversely, the *O*-GlcNAcase protein itself, which possesses similar specific activity to CBP and shares a HAT motif with the protein (see below) must be the source of the HAT activity. The bifunctionality of the protein as an *O*-GlcNAcase (30) and an acetyltransferase prompted us to change the name of the protein to NCOAT, giving no priority to either activity.

Similar results between NCOAT and p300/CBP are also seen when visualized by SDS-PAGE using free histones as substrates. Labeling of all four core histones is detectable when NCOAT is used in this assay just as seen for p300 (Fig. 1*E*). NCOAT actually appears to have a greater efficiency in labeling the histones when compared with p300 in

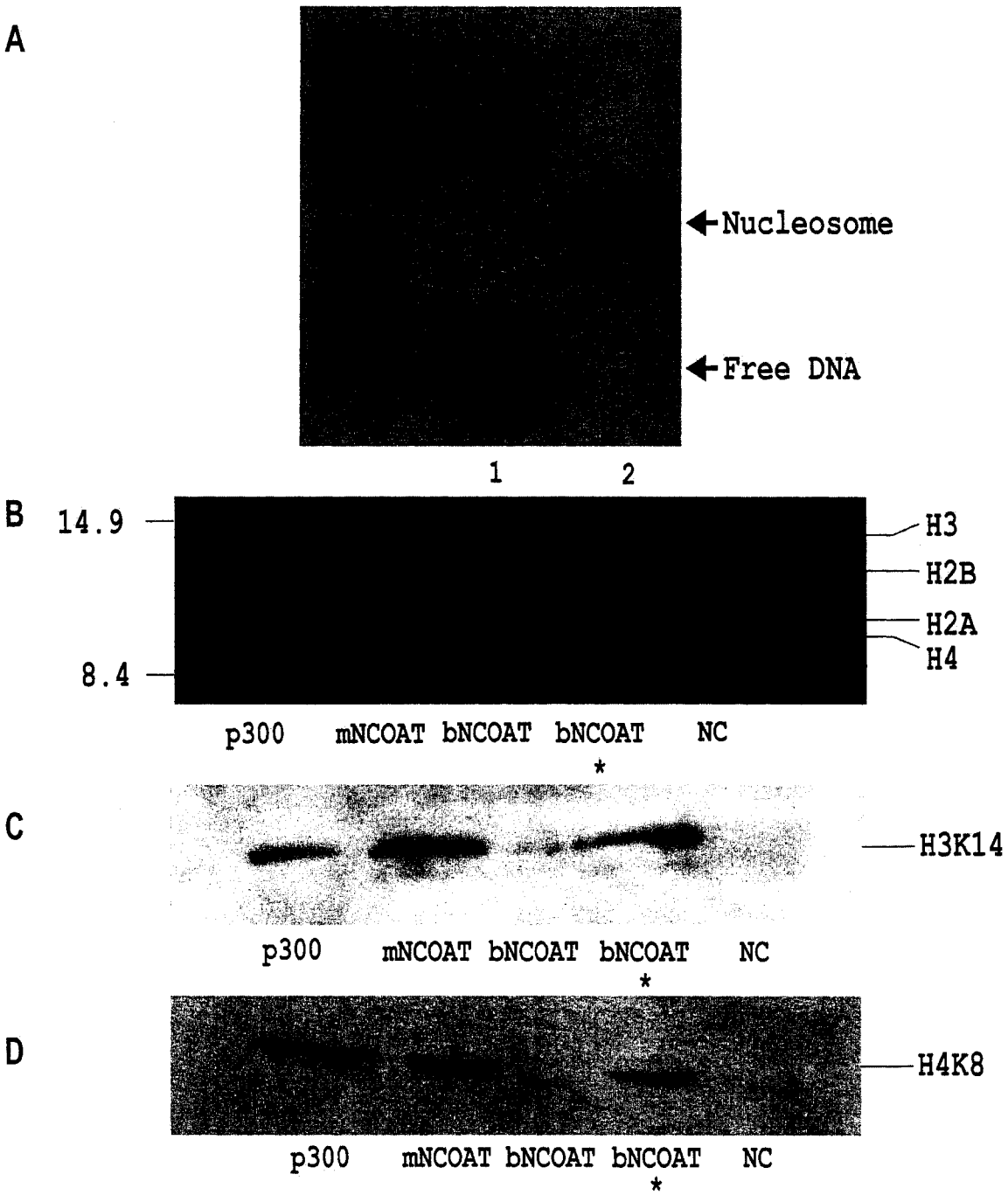
this assay. This greater activity may perhaps be due to the ability of NCOAT to acetylate a greater number of sites on the histones or can be a function of the fact that only the HAT domain of p300 was used in these assays, whereas the full-length p300 may possess greater activity. No HAT activity was measured when using NCOAT expressed in bacteria. The results in either assay, when NCOAT was expressed in bacteria, were similar to those seen when histone substrate was incubated with either GST or in the absence of any enzyme, a finding that may be indicative of why *O*-GlcNAcase/NCOAT may not have been discovered to have HAT activity to date.

These results indicate that NCOAT, when expressed in a mammalian system, possesses acetyltransferase activity for both a synthetic histone substrate and for free core histones. The total lack of activity of the bacterially expressed NCOAT is the likely result of the absence of an as yet uncharacterized post-translational modification(s) that is present on the material expressed in mammalian cells. This likelihood for the mammalian enzyme can be demonstrated when the bacterially expressed enzyme is incubated with a mammalian whole cell lysate prior to the activity assay. After such treatment on resin followed by washing in radioimmune precipitation assay buffer, full HAT activity can be observed in this protein (Fig. 1*B, lane 8*). The hexosaminidase activity of the enzyme was also considerably less when the full-length protein was expressed in bacteria compared with mammalian cells, although it was not completely absent as seen in the HAT assay. The hexosaminidase activity of the bacterial protein similarly increased to the level of mammalian-expressed protein following incubation with mammalian cell extract (Fig. 1*D*).

NCOAT Can Acetylate Oligonucleosomal Substrates—HAT proteins are classically described as being either type A, those that acetylate histones within chromatin, or

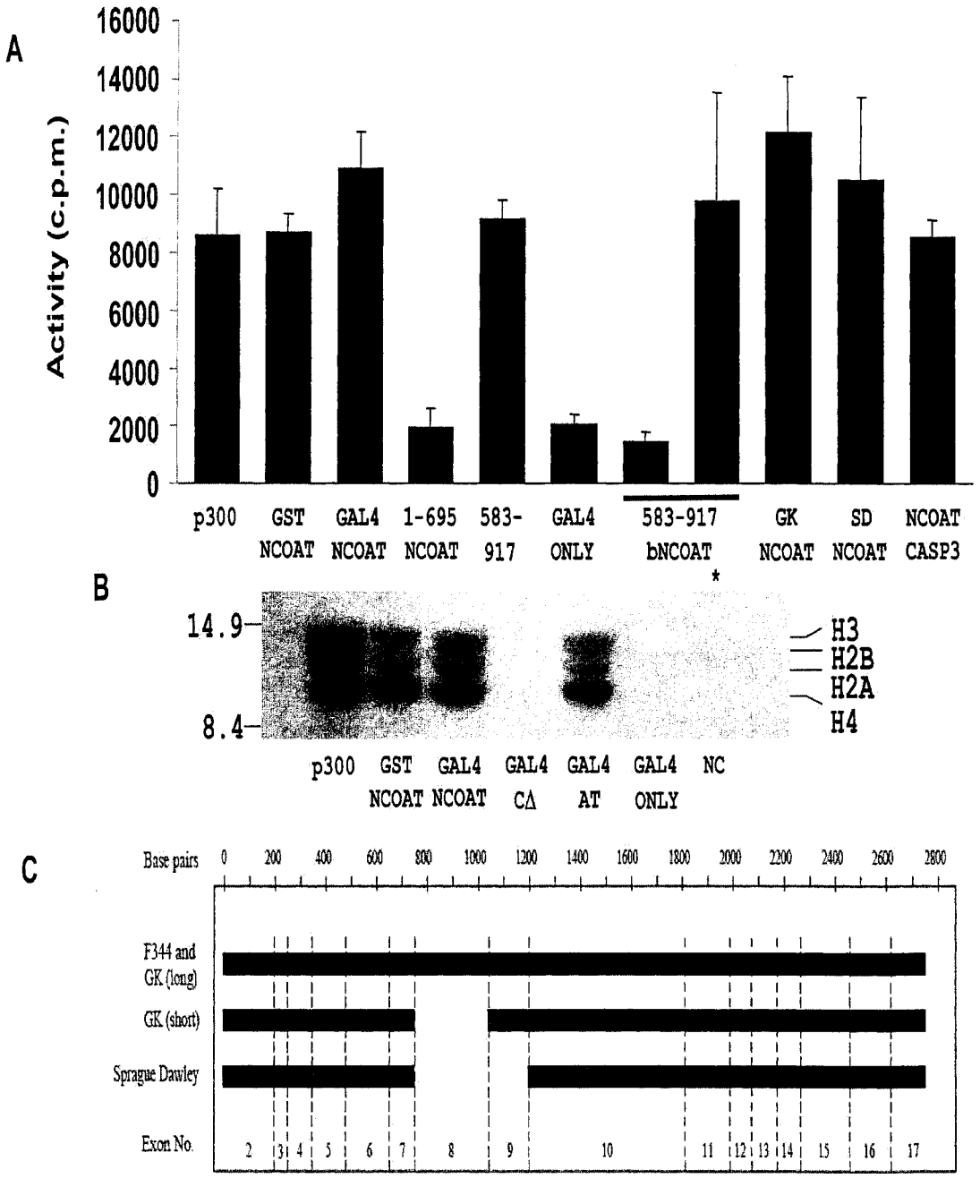
type B, those that acetylate free histones within the cytoplasm. For NCOAT to act as theorized on gene transcription directly, it would need to act as a type A HAT. As described above, NCOAT, like other type B HATs, can acetylate free core histones. We next tested its ability to act upon nucleosomal histones. For this, histones were reconstituted into oligonucleosome arrays for use as substrates in the HAT assays. Successful formation of these substrates was determined by gel shift assay (Fig. 2*A*). The shifted oligonucleosomes were purified from these gels and incubated with p300 or NCOAT in the HAT assay and run on 15% SDS-PAGE. NCOAT acetylated all four core histones even when bound by double-stranded DNA in the context of reconstituted nucleosomes (Fig. 2*B*). Again, mNCOAT and bNCOAT incubated with mammalian whole cell extract were active, while bNCOAT not preincubated with extract was not active. These HAT assays were also repeated using 10 μM unlabeled acetyl-CoA. Western blots of the subsequent gels were then probed with antibodies for specific acetylated lysine residues to determine some of the modified residues. Lysine 14 on histone 3 was efficiently acetylated by NCOAT in these assays (Fig. 2*C*). Lysine 8 on histone 4 was also acetylated by NCOAT under the same conditions (Fig. 2*D*). In both cases, mNCOAT was active, while bNCOAT required exposure to mammalian whole cell extract. When these reactions were probed with an antibody to lysine 16 on histone 4, we were unable to detect any acetylation of this particular residue by NCOAT (data not shown). While other modified lysine residues have not been mapped, these data suggest that NCOAT is a type A HAT capable of modifying the histones within chromatin. While we cannot conclude at this time how promiscuous NCOAT is as an acetyltransferase, it is likely to have specific targets in the context of nucleosomes, as do all other previously characterized HATs. Further selectivity is likely provided for by the complex(es) in which NCOAT is bound at the promoter.

FIG. 2. **NCOAT has the ability to acetylate oligonucleosomal substrates.** *A*, gel mobility shift assay. 186-bp α-satellite DNA from HeLa cells (*lane 1*) was reconstituted with histone octameric core particles by salt dialysis (*lane 2*). *B*, autoradiograph of the HAT assay using oligonucleosome substrates. The oligonucleosome particles from *A* were excised from the gel and treated with mammalian-expressed or bacterially expressed NCOAT in the presence of [^{14}C]acetyl-CoA. The bacterially expressed protein was also pretreated with a mammalian whole cell lysate prior to the assay (indicated by an *asterisk* in *B-D*). The negative control (*NC*) is the assay performed in the absence of enzyme. *C*, the same assay was repeated with 10 μM unlabeled acetyl-CoA and then probed using anti-acetylated histone 3, lysine 14 antibody. *D*, Western blot of the same assay using anti-acetylated histone 4, lysine 8 antibody.



The Active Site for Acetyltransferase Activity Lies in the C Terminus of the Enzyme—To determine the region of NCOAT containing HAT activity, we constructed an *O*-GlcNAcase with a deletion of the C-terminal third of the enzyme, thus removing the region identified by SMART analysis to have AT-like composition. This enzyme was expressed as a GAL4 fusion protein and placed in both HAT assays. When the C-terminal 221 amino acids were removed from NCOAT the enzyme no longer exhibited HAT activity in either assay (Fig. 3), suggesting that this portion of the enzyme either comprises the active site or contains a feature requisite for the ability of NCOAT to function as an acetyltransferase. To address this question a GAL4 fusion protein was made using only the C-terminal third of NCOAT to determine whether, like many other AT domains, the AT domain of NCOAT alone can confer HAT activity to an otherwise non-HAT peptide, the GAL4 DNA binding domain. As shown in Fig. 3, the C-terminal 334 amino acids of NCOAT were sufficient to introduce HAT activity to the GAL4-DNA binding domain when the peptide was expressed in mammalian cells. GAL4 alone had no significant activity in either assay. Furthermore, this domain of NCOAT alone, when expressed as a GAL4 fusion, exhibited a full retention of activity when compared with that of the full-length NCOAT or to the positive control for this assay, p300. The same 334-amino acid peptide, when expressed in bacteria as a GST fusion protein failed to show HAT activity, just as seen when the full-length enzyme was expressed in bacteria. However, once again full enzymatic activity was observed after incubation of this peptide with a mammalian whole cell lysate (Fig. 3*A, lane 8*). These results indicate that the AT domain of NCOAT lies in the C-terminal third of the enzyme and that this domain can act autonomously when expressed in a mammalian cell system.

FIG. 3. **The region of NCOAT responsible for HAT activity lies in the C terminus.** *A*, filter binding assay. GAL4-mNCOAT exhibited activity rivaling that for the GST fusion used previously. The same construct lacking the C terminus of the enzyme showed background levels of activity, while a GAL4 fusion with only the C-terminal third was fully active. A bacterial construct of the C-terminal third was also tested in the presence or absence of mammalian whole cell lysate (*asterisk*). Splice variant proteins and a caspase 3-cleaved enzyme were similarly tested. *GK*, Goto Kakazaki; *SD*, Sprague-Dawley. *B*, select reactions were repeated using core histones and viewed by autoradiograph. *NC*, negative control. *C*, NCOAT variants. Alignment of three rat NCOAT clones derived from F344 (parental), Goto Kakazaki (*GK*), and Sprague-Dawley rat brains revealing protein isoforms resulting from alternative splicing.

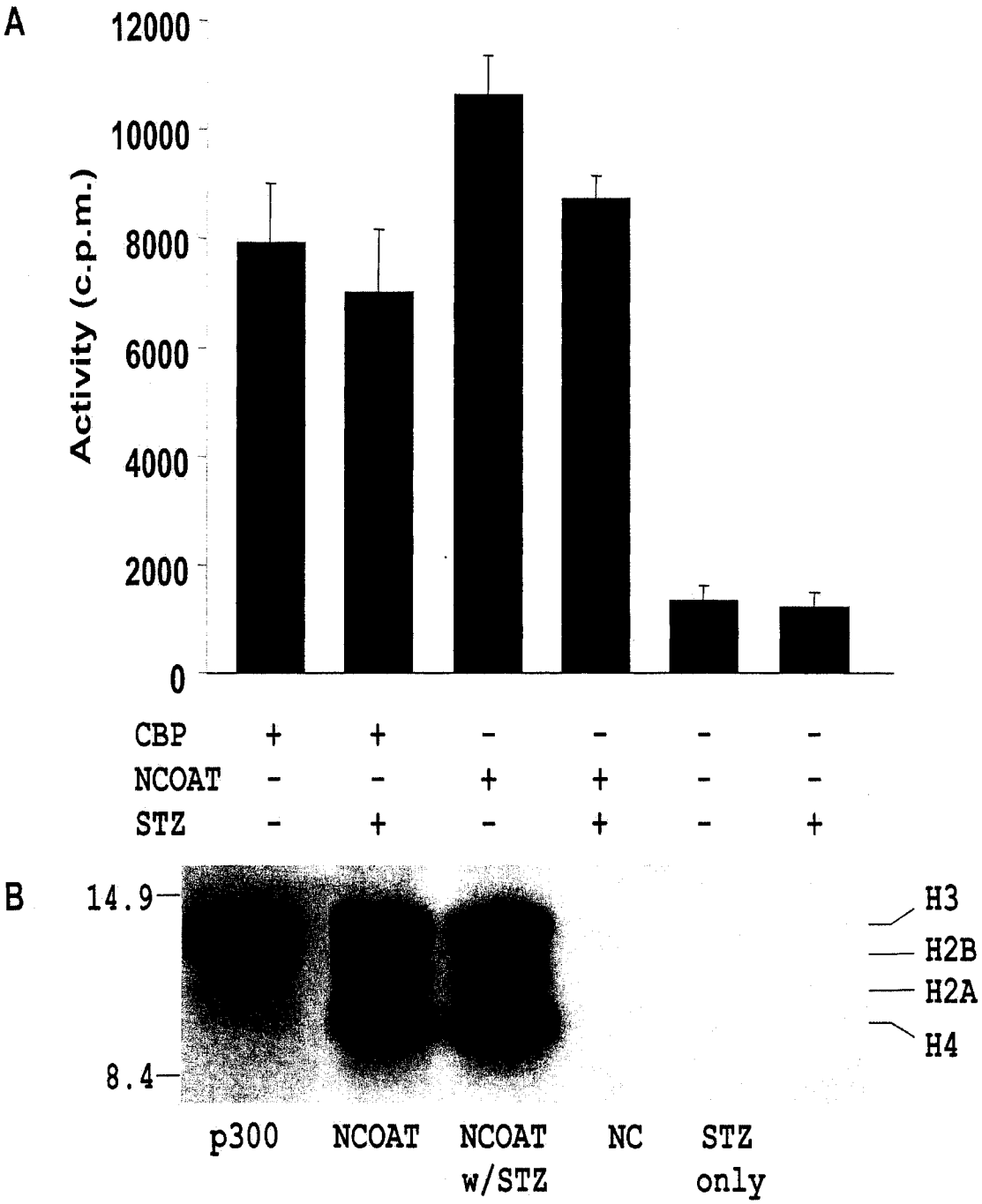


Identification of Two Splice Variants Containing HAT Activity but Not Hexosaminidase Activity—Over the course of our study, the rat NCOAT cDNA was cloned by reverse transcription-PCR from RNA derived from the brains of Spraque-Dawley and Goto-Kakazaki rats. We noted two subtle size variants. One cDNA corresponded in size to the full-length mouse NCOAT, while the two variants were smaller. These reverse transcription-PCR products were sequenced. Whereas the full-length rat NCOAT (GenBank™ accession number NM_131904) was nearly identical to its mouse counterpart (GenBank™ accession number AF132214), the shorter variants had missing sequences corresponding to exon 8 in the Goto Kakazaki rat and exons 8 and 9 in the Sprague-Dawley rat (Fig. 3*C*). Because the missing regions in these variants correspond exactly to exon boundaries, we believe they may result from alternative splicing of the gene. These variants were expressed as proteins of 90 kDa for the Goto Kakazaki variant due to the deletion of amino acids 250-345 and 84 kDa for the Sprague-Dawley variant due to the deletion of amino acids 250-398. Each of these variants was tested for hexosaminidase activity and was found to be catalytically inactive (data not shown). Both variant enzymes, however, were able to completely catalyze the addition of acetyl groups to the histones when compared with the full-length enzyme (Fig. 3*A*). This finding lends to the validity that the HAT active site resides in the C terminus of the enzyme, since the resultant proteins are able to retain full activity. These findings also raise the possibility that organisms may utilize alternative splicing to regulate the levels of the enzymatic activities of this bifunctional enzyme within a particular cell type or cell compartment. Interestingly, a splice variant resembling our C-terminal truncation missing the AT domain, has been discovered as well and retains considerable activity in hyaluronidase assays (31), further indicating the significance of splicing to favor one cellular activity over the other.

In addition to splicing, *O*-GlcNAcase has also been shown recently to be a competent substrate for caspase 3 cleavage, and that the processed product retains *O*-GlcNAcase activity (27). It has been reasoned that cleavage of the protein is an additional mechanism to control by separation the functions of the hexosaminidase domain from the potential AT domain (27). To independently address this possibility, we treated NCOAT with caspase 3 prior to our assays. Successful cleavage was monitored by SDS-PAGE and Coomassie staining. Such treatment did not affect the HAT activity of NCOAT (Fig. 3*A, lane 9*).

The HAT Activity of NCOAT Is Unchanged after Treatment with Streptozotocin—We and others have previously shown that the hexosaminidase activity of the enzyme can be inhibited *in vitro* with the diabetogenic drug, STZ (13, 25, 32, 33). Since STZ is a GlcNAc analog, its ability to inhibit the *O*-GlcNAcase function of NCOAT is not surprising. To rule out the possibility that STZ, which is an NO donor (34), can be targeted to the NCOAT protein, then exert an inhibitory effect on both enzymatic functions, we treated NCOAT with 0.1 M STZ for 1 h at 37 °C prior to performing either HAT assay. This concentration of STZ, which is sufficient to abolish hexosaminidase activity (25, 32), had a negligible effect on the acetyltransferase activity of NCOAT (Fig. 4, *A* and *B*). The slight decrease in activity observed by scintillation counting appears solely the consequence of the incubation conditions or to a trace amount of nonspecific involvement, since CBP and p300 also show a marginal loss of HAT activity when pretreated with STZ. The inhibitory effects of STZ on NCOAT were concluded therefore to be specific only for the hexosaminidase activity of the enzyme and not the result of a nonspecific, global effect such as oxidative damage to the enzyme. It seems likely that the

FIG. 4. **The HAT activity of NCOAT is unaffected by treatment with streptozotocin.** *A*, filter binding assay. Either CBP (*lanes 1* and *2*) or NCOAT (*lanes 3* and *4*) were left untreated or incubated with STZ prior to performing the HAT assay. Histone substrate in the absence of enzyme (*lane 5*) was also pretreated with STZ (*lane 6*). *B*, visualization of the HAT assays using core histones and autoradiography. Molecular mass markers in kDa are at the *left* of the gel. *NC*, negative control.



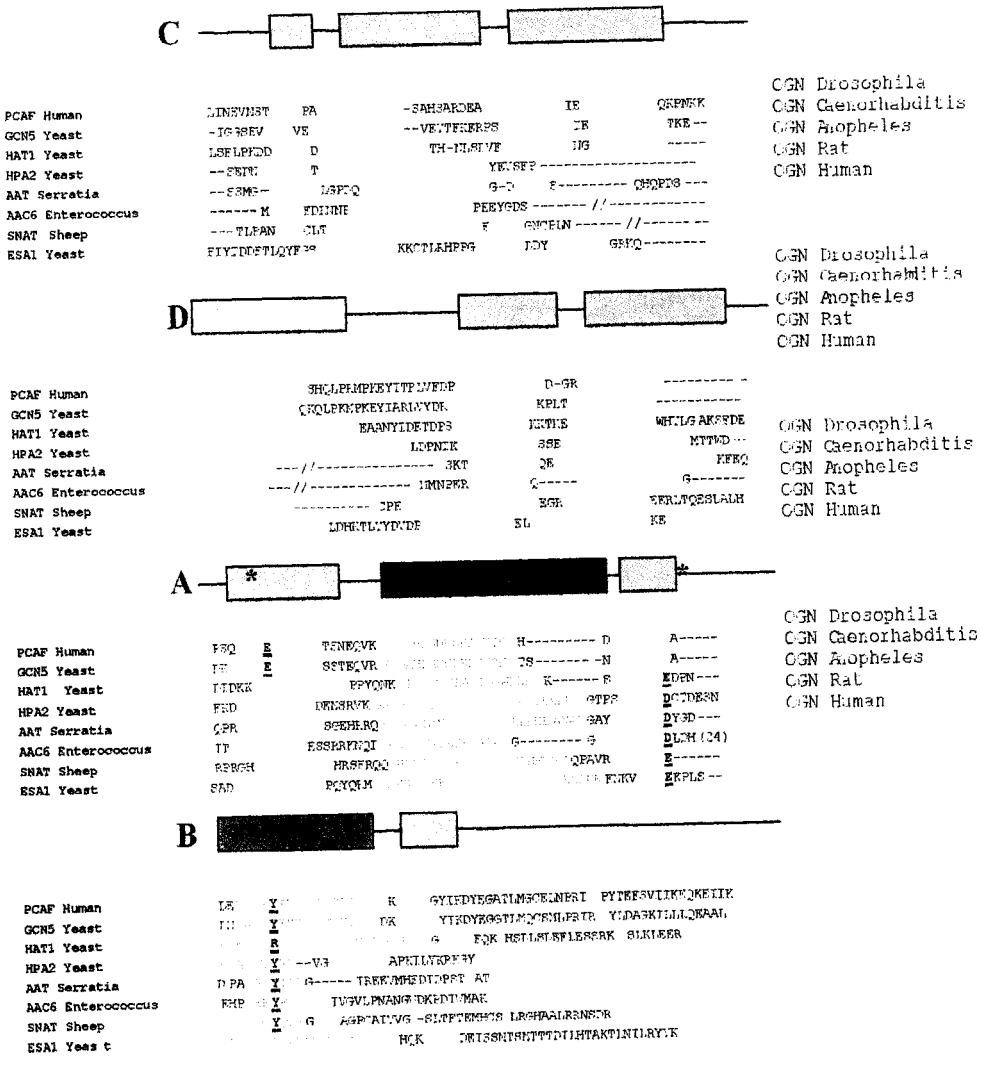
GlcNAc analog streptozotocin is acting directly on the *O*-GlcNAcase active site and that the C-terminal acetyltransferase domain is, in the active enzyme, completely distinct in function from the N-terminal *O*-GlcNAcase domain. Such functional independence is consistent with our observations using splice variants, caspase, or the HAT domain alone and also support splice variation and caspase cleavage as plausible methods of cellular control of these activities.

Identification of Residues Essential for HAT Activity—The C-terminal acetyltransferase domain of NCOAT was further characterized. Eight of the ten acetyltransferases that SMART computer analysis identified as having similar domain composition to NCOAT have been crystallized, and a detailed map of their active site secondary structures is known. A comparison of the active sites of these enzymes shows that, while containing only low levels of sequence identity (≈15%) and similarity (≤50%), they share a highly conserved structural fold (8). Such low levels of sequence homology make identification of new acetyltransferases by sequence alone difficult. A more defining feature of acetyltransferases is their overall active site core architecture. This core structure includes four universally present motifs, designated motifs C, D, A, and B (Fig. 5), whose lengths and sequences are variable. The intervening loops are also variable, nevertheless, the overall structure of these AT is superimposable in each case (6-8, 35, 36). All ATs contain the same arrangement of α-helixes and β-sheets as indicated in Fig. 4*A*. Using secondary structure prediction programs (19, 28, 29), the C terminus of NCOATs from five different species can be predicted to contain this exact motif pattern (Fig. 5), providing further evidence that the AT active site of NCOAT lies in the C-terminal third of the enzyme as indicated by measurements of activity. Within this stretch, there are seven invariant

FIG. 5. **NCOAT possesses similar active site composition and critical residues as several previously characterized acetyltransferases.** On the *left* are the secondary structure alignments of eight crystallized acetyltransferases. *Orange bars* indicate β-sheets. *Green, pink, blue*, and *red bars* show the four conserved α-helices. Published sites (with exceptions noted in the text) of residues confirmed necessary for catalysis are shown by an *asterisk*. The *right* shows the predicted secondary structures of NCOATs from five different species (19-21). Conserved residues corresponding to critical residues in acetyltransferases are likewise indicated by an *asterisk*. *OGN, O*-GlcNAcase.

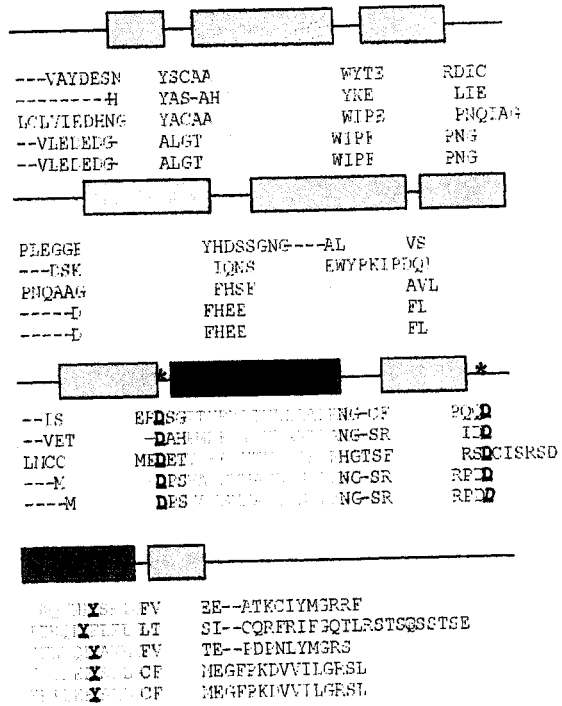
Acetyl Transferase Catalytic Domain

(secondary structure solved via crystallization)



NCOAT Putative AT Domain

(predicted secondary structural motifs)



residues among the secondary structure alignments. Three of these are glycines, there is one invariant leucine, two aspartic acids (at positions 853 and 884 for mouse NCOAT), and one tyrosine residue (at position 891), which is exchanged with an adjacent phenylalanine in the *Caenorhabditis elegans* NCOAT. The amino acid conservation at these sites suggests their functional importance to the activity of the enzyme.

Examination of the residues identified to be necessary for catalysis in the majority of these eight crystallized acetyltransferases reveals that a carboxylate-containing residue lying on the back end of the β-sheet C-terminal to helix A is the critical residue responsible for the initial base step of the enzymatic reaction. In this step, a proton is abstracted from the substrate lysine ε-amino group in the enzyme-lysine-coenzyme A ternary complex (37). Human PCAF and yeast Gcn5 are exceptions in this case, with these enzymes having their critical base carboxylate in the middle of the β-sheet N-terminal to helix A (38, 39). Interestingly, it has been reported that, despite this difference in sequence position, these critical base residues superimpose identically in the three-dimensional structures of each of these enzymes (35). After the deprotonated lysine directly attacks the acetyl group of acetyl-CoA, a general acid is needed to donate a proton to the CoA sulfur atom, to facilitate the departure of the thiolate leaving group. In all but two cases (8, 40) the catalytic acid is a tyrosine situated near the middle of helix B (8, 39).

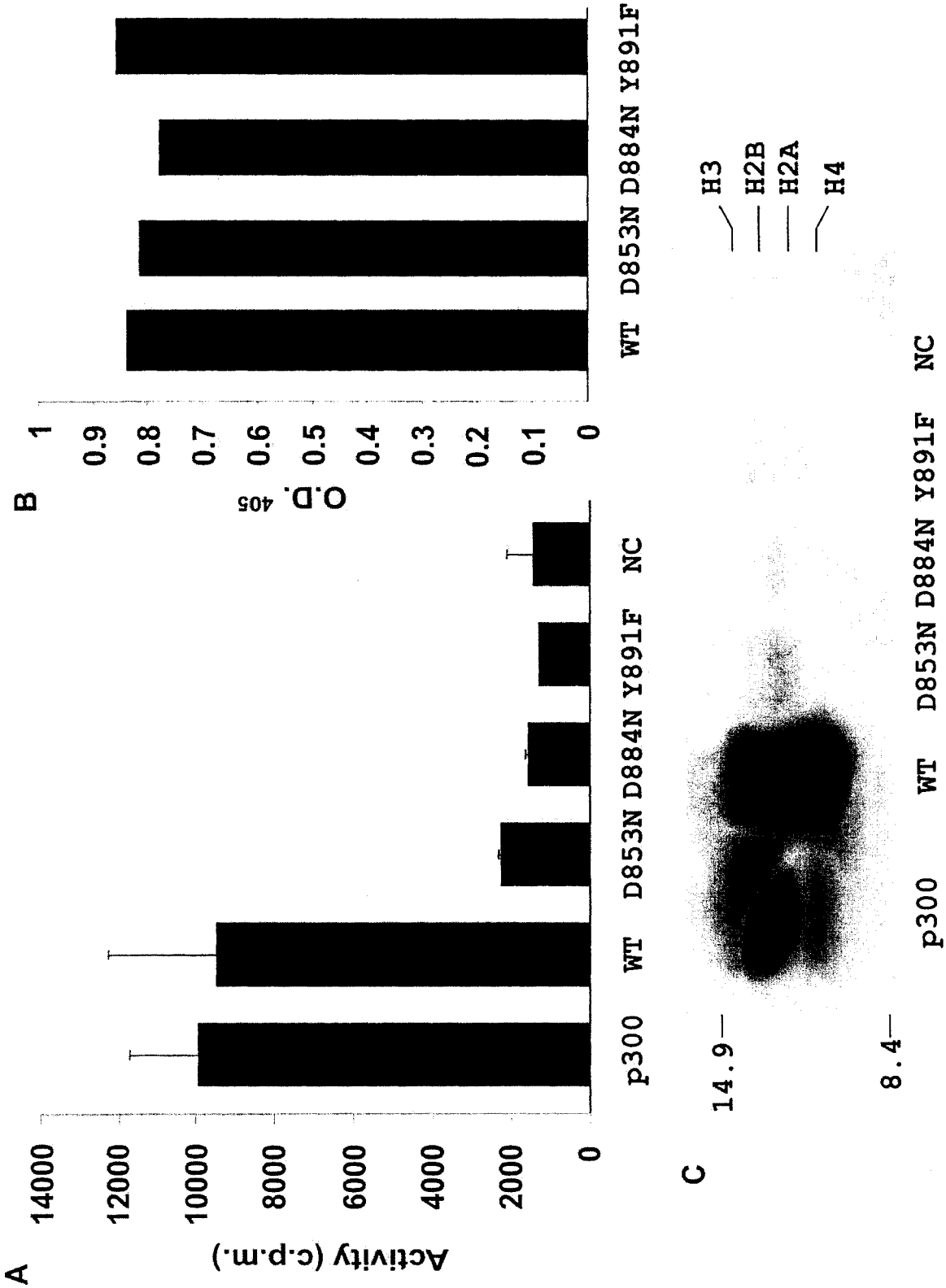
Since the invariant aspartic acids and the tyrosine in these enzymes lie at corresponding positions in the predicted secondary structures of NCOATs (Fig. 5), we wished to determine whether these residues indeed were necessary for catalysis, as they are in their crystallized counterparts. Each aspartic acid was independently mutated to an asparagine, and the tyrosine to a phenylalanine, so that the mutant proteins could be tested for acetyltransferase activity. In the filter binding assay, the counts observed for each mutant

protein matched that of the background (Fig. 6*A*). These mutant enzymes displayed no change on hexosaminidase activity (Fig. 6*B*), demonstrating again the independence of the two domains and suggesting that the lack of HAT activity observed by the mutation of these residues is not the result of a gross alteration of total protein conformation. Each mutant enzyme was also unable to label any of the core histones as visualized by autoradiograph (Fig. 6*C*). These results demonstrate the dependence of NCOAT on these residues for its ability to catalyze acetyltransfer. Because their respective positions in the predicted secondary structure of NCOATs are comparable with the essential catalytic acid and base in other AT enzymes, it is reasonable to expect that either Asp^{853} or Asp^{884} for mouse NCOAT is the catalytic acid in the reaction and the tyrosine at position 891 is the general base. Currently, however, in the absence of crystal structure, their exact roles cannot yet to be certified.

Discussion

Several transcriptional activating cofactors have been recently identified that contain a domain conferring HAT activity *in vitro*. HAT activity is required to open the chromatin structure as part of the process of gene activation. Recently, we added *O*- GlcNAc modification of the transcription apparatus by OGT as part of the repression process (12), therefore making the removal of these sugars also necessary for transcriptional activation to occur. Here, we demonstrate that *O*-GlcNAcase, the only enzyme encoded in the genome that can remove these sugars, is bifunctional; it contains a domain with intrinsic HAT activity. Because the protein has these two enzymatic activities, we have renamed it NCOAT.

FIG. 6. **Relative HAT activity of each point mutant enzyme.** *A*, filter binding assay. Each predicted critical residue from Fig. 4 was individually mutated and placed in the HAT assay with synthetic H4 tail substrate. *B*, hexosaminidase assay using each point mutant enzyme. *C*, mutant proteins were also incubated with core histones and visualized by SDS-PAGE and autoradiography. *WT*, wild type.



Saccharomyces cerevisiae cells have been shown to use HATs and HDACs to regulate gene transcription (41). In these single cell organisms, however, the post-translation modification of proteins with *O*-GlcNAc has not been observed. Nevertheless, yeast cells use a combination of proteins, including Ssn6-Tup1, in the repression of several genes (42, 43). Interestingly, Ssn6, while having no known enzymatic function, has TPR domains very similar to OGT. More recently, it has been shown that Ssn6, through its TPR domains, associates with corepressors and HDACs (44), further suggesting a similarity of OGT to Ssn6. We speculate that an Ssn6-like TPR protein fused with a glycosyltransferase domain to form OGT during evolution. This fusion may have been necessitated by the need for more stringent repression of gene expression when organisms became multicellular with differentiated cell functions. The enzymatic function of OGT would provide an additional covalent modification needed to cooperate with HDAC in corepression complexes. To activate genes repressed by OGT and HDACs, both the enzymatic removal of the sugars and addition of acetyl groups would therefore be required. That both functions reside in one protein, NCOAT, lends further credence to the importance of both modifications in metazoan gene expression.

We further characterized the HAT domain of NCOAT. It has the ability to acetylate all four core histones when either free or bound by DNA in the context of oligonucleosome arrays. Moreover, it does so with a specific activity rivaling that of the global transcriptional coactivators, p300 and CBP. The capacity of this dual function enzyme to catalyze its two activities, not surprisingly, appears to be under tight control, as displayed by the discrepancy between the levels of intrinsic activities of the same enzyme expressed in a mammalian *versus* a bacterial system. The recovery of optimal activity by the bacterially expressed enzyme when incubated with a mammalian whole cell lysate is sugges-

tive of a currently uncharacterized post-translational modification(s) that correlate with the activation of the enzymatic domains so that they function proficiently. We have also discovered two naturally occurring splice variants of NCOAT, which lack hexosaminidase activity but retain HAT activity, implying that alternative editing at the RNA level may also govern the two relative activity levels of NCOAT in a cell. Furthermore, an additional level of NCOAT control can result from separation of the two activities by a proapoptotic factor, caspase 3.

The entirety of the HAT domain lies in the C-terminal third of the protein where, although it shows weak sequence homology, NCOAT is predicted to contain the characteristic four-helix bundle, which comprises the conserved core within acetyltransferase active sites. This places the active site of the HAT domain at the opposite end of the protein from the N-terminal hexosaminidase domain. While the two domains may reside in the same protein, in the active state(s) of the enzyme the AT domain does not rely on the N-terminal *O*-GlcNAcase domain. STZ treatment only affects the *O*-GlcNAcase domain, while leaving the AT domain functional. The splice variants lacking the hexosaminidase activity retain HAT activity. Liberation of the AT domain by caspase 3 or expression of the C-terminal third of the protein alone in mammalian cells results in a fully functional acetyltransferase. The bacterially expressed C-terminal third is able to fully recover activity when incubated with a mammalian whole cell lysate, indicating that at least one of the proposed activating post-translational modifications lies within this region. While we have not yet determined the nature of these covalent modifications, this apparent dependence on modification emphasizes the tight control exerted on NCOAT by upstream signaling pathways.

Within the C-terminal active site we have identified three residues that are critical for efficient acetyl transfer. For mouse NCOAT these are Asp^{853}, Asp^{884}, and Tyr^{891}. The acetyltransferases found by SMART analysis to resemble NCOAT are known to utilize one of two distinct mechanisms. The principle mechanism involves an ordered Bi-Bi reaction in which an active site carboxylate base deprotonates the substrate lysine ε-amino group, then, in a single step, the acetyl group is transferred from the reactive thioester group of acetyl-CoA to the uncharged substrate lysine. An active site tyrosine (with exception) then donates a proton to facilitate CoASH departure (38, 45, 46). The other mechanism involves the acetyl transfer from CoA to an enzyme nucleophile, creating a covalent intermediate, and then to the substrate (40, 47). We were unable to detect evidence for such a self-acetylated intermediate. A mutation to the only free cysteine in the active site, the most likely site to accept an acetyl group, does not impair HAT activity (data not shown). We therefore propose that NCOAT undergoes the single step reaction described for those members of the Gcn5, HAT1, and GNAT family of acetyltransferases. This mechanism, and the placement of conserved residues in positions corresponding to those shown necessary for catalysis in these enzymes, suggest that the tyrosine at position 891 acts as the general acid involved in NCOAT acetyltransfer. Likewise, the aspartic acids at position 853 or 884 are promising candidates as the necessary catalytic base.

Given the mechanism, and the fact that NCOAT can acetylate all four core histones, like p300 and CBP, other substrates for NCOAT acetylation potentially exist as they do for these other enzymes. Much like the diversity of *O*-GlcNAcylated proteins, which exist both in the nucleus and cytoplasm (48), NCOAT may be critical for an increasing number of cellular events requiring acetyl modification. The diverse nuclear as well as cytoplasmic activities that are modulated by acetylation make this AT domain

perhaps as important as the *O*-GlcNAcase domain, justifying the change of the name of this bifunctional protein to NCOAT.

In summary, NCOAT contains versatile enzymes. Its *O*-GlcNAcase can activate proteasomes (49), but another compelling idea about the function of NCOAT is in the reversibility of gene repression. The recent finding that OGT is specifically recruited by mSin3A along with HDAC to contribute to the corepression of eukaryotic gene expression (12) implies that derepression requires the converse enzymes, an *O*-GlcNAcase and HATs, to reverse the protein modifications involved in repression. Although the AT activity of NCOAT has yet to be tested *in vivo*, we have come closer to this goal by showing that NCOAT specifically modifies some important lysine residues in reconstituted oligonucleosomes. Thus NCOAT contains two activities required to activate gene transcription. How NCOAT finds its way to repressed genes undergoing a switch to activation will become the subject of much future research.

References

1. Wolffe, A. P. (ed) (1992) *Chromatin: Structure and Function*, Academic Press, London, UK
2. Strahl, B. D., and Allis, C. D. (2000) *Nature* **403**, 41-45
3. Turner, B. M., and O'Neill, L. P. (1995) *Semin. Cell Biol.* **6**, 229-236
4. Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998) *Mol. Cell Biol.* **18**, 4629-4638
5. Richards, E. J., and Elgin, S. C. (2002) *Cell* **108**, 489-500
6. Marmorstein, R. (2001) *J. Mol. Biol.* **311**, 433-444
7. Wolf, E., Vassilev, A., Makino, Y., Sali, A., Nakatani, Y., and Burley, S. K. (1998) *Cell* **94**, 439-449

8. Angus-Hill, M. L., Dutnall, R. N., Tafrov, S. T., Sternglanz, R., and Ramakrishnan, V. (1999) *J. Mol. Biol.* **294**, 1311-1325

9. Hassig, C. A., Fleischer, T. C., Billin, A. N., Schreiber, S. L., and Ayer, D. E. (1997) *Cell* **89**, 341-347

10. 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

11. Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (1997) *Cell* **89**, 357-364

12. Yang, X., Zhang, F., and Kudlow, J. E. (2002) *Cell* **110**, 69-80

13. Yang, X., Su, K., Roos, M. D., Chang, Q., Paterson, A. J., and Kudlow, J. E. (2001) *Proc. Natl. Acad. Sci. U S A* **98**, 6611-6616

14. Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993) *J. Biol. Chem.* **268**, 10416-10424

15. Lamarre-Vincent, N., and Hsieh-Wilson, L. C. (2003) *J. Am. Chem. Soc.* **125**, 6612-6613

16. Perissi, V., Aggarwal, A., Glass, C. K., Rose, D. W., and Rosenfeld, M. G. (2004) *Cell* **116**, 511-526

17. Li, X., Oghi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., Nigam, S. K., Aggarwal, A. K., Maas, R., Rose, D. W., and Rosenfeld, M. G. (2003) *Nature* **426**, 247-254

18. Comtesse, N., Maldener, E., and Meese, E. (2001) *Biochem. Biophys. Res. Commun.* **283**, 634-640

19. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) *Proc. Natl. Acad. Sci. U S A* **95**, 5857-5864

20. Ausubel, J., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (eds) (1990) *Current Protocols in Molecular Biology*, Vol. 1, pp. 2.0.1-3.19.8, Greene Publishing and Wiley Interscience, New York

21. Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., and Luger, K. (2004) *Methods Enzymol.* **375**, 23-44

22. Tanaka, Y., Tawaramoto-Sasanuma, M., Kawaguchi, S., Ohta, T., Yoda, K., Kurumizaka, H., and Yokoyama, S. (2004) *Methods* **33**, 3-11

23. Brownell, J. E., and Allis, C. D. (1995) *Proc. Natl. Acad. Sci. U S A* **92**, 6364-6368

24. Bannister, A. J., and Kouzarides, T. (1996) *Nature* **384**, 641-643

25. Roos, M. D., Xie, W., Su, K., Clark, J. A., Yang, X., Chin, E., Paterson, A. J., and Kudlow, J. E. (1998) *Proc. Assoc. Am. Physicians* **110**, 422-432

26. Dong, D. L., and Hart, G. W. (1994) *J. Biol. Chem.* **269**, 19321-19330

27. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) *J. Biol. Chem.* **277**, 1755-1761

28. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) *Bioinformatics* **14**, 892-893

29. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* **13**, 222-245

30. Gao, Y., Wells, L., Comer, F. I., Parker, G. J., and Hart, G. W. (2001) *J. Biol. Chem.* **276**, 9838-9845

31. Heckel, D., Comtesse, N., Brass, N., Blin, N., Zang, K. D., and Meese, E. (1998) *Hum. Mol. Genet.* **7**, 1859-1872

32. Hanover, J. A., Lai, Z., Lee, G., Lubas, W. A., and Sato, S. M. (1999) *Arch. Biochem. Biophys.* **362**, 38-45

33. Konrad, R. J., Mikolaenko, I., Tolar, J. F., Liu, K., and Kudlow, J. E. (2001) *Biochem. J.* **356**, 31-41

34. Turk, J., Corbett, J. A., Ramanadham, S., Bohrer, A., and McDaniel, M. L. (1993) *Biochem. Biophys. Res. Commun.* **197**, 1458-1464

35. Yan, Y., Barlev, N. A., Haley, R. H., Berger, S. L., and Marmorstein, R. (2000) *Mol. Cell* **6**, 1195-1205

36. Dutnall, R. N., Tafrov, S. T., Sternglanz, R., and Ramakrishnan, V. (1998) *Cell* **94**, 427-438

37. Tanner, K. G., Trievel, R. C., Kuo, M. H., Howard, R. M., Berger, S. L., Allis, C. D., Marmorstein, R., and Denu, J. M. (1999) *J. Biol. Chem.* **274**, 18157-18160

38. Langer, M. R., Tanner, K. G., and Denu, J. M. (2001) *J. Biol. Chem.* **276**, 31321-31331

39. Poux, A. N., Cebrat, M., Kim, C. M., Cole, P. A., and Marmorstein, R. (2002) *Proc. Natl. Acad. Sci. U S A* **99**, 14065-14070

40. Yan, Y., Harper, S., Speicher, D. W., and Marmorstein, R. (2002) *Nat. Struct. Biol.* **9**, 862-869

41. Vogelauer, M., Wu, J., Suka, N., and Grunstein, M. (2000) *Nature* **408**, 495-498

42. Watson, A. D., Edmondson, D. G., Bone, J. R., Mukai, Y., Yu, Y., Du, W., Stillman, D. J., and Roth, S. Y. (2000) *Genes Dev.* **14**, 2737-2744

43. Smith, R. L., and Johnson, A. D. (2000) *Trends Biochem. Sci.* **25**, 325-330

44. Davie, J. K., Edmondson, D. G., Coco, C. B., and Dent, S. Y. (2003) *J. Biol. Chem.* **278**, 50158-50162

45. De Angelis, J., Gastel, J., Klein, D. C., and Cole, P. A. (1998) *J. Biol. Chem.* **273**, 3045-3050

46. Lewendon, A., Murray, I. A., Shaw, W. V., Gibbs, M. R., and Leslie, A. G. (1994) *Biochemistry* **33**, 1944-1950

47. Thompson, S., Mayerl, F., Peoples, O. P., Masamune, S., Sinskey, A. J., and Walsh, C. T. (1989) *Biochemistry* **28**, 5735-5742

48. Wells, L., Vosseller, K., and Hart, G. W. (2001) *Science* **291**, 2376-237849.

49. Zhang, F., Su, K., Yang, X., Bowe, D. B., Paterson, A. J., and Kudlow, J. E. (2003) *Cell* **115**, 715-725

THE LOCATION AND CHARACTERIZATION OF THE *O*-GlcNAcase ACTIVE SITE

by

CLIFFORD TOLEMAN, ANDREW J. PATERSON, RONALD SHIN, AND JEFFREY E. KUDLOW

Submitted to *Journal of Biological Chemistry*

Format adapted for dissertation

Abstract

Nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase is a bifunctional nucleo-cytoplasmic protein with both *O*-GlcNAcase and histone acetyltransferase domains. The *O*-GlcNAcase portion catalyzes the removal of *O*-GlcNAc modifications from proteins and resides in the N-terminal third of the enzyme. The recognition of the substrate GlcNAc suggests that the *O*-GlcNAcase is related in structure and catalytic mechanism to chitinases, hexosaminidases, and hyaluronidases. These families of glycosidases all possess a catalytic doublet of carboxylate-containing residues that orient and use the *N*-acetyl group for catalysis. Indeed, we show that the *O*-GlcNAcase does also possess the catalytic doublet motif shared among these enzymes and that the two essential residues noted above are aspartic acids at positions 175 and 177, respectively, in mouse *O*-GlcNAcase. In addition, a conserved cysteine at 166 and a conserved aspartic acid at 174 were also found to play a role in catalysis. Given this information, we propose that the *O*-GlcNAcase active site resembles those of the above glycosidases and that *O*-GlcNAcase carries out catalysis through a transition state in the same substrate-assisted acid-base manner used by these enzymes. Further evidence of this transition state was provided by streptozotocin, a selective inhibitor of the enzyme and GlcNAc analog. Streptozotocin appears to be converted by the enzyme to a more stable transition-state analog, an inhibitory compound which is up to 10^5 more potent than the original ligand, thus providing insight into the mode of catalysis of the *O*-GlcNAcase enzyme.

Introduction

O-GlcNAcylation is a ubiquitous post-translational modification in which the monosaccharide *N*-acetylglucosamine (GlcNAc) is linked by the enzyme β-*N*-

acetylglucosaminyltransferase (OGT) to specific serine or threonine hydroxyls on target protein (1). These modifications can be selectively removed by the enzyme β-*N*-acetylglucosaminidase (*O*-GlcNAcase), which hydrolyzes the glycosidic bond and liberates the free protein and GlcNAc (2-4). *O*-GlcNAcase activity can be selectively inhibited by the diabetogenic drug streptozotocin (STZ) (5), a GlcNAc analog that can cause the accumulation of *O*-GlcNAcylated proteins in a cell and result in cell death (6). Although STZ is a GlcNAc analog and can reasonably be predicted to target the *O*-GlcNAcase active site, currently its exact mechanism of action on *O*-GlcNAcase is unknown.

Close to 200 proteins to date have been shown to be *O*-GlcNAc-modified (7), with these proteins playing a role in a number of processes of fundamental importance to cell survival, including cell cycle control and cell division (8-11), import and export through the nuclear pore (9,12), transcription (13-17), translation (13), proteosomal degradation (18), and others. The role that this modification plays is not completely understood in many cases; however, evidence suggests that it plays many of the same roles that phosphorylation does such as the regulation of protein-protein interactions (9, 19, 20) and protein stability (15, 21) and the induction of conformational changes that result in an alteration of protein function. Unlike the many phosphatases that exist in a cell, however, *O*-GlcNAcase is the only nuclear and cytosolic enzyme known to be capable of removing *O*-GlcNAc additions. The *O*-GlcNAcase enzyme resides in the N-terminal portion of a larger protein termed nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase (NCOAT) which contains an acetyltransferase (AT) domain in the C terminus (22). At least in some circumstances, NCOAT associates with OGT to form the *O*-GlcNAczyme, an enzyme system involved in the control of transcription.

To date, the *O*-GlcNAcase domain has not been sufficiently characterized and has been assigned to glycoside hydrolase family 84 with enzymes of similar sequence (23, 24). It is likely, though, that the *O*-GlcNAcase is highly related in both catalytic mechanism and active site composition to those glycoside hydrolases in families 18, 20, and 56. These enzymes, chitinases, β-hexosaminidases, and hyaluronidases, respectively, have been more extensively characterized; several members have had their crystal structures solved. These structures as well as mutagenesis and product analyses have demonstrated that these enzymes, which are similar to *O*-GlcNAcase in that they recognize GlcNAc as their core substrate, have a conserved and highly proficient mechanism of catalysis (25).

These enzymes catalyze their respective reactions through an acid-base mechanism, which requires both a proton donor and a nucleophile. In the course of the reaction, the 2-acetamido group of the substrate itself becomes the required nucleophile (26, 27), thereby participating in the reaction. The *O*-GlcNAcase enzyme behaves in this same substrate-assisted mechanism. While mutagenesis evidence is provided for this mechanism, the putative transient transition state of the normal *O*-GlcNAc substrate is unstable and cannot be isolated. However, the GlcNAc analog, STZ, because of its chemical structure, competitively inhibits the enzyme, which converts the compound into a stable transition state analog. This paper utilizes these two independent methods to probe into the mechanism of catalysis of the *O*-GlcNAcase portion of NCOAT.

Experimental Procedures

Plasmids and Site-directed Mutagenesis—The construction of a pUC118-pTM hybrid expression vector containing the full-length and splice variant mouse NCOAT

(mNCOAT) was designed as described previously (22). The NCOAT gene was modified by site-directed mutagenesis using a four-primer cassette strategy to introduce the substitution and the following selection sites: BamHI for C166S, ClaI for D174N and D175A, and AseI for D177N. PCR amplifications were performed using pBluescript-mNCOAT as template and the following oligonucleotides (restriction sites in bold, mutated residues underlined): for C166S; 5'-CTCTCAGTTT**GGATCC**AGGTCGTTTGC-3'; for D174N; 5'-TTGCTTTTTGACGAT**ATCGAT**CATAATATG-3'; for D175A; 5'-TTGCTTTTTGATGCT**ATCGAT**CATAATATG-3'; for D177N; 5'-TTGCTTTTTGATGAT**ATTAAT**CATAATATG-3'. Flanking oligonucleotides were as follows: upstream, 5'-CATGGATCCGAGATGTATTCAGTGG-3', and downstream, 5'-TCTTCAATAGATTCCACTGG-3'. PCR and recombinant DNA manipulations were performed according to standard procedures. For C166S and D177N, each ligation mixture was transformed into DH5α cells (Invitrogen) and screened for the correct insert. The D174N and D175A mutant ligations were transformed into DM1 cells (Gibco) for selection by ClaI.

Protein Expression—The wild type, mutant, and splice variant transformants were grown overnight at 37 °c in 500 ml TB medium containing 50 mg ampicillin. The plasmid DNA from the lysed cell culture was collected by a cesium chloride/ethidium bromide gradient and purified by ethanol precipitation. Plasmid DNA was transfected into BSC-40 cells by electroporation using 20 μg of DNA. Cells were cultured overnight in Dulbecco's modified Eagle's medium (Gibco) with 10% fetal bovine serum, 50 mg penicillin, and 400 mg gentamicin. Plates were then infected with VTF-7 vaccinia virus

for 2 h in 2 ml serum-free media, after which 10% fetal bovine serum-containing media was returned; the cells were then cultured overnight before harvesting.

Protein Purification—Harvested cells were lysed as described previously (22) and then stored at -80 °c until use. An aliquot of the lysate was loaded onto glutathione-Sepharose beads (Amersham) and incubated at 4 °c for 1 h. After washing, the glutathione-*S*-transferase (GST) fusion proteins were eluted from the beads for 1 h at room temperature using 30 mM reduced glutathione. Purified fusion proteins were visualized by 8% SDS-PAGE followed by Coomassie staining, and equal quantities of protein were used in subsequent assays as confirmed by Bio-Rad protein assay. Purified NCOAT to be incubated with STZ was left resin bound for washing.

O-GlcNAcase Assay and STZ Treatments—The hexosaminidase assay was performed as described by Dong and Hart (4). When varying substrate levels were called for, these were diluted appropriately to maintain an equal reaction volume. For each reaction, 10 μl (1 μmol) of purified wild type, splice variant, or mutant NCOAT were used per 100 μl reaction. Where indicated, 10 μl of 0.1-M *N*-ethylmaleimide (Sigma) were added to a 90-μl protein suspension in 50 mM sodium cacodylate buffer to maintain a consistent 100-μl total reaction volume. For STZ treatments, a stock solution of 1 M STZ (or an appropriate 10X solution for the inhibition kinetic studies) was dissolved in H_2O; 10 μl were added to resin-bound protein, keeping a total final volume of 100 μl. Reactions were carried out for 30 min at 37 °c, after which the beads were washed and placed in the hexosaminidase assay. The beads were washed and trypsinized using 2 μl trypsin (2 μg) in 100 μl 25-mM $NaHCO_3$ for 4 h at 37 °c. The supernatants were collected and reincu-

bated with NCOAT as above. 0.1 M STZ alone was also trypsin treated and incubated with fresh NCOAT. For molecular weight determinations, either 0.1 M fresh, trypsin-treated, or enzyme-treated STZ was prepared as a 100-µl solution and injected into a liquid chromatograph tandem mass spectrometer (PE Sciex). Molecular weights were also determined for 100 µl STZ boiled in water for 5 min and for 100 µl STZ dissolved in NaOAc pH 4.2 buffer for 30 min at room temperature. To monitor structural changes, the same reactions were repeated with the following exceptions to obtain higher quantities for study. Bulk 500-ml cultures of NCOAT were lysed, and the entire lysate was purified on 1 ml of resin. 1 ml of 1 M STZ was used for incubation. Following washes, bound STZ was liberated by heat denaturation at 100°c for 5 min. The protein-STZ solutions were then passed through Micropure-EZ enzyme removers (Amicon). 1 mM fresh STZ or enzyme-incubated STZ was suspended in 100% D_2O for one dimensional NMR analysis using a Bruker Avance (DRX) 600-MHz spectrophotometer equipped with a TXI probe and using 128 scans at 298 °K. Samples were then lyophilized and suspended in deuterated dimethylsulfoxide for additional one- and two dimensional NOESY analysis using 100-msec mixing times, 128 scans, and 256 increments under identical conditions. Data were processed using Bruker's XWIN-NMR v3.5 software and further processed by NMRPIPE (NIH).

Free Cysteine Identification—To locate free cysteines, biotinylated *N*-(hexyl)-3'-(2'-pyridyldithio) propionamide (HPDP) (Pierce) was prepared fresh as a 4-mM solution in dimethylsulfoxide from a 50-mM stock in dimethylformamide. Purified wild type NCOAT was eluted in 50 µl of glutathione elution buffer, and 50 µl of biotinylated HPDP stock solution was added. Reactions were incubated in the dark at room tempera-

ture for 1 h with shaking. Samples were then digested with the addition of 2 µl trypsin (Roche) for 4 h at 37 °c. Biotinylated fragments were passed over a 100-µl avidin column (Pierce) for 1 h at 4 °c with shaking. Bound peptides were eluted with 50 mM dithiothreitol for 1 h at room temperature. The eluant was analyzed by matrix-assisted laser desorption/ionization mass spectroscopy.

Active Site Characterization—For pH analyses, stock solutions of 50 mM sodium cacodylate buffer were made and adjusted so that each pH value indicated was the final pH after 10 µl of elution buffer (Tris-HCl, pH 8.0) was added to 90 µl of sodium cacodylate buffer. For kinetic analyses, 0.25 µl, 0.5 µl, 1 µl, 2 µl, 3 µl, and 4 µl of 0.1-M *p*NP-GlcNAc were added to independent reactions. V_{max} and K_m were calculated using non-linear curve fitting with GraphPad Prism version 4.00 for Windows (GraphPad Software, San Diego California, www.graphpad.com). Bio-Rad protein assay for enzyme concentration allowed for the calculation of k_{cat} value on the specific activity of each purified protein at V_{max} for *p*NP-GlcNAc, assuming an M_r of 130,000 for the GST fusion protein. The addition of sodium azide was performed by making stock solutions of 50 mM sodium cacodylate with 0.5 M, 1 M, 1.5 M, 2 M, 3 M, or 4 M NaN_3 prior to performing the hexosaminidase assay.

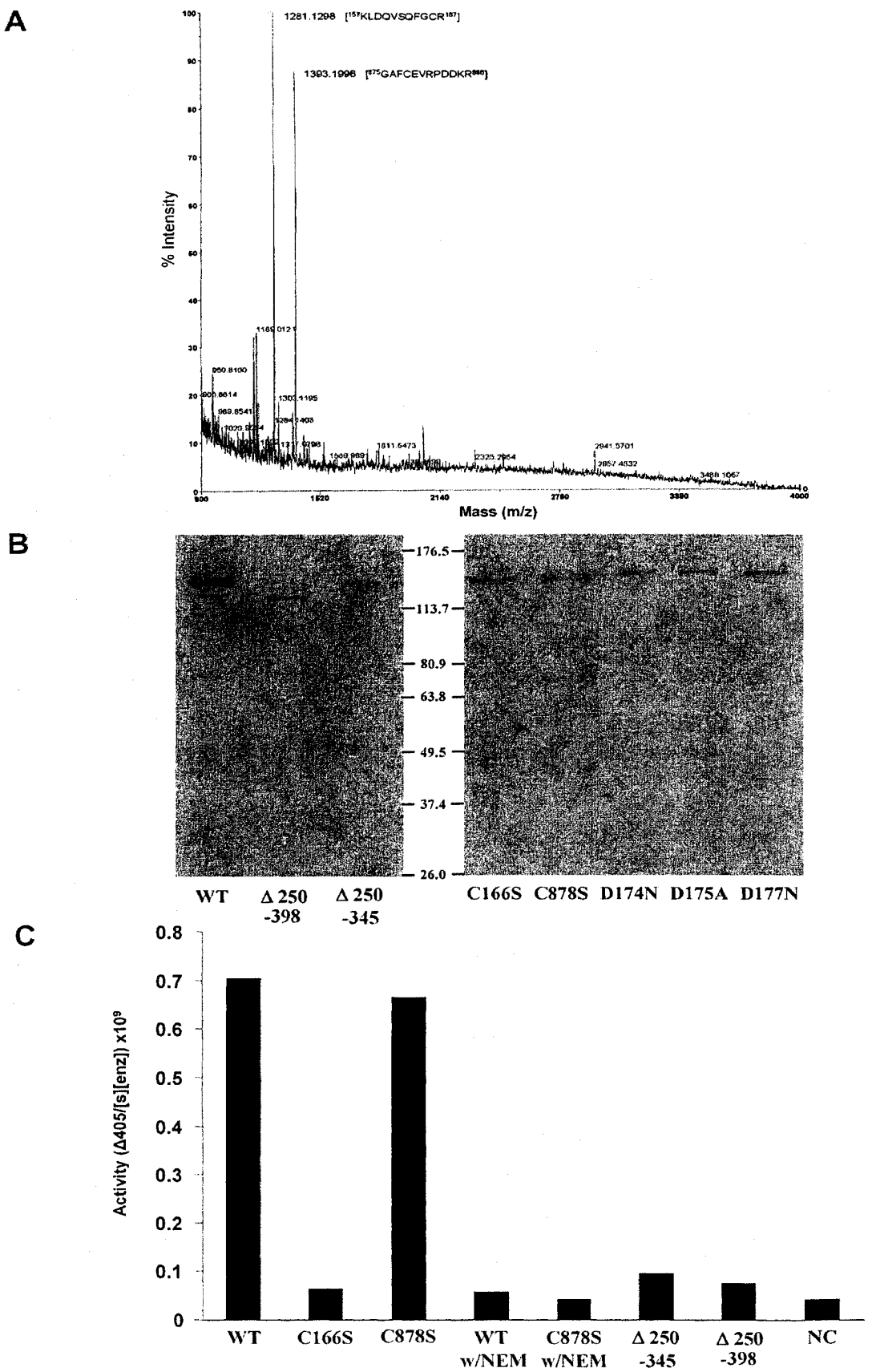
Results

The Active Site For O-GlcNAcase Activity Lies in the Enzyme's N Terminus—Previous characterization of the neutrally acting *O*-GlcNAcase had revealed that the enzyme's *in vitro* enzymatic activity can be inhibited to a substantial degree by the

sulfhydril alkylating reagent *N*-ethylmaleimide, indicating that there is a free, critical cysteine in or around the enzyme's active site (4). In order to identify this residue and zero in on the region of NCOAT that possesses the active site, we cloned and expressed the mouse NCOAT and treated it with a sulfhydril-specific reagent, biotinylated HPDP, to biotinylate any accessible free cysteine residue in the native enzyme. Treatment of NCOAT with this reagent, just as with *N*-ethylmaleimide, was able to inhibit enzymatic activity, suggesting again that, among others, the cysteine residue required for catalysis was modified (data not shown). A tryptic digest, followed by affinity chromatography with avidin, allowed for the purification of those peptides that had previously contained a free cysteine thiol. The HPDP-cysteine bond, being a disulfide linkage, is readily reversible with dithiothreitol. The resulting purified, free peptide mixture was analyzed by matrix-assisted laser desorption/ionization mass spectroscopy and revealed the presence of two peaks corresponding to peptides between residues 157-167 and 875-886 in the amino acid sequence of mouse NCOAT, both containing a single cysteine (Fig. 1*A*). Since mouse NCOAT contains 19 cysteines, at least one additional non-disulfide-bonded cysteine must exist. However, since this residue(s) is not readily accessible to HPDP, it was deemed equally unreachable by *N*-ethylmaleimide and was therefore ignored.

The cysteine residues found in these two peptides, Cys^{166} and Cys^{878}, were independently mutated to serines. The wild type and mutant proteins were then cloned and expressed, with expression and purity confirmed by Coomassie staining (Fig. 1*B*). The proteins were then used in a neutral hexosaminidase assay utilizing the colorimetric compound *p*NP-GlcNAc as substrate. The mutation of Cys^{878} to serine (C878S) gave rise to an enzyme with level of *O*-GlcNAcase activity comparable with that of the wild type enzyme, as measured by the spectrophotometric absorbance of the liberated *p*-nitrophenol

FIG. 1. **Relative enzyme activities of the Cys-to-Ser mutants and splice variants of NCOAT.** *A*, mass spectra showing the two isolatable peptides with free cysteines. *B*, the expression and purity of glutathione *S*-transferase-tagged NCOAT proteins were confirmed by Coomassie staining. Wild type enzyme (*WT*). *C*, *O*-GlcNAcase activity was tested by incubation at 37 °c for 30 min with the colorimetric substrate *para*-nitrophenyl-GlcNAc. Each reaction was terminated with 0.5 M sodium carbonate buffer and activity was measured as a change in absorbance by spectrophotometer at 405 nm. Each unit is described as the amount of *para*-nitrophenol cleavage when 1 mmol of *para*-nitrophenyl-GlcNAc is used in the presence of 1 μmol of enzyme. *para*-nitrophenol cleavage in the absence of any enzyme is used as negative control (*NC*). The wild type (*WT*) and C878S mutant were also pretreated with 10 mM *N*-ethylmaleimide (*NEM*) for 30 min at 37 °c prior to the assay.



group at 405 nm (Fig. 1*C*). In contrast, the C166S mutant displayed little to no activity over that of the substrate alone in the absence of enzyme. We concluded therefore that Cys^{166} is the target residue when sulfhydril reagents are used to inhibit *O*-GlcNAcase function. Additional evidence is provided by the fact that *N*-ethylmaleimide can still inhibit the enzymatic function of the C878S mutant protein, where the only remaining, accessible free cysteine thiol is that at Cys^{166} (Fig. 1*C*).

We have previously reported the identification of two naturally occurring splice variant isoforms of NCOAT that retain acetyltransferase activity (22). The first results in a protein missing amino acids 250-345 due to the removal of exon 8, and the other lacks amino acids 250-398 in the protein due to the specific deletion of exons 8 and 9. These variant proteins were cloned; expressed (Fig. 1*B*); and similarly assessed in the *O*-GlcNAcase assay, with their relative activities compared to that of the full-length enzyme. As shown in Fig. 1*C*, neither variant possessed any appreciable hexosaminidase activity over the background, and this loss of activity could not be overcome with increasing substrate concentration. The lack of enzymatic activity in these two constructs may be the result of a misfolded product or may be due to the absence of a modification requisite for activity. However, the N termini of NCOATs across species (between amino acids 54 and 391 for mouse NCOAT) are conserved; in addition there is significant homology between amino acids 63 and 331 of NCOAT and the Mu toxin of *Clostridium perfringes*, a hyaluronidase precursor. Therefore, it is quite likely that these two variants lack enzymatic activity due to the absence of a portion of their active site. Together, these data strongly suggest that the *O*-GlcNAcase active site can be assigned to a region between the cysteine residue at amino acid 166 and the common missing sequence in the two variants tested, residues 250-345. At this time, we have not explored the outer

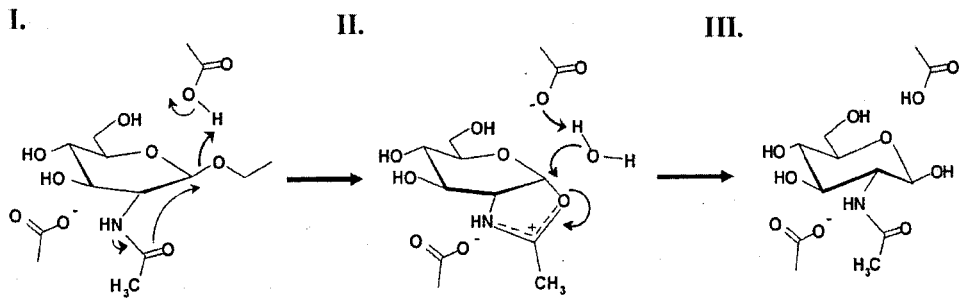
boundaries of the active site or the minimal region therein that is necessary and sufficient for full activity.

Mutation of the Aspartic Acid Residues at Positions 174, 175, and 177Impair in Vitro O-GlcNAcase Activity—Since NCOAT shares a core substrate, *N*-acetylglucosamine, with chitinases, hexosaminidases, and hyaluronidases, we speculated that NCOAT shares the catalytic mechanism that is invariant among these family 18, 20, and 56 glycosidases (25, 28, 29). The mechanism is depicted in Fig. 2*A*. In order to determine the sites of the residues required for catalysis, the first being the acid/base and the second being responsible for orienting the 2-acetamido group, a number of previously characterized glycosidases were examined (30-32). In each of these enzymes, it was individually reported, the residue responsible for orienting the *N*-acetyl nucleophile was an aspartic acid, and the catalytic acid/base was either an aspartate or a glutamate; these two key carboxylates are separated by a single hydrophobic amino acid (ψ). Examination of a large number of yet uncharacterized chitinases and hyaluronidases revealed the presence of this same pattern, a catalytic doublet with the motif D-ψ-D/E (Fig. 2*B*). This doublet is not present in hexosaminidases; instead the two essential carboxylates are found adjacent to one another in the primary sequence (33, 34). The significance of this difference is not yet fully understood.

The D-ψ-D/E motif not surprisingly can be found in each representative NCOAT examined (Fig. 2*B*). The motif consists of two aspartic acids at positions 175 and 177 in the mouse and rat NCOAT, directly within the putative active site and immediately downstream of the cysteine residue demonstrated above to be involved in catalysis. We performed site-directed mutagenesis to address the possibility that these two residues are

FIG. 2. **Residues essential for *O*-GlcNAcase catalysis.** *A*, acid-base substrate-assisted catalytic mechanism for family 18, 20, and 56 glycosidases. Two carbonyl-containing amino acids are ideally situated in the active site to allow both the substrate and one water molecule to enter between them (*I*). The first acts as a general acid, donating a proton to the oxygen on the C1 carbon and cleaving the glycosidic linkage. The second orients the *N*-acetyl group into position for nucleophilic attack at the anomeric center, resulting in the formation of a high-energy oxazolinium ion intermediate (*II*). The transition state is then hydrolyzed when the first carbonyl abstracts a proton from an incoming water molecule, with the resulting hydroxyl ion attacking the anomeric center, hydrolyzing the bicyclic intermediate, and liberating the free GlcNAc (*III*). *B*, mapping the residues essential for catalytic activity in family 18 chitinases and family 56 hyaluronidases. The position of the D-Ψ-D/E motif is shown for several representative chitinases and hyaluronidases. The location of the nucleophile-positioning residue is represented by ● and that of the catalytic acid/base by ■. Those enzymes where the roles of these residues have been confirmed are indicated with an *asterisk*; all other enzymes were chosen arbitrarily to demonstrate the presence of this motif among other family members. The presence of this motif is also shown in select NCOATs. The site of the previously mutated cysteine is shown with a ◆. *C*, relative *O*-GlcNAcase activity of wild type and aspartic acid mutant enzymes at pH 7.0. Wild type and mutant NCOAT were expressed and assayed at pH 7.0 and 37 °c for 30 min as described.

A



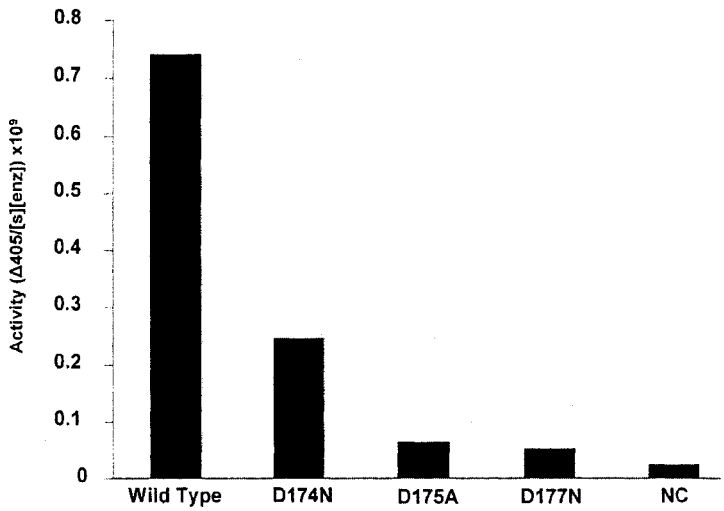
B

Chitinase:	● ■
Manduca Sexta *	^{134}KKYDFDGLDLDWEYPGAADRGGSFSDKDKFLYLVQEL^{170}
Serratia marcescens *	^{131}KDYGFDGYDIDWEYPQAAEYDGFIAALQEIRTLLNQQ^{167}
Saccharomyces cerevisiae	^{211}FRLGFDGIDLDWEFPGNNESEPRGYLKLVRMLRLKLN^{247}
Coccidioides immitus *	^{159}KDLGFDGIDIDWEYPEDEKQANDFVLLLKACREALDA^{195}
Caenorhabditis elegans	^{167}RTWGFDGIDIDWEYPSGATDMANYVALVKELKAACES^{203}
Aeromonas caviae *	^{303}TWKFFDGYDIDWEFPGGQGANPSLGGPNDGATYVVLM^{339}
Hevea brasiliensis *	^{141}GDAVLDGIDFDIEHGSTLYWDDLARYLSAYSKQGKKV^{177}

Hyaluronidase:	● ■
Macaca fascicularis *	^{136}PVDNLGMAVIDWEEWRPTWARNWKPKDVYKNRSIELV^{172}
Caenorhabditis elegans	^{123}DENFNGIAVIDIEEFRPMWELSWGPFSVYKTESIRLT^{159}
Mus musculus	^{117}GSSFAGLAVLDWEEWYPLWAGNWGPHRQVYLAASWVW^{153}
Apis mellifera *	^{101}DKSFPGVGVIDFESWRPIFRQNWASLQPYKKLSVEVV^{137}
Xanthamonas campestris	^{287}GVRSFYVAFDDIEYADEAAFGPSGEQAAATAQAKLLN^{328}
Clostridium perfringens	^{262}GVRSFAILWDDIENRSGVQQAEVLNRFNKEFIKNKEG^{298}
Enterococcus faecium	^{246}GVFQFGLLMDDIDYQLKGAAKRRFREPAFAHAYLVNR^{282}

NCOAT:	◆ ● ■
Drosophila melanogaster	^{123}GCEAYALLFDDAESELSKADKEVFQTFANAHVSVTNE^{159}
Anopheles gambiae	^{109}GCKAFALLFDDIEPEMSKPDKEVFQSFAHAQVSVTNE^{145}
Mus musculus	^{165}GCRSFALLFDDIDHNMCAADKEVFSSFAHAQVSITNE^{201}
Caenorhabditis elegans	^{123}GCDSFAVLFDDIEVQMQDEDQKQFTSFAHAQVHIANT^{159}

C



important for catalysis. Asp^{177} was changed to an asparagine, a conservative mutation that removes the charge and is non-ionizable. Asp^{175} was replaced with an alanine rather than an asparagine, however, due to previous reports suggesting that both the size and the charge of this residue are implicated in its role of orienting the 2-acetamido group for nucleophilic attack (Fig. 2*A*). A D-to-A mutation eliminates both of these features. In addition, a third aspartic acid, Asp^{174}, was also changed to an asparagine. Since this is also a conserved residue across NCOATs (Fig. 2*B*), it may also be involved in catalysis; however its mutation should exhibit an effect much less debilitating than that of the proteins with mutations in the two catalytically essential residues.

The mutant enzymes were similarly cloned and expressed, with purity again visualized by Coomassie staining (Fig. 1*B*). Each enzyme was then used in the hexosaminidase assay under the same neutral conditions. Under these conditions, the D175A and D177N mutants were nearly devoid of activity; however, the D174N mutant was able to retain activity, although at nearly one-quarter that of the wild type enzyme (Fig. 2*C*). Each mutant protein retained full histone acetyltransferase activity (data not shown). These results are indicative of the relative importance of each residue in *O*-GlcNAcase catalysis. Both Asp^{175} and Asp^{177} are absolutely critical, while Asp^{174} plays a significant but not essential role.

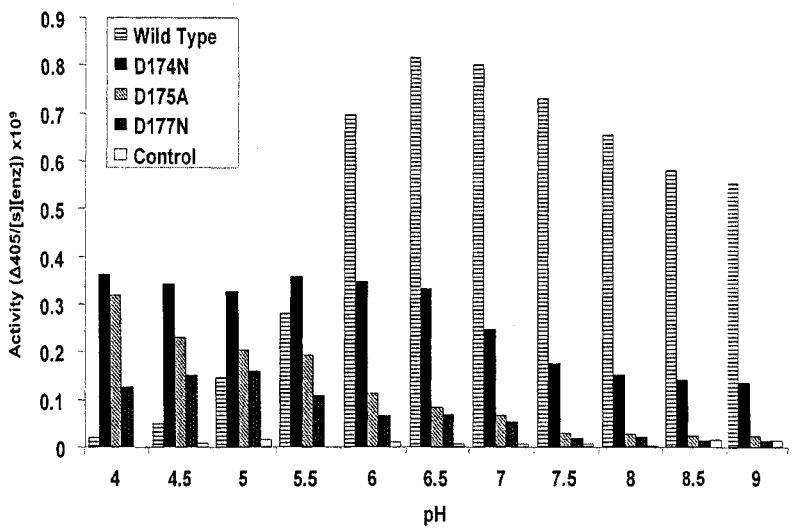
pH Profile Analysis of the Wild Type and Mutant Enzymes—The pH profiles of the wild type and each aspartic acid mutant NCOAT were determined by the measurement of the *para*-nitrophenol cleavage from *p*NP-GlcNAc in the hexosaminidase assay in buffers at the indicated pH values. The results are shown in Fig. 3*A*. The pH optimum for the wild type enzyme was found to be between pH 6.5 and 7.0, in agreement with previ-

ous characterization of NCOAT's being a neutrally acting enzyme (4). The D174N mutant maintained a reduced level of activity at the wild type enzyme's optimal pH as well as in the basic pH range. However, it achieved greater levels of activity than the wild type did at more acidic pHs and exhibited a shift in pH optimum to the pH 5-5.5 range. Proteins evolvespecialized environments surrounding their catalytically functional groups, most often to maintain a particularly charged state so that the enzyme can function at its optimal pH (35). Asp^{174} is likely involved in the maintenance of this micro-pH environment so that *O*-GlcNAcase activity is optimal at a neutral pH. The two catalytically essential aspartic acids must remain properly charged in order to carry out their roles when substrate enters. The first, Asp^{175}, must be unprotonated to assume a basic role and hydrogen bond the amide proton to facilitate the formation of the oxocarbenium intermediate; the second residue, Asp^{177}, must be protonated to serve the initial acid function (29, 36). We therefore propose that Asp^{174} plays a role in mediating the charges in the active site microenvironment, ensuring the enzyme is ready to function when substrate enters.

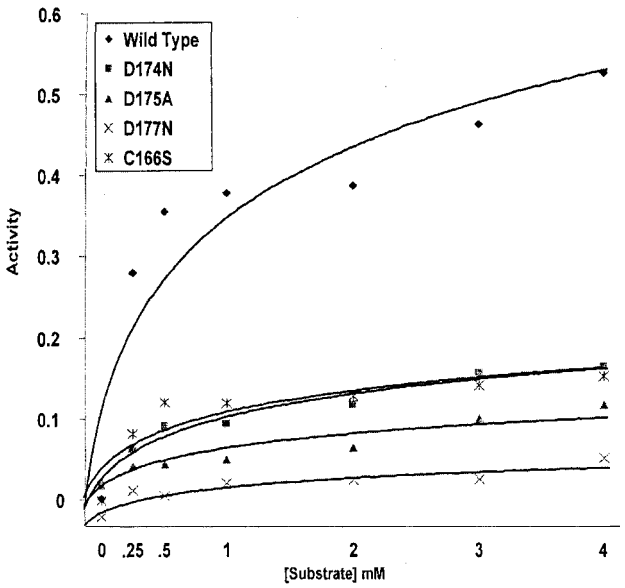
The D175A and D177N mutants, catalytically defunct at neutral pH, both regain an appreciable level of activity in pH solutions below 5.5, displaying their importance in maintaining positive charges for catalysis. The result is predictable for the D177N mutant since it acts as the catalytic acid in the first step of the reaction, the hydrolysis of the sugar linkage (Fig. 2*A*). The replacement of an alanine for the aspartate at position 175 may affect the charge of the Asp^{177} side chain, thus altering the pH optima for this mutant enzyme as well. These two mutants also effectively eliminate the basic limb of the pH profile seen with the wild type enzyme. These shifts in pH profile serve as confirmation that each residue plays a catalytic role. Interestingly, rather than having a bell-shaped curve indicative of the ionization of two residues in the active site, which is seen in a

FIG 3. **pH profiles and kinetic analyses for wild type and mutant NCOAT.** *A*, catalytic activity for wild type and mutant NCOATs were measured in 50 mM sodium cacodylate buffer with 5% bovine serum albumin ranging from pH 4.0 to 9.0 using *para*-nitrophenyl-GlcNAc as substrate as described in "Experimental Procedures." *B*, Michaelis-Menten kinetics of hydrolysis of *para*-nitrophenyl-GlcNAc by wild type and mutant NCOAT. Wild type and mutant NCOAT were independently incubated with the indicated increasing concentrations of *para*-nitrophenyl-GlcNAc. *C*, kinetic parameters of wild type and mutant NCOAT with *para*-nitrophenyl-GlcNAc (*pNP-GlcNAc*) substrate [a].

A



B



C

Enzyme	V_{max} (% of WT)	K_m (mM) [a]	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\ mM^{-1}$)
Wild Type	100.0%	.21 ± .01 [b]	72.5 ± 9.6	345 ± 79.4
Asp^{174}→Asn	33.3%	.48 ± .05	23.1 ± 1.6	48.1 ± 11.1
Asp^{175}→Ala	23.7%	.87 ± .09	16.3 ± .55	18.7 ± 3.5
Asp^{177}→Asn	10.3%	1.7 ± .18	7.9 ± .7	4.6 ± 1.2
Cys^{166}→Ser	30.5%	.19 ± .01	28.9 ± 1.9	152 ± 46.7

[a] for pNP-GlcNAc substrate

[b] standard deviation

large number of glycosidases (29), the pH profile of the wild type enzyme has a discernable shoulder in the basic pH range. This phenomenon is typically associated with the existence of a third ionizable residue in the enzyme's active site (29, 37), consistent with our finding that Asp^{174} is involved in catalysis in addition to Asp^{175} and Asp^{177}.

Kinetic Analysis of the Wild Type and Mutant Enzymes—To further characterize the exact roles of these three aspartic acid residues in catalysis as well as that for the Cys^{166} residue, the kinetic parameters of each enzyme were determined. The results of a Michaelis-Menten plot using increasing *p*NP-GlcNAc substrate are shown in Fig. 3*B*. The resulting data were used to determine the V_{max}, K_m, and k_{cat} values for the wild type and each mutant enzyme (Fig. 3*C*). Our model based on the characterization of similar enzymes predicts that both Asp^{175} and Asp^{177} should be integral in both substrate binding and catalysis. Indeed, the proposed *N*-acetyl-orienting residue, when changed to an alanine, has a 4-fold increase in K_m and an 80% decrease in k_{cat} when compared to those values for the wild type enzyme. The predicted acid/base catalyst Asp^{177}, when mutated to an asparagines, has a K_m eight times that of the wild type and a decrease in k_{cat} to 10% of the wild type. These changes can be seen on varying scales with a select number of analogous mutations in several similar enzymes (32, 38, 39). While Asp^{175} and Asp^{177} are required for both substrate binding and catalysis, the aspartic acid mutant D174N possessed only a 2-fold difference in K_m value from that of the wild type, demonstrating that substrate binding is only slightly affected with this mutation. We believe that, since Asp^{174} is involved in maintaining Asp^{175} and Asp^{177} in their unprotonated and protonated forms, respectively, and since these residues must be in these states in order to allow substrate entry (29, 36), it is likely that the mutation of Asp^{174} to an asparagine results in an

impairment of the enzyme in keeping these residues in their substrate-ready state. Therefore, we postulate that the small K_m effect seen in the D174N mutant is due to an indirect effect on substrate binding and is not necessarily the result of Asp^{174} needing to make specific substrate contacts. The C166S mutant did not show an appreciable change in its K_m, indicating that it plays no role in substrate binding. Each of these latter two mutants had a generously lower k_{cat} value (32% and 40% that of the wild type, respectively), however, indicating that both Asp^{174} and Cys^{166} again play a role in catalysis but are not an absolute requirement for activity.

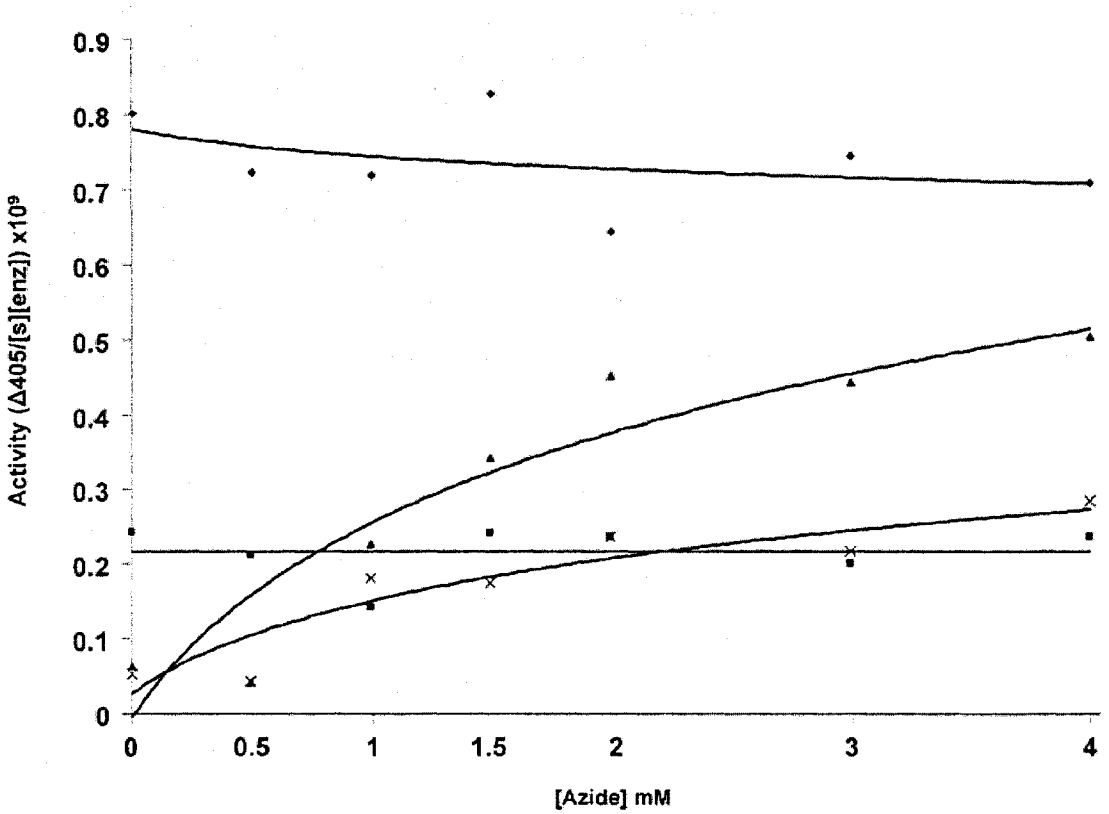
Chemical Rescue of the D175A Mutant NCOAT—We hypothesize that Asp^{175} has a role in assisting the *N*-acetyl group for nucleophilic attack and the generation of an oxazolinium ion intermediate. If this hypothesis is correct, the D175A mutant enzyme might be chemically rescued with the addition of an exogenous nucleophile. In this test, the addition of a strong anion should allow the D175A mutant enzyme to recover an appreciable level of activity while exhibiting little effect on the wild type and other mutant proteins. This phenomenon has been documented previously for analogous mutations in similar retaining glycosidases (29, 37, 40). The transition state can still be achieved when the acid/base residue is present; however, the reaction's completion is impaired in the D175A mutant. The chemical rescue is therefore due to the anion species attacking the anomeric carbon of the oxazolinium intermediate (Fig. 2*A*, *step II*) without the need for a general base (29, 37). The addition of the exogenous nucleophile increases the enzyme's k_{cat} because it is the second catalytic step, the hydrolysis of the oxazolinium ion intermediate, which is rate-limiting (29, 37).

Because sodium azide is small, it may fit into the vacant anion binding site left by the change of the aspartate carboxylate to the methyl group of alanine (37). As predicted, the addition of any concentration of NaN_3 to the reaction mixture had no effect on the wild type enzyme when compared to reactions carried out in the absence of azide (Fig. 4). Similarly, the D174N mutant activity remained unchanged when NaN_3 was added. The D177N mutant, the acid/base residue in the reaction, did regain up to a 5-fold level of activity in a dose-dependent manner with the addition of sodium azide, with the maximal effect at 2 M or greater. The activity increase is the result of the strong negative charge of N_3^-, which as stated above can replace the need of the hydroxyl in attacking the anomeric carbon to reopen the transition state. The first step of the reaction, the acid attack on the C1 oxygen, is still impaired in the D177N mutant, but may proceed to some extent if a good anomeric leaving group, like *para*-nitrophenol, is present (37). Thus the 4- to-5-fold recovery of activity is significant but is still only one-fourth that of the wild type enzyme. The partial recovery of activity then serves as additional support for the notion that Asp^{177} is the catalytic acid/base for *O*-GlcNAcase activity.

The major effect of the exogenous nucleophile, azide, was on the D175A mutant. This recovery with azide was also dose-dependent, with an 8-fold rescue of activity seen with the addition of 4 M azide. The azide gives a 70% recovery of activity to the D175 mutant when compared to the wild type enzyme. The chemical recovery of such a significant amount of activity in the D175A mutant in the presence of an exogenous nucleophile serves to confirm Asp^{175}'s role in mediating the nucleophilic step and the generation of the oxazolinium ion intermediate to be hydrolyzed during catalysis.

Incubation with NCOAT Makes STZ a More Potent Inhibitory Compound— Inhibitors have become invaluable tools in the study of catalytic mechanisms in glycosidase

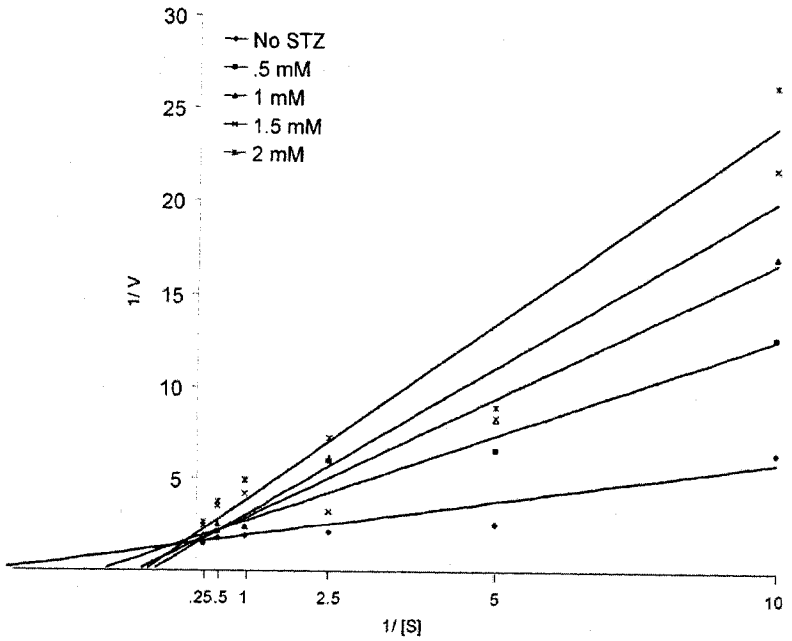
FIG. 4. **Activity of wild type and mutant NCOAT in the presence of increasing concentrations of azide.** The activity of the wild type (♦), D174N (■), D175A (▲) and D177N (✕) mutants was measured as a function of *para*-nitrophenol hydrolysis in the presence of the indicated concentrations of sodium azide.



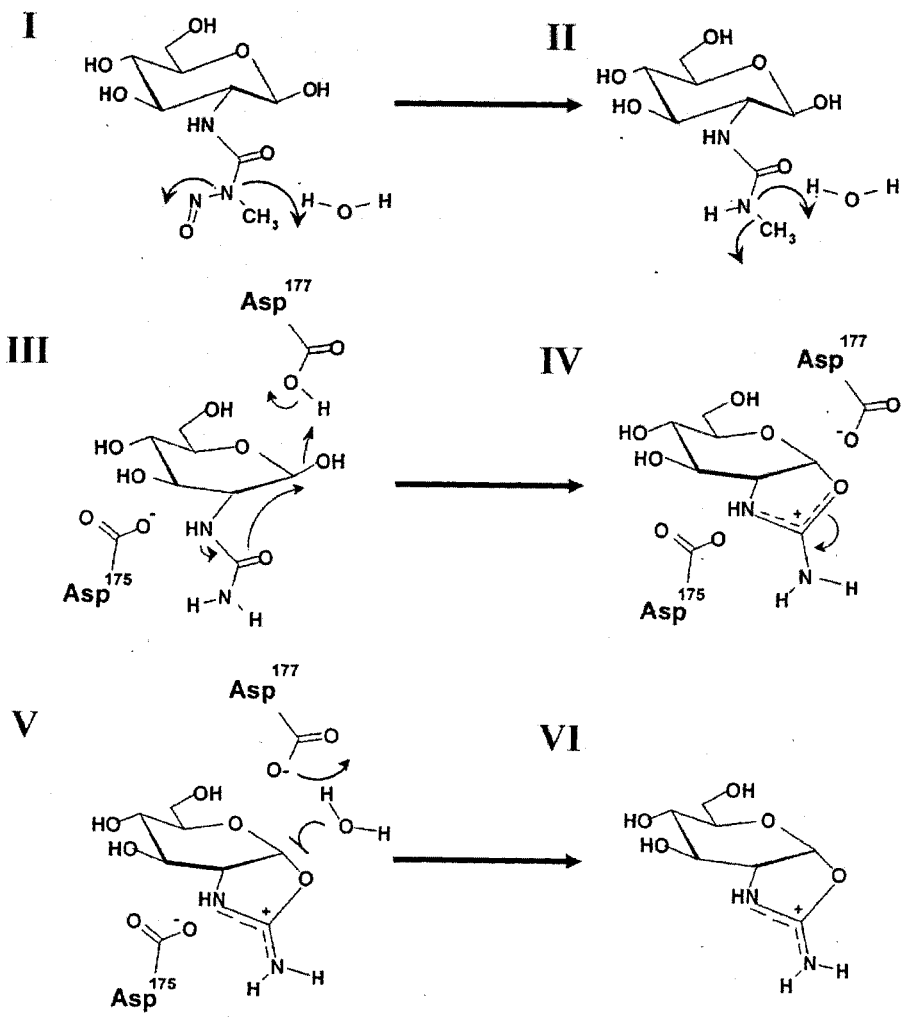
enzymes, the most potent of which are compounds that mimic the enzyme's transition state. These analogs bind with higher affinity within the active site and competitively inhibit the enzyme. For example, allosamidin, a pseudotrisaccharide, is a transition-state mimetic and a selective inhibitor of family 18 chitinases (27) also NAG-thiazoline is a potent inhibitor of lysosomal β-hexosaminidases because it simulates the oxazoline reaction intermediate (26). Both have been used gainfully to provide evidence for the substrate-assisted mechanism proposed for their respective enzymes and as a means for confirming the intermediate structure. The selective *O*-GlcNAcase inhibitor, STZ, displays classic competitive inhibition kinetics (Fig. 5*A*). Since STZ is a GlcNAc analog, it or a derivative of STZ may target the *O*-GlcNAcase active site of NCOAT and become a transition state analog. When STZ is dissolved in water, it rapidly and spontaneously liberates its NO group (≤5 min) (41), a reaction that can be observed by various methods, including mass spectroscopy. In addition, STZ is a known methyl donor (42), releasing the terminal CH_3 from the native compound at some point during incubation, either prior to exposure or in the presence of the active site. Regardless of the time of its release, the compound isolated from the enzyme contains no NO or CH_3. NO and CH_3 are replaced by protons (Fig. 5*B*, *steps I, II*). If the enzyme attempts to catalyze STZ as if it were normal substrate, the catalytic mechanism displayed in Fig. 2*A* predicts that the reaction is not favored to advance past the transition state, as depicted in Fig. 5*B*. That is, Asp^{177} will donate a proton to the oxygen atom on the C1 carbon, liberating a water molecule. The oxygen formerly in the *N*-methyl-nitrosourea chain of the STZ attacks the C1 carbon; the partial charges observed in the normal reaction's transition state (Fig. 2*A*, *step II*), can now be transferred between the neighboring amines (Fig. 5*B*, *step IV*), which have much higher nucleophilic character than the ether oxygen (or the methyl group in the natural

FIG. 5. **Streptozotocin as an inhibitor of *O*-GlcNAcase activity.** *A*, Lineweaver-Burk plot demonstrating STZ as a competitive inhibitor of *O*-GlcNAcase activity. *B*, hypothetical reactions and products when STZ is incubated with NCOAT. When STZ is dissolved in H_2O, NO is lost and replaced by a proton (*I*). STZ also releases its CH_3 (*II*). STZ is targeted to the *O*-GlcNAcase active site, where the enzyme attempts to catalyze it. As in Fig. 2, the active site acid donates a proton to the oxygen atom on the C1 carbon; also the oxygen atom of the urea side chain performs a nucleophilic attack on the anomeric center, generating a bicyclic intermediate (*III*). The partial charge on the oxygen is transferred between the neighboring amines (*IV*). When the active site base now deprotonates H_2O, the rate of attack by the resulting hydroxyl on the C1 carbon is reduced since charge on ether oxygen is not favored (*V*). The resulting product mimics the natural ligand transition state (*VI*).

A



B

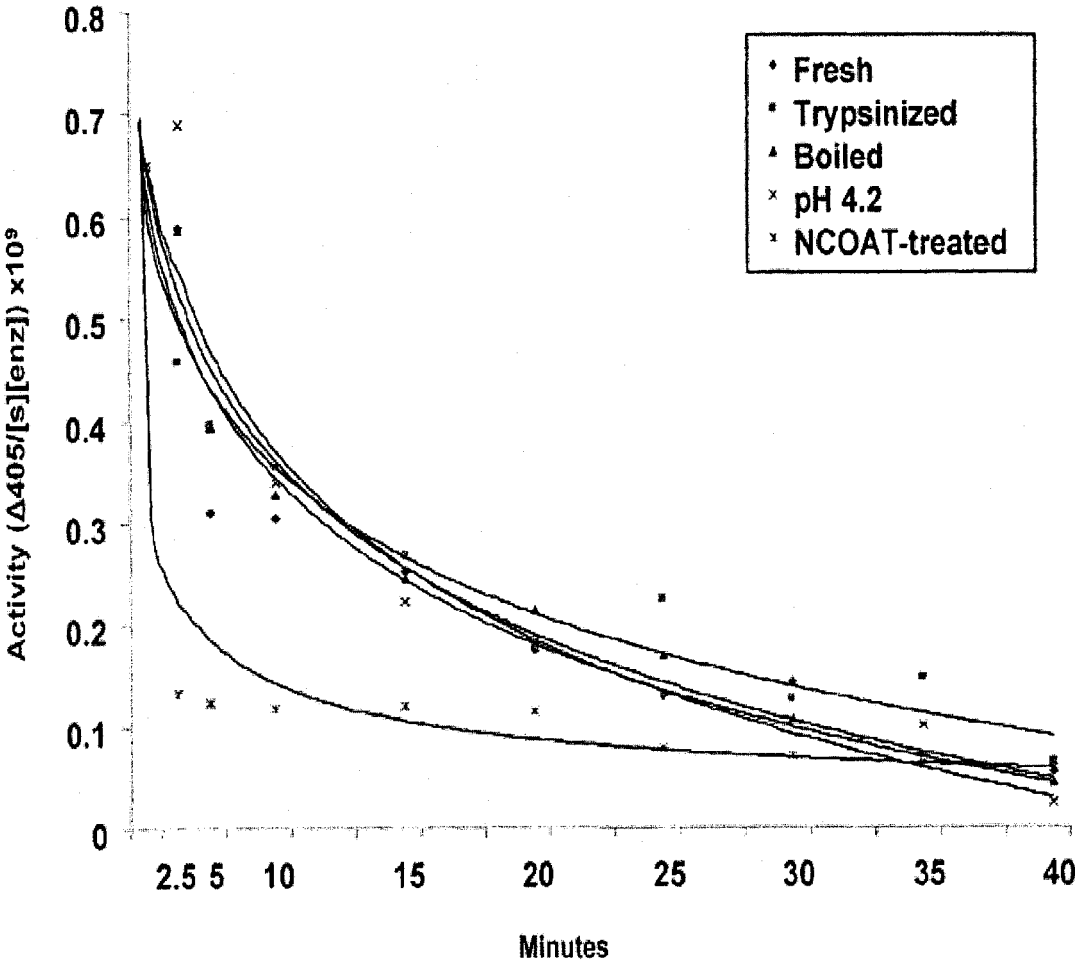


substrate) does. The compound is more stable as a result because hydroxyl attack on the C1 carbon will only break the C1-O bond when the less basic ether can draw the charge from its preferred resonance between nitrogen groups (Fig. 5*B step V*). If Fig. 5 is correct, then the resulting enzyme-assisted product should resemble the natural ligand transition state (with an NH_2 taking the place of a CH_3), resulting in a withdrawal of charge on the oxygen atom. Such an analog would be catalyzed to completion at a much slower rate and would thereby inhibit the enzyme. More importantly, the stability of this compound could allow for the isolation and validation of Fig. 5.

Withers and colleagues observed such a phenomenon with the inhibitory reactant and NAG-thiazoline product of β-hexosaminidase (26). We incubated 0.1 M STZ with resin-bound NCOAT, washed the unbound compound from the protein, and released the bound product from the trypsin-digested NCOAT. The resultant compound, without NO and CH_3 (Fig. 5*B*), could inhibit a fresh batch of NCOAT in 2 min at a concentration estimated to be less than 1 μM; however, it took 40 min to fully inhibit activity using up to 100 mM fresh STZ (10^5 times more potent) (Fig. 6). Trypsinized NCOAT incubated with fresh enzyme had no effect on activity (data not shown). Whether STZ was boiled, treated with trypsin, or preincubated in an acidic buffer, the compound had the same inhibition kinetics that fresh STZ had when it was then incubated with NCOAT prior to the assay (Fig. 6).

STZ Inhibits O-GlcNAcase Activity by Being an Enzyme-assisted Transition State Analog—The stability of the putative transition-state compound allowed us to determine its molecular mass. The same reactions were repeated, but the compound released after incubation with enzyme was analyzed by mass spectroscopy. As shown in Fig. 7*A*, fresh STZ has a molecular weight of 265.2 (observed as 266 by mass spectroscopy). STZ with

FIG. 6. **Rate of *O*-GlcNAcase inhibition using fresh versus enzyme pre-treated Streptozotocin.** The time course of inhibition using fresh STZ is indicated by ◆, STZ treated with trypsin only by ■, Boiled STZ by ▲, STZ incubated in acidic buffer to liberate NO by **X**, and STZ incubated with NCOAT and then liberated with trypsin and treated with fresh enzyme by ж. Trypsin-treated enzyme alone had no effect on future samples.



water loss is also observable [molecular weight (MW) 248]. STZ dissolved in water only or subjected to trypsin treatment yielded a subfraction of compound with an MW of 237, consistent with NO loss (Fig. 7, *B*, and *C*). Since trypsinization is carried out under basic conditions, which would deter H^+-induced NO release from STZ (43), we presume the NO-minus fraction to be a light- or time-dependent product or the result of the reaction temperature (37 °c), regardless of the alkaline conditions. We could not detect CH_3 loss (252 or 223 MW) by mass spectroscopy under these conditions, so conclusions about exactly when it is liberated could not be determined; however, as stated above, the inhibitory compound isolated from the enzyme had no detectable CH_3 group (Fig. 7*D*, Fig. 8, *B* and *D*). When STZ was incubated with NCOAT and then liberated by trypsin, the dominant peak was 206 MW (Fig. 7*D*), as predicted for the transition-state analog shown in Fig. 5, *step VI*. In addition, the native 266 peak is reduced to background, consistent with the proposal that the enzyme is in fact catalyzing the original STZ compound. The two point mutants, D175A and D177N, which are virtually incompetent catalytically, could not generate a peak corresponding to 206 MW when incubated with STZ. While unbound STZ is washed away, a small population of bound yet uncatalyzed ligand is present. Though they were present in a reduced quantity, the primary species observable with the spectrometer were that of the native compound and that without H_2O (Fig. 7, *E* and *F*). These data reinforce the notion that NCOAT itself is responsible for converting STZ to a new species with a molecular mass compatible with the reaction shown in Fig. 5.

Structural data using NMR were obtained on the 206-MW compound formed by the enzyme from STZ to assist our MS deductions. Again, the same reactions were completed, but scaling up required that the enzyme be denatured by heating. Boiling did re-

FIG. 7. **Molecular weight changes of streptozotocin after incubation with NCOAT.** *A*, mass spectra of fresh STZ, with its typical ionization products. *B*, STZ treated with trypsin only. *C*, STZ boiled for 5 min. *D*, mass spectra observable when STZ is incubated with NCOAT and then liberated. *E*, the mass spectra of STZ incubated with the D175A mutant. F, spectra after incubation with the D177N mutant NCOAT.

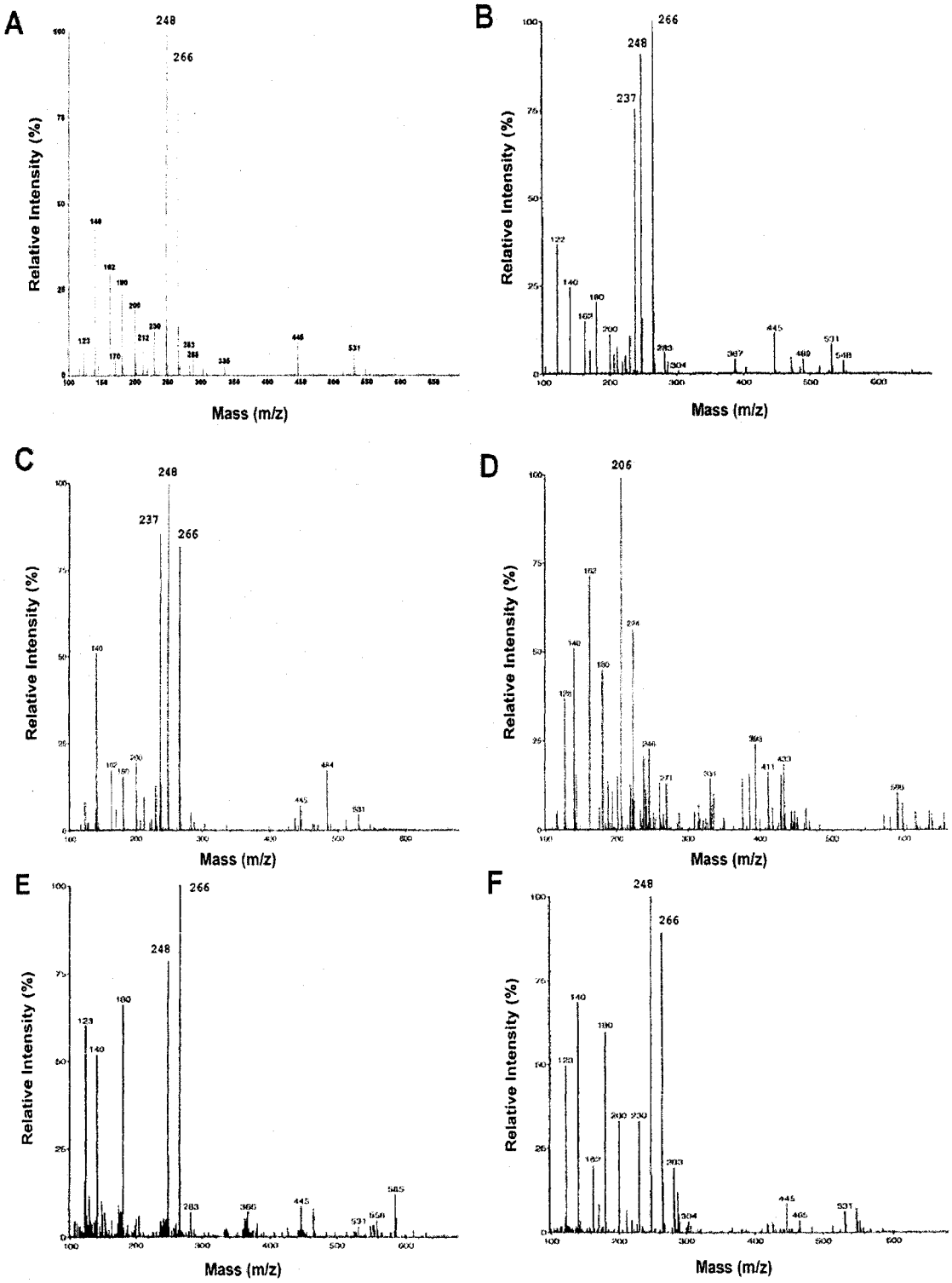
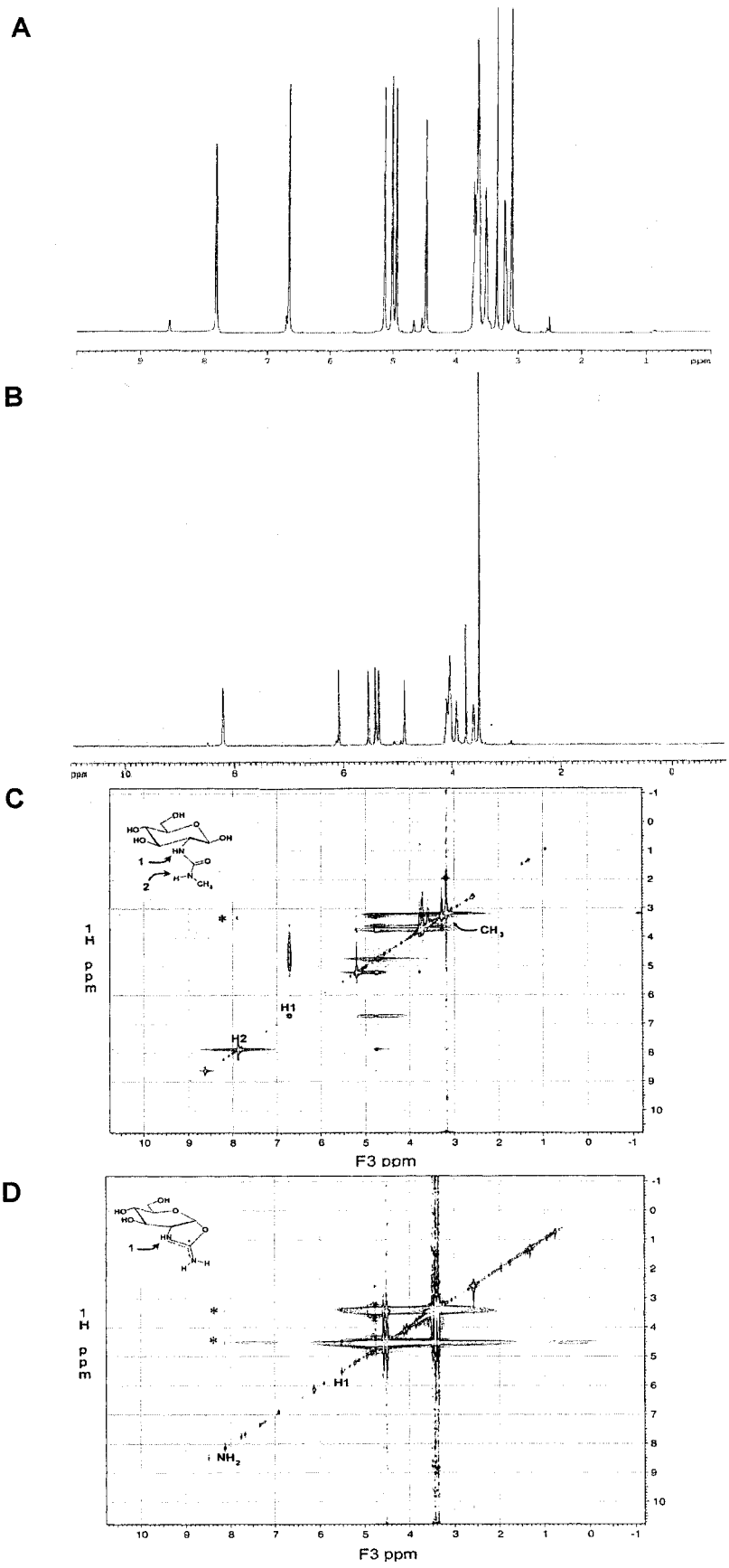


FIG. 8. **NMR structural analysis after streptozotocin is incubated with NCOAT.** *A,* one-dimensional NMR spectra for fresh STZ. *B*, one-dimensional NMR spectra for STZ incubated with NCOAT and then liberated with heat denaturation. *C*, two-dimensional NMR spectra for fresh STZ. *N*-methylurea protons are indicated with the accompanying *inset*. *D*, two-dimensional NMR spectra for STZ incubated with NCOAT and then liberated. Amine protons are indicated with the accompanying *inset*. Cross-talk peaks of interest in the two-dimensional spectra are indicated by an *asterisk*.



sult in the rapid release of NO (Fig. 7*C*) but otherwise gave results nearly identical to those seen for the trypsin spectra seen in Fig. 7*B*. Since there is not much difference between the starting and the enzyme-treated 206-MW compounds with respect to the positions of the exchangeable OH and NH protons, the differences in the NMR spectra would be expected to be subtle. Given the structure of the two compounds, however, two distinct observations could be anticipated. First, the 206-MW species would have a clearly discernable disappearance of the CH_3 peak. Second, the nitrogen-bound protons on the *N*-methyl-urea side-chain would shift their respective positions as a result of their interaction with a new environment of protons. Fig. 8 shows the one- and two-dimensional NMR spectra of the starting and enzyme-treated compounds. A strong methyl peak is observable at 3.1 ppm in the fresh STZ sample (Fig. 8, *A* and *C*). This peak is clearly absent in the sample treated and liberated from NCOAT (Fig. 8,*B* and *D*). The peaks representing the sugar ring, between 5.1 and 3.3 ppm in each sample, also shift accordingly. The nitrogen-bound protons visualized at their respective 6.7- and 7.8-ppm positions in the untreated sample (peak assignments determined by one-dimensional NMR) markedly shift to 6.2 and 8.2 ppm in the STZ sample analyzed after treatment with NCOAT. In addition, the proton that replaces the NO group (H2 in panel *C*) exhibits cross talk with the methyl group in the untreated sample (indicated by an *asterisk*), as would be expected from its proximity (Fig. 8 *C*). This cross-talk peak is missing in the treated sample; however, the NH_2 protons begin to display interaction with other sugar ring protons (Fig 8*D*). These latter observations are consistent with these protons being drawn closer to the sugar ring during the enzyme-assisted bicyclization process. The NMR data are not conclusive alone; however, when taken with the MS data and functional enzymology, they

lead to the preeminent conclusion of the formation of the proposed transition state analog by the action of the enzyme on STZ (Fig. 5).

Discussion

While STZ inhibits *O*-GlcNAcase activity (5), its existence has been known much longer than that of *O*-GlcNAcylation. Therefore, many valid attributes of this relatively simple compound have been ascertained that may contribute to its diabetogenicity (42, 44, 45). Nevertheless, the inhibition of *O*-GlcNAcase may be the principle means of its biological action. For this to be the case, several important properties must be explained. First, STZ itself is not a potent inhibitor of *O*-GlcNAcase. Second, intact STZ exists transiently in solution, having a half-life of only minutes (43). However, the inhibition of the *O*-GlcNAcase can last at least a day *in vivo* (46). The observation that STZ is converted by the enzyme to a transition-state analog that is stable to heat treatment and trypsin digestion of the enzyme may explain these paradoxes of potency and stability.

The results of this study have allowed for the positive identification of the *O*-GlcNAcase active site region. The active site resides in the N-terminal third of the NCOAT protein; it is likely, given the conservation among NCOAT proteins and the strong similarity to the *Clostridium perfringens* hyaluronidase precursor, Mu toxin (BLAST score 97.4), that the entire *O*-GlcNAcase active site lies in a stretch between amino acids 63 and 283, a span that had previously been predicted to be the enzyme's active site in a review by Hanover (47). This likelihood not only explains why *N*-ethylmaleimide can inhibit *O*-GlcNAcase activity, since there is a free cysteine at position 166 within the active site, but would also explain why each splice variant protein tested was inactive in the *O*-GlcNAcase *in vitro* assay, since they both lacked a region

between residues 250 and 345 that partially overlaps the active site (Fig. 1). Interestingly, Structural Classification of Protein analysis predicts NCOAT to have a motif resembling a triose phosphate isomerase barrel starting at residue 64 (48). The triose phosphate isomerase α/β_8 motif is found in over 25 protein superfamilies, and its presence has been confirmed (routinely lacking the eighth β-strand) in the active site structures of family 18, 20, and 56 glycosidases (31, 32, 49). Secondary structure predictions can be used to identify the probable presence of each independent α- and β-motif starting from residue 64 and continuing to residue 359 (data not shown). Finally, this active site assignment is consistent with several other previous findings. First, an identified splice variant of this enzyme, which lacks the C-terminal third and the HAT domain of the enzyme, is able to remain active in hyaluronidase assays (50). A similar C-terminal truncated NCOAT was constructed and tested during these experiments and was found to retain one-third of the activity of the wild type (data not shown). Second, it had been found that NCOAT is a competent substrate for caspase-3 cleavage and that the resultant protein, processed into two nearly equal halves, is still functional in *O*-GlcNAcase assays (2).

To further characterize the active site, we have also examined the role(s) of candidate residues in the active site that have been shown to be involved in catalysis in analogous enzymes. Since NCOAT recognizes *N*-acetylglucosamine as its core substrate, its mechanism of catalysis is likely the same as or markedly similar to those of other enzymes which catalyze the hydrolysis of GlcNA,c such as chitinases (hydrolyze polymers of β-1,4-linked GlcNAc), hexosaminidases (remove β-1,4-linked GlcNAc from nonreducing ends of oligosaccharides), and hyaluronidases (cleave the glucuronic acid β-1,3-GlcNAc disaccharides in hyaluronan). The mechanism of catalysis is highly conserved among these enzyme families and requires the presence of two principally essential resi-

dues (Fig. 2). We noted that, among chitinases (family 18 glycosidases) and hyaluronidases (family 56), these two key residues, the nucleophilic attack mediator and catalytic acid/base, are routinely carboxylic acid-containing residues in a catalytic doublet separated by a single hydrophobic amino acid. This shared catalytic motif is also present among NCOATs. In mouse NCOAT, this catalytic doublet consists of two aspartic acid residues at positions 175 and 177 in the primary sequence. Not coincidentally, these two residues also lie in the predicted β4-strand of the triose phosphate isomerase barrel, a location corresponding to those of analogous residues in family members that have been crystallized (39, 49, 51).

When either Asp^{175} was converted to an alanine or Asp^{177} was changed to an asparagine, the resulting mutant enzyme was catalytically dead in the *in vitro* *O*-GlcNAcase assay. The mutant enzyme's K_m increased significantly, and the k_{cat} reduced dramatically in each case; these findings indicate that these two residues are both intimately involved in substrate binding and catalysis. As further confirmation that each of these residues is indispensable for catalysis, each enzyme was able to recover activity when acidic pH buffers were used in the assay. When an exogenous nucleophile was added, the Asp^{177} mutant recovered 40% activity, indicative of the residue's importance in at least one step of the acid/base reaction. Asp^{175} was able to recover up to 70% of the activity when a 4-M solution of an exogenous nucleophile was added to the system, while its activity was essentially dead in the solution's absence; these findings strongly suggest a nucleophilic role for this amino acid. Taken together, these data lead us to propose that these two residues in NCOAT play the roles of the two catalytically necessary residues in those enzymes in glycosidase families 18 and 56 and that NCOAT follows the same catalytic mechanism that these same enzymes do (Fig. 2*A*).

Finally, the Asp^{174} residue, lying immediately N-terminal to the catalytic doublet, is conserved among NCOATs. The mutation of this residue to an asparagine impairs the enzymatic activity at neutral pH and converts the enzyme toward an acidic optimum. Several glycosidase families have conserved catalytic carboxylates in their active sites in addition to those involved directly in catalysis (40, 52). Upon characterization, these conserved carboxylates were found to play a role in maintaining the proper charge on other residues within the active site. It is likely that this Asp^{174}, due to its proximity to both of the essential carboxylates in the catalytic doublet, similarly is involved in regulating the pK_a of the active site hydrophobic environment and aids in keeping both members of the catalytic doublet in a substrate-ready state.

The role of the free cysteine residue is more problematical. It was previously demonstrated that sulfhydril-specific reagents could inhibit the enzymatic activity of NCOAT (4). This observation is indicative of a free cysteine important to the normal functioning of the enzyme. Since a cysteine modification results in the conversion of the S-H to an S-R, where R depends on the sulfhydril reagent used, it is possible that this R group, which is larger than an H, may impart electronic, polar, or steric constraints not consistent with the normal active site architecture. By replacing an active site cysteine (Cys^{166}) with a serine, we addressed directly the importance of this residue in catalysis. This residue is not involved in *O*-GlcNAc substrate binding, as indicated by its K_m value's rivaling that of the wild type enzyme. However, the one-third reduction in k_{cat} when compared to wild type verifies its role, direct or indirect, in catalysis. Currently, with the absence of crystal structure, a distinct role for this residue in catalysis cannot be assigned.

While the mutagenesis and structural studies on other glycosidases provide evidence for the catalytic mechanism of the *O*-GlcNAcase, additional evidence is obtainable using an inhibitor of the enzyme. STZ, a GlcNAc analog and specific, competitive inhibitor of *O*-GlcNAcase activity (Fig. 5), was used to acquire a stable transition state mimetic inhibitory compound. NCOAT itself is capable of transforming STZ to this stable compound, which has a molecular weight of 206 and a structure consistent with that of the bicyclic compound predicted by the catalytic mechanism, with no NO or CH_3 groups present. Both residues 175 and 177, which are required by the enzyme for proper catalysis with the native substrate, are also required to convert STZ to its more potent transition state analog. Similar to the work on other glycosidases and their respective transition state analog inhibitors, these studies with STZ serve to endorse our model for the catalytic mechanism of NCOAT. These findings also demonstrate exactly how STZ performs its inhibitory function on NCOAT. It is not via an oxidative damage mechanism or *S*-nitrosation since STZ free of NO can still inhibit the enzyme (Fig. 6, 7, *B* and *C*) and since the liberated NO is removed from the system through washing. These data leave the enzyme-assisted formation of a transition state analog from STZ as the most plausible mechanism of inhibition.

Chemical inhibition of *O*-GlcNAcase by STZ reproduces several hallmark features of diabetes and Alzheimer's disease (46, 53). For diabetes, both β-cell death and insulin secretory defects result from STZ treatment (54). While STZ releases NO (43) and methyl groups (43), high amounts of oxidative damage or DNA alkylation would be required to induce apoptosis, and the β-cell specificity remains unexplained. Because high levels of OGT exist in β-cells (5) and because *O*-GlcNAc accumulation in these cells occurs in a glucose-dependent fashion at physiological concentrations (6), we favor

the inhibition of *O*-GlcNAcase as the premier basis for STZ-induced β-cell failure. OGT is also expressed highly in the pituitary and hippocampus (54), and this suggests how STZ might target these tissues as well. The new data presented in this paper further this contention by indicating how the inhibition of *O*-GlcNAcase by STZ could be sustained and potent. The potent transition-state analog produced from STZ in the cell could persistently inhibit newly synthesized *O*-GlcNAcase when freed during normal protein turnover/degradation. β-cell death and diabetes as well as hippocampal cell death and Alzheimer's disease could then result from the inhibition of *O*-GlcNAcase, the only nucleocytosolic enzyme capable of removing the glucose-dependent *O*-GlcNAc modifications that participate in so many vital cellular processes (proteasome, gene repression (18, 22)).

References

1. Wells, L., Vosseller, K., and Hart, G. W. (2001) *Science* **291**, 2376-2378
2. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) *J. Biol. Chem.* **277**, 1755-1761
3. Gao, Y., Wells, L., Comer, F. I., Parker, G. J., and Hart, G. W. (2001) *J. Biol. Chem.* **276**, 9838-9845
4. Dong, D. L., and Hart, G. W. (1994) *J. Bio. Chem.* **269**, 19321-19330
5. Roos, M. D., Xie, W., Su, K., Clark, J. A., Yang, X., Chin, E., Paterson, A. J., and Kudlow, J. E. (1998) *Proc. Assoc. Am. Physicians* **110**, 422-432
6. Liu, K., Paterson, A. J., Chin, E., and Kudlow, J. E. (2000) *Proc. Natl. Acad. Sci. U S A* **97**, 2820-2825
7. Gupta, R., Birch, H., Rapacki, K., Brunak, S., and Hansen, J. E. (1999) *Nucleic Acids Res.* **27**, 370-372
8. Haltiwanger, R. S., Kelly, W. G., Roquemore, E. P., Blomberg, M. A., Dong, L. Y., Kreppel, L., Chou, T. Y., and Hart, G. W. (1992) *Biochem. Soc. Trans.* **20**, 264-269

9. Hart, G. W., Haltiwanger, R. S., Holt, G. D., and Kelly, W. G. (1989) *Ciba Found. Symp.* **145**, 102-112, discussion 112-118

10. Kearse, K. P., and Hart, G. W. (1991) *Arch. Biochem. Biophys.* **290**, 543-548

11. Chou, C. F., and Omary, M. B. (1993) *J. Biol. Chem.* **268**, 4465-4472

12. Haltiwanger, R. S., Blomberg, M. A., and Hart, G. W. (1992) *J. Biol. Chem.* **267**, 9005-9013

13. Comer, F. I., and Hart, G. W. (1999) *Biochim. Biophys. Acta* **1473**, 161-171

14. Yang, X., Zhang, F., and Kudlow, J. E. (2002) *Cell* **110**, 69-80

15. Cheng, X., and Hart, G. W. (2001) *J. Biol. Chem.* **276**, 10570-10575

16. Yang, X., Su, K., Roos, M. D., Chang, Q., Paterson, A. J., and Kudlow, J. E. (2001) *Proc. Natl. Acad. Sci. U S A* **98**, 6611-6616

17. Jackson, S. P., and Tjian, R. (1988) *Cell* **55**, 125-133

18. Zhang, F., Su, K., Yang, X., Bowe, D. B., Paterson, A. J., and Kudlow, J. E. (2003) *Cell* **115**, 715-725

19. Roos, M. D., Su, K., Baker, J. R., and Kudlow, J. E. (1997) *Mol. Cell. Biol.* **17**, 6472-6480

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

21. Han, I., and Kudlow, J. E. (1997) *Mol. Cell. Biol.* **17**, 2550-2558

22. Toleman, C., Paterson, A. J., Whisenhunt, T. R., and Kudlow, J. E. (2004) *J. Biol. Chem.* **279**, 53665-53673

23. Henrissat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.* **7**, 637-644

24. Henrissat, B. (1998) *Biochem. Soc. Trans.* **26**, 153-156

25. Rye, C. S., and Withers, S. G. (2000) *Curr. Opin. Chem. Biol.* **4**, 573-580

26. Knapp, S., Vocadlo, D., Gao, Z., Kirk, B., Lou, J., Withers, S. G. (1996) *J. Am. Chem. Soc.* **118**, 6804-6805

27. Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., and Dijkstra, B. W. (1995) *Biochemistry* **34**, 15619-15623

28. Zechel, D. L., and Withers, S. G. (2000) *Acc. Chem. Res.* **33**, 11-18

29. Ly, H. D., and Withers, S. G. (1999) *Annu. Rev. Biochem.* **68**, 487-522

30. Yudin, A. I., Li, M. W., Robertson, K. R., Cherr, G. N., and Overstreet, J. W. (2001) *Reproduction* **121**, 735-743

31. Markovic-Housley, Z., Miglierini, G., Soldatova, L., Rizkallah, P. J., Muller, U., and Schirmer, T. (2000) *Structure Fold Des.* **8**, 1025-1035

32. Bokma, E., Rozeboom, H. J., Sibbald, M., Dijkstra, B. W., and Beintema, J. J. (2002) *Eur. J. Biochem.* **269**, 893-901

33. Mark, B. L., Wasney, G. A., Salo, T. J., Khan, A. R., Cao, Z., Robbins, P. W., James, M. N., and Triggs-Raine, B. L. (1998) *J. Biol. Chem.* **273**, 19618-19624

34. Hou, Y., Vocadlo, D. J., Leung, A., Withers, S. G., and Mahuran, D. (2001) *Biochemistry* **40**, 2201-2209

35. Mehler, E. L., Fuxreiter, M., Simon, I., and Garcia-Moreno, E. B. (2002) *Proteins* **48**, 283-292

36. Sierks, M. R., and Svensson, B. (1992) *Protein Eng.* **5**, 185-188

37. Williams, S. J., Mark, B. L., Vocadlo, D. J., James, M. N., and Withers, S. G. (2002) *J. Biol. Chem.* **277**, 40055-40065

38. Lin, F. P., Chen, H. C., and Lin, C. S. (1999) *I.U.B.M.B. Life* **48**, 199-204

39. Prag, G., Papanikolau, Y., Tavlas, G., Vorgias, C. E., Petratos, K., and Oppenheim, A. B. (2000) *J. Mol. Biol.* **300**, 611-617

40. MacLeod, A. M., Tull, D., Rupitz, K., Warren, R. A., and Withers, S. G. (1996) *Biochemistry* **35**, 13165-13172

41. Turk, J., Corbett, J. A., Ramanadham, S., Bohrer, A., and McDaniel, M. L. (1993) *Biochem. Biophys. Res. Commun.* **197**, 1458-1464

42. Elsner, M., Guldbakke, B., Tiedge, M., Munday, R., and Lenzen, S. (2000) *Diabetologia* **43**, 1528-1533

43. Kwon, N. S., Lee, S. H., Choi, C. S., Kho, T., and Lee, H. S. (1994) *FASEB J.* **8**, 529-533

44. Kroncke, K. D., Fehsel, K., Sommer, A., Rodriguez, M. L., and Kolb-Bachofen, V. (1995) *Biol. Chem. Hoppe Seyler* **376**, 179-185

45. Cardinal, J. W., Margison, G. P., Mynett, K. J., Yates, A. P., Cameron, D. P., and Elder, R. H. (2001) *Mol. Cell. Biol.* **21**, 5605-5613

46. Liu, K., Paterson, A. J., Zhang, F., McAndrew, J., Fukuchi, K., Wyss, J. M., Peng, L., Hu, Y., and Kudlow, J. E. (2004) *J. Neurochem.* **89**, 1044-1055

47. Hanover, J. A. (2001) *FASEB J.* **15**, 1865-1876

48. Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001) *J. Mol. Biol.* **313**, 903-919

49. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G., and James, M. N. (2001) *J. Biol. Chem.* **276**, 10330-10337

50. Heckel, D., Comtesse, N., Brass, N., Blin, N., Zang, K. D., and Meese, E. (1998) *Hum. Mol. Genet.* **7**, 1859-1872

51. Hollis, T., Monzingo, A. F., Bortone, K., Ernst, S., Cox, R., and Robertus, J. D. (2000) *Protein Sci.* **9**, 544-551

52. MacLeod, A. M., Lindhorst, T., Withers, S. G., and Warren, R. A. (1994) *Biochemistry* **33**, 6371-6376

53. Konrad, R. J., Mikolaenko, I., Tolar, J. F., Liu, K., and Kudlow, J. E. (2001) *Biochem. J.* **356**, 31-41

54. Liu, K., Paterson, A. J., Konrad, R. J., Parlow, A. F., Jimi, S., Roh, M., Chin, E., Jr., and Kudlow, J. E. (2002) *Mol. Cell. Endocrinol.* **194**, 135-146

SUMMARY

The Nomenclature of Nuclear Cytoplasmic O-GlcNAcase and Acetyltransferase

O-GlcNAcase was initially characterized in 1994 by Dong and Hart (4). They reported its specificity for *O*-GlcNAc substrates and biochemically characterized its activity in the presence of metal ion effectors by varying pH and a range of common glycosidase inhibitors. Independently, in 1998, the same enzyme was described by Heckel *et al.* (38), who purified the enzyme from the screening of meningioma expressing libraries, and therefore referred to the enzyme as meningioma expressed antigen 5. Due to the nature of the meningiomas in study, the enzyme was examined for hyaluronidase activity and was found to competently cleave hyaluronic acid substrates. Since hexosaminidases and hyaluronidases both recognize and catalyze GlcNAc substrates, these independent findings are not surprising. The name *O*-GlcNAcase has taken precedent over meningioma expressed antigen 5 for the "hyaluronoglucosaminidase" characterized by both groups. The results of the works presented here demonstrate that *O*-GlcNAcase also possesses a separate acetyltransferase activity for histone substrates. In order to prevent the prioritized assignment of one of these activities over the other in terms of overall cellular importance, we hereafter refer to the enzyme as NCOAT (nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase) to clearly portray its two activities.

Characterization of the O-GlcNAcase Active Site of NCOAT

The knowledge that the sulfhydril-specific alkylating reagent *N*-ethylmaleimide inhibits *O*-GlcNAcase activity suggests that the enzyme has a free cysteine residue in or around its active site (4). This information was used as the starting point to narrow in on the active site's location. The use of a biotinylated sulfhydril reagent, *N*-(hexyl)-3'-(2'-pyridyldithio) propionamide, allowed us to purify peptides that contained a free cysteine in the enzyme's native structure and to utilize matrix-assisted laser desorption/ionization mass spectroscopy to determine their sites in the primary sequence of the enzyme. Two free cysteines were discovered and subsequently mutated to serines. Only one of these two mutant enzymes, the Cys^{166}→Ser enzyme, lacked enzymatic activity in the hexosaminidase assay, thus providing the first biochemical evidence of the active site's proximity. When two naturally occurring splice variants were used in the assay, they both were catalytically defunct, leading us to the conclusion that the enzyme's active site overlapped the cysteine residue at position 166 as well as the common region missing from the two splice variants, amino acids 250-345. This finding is supported by several independent observations and lines of evidence. First, the region between amino acids 63 and 283 for mouse NCOAT displays strong similarity to hyaluronidase active sites as determined by SMART protein analysis (39). NCOATs across species are also highly conserved in sequence within this area. Another protein analysis program, Structural Classification of Protein, predicts NCOAT to have a triose phosphate isomerase α/β_8 barrel corresponding to this region, with the same secondary structure pattern seen in chitinase, hexosaminidase, and hyaluronidase active sites (64, 65). Second, an identified splice variant of this enzyme, which lacks the C-terminal third and the HAT domain of the enzyme, retains significant activity in hexosaminidase assays (14). Third, NCOAT is a sub-

strate for cleavage by the protease caspase-3; the resulting protein, possessing only the N-terminal 350 amino acids, is similarly active in *O*-GlcNAcase assays.

Having determined that the location and architecture of the *O*-GlcNAcase active site resemble those of chitinases, hexosaminidases, and hyaluronidases allows for the assumption that the catalytic mechanism of NCOAT is particularly similar to these enzymes. These enzymes have a conserved mechanism whereby a carboxylic acid-containing residue donates a proton to the oxygen atom of the *O*-linkage, breaking the glycosidic bond. The 2-acetamido side-chain of the GlcNAc substrate then acts as a nucleophile, attacking the anomeric carbon to form an oxazolinium ion intermediate. The nucleophilic attack is aided by a second carbonyl, which manipulates the side-chain as well as stabilizes the transition state charge. These two critical residues in these enzymes reside in a D-ψ-D/E motif that is found in the mouse *O*-GlcNAcase at amino acids 175-177. Mutation of these residues resulted in inactive enzymes, thereby establishing their importance for *O*-GlcNAcase activity. By using assays involving increasing substrate concentration, varying pH, and nucleophilic chemical rescue, we have confirmed the roles of these two residues; the aspartic acid at 175 plays a role in mediating the nucleophilic attack and the aspartic acid at 177 is the critical acid/base. The aspartic acid at position 175, when mutated to an alanine, resulted in a protein with a shift to an acidic pH optimum, a 4-fold increase in the enzyme's K_m, and an 80% decrease in k_{cat}; these findings confirm its role in substrate binding and *O*-linked GlcNAc catalysis. This mutant enzyme also had the ability to recover up to 70% of the wild type level of activity when an exogenous nucleophile, azide, was added to the reaction, validating its nucleophilic role in catalysis. When the aspartic acid residue at position 177 was mutated to an aspartate, the protein also exhibited a marked shift to an acidic pH optimum. The $Asp^{177}\rightarrow Asn$

mutant had a K_m eight times that of the wild type enzyme and a decrease in k_{cat} to 10% of that of the wild type, once again demonstrating its importance in substrate binding and catalysis. A 5-fold recovery of activity in the presence of negatively charged azide exhibited its role in the basic step of the enzymatic reaction. The cysteine residue at position 166, mentioned above, and a mutation of the aspartic acid at 174 were also tested. The Cys^{166}→Ser mutant was determined to be involved indirectly in catalysis on the basis of its significant decrease in $k_{cat,}$ with no change in K_m. The Asp^{174}→Asn mutant displayed kinetic parameters similar to those of the Cys^{166}→Ser mutant and showed a dramatic optimal pH shift to the acidic range, demonstrating that it also plays a role in catalysis as well as in maintaining the charges on the surrounding residues so they can function at neutral pH. These data provide a platform from which specific drugs may be designed to inhibit *O*-GlcNAcase activity, facilitating future studies of any pathway under the influence of *O*-GlcNAc modification.

Mechanism of O-GlcNAcase Inhibition by Streptozotocin

Having established the mechanism of *O*-GlcNAcase catalysis and having determined the amino acid residues within the active site which are principally responsible for this mechanism provide a foundation for examining how the diabetogenic drug STZ can inhibit this activity. Since STZ is a GlcNAc analog, it can reasonably be expected to target the *O*-GlcNAcase active site of NCOAT, where the enzyme will attempt to catalyze the compound as if it were a normal *O*-GlcNAcylated substrate.

STZ has previously been revealed to spontaneously release NO as well as to donate its terminal methyl group (CH_3) to specific targets (66, 67). The NO and CH_3 group loss occurs independently of the presence of NCOAT, with the NO loss starting after 5

min when added to a neutral or acidic solvent and the CH_3 group lost over several hours. The resulting ligand, when incubated with *O*-GlcNAcase, gets converted into a compound which can inhibit a fresh batch of enzyme at a concentration, 1 μM, which is 10^5 times more potent than the native compound. This inhibition takes less than 5 min, whereas it takes 0.1 M of the original ligand 40 min to completely inhibit the enzyme. This new chemical compound is the result of the enzyme's catalytic mechanism, which converts STZ into a transition state analog which is structurally stable in the enzyme's active site. The resulting compound has an ionized molecular weight of 206, consistent with that predicted to be the case when the enzyme catalyzes the substrate into the transition state analog, where its only divergence from the natural ligand transition state is the replacement of a CH_3 with an NH_2. Structural observations by NMR also reinforce the conclusion that STZ is converted by the enzyme into the bicyclic transition state mimetic. Glycomimetics, often designed to resemble the transition state of a catalyzed substrate, are emerging as important tools in the structure/function study of the enzymes they inhibit. They are also evolving as candidates for therapeutic compounds when the enzyme they inhibit has been validated as disease-causing or restorative when impaired. Glycomimetics are currently showing considerable promise as chemotherapeutics for the treatment of viral infections (68), cancer (69), and osteoarthritis (70) as well as for metabolic disorders such as Tay-Sachs and Sandhoff's disease (71, 72) and diabetes (73). The work presented here provides a valuable tool for the study of inducible diabetes in laboratory animals.

Characterization of the Acetyltransferase Activity and Domain of NCOAT

Published results by Comtesse *et al.* (39) had reported the observation that NCOAT possesses a domain in its C terminus that has similarity to the active sites of acetyltransferases, though at that time a corresponding activity had not been established. The results of the work detailed here demonstrate that NCOAT does in fact possess acetyltransferase activity *in vitro* for a synthetic histone H4 tail substrate as well as for all four core histones, both free and in the context of reconstituted oligonucleosomes. Moreover, it does so with site specificity, as NCOAT can acetylate histone H3 on lysine 14 and histone H4 on lysine 8 but not on lysine 16.

The active site for this activity was found to reside, as predicted, in the C-terminal third of the protein, where it displays the exact pattern of α-helices and β-sheets seen in the crystal structures for nine different acetyltransferases that are predicted to have similarity to NCOAT. The facts that STZ does not affect acetyltransferase activity and that the C-terminal third of the enzyme alone is sufficient to catalyze acetyltransfer dictate that this domain, like other acetyltransferase domains, acts independently of the rest of the protein.

The acetyltransferase domain of NCOAT was also ascertained to maintain the same catalytic mechanism that the Gcn5, HAT1, and GNAT acetyltransferases families do. These enzymes catalyze acetyltransfer by having an active site carboxylate-containing residue acting as a base that deprotonates the target lysine residue, which invariably enters the active site in its NH_3^+ form. This deprotonation causes the ε-amino group of the lysine to attack and bind to the acetyl group of acetyl-CoA. The reaction is then drawn to completion when an active site tyrosine (with exception) donates a proton to the coenzyme A sulfur atom to generate the thiolate leaving group. NCOAT's C termi-

nus possesses the necessary carboxylate-containing residue and tyrosine in positions nearly identical to (if not identical to) these enzymes in the predicted secondary structure of the active site. When placed in histone acetyltransferase assays, the proteins with mutations to these residues, the tyrosine at position 891 and the aspartate at either position 853 or position 884, were found to retain little to no activity over background, thereby establishing these residues as essential for catalysis and validating the above mechanism as the means by which NCOAT can catalyze acetyltransfer.

NCOAT's possession of two catalytic activities gives the enzyme the potential to have a profound impact on gene transcriptional regulation. It is well documented that areas of repressed gene transcription have a tendency to be hypoacetylated (74-76). In addition, areas of repressed chromatin also have shown a tendency to be hyperglycosylated (21); also, *O*-GlcNAc addition to transcriptional activators such as Sp1, cyclic 3',5' adenosine monophosphate responsive element-binding protein, and RNA polymerase II has been shown to impair their ability to act as genetic coactivators (23, 77, 78). Unpublished data from our laboratory have shown that OGT can bind very strongly to NCOAT in pull-down assays using either overexpressed or endogenous proteins. The fact that OGT has been shown to be recruited to repressed chromatin by mSin3A (21) immediately leads to the implication that NCOAT will be found at these sites as well. NCOAT has the ability to remove the repressive *O*-GlcNAc modifications on transcriptional activators as well as the ability to acetylate histones and open up the chromatin for transcriptional activity, making NCOAT a player of potentially paramount importance in gene regulation.

GENERAL LIST OF REFERENCES

1. Olson, A. L., and Pessin, J. E. (1996) *Annu. Rev. Nutr.* **16**, 235-256

2. Kaneto, H., Xu, G., Song, K. H., Suzuma, K., Bonner-Weir, S., Sharma, A., and Weir, G. C. (2001) *J. Biol. Chem.* **276**, 31099-31104

3. Liu, K., Paterson, A. J., Chin, E., and Kudlow, J. E. (2000) *Proc. Natl. Acad. Sci. U S A* **97**, 2820-2825

4. Dong, D. L., and Hart, G. W. (1994) *J. Biol. Chem.* **269**, 19321-19330

5. Torres, C. R., and Hart, G. W. (1984) *J. Biol. Chem.* **259**, 3308-3317

6. Kreppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) *J. Biol. Chem.* **272**, 9308-9315

7. Kreppel, L. K., and Hart, G. W. (1999) *J. Biol. Chem.* **274**, 32015-32022

8. 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

9. Haltiwanger, R. S., Blomberg, M. A., and Hart, G. W. (1992) *J. Biol. Chem.* **267**, 9005-9013

10. Hanover, J. A., Lai, Z., Lee, G., Lubas, W. A., and Sato, S. M. (1999) *Arch. Biochem. Biophys.* **362**, 38-45

11. Roos, M. D., Xie, W., Su, K., Clark, J. A., Yang, X., Chin, E., Paterson, A. J., and Kudlow, J. E. (1998) *Proc. Assoc. Am. Physicians* **110**, 422-432

12. Lubas, W. A., Frank, D. W., Krause, M., and Hanover, J. A. (1997) *J. Biol. Chem.* **272**, 9316-9324

13. Konrad, R. J., Zhang, F., Hale, J. E., Knierman, M. D., Becker, G. W., and Kudlow, J. E. (2002) *Biochem. Biophys. Res. Commun.* **293**, 207-212

14. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) *J. Biol. Chem.* **277**, 1755-1761

15. Gupta, R., Birch, H., Rapacki, K., Brunak, S., and Hansen, J. E. (1999) *Nucleic Acids Res.* **27**, 370-372

16. Hart, G. W., Haltiwanger, R. S., Holt, G. D., and Kelly, W. G. (1989) *Ciba Found. Symp.* **145**, 102-112, discussion 112-118

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

18. Kearse, K. P., and Hart, G. W. (1991) *Arch. Biochem. Biophys.* **290**, 543-548

19. Chou, C. F., and Omary, M. B. (1993) *J. Biol. Chem.* **268**, 4465-4472

20. Comer, F. I., and Hart, G. W. (1999) *Biochim. Biophys. Acta* **1473**, 161-171

21. Yang, X., Zhang, F., and Kudlow, J. E. (2002) *Cell* **110**, 69-80

22. Cheng, X., and Hart, G. W. (2001) *J. Biol. Chem.* **276**, 10570-10575

23. Yang, X., Su, K., Roos, M. D., Chang, Q., Paterson, A. J., and Kudlow, J. E. (2001) *Proc. Natl. Acad. Sci. U S A* **98**, 6611-6616

24. Jackson, S. P., and Tjian, R. (1988) *Cell* **55**, 125-133

25. Zhang, F., Su, K., Yang, X., Bowe, D. B., Paterson, A. J., and Kudlow, J. E. (2003) *Cell* **115**, 715-725

26. Zachara, N. E., and Hart, G. W. (2004) *Biochim. Biophys. Acta* **1673**, 13-28

27. Han, I., and Kudlow, J. E. (1997) *Mol. Cell. Biol.* **17**, 2550-2558

28. Roos, M. D., Su, K., Baker, J. R., and Kudlow, J. E. (1997) *Mol. Cell. Biol.* **17**, 6472-6480

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

30. Parker, G. J., Lund, K. C., Taylor, R. P., and McClain, D. A. (2003) *J. Biol. Chem.* **278**, 10022-10027

31. Federici, M., Menghini, R., Mauriello, A., Hribal, M. L., Ferrelli, F., Lauro, D., Sbraccia, P., Spagnoli, L. G., Sesti, G., and Lauro, R. (2002) *Circulation* **106**, 466-472

32. Comer, F. I., and Hart, G. W. (2001) *Biochemistry* **40**, 7845-7852

33. Neufeld, E. F. (1989) *J. Biol. Chem.* **264**, 10927-10930

34. Farook, V., Bogardus, C., and Prochazka, M. (2002) *Mol. Genet. Metab.* **77**, 189-193

35. Kudlow, J. (2002) *Mol. Genet. Metab.* **77**, 1-2

36. Hanover, J. A. (2001) *FASEB J* **15**, 1865-1876

37. Appel, R. D., Bairoch, A., and Hochstrasser, D. F. (1994) *Trends Biochem. Sci.* **19**, 258-260

38. Heckel, D., Comtesse, N., Brass, N., Blin, N., Zang, K. D., and Meese, E. (1998) *Hum. Mol. Genet.* **7**, 1859-1872

39. Comtesse, N., Maldener, E., and Meese, E. (2001) *Biochem. Biophys. Res. Commun.* **283**, 634-640

40. Hickman, A. B., Namboodiri, M. A., Klein, D. C., and Dyda, F. (1999) *Cell* **97**, 361-369

41. Angus-Hill, M. L., Dutnall, R. N., Tafrov, S. T., Sternglanz, R., and Ramakrishnan, V. (1999) *J. Mol. Biol.* **294**, 1311-1325

42. Dutnall, R. N., Tafrov, S. T., Sternglanz, R., and Ramakrishnan, V. (1998) *Cell* **94**, 427-438

43. Poux, A. N., Cebrat, M., Kim, C. M., Cole, P. A., and Marmorstein, R. (2002) *Proc. Natl. Acad. Sci. U S A* **99**, 14065-14070

44. Yan, Y., Barlev, N. A., Haley, R. H., Berger, S. L., and Marmorstein, R. (2000) *Mol. Cell* **6**, 1195-1205

45. Wolf, E., Vassilev, A., Makino, Y., Sali, A., Nakatani, Y., and Burley, S. K. (1998) *Cell* **94**, 439-449

46. Clements, A., Rojas, J. R., Trievel, R. C., Wang, L., Berger, S. L., and Marmorstein, R. (1999) *EMBO J.* **18**, 3521-3532

47. Wybenga-Groot, L. E., Draker, K., Wright, G. D., and Berghuis, A. M. (1999) *Structure Fold Des.* **7**, 497-507

48. Marmorstein, R. (2001) *J. Mol. Biol.* **311**, 433-444

49. Yan, Y., Harper, S., Speicher, D. W., and Marmorstein, R. (2002) *Nat. Struct. Biol.* **9**, 862-869

50. Wolf, E., De Angelis, J., Khalil, E. M., Cole, P. A., and Burley, S. K. (2002) *J. Mol. Biol.* **317**, 215-224

51. Chinenov, Y. (2002) *Trends Biochem. Sci.* **27**, 115-117

52. Zechel, D. L., and Withers, S. G. (2000) *Acc. Chem. Res.* **33**, 11-18

53. Rye, C. S., and Withers, S. G. (2000) *Curr. Opin. Chem. Biol.* **4**, 573-580

54. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G., and James, M. N. (2001) *J. Biol. Chem.* **276**, 10330-10337

55. Piszkiewicz, D., and Bruice, T. C. (1968) *J. Am. Chem. Soc.* **90**, 2156-2163

56. Yudin, A. I., Li, M. W., Robertson, K. R., Cherr, G. N., and Overstreet, J. W. (2001) *Reproduction* **121**, 735-743

57. Markovic-Housley, Z., Miglierini, G., Soldatova, L., Rizkallah, P. J., Muller, U., and Schirmer, T. (2000) *Structure Fold Des.* **8**, 1025-1035

58. Bokma, E., Rozeboom, H. J., Sibbald, M., Dijkstra, B. W., and Beintema, J. J. (2002) *Eur. J. Biochem.* **269**, 893-901

59. Pearl, J., Elster, E., and Golocovsky, M. (2002) *Am. Surg.* **68**, 886-888

60. Yamamoto, H., Uchigata, Y., and Okamoto, H. (1981) *Nature* **294**, 284-286

61. Cardinal, J. W., Margison, G. P., Mynett, K. J., Yates, A. P., Cameron, D. P., and Elder, R. H. (2001) *Mol. Cell. Biol.* **21**, 5605-5613

62. Saini, K. S., Thompson, C., Winterford, C. M., Walker, N. I., and Cameron, D. P. (1996) *Biochem. Mol. Biol. Int.* **39**, 1229-1236

63. Jaffrey, S. R., Erdjument-Bromage, H., Ferris, C. D., Tempst, P., and Snyder, S. H. (2001) *Nat. Cell Biol.* **3**, 193-197

64. Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., and Dijkstra, B. W. (1995) *Biochemistry* **34**, 15619-15623

65. Hollis, T., Monzingo, A. F., Bortone, K., Ernst, S., Cox, R., and Robertus, J. D. (2000) *Protein Sci.* **9**, 544-551

66. Elsner, M., Guldbakke, B., Tiedge, M., Munday, R., and Lenzen, S. (2000) *Diabetologia* **43**, 1528-1533

67. Turk, J., Corbett, J. A., Ramanadham, S., Bohrer, A., and McDaniel, M. L. (1993) *Biochem. Biophys. Res. Commun.* **197**, 1458-1464

68. Taylor, D. L., Sunkara, P. S., Liu, P. S., Kang, M. S., Bowlin, T. L., and Tyms, A. S. (1991) *AIDS* **5**, 693-698

69. Goss, P. E., Baker, M. A., Carver, J. P., and Dennis, J. W. (1995) *Clin. Cancer Res.* **1**, 935-944

70. Liu, J., Shikhman, A. R., Lotz, M. K., and Wong, C. H. (2001) *Chem. Biol.* **8**, 701-711

71. Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S., and James, M. N. (2003) *J. Mol. Biol.* **327**, 1093-1109

72. Maier, T., Strater, N., Schuette, C. G., Klingenstein, R., Sandhoff, K., and Saenger, W. (2003) *J. Mol. Biol.* **328**, 669-681

73. Balfour, J. A., and McTavish, D. (1993) *Drugs* **46**, 1025-1054

74. Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998) *Mol. Cell. Biol.* **18**, 4629-4638

75. Turner, B. M., and O'Neill, L. P. (1995) *Semin. Cell Biol.* **6**, 229-236

76. Richards, E. J., and Elgin, S. C. (2002) *Cell* **108**, 489-500

77. Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993) *J. Biol. Chem.* **268**, 10416-10424

78. Lamarre-Vincent, N., and Hsieh-Wilson, L. C. (2003) *J. Am. Chem. Soc.* **125**, 6612-6613

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM
DOCTOR OF PHILOSOPHY

Name of Candidate Clifford A. Toleman

Graduate Program Cell Biology

Title of Dissertation Characterization of NCOAT, A Bifunctional Enzyme with *O*-GlcNAcase and Acetyltransferase Activity

I certify that I have read this document and examined the student regarding its content. In my opinion, this dissertation conforms to acceptable standards of scholarly presentation and is adequate in scope and quality, and the attainments of this student are such that he may be recommended for the degree of Doctor of Philosophy.

Dissertation Committee:

Name		Signature
Jeffrey E. Kudlow	**, Chair**	
Xinbin Chen		
Jeffrey A. Engler		
Rakesh P. Patel		
Bradley K. Yoder		

Director of Graduate Program

Dean, UAB Graduate School

Date