

Crystal Structure of the Histone Acetyltransferase Hpa2: A Tetrameric Member of the Gcn5-related N-acetyltransferase Superfamily

Melinda L. Angus-Hill[†], Robert N. Dutnall^{†*}, Stefan T. Tafrov, Rolf Sternglanz and V. Ramakrishnan*

Department of Biochemistry
University of Utah School of
Medicine, Salt Lake City
UT 84132, USA

We report the crystal structure of the yeast protein Hpa2 in complex with acetyl coenzyme A (AcCoA) at 2.4 Å resolution and without cofactor at 2.9 Å resolution. Hpa2 is a member of the Gcn5-related N-acetyltransferase (GNAT) superfamily, a family of enzymes with diverse substrates including histones, other proteins, arylalkylamines and aminoglycosides. *In vitro*, Hpa2 is able to acetylate specific lysine residues of histones H3 and H4 with a preference for Lys14 of histone H3. Hpa2 forms a stable dimer in solution and forms a tetramer upon binding AcCoA. The crystal structure reveals that the Hpa2 tetramer is stabilized by base-pair interactions between the adenine moieties of the bound AcCoA molecules. These base-pairs represent a novel method of stabilizing an oligomeric protein structure. Comparison of the structure of Hpa2 with those of other GNAT superfamily members illustrates a remarkably conserved fold of the catalytic domain of the GNAT family even though members of this family share low levels of sequence homology. This comparison has allowed us to better define the borders of the four sequence motifs that characterize the GNAT family, including a motif that is not discernable in histone acetyltransferases by sequence comparison alone. We discuss implications of the Hpa2 structure for the catalytic mechanism of the GNAT enzymes and the opportunity for multiple histone tail modification created by the tetrameric Hpa2 structure.

© 1999 Academic Press

Keywords: histone acetyltransferase; Gcn5-related N-acetyltransferase superfamily; acetyl coenzymeA; protein oligomer; X-ray crystallography

*Corresponding authors

Introduction

Histone acetylation has become recognized as an important factor governing gene expression *via* its effects on chromatin structure and assembly (Kuo & Allis, 1998). In the last few years several genes encoding enzymes responsible for this modification, histone acetyltransferases (HATs) have been identified and cloned including Hat1 (Kleff *et al.*, 1995; Parthun *et al.*, 1996; Verreault *et al.*, 1998), Gcn5 (Brownell *et al.*, 1996), P/CAF (Yang *et al.*, 1996), Esa1 (Clarke *et al.*, 1999; Smith *et al.*, 1998), and p300/CBP (Bannister & Kouzarides, 1996; Ogryzko *et al.*, 1996). The activity of these enzymes varies according to how many of the core histones and which particular lysine side-chain(s) they modify, and whether or not they modify histones when bound to DNA in the context of the nucleosome. Some HATs, notably p300/CBP, are also capable of modifying non-histone substrates, which makes

[†]These authors contributed equally to this work
Present addresses: M. L. Angus-Hill, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA; R. N. Dutnall, Department of Biology, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA; S. T. Tafrov and R. Sternglanz, Department of Biochemistry and Cell Biology, SUNY, Stony Brook, NY 11794, USA; V. Ramakrishnan, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Abbreviations used: AAC6, aminoglycoside 6'-N-acetyltransferase; AcCoA, acetyl coenzymeA; AAT, aminoglycoside 3-N-acetyltransferase; GNAT, Gcn5-related N-acetyltransferase; HAT, histone acetyltransferase; MAD, multiple wavelength anomalous diffraction; NAT, N-acetyltransferase; RMSD, root mean square deviation; SNAT, serotonin N-acetyltransferase.

E-mail address of the corresponding author:
rdutnall@biomail.ucsd.edu

it possible that these enzymes influence transcriptional activity *via* mechanisms that do not involve chromatin (Boyes *et al.*, 1998; Gu & Roeder, 1997; Imhof *et al.*, 1997; Waltzer & Bienz, 1998). It is therefore important to know what the true substrates of these enzymes are *in vivo* and to understand how they achieve specificity for particular lysine residues. Both Gcn5 and Esa1 have been shown to acetylate histones *in vivo* (Clarke *et al.*, 1999; Kuo *et al.*, 1998).

Sequence analysis has shown that many of the known or presumed HATs belong to a large family of *N*-acetyltransferases (NATs) called the Gcn5-related *N*-acetyltransferase (GNAT) superfamily (Neuwald & Landsman, 1997). This superfamily spans all kingdoms of life and includes enzymes with diverse substrates and cellular functions. These enzymes can acetylate the N-terminal α -amino group of proteins, the ϵ -amino group of lysine residues, aminoglycoside antibiotics, spermine/spermidine or arylalkylamines such as serotonin. There is thus considerable interest in structures of proteins from this family because of their roles in gene regulation, antibiotic resistance and hormonal regulation of circadian rhythms. The GNAT superfamily is characterized by four conserved sequence motifs (A-D) arranged in the order C-D-A-B in primary sequence. The structures of the HAT enzyme Hat1 from *Saccharomyces cerevisiae* (Dutnall *et al.*, 1998) and the aminoglycoside 3-*N*-acetyltransferase (AAT) from *Serratia marcescens* (Wolf *et al.*, 1998) provided the first look at representative proteins and shed light on their relationship to the GNAT superfamily. Since then the structures of serotonin *N*-acetyltransferase (SNAT) (Hickman *et al.*, 1999a,b) and of an aminoglycoside, 6'-*N*-acetyltransferase (AAC6) (Wybenga-Groot *et al.*, 1999) have added to our understanding of this protein family.

Hpa2 is a GNAT superfamily member from *S. cerevisiae* that is able to acetylate specific lysine residues of histones H3 and H4 *in vitro* (Lys4 and Lys14 of histone H3, Lys5 and Lys12 of histone H4), showing a preference for Lys14 of histone H3 (S.T.T. and R.S., unpublished results). Hpa2 will also autoacetylate itself in an intermolecular reaction. Deletion of the Hpa2 gene causes no obvious growth defect. Hpa2 shares a high degree of similarity with a second yeast protein, Hpa3. Remarkably, despite sharing nearly 49% sequence identity and 81% similarity over 156 residues, Hpa3 is barely able to acetylate histones *in vitro*. However, like Hpa2, Hpa3 will autoacetylate itself, demonstrating that it is nevertheless an active acetyltransferase. This presents an ideal opportunity to examine factors governing the substrate specificity of these enzymes. Here we show that Hpa2 forms a stable dimer in solution and that in the presence of the cofactor AcCoA two dimers associate to form a tetramer. We describe the structure of the tetrameric Hpa2-AcCoA complex and compare the structure to those of other GNAT

superfamily members. This has allowed us to define a conserved core structural unit in the catalytic domain of the GNAT superfamily. The structure of Hpa2 represents the first oligomeric structure of a GNAT superfamily member and illustrates a novel method of protein oligomer stabilization *via* base-pair interactions between the bound cofactor molecules. We also discuss the implications of the tetrameric Hpa2 structure for histone substrate binding.

Results and Discussion

Hpa2 is a stable dimer that forms a tetramer in the presence of acetyl CoA

Several initial observations indicated that Hpa2 might form dimers or higher oligomers, including the elution profile during gel filtration chromatography. We therefore decided to examine Hpa2 by sedimentation equilibrium analysis because it is a rigorous method for determining native molecular mass and macromolecular interactions (Laue, 1995). In the absence of AcCoA the sedimentation profile of Hpa2 could best be fit by a dimer-tetramer model fixing the molecular weight to the size of a dimer (36,406.6 Da) which yielded a K_d of 203(\pm 10) μ M (Figure 1(a); see also Materials and Methods).

In the presence of AcCoA the sedimentation profile of Hpa2 could best be fit by a single species model yielding a molecular weight of 71,454(\pm 2954) Da (expected value for an Hpa2 tetramer is 72,813.2 Da) (Figure 1(b)). Even at the lowest possible protein concentrations used, very little dimer species could be detected and we were unable to determine a dissociation constant for the tetramer. These results indicate that Hpa2 is a stable dimer in solution and that AcCoA stabilizes formation of a tetrameric Hpa2 species.

Structure determination

We solved the structure of the Hpa2-AcCoA complex using multiple wavelength anomalous diffraction (MAD) on a single crystal of selenomethionyl protein (Table 1 and Materials and Methods). Hpa2 crystallizes in the orthorhombic space group $C222_1$ with four Hpa2 molecules in the asymmetric unit. Each Hpa2 monomer contains three methionine residues (mass spectrometry and N-terminal sequencing revealed that the N-terminal methionine residue is absent) and all 12 selenium sites were located in Patterson maps using the program SOLVE. The solvent flattened electron density maps were of excellent quality (Figure 2) and revealed clearly interpretable density for nearly all residues of the four protein molecules. No electron density could be located for part of the N terminus of each monomer but otherwise all residues were clearly defined. The model includes residues 7-156 of monomer A, residues 8-156 of monomers B and C, and residues 5-156 of mono-

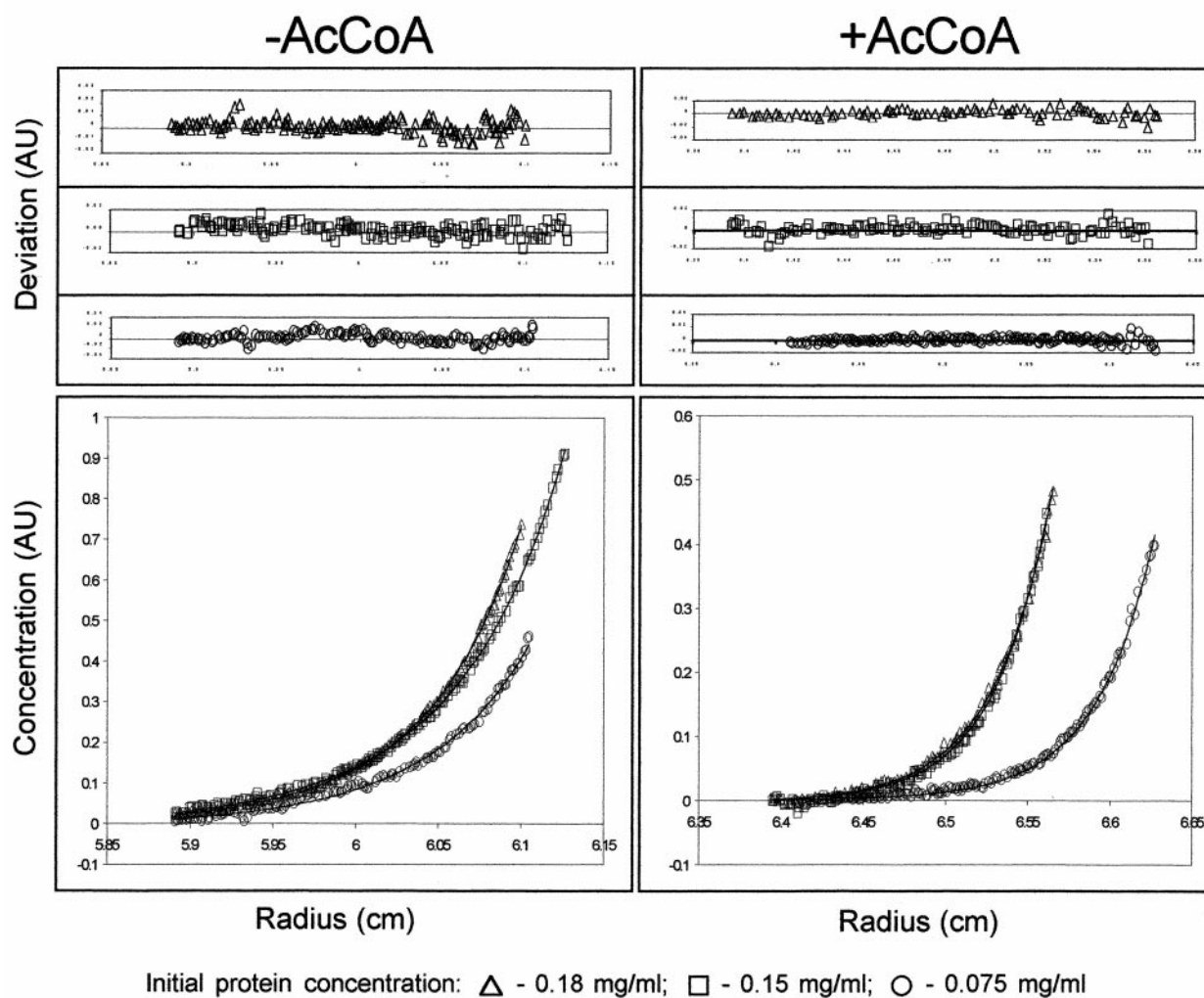


Figure 1. Equilibrium sedimentation analysis of Hpa2. Representative data sets are shown for Hpa2 in the absence (left) or presence of AcCoA (right). Three different initial protein concentrations are shown as indicated in the key. The lower panels show concentration distribution plots along with the fitted curves assuming a dimer-tetramer equilibrium (-AcCoA) or a single species (+AcCoA). The upper panels show separate plots of residual deviations. Note that in the presence of AcCoA two of the data curves (0.18 mg/ml and 0.15 mg/ml) almost coincide.

mer D, all atoms of the four AcCoA molecules and 190 solvent molecules. It has been refined to 2.4 Å resolution with an R value of 16.8% and free- R value of 24.4% (Table 1).

Independently, we have also solved the structure of Hpa2 without AcCoA using MAD (Table 2 and Materials and Methods). This model has been refined to a resolution of 2.9 Å with an R -value of 19.0% and a free- R value of 27.3%. The crystals also belong to space group $C22_1$ with similar unit cell dimensions and a single tetramer in the asymmetric unit. Overall the conformation of this tetramer is very similar to that observed in the presence of AcCoA. The difference in resolution and presence of crystal packing forces mean that caution must be used when interpreting any differences and so further discussion here will focus on the structure of the Hpa2-AcCoA complex.

Overview of the structure

In the presence of AcCoA Hpa2 forms a tetrameric structure composed of a dimer of dimers with approximate 22_2 point group symmetry (Figure 3). It can be likened to a flattened square with overall dimensions of $\sim 85 \text{ \AA} \times 80 \text{ \AA}$ (Figure 3(a)), which if viewed along one side has a dumb-bell shape varying in dimension from $\sim 40 \text{ \AA}$ at it widest to $\sim 10\text{-}20 \text{ \AA}$ at the dimer-dimer interface (Figure 3(b)). If viewed along the other side of the square it has a more constant thickness.

Description of the structure of the Hpa2 monomer

Each monomer of Hpa2 has a very similar, compact $\alpha\beta$ structure (Figures 3(c) and 4). The RMSD for C^α atoms of residues 8-156 is between 0.251

Table 1. Data collection, phasing and refinement statistics for the Hpa2-AcCoA complex

Crystal information				
Space group:	C222 ₁			
Unit cell dimensions: (Å)	<i>a</i> = 145.6	<i>b</i> = 184.1	<i>c</i> = 70.0	
Estimated solvent content:	~55% (<i>V_m</i> ~3.2 Å ³ /dalton) ^a ; 4 monomers (1 tetramer)/asymmetric unit			
Data collection				
	λ_1	λ_2	λ_3 (native)	λ_4 (native)
Wavelength (Å)	0.9803	0.9801	0.93	1.10
Resolution (Å)	2.6	2.6	2.6	2.4
<i>f</i> ' of selenium (electrons)	-9.52	-7.35	-2.19	-2.19
<i>f</i> " of selenium (electrons)	3.15	5.92	3.46	3.46
<i>I</i> / σ (highest resolution shell)	2.75	3.0	3.0	7.66
Total reflections	178,391	178,340	188,637	456,328
Unique reflections	29,729	29,721	29,983	37,150
Completeness (%)	98.7	98.7	98.9	100.0
<i>R_{sym}</i> (total/highest resol. shell) ^b	0.094/0.299	0.087/0.230	0.096/0.309	0.045/0.113
Phasing (SOLVE)				
<i>R_{cullis}</i> ^c	0.62	Phasing power ^d	centric	0.69
Figure of merit	0.49		acentric	0.79
Map correlation coefficients ^e :	SOLVE		0.582	
	SOLVE+SOLOMON		0.793	
Refinement (X-PLOR)				
Resolution range:	20.0-2.4 Å	<i>R</i> ^f	0.168	
<i>I</i> / σ cutoff:	0.0	<i>R_{free}</i> ^g	0.244	
	Number of non-hydrogen atoms:	Average <i>B</i> -factor value (Å ²)		
Protein:	5011	21.3 (main-chain); 24.2 (side-chain)		
Acetyl CoA:	204	28.8		
Water molecules:	190	29.8		
RMS deviations from ideal geometry:		Bond lengths (Å):	0.007	
		Bond angles (°):	1.1	
Residues within most allowed regions of Ramachandran plot (%):			99.8	

^a *V_m* = Matthew's coefficient (volume of asymmetric unit/molecular weight).

^b *R_{sym}* = $\sum_h \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l I_{hl}$, where $\langle I_h \rangle$ is the mean of the observations *I_{hl}* of reflection *h*.

^c *R_{cullis}* = $\sum |E| / \sum |F_i| - |F_{10}|$, where *E* is the lack of closure.

^d Phasing power = $(|F_H(\text{calc})|/|E|)$, where *F_H*(calc) is the calculated anomalous difference and *E* is the lack of closure.

^e Map correlation coefficient was calculated between the calculated map for the final refined model and the experimental maps before and after solvent flattening.

^f *R* = $\sum (|F_p(\text{obs})| - |F_p(\text{calc})|) / \sum |F_p(\text{obs})|$.

^g *R_{free}* = *R*-factor for a selected subset (10%) of the reflections that was not included in prior refinement calculations.

and 0.390 Å, a value within the expected experimental coordinate error at this resolution (Luzzatti, 1952). The core of the structure contains a central mixed five-stranded sheet structure (strands β 1- β 5). Strands β 1- β 4 are arranged in an anti-parallel fashion while strands β 4 and β 5 interact in a parallel fashion but only at their amino-terminal ends. At the other end, they are splayed apart because of a β -bulge in strand β 4 formed by residue N74 of strand β 3 and N91 and D92 of strand β 4. The central sheet is flanked on each side by two α -helices. Helices α 1 and α 2 are on one side of the sheet with helix α 1 lying almost flat against and perpendicular to the direction of the strands, while helices α 3 and α 4 are on the other side with helix α 3 cupped within the curved face of the sheet.

Hpa2 forms dimers *via* an extensive, interdigitated, and largely hydrophobic interface

The two monomers in each Hpa2 dimer (monomers A + B or C + D) form an extensive interface

that buries approximately 2900 Å² of surface area per monomer. Most of the secondary structure elements of the monomer contribute residues involved in dimer contacts (Figure 4). A large part of the interface is formed by two projections from the core part of the monomer structure. The first projection is formed by the C-terminal end of strand β 3, turn β 3 β 4, and the N-terminal end of strand β 4, while the second is formed by strand β 7 (Figure 3(c)). Together with strands β 5 and β 6 they form a barrel-like structure containing ten strands (Figure 3(d)) in which the component strands of the barrel interdigitate. Most significantly, strand β 7 from each monomer interacts between strands β 5 and β 6 of the opposite monomer, which also extends the central sheet structure by two strands. The two projections also interact with residues from helices α 1 and α 2, turn α 1 α 2, turn α 2 β 2 and helices α 3 and α 4 of the opposite monomer. Most of the interactions between monomers, excluding the main-chain interactions formed between strands, are hydrophobic in nature although a few polar interactions are also made.

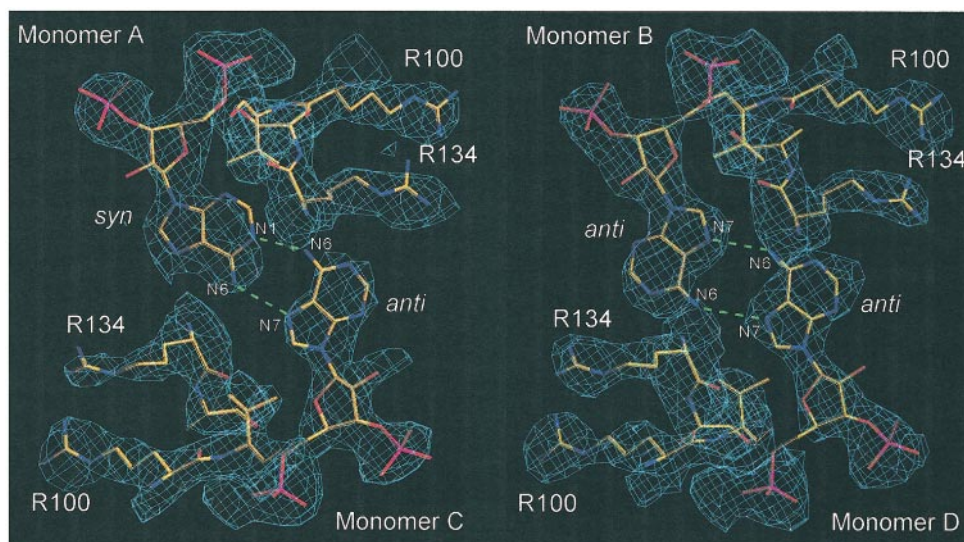


Figure 2. Representative electron density map showing adenine-adenine base-pairing between Hpa2 dimers. Solvent flattened experimental electron density map at 2.6 Å resolution contoured at 2 σ . Two regions of the map are shown (see the text for details); the *syn-anti* base-pair between the AcCoA molecules bound by monomers A and C (left panel) and *anti-anti* base-pair between AcCoA bound by monomers B and D (right panel). Selected neighboring protein side-chains are also shown. Green broken lines indicate hydrogen bonds.

The characteristics of the dimer interface described above, particularly the extent of the interface and the interdigitation of structural elements, make it highly unlikely that any Hpa2 protein would exist as a monomer in solution. This would account for the absence of monomer species in equilibrium sedimentation studies and the elution profile of Hpa2 from a gel filtration column.

The Hpa2 tetramer is formed by a largely polar interface between dimers which includes adenine: adenine base-pair interactions between acetyl CoA molecules

Our equilibrium sedimentation analysis shows that Hpa2 forms a tetramer in the presence of AcCoA. The Hpa2-AcCoA complex structure shows that two Hpa2 dimers associate *via* interactions between structure elements at the C terminus of each monomer (helix $\alpha 4$, strands $\beta 6$ and $\beta 7$; Figures 3 and 4). This interface buries approximately 800 Å² per dimer and is predominantly polar; including a cluster of salt bridges between symmetrically related residues E130 and K145 at the center of the interface. A remarkable feature of the interface is the presence of base-pair interactions between the adenine moieties at the end of each AcCoA molecule, one between the AcCoA bound by monomers A and C and the other between monomers B and D (Figures 2, 3(a) and 3(b)). However, unlike other interactions between the four monomers, these base-pairs are not symmetrical. In the AcCoA molecule bound by monomer A, the adenine adopts a *syn* conformation with respect to the ribose ring, whereas in the others it is in an *anti* conformation. Thus one base-

pair is an asymmetric *syn-anti* base-pair with hydrogen bonds between the N6 and N1 of one adenine (*syn*) and the N6 and N7 of the other, while the other is a symmetrical *anti-anti* base-pair involving the N6 and N7 of both bases. Each adenine base may also be contacted by a hydrogen bond from the side-chain of R134 from the opposite dimer (for example, the N^ε of R134 from monomer A is 3.72 Å from the N1 nitrogen of the adenine of the AcCoA bound by monomer C). These base-pairs and inter-dimer contacts involving the AcCoA may account for the stabilizing effect of AcCoA on tetramer formation observed by equilibrium sedimentation analysis. While other examples of oligomeric AcCoA-binding proteins exist (Engel & Wierenga, 1996), this is, to our knowledge, the first example of an oligomeric structure stabilized by base-pair interactions between AcCoA molecules.

The conformation of the base in the AcCoA bound by monomer A is the major difference between the four monomers in the Hpa2 tetramer. The conformation of protein side-chains, including those in the vicinity of the adenine bases is extremely similar (compare for example, the conformations of R100 and R134 in Figure 2). It seems probable that in solution all four adenine bases would be in the energetically more favorable *anti* conformation and that the tetramer would therefore be symmetrical. The *syn* orientation appears to be due to crystal packing forces as the local environment of the two base-pairs in the crystal lattice is different. The adenine ring of the AcCoA bound by monomer A is close to the N terminus of a monomer from an adjacent Hpa2 tetramer related by crystal symmetry, although a small space exists between them. Electron density for the

first few residues of this protein monomer is not visible, but it is clear from biochemical analysis that these residues exist in the crystal, presumably existing in multiple conformations. It is most likely that these residues occupy the space described above, leading to the exclusion of the *anti* conformation of the adenine base.

A common core fold for the catalytic domain of the GNAT superfamily

Comparison of the structure of Hpa2 with the structures of other GNAT superfamily members shows that they share a common structural fold around which a more variable repertoire of structures is built (Figures 4 and 5) (Modis & Wierenga, 1998). This core fold features the four conserved sequence motifs of the GNAT family and comprises a central highly curved five-stranded β -sheet (strands β 1- β 5 of Hpa2) flanked on both sides by helical segments (helices α 1 and α 3 of Hpa2).

Comparison of RMSD values for superposition of the C^α atoms of structurally equivalent residues with the degree of conservation between these proteins shows that they share a very similar core fold

structure despite a low level of sequence identity (Table 3). This is a feature of the GNAT family, which contains very few invariant residues (Neuwald & Landsman, 1997). The degree of conservation of the four GNAT sequence motifs also varies considerably. Whereas motif A, the longest and most highly conserved motif, is universally present, motifs B, C and D are not always identified by the multiple alignment-database search program used to discover the GNAT family (Neuwald *et al.*, 1997). Nearly all members contain a motif B and D, but motif C is often not identifiable, for example in HAT members of the family (Neuwald & Landsman, 1997). It is therefore useful to address these motifs individually and briefly describe the way in which they contribute to the common structural fold.

Motifs D and A

These motifs are the most similar in structure. Motif A forms a $\beta\alpha\beta$ unit that contributes many of the most critical contacts to the AcCoA molecule. It also contains most of the residues that are invariant across the GNAT family, notably the motif

Table 2. Data collection, phasing and refinement statistics for the Hpa2 (without AcCoA)

<i>Crystal information</i>			
Space group:	C222 ₁		
Unit cell dimensions: (Å)	<i>a</i> = 147.5	<i>b</i> = 187.1 (1 tetramer)	<i>c</i> = 68.7
Estimated solvent content:	~55% ($V_m \sim 3.2 \text{ \AA}^3/\text{dalton}$) ^a ; 4 monomers/asymmetric unit		
<i>Data collection</i>			
	λ_1	Beamline X12C	λ_3 (native)
Wavelength (Å)	0.9790	λ_2	0.93
Resolution (Å)	2.9	2.9	2.9
<i>f'</i> of selenium (electrons)	-9.52	-7.35	-2.19
<i>f''</i> of selenium (electrons)	3.15	5.92	3.46
<i>I</i> / σ (highest resolution shell)	4.7	4.1	3.2
Total reflections	150,487	150,470	168,213
Unique reflections	21,537	21,529	21,539
Completeness (%)	99.5	99.5	99.7
<i>R</i> _{sym} (total/highest resol. shell) ^b	0.062/0.163	0.072/0.184	0.084/0.231
<i>Phasing (SOLVE)</i>			
<i>R</i> _{cullis} ^c 0.61	Phasing power ^d	centric	0.74
Figure of merit 0.50		acentric	0.87
Map correlation coefficients ^e :	SOLVE		0.54
	SOLVE+SOLOMON		0.76
<i>Refinement (X-PLOR)</i>			
Resolution range:	20.0-2.9 Å	<i>R</i> ^f	0.190
<i>I</i> / σ cutoff:	0.0	<i>R</i> _{free} ^g	0.273
Number of non-hydrogen atoms:	Protein:	4980	
Average <i>B</i> -factor value (Å ²):	23.9 (main-chain); 25.4 (side-chain)		
RMS deviations from ideal geometry:			
	Bond lengths (Å):		0.007
	Bond angles (°):		1.2
Residues within most allowed regions of Ramachandran plot (%):			
			98.2

^a V_m = Matthew's coefficient (volume of asymmetric unit/molecular weight).

^b $R_{sym} = \sum_h \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean of the observations I_{hl} of reflection h .

^c $R_{cullis} = \sum |E| / \sum ||F_l| - |F_0||$, where E is the lack of closure.

^d Phasing power = $(|F_H(\text{calc})| / |E|)$, where $F_H(\text{calc})$ is the calculated anomalous difference and E is the lack of closure.

^e Map correlation coefficient was calculated between the calculated map for the final refined model and the experimental maps before and after solvent flattening.

^f $R = \sum (|F_p(\text{obs})| - |F_p(\text{calc})|) / \sum |F_p(\text{obs})|$.

^g R_{free} = *R*-factor for a selected subset (10%) of the reflections that was not included in prior refinement calculations.

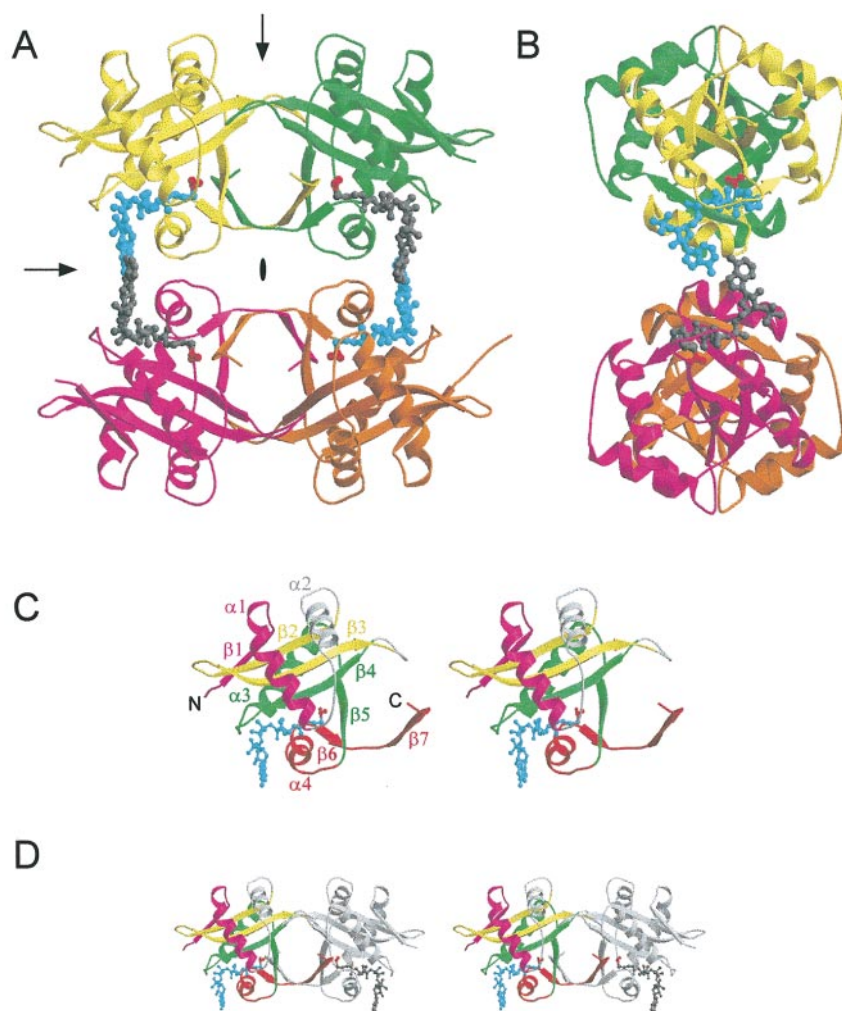


Figure 3. Overview of the structure of the Hpa2-AcCoA complex. (a) and (b) Structure of the Hpa2 tetramer. Each monomer of Hpa2 is colored differently for contrast (monomer A, yellow; monomer B, green; monomer C, magenta; monomer D, orange). The AcCoA molecules are shown in a ball-and-stick representation, colored either cyan or grey for contrast. The acetyl group of each cofactor molecule is colored red to highlight the position of the active sites. Arrows or ovals indicate pseudo-2-fold symmetry axes that lie in the plane of the paper or perpendicular to the plane of the paper, respectively. In (b) the view is from the left side of the tetramer with respect to (a) and the AcCoA molecules bound by monomers B and D have been omitted for clarity. (c) Stereo view of the Hpa2 monomer. The four conserved GNAT motifs are colored magenta (C), yellow (D), green (A) and red (B). The N and C termini, and secondary structure elements are indicated. (d) Stereo view of the Hpa2 dimer. One monomer is colored as in (c) to show contribution of GNAT motifs to dimer interface.

Q/R-x-x-G-x-G/A (see below). This combination of structural and functional importance would explain why it is the most easily recognized and universally present motif. Although Motif D does not contribute any AcCoA binding residues it undoubtedly forms an important part of the framework of the structure. This makes it unlikely that it would be absent in some structures; it is more likely that it simply falls below the detection level employed in database searches.

Motif C

The structural comparison reveals that Hpa2 and Hat1 do in fact contain a structure similar to motif C of other GNAT superfamily members. Motif C of AAT, SNAT and AAC6 forms a strand-turn-helix unit (Figures 4 and 5). This corresponds to strand $\beta 1$ and helix $\alpha 1$ of Hpa2 or strand $\beta 11$ and helices $\alpha 6$ and $\alpha 7$ of Hat1. Although structurally similar there is very little sequence similarity between Hpa2, Hat1 and the other proteins in this region (Figure 4) which would explain how it escapes detection in database searches. There is also some variation in the length and orientation of

the helical segment of this motif, most notably for helices $\alpha 6$ and $\alpha 7$ of Hat1. However, a remarkable feature is that the C-terminal end of the helical segment occupies a very similar position with respect to the central sheet structure, forming one side of the channel occupied by the AcCoA molecule. In Hpa2, Hat1 and SNAT residues from the C-terminal end of this helix make contacts to the AcCoA. It is also interesting to note that comparison of the structure of SNAT with and without substrate reveals that this motif can undergo a dramatic conformational change upon binding a bisubstrate analog (Hickman *et al.*, 1999b). This does not appear to be the case with Hpa2 where the structure of this motif is unaltered upon binding AcCoA indicating a possible difference in conformational variability between GNAT members upon substrate binding.

Motif B

This is the most C-terminal motif, following immediately after the final strand of the central sheet. There is a great deal of variation in the structure formed by this motif and it is therefore poss-

Table 3. Structure and sequence similarity between Hpa2, Hat1 and AAT

Comparison	Number of residues	C α RMSD (Å) ^a	Sequence similarity (%)	
			Identity	Similarity
Hpa2 <i>versus</i> AAT	88	1.51	23.9	48.9
AAC6 <i>versus</i> AAT	85	1.57	27.1	58.8
AAT <i>versus</i> SNAT	83	1.73	15.7	50.6
Hat1 <i>versus</i> SNAT	92	2.03	9.8	43.5
Hpa2 <i>versus</i> Hat1	101	2.17	11.9	38.6
Hpa2 <i>versus</i> SNAT	99	2.24	13.1	41.4
AAC6 <i>versus</i> Hat1	95	2.44	7.4	41.1
Hat1 <i>versus</i> AAT	89	2.49	9.0	43.8
AAC6 <i>versus</i> Hpa2	108	n.a.	22.2	51.9
AAC6 <i>versus</i> SNAT	96	n.a.	19.8	44.8

^a Number of structurally equivalent residues compared. For comparisons with AAT this only includes GNAT motifs C, D and A, otherwise all GNAT motifs are included.

ible that the apparent sequence conservation in this motif is purely coincidental and does not reflect a common structural origin (Wybenga-Groot *et al.*, 1999). Among the proteins considered in this comparison Hpa2, SNAT and AAC6 appear to be the most similar with respect to motif B.

A common function of the GNAT family members is to bind AcCoA and it appears that much of the common fold is involved in creating a stable cofactor binding platform. This is particularly true for motifs D and A, which would account for the invariance in the way that they contribute to the fold.

Motifs B and C may be more variable for a number of reasons. Firstly, they make up the more exterior portions of the catalytic domain and so could be influenced by the need to interact with adjacent polypeptide domains. In Hat1 helices $\alpha 6$ and $\alpha 7$ of motif C are involved in interactions with the N-terminal domain of the protein (Figure 5). In some GNAT family members the catalytic HAT domain is embedded in a much larger protein and would likely interact with other domains of the protein, possibly providing a way of regulating the activity of the HAT domain. Secondly, structural differences may be related to differences in the oligomeric states of GNAT superfamily members. In Hpa2 motif B makes up a large part of the dimer interface and the entire tetramer interface (Figure 4). Wolf *et al.* (1998) observed that AAT also forms a dimer in the crystal lattice which involves motif B, but there are several differences in the nature of the Hpa2 and AAT dimers. Other members of the GNAT family have been shown to dimerize, such as spermidine/spermine NAT (Coleman *et al.*, 1996) and ARD1, a protein NAT from yeast (Park & Szostak, 1992).

Structural variation of motifs B and C may also reflect the diverse substrates of the GNAT superfamily. Motifs B and C, and the variable structure between motifs C and D, contribute to differences in architecture around the active site. In Hat1 and Hpa2 these regions contribute to a channel that could bind a histone substrate (Dutnall *et al.*, 1998) while in AAT and AAC6 they contribute to the walls of an acidic slot which is a likely binding site

for an aminoglycoside molecule (Wolf *et al.*, 1998; Wybenga-Groot *et al.*, 1999). As noted previously, motif C of SNAT undergoes a conformational change upon substrate binding which helps to create a binding site for the substrate serotonin (Hickman *et al.*, 1999b).

A conserved mode of acetyl CoA binding

The conformation of the AcCoA bound by each monomer of Hpa2 and orientation with respect to the protein is virtually indistinguishable from that observed in the other structures of GNAT superfamily members. For example, if the structures of Hpa2 and Hat1 are superimposed using the common core structural fold, the RMSD for the AcCoA molecules (excluding the 3'-phosphoribose and adenine base) is ~ 0.3 Å, well within the expected coordinated error of each structure.

In Hpa2, as in the other structures, most of the hydrogen bonding interactions with the cofactor are highly conserved and made *via* main-chain groups rather than side-chains (Figure 6). This would explain why these proteins bind cofactor in the same way despite the low degree of sequence conservation. The conserved main-chain contacts include those from the invariant Q/RxxGxG/A segment in motif A. This segment forms a loop prior to, and part of the first turn of the helix of motif A (residues 100-104 of Hpa2). Most of the residues in this loop contribute to a network of hydrogen bond interactions with the α and β phosphate oxygen atoms *via* main-chain groups, which includes a conserved solvent molecule interaction (Figure 6). We could not locate electron density for this loop (residues 96-104) in the structure of Hpa2 solved in the absence of cofactor, indicating that this region may only become structured upon cofactor binding. For many GNAT superfamily members mutation of even a single residue within this motif is sufficient to drastically reduce or abolish enzyme activity (discussed by Dutnall *et al.*, 1998; Wolf *et al.*, 1998). It therefore appears that this network of interactions is a critical determinant of AcCoA binding.

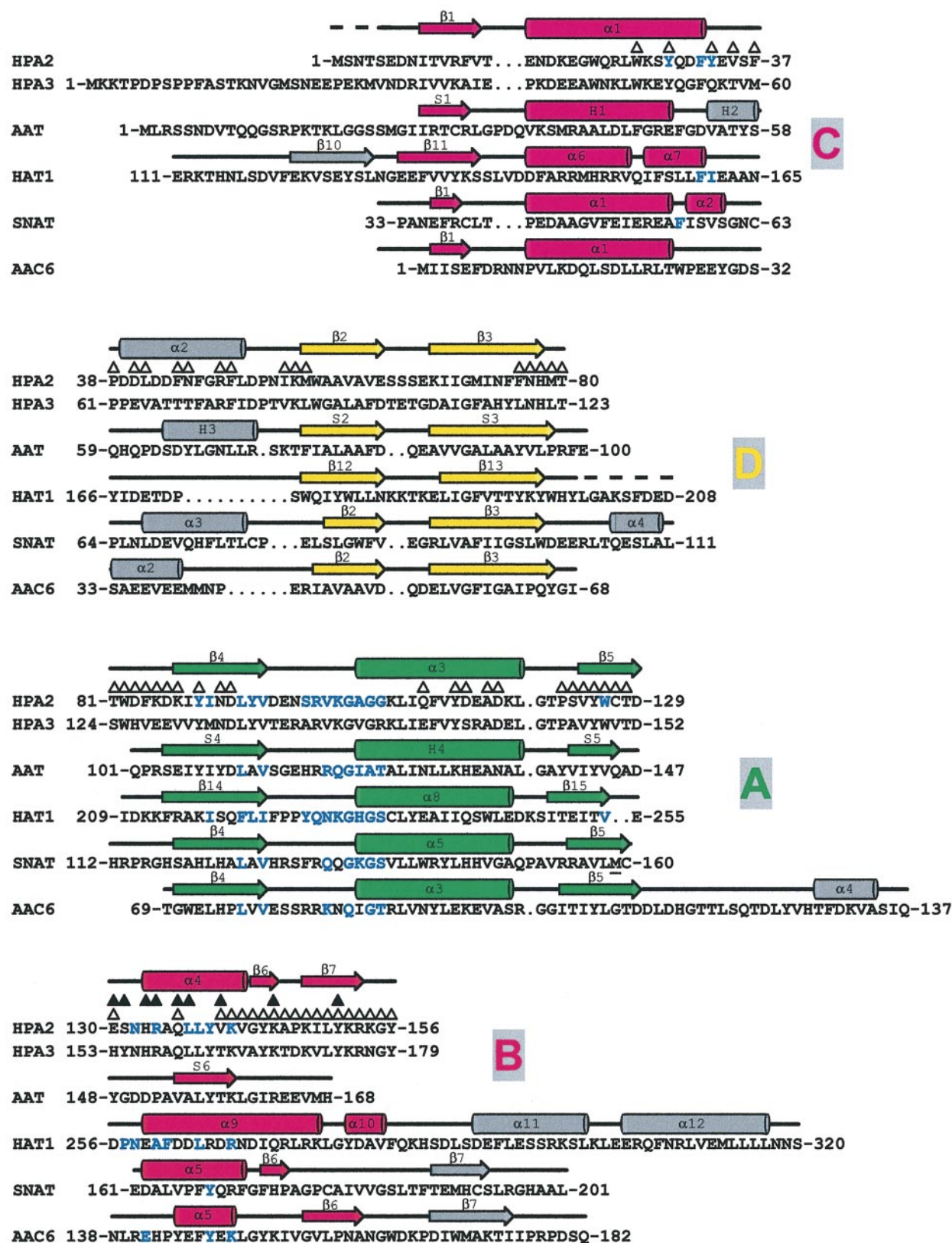


Figure 4. Sequence alignment of Hpa2, Hpa3 and other GNAT superfamily members. Sequences of Hpa2 (Genbank accession number 786308), Hpa3 (Genbank accession number 603252), *Serratia marcescens* aminoglycoside 3-*N*-acetyltransferase (AAT); the C-terminal domain of Hat1 (residues 111-320), sheep serotonin acetyltransferase (SNAT; also called arylalkylamine *N*-acetyltransferase), and *Enterococcus faecium* aminoglycoside 6'-*N*-acetyltransferase (AAC6) are shown aligned according to the conserved GNAT motifs. Secondary structure elements for each protein are indicated above the sequences colored according to GNAT sequence motif. Structure elements outside of these motifs are colored grey. Residues highlighted in blue are involved in contacts to AcCoA. Symbols above the sequence of Hpa2 indicate residues involved in dimer formation (Δ) and tetramer formation (\blacktriangle).

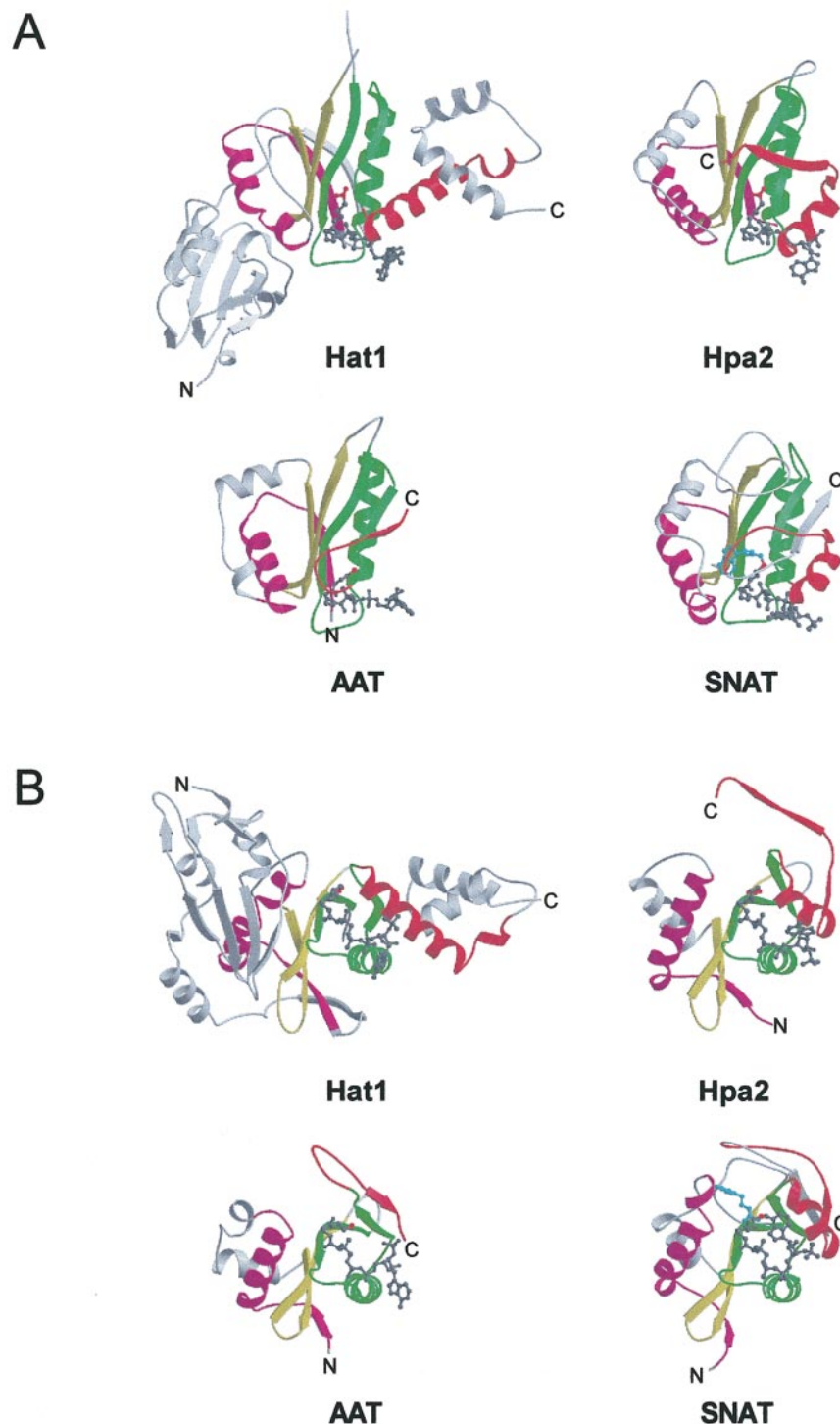


Figure 5. Comparison of the structures of Hat1, Hpa2, AAT and SNAT. Two schematic views of each structure are shown in (a) and (b). The structures are superimposed according to the common fold of the catalytic domain but shown side-by-side for easier comparison. Segments of the structure are colored according to the GNAT sequence motifs. N and C termini of each protein are indicated. In (b) each structure has been rotated by 90° along the horizontal axis.

Implications for catalysis in the GNAT family

The GNAT family does not contain a universally conserved residue that could be imagined to play a catalytic role. In the structures solved to date there

is also no cysteine side-chain near to the acetyl group that could act as an intermediate acceptor in an acyl transfer mechanism. Instead it seems more likely that the acetylation reaction proceeds *via* a direct nucleophilic attack by the amino group of

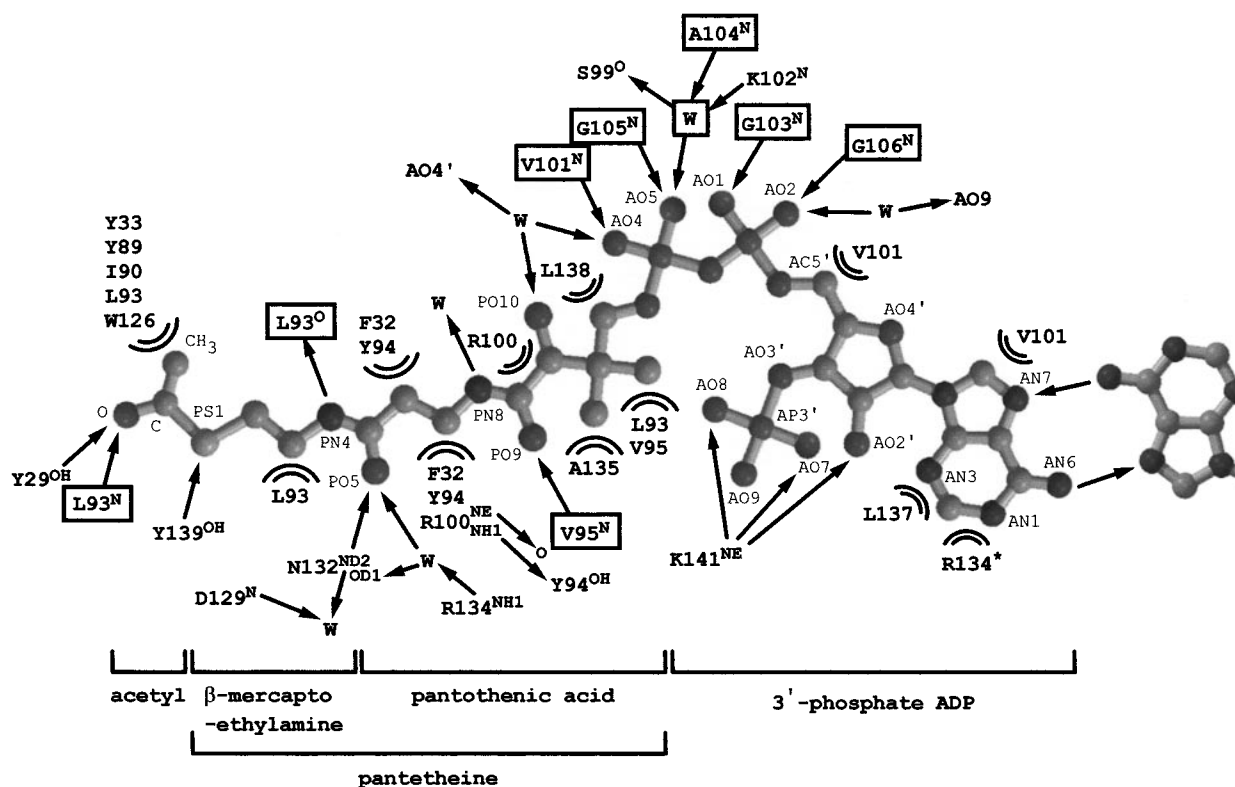


Figure 6. Conservation of Hpa2-AcCoA interactions. Schematic representation of Hpa2-AcCoA contacts highlighting conservation with other GNAT superfamily members. The AcCoA molecule is shown stretched out for clarity with contacts indicated by arrows (hydrogen bonds) or concentric semi-circles (hydrophobic contacts). Contacts that are highly conserved with the other GNAT superfamily member structures are boxed. The asterisk indicates a contact from a second Hpa2 monomer.

the substrate on the carbonyl carbon of the acetyl group. Each enzyme may facilitate catalysis by ordering the substrates for reaction and by promoting an appropriate charge state on the amino group to be modified.

A central problem for GNAT family members is the high pK_a value of the substrate amino group, requiring the enzyme to deprotonate this at some stage prior to acyl transfer. For Hat1, AAT, and AAC6 the protein surface in the vicinity of the acetyl group has a negative electrostatic potential (Dutnall *et al.*, 1998; Wolf *et al.*, 1998; Wybenga-Groot *et al.*, 1999). In principal this would favor a positive charge on the amino group and so for each of these proteins a mechanism may exist to transfer a proton from the amino group during substrate binding. Support for this comes from the structure of the SNAT-bisubstrate analog complex where a chain of solvent molecules runs from the active site to bulk solvent *via* a pair of histidine residues (Hickman *et al.*, 1999b). Hickman *et al.* (1999b) propose that this could act as a sink or a conduit to transfer charge away from the substrate. In Hat1 and other enzymes it is possible that acidic side-chains in the vicinity of the active site could also play a role in proton transfer. For Hpa2 the situation seems more clear cut in that the protein surface around the active site is characterized by a positive electrostatic potential (Figure 7(a)), which

would favor an uncharged state. Two conserved features of the GNAT superfamily may help catalysis. Firstly, in each structure there are a number of main-chain carbonyl groups without hydrogen bonding partners in the active site. These could act in a proton transfer pathway either directly or by helping to locate water molecules. Secondly, there is a small hydrophobic pocket around the acetyl group. Once the amino group is deprotonated this pocket would then stabilize the neutral charge while the substrate is bound to the enzyme.

Hpa2 shows similarity with SNAT with respect to the positioning of a tyrosine side-chain within hydrogen bonding distance of the sulfur atom of CoA. Tyr168 of SNAT may play a role in the reprotonation of the thiolate leaving group and mutagenesis of this residue has a dramatic effect on the activity of SNAT (Hickman *et al.*, 1999b). In Hpa2, Tyr139 is positioned in a very similar fashion, suggesting that it too may play a role in catalysis. Comparison of the sequences of the GNAT superfamily members described here shows that five of them have a tyrosine at this position (Figure 4). In the structure of AAC6, Tyr147 is also within hydrogen binding distance of the sulfur atom. However, Tyr157 of AAT is not positioned appropriately, although it cannot be ruled out that it would be so after substrate binding or that the difference is because this structure contains CoA

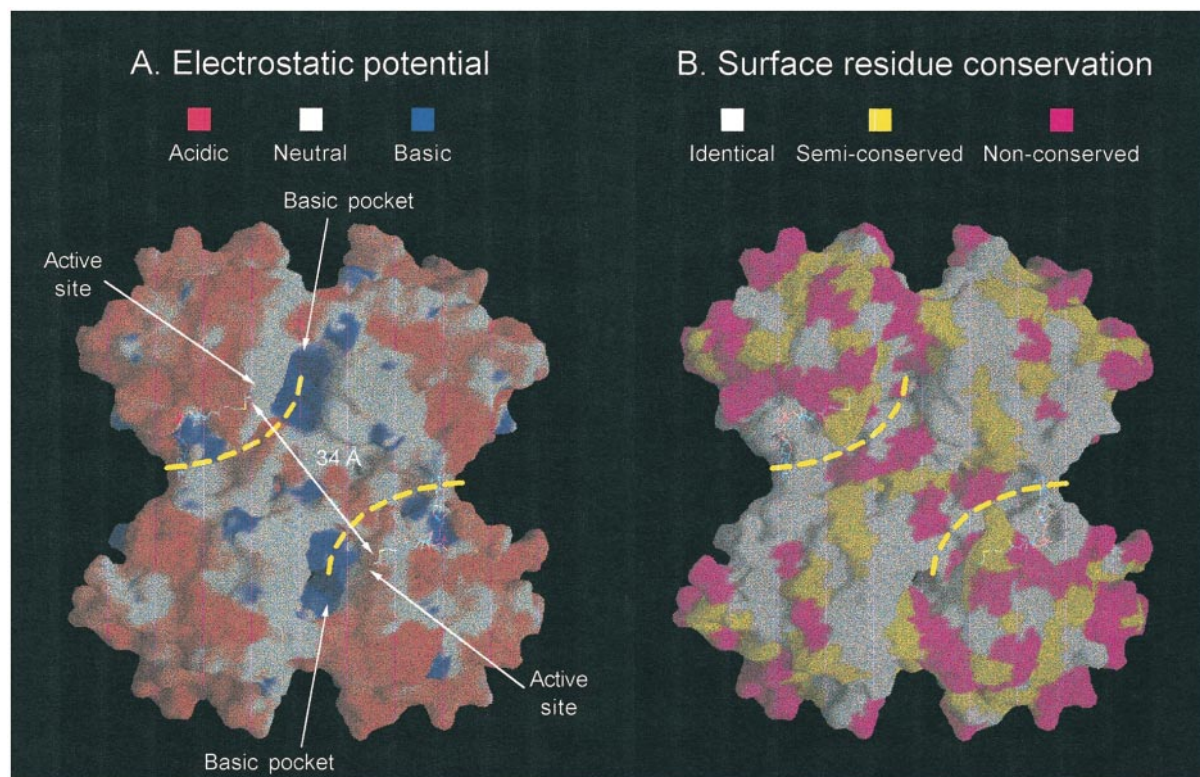


Figure 7. Surface chemistry of Hpa2 tetramer. (a) Molecular surface of Hpa2 tetramer colored according to electrostatic potential as indicated. The basic pocket adjacent to the active sites is indicated. (b) Molecular surface colored according to sequence conservation between Hpa2 and Hpa3 as indicated. In both (a) and (b) the view of the tetramer is as in Figure 3(a) and the surface is shown partially transparent to show the location of the AcCoA molecules of the two monomers whose active site is exposed on this face of the tetramer. The yellow broken lines indicate possible peptide binding channels. Note that the molecular symmetry of Hpa2 means that the upper left and lower right quarters of the surface show the active site face of an Hpa2 monomer, whereas the upper right and lower left quarters show the opposite face of the protein.

rather than AcCoA. In Hat1 the corresponding residue is Arg265, but this also points away from the sulfur atom. It is therefore possible that as with the earlier steps of the reaction, GNAT superfamily members can use alternative mechanisms to reprotonate the CoA.

Implications for substrate binding and histone acetylation by Hpa2

In vitro Hpa2 can acetylate histones H3 and H4 at specific lysine residues (Lys4 and Lys14 of histone H3; Lys5 and Lys12 of histone H4) with a preference for Lys14 of histone H3. Examination of the structure of Hat1 led us to propose a model for histone substrate binding, which could account for its observed specificity (Dutnall *et al.*, 1998). A channel of varying width and depth crosses the surface of the Hat1 protein, which can accommodate a six to seven residue peptide in an extended conformation. A possible peptide binding channel is also visible on the surface of Hpa2 (Figure 7). Hpa2 and Hat1 differ in the region of the active site that has potential consequences for substrate binding. The

architecture of Hpa2 creates a pocket adjacent to the active site, which is closed off on two sides by the second Hpa2 monomer in each dimer (Figure 7(a)). This pocket places restriction on the conformation of a polypeptide backbone in the vicinity of the active site. It is therefore possible that Hpa2 is able to discriminate between potential substrates *via* an "indirect read-out" mechanism whereby the only lysine side-chains that can enter the active site are those where the surrounding polypeptide can adopt a conformation that can fit into the pocket.

As described in the Introduction Hpa2 shares a high degree of sequence similarity with another yeast protein, Hpa3, which means that the structure of an Hpa3 monomer will be very similar to that of Hpa2 (Figure 4). However, unlike Hpa2, Hpa3 is barely able to acetylate histones and therefore residues that are not conserved between these proteins are potentially involved in substrate specificity. Consistent with this, it is noticeable that non-conserved surface residues map mostly to the same face of the protein as the active site (Figure 7(b)). These differences, as well as possible

subtle differences in the orientation of the monomers (such as observed for Hpa2 and AAT) may change the architecture of the active site such that histone binding by Hpa3 is precluded. Structural analysis of Hpa2-peptide complexes and of Hpa3 will be required to investigate this further.

The tetrameric structure of Hpa2 has consequences for peptide binding which distinguish it from monomeric enzymes such as Hat1. Firstly, the walls of the pocket around each active site are contributed by both monomers of an Hpa2 dimer. It is therefore likely that both subunits will provide substrate binding residues. Secondly, two active sites are exposed on each of the square faces of the Hpa2 structure. The direct linear distance between these active sites is approximately 34 Å. Hpa2 can acetylate Lys4 and Lys14 of histone H3 or Lys5 and Lys12 of histone H4. Thus it is at least conceivable that for histone H3 a single tail could span the two active sites so that Lys4 and Lys14 could be acetylated simultaneously. However, the general architecture of Hpa2 makes us believe that this is an unlikely scenario.

It is more likely that Hpa2 could coordinately modify lysine residues in two separate histone tails. Histones H3 and H4 form a heterotetrameric structure in the nucleosome or under physiological conditions in the absence of DNA (van Holde, 1989). The arrangement of the Hpa2 tetramer and the structure of the nucleosome (Luger *et al.*, 1997) makes it unlikely that Hpa2 could modify all four of the histone H3 and histone H4 tails. However, it is quite conceivable that Hpa2 could modify both histone H3 tails, or both histone H4 tails, or a histone H3 and a histone H4 tail. The Hpa2 structure therefore illustrates a possible mechanism for ensuring that both copies of a particular histone in a nucleosome are modified in the same way. It also illustrates how a modification of histone H3 could be linked to a modification of histone H4. Furthermore it illustrates how modifications of two or more separate nucleosomes could be coordinated. Such coordinated modifications could play a role in transcriptional activation, chromatin assembly or the inheritance of particular chromatin states.

Concluding remarks

The structure of Hpa2 illustrates that the highly conserved GNAT catalytic fold is ideally suited to forming dimeric and tetrameric structures. For Hpa2 a tetrameric structure is stabilized surprisingly by base-pair interactions between bound AcCoA molecules. Formation of a tetrameric structure has consequences for mechanisms of substrate binding and for possible allosteric regulation of enzyme activity. Future structural studies of substrate complexes of these enzymes, as well as other representative members, promise to yield further insights into the remarkable diversity of the GNAT family.

Materials and Methods

Protein expression and purification

Hpa2 was overexpressed in *Escherichia coli* strain BL21(DE3) using the T7 polymerase expression system (Studier *et al.*, 1990). Selenomethionine-substituted protein was produced by using the *E. coli* methionine auxotrophic strain B834(DE3) grown in a defined medium containing seleno-L-methionine as described previously (Ramakrishnan *et al.*, 1993). Protein was purified from soluble extracts through a combination of anion exchange, hydroxyapatite and gel-filtration chromatography. The protein was dialyzed into 5 mM Hepes (pH 7.8), 100 mM NaCl, 1 mM DTT and concentrated to 20 mg/ml.

Equilibrium sedimentation

Equilibrium sedimentation analysis was carried out using a Beckman XL-A analytical centrifuge. Samples were spun at 22,000 rpm at 20 °C and monitored by scanning at 280 and 360 nm. Absorbance profiles were compared at different times to verify that the samples had reached equilibrium. Eight different data sets were collected in the absence of AcCoA at protein concentrations ranging from 400 µg/ml to 40 µg/ml in 5 mM Hepes (pH 7.8), 50 mM NaCl. These data sets were globally fit to a dimer/tetramer model using non-linear least squares analysis (McRorie & Voelker, 1993). To test the effect of AcCoA, four data sets were collected using a 1:1 molar ratio of Hpa2 monomer: AcCoA and protein concentrations from 180 µg/ml to 40 µg/ml. AcCoA was included in the reference cell at the same concentration as in the sample cell. These data sets were globally fit to a single species model. In both cases solvent density was measured to be 1.004 g/ml and the partial specific volume was estimated to be 0.732 ml/g based on amino acid composition (McRorie & Voelker, 1993). We also subtracted the absorbance measured at 360 nm to correct for aberrations in the sample cell. The presence of Hpa2-AcCoA binding equilibria results in an uneven distribution of AcCoA in the sample cell which cannot be mimicked in the reference cell under the conditions of this experiment. Also AcCoA absorbs UV radiation significantly at the wavelength used to monitor protein, which means that the AcCoA binding equilibria cannot be removed by adding a vast excess of cofactor.

Crystallization and data collection

Prior to crystallization, Hpa2 protein (10 mg/ml) was incubated with a twofold molar excess of AcCoA at room temperature for one hour in 5 mM Hepes (pH 7.8), 50 mM NaCl, 1 mM DTT. Crystals were grown at 4 °C by hanging drop vapor diffusion against 8% (w/v) PEG4000, 0.1 M Mes (pH 6.9), 15 M Ca(CH₃CO₂)₂. For analysis crystals were transiently soaked in cryoprotectant solutions containing 8% PEG4000, 0.1 M Mes (pH 6.0), 15 M Ca(CH₃CO₂)₂, twofold molar excess of AcCoA, and increasing concentrations of PEG400 (up to 30%), and flash-frozen in liquid nitrogen. Crystals of Hpa2 in the absence of AcCoA were grown by hanging drop vapor diffusion against 15% PEG4000, 0.1 M Tris (pH 8), 0.15 M Ca(CH₃CO₂)₂, 20% (v/v) glycerol at 4 °C and frozen directly in liquid nitrogen.

MAD data sets were collected on single crystals of selenomethionine-substituted Hpa2 or Hpa2 + AcCoA at the Brookhaven National Laboratory Synchrotron Light

Source on beamlines X12-C and X25. The crystals were maintained at 100 K using an Oxford Cryostream and data were collected in 1° oscillations. For the Hpa2 + AcCoA crystal data sets were collected at four wavelengths in single sweeps: λ_1 was the inflection wavelength at 0.9803 Å, λ_2 was at the K-edge of selenium at 0.9801 Å, and λ_3 and λ_4 were remote wavelengths (0.93 Å and 1.10 Å, respectively). For the Hpa2 crystal three data sets were collected at the selenium inflection (λ_1), K-edge (λ_2) and a remote wavelength (λ_3). Data were processed using the programs DENZO and SCALEPACK (Otwinowski & Minor, 1997).

Structure determination and refinement

The MAD data were phased using the program SOLVE (Terwilliger & Berendzen, 1999). These phases were used to calculate electron density maps using the CCP4 program suite (Collaborative Computational Project Number 4, 1994) which were solvent flattened using SOLOMON (Abrahams & Leslie, 1996) with a calculated solvent value of 55%. Model building was carried out using the program O (Jones *et al.*, 1991). For the Hpa2-AcCoA complex refinement was carried out against the λ_4 data from 6.0-2.4 Å using the program XPLOR (Brünger, 1996). After an initial round of positional, restrained individual *B*-factor and torsion angle refinement, the four AcCoA molecules were placed in the model followed by rounds of positional, restrained individual *B*-factor and simulated annealing interspersed with inspection of $2F_o - F_c$ and $F_o - F_c$ maps and manual adjustment as necessary. In all calculations, Bijvoet pairs were kept separate and the anomalous scattering terms of selenium were included in the calculation of structure factors. Solvent molecules were built into the model on the basis of observable electron density and reasonable hydrogen bonding partners but only retained if they refined with a *B*-factor lower than 60 Å². In the final round of refinement, all data from 20.0-2.4 Å were included after applying a bulk solvent correction. Refinement of the structure of Hpa2 alone proceeded as for the Hpa2-AcCoA complex with the exception that it used data to 2.9 Å and no solvent molecules were included.

Protein Data Bank accession number

Coordinates for the Hpa2 structures described here have been deposited with the Protein Data Bank with accession codes 1QSM (Hpa2 + AcCoA) and 1QSO (Hpa2 alone).

Acknowledgments

We thank William Clemons, Robert M. Sweet and Lonny Berman for assistance with data collection. We also thank Adrian Hahn for construction of the Hpa2 expression clone, and Lisa Joss for invaluable assistance with equilibrium sedimentation data analysis. This work was supported by NIH grants GM 42796 (to V.R.) and GM 28220 and GM5641 (to R.S.). Beamlines X12C and X25 at the NSLS at Brookhaven are supported by the United States Department of Energy Offices of Health and Environmental Research and of Basic Sciences, and by the National Science Foundation.

References

- Abrahams, J. P. & Leslie, A. G. W. (1996). Methods used in the structure determination of bovine mitochondrial F₁ATPase. *Acta Crystallog. sect. D*, **52**, 30-42.
- Bannister, A. J. & Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature*, **384**, 641-643.
- Boyes, J. P. B., Nakatani, Y. & Ogryzko, V. (1998). Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature*, **396**, 594-598.
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y. & Allis, C. D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, **84**, 843-851.
- Brünger, A. T. (1996). *X-PLOR Version 3.843: A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
- Clarke, A. S., Lowell, J. E., Jacobson, S. J. & Pillus, L. (1999). Esa1p is an essential histone acetyltransferase required for cell cycle progression. *Mol. Cell Biol.* **19**, 2515-2526.
- Coleman, C. S., Huang, H. & Pegg, A. E. (1996). Structure and critical residues at the active site of spermidine/spermine-N1-acetyltransferase. *Biochem. J.* **316**, 697-701.
- Collaborative Computational Project Number 4. (1994). The CCP4 suite: Programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760-763.
- Dutnall, R. N., Tafrov, S. T., Sternglanz, R. & Ramakrishnan, V. (1998). Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily. *Cell*, **94**, 427-438.
- Engel, C. & Wierenga, R. (1996). The diverse world of coenzyme A binding proteins. *Curr. Opin. Struct. Biol.* **6**, 790-797.
- Gu, W. & Roeder, R. G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell*, **90**, 595-606.
- Hickman, A. B., Klein, D. C. & Dyda, F. (1999a). Melatonin biosynthesis: the structure of serotonin N-acetyltransferase at 2.5 Å resolution suggests a catalytic mechanism. *Mol. Cell*, **3**, 23-32.
- Hickman, A. B., Nambodiri, M. A., Klein, D. C. & Dyda, F. (1999b). The structural basis of ordered substrate binding by serotonin N-acetyltransferase: enzyme complex at 1.8 Å resolution with a bisubstrate analog. *Cell*, **97**, 361-369.
- Imhof, A., Yang, X. J., Ogryzko, V. V., Nakatani, Y., Wolffe, A. P. & Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. *Curr. Biol.* **7**, 689-692.
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110-119.
- Kleff, S., Andrulis, E. D., Anderson, C. W. & Sternglanz, R. (1995). Identification of a gene encoding a yeast histone H4 acetyltransferase. *J. Biol. Chem.* **270**, 24674-24677.
- Kuo, M. H. & Allis, C. D. (1998). Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays*, **20**, 615-625.
- Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E. & Allis, C. D. (1998). Histone acetyltransferase activity

- of yeast Gcn5p is required for the activation of target genes in vivo. *Genes Dev.* **12**, 627-639.
- Laue, T. M. (1995). Sedimentation equilibrium as thermodynamic tool. *Methods Enzymol.* **259**, 427-452.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution [see comments]. *Nature*, **389**, 251-260.
- Luzzatti, P. V. (1952). Traitement Statistique des Erreurs dans la Determation des Structures Cristallines. *Acta Crystallog.* **5**, 802-810.
- McRorie, D. K. & Voelker, P. J. (1993). *Self Associating Systems in the Analytical Ultracentrifuge*, Beckman Instruments, Inc., Palo Alto, CA.
- Modis, Y. & Wierenga, R. (1998). Two crystal structures of N-acetyltransferases reveal a new fold for CoA-dependent enzymes. *Structure*, **6**, 1345-1350.
- Neuwald, A. F. & 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.
- Neuwald, A. F., Liu, J. S., Lipman, D. J. & Lawrence, C. E. (1997). Extracting protein alignment models from the sequence database. *Nucl. Acids Res.* **25**, 1665-1677.
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*, **87**, 953-959.
- Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-325.
- Park, E. C. & Szostak, J. W. (1992). ARD1 and NAT1 proteins form a complex that has N-terminal acetyltransferase activity. *EMBO J.* **11**, 2087-2093.
- Parthun, M. R., Widom, J. & Gottschling, D. E. (1996). The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell*, **87**, 85-94.
- Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219-223.
- Smith, E. R., Eisen, A., Gu, W., Sattah, M., Pannuti, A., Zhou, J., Cook, R. G., Lucchesi, J. C. & Allis, C. D. (1998). ESA1 is a histone acetyltransferase that is essential for growth in yeast. *Proc. Natl Acad. Sci. USA*, **95**, 3561-3565.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 61-89.
- Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D. Biol. Crystallog.* **55**, 849-861.
- van Holde, K. E. (1989). *Chromatin*, Springer-Verlag, New York.
- Verreault, A., Kaufman, P. D., Kobayashi, R. & Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr. Biol.* **8**, 96-108.
- Waltzer, L. & Bienz, M. (1998). Drosophila CBP represses the transcription factor TCF to antagonize Wingless signalling. *Nature*, **395**, 521-525.
- Wolf, E., Vassilev, A., Makino, Y., Sali, A., Nakatani, Y. & Burley, S. K. (1998). Crystal structure of a GCN5-related N-acetyltransferase: *Serratia marcescens* aminoglycoside 3-N-acetyltransferase. *Cell*, **94**, 439-449.
- Wybenga-Groot, L. E., Draker, K., Wright, G. D. & Berghuis, A. M. (1999). Crystal structure of an aminoglycoside 6'-N-acetyltransferase: defining the GCN5-related N-acetyltransferase superfamily fold. *Structure*, **7**, 497-507.
- Yang, X. J., Ogryzko, V. V., Nishikawa, J., Howard, B. H. & Nakatani, Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature*, **382**, 319-324.

Edited by T. Richmond

(Received 28 June 1999; received in revised form 22 October 1999; accepted 26 October 1999)