THE PROTEIN SOCIETY

METHODS AND APPLICATIONS

Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related *N*-acetyltransferases with unknown substrate specificity

Misty L. Kuhn,[1] Karolina A. Majorek,[2] Wladek Minor,[2] and Wayne F. Anderson[1]*

[1]Department of Pharmacology and Cellular Biology, Center for Structural Genomics of Infectious Diseases, Northwestern Feinberg School of Medicine, Chicago, Illinois 60611

[2]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908

Received 22 September 2012; Revised 15 November 2012; Accepted 15 November 2012
DOI: 10.1002/pro.2199
Published online 26 November 2012 proteinscience.org

Abstract: Due to a combination of efforts from individual laboratories and structural genomics centers, there has been a surge in the number of members of the Gcn5-related acetyltransferasesuperfamily that have been structurally determined within the past decade. Although the number of three-dimensional structures is increasing steadily, we know little about the individual functions of these enzymes. Part of the difficulty in assigning functions for members of this superfamily is the lack of information regarding how substrates bind to the active site of the protein. The majority of the structures do not show ligand bound in the active site, and since the substrate-binding domain is not strictly conserved, it is difficult to predict the function based on structure alone. Additionally, the enzymes are capable of acetylating a wide variety of metabolites and many may exhibit promiscuity regarding their ability to acetylate multiple classes of substrates, possibly having multiple functions for the same enzyme. Herein, we present an approach to identify potential substrates for previously uncharacterized members of the Gcn5-related acetyltransferase superfamily using a variety of metabolites including polyamines, amino acids, antibiotics, peptides, vitamins, catecholamines, and other metabolites. We have identified potential substrates for eight bacterial enzymes of this superfamily. This information will be used to further structurally and functionally characterize them.

Keywords: GNAT; Gcn5-related acetyltransferase; *N*-acetyltransferase; Ellman's reagent; broad-substrate screen; acetylation of antibiotics

Introduction

The General control of nutrition 5 (Gcn5)-related *N*-acetyltransferase (GNAT) superfamily contains enzymes that transfer an acetyl group from acetyl coenzyme A (AcCoA) to an acceptor substrate and release both CoA and the acetylated product. Members of this superfamily include histone acetyltransferases,

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Wayne F. Anderson, Department of Pharmacology and Cellular Biology, Northwestern Feinberg School of Medicine, 303 E. Chicago Ave., Chicago, IL 60611. E-mail: wf-anderson@northwestern.edu

aminoglycoside *N*-acetyltransferases, glucosamine 6-phosphate acetyltransferases, and serotonin *N*-acetyltransferases, to name a few.[1] GNATs exist in all three kingdoms of life. They are important because they tightly control a wide variety of orchestrated cellular processes including regulation of transcription and metabolic flux by acetylating the proteins that perform these functions. They also aid in bacterial drug resistance mechanisms by acetylating and rendering antibiotics inactive.[2,3]

There are currently 200 three-dimensional structures of GNATs in the Protein Data Bank (PDB). Seventy-five percent of these structures have been determined since 2005, and over half are from structural genomics-related efforts. Although the number of structures of this superfamily is increasing steadily, functional studies to determine their substrate specificity have lagged. The task of predicting or identifying their substrates is complex for several reasons. Substrate binding domains of GNATs are not conserved; thus, the preferred substrate cannot be predicted from the three-dimensional structures. Moreover, GNATs are capable of acetylating a wide variety of substrates including metabolic intermediates, antibiotics, and proteins,[1,4] and some GNATs have the capacity to acetylate both antibiotics and proteins.[5] Once a substrate is known for a particular GNAT, several approaches exist to further characterize the enzymes. Several GNATs have been kinetically characterized, including GNATs that use protein substrates,[4,6–8] polyamine substrates,[9] and other metabolites.[8,10,11] Biochemical approaches for addressing the issue of substrate specificity of uncharacterized GNATs, however, have been limited.[4]

To increase the number of GNATs that have been both functionally and structurally characterized, we designed a broad-substrate screen that has the capacity to identify a potential substrate for nearly any member of the GNAT superfamily. The screen is composed of 95 different compounds including a variety of antibiotics, metabolites, amino acids, and commercially available peptides. To test the screen, we selected 10 kinetically uncharacterized GNATs with low sequence identity between them (18% or less). Nine of the 10 enzymes are from pathogenic bacteria, including *Agrobacterium tumefaciens, Pseudomonas aeruginosa, Staphylococcus aureus, Clostridium difficile,* and *Vibrio cholerae*, and the 10th is from the thermophilic archaea *Thermoplasma acidophilum.* Using this approach, we have identified several new compounds that can be acetylated by GNATs and at least one potential substrate for further characterization for each GNAT tested. We also learned that some GNATs acetylate more than one class of substrate and therefore potentially have multiple functions. To our knowledge, this combination of substrates has not been used prior to this work. This assay is quick, easy, cost-effective, requires minimal equipment to perform, has been formatted for high-throughput processes, and can be used in a variety of settings, including teaching laboratories. To tackle the problem of substrate identification for members of the GNAT superfamily, a combination of methods should be used. The approach presented here is our contribution.

Results

For this assay, we chose a variety of substrates, including polyamines, amino acids, antibiotics (aminoglycosides, penicillins, polypeptides, sulfonamides, aminonucleosides, and cephalosporins), nucleosides, coenzymes, vitamins, catecholamines, organic building blocks, antioxidants, small peptides, and other metabolites. These compounds were selected based on known substrates or classes of substrates of general acetyltransferases. Using several iterations, nearly 150 compounds from these categories were screened to determine the best substrates to include in one 96-well format assay. Substrates that were included in the final format were chosen based on the following criteria: (1) whether or not activity for any of the 10 proteins was detectable against it, (2) the compound was soluble in an assay-compatible solvent and did not give a high background, (3) if a compound was known to be a substrate for a particular acetyltransferase, and (4) each class of substrate had several representative compounds with differing chemical properties. A detailed list of substrates chosen for this final screen is shown in Table I. To test the possibility that GNATs could perform *O*-acetylation, we performed a second screen of metabolites that lack primary amino groups but contain hydroxyl groups such as sugars and alcohols. None of the enzymes tested acetylated these compounds, but this combination of substrates could be used to study other acetyltransferases that preferentially catalyze *O*-acetylation reactions. Their identities are detailed in Supporting Information Table I.

Since we were testing activity with a wide variety of substrates, we chose an assay method that allowed us to detect the production of CoA regardless of the other substrate involved in the reaction. We detect CoA spectrophotometrically with 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB, also known as Ellman's reagent), which reacts in a 1:1 ratio with the free sulfhydryl of CoA. This method of detection has a few limitations that were considered as we designed our screen. First, we could not include any compound that contains a free sulfhydryl group (e.g. the amino acid cysteine) because it could react with DTNB. Additionally, proteins were dialyzed into buffer lacking reducing agents like β-mercaptoethanol or dithiothreitol to remove the possibility that DTNB would react with their free sulfhydryl groups and increase the background. Second, we could not include the amino acid

Table I. *Plate Layout and Substrates for the Broad-substrate Screen*

Number	Substrate	Number	Substrate
1	Cadaverine	49	Kanamycin B
2	Putrescine	50	Antibiotic G418
3	Spermine	51	Streptomycin
4	N^1-Acetylspermine	52	Tobramycin
5	Spermidine	53	Apramycin
6	Agmatine	54	Ampicillin
7	N^8-Acetylspermidine	55	Sulfacetamide
8	Thialysine	56	Puromycin
9	*N*-α-Acetyl-L-lysine	57	Chloramphenicol
10	*N*-ε-Acetyl-L-lysine	58	Polymyxin B
11	L-Lysine	59	Bacitracin
12	Poly-L-Lysine	60	Colistin
13	L-Glutamine	61	Blasticidine S
14	*N*-α-Acetyl-L-glutamine	62	Cephalexin
15	L-Asparagine	63	7-Aminocephalosporanic acid
16	*N*-α-Acetyl-L-asparagine	64	Adenosine
17	L-Arginine	65	Guanosine
18	*N*-α-Acetyl-L-arginine	66	Cytidine
19	L-Aspartic acid	67	NAD
20	*N*-Acetyl-aspartic acid	68	NADP
21	L-Glutamic acid	69	Thiamine pyrophosphate
22	*N*-Acetyl-glutamic acid	70	Cyclic AMP
23	Glycine	71	Urea
24	*N*-Acetylglycine	72	Guanidine
25	L-Alanine	73	L-Ornithine
26	*N*-Acetyl-L-alanine	74	Allantoin
27	L-Serine	75	Pyridoxamine
28	*N*-Methyl-L-serine	76	Nicotinamide
29	*O*-Acetyl-L-serine	77	Thiamine
30	L-Threonine	78	Folic acid
31	*N*-Methyl-L-threonine	79	4-Aminobenzoic acid
32	L-Methionine	80	AICAR
33	*N*-Acetyl-L-methionine	81	Chitosan
34	L-Tryptophan	82	D-Glucosamine
35	*N*-Acetyl-L-tryptophan	83	Glucosamine 6-phosphate
36	L-Phenylalanine	84	Dopamine
37	*N*-Acetyl-L-phenylalanine	85	(−)-Epinephrine
38	L-Tyrosine	86	Tyramine
39	*N*-Acetyl-L-tyrosine ethyl ester	87	Serotonin
40	L-Valine	88	Creatine
41	L-Isoleucine	89	4-Aminobutyrate ethyl ester
42	*N*-Acetyl-isoleucine	90	5-Amino-4-imidazolecarboxamide
43	L-Leucine	91	Pterine
44	*N*-Acetyl-L-leucine	92	(−)-Glutathione (oxidized)
45	L-Homoserine	93	*N*-Phenylacetyl-Gly-Lys
46	L-Citrulline	94	Acetyl-Ser-Asp-Lys-Pro
47	Neomycin	95	Asp-Phe methyl ester (aspartame)
48	Gentamicin	96	Enzyme blank

histidine in the screen because it can hydrolyze AcCoA in the absence of enzyme. Third, we performed assays at pH 8.0 to increase the reactivity of amino groups while limiting the background hydrolysis of DTNB at more alkaline pH.[12,13] This pH is optimal for this assay and also a near optimal choice for the activity of GNATs we subsequently characterized. Finally, compounds with an inherently intense yellow or orange color like riboflavin, tetracycline, and FAD were tested, but were not included in the final screen due to their high contaminating absorbance at 415 nm. It was imperative to remove the polyhistidine tag from the purified proteins before screening because for many GNATs we have observed that the affinity for AcCoA decreased significantly in its presence (unpublished data).

All proteins that were used in this study had not been characterized prior to this work, and the information regarding their possible functions was limited to the following annotations: ATU2258 (acetyltransferase), CD1211 (*N*-acetyltransferase), PA2271 (GNAT acetyltransferase), PA2578 (acetyltransferase including *N*-acetylases of ribosomal proteins), PA3944 (hypothetical protein, GNAT acetyltransferase), PA5475 (hypothetical protein, GNAT acetyltransferase), SAOL0519 (acetyltransferase GNAT family), SACOL 1063 (acetyltransferase GNAT family), Ta0374 (hypothetical protein, *N*-acetyltransferase superfamily),

Table II. *Specific Activity[a] (nmol/min/mg) of each GNAT Toward Selected Substrates*

Substrate	SACOL1063	VCA0947	CD1211	ATU2258	PA3944	PA2271	PA2578	PA5475
Spermine		1400				81.9		
N^1-Acetylspermine		497						
Spermidine		1400				134		
N^8-Acetylspermidine				109		61.5		
Thialysine				125				
N-α-Acetyl-L-glutamine					81.9			
Glycine					155			
L-Serine			1150					
O-Acetyl-L-serine			605					
L-Threonine	1250		1440					
L-Methionine			981					
L-Tryptophan	303		134					
L-Tyrosine	50.9							
L-Valine			57.5					
L-Isoleucine			308					
L-Leucine			256					
L-Homoserine			64.2					
Puromycin				58.6	76.3			
Chloramphenicol				84.1	69.7		891	180
Polymyxin B				58.6	385			
Colistin					357	10.0		
7-Aminocephalosporanic acid			149			83.0		
Thiamine pyrophosphate			73.0					
Thiamine			117			55.3		
D-Glucosamine					57.5			
Glucosamine 6-phosphate				236				
Dopamine				214	161			
Tyramine								
Serotonin					75.2			
4-Aminobutyrate ethyl ester				62.0				
N-Phenylacetyl-Gly-Lys								
Asp-Phe methyl ester					822			

[a] The concentration of free thiol of CoA produced during the assay was determined via its reaction with DTNB, which was used to calculate the specific activity of the enzyme towards each substrate.

and VCA0947 (spermine/spermidine N1-acetyltrans ferase).

A wide range of activity (low μmol/min/mg to low nmol/min/mg) was detected for the 10 GNATs, and the difference in activity between the enzymes with highest and lowest activity was two orders of magnitude. Enzymes displaying the highest activities were CD1211, VCA0947, and SACOL1063 (∼1200–1400 nmol/min/mg), PA2578 and PA3944 (∼800–900 nmol/min/mg), and ATU2258, PA5475, and PA2271 (∼100–300 nmol/min/mg). The enzymes with the lowest activity against all tested substrates were TA0374 (<10 nmol/min/mg) and SACOL0519 (∼10 nmol/min/mg). A substrate was considered to be a good candidate for further testing if the enzyme produced a specific activity of 50 nmol/min/mg or higher in its presence (Table II). Data for each enzyme against all substrates is shown in Supporting Information Table II.

The GNATs screened showed a myriad of activities. Some had a preference for one specific class of substrate, while others were promiscuous amongst classes. Enzymes with a preference for only one class of substrate were SACOL1063, VCA0947, PA2578, and PA5475. Those that used multiple classes were CD1211, ATU2258, PA3944, and PA2271 (Table III). The highest activities were seen for six classes of substrates, including polyamines, amino acids, antibiotics, catecholamines, peptides, and other metabolites. A summary of class preferences for each enzyme is shown in Table III.

SACOL1063 from *Staphylococcus aureus* preferred acetylating amino acids and displayed the highest activity for L-threonine. It also could acetylate L-tryptophan, and to a much lesser extent L-tyrosine. VCA0947 from *Vibrio cholerae* preferred the polyamines spermine or spermidine and could acetylate N^1-acetylspermine, which was consistent with its functional annotation. CD1211 from *Clostridium difficile* primarily used amino acids L-threonine, L-serine, L-methionine, *O*-acetyl- L-serine, L-isoleucine, and L-leucine, but could use antibiotics and coenzymes or vitamins to a lesser degree. ATU2258 from *Agrobacterium tumefaciens* preferred the metabolite glucosamine 6-phosphate or the catecholamine dopamine. It also could use the lysine analog thialysine, the polyamine N^8-acetylspermidine, antibiotics chloramphenicol, puromycin, and polymyxin B, and the organic building block 4-aminobutyrate ethyl ester. PA3944 from *Pseudomonas aeruginosa* preferred the peptide Asp-Phe methyl ester (or aspartame) and the peptide antibiotics polymyxin

Table III. *Summary of Substrate Class Preferences for GNATs Tested*

Protein	Polyamine	Lysine or analog	Amino acid	Antibiotic	Coenzyme or vitamin	Catecholamine	Building blocks	Other metabolite	Peptide
SACOL1063			1						
VCA0947	1								
CD1211			1	2	3				
ATU2258	2	2		3		1	4	1	
PA3944			3	2		3		4	1
PA2271	1			2	3				
PA2578				1					
PA5475				1					

The number indicates the substrate preferences for each GNAT, where the highest to lowest preference are: (1) > (2) > (3) > (4).

B and colistin. Other substrates like the catecholamine dopamine, and amino acids glycine and *N*-α-acetyl-L-glutamine were used and displayed lower activity. PA2271 from *Pseudomonas aeruginosa* used a variety of substrates including the polyamines like spermidine and spermine, the cephalosporin building block 7-aminocephalosporanic acid, and vitamins and cofactors thiamine and thiamine pyrophosphate. PA2578 and PA5475 from *Pseudomonas aeruginosa* both preferred the antibiotic chloramphenicol. This result is inconsistent with the functional annotation of PA2578; however, further testing against peptides that mimic ribosomal proteins would be necessary to confirm this discrepancy in annotation.

More intensive studies regarding optimal assay conditions (i.e. pH, temperature, and substrate concentration) for the SACOL0519 and TA0374 enzymes are necessary to determine their kinetic parameters and confirm their substrate utilization. SACOL0519 from *Staphylococcus aureus* was only active against the catecholamine dopamine (10.3 nmol/min/mg), which is below the lower limit of the assay. Additionally, the TA0374 enzyme reproducibly displayed activities lower than 10 nmol/min/mg toward amino acids L-tryptophan, L-leucine, and L-lysine, and catecholamines dopamine and tyramine. This enzyme is from an acidophilic archaea that has a preference for acidic pH and high temperatures, which are not the general assay conditions of this screen. However, the screen is versatile enough that the assay conditions can be adjusted to meet the specific requirements of a particular protein.

Discussion

Our goal was to develop a screen to identify potential substrates for uncharacterized GNATs from a variety of organisms to aid in functional and structural characterization of 10 members of this superfamily. Analyzing several classes of substrates, we have found at least one substrate that can be used for further studies for nearly all of the enzymes tested. We propose that these enzymes display characteristics of the following types: N-terminal acetyltransferases (NATs), histone/protein lysine acetyltransferases (KATs), spermidine/spermine N1-acetyltransferases (SSATs), Glucosamine 6-phosphate *N*-acetyltransferases, arylalkylamine *N*-acetyltransferases (aaNATs), puromycin *N*-acetyltransferases, and possibly chloramphenicol acetyltransferases (CATs).

Protein acetylation by GNATs occurs via N-terminal amino acid (Nα) acetylation, *N*ε-lysine acetylation, or *O*-acetylation of serine and threonine residues on proteins (summarized in Yang *et al.*[14]). N-terminal protein acetylation typically occurs on methionine, alanine, glycine, serine, or threonine residues that are exposed after methionine aminopeptidase cleavage. This type of cleavage usually occurs when these residues precede glycine, alanine, serine, threonine, proline, valine, or cysteine residues. In contrast, the peptidase does not cleave after methionine when aspartic acid, glutamic acid, phenylalanine, histidine, isoleucine, lysine, leucine, methionine, asparagine, glutamine, arginine, tryptophan, or tyrosine is second; thus, these amino acids are not likely to be *in vivo* substrates for Nα acetylation.[15] In this study, we have identified GNATs that can acetylate amino acids. Some of these GNATs could acetylate multiple amino acids, while others acetylated a limited number. The combination of amino acids that were acetylated by the GNATs tested were L-threonine, L-serine, L-methionine, glycine, L-leucine, L-isoleucine, L-valine, L-tyrosine, and L-tryptophan. Since the GNATs were unable to *O*-acetylate the *N*-methylated or *N*-acetylated versions of the amino acids, we assume that these enzymes are involved in N-terminal acetylation of proteins rather than *O*-acetylation of serine or threonine. The physiological relevance of L-leucine, L-isoleucine, L-tyrosine, and L-tryptophan *N*-acetylation needs to be assessed since these residues are typically not exposed after methionine aminopeptidase activity. Four of the proteins tested were capable of acetylating thialysine, 4-aminobutyrate ethyl ester, and/or the peptides *N*-phenylacetyl-Gly-Lys and Asp-Phe methyl ester (also known as aspartame). *N*ε-acetylation of lysine, peptides, and the lysine analog thialysine is also a known function of GNATs and we hypothesize that these likely function as *N*ε-lysine acetyltransferases. 4-aminobutyrate ethyl ester is an organic building block that is structurally

similar to L-lysine and offers a potential explanation as to why this substrate was acetylated.

None of the GNATs we studied acted like aminoglycoside *N*-acetyltransferases (AgNATs) by acetylating aminoglycosides. However, our screen has revealed a new class of antibiotics that can be acetylated *in vitro*. These include the cationic peptide antibiotics polymyxin B and colistin, which were acetylated by GNATs from Gram-negative bacteria. These antibiotics primarily target Gram-negative bacteria by permeabilizing their cell membranes, so it is unclear if the GNATs tested here are properly localized *in vivo* to acetylate the antibiotics. These enzymes may be acetylating polymyxin B and colistin *in vitro* because they mimic cationic peptides of proteins rather than actually acetylating the antibiotics *in vivo*. However, bacteria are becoming resistant to polymyxin B and colistin and the mechanism of resistance is not completely clear. Clinical strains of *Pseudomonas aeruginosa* have recently been identified as acquiring resistance to these antibiotics,[16,17] and a lipid IVA acyltransferase gene *msbB* from *Vibrio cholerae* was shown to be necessary for polymyxin B resistance in clinical strains.[18] Therefore, more thorough characterization of the GNATs studied here is necessary and will be performed in future work to ascertain their true *in vivo* role.

Some enzymes we studied were capable of acetylating 7-aminocephalosporanic acid, glucosamine 6-phosphate, puromycin, or dopamine. 7-aminocephalosporanic acid is the hydrolysis product of cephalosporin C, which is found in *Streptomyces* and is used chemically as a building block to produce cephalosporin derivatives. It is therefore reasonable that this moiety is acetylated because it may be an endogenous function for certain bacteria. Acetylation of glucosamine 6-phosphate has been identified previously for GNATs, and *N*-acetyl-D-glucosamine 6-phosphate is used in UDP-*N*-acetylglucosamine biosynthesis for the bacterial cell wall. To our knowledge, bacterial GNATs that can acetylate puromycin have only been identified in *Streptomyces*. The puromycin *N*-acetyltransferase was found in *Streptomyces alboniger*[19] and has been used as a reporter gene, similar to CAT reporter genes. Additionally, an *N*-acetyltransferase that can acetylate L-DOPA (L-dihydroxyphenyalanine), a precursor of dopamine, was found in *Streptomyces akiyoshiensis*.[20] However, we do not know of a bacterial *N*-acetyltransferase that uses dopamine, and the ones described only exist in eukaryotes.

On the surface, the physiological relevance of bacterial dopamine (or other catecholamine) acetylation seems implausible. However, several studies have shown that these compounds exist even in organisms without a nervous system (e.g. bacteria, fungi, and plants; reviewed in Roshchina[21]). It is believed that these compounds may be used as signaling molecules for communication, as a defense mechanism, or for growth and development (see Roshchina[21]). In fact, plants produce increased amounts of dopamine under high stress conditions[22] and it is believed that some plants have used it to resist fungal infection.[23] It has been shown that dopamine can inhibit bacterial growth, whereas *N*-acetyldopamine is ineffective.[24] Therefore, some bacteria like the plant pathogen *Agrobacterium* (this study) may have developed the ability to *N*-acetylate dopamine as a mechanism for survival.

We were surprised by the results that PA2578 and PA5475 both preferentially *O*-acetylate chloramphenicol since they do not have sequence similarity with the well characterized CATs. The fact that they acetylate chloramphenicol but none of the other compounds we tested for *O*-acetylation (Supporting Information Table II) may indicate that these GNATs indeed act as CATs. Although a thorough kinetic characterization is necessary to validate the results and determine the mechanism of *O*-acetylation, it may be plausible from recent literature that GNATs can perform either *N*- and/or *O*-acetylation reactions. For instance, an aminoglycoside acetyltransferase from the bacterium *Streptomyces albulus* was shown to both *N*- and *O*-acetylate aminoglycoside antibiotics when an amino or hydroxyl group was present at the 6′ position.[25] Also, an AgNAT from *Mycobacterium tuberculosis* can both *N*- and *O*-acetylate certain aminoglycosides at the 2′ position.[26] In humans, an *N*-acetyltransferase 2 in hepatocytes was identified that can catalyze both *N*- and *O*-acetylation reactions,[27] and the mechanism of an arylamine *N*-acetyltranferase that *O*-acetylates hydroxylated heterocyclic amines in humans has been studied computationally.[28] In general, we have assumed that GNATs should *N*-acetylate substrates, but it seems *O*-acetylation may be more common than once thought.

Although it has been shown that GNATs are capable of both *N*- and *O*-acetylation, it is not known if GNATs are also capable of *S*-acetylation. Unfortunately, we were not able to test this possibility against any compounds with free sulfhydryl groups due to the limitations of this assay. Therefore, further experiments to test acetyltransferases for *S*-acetylation should be performed using a different assay. This would require a more detailed kinetic analysis because of the potential for a high background due to non-enzymatic acetylation of either free cysteine or cysteine-containing peptides.[29]

Polyamine acetylation is also a known function for GNATs. Three of the acetyltransferases tested in this study could use polyamines as substrates, two of which preferred spermidine or spermine as substrates. We obtained contradictory results for the archaeal protein TA0374 regarding its proposed functional role. NMR studies of TA0374 showed it was possible to acetylate spermidine at 5 m*M* substrate concentration (AcCoA and spermidine). However, a full kinetic characterization was not performed and

spermidine was not present in the active site of the crystal structure.[30] We did not observe activity with spermine and spermidine even at higher concentrations of the polyamine or enzyme (data not shown). Instead, we detected activity against amino acids and catecholamines. To confirm these findings, additional assays are necessary to determine whether this enzyme is more active at lower pH and higher temperatures, as well as if it could potentially acetylate polyamines under these conditions. Our screening method is flexible enough and the assay conditions can be adjusted to accommodate these types of preferences.

From our results, it seems there is an array of GNATs present in each bacterium that catalyze a diverse set of acetylation reactions, and some may perform multiple functions. This diversity implies that bacteria regularly perform a series of choreographed acetylation reactions to maintain cellular homeostasis and enable survival. Our broad screen provides a starting point for further kinetic characterization to determine the potentially physiologically relevant ligand(s) for each enzyme. At minimum, it narrows the possible classes of substrates for analysis to a more manageable number. Once a substrate is known for an enzyme, the screen can be tailored to fit that particular type of GNAT. For instance, buffer, salt concentration, pH, time of reaction, temperature, and types of substrates that will be used for additional characterizations can be adjusted. This rapid type of approach can also be extended to and tailored to study enzymes from other classes where their structures may be known but their specific function and specificity is not known. These types of methods to functionally annotate large numbers of proteins is becoming more and more important as we quickly increase the number of three-dimensional structures that are being determined. A full characterization of each enzyme is then necessary to determine which potential substrates are preferred and/or physiologically relevant. Further structural and functional characterization is currently underway for the enzymes used in this study and will be detailed in subsequent publications.

Materials and Methods

Gene cloning and protein purification

All genes were cloned into the pMCSG7 vector at JCVI using previously described protocols[31,32] and expressed in BL21-CodonPlus(DE3)-RIL cells (Stratagene). The vector is ampicillin resistant and produces protein with an N-terminal poly-histidine tag followed by a tobacco etch virus (TEV) protease cleavage site, and the expression cells are chloramphenicol resistant. All proteins are targets from either the Center for Structural Genomics of Infectious Diseases (CSGID) or the Midwest Center for Structural Genomics (MCSG) and are identified by the target ID where IDP codes for targets in CSGID and APC for MCSG. The following clones were selected for characterization: SACOL1063 (IDP00739) *Staphylococcus aureus* subsp. aureus COL, SACOL0519 (IDP00844) *Staphylococcus aureus* subsp. aureus COL, VCA0947 (IDP01616) *Vibrio cholerae* O1 biovar El Tor str. N16961, CD1211 (IDP05291) *Clostridium difficile* 630, ATU2258 (APC5884) *Agrobacterium tumefaciens* str. C58, PA3944 (APC6869) *Pseudomonas aeruginosa* PA01, PA2271 (APC6926) *Pseudomonas aeruginosa* PA01, PA2578 (APC6946) *Pseudomonas aeruginosa* PA01, PA5475(APC6949) *Pseudomonas aeruginosa* PA01, and Ta0374 (APC61169.1) *Thermoplasmaacidophilum* DSM 1728. All sequences are in Supporting Information Table III.

Cells containing the plasmid for expression were grown in terrific broth (TB), expressed, and harvested as previously described.[33] After centrifugation, cells from 1.5 L culture were resuspended in 100 mL lysis buffer composed of 1.5 m*M* magnesium acetate, 1 m*M* $CaCl_2$, 250 m*M* NaCl, 100 m*M* ammonium sulfate, 40 m*M* Na_2HPO_4, 3.25 m*M* citric acid, 5% (v/v) glycerol, 5 m*M* imidazole, and 0.08% (w/v) *n*-dodecyl β-D-maltoside (DDM). Sonication and centrifugation to obtain the clarified lysate were performed like before.[33] Proteins were purified using 5 mL HisTrapFF columns (GE Healthcare) on an AKTAxpress system (GE Healthcare) using the following protocol. Lysate was loaded onto the column at a flow rate of 5 mL/min, washed with loading buffer (10 m*M* TrisHCl pH 8.3 and 500 m*M* NaCl), washed with loading buffer plus 25 m*M* imidazole, and eluted with loading buffer plus 500 m*M* imidazole. The eluted protein was then loaded onto a HiPrep200 26/60 gel-filtration column (GE Healthcare) in loading buffer at a flow rate of 3.2 mL/min and fractions were collected in a 96 deep well block. The purity of the protein was determined via SDS-PAGE and all were purified to near homogeneity. Following purification, the N-terminal poly-histidine tag was removed using a TEV protease with an N-terminal poly-histidine tag.[34] Purified protein was combined with TEV protease in a 20:1 ratio in a 10K MWCO Slide-A-Lyzer dialysis cassette (Pierce) and dialyzed in 1 L of loading buffer plus 5 m*M* β-mercaptoethanol and 1 m*M* EDTA at 37°C for 1 hour then 4°C overnight. The following day, the buffer was changed to loading buffer using three changes of buffer (1 L) every 30 minutes. The cleaved protein was purified by loading the dialysate onto a 5 mL HisTrapFF column, washing with loading buffer, and then washing with loading buffer plus 25 m*M* imidazole until the absorbance of the peak at 280 nm was near zero. TEV and the poly-histidine tag from the cleaved protein were bound to the column and the eluate contained the cleaved protein of interest. The cleaved protein was then concentrated to 10–20 mg/mL, aliquoted, and stored at −80°C until use. Extinction coefficients and molecular weights of the proteins are listed in Supporting Information Table III.

Chemicals and stock solutions

All chemicals were purchased from Sigma-Aldrich at the highest quality available. Compounds used for the screen and the information regarding their stock concentrations and how they were solubilized are included in Supporting Information Table IV. Some of the compounds are toxic and need to be handled carefully; therefore, all chemicals were prepared in a hood or with a mask to ensure that they were not inhaled. Proper waste disposal was necessary, especially the Ellman's reagent which is toxic to aquatic life.

Enzyme assay

Assays were performed in 96-well clear polystyrene flat-bottom plates in a reaction volume of 50 μL. The reaction mixture contained 50 m*M* TrisHCl pH 8.0, 0.5 m*M* AcCoA, and 2.5 m*M* substrate. The exceptions were 1.9 and 0.4 m*M* for AICAR and acetyl-Ser-Asp-Lys-Pro, respectively. Additionally, the final concentration of poly-L-lysine and chitosan in the reactions was 4 mg/mL. One microgram of enzyme was used to initiate the reaction, and it was diluted in 50 m*M* TrisHCl pH 8.0 and 500 m*M* NaCl to a concentration of 0.1 mg/mL directly before use. The reaction proceeded for 10 min at 33°C and was stopped using 50 uL of 100 m*M* TrisHCl pH 8.0 and 6*M* guanidine HCl.[11] To detect the product CoA, 200 uL of a solution of 0.2 m*M* 5,5′-Dithiobis(2-nitrobenzoic acid), 100 m*M* TrisHCl pH 8.0, and 1 m*M* EDTA was added to each well for 10 min at room temperature. The absorbance was measured at 415 nm in a Biotek microplate reader equipped with filters. A separate microplate contained standards of CoA in duplicate for each assay, which were in the linear range of detection. All assays were performed in duplicate with an error of $<10\%$ between the two trials. Two types of plates with identical assay conditions were used for each assay; one type of plate was the control plate without enzyme and the other was the reaction plate with enzyme. Proper controls were included within these two types of plates. The control plate contained a blank with no substrate and no enzyme (to detect spontaneous hydrolysis of AcCoA) as well as a blank for each substrate plus AcCoA in the absence of enzyme. The reaction plate contained a blank of enzyme in the presence of AcCoA with no additional substrate (some enzymes hydrolyze AcCoA in the absence of other substrates). All reactions were set by hand using multi-channel pipettors; however, the design is formatted for use with automatic handlers.

We noticed that if the reactions containing dopamine or serotonin were incubated with the stop solution for long periods of time a background purple/brown color appeared. This color change is likely due to the fact that catecholamines become oxidized at alkaline pH and their oxidation products produce colored compounds that can polymerize or react with free sulfhydryls on the protein to form colored precipitate.[35–37] Therefore, it was necessary to stop these reactions, immediately add the Ellman's reagent, and measure the absorbance for the reactions within the specified timeframe.

Electronic supplementary material

Four supporting information tables are included in the electronic edition of the journal. A list of compounds used to test *O*-acetylation is in Supporting Information Table I. For each GNAT studied the activity in nmol/min/mg for each substrate is included in Supporting Information Table II. The protein sequences, locus tags, and NCBI accession numbers, extinction coefficients and molecular weights for each protein are included in Supporting Information Table III. Finally, a table with information regarding stock solution concentrations, the method for solubilization, and Sigma product numbers for each compound in the broad-substrate screen is included in Supporting Information Table IV.

Acknowledgments

This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN272200700058C. We would like to sincerely thank Sergii Pshenychnyi, Ievgeniia Dubrovska, and James R. Winsor for their technical assistance, Ekaterina Filippova and Daniel Becker for their helpful discussions, and Alan Wolfe for his revisions of the manuscript.

References

1. Vetting MW, de Carvalho LPS, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433:212–226.
2. Dyda F, Klein DC, Hickman AB (2000) Gcn5-related N-acetyltransferases: A structural overview. Annu Rev Biophys Biomol Struct 29:81–103.
3. Neuwald AF, Landsman D (1997) Gcn5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. Trends Biochem Sci 22:154–155.
4. Yu M, de Carvalho LPS, Sun G, Blanchard JS (2006) Activity-based substrate profiling for Gcn5-related N-acetyltransferases: The use of chloroacetyl-coenzyme A to identify protein substrates. J Am Chem Soc 128: 15356–15357.
5. Vetting MW, Magnet S, Nieves E, Roderick SL, Blanchard JS (2004) A bacterial acetyltransferase capable of regioselective N-acetylation of antibiotics and histones. Chem Biol 11:565–573.
6. Thao S, Escalante-Semerena JC (2011) Biochemical and thermodynamic analyses of *Salmonella enterica* Pat, a multidomain, multimeric *N*ε-lysine acetyltransferase involved in carbon and energy metabolism. MBio 2.
7. Ronen M, Biochemical and structural characterization of recombinant histone acetyltransferase proteins. In:

Allis CD, Carl W, Eds. (2003) Methods in Enzymology. Academic Press, New York, pp 106–119.
8. Wybenga-Groot LE, Draker K-A, Wright GD, Berghuis AM (1999) Crystal structure of an aminoglycoside 6′-N-acetyltransferase: Defining the Gcn5-related N-acetyltransferase superfamily fold. Structure 7:497–507.
9. Lin H-J, Lien Y-C, Hsu C-H (2010) A high-throughput colorimetric assay to characterize the enzyme kinetic and cellular activity of spermidine/spermine N1-acetyltransferase 1. Anal Biochem 407:226–232.
10. Brooke EW, Davies SG, Mulvaney AW, Pompeo F, Sim E, Vickers RJ (2003) An approach to identifying novel substrates of bacterial arylamine N-acetyltransferases. Biorrg Med Chem 11:1227–1234.
11. Zhou Y, Xin Y, Sha S, Ma Y (2011) Kinetic properties of *Mycobacterium tuberculosis* bifunctional GlmU. Arch Microbiol 193:751–757.
12. George L E (1958) A colorimetric method for determining low concentrations of mercaptans. Arch Biochem Biophys 74:443–450.
13. Danehy JP, Elia VJ, Lavelle CJ (1971) Alkaline decomposition of organic disulfides. IV. Limitation on the use of Ellman's reagent. 2,2′-dinitro-5,5′-dithiodibenzoic acid. J Org Chem 36:1003–1005.
14. Yang Y-Y, Hang HC (2011) Chemical approaches for the detection and synthesis of acetylated proteins. ChemBioChem 12:314–322.
15. Walker KW, Bradshaw RA (1999) Yeast methionine aminopeptidase I. J Biol Chem 274:13403–13409.
16. Lee J-Y, Lim MH, Heo ST, Ko KS (2012) Repeated isolation of *Pseudomonas aeruginosa* isolates resistant to both polymyxins and carbapenems from one patient. Diagn Microbiol Infect Dis 72:267–271.
17. Fernández L, Gooderham WJ, Bains M, McPhee JB, Wiegand I, Hancock REW (2010) Adaptive resistance to the "last hope" antibiotics polymyxin B and colistin in *Pseudomonas aeruginosa* is mediated by the novel two-component regulatory system ParR-ParS. Antimicrob Agents Chemother 54:3372–3382.
18. Matson JS, Yoo HJ, Hakansson K, DiRita VJ (2010) Polymyxin B resistance in El Tor *Vibrio cholerae* requires lipid acylation catalyzed by MsbB. J Bacteriol 192:2044–2052.
19. Vara J, Perez-Gonzalez JA, Jimenez A (1985) Biosynthesis of puromycin by *Streptomyces alboniger*: Characterization of puromycin N-acetyltransferase. Biochemis try 24:8074–8081.
20. Rodriguez-Juarez RC. Isolation, partial purification and characterization of an N-acetyltransferase from *Streptomyces akiyoshiensis* L-138. (2003) Department of Chemistry and Biochemistry. University of Lethbridge, Lethbridge, Alberta, Canada, Master Thesis.
21. Roshchina VV, Evolutionary considerations of neurotransmitters in microbial, plant, and animal cells. In: Lyte M, Freestone PPE, Eds. (2010) Microbial endocrinology: Interkingdom signaling in infectious disease and health. New York: Springer, pp 17–52.
22. Swiedrych A, Lorenc-Kukula K, Skirycz A, Szopa J (2004) The catecholamine biosynthesis route in potato is affected by stress. Plant Physiol Biochem 42:593–600.
23. Odjakova M, Hadjivanova C (1997) Animal neurotransmitter substances in plants. Bulg J Plant Physiol 23: 94–102.
24. Zhao P, Li J, Wang Y, Jiang H (2007) Broad-spectrum antimicrobial activity of the reactive compounds generated in vitro by *Manduca sexta* phenoloxidase. Insect Biochem Mol Biol 37:952–959.
25. Hamano Y, Hoshino Y, Nakamori S, Takagi H (2004) Overexpression and characterization of an aminoglycoside 6′-*N*-acetyltransferase with broad specificity from an ε-poly-L-lysine producer, *Streptomyces albulus* IFO 14147. J Biochem 136:517–524.
26. Vetting MW, Hegde SS, Javid-Majd F, Blanchard JS, Roderick SL (2002) Aminoglycoside 2′-N-acetyltransferase from *Mycobacterium tuberculosis* in complex with coenzyme A and aminoglycoside substrates. Nat Struct Mol Biol 9:653–658.
27. Doll MA, Zang Y, Moeller T, Hein DW (2010) Codominant expression of N-acetylation and O-acetylation activities catalyzed by N-acetyltransferase 2 in human hepatocytes. J Pharmacol Exp Ther 334:540–544.
28. Lau EY, Felton JS, Lightstone FC (2006) Insights into the O-acetylation reaction of hydroxylated heterocyclic amines by human arylamine N-acetyltransferases: A computational study. Chem Res Toxicol 19:1182–1190.
29. Dormeyer W, Ott M, Schnölzer M (2005) Probing lysine acetylation in proteins. Mol Cell Proteomics 4: 1226–1239.
30. Filippova EV, Shuvalova L, Minasov G, Kiryukhina O, Zhang Y, Clancy S, Radhakrishnan I, Joachimiak A, Anderson WF (2011) Crystal structure of the novel PaiA N-acetyltransferase from thermoplasma acidophilum involved in the negative control of sporulation and degradative enzyme production. Proteins 79: 2566–2577.
31. Dieckman L, Gu M, Stols L, Donnelly MI, Collart FR (2002) High throughput methods for gene cloning and expression. Protein Express Purif 25:1–7.
32. Stols L, Gu M, Dieckman L, Raffen R, Collart FR, Donnelly MI (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. Protein Express Purif 25: 8–15.
33. Sanville Millard C, Stols L, Quartey P, Kim Y, Dementieva I, Donnelly MI (2003) A less laborious approach to the high-throughput production of recombinant proteins in *Escherichia coli* using 2-liter plastic bottles. Protein Express Purif 29:311–320.
34. Kapust RB, Waugh DS (2000) Controlled intracellular processing of fusion proteins by TEV protease. Protein Express Purif 19:312–318.
35. Bisaglia M, Mammi S, Bubacco L (2007) Kinetic and structural analysis of the early oxidation products of dopamine. J Biol Chem 282:15597–15605.
36. Richter HW, Waddell WH (1983) Mechanism of the oxidation of dopamine by the hydroxyl radical in aqueous solution. J Am Chem Soc 105:5434–5440.
37. Walker JRL, Ferrar PH (1998) Diphenol oxidases, enzyme-catalysed browning and plant disease resistance. Biotechnol Genet Eng Rev 15:457–498.