Annu. Rev. Biophys. Biomol. Struct. 2000. 29:81–103
Copyright © 2000 by Annual Reviews. All rights reserved

GCN5-Related N-Acetyltransferases: A Structural Overview

Fred Dyda,[1] David C. Klein,[2] and Alison Burgess Hickman[1]

[1]*Laboratory of Molecular Biology, National Institute of Diabetes, Digestive, and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892; e-mail: dyda@ulti.niddk.nih.gov, ahickman@helix.nih.gov*

[2]*Laboratory of Developmental Neurobiology, National Institute of Child Health and Development, National Institutes of Health, Bethesda, Maryland 20892; e-mail: klein@helix.nih.gov*

Key Words melatonin, histone acetylation, GNAT, aminoglycoside, acetyl coenzyme A

■ **Abstract** Hundreds of acetyltransferases exist. All use a common acetyl donor—acetyl coenzyme A—and each exhibits remarkable specificity for acetyl acceptors, which include small molecules and proteins. Analysis of the primary sequences of these enzymes indicates that they can be sorted into several superfamilies. This review covers the three-dimensional structures of members of one of these superfamilies, now referred to in the literature as the GCN5-related N-acetyltransferases (GNAT), reflecting the importance of one functional category, the histone acetyltransferases. Despite the diversity of substrate specificities, members of the GNAT superfamily demonstrate remarkable similarity in protein topology and mode of acetyl coenzyme A binding, likely reflecting a conserved catalytic mechanism.

CONTENTS

INTRODUCTION 82
Summary of Structures Determined to Date 84
Sequence and Structure-Based Alignment of GNAT Enzymes 86
DESCRIPTION OF THE STRUCTURE 86
Overview of the Structural Fold 86
The N-Terminal Region and Motif C 88
Motifs D and A 89
Motif B and the Completion of the AcCoA Binding Site 93
SUBSTRATE BINDING 94
The Conformation of Bound AcCoA/CoA 94
Acceptor Substrate Binding 95
CHEMICAL CATALYSIS 96

1056-8700/00/0610-0081$14.00

The Need for a General Base? 96
The Role of Proton Wires in Amine Deprotonation 98
Substrate Positioning 99
SUMMARY 100

INTRODUCTION

The transfer of an acetyl group from one molecule to another is a fundamental biochemical process. Several unrelated classes of enzymes catalyze such a reaction, and the focus of this review is the recent burst of three-dimensional structural information about one of these classes, the GCN5-related N-acetyltransferases (GNAT), which catalyze the transfer of the acetyl group from acetyl coenzyme A (AcCoA, the "donor") to a primary amine (the "acceptor"). Over the course of a twelve-month period spanning 1998–1999, eleven experimentally determined three-dimensional structures of GNAT superfamily members have been reported, elucidated both by single-crystal X-ray diffraction and by NMR spectroscopy (2, 7, 11, 12, 20, 34, 39, 43, 45). From the perspective of structural biology, the GNAT superfamily is a case study in how a common acetyltransferase domain evolved to serve a wide variety of functions. What makes this superfamily especially interesting is the large assortment of acceptor substrates carrying a primary amine that are capable of being acetylated, and which the different GNAT members have to interact with and recognize.

Even a cursory inspection of three-dimensional protein structures in the Protein Data Bank (http://www.rcsb.org/pdbl) reveals a large number of different ways in which AcCoA and CoA can be bound to protein molecules. In addition to the diverse binding modes, the protein subunits that are responsible for AcCoA binding also display high diversity in their folds (8). The GNAT superfamily adds yet other variations, both to AcCoA conformation and binding modes and to the variety of known protein folds. The polypeptide of the N-acetyltransferase domain that is responsible for AcCoA binding has a unique fold, as noted in the reports of the two first structures that were solved from the superfamily (7, 43). All the subsequent structure determinations have demonstrated nearly the same protein topology and AcCoA binding mode. There is no detectable similarity in either primary sequence or three-dimensional structure between enzymes in the GNAT superfamily and other AcCoA-dependent acetyltransferases such as chloramphenicol acetyltransferase (19) and dihydrolipoyl transacetylase (23). [To avoid confusion, it is important to keep this in mind, particularly since in the three-dimensional database SCOP (29), the family of CoA-dependent acetyltransferases contains only these two enzymes.]

Several years before three-dimensional structural information became available, two conserved sequence motifs (motifs A and B) were noticed between the yeast MAK3 protein and N-acetyltransferases (38); they were later identified in other N-acetyltransferases, including the arylalkylamine N-acetyltransferase (AANAT)

family (3, 22). More recently, Neuwald & Landsman extended and generalized this observation to a large class of >140 related proteins, which they designated the GNAT superfamily (30). They also identified two additional motifs (motifs C and D) present in many GNAT members. The structures reviewed here represent three functional families within the GNAT superfamily and act on three general groups of substrates: histones, aminoglycosides, and arylalkylamines. Aspects of some of these structures have recently been reviewed elsewhere (28, 35).

Among the members of the GNAT superfamily, perhaps the most intense current interest swirls around a group of N-acetyltransferases involved in the acetylation of histones at specific lysine residues, a process that is required for transcriptional activation and that has been implicated in chromatin assembly and DNA replication (10, 27). Of the large number of histone N-acetyltransferase (HAT) enzymes, many turn out to be previously identified transcriptional activator proteins and most likely act in the context of large multiprotein assemblies. Probably the best studied group of HAT proteins are the members of the GCN5 family (indeed, the name of the superfamily derives from these prominent members), on which a wealth of biochemical and functional information is available.

Another important N-acetyltransferase activity is that contributing to the emergence of antibiotic resistance among pathogenic bacteria. N-acetylation of a particular amino group is one of the chemical modifications of aminoglycoside antibiotics that can result in decreased affinity of the drug for its target—a process therefore advantageous for the bacteria. These modifications are carried out by a class of bacterial aminoglycoside N-acetyltransferases (AAC) that can be divided into further subclasses depending on the regiospecificity of acetyl transfer.

Much recent attention has also focused on the arylalkylamine N-acetyltransferase (serotonin N-acetyltransferase, AANAT) family. Members of this family have been found only in vertebrates. AANAT catalyzes the penultimate step in the synthesis of the circadian neurohormone melatonin from serotonin (16). The circulating levels of melatonin are correlated with the light-dark cycle, with high levels of melatonin occurring only at night. Melatonin plays an important role in the proper coordination of the sleep-wake cycle as well as in the adaptation to seasonal changes. There is special interest in AANAT because it exhibits a large light-dark rhythm in activity that controls the rhythm in circulating levels of melatonin. In addition, changes in active protein levels can be very rapid: Light exposure at night causes a decrease in AANAT activity with a half-life of only 3 min (17). AANAT activity has been found to be closely correlated with the amount of AANAT protein. This is regulated by control of proteasomal proteolysis and also by controlling synthesis through regulation of AANAT mRNA levels (9, 16).

Other members of the GNAT superfamily play a variety of anabolic and catabolic roles. For instance, an N-acetyltransferase has been identified that is essential in the synthesis of UDP-N-acetylglucosamine, an essential metabolite in both prokaryotes and eukaryotes (26). An example of catabolic involvement is spermidine/spermine N-acetyltransferase (22). Upon acetylation of the N^1 position of spermidine or spermine, the acetylated polyamines are excreted from the cell or

subsequently metabolized by the FAD-requiring polyamine oxidase. This is by no means an exhaustive list of all the characterized N-acetyltransferases, and there are also several functionally uncharacterized proteins that are clearly members of the GNAT superfamily (3, 30). It is likely that as more sequence information becomes available, the already substantial functional diversity will expand even further.

Summary of Structures Determined to Date

As of mid-1999, there have been eleven reported three-dimensional structures of GNAT superfamily members. Protein particulars, constructs used for structure determination, structural resolution, and Protein Data Bank (PDB, http://www.rcsb.org/pdb/) ID codes are listed in Table 1.

Histone N-Acetyltransferases Of the known structures, seven correspond to histone N-acetyltransferases, or HATs. The structure of Hat1 from *Saccharomyces cerevisiae* (yHat1) was determined complexed with AcCoA (7). The construct used to obtain well-diffracting crystals, despite missing 54 carboxy-terminal residues, is fully active in vitro. PCAF (p300/CBP-associating factor) is a transcriptional activator that can acetylate nucleosomal histone substrates, as well as other transcriptional activators such as p53 (21). Full-length human PCAF is an 832-residue molecule containing an N-terminal domain that interacts with transcriptional activators, a C-terminal bromodomain that can interact with acetylated lysine (for a recent NMR structure, see Reference 5), and a central HAT domain. The crystal

TABLE 1 Structurally characterized members of the GNAT superfamily

Protein[a]	Source	Residues in construct	Substrates	PDB ID	Resolution (Å)	Reference
AANAT	Ovine	28–201	—	1b6b	2.5	(11)
		28–201	Bisubstrate	1cjw	1.8	(12)
Hat1	Yeast	1–320	AcCoA	1BOB	2.3	(7)
GCN5	*Tetrahymena*	47–210	CoA	5GCN	b	(20)
		48–210	—	1QST	1.7	(34)
		48–210	AcCoA	1QSR	2.0	(34)
		48–210	CoA/peptide	1QSN	2.2	(34)
GCN5	Yeast	99–262	—	1ygh	1.9	(39)
PCAF	Human	493–658	CoA	1cm0	2.3	(2)
AAC(3)	*S. marcescens*	1–168	CoA	1bo4	2.3	(43)
AAC(6′)	*E. faecium*	1–182	AcCoA	1B87	2.7	(45)

[a]AANAT, arylalkylamine N-acetyltransferase; PCAF, p300/CBP-associating factor; Hat, histone N-acetyltransferase; AAC, aminoglycoside N-acetyltransferase.

[b]Structure determined by NMR.

structure of this ∼200-residue HAT domain was solved in complex with CoA (2). Highly homologous with the PCAF HAT domain are the GCN5 proteins from various species. The crystal structure of the HAT domain of yeast GCN5 (yGCN5) was solved in absence of substrates (39), whereas the solution structure of the *Tetrahymena* GCN5 HAT domain (tGCN5) has been solved with NMR spectroscopy as a complex with CoA (20). Essentially the same version of tGCN5 was also used in a three-part crystallographic study in which the structure was solved in the uncomplexed form, complexed with AcCoA, and—most importantly—in a ternary complex with CoA and an 11-residue peptide centered around the reactive lysine of yeast histone H3 (34).

Aminoglycoside N-Acetyltransferases The structures of two aminoglycoside N-acetyltransferases exhibiting different specificity for acetyl transfer to aminoglycoside antibiotics have been determined. 3-N-acetyltransferase [AAC(3)] from *Serratia marcescens* was solved complexed with CoA (43). Nine dispensable residues were truncated from the C terminus to obtain diffraction-quality cocrystals. Although no in vitro N-acetyltransferase activity could be demonstrated for this construct, in an in vivo experiment it was deemed sufficient to produce gentamicin resistance. Of all the known GNAT structures, this is the only one that forms a distinguishable dimer in the crystal lattice, with monomers related to each other by a nearly perfect twofold rotation axis. Although the surface area buried by this dimer is relatively large (880 $Å^2$), there is no direct experimental evidence to confirm that the biologically active species is indeed a dimer in solution, and the physiological significance of the observed dimer is unclear. The crystal structure of aminoglycoside 6′-N-acetyltransferase [AAC(6′)] complexed with AcCoA was solved using the full-length protein (45), the only example within the group of currently available GNAT structures for which a nontruncated construct was used. Although analytical gel filtration data indicated that AAC(6′) is a dimer in solution, it was not possible to select the biologically relevant oligomer from the large number of dimers that were created by the high (cubic) crystallographic symmetry. Considering the ambiguities regarding quaternary organization, or that in most cases the biologically relevant unit is a monomer, or that the structures solved involve only the catalytic domains of multidomain proteins, we do not further consider here the issue of oligomerization.

Serotonin N-Acetyltransferase Serotonin N-acetyltransferase (AANAT) represents the third class of GNAT enzymes that have been structurally characterized, and the structure of ovine AANAT has been solved both in the uncomplexed form (11) and complexed with a bisubstrate analog (12). The full-length enzyme contains 207 amino acids but for successful crystallization was truncated at both termini, resulting in a fully active 28–201 construct. Together with the tGCN5 ternary complex, AANAT provides the most complete source of structural information to date, as the bisubstrate analog used in the cocrystallization is a very close chemical approximation of the intermediate that is present during the acetyl

transfer. To date, among the GNAT complexes, this AANAT structure is the one solved at the highest resolution (1.8 Å), owing to the exceptional crystal quality resulting from the stabilizing effect of the bisubstrate compound.

A Structurally Related Cousin Two recent structure determinations of N-myristoyltransferase, one of the uncomplexed enzyme (42) and the other of a ternary complex with myristoyl-CoA and a peptide (1), reveal a myristoyl-CoA binding domain with the same fold and mode of cofactor binding as those of the GNAT superfamily members. It seems likely that a common ancestral protein served as the precursor for both groups of enzymes, but further discussion of N-myristoyltransferase is beyond the scope of this review as it is not specifically an N-acetyltransferase.

Sequence and Structure-Based Alignment of GNAT Enzymes

With the several different three-dimensional structures representing three distinct families within the GNAT superfamily, we are now in a position to reexamine the N-acetyltransferase domain in light of this wealth of structural information. The aligned three-dimensional structures are shown in Figure 1 (see color insert) and the corresponding alignment of primary sequences in Figure 2 (see color insert). The four regions highlighted in both figures in the same colors correspond to the sequence motifs identified first by Tercero et al (38) and subsequently by Neuwald & Landsman (30). As a result of this two-step historical evolution, the order in which the conserved motifs follow each other in the primary sequences is C, D, A, and B. The most impressive conservation across the superfamily is in motifs A and B, whereas the weakest is in motif C. These four regions roughly comprise what we refer to as the N-acetyltransferase domain.

It is most remarkable to observe that, despite the functional variation across the superfamily, the protein topology is nearly identical. This is a clear example of the observation made by Petsko et al (31) that protein families seem to evolve from ancestral molecules that catalyze the necessary chemistry (in this case, transfer of an acetyl group from AcCoA), and that further mutations that do not change the topology will enable the protein to bind alternative substrates. As we shall see, deviations from the common topology of the GNAT superfamily members are most likely related to the need to accommodate a wide range of primary amine-carrying acceptor substrates, which varies from small compounds such as serotonin to large protein molecules such as histones.

DESCRIPTION OF THE STRUCTURE

Overview of the Structural Fold

From a topogical point of view, the N-acetyltransferase domain folds around a central, mixed β sheet that is largely built up of antiparallel strands; the only

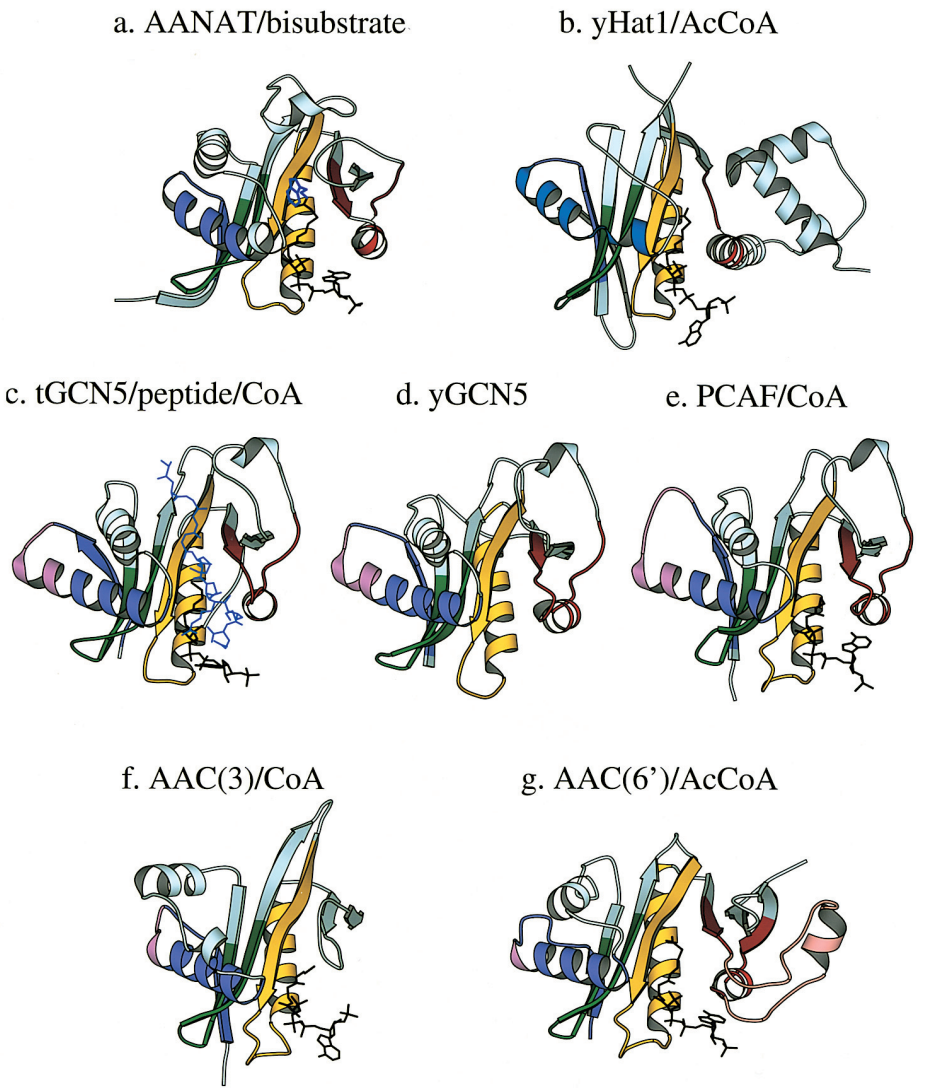

Figure 1 Alignment of the three-dimensional structures of GCN5-related N-acetyltransferases. *AANAT*, serotonin N-acetyltransferase; *PCAF*, p300/CBP-associating factor; *yHat*, histone N-acetyltransferase; *AAC*, aminoglycoside N-acetyltransferase. For AANAT (*a*), the complex with the bisubstrate analog is shown (indole ring in *blue*), while for tGCN5 (*c*), the ternary complex with CoA and an 11-residue peptide (in *blue*) is shown. The *black lines* indicate coenzyme A or acetyl coenzyme A. The four conserved motifs of the GNAT superfamily—C, D, A, and B—are shown in *purple*, *green*, *yellow*, and *red*, respectively. Structural elements that represent insertions into motifs C and B are shown in *lilac* and *pink*, respectively. Helix $\alpha 1$ of yHat1 is shown in *blue* to indicate the positional shift of this helix with respect to the other aligned structures. The figure was prepared using MOLSCRIPT (18).

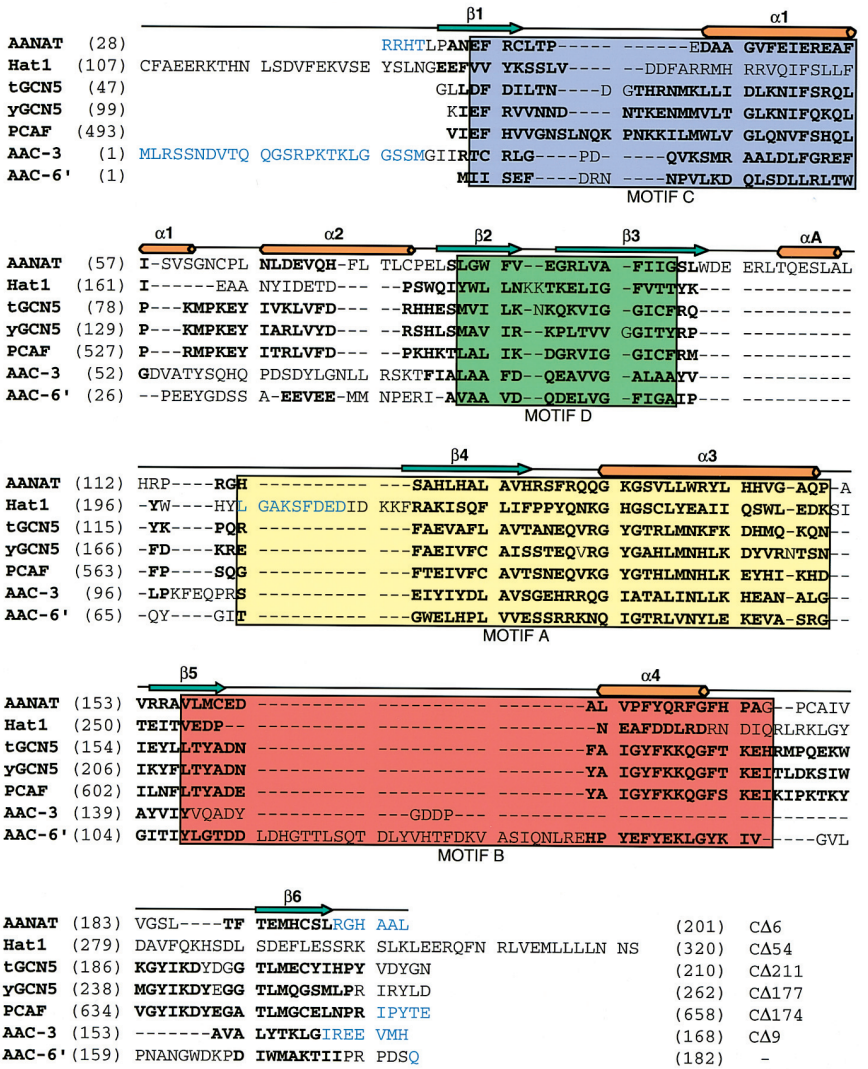

Figure 2 Primary sequences of structurally aligned GCN5-related N-acetyltransferases. Structures were optimally aligned to that of PCAF using "O" (13); reidues whose Cα positions are within 3.8 Å of those of PCAF are shown in *bold* lettering. For AANAT, the bisubstrate-bound form was used, and for tGCN5, that of the ternary complex was used. The residues shown correspond to those present in the constructs used for structure determination, with the exception of yHat1, which has an additional N-terminal domain (residues 1–106). Residues in *blue* are disordered regions that could not be located in the electron density maps. The secondary structural elements of AANAT are shown above the aligned resiudes. The *colored boxes* correspond to the regions designated motifs C, D, A, and B by Neuwald & Landsman (30), and are shown in the same colors used in Figure 1. The final column indicates the number of residues truncated from the C terminus of each construct used for structure determination.

exception is a structurally conserved four-residue-long parallel stretch. The most common number of strands in the β sheet is six, although AAC(6′) has an extra C-terminal strand (45), whereas yHat1 has an additional strand at the N-terminal end of the domain (7). The absence of this N-terminal strand in the other proteins might be a result of the use of N-terminally truncated constructs or disorder in the crystal lattice that renders the electron density in this segment weaker than the noise level and therefore invisible. On the other hand, this additional strand is not present in AAC(6′), a full-length construct for which all the residues are resolved in the crystal structure (45). Given this fact, we start the strand numbering of the domain according to the first strand of AAC(6′), since this is the first common β strand in all the GNAT structures. (The numbering is shown in Figure 3, with the structure of the AANAT/bisubstrate analog complex serving as the representative structure for the superfamily.)

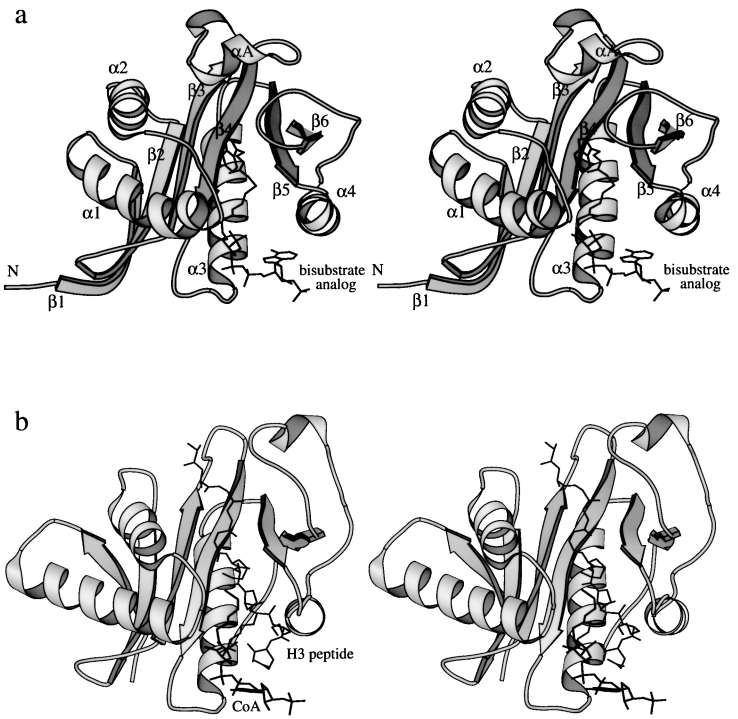

Figure 3 (*a*) Stereo view of serotonin N-acetyltransferase (AANAT) bound to a bisubstrate inhibitor. The common structural elements among the members of the GNAT superfamily are labeled and correspond to the elements shown in Figure 2. AANAT contains a unique insertion between strands $\beta 3$ and $\beta 4$, and the helix contained within this subdomain is labeled αA. (*b*) Stereo view of the tGCN5 ternary complex with CoA and an 11-residue H3 peptide, in the same orientation. The figure was prepared using MOLSCRIPT (18).

The N-Terminal Region and Motif C

The β sheet that forms the core of the N-acetyltransferase domain can be considered as being composed of two parts, one encompassing the first four strands that are nearly identical in all the structures and the second formed by the last two strands [or three, for AAC(6′)]. Generally, in the loop connecting the first and the second strands, two helices run antiparallel to each other and nearly perpendicular to the direction of the first four strands. yHat1 is missing the second helix but not the first (Figure 1*b*), whereas AAC(3) has three helices in this region (Figure 1*g*). The absence of $\alpha 2$ in yHat1 might be due, in part, to the fact that the position of $\alpha 1$ of yHat1 is shifted with respect to the other GNAT structures, due to the proximity of an additional N-terminal domain (residues 1–111, not shown in Figure 1*b*) that is not present in the other structures (7). This ∼4 Å shift in the direction of where $\alpha 2$ would be, as inferred from the rest of the structures, could destabilize its conformation to the extent that the helix is lost. (Note that this shift is also reflected in Figure 1*b* in the lack of alignment of $\alpha 1$ of yHat1 with the analogous helices in the other GNAT structures.) Unfortunately, as none of the other HATs was crystallized with regions other than the catalytic domain present, we will have to wait until such structures are elucidated to see if the absence of $\alpha 2$ in intact HAT proteins is a general feature of the family. A particularly important position on $\alpha 1$ is located on its last turn, represented by Phe-56 in AANAT. In all the structures, this is a large, hydrophobic side chain (Phe twice, Trp once, in two cases Leu and one Ile) that plays an important role in ensuring the characteristically bent conformation of AcCoA (discussed below).

There are an ample number of primarily (but not exclusively) hydrophobic interactions between helices $\alpha 1$ and $\alpha 2$ and the four-stranded sheet below. These very likely contribute to the rigidity of the first four strands of the β sheet and probably explain why this sheet is essentially unchanged from one molecule to another. Similarly, there are several (again, primarily hydrophobic) interactions between the two helices themselves. These sets of interactions, in conjunction with the hydrogen-bonding network between the four antiparallel strands, result in a rigid and compact subdomain in this region of the molecule.

It is obvious from the three-dimensional alignment of the different structures that there is a much better positional match between the first helices ($\alpha 1$) than between the second ones, where yHat1 can be regarded as sitting at the far end of the structural spectrum with $\alpha 2$ completely missing. It is not surprising, then, that sequence motif C includes the first helix. What is perhaps more interesting is that the HAT family, including the GCN5 enzymes, was initially thought to be missing this motif altogether (30). As shown in the sequence alignment in Figure 2, this may in part be due to the variable length of the turn (shown in light purple in Figure 1) between the first β strand and the first helix.

Where the second helix ($\alpha 2$) is present, it is generally shorter than $\alpha 1$ and shows more variation in position between the structures. Notably, it is significantly different in the two AANAT crystal structures (11, 12). In the uncomplexed form,

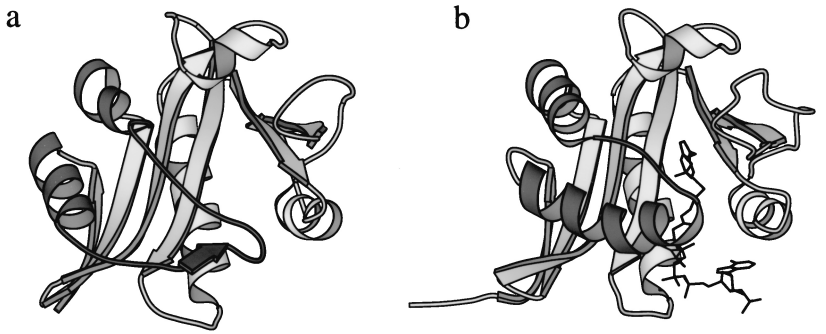

Figure 4 Comparison of the uncomplexed (*a*) and bisubstrate analog-bound (*b*) forms of serotonin N-acetyltransferase. Note the shorter helices $\alpha 1$ and $\alpha 2$, and the additional β strand in the uncomplexed form; the region of the conformational change is highlighted in dark gray. The figure was prepared using MOLSCRIPT (18).

both $\alpha 1$ and $\alpha 2$ are fairly short, whereas in the complex form of AANAT with the bisubstrate analog, both helices extend in length (Figure 4). The shorter helices in the uncomplexed form allow the polypeptide chain to assume a more extended conformation. Indeed, it reaches far enough out and down into the AcCoA binding site to form a short β strand that is not present in the complex form and that occupies part of the binding site. It is the binding of AcCoA that selects a conformational state in which the polypeptide has been reordered, creating the binding site for serotonin (12). This structural result is in good accord with earlier kinetic work that showed that AANAT binds its substrates in an ordered fashion, with AcCoA binding increasing the enzyme's affinity for serotonin substantially, and that a conformational change may be involved (4). The important observation is that, at least in the case of AANAT, this part of the molecule is clearly subject to conformational changes, and this change is a function of whether the AcCoA binding site is occupied or not.

Motifs D and A

The polypeptide chain after $\alpha 2$ completes the first four strands of the sheet, running through sequence motif D that includes most of strands $\beta 2$ and $\beta 3$, and turns into strand $\beta 4$ where motif A, the longest and most highly conserved motif, starts. Note that the regions corresponding to motifs C and D, despite their separation in the primary sequence, are located next to and interact with each other three-dimensionally, forming the rigid subdomain that comprises the first half of the molecule. Strand $\beta 4$ plays a crucial role in AcCoA binding and contains residues that are important for catalytic activity. In the region between strands $\beta 3$ and $\beta 4$, most of the structures contain a short turn. As can be seen in Figure 1, there are two exceptions. In yHat1, there is an $\sim$13-residue extension between $\beta 3$ and $\beta 4$ that cannot be aligned with the other structures; nine of these amino acids are disordered

and therefore invisible in the crystal structure (7). In AANAT, there is also an ∼13-residue-long extension, but in this case, this region forms a small subdomain with a very well defined three-dimensional structure (11), folded around Trp-99, a conserved residue in AANATs. Mutation of Trp-99 to Ala results in a protein that is completely insoluble (AB Hickman, unpublished data), presumably because of gross folding problems. In AANAT, this subdomain contains a short two-turn helix, not seen in the other structures (labeled αA in Figure 3). This helix defines one side of the hydrophobic substrate binding pocket for serotonin and therefore likely plays a crucial role in determining substrate specificity. With the help of this short helix, AANAT manages to exclude the hydrophobic indole moiety of serotonin from solvent and surround it with an essentially entirely hydrophobic environment. In the other structures, the absence of such a subdomain leaves this part of the molecule more accessible to solvent, consistent with the larger-sized or more hydrophilic substrates (relative to serotonin) to be acetylated by the other GNAT enzymes.

Strand β4 and the structural element directly following it, α3, form the essence of the AcCoA binding site. Strand β4 starts with a four-residue stretch that runs parallel to β5, forming the only parallel β structure in the GNAT catalytic domain. Motif A begins just before the parallel stretch of β4 and extends to the carboxy end of α3. The parting of the two parallel strands results in a wedge-like opening in the center of the protein where AcCoA binds. A characteristic feature of the GNAT superfamily is a β bulge, located in the middle of β4, immediately following the short parallel strand segment. In a regular β strand, the peptide planes alternate in orientation from one residue to the next. This alternation projects each sequential main-chain carbonyl oxygen to opposite sides of the strand. This places the carbonyl groups in the appropriate location and register to form hydrogen bonds with similarly alternating main-chain amide groups of adjacent strands. A β bulge is an exception from this pattern, whereby two residues occupy the space in a β sheet normally occupied by one.

The "classic" β bulge (33) observed in β4 of GNAT enzymes is created by two residues that directly follow each other in the sequence turning their carbonyls to the same side of the β strand, in this case toward the cofactor binding site (shown in Figure 5 for AANAT). Correspondingly, the amides of these two adjacent residues point to the other side of the β strand, placed such that they can form a bifurcated hydrogen bond with a main-chain carbonyl from β3. Thus, the two strands remain held together albeit locally distorted. It is likely that the β bulge is a significant feature breaking the parallel segment between β4 and β5, and hence in the formation of the AcCoA binding site. Among the GNAT structures determined to date, there is one exception to this observation: AAC(6′) has a Pro residue inserted at the analogous position in β4, resulting in a markedly different hydrogen-bonding pattern (45). Since the amide nitrogen of Pro is not available for hydrogen bonding, the introduction of a Pro into β4 results in an irregular hydrogen-bonding pattern between β3 and β4, with the two adjacent carbonyls from β4 now pointing toward β3. Since the α-carbon trace of AAC(6′) is essentially identical to the others

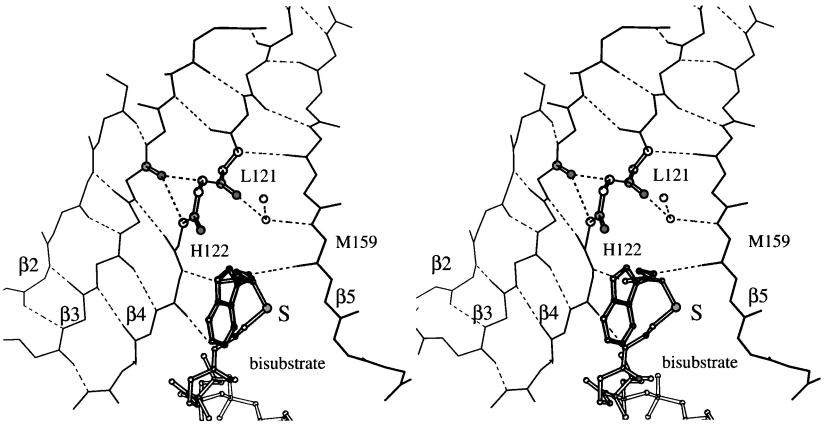

Figure 5 Stereo view of the β bulge region of serotonin N-acetyltransferase (AANAT). *Dashed lines* represent hydrogen bonds. The two *circles* represent the most deeply located water molecules in the proton wire. The figure was prepared using MOLSCRIPT (18).

in the region of the β bulge, insertion of a proline residue into $\beta 4$ seems to be a different means of achieving essentially the same structural end.

Immediately downstream of the β bulge are three residues that form main-chain hydrogen bonds with the AcCoA substrate, starting with the amide of the residue immediately after the bulged residue. In the cases where the structures include AcCoA, this residue is directly hydrogen bonded to the carbonyl of the acetyl group. Indeed, AcCoA interacts with $\beta 4$ much as if it were the next strand in a β sheet. The acetyl and pantetheine moieties of AcCoA project carbonyl and amide groups to both sides and are separated by the right distance to hydrogen bond with an adjacent β strand. (The structure of AcCoA is shown in Figure 6*a*.)

In addition to the β bulge, the second characteristic local structural feature of the GNAT superfamily in this region is the pyrophosphate binding pocket at the amino end of $\alpha 3$; this helix is part of a $\beta\alpha\beta$ motif resembling the Rossman fold found in many nucleotide-binding proteins (32). Helix $\alpha 3$ is the longest helix in the structure, and the helix dipole must contribute significantly to phosphate binding. Indeed, it is quite impressive that among the different molecules, both substrate conformation and mode of pyrophosphate binding are essentially identical down to the most intimate details. (The overlain cofactor structures are shown in Figure 7, viewed from the same orientation as in Figure 1.) The only significant differences are seen in the NMR structure of the tGCN5/CoA complex (20), but since this is not consistent with the X-ray results for tGCN5 and the other members of the superfamily, we do not analyze this result further in this review.

One unusual feature of the coenzyme A binding mode is the lack of direct hydrogen bonds by Lys or Arg side chains to the phosphates, even though such salt bridges are frequently seen in other types of CoA-binding proteins (8). The

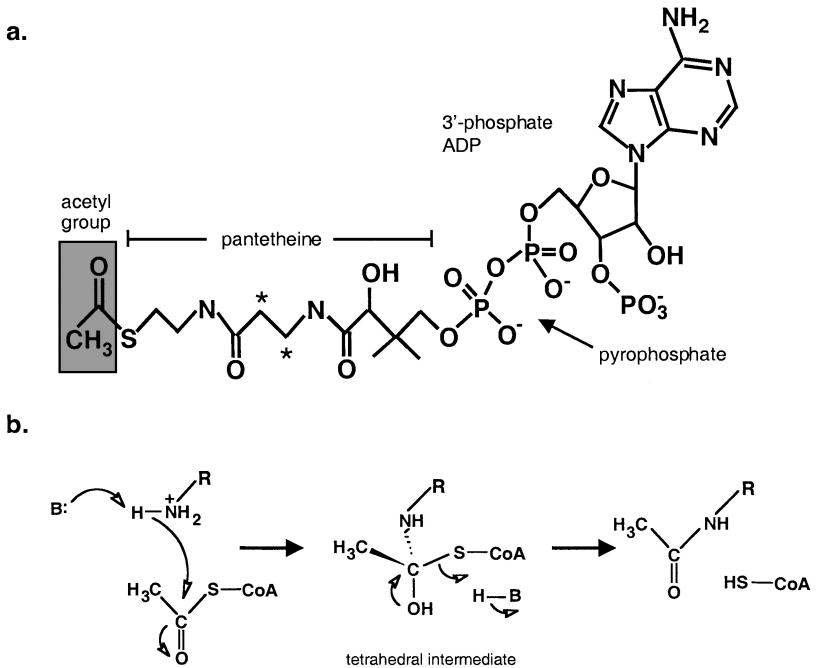

Figure 6 (*a*) The chemical structure of acetyl coenzyme A. The characteristic sharp bend observed in the bound forms of AcCoA and CoA in GNAT superfamily members is around the bond highlighted by the starred carbon atoms. (*b*) The general reaction catalyzed by GCN5-related N-acetyltransferases, showing the presumed tetrahedral intermediate that results from nucleophilic attack of a primary amine on the acyl carbon of the acetyl group.

pyrophosphate group is deeply buried, and with one exception, all the polar interactions between the phosphates and the protein involve main-chain amides. The amides of all the six residues that form the pyrophosphate binding pocket are hydrogen bonded to the phosphate oxygen atoms. Four of these bonds are direct, two are water mediated, and these water molecules are conserved in all the X-ray structures. This situation is somewhat reminiscent of that seen for succinyl-CoA synthetase in which the pyrophosphate is also hydrogen bonded to main-chain amides at the N-terminus of an α helix (44). In this case, however, the binding site for CoA is located between two adjacent protein subunits.

Among the GNAT structures, the only side-chain interaction with the pyrophosphate group is between the α-phosphate and the side-chain hydroxyl of a residue on the first turn of α3. With the exception of yGCN5, in the structures determined to date, this residue is either a Ser or a Thr and forms the same interaction with the phosphate. In yGCN5, the analogous residue is an Ala, but as this structure was solved without substrate (39), it is not yet clear how this affects phosphate binding. It should be noted that this residue is not conserved across the superfamily.

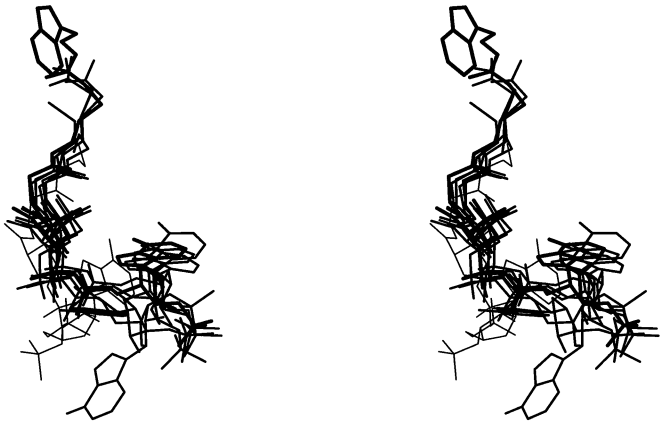

Figure 7 Stereo view of the superposition of all the observed conformations of bound coenzyme A–related substrates. The bisubstrate analog bound to AANAT is shown in the thickest line, and its indole moiety can be seen at the *upper left*; the thinnest line represents the conformation observed in the NMR structure (20). The view is the same as in Figure 1.

Motif B and the Completion of the AcCoA Binding Site

Helix $\alpha 3$, consisting of five turns, leads into $\beta 5$, and the turn between these two structural elements is where motif A ends. Strand $\beta 5$ starts by pairing up with $\beta 4$ to form the short parallel segment of the β sheet, and subsequently parts from it to define the other side of the wedge-like AcCoA binding site. It is not surprising, therefore, that strand $\beta 5$ is also where another highly conserved region, motif B, starts. With the exception of the main-chain carbonyl oxygen of one residue (Met-159 in AANAT; see Figure 5), there are no hydrogen-bond contacts at all between $\beta 5$ and the substrates. However, this particular carbonyl is likely to play an important role in enzymatic action, as it is hydrogen bonded to the substrate amine in the AANAT/bisubstrate analog complex (12). The carbonyl of the analogous residue in tGCN5, Tyr-160, is in the appropriate position to do the same in the ternary complex of tGCN5 (34). Despite the lack of direct hydrogen bonds, $\beta 5$ and the helix that follows it, $\alpha 4$, do make several interactions with the AcCoA substrate. In particular, there are several large hydrophobic residues located on helix $\alpha 4$ that contact either the substrate or side chains from $\beta 4$, defining and stabilizing this bottom part of the donor substrate binding site. Another residue in $\alpha 4$ that appears to play a role in AcCoA binding is a highly conserved basic residue, corresponding to Arg-170 in AANAT (note this position is shifted downstream by one in yHat1). In some structures, this residue forms a salt bridge to the 3′ phosphate of AcCoA, whereas in other structures, it is in a position to potentially do so.

Another important position on $\alpha 4$ is that corresponding to Tyr-168 in AANAT. In all the aligned structures (allowing for a one-residue upstream shift in yHat1), this is a large hydrophobic residue, either Tyr or Phe. It contacts AcCoA from the

direction of the protein interior and is likely to play an important role in the proper positioning of the acetyl group for the transfer reaction to occur (discussed below). In some cases, it may also have a role in the resolution of the reaction intermediate. Its importance was demonstrated in AANAT, where the Tyr-to-Phe mutation was shown to have a substantial negative effect on catalytic efficiency (12). Indeed, of all the point mutations that were made in the vicinity of the AANAT active site, the effect of this mutation was the most detrimental.

The one structure that deviates from the description above regarding $\alpha 4$ is AAC(3). Starting at the end of $\beta 5$, its secondary structure is significantly different from that of the other superfamily members (43). In particular, AAC(3) is completely missing $\alpha 4$. Instead, $\beta 5$ is immediately followed by a short antiparallel β strand (Figure 1*f*).

With the downstream end of $\alpha 4$, motif B ends as well, and from this point to the C-termini, the structures show more divergence than in the parts examined so far. This is probably due in part to the fact that the C-terminal region contains a loop of varying length and position. For example, in AAC(6′), there is a long insertion between $\beta 5$ and $\alpha 4$, whereas in the GCN5 enzymes, the loop follows $\alpha 4$ and contains 20 residues. In both the ternary complex form of tGCN5 (34) and in the AANAT/bisubstrate analog complex (12), this loop contacts the acceptor substrate. Consistent with the notion that this loop is a structural component of the binding site for acceptor substrates, in yGCN5, this region is where some of the most debilitating mutations were located (40). Thus, the structural diversity seen in this region may be a reflection of the variety of acceptor substrates that the different GNAT members have evolved to acetylate. A notable variation is yHat1 (7): after $\alpha 4$, the polypeptide does not return to the substrate binding surface but instead forms a helical bundle not seen in the other structures (Figure 1*b*). Given the lack of specific mutational data, the significance of this difference is not known. Finally, in the structures other than yHat1, the GNAT domain ends with a last β strand, $\beta 6$, that runs antiparallel to $\beta 5$.

SUBSTRATE BINDING

The Conformation of Bound AcCoA/CoA

As described earlier, AcCoA and CoA bind in the opening formed between the diverging strands $\beta 4$ and $\beta 5$, contacting protein atoms mainly from $\beta 4$ and $\alpha 4$, predominantly through main-chain interactions. Coenzyme A is bound in a characteristically sharply bent conformation (Figures 1, 7). This bend introduces an acute 65–70° angle between the two amide planes of the pantetheine group, and it seems to be the result of two specific hydrogen bonds to main-chain groups located on $\beta 4$ and of the steric influence of a large hydrophobic residue projecting from $\alpha 1$ (Phe-56 in AANAT). In the GNAT structures determined with co-crystallized AcCoA (or the bisubstrate analog for AANAT), in most cases the carbonyl of

the acetyl group is hydrogen bonded to the main-chain amine of a residue just downstream of the β bulge on β4 (Leu-124 in AANAT). The one exception is AAC(6′), but in this case the crystallographic temperature factors at this end of AcCoA are significantly larger than those of the surrounding protein atoms (50–60 versus <5), so it is likely that the positional accuracy of the acetyl group is low (45). Apart from this variation, the conformation of AcCoA, up to and including the pyrophosphate group, is essentially identical in all the structures. In the CoA complexes, there is some variation in the position of the sulphur atom, but this is most likely due to the lack of the specific hydrogen bond stabilizing the position of the (missing) acetyl group.

In contrast, there is significant variation in the position and conformation of the ribose possessing the 3′ phosphate and the base. This correlates with the relatively few number of interactions in which this part of CoA engages with the protein. In some cases, the quality of the electron density is probably not sufficient for accurate determinations; nevertheless, in all the X-ray structures, the glycosidic bond is in the anti conformation, and the sugar pucker is the corresponding C2′ endo configuration.

Acceptor Substrate Binding

To date, only two structure determinations have been carried out that include the acceptor substrate. The first is the bisubstrate analog complex form of AANAT (12). In this case, a chemically stable compound was synthesized by covalently linking an N-acetylated substrate, tryptamine, to CoA (14). This bisubstrate analog, a close approximation of the presumed reaction intermediate (Figure 6*b*), is a potent inhibitor of AANAT ($IC_{50} \approx 150$ nM). More recently, tGCN5 was crystallized in a ternary complex form with CoA and a peptide containing the reactive lysine of histone H3 (34). These two cases reveal the essence of the situation: although the details of acceptor binding vary slightly, in both AANAT and tGCN5, the N-acetyltransferase domain utilizes the same regions to make key interactions with the amine-containing substrate (Figure 8, see color insert).

Critical for interaction are two loops mentioned earlier: one located between α1 and α2, and another just preceding the last β strand of the domain. The role of these two loops in acceptor substrate binding was initially suggested based on the uncomplexed structure of AANAT (11), and their importance was underscored by the alanine-scanning mutational work in yGCN5 in which a stretch of six residues just after α1 and a three-residue stretch just before β6 were demonstrated to be essential for its in vivo transcriptional function (40).

As discussed earlier, the upstream end of α2, including the loop between α1 and α2, plays an important role in the binding of the acceptor substrate. For example, in AANAT, Pro-64, an important residue that defines one side of the serotonin binding pocket (Figure 8*a*), is located in this region (12). Similarly, in the ternary complex of tGCN5, this same region spanning the connecting loop and most of α2 (Pro-78 to Val-88) makes numerous intimate contacts with the peptide

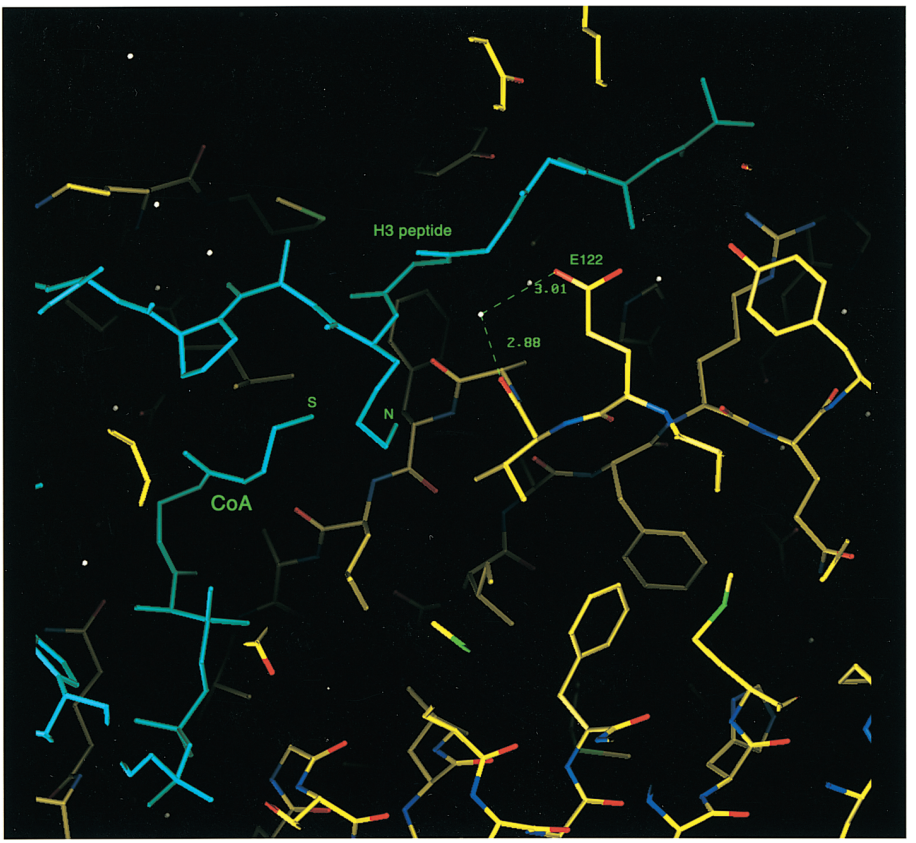

Figure 8 (*a*) The proton wire of serotonin N-acetyltransferase (AANAT). The crystallographically observed water molecules are shown as *white spheres*, the bisubstrate inhibitor in *blue*.

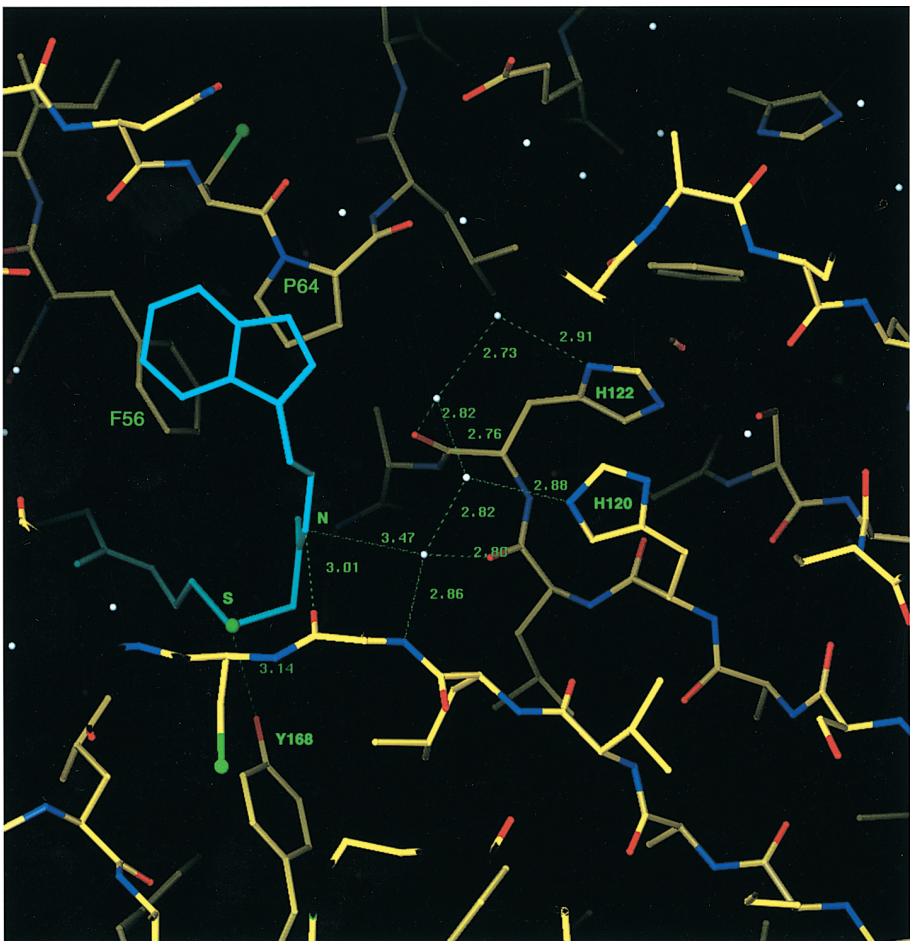

Figure 8 (*b*) Close-up of the active site region of the tGCN5 ternary complex. Water molecules are represented as *white spheres*, and the substrates are shown in *blue*.

substrate (34). In both proteins, the two loop regions are located on the same surface of the molecule and define a cleft located directly above the acetyl group of AcCoA. Both serotonin, the acceptor substrate of AANAT, and the histone H3 peptide substrate of yGCN5 lie in this cleft. In AANAT, the indole moiety of serotonin is almost entirely enclosed in the cleft whose "third wall" is contributed by the additional helix (αA), unique to AANAT, located between β3 and β4. In tGCN5, there is only a short turn between β3 and β4, and in this case the cleft is long and open on both ends, allowing the peptide to fit in such that all 11 residues are in contact with the protein (compare Figure 1*a* and *c*). Although the tGCN5-peptide interactions are about evenly distributed between van der Waals interactions and hydrogen bonds (interestingly, with the notable exception of the primary amine of the reactive Lys, these interactions mostly involve peptide main-chain groups), in AANAT, serotonin is held in place entirely by van der Waals–type contacts provided by residues in the three converging loops. Indeed, the indole ring is sandwiched between Pro-64 on one side and Phe-188 on the other. Since the NE1 nitrogen of the indole is hydrogen bonded to a water molecule, AANAT apparently does not exploit this group for substrate recognition. It should be noted, however, that in the experimental structure determination, tryptamine rather than serotonin is tethered to CoA. Tryptamine differs from serotonin in lacking the 5′-OH substitution on the indole, which—if present—would be in an ideal position to hydrogen bond to the side chain of Ser-60.

From the mapped acetylation sites of yGCN5, only two positions around the reactive lysine seem to be conserved, indicating a G-K-X-P consensus sequence for peptide recognition (34). Indeed, in the ternary complex, other than the reactive lysine, only this nearby Pro residue appears to be in extensive hydrophobic contact with the protein. Perhaps somewhat surprisingly, this proline residue also contacts the pantetheine moiety of AcCoA. It seems, therefore, that AcCoA contributes to peptide recognition and binding, or—stated another way—that the formation of a GCN5/AcCoA complex must be required prior to peptide binding. This is an interesting variation of what is observed for AANAT, where the binding of AcCoA introduces a major conformational change, resulting in the completion of the serotonin binding site, and only then in the binding of serotonin. In contrast, there are only modest movements of tGCN5 upon AcCoA binding (34).

CHEMICAL CATALYSIS

The Need for a General Base?

In considering catalytic mechanism, there are two distinct ways in which AcCoA-dependent acetyltransferases could catalyze the transfer of the acetyl group. In one, a ping-pong mechanism, the acetyl group is transiently transferred to a suitably located Cys residue of the enzyme, forming a covalently bound acetylated enzyme intermediate. The enzyme subsequently catalyzes the transfer of the acetyl group from the Cys residue to the acceptor substrate (see, for example, 6, 41). In the

second possible mechanism, the acetyl group is transferred directly from AcCoA to the acceptor via direct nucleophilic attack by the primary amine on the acyl-carbon (Figure 6*b*). This second possibility obviously requires the formation of a ternary complex between the enzyme, AcCoA, and the acceptor substrate.

Several lines of biochemical data indicate that the GNAT superfamily members use the direct acetyl transfer mechanism. These include kinetic experiments (4, 37), the failure to identify covalently bound intermediates (7), and the inability to inactivate yGCN5 using reagents that block thiol groups (37). In addition, most of the structural work failed to identify suitably located Cys residues near the acetyl group of AcCoA, and the case for direct nucleophilic attack was further strengthened by the recent direct crystallographic evidence for the existence of ternary complexes (12, 34). Taken together, these data strongly favor direct acetyl transfer as the catalytic mechanism for the GNAT superfamily of enzymes.

For direct nucleophilic attack to occur, the primary amine must be in an uncharged form, and given the high pKa values [$\sim$10 for both arylalkylamines (24) and lysine (36)], it is very likely that GNAT enzymes must provide some mechanism of deprotonation, presumably involving an amino acid near the active site that can act as a general base. Further underscoring the need for a mechanism of proton removal, it was demonstrated for AANAT that the protonated form of the acceptor substrate preferentially binds to the enzyme (15). Extensive mutational work aimed at identifying the general base has been carried out in two systems, yGCN5 and AANAT. For yGCN5, the results strongly implicate Glu-173 as the general base (37), a residue that is conserved in the GCN5 family, including PCAF. The conservative replacement of Glu with Gln resulted in a 320-fold reduction of k_{cat}. Furthermore, pH profile experiments in the alkaline regime showed a decreased effect of this mutation, an expected result considering the increase in the relative concentration of the nonenzymatically deprotonated amine.

In contrast to the compelling evidence implicating Glu-173 as the general base in yGCN5, the results for AANAT are not as straightforward. Substitution of His-120 and His-122 in ovine AANAT, two residues previously suggested as possible catalytic residues (16), does not result in a significant decrease in k_{cat} (12). Looking at the three-dimensionally aligned structures, the situation becomes even more intriguing. His-120 of AANAT and Glu-122 of tGCN5 (the equivalent residue to Glu-173 of yGCN5) are in an identical location, just one residue upstream of the β bulge on β4. Although this position is clearly in the vicinity of the active site, it is still too far from the observed positions of the amine nitrogen for direct proton transfer to occur. In the bisubstrate analog complex form of AANAT, NE2 of His-120 is 7.5 Å away from the amine nitrogen, and in the ternary complex of tGNC5, the distance between the primary amine and OE1 of Glu-122 is $>$8 Å. For the other structurally characterized members of the GNAT superfamily, the situation regarding a general base is murky. While there is a Glu (Glu-72) in AAC(6′) at the analogous position, mutational data are not available. The corresponding residue in yHat1 is Lys-216, whereas in AAC(6′), it is Tyr-109. These residues

are not likely to act as general bases. In yHat1, there are two Asp residues on β5, located close to the acetyl group, one of which (Asp-255) appears well positioned for the task. However, it is also close to Lys-216, with which an interaction might interfere with a potential catalytic role. In AAC(3), there is a suitably located Asp (Asp-110) just one residue downstream of the β bulge, but again there are no mutational data available to support a catalytic role. Thus it appears that there is no conserved general base at the equivalent position in the active sites of these enzymes that would serve as a hallmark of catalysis.

The Role of Proton Wires in Amine Deprotonation

From a purely mechanistic point of view, the tGCN5 and AANAT complexes offer complementary—or dynamic—views of the catalytic process. The ternary complex of tGCN5 represents a snapshot of the reaction just prior to nucleophilic attack, since the system is frozen at this state due to the lack of the acetyl group on the donor substrate (34). In the AANAT complex, the snapshot is of the tetrahedral reaction intermediate, a consequence of the covalent linkage between the two substrates (12). It is reassuring that both structures offer a consistent view of how proton removal from the primary amine is likely to be effected by the enzyme.

In AANAT, there is an array of well-located water molecules, with low crystallographic temperature factors, that form a continuous chain or "proton wire" connecting the amine nitrogen both to internal exchangeable groups (including His-120 and His-122) and to the surrounding solvent (Figure 8). Such proton wires have been previously observed in biological systems (25). The first water molecule in the chain, located deepest in the protein interior, is coordinated by three hydrogen bonds: one to the carbonyl oxygen of Leu-121, the upstream residue of the β bulge on β4; the second to the amide nitrogen of Met-159 on the adjacent strand, β5; and the third to the next water molecule in the proton wire. In its observed location, this water molecule is well positioned to hydrogen bond to the primary amine of the substrate. In the experimentally determined structure, this hydrogen bond cannot exist, however, since the analog contains a secondary amine, already hydrogen bonded to the carbonyl oxygen of Met-159, a reflection of the fact that the analog is a representation of the chemistry after the initial nucleophilic attack by the primary amine. Nevertheless, this water molecule is obviously ideally suited to accept the extra proton and to channel it away to either of the two local imidazoles or to the open solvent through the chain of crystallographically visible water molecules. It is also worth noting that the two most deeply located water molecules of the proton wire are held in place by hydrogen bonds to the two carbonyl oxygens that are projected into the active site area by the β bulge (Figure 5), granting additional functional significance to this well-conserved structural feature.

In the ternary complex of tGCN5, a buried water molecule connects the location of the primary amine to the side chain of Glu-122 (34). This water molecule also

has a low crystallographic temperature factor; thus its location is well determined (see Figure 8*b*). Its position is structurally identical to that of the first water molecule of the proton wire in the AANAT complex, and it is also held in place by a hydrogen bond to the carbonyl oxygen of the first β bulged residue. Allowing for the larger conformational flexibility permitted by the longer aliphatic side chain of lysine, this primary amine could donate its proton to this water molecule, which could subsequently transfer it to Glu-122. Thus, the two ternary complexes provide a structural alternative to direct proton abstraction by a catalytic general base: The water molecules observed in both structures most likely serve as a conduit for proton transfer away from the primary amine.

Substrate Positioning

In the absence of a conserved general base in the active sites of GNAT enzymes, what does seems to be conserved, on the other hand, is the relative positioning of the reactive groups (i.e. the primary amine and the acetyl group) by the enzyme. In the three-dimensionally aligned structures, the positions of the acyl carbons of AcCoA are nearly identical, maintained by virtue of an important hydrogen bond to a main-chain amide. In the case of the AANAT/bisubstrate complex, the primary amine of the "acceptor substrate" is perforce located close to the acyl group through covalent bonds. In the case of the tGCN5 ternary complex, the primary amine of the reactive Lys is not in precisely the same location as that of the amine of the bisubstrate. However, this may well be a consequence of experimental design: to trap a ternary complex prior to catalysis, CoA was used in place of AcCoA, and the hydrogen bond anchoring the acyl group in place cannot exist. Nevertheless, it is clear that only slight movements of the lysine side chain are required for the primary amine of the tGCN5 ternary complex to occupy the same position as the bisubstrate analog amine group in AANAT. Thus, despite the significant evolutionary distance between the different members of the GNAT superfamily, the positioning of the primary amine group is likely to be a conserved feature of catalysis. In this light, we note that the detrimental Tyr-168–to–Phe replacement in AANAT (discussed above) could affect this delicate positioning, and that this in turn could be responsible for the observed significant decrease in reaction rate (12).

Although there is no conserved catalytic base in the active sites of GNAT enzymes, it is not clear to what extent critical residues in the general vicinity of the acetyl group affect catalysis. Acetyl transfer from AcCoA to a primary amine in aqueous solution is energetically favorable as evidenced by measurable nonenzymatic background rates (37). The question remains to what extent the enzymatic rate acceleration can be attributed to appropriate substrate positioning.

Upon bringing the two substrates into the appropriate position for catalysis, the AANAT/bisubstrate complex offers insight into the next step in the reaction, the resolution of the tetrahedral intermediate. It is clear from chemical considerations

that the resolution in the direction of the desired products could be facilitated by providing a proton to the sulfur of CoA since a negatively charged thiolate is not an ideal leaving group. In AANAT, Tyr-168 is located 3.1 Å from the sulfur in a position that would be ideal for donating a proton. Interestingly, there is a Tyr residue in an identical position in AAC(6′), although there are no mutational data available for this enzyme.

The final step in the catalytic cycle is product release. The environment around the primary amine is polar, and upon formation of acetylserotonin, the dramatic increase in hydrophobicity may serve to eject the acetylated product from the protein. The emptied indole binding pocket would lose its stability, leading to a conformational state that could effect the release of CoA. It is not clear, however, that hydrophobicity is necessarily the driving force across the superfamily of GNAT enzymes; for tGCN5, steric constraints have been invoked to explain product release (34).

SUMMARY

The three-dimensional structures determined for various members of three families of N-acetyltransferases of the GNAT superfamily demonstrate a remarkable consistency in protein topology. It appears that a central β sheet, which incorporates elements from all four conserved sequence motifs identifed prior to any structural information, serves as a scaffold to appropriately place residues to form two substrate binding sites. The binding site for the donor substrate, AcCoA, is remarkably conserved, and the same three-dimensionally aligned residues are involved in holding AcCoA in place. In particular, the crucial acyl group is hydrogen bonded in precisely the same manner among all the characterized members of the family, defining a conserved enzyme active site. The binding site for the acceptor substrate has been characterized in only two cases, but it is clear that the same regions of the protein are involved: two (or more) analogous loops of varying length and differing conformation are arrayed across the protein surface to create a substrate binding cleft that positions the reactive primary amine close to the acetyl group of AcCoA. The structural work to date provides intriguing glances into mechanisms of catalysis, unequivocably demonstrating that there is no universally conserved, general base to directly abstract a proton from the primary amine substrates, yet offering in its place a potential pathway for proton removal involving crystallographically located water molecules and a proton wire.

It is hoped that future work in this field will provide more information on modes of acceptor substrate binding, particularly for those members whose substrates differ from those characterized to date. It will also be instructive, from the point of view of regulation of activity, to obtain structures of the HAT domains, both in the context of their intact proteins and as part of the multiprotein complexes involved in transcriptional regulation.

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

1. Bhatnagar RS, Futterer K, Farazi TA, Korolev S, Murray CL, et al. 1998. Structure of N-myristoyltransferase with bound myristoylCoA and peptide substrate analogs. *Nat. Struct. Biol.* 5:1091–97
2. Clements A, Rojas JR, Trievel RC, Wang L, Berger SL, Marmorstein R. 1999. Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A. *EMBO J.* 18:3521–32
3. Coon SL, Roseboom PH, Baler R, Weller JL, Namboodiri MAA, et al. 1995. Pineal serotonin N-acetyltransferase: expression cloning and molecular analysis. *Science* 270:1681–83
4. De Angelis J, Gastel J, Klein DC, Cole PA. 1998. Kinetic analysis of the catalytic mechanism of serotonin N-acetyltransferase (EC 2.3.1.87). *J. Biol. Chem.* 273:3045–50
5. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399:491–96
6. Dupret J, Grant DM. 1992. Site-directed mutagenesis of recombinant human arylamine N-acetyltransferase expressed in *Escherichia coli. J. Biol. Chem.* 267:7381–85
7. Dutnall RN, Tafrov ST, Sternglanz R, Ramakrishnan V. 1998. Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily. *Cell* 94:427–38
8. Engel C, Wierenga R. 1996. The diverse world of coenzyme A binding proteins. *Curr. Opin. Struct. Biol.* 6:790–97
9. Gastel JA, Roseboom PH, Rinaldi PA, Weller JL, Klein DC. 1998. Melatonin production: proteolysis in serotonin N-acetyltransferase regulation. *Science* 279:1358–60
10. Grunstein M. 1997. Histone acetylation in chromatin structure and transcription. *Nature* 389:349–52
11. Hickman AB, Klein DC, Dyda F. 1999. Melatonin biosyntheses: the structure of serotonin N-acetyltransferase at 2.5 Å resolution suggests a catalytic mechanism. *Mol. Cell.* 3:23–32
12. Hickman AB, Namboodiri MAA, Klein DC, Dyda F. 1999. The structural basis of ordered substrate binding by serotonin N-acetyltransferase: enzyme complex at 1.8 Å resolution with a bisubstrate analog. *Cell* 97:361–69
13. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47:110–19
14. Khalil EM, Cole PA. 1998. A potent inhibitor of the melatonin rhythm enzyme. *J. Am. Chem. Soc.* 120:6195–96
15. Khalil EM, De Angelis J, Cole PA. 1998. Indoleamine analogs as probes of the substrate selectivity and catalytic mechanism of serotonin N-acetyltransferase. *J. Biol. Chem.* 273:30321–27
16. Klein DC, Coon SL, Roseboom PH, Weller JL, Bernard M, et al. 1997. The melatonin rhythm-generating enzyme: molecular regulation of serotonin N-acetyltransferase in the pineal gland. *Recent Prog. Horm. Res.* 52:307–58
17. Klein DC, Weller JL. 1972. Rapid light induced decrease in pineal serotonin N-acetyltransferase activity. *Science* 177:532–33
18. Kraulis PJ. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–50
19. Leslie AGW, Moody PCE, Shaw VW. 1988. Structure of chloramphenicol

acetyltransferase at 1.75 Å resolution. *Proc. Natl. Acad. Sci. USA* 85:4133–37
20. Lin Y, Fletcher CM, Zhou J, Allis CD, Wagner G. 1999. Solution structure of the catalytic domain of GCN5 histone acetyltransferase bound to coenzyme A. *Nature* 400:86–89
21. Liu L, Scolnick DM, Trievel RC, Zhang HB, Marmorstein R, et al. 1999. p53 sites acetylated in vitro by PCAF and p300 are acetylated in vivo in response to DNA damage. *Mol. Cell. Biol.* 19:1202–9
22. Lu L, Berkey KA, Casero RA Jr. 1996. RGFGIGS is an amino acid sequence required for acetyl coenzyme A binding and activity of human spermidine/spermine N^1 acetyltransferase. *J. Biol. Chem.* 271: 18920–24
23. Mattevi A, Obmolova G, Kalk KH, Westphal AH, de Kok A, Hol WGJ. 1993. Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Å resolution. *J. Mol. Biol.* 230:1183–99
24. *Merck Index.* 1976. p. 1095. Rahway, NJ: Merck & Co. 9th ed.
25. Meyer E. 1992. Internal water molecules and H-bonding in biological macromolecules: a review of structural features with functional implications. *Protein Sci.* 1:1543–62
26. Mio T, Yamada-Okabe T, Arisawa M, Yamada-Okabe H. 1999. *Saccharomyces cerevisiae* GNA1, an essential gene encoding a novel acetyltransferase involved in UDP-*N*-acetylglucosamine synthesis. *J. Biol. Chem.* 274:424–29
27. Mizzen CA, Allis CD. 1998. Linking histone acetylation to transcriptional regulation. *Cell. Mol. Life Sci.* 54:6–20
28. Modis Y, Wierenga R. 1998. Two crystal structures of N-acetyltransferases reveal a new fold for CoA-dependent enzymes. *Structure* 6:1345–50
29. Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247:536–40
30. Neuwald AF, Landsman D. 1997. GCN5-related histone *N*-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem. Sci.* 22:154–55
31. Petsko GA, Kenyon GL, Gerlt JA, Ringe D, Kozarich JW. 1993. On the origin of enzymatic species. *Trends Biol. Sci.* 18:372–76
32. Rao ST, Rossman MG. 1973. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* 76:241–56
33. Richardson JS. 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34:167–330
34. Rojas JR, Trievel RC, Zhou J, Mo Y, Li X, et al. 1999. Crystal structure of *Tetrahymena* GCN5 with bound coenzyme-A and histone H3 peptide. *Nature* 401:93–98
35. Sternglanz R, Schindelin H. 1999. Structure and mechanism of action of the histone acetyltransferase Gcn5 and similarity to other *N*-acetyltransferases. *Proc. Natl. Acad. Sci. USA* 96:8807–8
36. Stryer L. 1988. *Biochemistry*, p. 21. New York: Freeman. 3rd ed.
37. Tanner KG, Trievel RC, Kuo M, Howard RM, Berger SL, et al. 1999. Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. *J. Biol. Chem.* 274:18157–60
38. Tercero JC, Riles LE, Wickner RB. 1992. Localized mutagenesis and evidence for post-transcriptional regulation of MAK3. *J. Biol. Chem.* 267:20270–76
39. Trievel RC, Rojas JR, Sterner DE, Venkataramani RN, Wang L, et al. 1999. Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator. *Proc. Natl. Acad. Sci. USA* 96:8931–36
40. Wang L, Liu L, Berger SL. 1998. Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional

function in vivo. *Genes Dev.* 12:640–53
41. Watanabe M, Sofuni T, Nohmi T. 1992. Involvement of Cys69 residue in the catalytic mechanism of *N*-hydroxyarylamine *O*-acetyltransferase of *Salmonella typhimurium*. *J. Biol. Chem.* 267:8429–36
42. Weston SA, Camble R, Colls J, Rosenbrock G, Taylor I, et al. 1998. Crystal structure of the anti-fungal target N-myristoyl transferase. *Nat. Struct. Biol.* 5:213–21
43. Wolf E, Vassilev A, Makino Y, Sali A, Nakatani Y, Burley SK. 1998. Crystal structure of a GCN5-related N-acetyltransferase: *Serratia marcescens* aminoglycoside 3-N-acetyltransferase. *Cell* 94:439–49
44. Wolodko WT, Fraser ME, James MNG, Bridger WA. 1994. The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5 Å resolution. *J. Biol. Chem.* 269:10883–90
45. Wybenga-Groot LE, Draker K, 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