HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2014 March 01.

Published in final edited form as:

Nat Struct Mol Biol. 2013 September ; 20(9): 1098–1105. doi:10.1038/nsmb.2636.

Molecular Basis for Amino-Terminal Acetylation by the Heterodimeric NatA Complex

Glen Liszczak[1,2], **Jacob M. Goldberg**[2], **Havard Foyn**[3], **E. James Petersson**[2], **Thomas Arnesen**[3], and **Ronen Marmorstein**[1,2]

[1]Program in Gene Expression and Regulation, The Wistar Institute, Philadelphia, PA 19104, USA

[2]Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA

[3]Department of Molecular Biology, University of Bergen, N-5020 Bergen, Norway

[4]Department of Surgery, Haukeland University Hospital, N-5021 Bergen, Norway

SUMMARY

Amino-terminal acetylation is ubiquitous among eukaryotic proteins and controls a myriad of biological processes. Of the N-terminal acetyltransferases (NATs) that facilitate this co-translational modification, the heterodimeric NatA complex harbors the most diversity for substrate selection and modifies the majority of all amino-terminally acetylated proteins. Here, we report the X-ray crystal structure of the 100 kDa holo-NatA complex from *Schizosaccharomyces pombe* in the absence and presence of a bisubstrate peptide-CoA conjugate inhibitor, as well as the structure of the uncomplexed Naa10p catalytic subunit. The NatA-Naa15p auxiliary subunit contains 13 TPR motifs and adopts a ring-like topology that wraps around the NatA-Naa10p subunit, an interaction that alters the Naa10p active site for substrate-specific acetylation. These studies have implications for understanding the mechanistic details of other NAT complexes and how regulatory subunits modulate the activity of the broader family of protein acetyltransferases.

INTRODUCTION

The co-translational process of amino-terminal acetylation occurs on 85% of human proteins and ~50% of yeast proteins and mediates a wide range of biological processes including cellular apoptosis, enzyme regulation, protein localization and the N-end rule for protein degradation[1–6]. At least three N-terminal acetyltransferase (NAT) complexes that perform

Correspondence should be addressed to Ronen Marmorstein, marmor@wistar.org.

Accession codes. The coordinates of the structures have been deposited in the PDB under accession numbers 4KVX (Naa10p·AcCoA), 4KVO (Naa10p·Naa15p·AcCoA) and 4KVM (Naa10p·Naa15p·bisubstrate inhibitor).

Author Contributions

G.L. performed all of the structural and biochemical experiments described in the manuscript and J.M.G. carried out inhibitor synthesis. G.L. prepared manuscript figures, text and movies; H.F. carried out preliminary inhibition studies that led to experiments reported in the manuscript; R.M. designed and supervised experiments by G.L. and prepared manuscript text. T.A. supervised the experiments of H.F. and prepared manuscript text. E.J.P. supervised the experiments of J.M.G. All authors read and approved the submitted manuscript.

The authors declare no conflict of interest.

this modification, NatA, NatB and NatC, exist as heterodimers with one unique catalytic subunit and an additional unique auxiliary subunit that both activates the enzymatic component and anchors the complex to the ribosome during translation[7–13]. Aberrant expression of the proteins that make up the NatA complex has been observed in a number of cancer cell tissues; consequently, NAT enzymes are emerging targets for chemotherapeutic development[14–21].

The three NAT complexes are highly conserved from yeast to humans and are differentiated from one another on the basis of their substrate specificities[1,5,22–24]. NatA, which is composed of the catalytic NAA10 subunit and the auxiliary NAA15 subunit, is the most promiscuous of all NAT enzymes; it will traditionally acetylate an α-amino group on nascent peptide chains with an amino-terminal alanine, cysteine, glycine, serine, threonine, or valine residue[1,22,24,25]. Interestingly, recent studies demonstrate that NAA10 also exists as a monomer in cells and that it can acetylate the α-amino group of substrates with amino-terminal aspartic acid and glutamic acid residues that can be generated post-translationally, but will not acetylate traditional NatA substrates[26]. The NatB and NatC complexes acetylate the N-terminus of proteins with an amino-terminal methionine with further specificity dependent upon the identity of the second residue[13,23,25,27]. Despite the fact that NATs as well as many other lysine side-chain acetyltransferases require binding partners for optimal catalytic activity, no acetyltransferase has been structurally characterized in the presence of its activating partner[28–30].

Three additional NAT enzymes, NatD-NatF, have been identified and appear to be independently active. They also have a very limited set of biologically relevant substrates and are not well characterized across eukaryotes[31–35]. Currently, there is no structure of a NAT complex or any of the integral subunits and the mechanism of substrate-specific catalytic regulation by the auxiliary subunit remains uncharacterized. The only eukaryotic NAT for which structural data is available is the human NatE (NAA50) enzyme, which is independently active and has only one known biologically relevant substrate that contains a Met1-Leu2-Gly3-Pro4- amino-terminal sequence that is present in the X-ray crystal structure[1,36].

In this study we set out to determine the molecular basis for substrate binding, acetylation specificity and mode of catalysis of the NatA complex, and to determine the role of the auxiliary subunit in these activities. Towards this goal, we determined the X-ray crystal structures of the holo-NatA complex bound to AcCoA and a bisubstrate inhibitor, determined the structure of the Naa10p subunit bound to AcCoA in its uncomplexed form that preferentially acetylates a unique subset of substrates, and carried out structure-guided mutational analysis to derive structure-function correlations underlying NatA acetylation.

RESULTS

Overall Structure of NatA

In an attempt to prepare the NatA complex for X-ray structure determination, we overexpressed the human complex in Sf9 infected insect cells and the orthologous proteins from *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* in insect cells and

bacteria, respectively. We found that coexpression of the *Schizosaccharomyces pombe* full-length Naa15p subunit (residues 1–729) with a C-terminal truncation construct of the Naa10p subunit (residues 1–156 out of 177 total residues) produced a soluble, active, heterodimeric complex that could be purified to homogeneity and crystallized for use in structural studies. Crystals of this NatA complex were formed in the presence of AcCoA in the P1 space group with 4 heterodimers in the asymmetric unit and the structure was determined using multiwavelength anomalous diffraction with a combination of a K_2PtBr_4 heavy atom soaked crystal and selenomethionine-derivatized protein (Table 1). The dataset was collected to 3.15 Å resolution and was refined to R_{work} and R_{free} values of 22.4% and 25.9%, respectively.

The structure of the NatA complex revealed that the auxiliary Naa15p subunit is composed of 37 α-helices ranging from 8 to 32 residues in length, among which 13 conserved helical bundle tetratricopeptide repeat (TPR) motifs can be identified, which are often employed for protein-protein interactions (Supplementary Fig. 1)[37]. While a number of these TPR-helices are involved in interaction with Naa10p, we suspect that the additional TPR motifs are important for interaction with other NatA binding partners such as the ribosome, NatE and the HYPK chaperone[8,38]. The Naa15p helices arrange themselves into a ring-like tertiary structure that wraps completely around the Naa10p subunit, burying a total of 3,520 $Å^2$ of the solvent accessible surface area of the enzyme (Fig. 1a,b **and** Supplementary Movie 1). This fold is reminiscent of the importin-β structure, which employs 19 tandem HEAT repeats to form a solenoid tertiary structure that surrounds its binding partner, importin-α[39]. Indeed, it appears that Naa15p stability depends on the binding of Naa10p. This is consistent with the observation that recombinant Naa15p undergoes significant degradation and precipitation when purified as a monomer.

The catalytic Naa10p subunit adopts a Gcn5-related N-acetyltransferase (GNAT) fold[40] containing a central AcCoA binding region and flanking N- and C- terminal segments that are similar to the structure of NAA50. The most intimate interactions between the two proteins are made between several Naa15p helices and the N-terminal α1-loop-α2 segment of Naa10p (Fig. 1c,d), where the structure has relatively low B-factors and is well ordered (Supplementary Fig. 2). This region features one large hydrophobic interface between residues that span the entire length of Naa10p-α2 (Leu28, Leu32, Tyr33, Ile36, Trp38) and several residues of Naa10p-α1 (Ile8 and Leu11) with residues in Naa15p-α29 (Trp526, Phe533, Phe536) and α30 (Leu549, Trp552); and hydrogen bonds between Naa10p-Lys29 and Naa10p-Gln15 with Naa15p-Asp532 and Naa15p-Gln491, respectively. Single point mutations in this region did not break up the complex and had only modest effects on substrate binding and catalysis (Table 2), likely due to the extensiveness of the interface. A smaller hydrophobic interface is formed between Naa10p-His20 with Naa15p-Phe449 and Naa15p-Trp494 and is supplemented with a hydrogen bond between Naa10p-Gln25 and Naa15p-Arg448. This region of Naa15p directly stabilizes the position of the Naa10p-α1 helix as well as the Naa10p-α1-α2 loop and as a result is crucial for proper complex formation. This is evident from the observation that alanine point mutations at either Naa15p-Arg448 or Naa15p-Phe449 were able to disrupt NatA complex formation. Several

additional scattered intermolecular interactions serve to supplement the Naa10p-Naa15p interface (Supplementary Fig. 3).

Molecular Basis for Naa15p Modulation of Naa10p Acetylation

To explore the molecular basis for Naa15p modulation of Naa10p acetyltransferase activity, we determined the X-ray crystal structure of Naa10p in the absence of Naa15p, for comparison with the holo-NatA complex (Table 1). The structure of Naa10p (residues 1–156) was determined to 2.00 Å resolution using a combination of single wavelength anomalous diffraction and molecular replacement (MR; model: NAA50) to phase data collected on a selenomethionine-derivatized Naa10p protein. An alignment of the complexed and uncomplexed forms of Naa10p revealed that the α1-loop-α2 segment assumes a substantially different conformation in the presence of Naa15p (Fig. 2a). Notably, this conformational change is driven by the movement of several hydrophobic residues in Naa10p-α2 (Leu28, Leu32, Ile36), which make intramolecular interactions with residues in Naa10p-α1 and β3 (Ile8, Leu11, Met14, Tyr55, Tyr57) in apo-Naa10p but shift to make alternative intermolecular interactions with helices of Naa15p in the NatA complex (Figs. 1c and 2b). As a result of this interaction, the C-terminal region of the α1 helix undergoes an additional helical turn, which helps to reposition the α1-α2 loop. Interestingly, docking of the apo-Naa10p structure into the corresponding binding pocket of Naa15p showed a clash between the Naa10p α1-loop-α2 and Naa15p-Arg448 and Naa15p-Phe449 of α25 (Fig. 2c), the same interface that we have shown to be necessary for proper complex formation (Fig. 1d and Table 2).

As a result of the Naa15p-interaction along one side of the Naa10p-α1-loop-α2 region, residues on the opposite side of this loop region appear to adopt a specific conformation that is essential for catalysis towards traditional substrates (alanine, cysteine, glycine, serine, threonine, or valine) (Supplementary Movies 2 and 3). Specifically, residues Naa10p-Leu22 and -Tyr26 shift about 5.0 Å from surface exposed positions to buried positions in the active site and Naa10p-Glu24 moves by about 4.0 Å, drastically altering the landscape of the NatA active site (Fig. 2d). All of these residues are well ordered in both structures (Supplementary Fig. 4). Our comparison of the complexed and uncomplexed structures suggests that the auxiliary subunit induces an allosteric change in the Naa10p active site to an extent that is required for the mechanism of catalysis employed by the NatA complex, and likely represents an active NAT fold. Consistent with this hypothesis, a backbone alignment of key active site elements in active Naa10p with the corresponding region in the independently active human NAA50 that selects a Met1-Leu2- amino-terminal sequence showed a high degree of structural conservation (RMSD of 1.52 Å). The corresponding alignment of the complexed and uncomplexed forms of Naa10p showed less structural conservation, with an RMSD of 2.43 Å (Fig. 2e).

Substrate Peptide Binding and NatA Inhibition

To determine the molecular basis for substrate-specific peptide-binding by NatA, and in particular how, unlike most other NATs, it is able to accommodate a number of non-methionine amino-terminal substrates, we synthesized a bisubstrate conjugate in which CoA is covalently linked to a biologically relevant substrate peptide fragment with the sequence

Ser1-Ala2-Ser3-Glu4-Ala5 (CoA-SASEA). We performed inhibition studies with CoA-SASEA and with a control compound, acetonyl-CoA, which is a non-hydrolyzable AcCoA analog (Fig. 3a). IC_{50} determinations revealed that CoA-SASEA had an IC_{50} of 1.4±1.0 μM, while acetonyl-CoA had an IC_{50} of 380±10 μM (Fig. 3b). To assess the specificity of this inhibitor towards NatA, we also calculated IC_{50} values of CoA-SASEA and acetonyl-CoA with NAA50, a NAT that requires a substrate amino-terminal methionine residue. We found that CoA-SASEA has an IC_{50} of 11±2 μM while acetonyl-CoA has an IC_{50} of 130±12 μM (Fig. 3b). With NAA50, the addition of a SASEA-peptide portion was only able to increase the potency of acetonyl-CoA by about 10-fold, whereas with NatA this peptide addition exhibited an increase in potency of about 300-fold. The greater potency of acetonyl-CoA with NAA50 over Naa10p can be explained by the fact that AcCoA binds more strongly to NAA50 (K_m=27±2 μM) than it does to Naa10p (K_m=59±5 μM). The markedly higher potency of the CoA-SASEA inhibitor relative to acetonyl-CoA towards NatA, along with the observed NatA specificity that the peptide portion imparts on the inhibitor, suggested that CoA-SASEA interacts with the complex in a biologically relevant conformation and can be used to understand the molecular basis for amino-terminal substrate binding by NatA.

We were able co-crystallize NatA bound to CoA-SASEA and determine its X-ray crystal structure to 2.60 Å resolution using the refined NatA·AcCoA structure as a search model (Table 1 and Fig. 4a). The structure of the bisubstrate inhibitor-bound NatA complex revealed that the N-terminal Ser1 side chain makes van der Waals interactions with Naa10p residues Leu22, Glu24 and Tyr139 (Fig. 4b). The Naa10p-Tyr139 side chain also forms a hydrogen bond with the Ser1 backbone carbonyl group, and Naa10p-Tyr26 forms a hydrogen bond with the Ala2 backbone carbonyl. Only modest rearrangements were observed in Naa10p upon CoA-SASEA binding as compared to AcCoA binding (alignment not shown).

We used *in vitro* activity assays to kinetically characterize the wild type Naa10p monomer enzyme as well as wild type NatA and a series of NatA mutants (Table 2). These assays were performed in the presence of a substrate peptide with an amino-terminal sequence identical to that of the inhibitor, a Met1- amino-terminal substrate peptide that corresponds to the NAA50 *in vivo* substrate, and a substrate peptide with an amino-terminal Glu1-Glu2-Glu3- sequence that matches the sequence for γ-actin, the only known NAA10 *in vivo* substrate. These studies revealed that while monomeric Naa10p is inactive with the Met1- and Ser1- amino-terminal substrates, we are able to generate catalytic parameters with the Glu1- amino-terminal substrate (k_{cat}=0.19 ± 0.02 sec^{-1}; K_m=1.72 ± 0.25 μM) (Table 2). While previous studies were able to detect acetylation of the amino-terminal Glu1- substrate by the NatA complex[26], we were unable to detect activity towards this substrate. In contrast, we were able to determine kinetic parameters for the Ser1- amino-terminal substrate in the presence of a saturating amount of AcCoA (k_{cat}=3.0±0.5 s^{-1} and K_m=340±50 μM). The NatA-Naa10p-L22A point mutation exhibited a >99% decrease in catalytic efficiency (k_{cat}·K_m-1), suggesting that steric alterations to the amino-terminal residue-binding pocket greatly hinder productive substrate binding and catalysis (Fig. 4c and Table 2). Similarly, the NatA-Naa10p-E24A mutant had no measurable activity. The k_{cat} of the NatA-Naa10p-E24Q was reduced by 99% while the K_m was unaffected, which emphasizes the importance

of Glu24 in both substrate binding and catalysis (Table 2). The NatA-Naa10p-Y26A mutant resulted in a >99% decrease in catalytic efficiency, while the catalytic efficiency for the NatA-Naa10p-Y139A mutant could not be calculated due to the immeasurably high K_m (Table 2), which is likely caused by the fact that Tyr139 is involved in both a backbone interaction and the Ser1 side chain-binding pocket. To evaluate the active site of the uncomplexed Naa10p, we docked the bisubstrate inhibitor into the monomeric structure that is inactive towards the amino-terminal Ser1- substrate. Consistent with the particular importance of Naa10p residues Leu22, Glu24, Tyr26 and Tyr139 in NatA catalysis of traditional substrates, each of these residues are not in a position for productive substrate binding and catalysis in uncomplexed Naa10p, with the exception of Tyr139 (Fig. 4d).

Although all of the important catalytic residues noted above for NatA catalysis of the Ser1-amino terminal substrate appear out of place for catalysis by the Naa10p monomer, it is able to catalyze acetylation of the Glu1-amino-terminal substrate, suggesting that the acetylation reaction by monomeric Naa10p is catalyzed through an alternate mechanism relative to that of the NatA complex. To investigate the possibility that the Glu1-amino-terminal substrate does not require the Glu24 catalytic residue for acetylation, we performed an assay using this substrate and the NatA-E24A mutant. This mutant, which we found will not acetylate the amino-terminal Ser1- or Met1- amino-terminal substrates used in this study, is able to acetylate the Glu1- amino-terminal substrate with a rate of 0.65±0.08 sec^{-1}, which is ~22% of the rate of the wild type NatA complex with its cognate Ser1-amino terminal substrate peptide (Table 2). We suspect that the more open amino-terminal residue-binding pocket is able to accommodate an amino-terminal Glu1- substrate peptide and that the carboxylic acid group of the glutamate side chain occupies a position that is very similar to the corresponding group in the wild type NatA-Glu24 side chain; consequently, it seems to be able to play an equivalent role in catalysis. The recovery of activity observed when using this substrate-mutant combination further supports the importance of the Glu24 carboxylic acid moiety in the NatA catalytic mechanism.

To determine the molecular basis for the inability of NatA to acetylate N-terminal methionine and glutamate residues, we modeled an amino-terminal glutamate residue and docked the human NAA50 substrate into the active Naa10p site (Fig. 4e). This superposition revealed that Naa10p-Glu24 would clash with the substrate peptide glutamate and methionine residues, suggesting that Naa10p-Glu24 is particularly important for excluding these amino-terminal substrates. This is consistent with the importance of the corresponding human Naa50-Val29 residue for N-terminal methionine recognition and acetylation[36]. We also modeled substrate amino-terminal threonine and valine residues, the most bulky NatA substrates, into the active site to show that, similar to Serine, these side chains can also be accommodated without steric clash in the wild type NatA substrate-binding pocket (Fig. 4e).

Catalysis by NatA

Previous structural and kinetic studies on human NAA50 support a dual base mechanism in which NAA50-Tyr73 and -His112, located in the central core region and C-terminal segment of the catalytic domain, contribute to proton abstraction from the substrate α amino group to facilitate the acetylation reaction [36]. An active site alignment of these enzymes

shows that Naa10p also has two potential general bases in corresponding positions, Naa10p-His72 and -His111 (Fig. 5a). However, point mutations of H72A (k_{cat} = 2.4±0.3 sec^{-1}) and H111A (k_{cat} = 2.0±0.2 sec^{-1}) each had a negligible effect on the rate of catalysis, arguing against the importance of these residues in catalysis by Naa10p (Table 2). Notably, NAA50 also has a water molecule in the active site that likely plays a catalytic role, which is not observed in the Naa10p active site and further supports the idea that NatA employs a different catalytic strategy[36].

The NatA structures reveal that Glu24, located in the loop between the α1 and α2 helices in the N-terminal segment of the Naa10p catalytic domain, is in close proximity to the amino-terminal residue and could be playing a role in catalysis (Fig. 5a). This is consistent with our observation that the E24Q mutant shows a 99% reduction in k_{cat} and has a negligible effect on K_m, which suggests that the carboxylic acid moiety is essential for catalysis, but not required for substrate binding (Table 2). In both Gcn5 and ESA1, mutating the general base glutamate to a glutamine also results in a decrease in k_{cat} by ~2 orders of magnitude, which supports the possibility that Naa10p-Glu24 of NatA could be acting as a general base in the NatA acetylation reaction[41,42], whereby the buried glutamate residue harbors an increase in side chain pK_a that allows it to directly deprotonate the substrate amino group. Alternatively, the glutamate residue could act through an electrostatic mechanism causing an induced-dipole on the α-amino group that would increase the nucleophilicity of the substrate, or could be involved in proper positioning of the α-amino group within the active site. While NatA-Naa10p-Glu24 functions as an essential catalytic residue, it is also possible that NatA-Naa10p-Tyr139, located in the C-terminal segment of the Naa10p catalytic domain, also contributes to catalysis due to its location in the active site and the dramatic effect on catalysis observed for the NatA-Naa10p-Y139A mutant. Notably, NatA-Naa10p-Arg113, also in this C-terminal segment, forms an ionic interaction with Glu24. An observed 94% reduction in catalytic efficiency for the Naa10p-R113A mutant suggests that this residue may also be important for proper positioning of the Naa10p-Glu24, altering the pK_a of the Glu24 side chain, or it may have some other function in the NatA catalytic mechanism or substrate binding. A comparison of the catalytic efficiencies of potential key catalytic residues is shown in Figure 4c. A sequence alignment of the Naa10p catalytic subunits from a diverse group of eukaryotic species reveals that all of the α1-α2 loop residues, as well as Arg113 and Tyr139, are strictly conserved, which reinforces the importance of these residues in catalysis (Fig. 5b).

To further dissect the reaction mechanism, we also calculated K_i values of the CoA-SASEA inhibitor as well as its mode of inhibition towards both AcCoA and the substrate peptide. Our results (Supplementary Fig. 5) reveal that the CoA-SASEA is a competitive inhibitor versus AcCoA (K_i=1.1±0.1 μM) and a noncompetitive inhibitor versus the substrate peptide (K_i=8.9±0.1 μM). This pattern of inhibition supports an ordered BiBi model for catalysis, with AcCoA binding occurring before the substrate peptide. This type of mechanism was also recently reported for NAA50 in a study that used NMR experiments to show that AcCoA binding induces rearrangements in the peptide-binding pocket and that the peptide shows a far greater affinity toward the NAA50·AcCoA complex over the apo enzyme[43].

DISCUSSION

The data presented here reveals that Naa15p allosterically reconfigures the Naa10p active site for sequence-specific amino-terminal acetylation. Specifically, this study highlights a unique mechanism in which NatA substrate binding specificity is coupled to the catalytic strategy that is employed. This finding is supported by evidence that an amino-terminal residue of an incompatible substrate, such as a methionine, would have to displace Glu24, an essential catalytic residue, for binding to occur (Fig. 4e). Furthermore, upon modeling all bulky, β-branched NatA cognate substrate amino-terminal residues into the binding pocket, we find that none of them clash with the Glu24 residue, which further supports the importance of this residue and its role in substrate-specific acetylation (Fig. 4e). We also note that in the absence of Glu24, the Glu1-amino-terminal substrate can undergo acetylation, likely due to the fact that the side chain carboxylic acid moiety is able to functionally replace the corresponding group on the Glu24 side chain. This could also explain the activity observed for the Naa10p monomer in the presence of this substrate[26].

Comparison of the Naa10p active site with the corresponding region of NAA50 and other HATs shows significant differences that appear to underlie the fact that Naa15p binding is required for Naa10p acetyltransferase activity towards conventional substrates. In particular, three catalytically essential Naa10p residues that are located in the α1-α2 loop that is N-terminal to the catalytic core region, Leu22, Glu24, and Tyr26, are not correctly positioned for catalysis in the uncomplexed form of the enzyme. Remarkably, in the context of the NatA complex, the Leu22 and Tyr26 side chains shift from surface exposed positions into the active site while the Glu24 side chain is repositioned by ~5 Å to facilitate the observed coupling of substrate binding specificity and catalysis. These changes in the Naa10p conformation are essential for productive catalysis of traditional substrates. In contrast, the independently active NAA50 and other HATs have their key catalytic residues located on the opposite side of the substrate, either within the central core region or the segment C-terminal to the catalytic core region (Fig. 5c). It appears that Naa10p contains its key residues for catalysis on the opposite side of the substrate relative to NAA50 and the other HATs to allow for catalytic regulation through Naa15p binding.

A sequence alignment of the Naa10p catalytic subunits from a diverse group of eukaryotic species reveals that all of the α1-α2 loop residues, as well as Arg113 and Tyr139, are strictly conserved, which reinforces the importance of these residues in catalysis (Fig. 5b). In addition, based on our findings of Naa15p regulation of Naa10p activity, we propose that the catalytic subunits of the NatB and NatC NAT complexes are modulated for acetylation in similar ways by their auxiliary subunits. This hypothesis is consistent with the strict conservation of the Naa10p-Leu22, -Glu24, and Tyr26 residues across catalytic subunits of each of these NAT complexes (Fig. 5d), which also reinforces the importance of this loop for enzymatic activity of NAT enzymes. However, NatB and NatC both acetylate methionine containing amino-terminal substrates, so these complexes likely use other novel strategies to accommodate their respective substrates. Many protein lysine side-chain acetyltransferases, including HAT1[30], HBO1[29] and Sas2[28], also require auxiliary subunits for activated acetylation and it is of significant interest to determine if they also share similar strategies for catalytic regulation with heterodimeric NAT proteins.

The unique features of the NatA complex could be exploited for the development of NatA-specific inhibitors. Indeed, to our knowledge, the CoA-SASEA NatA inhibitor that we have prepared is the most potent NAT inhibitor that has been reported to date. This inhibitor may serve as a suitable probe for NatA inhibition in vitro, and although peptide-based inhibitors are generally not suitable inhibitors for cell-based studies due to their poor ability to penetrate cell membranes, the CoA-SASEA inhibitor may nonetheless serve as a starting point for the development of small molecule mimics with better pharmacological properties. Furthermore, our inhibition studies suggest that inhibitors targeting the peptide-binding pocket have the potential to specifically inhibit NatA relative to other NATs while AcCoA analogues will likely exhibit more promiscuous inhibition patterns. Given the observation that NatA subunits are often overexpressed in several different cancer tissues, NatA inhibitors may have therapeutic application[14,18].

Materials and Methods

Naa10p expression and purification

The full length Naa10p gene (encoding residues 1–177) was cloned out of the *Schizosaccharomyces pombe* genome (ATCC) and several C-terminal truncation constructs were engineered into a modified pETDUET vector containing a TEV protease-cleavable 6XHis-tag. All constructs were transformed into Rosetta (DE3)pLysS competent *E. coli* cells, grown to an OD_{600}=0.7 and induced with 0.5 mM IPTG at 16°C for ~16 hours. Cells were isolated by centrifugation and lysed by sonication in a buffer containing 25 mM Tris (pH 8.0), 500 mM NaCl, 10 mM β-ME and a complete, EDTA free protease inhibitor tablet (Roche). The solution was isolated and passed over Ni-resin (Thermo Scientific), which was subsequently washed with >20 column volumes of lysis buffer supplemented with 25 mM imidazole. The protein was eluted in the same buffer with an imidazole gradient (25–500 mM imidazole) and TEV protease was added to fractions containing the target protein for the duration of a 14-hour dialysis into lysis buffer. The solution was passed through an additional nickel column to remove TEV protease as well as any uncut Naa10p. The resin was then washed with ~ 7 column volumes of lysis buffer supplemented with 25 mM imidazole, which was pooled with the initial flow through. This solution was dialyzed into a buffer containing 25mM sodium citrate monobasic (pH 5.5), 100 mM NaCl and 2 mM DTT and loaded onto a 5 mL HiTrap SP ion exchange column (GE Healthcare). The protein was eluted in the same buffer using a salt gradient (100–750 mM NaCl). Peak fractions were concentrated to 8.0 $mg \cdot mL^{-1}$ (10kDa concentrator; Amicon Ultra, Millipore) as measured by UV_{280} absorbance (extinction coefficient = 18,000), for crystallization trials. The selenomethionine derivative was prepared by expressing Naa10p (residues 1–156) in minimal medium (Molecular Dimensions) containing 50 $\mu g \cdot L^{-1}$ of selenomethionine (Sigma), which was purified using the scheme described above.

Naa10p·Naa15p complex expression and purification

The full length Naa15p gene (encoding residues 1–729) was cloned out of the *Schizosaccharomyces pombe* genome and engineered into MCSI of a modified pETDUET vector containing a TEV protease-cleavable 6XHis-tag. DNA encoding residues 1–156 of Naa10p was engineered into MCSII of the corresponding vector. Cells were grown and the

binary protein complex was prepared essentially as described for Naa10p, except that the first of two peaks that eluted from the HiTrap SP ion exchange column was concentrated to 1mL, and loaded onto an s200prep gel filtration column (GE Healthcare) in a buffer containing 25 mM HEPES (pH 7.0), 200 mM NaCl and 1 mM TCEP. This protein was concentrated to 15 $mg \cdot mL^{-1}$ as measured by UV_{280} (extinction coefficient = 100,000) for crystallization trials. The selenomethionine labeled complex was obtained in as described above. Point-mutation proteins were generated using the Stratagene QuickChange protocol and protein harboring point mutations was expressed and purified using the protocols described above[44].

Naa10p crystallization and structure determination

Naa10p at a concentration of 8 $mg \cdot mL^{-1}$ was mixed with AcCoA at a 1:3 molar ratio. Initial crystallization hits were obtained from an Naa10p construct containing residues 1–161 (Naa10p (1–161)) using hanging drop vapor diffusion at 20°C with a well solution containing 16% PEG 8000 (Hampton Research), 10% ethylene glycol and 0.1M HEPES (pH 7.5) and yielded poor diffraction. These crystals were used to seed entire grid screens with Naa10p(1–156), which did not crystallize in the absence of the seed crystals. Large crystals of Naa10p(1–156) were obtained from this screen in a 1:1 mixture of protein and well solution (14% PEG 3350 and 0.1M bis-tris (pH 6.5)) using hanging drop vapor diffusion at 20°C. An additive screen (Hampton Research) revealed that these crystals could be reproduced in the absence of seed crystals when 10% glycerol was present as an additive in the crystallization drop. Final crystallization conditions of the Naa10p protein (construct 1–156) were optimized and include a drop with a 2.5:0.5:2.0 ratio of protein (8 $mg \cdot mL^{-1}$): glycerol (50%): well solution, mixed in that order. Selenomethionine crystals of this construct were obtained using identical conditions with a lower protein concentration (6.5 $mg \cdot mL^{-1}$). Diffraction quality crystals required 1–3 days to grow to maximum dimensions.

Datasets were collected at beamline X29A at the National Synchrotron Light Source (Brookhaven National Laboratory) processed using HKL2000[45]. Initial molecular replacement trials with the program Phaser were performed using several truncated variants of the hNAA50 structure (PDB ID: 3TFY) as a model, but were unable to provide electron density maps suitable for model building[46]. A dataset was then collected on the selenomethionine-labeled protein at the selenium peak wavelength (0.9795Å) but selenium sites could not be identified using single wavelength anomalous diffraction (SAD) or single isomorphous replacement with anomalous scattering (SIRAS). The selenium sites were located by performing molecular replacement (as described above) on the selenomethionine derivative dataset, contouring the output electron density map to 3.0σ, and placing selenium atoms at the resulting peaks in Coot [47]. These sites were input into Phaser along with the hNAA50 model, and an MR-SAD approach yielded high quality electron density maps. Initial model building was carried out using ARP/wARP and refinement was carried out using Phenix.refine (experimental phases excluded throughout refinement) while applying 2-fold NCS[48,49]. TLS was used in later stages of refinement (TLS groups defined using the open access server)[50,51]. Manual refinement and additional model building were performed using the program Coot. The selenomethionine containing crystals diffracted to a higher resolution (2.00 Å) than the wild type crystals (2.60 Å) and so the structure used in all

analyses was that of the SeMet-derivatized protein. The final model was checked for errors using a composite omit map generated by AutoBuild in the Phenix suite of programs[49].

Naa10p·Naa15p crystallization and structure determination

The Naa10p (1–156)·Naa15p (full length) complex was concentrated to 12 mg·mL^{-1} and mixed with 3X cofactor (AcCoA or bisubstrate inhibitor) and allowed to stand for 12–15 hours at 4°C. The initial crystallization hit was obtained using hanging drop vapor diffusion at 20°C in a well solution containing 10% PEG 4000, 10% isopropanol and 0.1M HEPES (pH 7.0) (all reagents from Qiagen). The best crystals were obtained by mixing protein:well at a 2.0 μL:2.0 μL ratio and grew to maximum dimensions in 7–10 days. These crystals were transferred into mother liquor supplemented with ethylene glycol in 5% increments (15%, 20%, 25%) for 1 minute at each concentration and finally flash-frozen in liquid nitrogen. The selenomethionine-labeled protein required that the protein buffer be supplemented with 1mM reduced L-Glutathione and 1mM oxidized L-Glutathione for crystallization to occur in conditions identical to those described for the native Naa10p·Naa15p complex. The crystals were cryoprotected in the stepwise manner described above. To derivatize native crystals with Pt (II), the crystals were transferred into mother liquor supplemented with 1mM K_2PtBr_4 for 20 minutes before cryoprotection using the method described for the native Naa10p·Naa15p complex.

Datasets were collected at beamlines X25 and X29A at the National Synchrotron Light Source (Brookhaven National Laboratory) and processed using HKL2000. For phasing, a total of 1,000 degrees (1° per frame) were collected on a single derivatized crystal (500 frames at the peak wavelength, and 500 frames at the inflection wavelength) to reach desirable completeness and redundancy in the P1 space group, and datasets were processed to 3.50 Å resolution. The two wavelength MAD data was used by the Hybrid Substructure Search in Phenix to identify a total of 8 heavy atom sites in the asymmetric unit and initial phases were obtained using SOLVE[49,52]. Density modification was performed in Resolve by applying 4-fold NCS (corresponding to each heterodimer) and using histogram matching and solvent flattening[53]. Initially, the Naa15p helices were built as polyalanine chains and the Naa10p molecules were manually placed into the density-modified map using the program Coot. The resulting structure was used as a molecular replacement model for the selenomethionine derivative dataset that was collected at the selenium peak wavelength and processed to 3.35 Å resolution. A total of 80 selenium sites were located manually in the asymmetric unit by contouring the map to 3.0σ and placing selenium atoms in the peaks in Coot. These sites were used to place methionine residues in Naa15p, and manual model building was completed in coot using the methionine residues as a guide. The final structure was used as a molecular replacement model for the high-resolution datasets (500° of data each) of the Naa10p·Naa15p·AcCoA (3.10 Å resolution) and the Naa10p·Naa15p·bisubstrate inhibitor (2.60 Å resolution) complexes. Refinement of the binary complex structures was carried out in Phenix and 4-fold NCS restraints were applied to two separate groups, the full-length Naa10p molecules, and residues 100–731 of Naa15p. The positioning of the first 5 N-terminal helices in the Naa15p subunit is slightly different in each heterodimer, so NCS could not be applied to this region. TLS was also used in later stages of refinement (TLS groups defined using the open access server). The backbone for the α20-α21 loop from

residues 356–377 can be traced, however it is largely disordered (average B-factor=196 $Å^2$) as it likely samples numerous conformations in the crystal. No density was observed for a loop from residues 633–641 and so this region was omitted from the final structure. Naa15p molecules were built starting between residues 1 and 5 depending on the chain and dataset. The final models were checked for errors using a simulated annealing composite omit map generated by AutoBuild in the Phenix suite of programs. All surface area and distance calculations, as well as three-dimensional alignment RMSDs and graphics were generated in PyMol (Delano Scientific LLC).

Acetyltransferase Assays

Naa10p·Naa15p acetyltransferase assays were carried out with 100 nM complex at 25°C for 3–120 min. (3 min for wild type protein and longer for mutant complexes that were catalytically deficient) in 100 mM Tris (pH 8.0) and 50 mM NaCl unless stated otherwise. A saturating amount of AcCoA (500 μM) was used in all enzymatic reactions, and the substrate peptide (NH_2-SASEAGVRWGRPVGRRRRP-COOH; GenScript) concentration was varied to determine steady-state catalytic parameters. The first seven residues of this peptide correspond to the amino-terminal sequence of a protein that has been shown to be highly amino-terminally acetylated in yeast while the C-terminal 12 residues form a basic tag that make the substrate compatible with our assay[1]. The Glu1- amino-terminal substrate peptide (NH_2- EEEIAALRWGRPVGRRRRP-COOH; Genscript) was designed with the same rationale, except that the first seven residues are from the Naa10p *in vivo* substrate, γ-actin. The Met1- amino-terminal substrate peptide (NH_2- MLGPEGGRWGRPVGRRRRP-COOH; Genscript) also follows the design rationale described above, with the first seven residues corresponding to the *in vivo* substrate, hnRNP F. In the assay, radiolabeled [^{14C}]acetyl-CoA (4 $mCi{\cdot}mmol^{-1}$; PerkinElmer Life Sciences) was mixed with the substrate peptide and allowed to incubate with enzyme in a 25 μL reaction volume at 25°C. To quench the reaction, 20 μL of the reaction mixture was added to P81 paper discs (Whatman), and the paper discs were immediately placed in wash buffer. Washes were then carried out three times in 10 mM HEPES (pH 7.5), with each wash lasting 5 min, to remove unreacted AcCoA. The papers were then dried with acetone and added to 4 ml of scintillation fluid, and the signal was measured using a Packard Tri-Carb 1500 liquid scintillation analyzer. Background control reactions were performed in the absence of enzyme. Reactions were also performed in the absence of the substrate peptide to ensure that any possible signal due to autoacetylation was negligible.

For the uncomplexed Naa10p subunit, reactions with the Ser1- amino-terminal peptide were carried out for 1 hour using 500 μM AcCoA, substrate peptide (100–1,500 μM) and up to 1 μM untagged enzyme. No activity could be detected. Reactions with the Glu1- amino-terminal peptide were performed at 25°C for 30 minutes in the presence of 200 nM enzyme and 500 μM AcCoA at various peptide concentrations (50–1,000 μM). For the NatA complex, substrate peptide K_m and V_{max} parameters were derived by titrating the substrate peptide at eight different concentrations ranging from 100 to 1,500 μm in the presence of 500 μM AcCoA. To obtain kinetic data for the point mutants L22A, Y26A, and Y139A, 500 nM of enzyme was used in the reaction. Additionally, the AcCoA K_m values for the wild type protein as well as the L22A mutant were determined by titrating the AcCoA at 8

different concentrations (ranging from 20–300 μM) in the presence of 1.5 mM substrate peptide. All reactions were performed in triplicate. All radioactive count values were converted to molar units using a standard curve created with known concentrations of radioactive acetyl-CoA added to scintillation fluid. The program GraphPad Prism version 5.01 was used for all data fitting to the Michaelis-Menten equation.

Inhibition assays were carried out in a similar manner as described above. These reactions were performed using 100nM wild type protein for 3 minutes in the buffer described for the previous assays. IC_{50} values for Naa10p were calculated by titrating inhibitor (CoA-SASEA: 0.01–120 μM and acetonyl-CoA: 5–10,000 μM) into reaction mixtures containing fixed concentrations of substrate peptide (500 μM) and AcCoA (200 μM). Similarly, IC_{50} values for NAA50 (expressed and purified as previously described[36]) were calculated by titrating inhibitor (CoA-SASEA: 0.01–240 μM and acetonyl-CoA: 5–10,000 μM) into reaction mixtures containing fixed concentrations of substrate peptide (1,200 μM) and AcCoA (200 μM). To clarify, peptide concentrations were held constant at K_m. To calculate the K_i of the CoA-SASEA inhibitor versus AcCoA, substrate peptide concentration was held constant at 1,200 μM and kinetic curves were generated using AcCoA concentrations ranging from 15–500 μM in the presence of inhibitor concentrations ranging from 0–15 μM. To calculate the K_i of the CoA-SASEA inhibitor versus substrate peptide, AcCoA concentration was held constant at 500 μM and kinetic curves were generated using peptide concentrations ranging from 200–1,000 μM in the presence of inhibitor concentrations ranging from 0–30 μM. All reactions described here were performed in duplicate and data were fit to a sigmoidal dose-response curve using GraphPad Prism version 5.01.

Bisubstrate inhibitor synthesis

The peptide SASEA was synthesized on a 100 μmol scale using standard Fmoc solid-phase synthesis techniques. Rink amide resin (100–200 mesh), Fmoc-Ala-OH, Fmoc-Glu(tBu)-OH, Fmoc-Ser(tBu)-OH, and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Novabiochem (currently EMD Millipore; Billerica, MA). All other reagents, including Sigmacote ®, were purchased from Sigma-Aldrich (St. Louis, MO). A clean, oven-dried glass reaction vessel (RV) was treated with Sigmacote and rinsed with dimethylformamide (DMF) prior to use. Rink amide resin (151 mg; substitution 0.66 $mmol \cdot g^{-1}$) was transferred to the RV and stirred with 6 mL DMF twice for 15 min. After each incubation, DMF was removed with vacuum suction. The beads were incubated with 20% piperidine in DMF (5 mL) for 20 min with magnetic stirring to remove the Fmoc protecting group from the resin. The spent solution was removed with vacuum suction and the beads were rinsed thoroughly with DMF. Excess DMF was removed with vacuum suction and the beads were stirred with a solution of Fmoc-Ala-OH and HBTU in DMF (5 equiv; 83 mM, 6 mL) and diisopropylethylamine (10 equiv; 175 μL). After 30 min, the spent solution was drained from the RV and the beads were washed extensively with DMF. Excess DMF was removed with vacuum suction. Subsequent amino acids were added iteratively using analogous Fmoc-deprotection and coupling procedures. After removal of the final N-terminal Fmoc-protecting group, bromoacetic acid was coupled to the peptide as an active *O*-acylisourea using carbodiimide chemistry. Bromoacetic acid (10 equiv; 139 mg, 1 mmol) was pre-activated with *N*,*N*′-diisopropylcarbodiimide (20 equiv;

309 μL, 2 mmol) in 2 mL dry CH_2Cl_2 for 20 min with stirring. The solvent was then removed under reduced pressure and the residue was transferred to the RV as a solution in 6 mL DMF. The resulting slurry was allowed to react for 1 h with constant stirring. The resin beads were then thoroughly washed with DMF and CH_2Cl_2. Residual solvent was removed with vacuum suction and the beads were dried under vacuum for 30 min. A cocktail of trifluoroacetic acid (TFA)-triisopropylsilane-water (38:1:1 volume-to-volume) was used for global deprotection and resin cleavage. The resin beads were incubated with fresh 6 mL portions of this solution for 60 min and then 30 min on a rotisserie. The cleavage solution was drained from the RV under N_2 flow after each incubation. The solutions were combined and concentrated under reduced pressure. The residue was dissolved in a $H_2O{\cdot}CH_3CN$ solution (3:1 volume-to-volume) and lyophilized. The crude, lyophilized powder was dissolved in 3 mL triethylammonium bicarbonate buffer, pH 8.00. A 1 mL portion of this solution was mixed with coenzyme A hydrate (25 mg, 33 μmol) and the resulting solution was allowed to react for 4 h at room temperature and then overnight at 4 °C. The peptide conjugate was purified by reversed-phase high-performance liquid chromatography (HPLC) using a Varian ProStar HPLC instrument outfitted with a diode array detector (currently Agilent Technologies; Santa Clara, CA) and Vydac 218TP C18 analytical and prep columns (Grace-Vydac; Deerfield, IL) with a linear solvent gradient. The aqueous phase was 0.1% TFA in H_2O and the organic phase was 0.1% TFA in CH_3CN. The gradient remained isocratic at 98% aqueous phase for 5 min, before ramping to 81% aqueous phase over 16 min, then to 0% aqueous phase over 6 min, then returning to 98% aqueous phase during a 9 min wash out period. The peptide conjugate eluted around 15 min with this method. Purified material was lyophilized. Matrix-assisted laser desorption ionization (MALDI) mass spectra were collected with a Bruker Microflex LRF MALDI-TOF mass spectrometer to confirm identity. MALDI-MS(-) m/z clacd for $C_{40}H_{65}N_{13}O_{26}P_3S^-$ $[M-H]^-$ 1268.31, found 1268.32.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by US National Institutes of Health (NIH) grant GM060293 (R.M.) and NIH training grant GM071339 (G.L.). We acknowledge the use of the Wistar Proteomics Core facility for the work reported here, which is supported in part by NIH grant CA010815. T.A. was supported by the Research Council of Norway and the Norwegian Cancer Society. We also acknowledge Marmorstein Lab members and E. Skordalakes for helpful discussions.

References

1. Arnesen T, et al. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. Proc Natl Acad Sci U S A. 2009; 106:8157–62. [PubMed: 19420222]
2. Forte GM, Pool MR, Stirling CJ. N-terminal acetylation inhibits protein targeting to the endoplasmic reticulum. PLoS Biol. 2011; 9:e1001073. [PubMed: 21655302]
3. Hwang CS, Shemorry A, Varshavsky A. N-terminal acetylation of cellular proteins creates specific degradation signals. Science. 2010; 327:973–7. [PubMed: 20110468]

4. Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA. N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. Science. 2011; 334:674–8. [PubMed: 21940857]
5. Starheim KK, Gevaert K, Arnesen T. Protein N-terminal acetyltransferases: when the start matters. Trends Biochem Sci. 2012; 37:152–61. [PubMed: 22405572]
6. Yi CH, et al. Metabolic regulation of protein N-alpha-acetylation by Bcl-xL promotes cell survival. Cell. 2011; 146:607–20. [PubMed: 21854985]
7. Arnesen T, et al. Identification and characterization of the human ARD1-NATH protein acetyltransferase complex. Biochem J. 2005; 386:433–43. [PubMed: 15496142]
8. Gautschi M, et al. The yeast N(alpha)-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides. Mol Cell Biol. 2003; 23:7403–14. [PubMed: 14517307]
9. Mullen JR, et al. Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. EMBO J. 1989; 8:2067–75. [PubMed: 2551674]
10. Park EC, Szostak JW. ARD1 and NAT1 proteins form a complex that has N-terminal acetyltransferase activity. EMBO J. 1992; 11:2087–93. [PubMed: 1600941]
11. Polevoda B, Brown S, Cardillo TS, Rigby S, Sherman F. Yeast N(alpha)-terminal acetyltransferases are associated with ribosomes. J Cell Biochem. 2008; 103:492–508. [PubMed: 17541948]
12. Starheim KK, et al. Identification of the human N(alpha)-acetyltransferase complex B (hNatB): a complex important for cell-cycle progression. Biochem J. 2008; 415:325–31. [PubMed: 18570629]
13. Starheim KK, et al. Knockdown of human N alpha-terminal acetyltransferase complex C leads to p53-dependent apoptosis and aberrant human Arl8b localization. Mol Cell Biol. 2009; 29:3569–81. [PubMed: 19398576]
14. Arnesen T, Thompson PR, Varhaug JE, Lillehaug JR. The protein acetyltransferase ARD1: a novel cancer drug target? Curr Cancer Drug Targets. 2008; 8:545–53. [PubMed: 18991565]
15. Arnesen T, et al. Induction of apoptosis in human cells by RNAi-mediated knockdown of hARD1 and NATH, components of the protein N-alpha-acetyltransferase complex. Oncogene. 2006; 25:4350–60. [PubMed: 16518407]
16. Lim JH, Park JW, Chun YS. Human arrest defective 1 acetylates and activates beta-catenin, promoting lung cancer cell proliferation. Cancer Res. 2006; 66:10677–82. [PubMed: 17108104]
17. Ren T, et al. Generation of novel monoclonal antibodies and their application for detecting ARD1 expression in colorectal cancer. Cancer Lett. 2008; 264:83–92. [PubMed: 18325661]
18. Kalvik TV, Arnesen T. Protein N-terminal acetyltransferases in cancer. Oncogene. 2013; 32:269–76. [PubMed: 22391571]
19. Yu M, et al. Correlation of expression of human arrest-defective-1 (hARD1) protein with breast cancer. Cancer Invest. 2009; 27:978–83. [PubMed: 19909012]
20. Fluge O, Bruland O, Akslen LA, Varhaug JE, Lillehaug JR. NATH, a novel gene overexpressed in papillary thyroid carcinomas. Oncogene. 2002; 21:5056–68. [PubMed: 12140756]
21. Yu M, et al. Immunohistochemical analysis of human arrest-defective-1 expressed in cancers in vivo. Oncol Rep. 2009; 21:909–15. [PubMed: 19287988]
22. Polevoda B, Sherman F. Composition and function of the eukaryotic N-terminal acetyltransferase subunits. Biochem Biophys Res Commun. 2003; 308:1–11. [PubMed: 12890471]
23. Polevoda B, Sherman F. N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. J Mol Biol. 2003; 325:595–622. [PubMed: 12507466]
24. Starheim KK, Gromyko D, Velde R, Varhaug JE, Arnesen T. Composition and biological significance of the human Nalpha-terminal acetyltransferases. BMC Proceedings. 2009; 3 (Suppl 6):S3. [PubMed: 19660096]
25. Polevoda B, Norbeck J, Takakura H, Blomberg A, Sherman F. Identification and specificities of N-terminal acetyltransferases from Saccharomyces cerevisiae. EMBO J. 1999; 18:6155–68. [PubMed: 10545125]

26. Van Damme P, et al. Proteome-derived peptide libraries allow detailed analysis of the substrate specificities of N(alpha)-acetyltransferases and point to hNaa10p as the post-translational actin N(alpha)-acetyltransferase. Mol Cell Proteomics. 2011; 10:M110 004580.
27. Van Damme P, et al. N-terminal acetylome analyses and functional insights of the N-terminal acetyltransferase NatB. Proc Natl Acad Sci U S A. 2012; 109:12449–54. [PubMed: 22814378]
28. Sutton A, et al. Sas4 and Sas5 are required for the histone acetyltransferase activity of Sas2 in the SAS complex. J Biol Chem. 2003; 278:16887–92. [PubMed: 12626510]
29. Iizuka M, Stillman B. Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. J Biol Chem. 1999; 274:23027–34. [PubMed: 10438470]
30. Parthun MR, Widom J, Gottschling DE. The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. Cell. 1996; 87:85–94. [PubMed: 8858151]
31. Polevoda B, Hoskins J, Sherman F. Properties of Nat4, an N(alpha)-acetyltransferase of Saccharomyces cerevisiae that modifies N termini of histones H2A and H4. Mol Cell Biol. 2009; 29:2913–24. [PubMed: 19332560]
32. Hole K, et al. The human N-alpha-acetyltransferase 40 (hNaa40p/hNatD) is conserved from yeast and N-terminally acetylates histones H2A and H4. PloS one. 2011; 6:e24713. [PubMed: 21935442]
33. Van Damme P, et al. NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. PLoS Genet. 2011; 7:e1002169. [PubMed: 21750686]
34. Song OK, Wang X, Waterborg JH, Sternglanz R. An Nalpha-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A. J Biol Chem. 2003; 278:38109–12. [PubMed: 12915400]
35. Evjenth R, et al. Human Naa50p (Nat5/San) displays both protein N alpha- and N epsilon-acetyltransferase activity. J Biol Chem. 2009; 284:31122–9. [PubMed: 19744929]
36. Liszczak G, Arnesen T, Marmorstein R. Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. J Biol Chem. 2011; 286:37002–10. [PubMed: 21900231]
37. Blatch GL, Lassle M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays. 1999; 21:932–9. [PubMed: 10517866]
38. Arnesen T, et al. The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of Huntingtin aggregation. Mol Cell Biol. 2010; 30:1898–909. [PubMed: 20154145]
39. Cingolani G, Petosa C, Weis K, Muller CW. Structure of importin-beta bound to the IBB domain of importin-alpha. Nature. 1999; 399:221–9. [PubMed: 10353244]
40. Vetting MW, et al. Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys. 2005; 433:212–26. [PubMed: 15581578]
41. Tanner KG, et al. Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. J Biol Chem. 1999; 274:18157–60. [PubMed: 10373413]
42. Berndsen CE, Albaugh BN, Tan S, Denu JM. Catalytic mechanism of a MYST family histone acetyltransferase. Biochemistry. 2007; 46:623–9. [PubMed: 17223684]
43. Evjenth RH, et al. Human protein N-terminal acetyltransferase hNaa50p (hNAT5/hSAN) follows ordered sequential catalytic mechanism: combined kinetic and NMR study. J Biol Chem. 2012; 287:10081–8. [PubMed: 22311970]
44. Braman J, Papworth C, Greener A. Site-directed mutagenesis using double-stranded plasmid DNA templates. Methods Mol Biol. 1996; 57:31–44. [PubMed: 8849992]
45. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
46. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
47. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–32. [PubMed: 15572765]

48. Langer G, Cohen SX, Lamzin VS, Perrakis A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nature Protocols. 2008; 3:1171–9. [PubMed: 18600222]

49. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–21. [PubMed: 20124702]

50. Painter J, Merritt EA. A molecular viewer for the analysis of TLS rigid-body motion in macromolecules. Acta Crystallogr D Biol Crystallogr. 2005; 61:465–71. [PubMed: 15809496]

51. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D Biol Crystallogr. 2006; 62:439–50. [PubMed: 16552146]

52. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr. 1999; 55:849–61. [PubMed: 10089316]

53. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr. 2000; 56:965–72. [PubMed: 10944333]

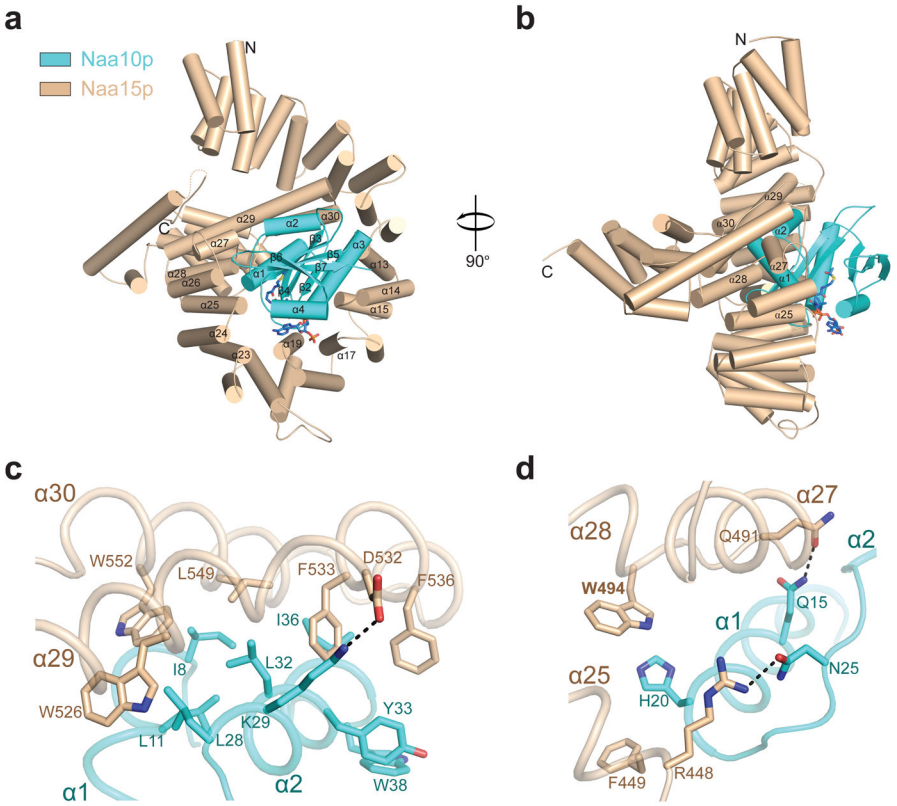

Figure 1.

Overall structure of the NatA complex bound to AcCoA. (**a**) The Naa10p (teal) and Naa15p (brown) subunits are shown in cartoon bound to AcCoA (CPK coloring and stick format). Only Naa15p helices that contact Naa10p are labeled. The dotted brown line represents a disordered loop region in Naa15p. The dimensions of the complex are 107 Å × 85 Å × 70 Å. (**b**) A 90° rotation of the view in (**a**). (**c**) A zoom view highlighting key residues that composes the predominantly hydrophobic interface between Naa10p-α1-α2 and Naa15p-α29-α30. (**d**) A zoom view of the intersubunit interface at the C-terminal region of Naa10p-α1 and the Naa15p-α25-α27-α28 helices.

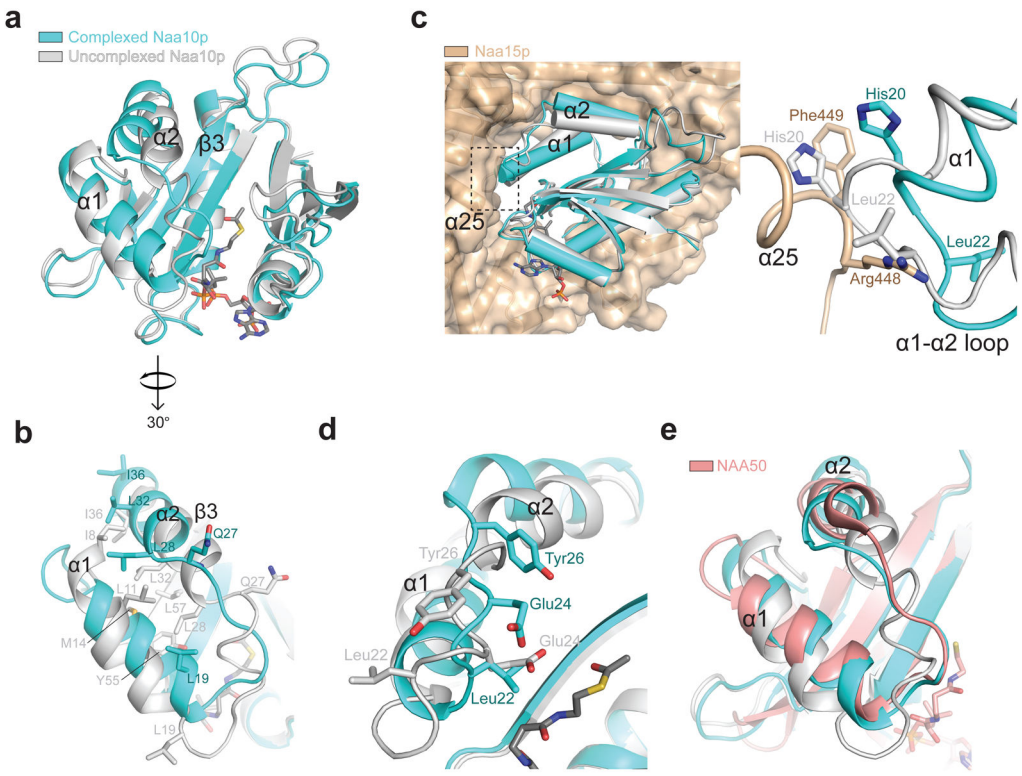

Figure 2.

Structure of the Naa10p monomer bound to AcCoA. (**a**) A structural alignment of the uncomplexed Naa10p monomer (silver cartoon) bound to AcCoA (gray stick format) with the active Naa10p from the NatA structure (teal cartoon). (**b**) A zoom view of the surface exposed Naa10p-α1-α2 region from the alignment shown in (**a**) that highlights residues in these two helices that undergo a conformational shift upon Naa15p association. (**c**) The alignment from (**a**) docked into the Naa15p subunit. The stretch of the α1- α2 loop in the Naa10p monomer that clashes with Naa15p is highlighted with a dashed box and a zoom view of this region is also shown. Secondary structural elements in the zoom view are shown in cartoon format and have been labeled. Key residue side chains are shown in stick format and have been labeled. (**d**) A 70° clockwise rotation of (**b**). The repositioning of the Naa10p-α1-loop-α2 upon complex formation and accompanying altered active site landscape are shown. Key repositioned residues are shown in stick format. (**e**) A structural alignment of the complexed and uncomplexed forms of Naa10p with hNAA50 (pink) bound to CoA (gray stick format).

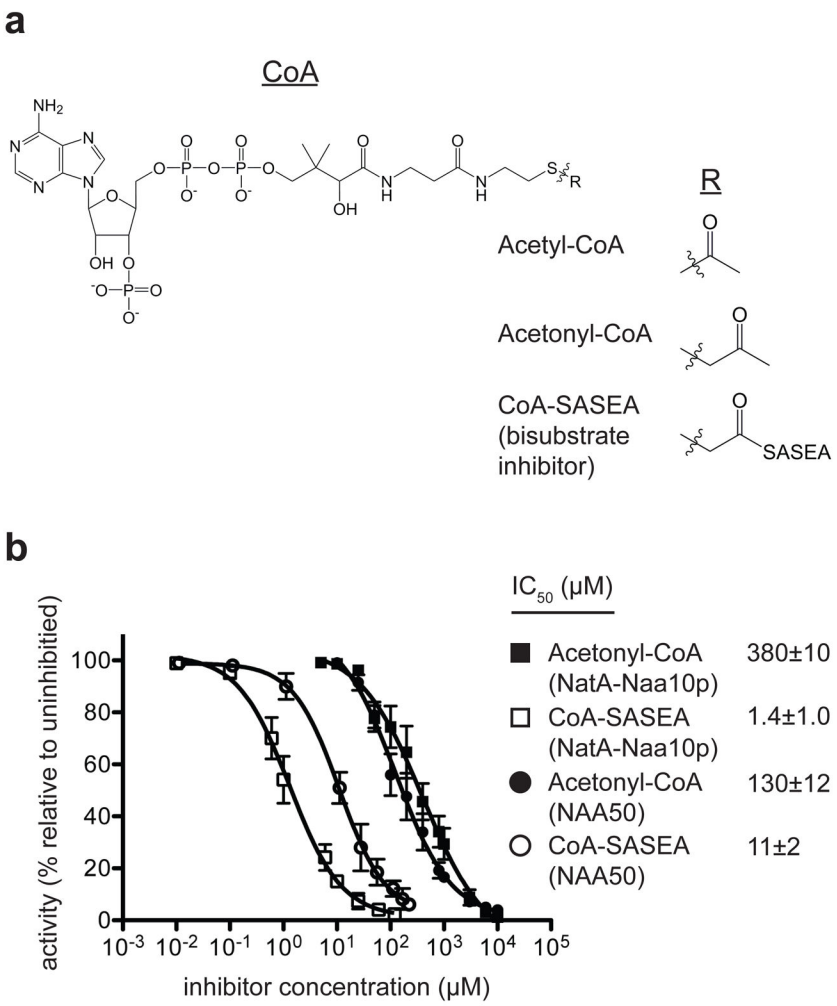

Figure 3.

Inhibitor structures and IC_{50} curves. (**a**) A ChemDraw representation of CoA is shown along with R-groups that correspond to the structure of AcCoA, acetonyl-CoA and the bisubstrate inhibitor (CoA-SASEA). (**b**) The dose response curves corresponding to the titration of CoA-SASEA and acetonyl-CoA into wild type Naa10p·Naa15p and hNAA50 acetyltransferase reactions. IC_{50} values for each inhibitor against each NAT are also indicated. Reactions were performed in triplicate and error bars represent the standard deviation of each measurement.

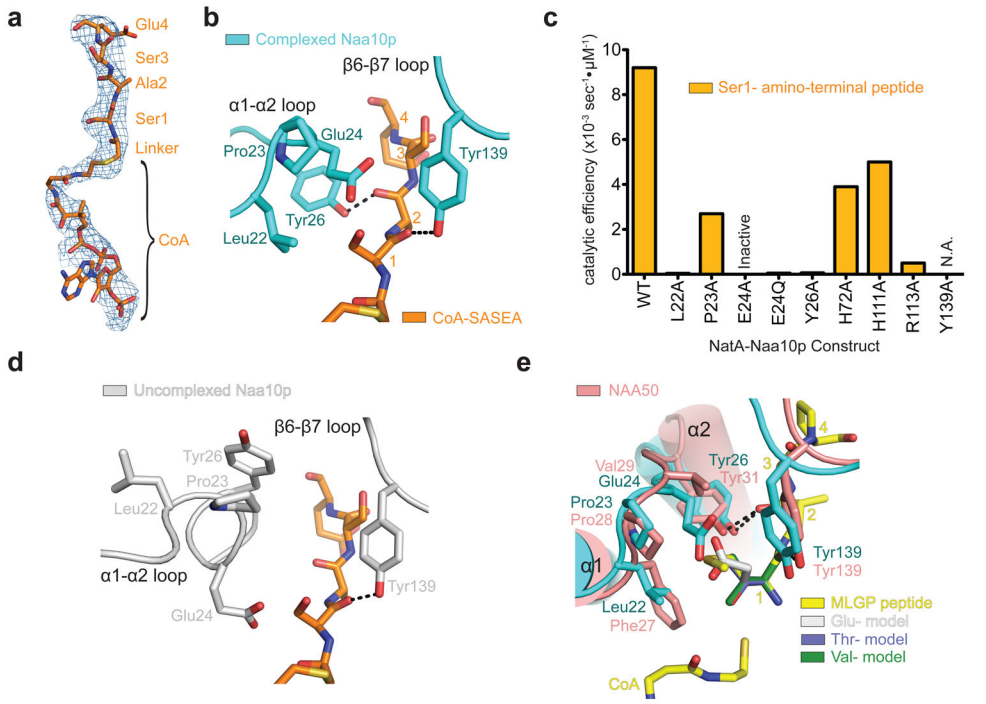

Figure 4.

Structure of the NatA complex bound to a bisubstrate inhibitor. (**a**) The structure of the bisubstrate inhibitor (orange) as it appears when bound to Naa10p in the context of the NatA complex. A simulated annealing omit map contoured to 1.5σ is shown in blue. Peptide residues are numbered relative to their position in the sequence. (**b**) A zoom view of the active site of the complexed Naa10p enzyme bound to the inhibitor. Residues that interact with the substrate peptide regions are shown in stick format and hydrogen bonds are shown as black dotted lines. (**c**) A bar graph showing the catalytic efficiencies of mutants. N.A. = This mutant has a K_m greater than 2,000 μM and catalytic efficiency cannot be calculated. (**d**) A model of the inhibitor docked into the uncomplexed form of Naa10p. Corresponding residues of the uncomplexed Naa10p structure that mediate substrate interactions in the complexed form of the enzyme are shown in stick format. (**e**) A model of the human NAA50 substrate and CoA (yellow stick format) docked into the active site of the complexed Naa10p. Glutamate (silver), valine (green) and threonine (blue) substrate amino-terminal side chains have also been docked into the peptide-binding pocket. Residues from each enzyme that have been shown to interact with their corresponding substrate peptides are shown in stick format.

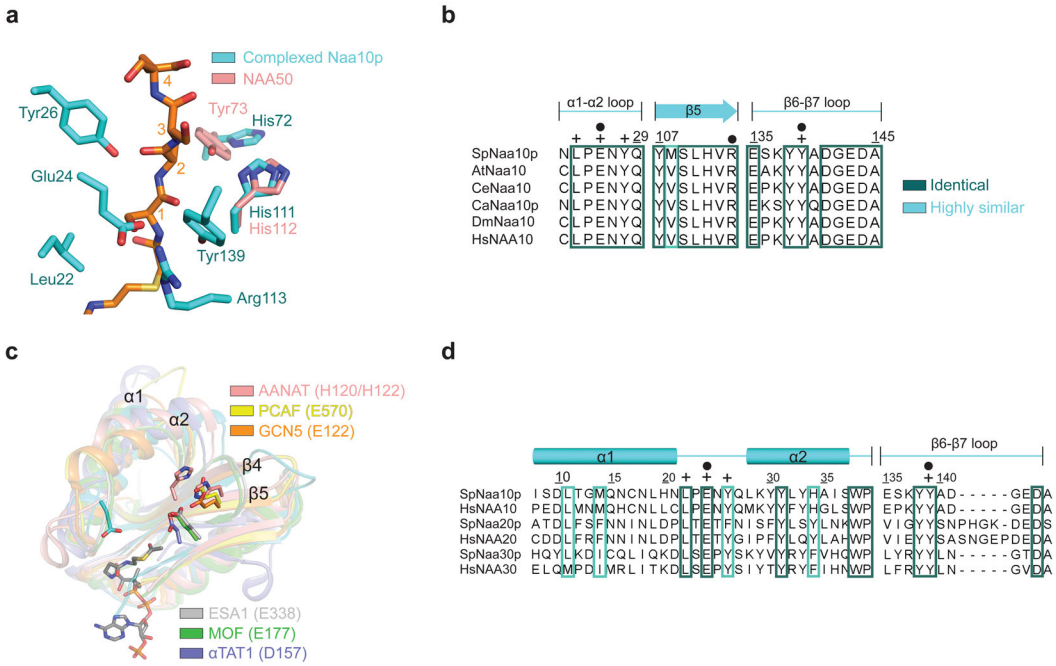

Figure 5.

The active site of the NatA complex. (**a**) The active site of the inhibitor-bound Naa10p subunit. The inhibitor (orange) and residues found to be important for catalysis (teal) are shown in stick format. The general bases of the hNAA50 enzyme (pink) and corresponding residues in Naa10p are also shown in stick format. (**b**) Sequence alignment of key features from Naa10 catalytic subunits from *Schizosaccharomyces pombe* (Sp), *Arabidopsis thaliana* (At), *Caenorhabditis elegans* (Ce), *Candida albicans* (Ca), *Drosophila melanogaster* (Dm), and *Homo sapiens* (Hs). Black circles (●) have been placed above proposed key catalytic residues and crosses (+) have been placed above proposed substrate binding residues. (**c**) Alignment of the NatA active site with the corresponding active sites of other protein and small molecule GNAT enzymes. Key catalytic residues for each enzyme are highlighted. (**d**) Sequence alignment of catalytic subunits of NAT complexes from NatA, NatB and NatC from *Schizosaccharomyces pombe* (Sp) and *Homo sapiens* (Hs). Black circles (●) have been placed above proposed key catalytic residues and crosses (+) have been placed above proposed substrate binding residues.

Table 1

Data collection, phasing and refinement statistics

	NatA-CoA-SASEA (native)[a]	NatA-CoA-SASEA (K2PtBr4)[a]		NatA-CoA-SASEA (SeMet)[a]	NatA-AcCoA (native)[a]	Naa10p-AcCoA (SeMet)[a]
Data collection						
Space group	P1	P1		P1	P1	P21
Cell dimensions						
a, *b*, *c* (Å)	81.439, 119.381, 134.063	81.021, 119.278, 133.520		80.388, 119.346, 133.054	80.739, 119.692, 132.024	40.974, 64.833, 60.723
α, *β*, *γ* (°)	80.200, 76.600, 70.425	80.305, 76.651, 70.384		79.265, 80.709, 70.450	80.284, 76.852, 70.651	90.000, 97.552, 90.000
		Peak	*Inflection*	*peak*		*Peak*
Wavelength	0.9795	1.0717	1.0722	0.9791	0.9795	0.9795
Resolution (Å)	50.00–2.60 (2.69–2.60)[b]	50.00–3.50 (3.63–3.50)	50–3.50 (3.63–3.50)	50.00–3.35 (3.47–3.35)	50.00–3.15 (3.26–3.15)	30.00–2.00 (2.03–2.00)
R_{sym}	5.9 (57.8)	6.6 (22.1)	6.0 (23.7)	11.2 (55.0)	7.6 (60.2)	11.0 (50.4)
$I / \sigma I$	20.8 (1.7)	22.3 (8.8)	26.9 (7.3)	8.9 (2.0)	17.2 (2.4)	36.4 (5.6)
Completeness (%)	97.3 (87.0)	99.2 (98.8)	99.1 (97.6)	99.1 (98.7)	99.0 (98.4)	98.7 (84.4)
Redundancy	4.7 (4.5)	5.9 (5.9)	5.8 (5.7)	3.3 (3.3)	3.9 (3.9)	10.9 (9.0)
Refinement						
Resolution (Å)	50.00–2.60 (2.69–2.60)				50.00–3.15 (3.26–3.15)	30.00–2.00 (2.03–2.00)
No. reflections	137,695				77,244	21,216
R_{work} / R_{free}[c]	23.64/26.92				21.70/24.60	18.25/23.08
No. atoms	28,758				28,431	2,788
Protein	28,051				28,115	2,456
Ligand/ion	308				204	102
Water	330				107	230
B factors (Å^2)						
Protein	59.9				86.0	24.7
Ligand/ion	58.7				88.7	22.5
Water	45.8				52.4	31.5
r.m.s. deviations						
Bond lengths (Å)	0.003				0.003	0.008
Bond angles (°)	0.860				0.790	1.167

[a] one crystal was used for data collection and refinement when applicable.

[b] Values in parentheses are for highest resolution shell.

[c] R_{free} was calculated using 5% of the reflection data.

Nat Struct Mol Biol. Author manuscript; available in PMC 2014 March 01.

Table 2

Catalytic parameters for wild type and mutant NatA as well as wild type monomeric Naa10p with various substrate peptides

Enzyme	Substrate (amino- terminus)	kcat s^{-1}	kcat Normalized to WT-NatA[b,c]	Km[a] μM	Km[a] Normalized to WT-NatA[b]
NatA-WT[d]	Ser1-	3.0±0.5	1.0	340±50	1.0
	Met1-	N.D.			
	Glu1-	N.D.			
NatA-Naa10p mutations					
H20A	Ser1-	1.5±0.2	0.50	920±110	2.7
L22A[d]	Ser1-	0.081±0.006	0.027	1,850±190	5.4
P23A	Ser1-	2.0±0.1	0.67	740±70	2.2
E24A	Ser1-	N.D.			
	Met1-	N.D.			
	Glu1-	0.65±0.08	0.22	250±60	0.74
E24D	Ser1-	0.42±0.04	0.14	750±40	2.2
E24Q	Ser1-	0.025±0.002	0.0083	440±50	1.3
Y26A	Ser1-	0.047±0.04	0.016	730±70	2.1
K29A	Ser1-	4.4±0.6	1.5	300±40	0.90
Y33A	Ser1-	0.44±0.02	0.15	510±40	1.5
K59A	Ser1-	1.5±0.1	0.50	760±70	2.2
E61A	Ser1-	1.1±0.1	0.37	410±60	1.2
E62A	Ser1-	1.1±0.1	0.37	590±70	1.7
H72A	Ser1-	2.4±0.3	0.80	600±70	1.8
R80A	Ser1-	0.85±0.04	0.28	770±70	2.3
H111A	Ser1-	2.0±0.2	0.67	330±30	0.97
R113A	Ser1-	0.19±0.02	0.063	380±30	1.1
Y139A	Ser1-	N.A.		>2,000	>5.9
K29A Y33A	Ser1-	0.61±0.07	0.20	340±80	1.0
K59A E61A	Ser1-	1.7±0.1	0.57	520±60	1.5
K59A E62A	Ser1-	0.73±0.5	0.24	410±30	1.2
NatA-Naa15p mutations					
R448A[e]					

Enzyme	Substrate (amino- terminus)	kcat s^{-1}	kcat Normalized to WT-NatA[b,c]	Km[a] μM	Km[a] Normalized to WT-NatA[b]
F449A[e]					
F474A	Ser1-	0.40±0.04	0.13	350±30	1.0
D532A	Ser1-	0.86±0.05	0.29	540±60	1.6
F533A	Ser1-	0.90±0.08	0.30	330±40	0.97
F536A	Ser1-	1.8±0.6	0.60	450±60	1.3
F533A F536A	Ser1-	3.0±0.3	1.0	500±50	1.5
Monomeric	Ser1-	N.D.			
Naa10p-WT	Met1-	N.D.			
	Glu1-	0.19±0.02	0.063	1,720±250	5.1

[a] The K_m values reported are for the substrate peptide described in the 'Substrate' column.

[b] WT (wild type).

[c] All normalizations reported are relative to wild type NatA catalytic parameters generated with the Ser1- amino-terminal substrate.

[d] The AcCoA K_m was calculated for these variants (WT=59±5 μM, L22A=53±6 μM).

[e] These mutations disrupted stable complex formation. Where k_{cat}=N.A., the K_m is greater than 2,000 μM and as a result a rate could not be calculated from our assay. Where k_{cat}=N.D., activity could not be detected in our assay.