Molecular basis for N-terminal acetylation by the heterodimeric NatA complex

Glen Liszczak¹,², Jacob M Goldberg², Håvard Foyn³, E James Petersson², Thomas Arnesen³,⁴ & Ronen Marmorstein¹,²

N-terminal acetylation is ubiquitous among eukaryotic proteins and controls a myriad of biological processes. Of the N-terminal acetyltransferases (NATs) that facilitate this cotranslational modification, the heterodimeric NatA complex has the most diversity for substrate selection and modifies the majority of all N-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 tetratricopeptide motifs and adopts a ring-like topology that wraps around the NatA-Naa10p peptide-CoA–conjugate inhibitor, as well as the structure of the uncomplexed Naa10p catalytic subunit. The NatA-Naa15p structure of the 100-kDa holo-NatA complex from Schizosaccharomyces pombe shows that the NatA substrate recognition domain is highly conserved across eukaryotes and appears to be independently active. They also have a more limited set of biologically relevant substrates and are not well characterized for substrate selection and specificity dependent upon the identity of the second residue. NatA, which is composed of the catalytic NAA10 subunit and the auxiliary NAA15 subunit, is the most promiscuous of all NAT enzymes; classically, it acetylates an α-amino group of substrates with an N-terminal alanine, cysteine, glycine, serine, threonine or valine residue. Notably, recent studies demonstrate that NAA10 also exists as a monomer in cells and that it can acetylate the α-amino group of substrates with N-terminal aspartate and glutamate residues that can be generated post-translationally, but it will not acetylate traditional NatA substrates. The NatB and NatC complexes acetylate the N termini of proteins with an N-terminal methionine with further specificity dependent upon the identity of the second residue. Although 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.

In an attempt to prepare the NatA complex for X-ray structure determination, we overexpressed the human complex.

RESULTS

Overall structure of NatA

In an attempt to prepare the NatA complex for X-ray structure determination, we overexpressed the human complex.

1Program in Gene Expression and Regulation, Wistar Institute, Philadelphia, Pennsylvania, USA. 2Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania, USA. 3Department of Molecular Biology, University of Bergen, Bergen, Norway. 4Department of Surgery, Haukeland University Hospital, Bergen, Norway. Correspondence should be addressed to R.M. (marmor@wistar.org).

Received 12 April; accepted 19 June; published online 4 August 2013; doi:10.1038/nsmb.2636
Table 1 Data collection and refinement statistics

|                              | NatA–CoA–SASEA (native) | NatA–CoA–SASEA (K2PtBr4) | NatA–CoA–SASEA (SeMet) | NatA–acyetyl CoA (native) | NatA10p–acyetyl CoA (SeMet) |
|------------------------------|--------------------------|---------------------------|------------------------|---------------------------|----------------------------|
|                              | P1                       | P1                        | P1                     | P1                        | P1                         |
| **Space group**              | P1                       | P1                        | P1                     | P1                        | P1                         |
| **Cell dimensions**          |                          |                           |                        |                           |                            |
| a, b, c (Å)                  | 81.439, 119.381          | 81.021, 119.278           | 80.388, 119.346        | 80.739, 119.692           | 40.974, 64.883             |
|                            | 134.063                  | 133.520                   | 133.054                | 132.024                   | 60.723                     |
| **α, β, γ (°)**              | 80.200, 76.600           | 80.305, 76.651            | 79.265, 80.709         | 80.284, 76.852            | 90.000, 97.552             |
|                            | 70.425                   | 70.384                    | 70.450                 | 70.651                    | 90.000                     |
| **Wavelength**               | 0.9795                   | 1.0717                    | 1.0722                 | 0.9791                    | 0.9795                     |
|                            |                          | 0.9791                    | 0.9791                 |                           |                            |
| **Resolution (Å)**           | 50.00–2.60               | 50.00–3.50                | 50.00–3.35             | 50.00–3.15                | 50.00–3.00                 |
|                            | (2.69–2.60)b             | (3.63–3.50)               | (3.63–3.50)            | (3.47–3.35)               | (3.26–3.15)                |
|                            |                          | (2.03–2.00)               | (2.03–2.00)            |                           |                            |
| **Rsym**                     | 5.9 (57.8)               | 6.6 (22.1)                | 6.0 (23.7)             | 11.2 (55.0)               | 7.6 (60.2)                 |
|                            | (3.63–3.50)              | (3.63–3.50)               | (3.47–3.35)            | (3.26–3.15)               | (2.03–2.00)                |
| **Completeness (%)**         | 97.3 (87.0)              | 99.2 (98.8)               | 99.1 (97.6)            | 99.1 (98.7)               | 98.9 (84.4)                |
| **Redundancy (%)**           | 4.7 (4.5)                | 5.9 (5.9)                 | 5.8 (5.7)              | 3.3 (3.3)                 | 3.9 (3.9)                  |
|                            |                          |                          |                        |                           |                            |
| **Data collection**          |                          |                           |                        |                           |                            |
| **No. reflections**          | 137,695                  | 77,244                    | 21,216                 | 77,244                    | 21,216                     |
|                            |                          | 21,216                    | 3.3 (3.3)              | 18.25 / 23.08             |
| **Rwork / Rfree**            | 23.64 / 26.92            | 21.70 / 24.60             | 18.25 / 23.08          | 21.70 / 24.60             | 18.25 / 23.08              |
| **No. atoms**                | 28,758                   | 28,431                    | 28,115                 | 28,115                    | 28,115                     |
| **Protein**                  | 28,051                   | 28,115                    | 28,115                 | 28,115                    | 28,115                     |
| **Ligand/ion**               | 308                      | 204                      | 204                    | 204                       | 204                        |
| **Water**                    | 330                      | 107                      | 107                    | 107                       | 107                        |
| **B factors (Å²)**           | 59.9                     | 86.0                     | 86.0                   | 86.0                      | 24.7                       |
|                            | 58.7                     | 88.7                     | 88.7                   | 88.7                      | 22.5                       |
|                            | 45.8                     | 52.4                     | 52.4                   | 52.4                      | 31.5                       |
| **B factors (Å²)**           |                          |                          |                        |                           |                            |
| **r.m.s. deviations**        | 0.003                    | 0.003                    | 0.003                  | 0.003                     | 0.008                      |
| **Bond lengths (Å)**         | 0.860                    | 0.790                    | 1.167                  | 0.790                     | 1.167                      |

*aOne crystal was used for data collection and refinement when applicable. SeMet, selenomethionine. *Values in parentheses are for highest-resolution shell. *Rfree was calculated with 5% of the reflection data.

in baculovirus-infected Sf9 insect cells and the orthologous proteins from *S. cerevisiae* and *S. pombe* in insect cells and bacteria, respectively. We found that coexpression of the *S. pombe* full-length Naa15p subunit (residues 1–729) with a C-terminal truncation construct of the Naa15p 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 acetyl CoA in the P1 space group with four heterodimers in the asymmetric unit, and the structure was determined by multilength anomalous diffraction with a combination of a K2PtBr4 heavy atom–soaked crystal and selenomethionine–derivatized protein (Table 1). The data set was collected to 3.15-Å resolution and the structure refined to *R* < sub = 23.6% and *R* < free = 26.9%, 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 tetraricopeptide repeat (TPR) motifs can be identified, which are often used in protein–protein interactions (Supplementary Fig. 1). Although a number of these TPR helices are involved in interactions 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. The Naa15p helices arrange themselves into a ring-like tertiary structure that wraps completely around the Naa10p subunit, burying a total of 3,520 Å2 of the solvent-accessible surface area of the enzyme (Fig. 1a,b and Supplementary Video 1). This fold is reminiscent of the importin-β structure, which uses 19 tandem HEAT repeats to form a solenoid tertiary structure that surrounds its binding partner, importin-α. Indeed, it appears that Naa15p stability depends on the binding of Naa10p. This is consistent with the observation that recombinant Naa15p undergoes marked degradation and precipitation when purified as a monomer.

The catalytic Naa10p subunit adopts a Gcn5-related N-acetyltransferase (GNAT) fold containing a central acetyl CoA–binding region and flanking N- and C-terminal segments that are similar to the corresponding regions 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 and Trp38) and several residues of Naa15p α1 (Ile8 and Leu11) and residues in Naa15p α29 (Trp526, Phe533 and Phe536) and α30 (Leu549 and Trp552), as well as hydrogen bonds...
A result of this interaction, the C-terminal region of the α helix with helices of Naa15p in the NatA complex (Figs. 1c and 1d). A 90° rotation of the view in a. Helices that are depicted in c and d are labeled. (c) Zoom view highlighting key residues that compose the predominantly hydrophobic interface between Naa10p α2–α3 and Naa15p α29–α30. (d) Zoom view of the intersubunit interface at the C-terminal region of Naa10p α1 and the Naa15p α25–α27–α28 helices.

Figure 1 Overall structure of the NatA complex bound to acetyl CoA. (a) Naa10p (teal) and Naa15p (brown) subunits, shown in cartoon, bound to acetyl CoA (CPK coloring and stick format). Only Naa15p helices that contact Naa10p are labeled. The dashed 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. Helices that are depicted in a and b are labeled. (c) Zoom view highlighting key residues that compose the predominantly hydrophobic interface between Naa10p α1–α2 and Naa15p α29–α30. (d) Zoom view of the intersubunit interface at the C-terminal region of Naa10p α1 and the Naa15p α25–α27–α28 helices.

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 of S. pombe in the absence of Naa15p, for comparison with the holo-NatA complex (Table 1). We determined the structure of Naa10p (residues 1–156) to 2.00-Å resolution, using a combination of single-wavelength anomalous diffraction and molecular replacement (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 and Ile36), which make intramolecular interactions with residues in Naa10p α1 and β3 (Ile8, Leu11, Met14, Tyr55 and Tyr57) in apo-Naa10p but shift to make alternative intramolecular interactions with helices of Naa15p in the NatA complex (Figs. 1c and 1d). 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. Notably, 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 of traditional substrates (alanine, cysteine, glycine, serine, threonine or valine) (Supplementary Videos 2 and 3). Specifically, Naa10p residues 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, substantially 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 by the NatA complex, and Naa10p is likely to represent an active GNAT 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 1-Met-Leu-2-N-terminal sequence shows a high degree of structural conservation (r.m.s. deviation of 1.52 Å). The corresponding alignment of the complexed and uncomplexed forms of Naa10p showed less structural conservation, with an r.m.s. deviation 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 nonmethionine N-terminal substrates, we synthesized a bisubstrate conjugate in which CoA is covalently attached to NAA50 over Naa10p can be explained by the stronger binding of acetyl CoA to NAA50 (Km = 27 ± 2 μM) than to Naa10p (Km = 59 ± 5 μM). The markedly higher potency of the CoA-SASEA inhibitor
Table 2 Catalytic parameters for wild-type and mutant NatA and wild-type monomeric Naa10p with various substrate peptides

| Enzyme          | Substrate (N terminus) | $k_{cat}$ (s$^{-1}$) | $k_{cat}$ (normalized to WT NatAb$^{-1}$) | $K_m$ (µM) (normalized to WT NatAb$^{-1}$) |
|-----------------|------------------------|----------------------|------------------------------------------|------------------------------------------|
| NatA WT$^{+}$   | Ser1                   | 3.0 ± 0.5            | 1.0                                      | 340 ± 50                                 | 1.0                                      |
|                 | Met1                   | ND                   |                                         |                                         |                                         |
|                 | Glu1                   | ND                   |                                         |                                         |                                         |
| NatA-Naa10p mutations | H20A             | 1.5 ± 0.2            | 0.50                                     | 920 ± 110                                | 2.7                                      |
|                 | L22A$^a$              | 0.081 ± 0.006        | 0.027                                    | 1,850 ± 190                              | 5.4                                      |
|                 | P23A                   | 2.0 ± 0.1            | 0.67                                     | 740 ± 70                                 | 2.2                                      |
|                 | E24A                   | ND                   |                                         |                                         |                                         |
|                 | Met1                   | ND                   |                                         |                                         |                                         |
|                 | Glu1                   | 0.65 ± 0.08          | 0.22                                     | 250 ± 60                                 | 0.74                                     |
|                 | E24D                   | 0.42 ± 0.04          | 0.14                                     | 750 ± 40                                 | 2.2                                      |
|                 | E24Q                   | 0.025 ± 0.002        | 0.0083                                   | 440 ± 50                                 | 1.3                                      |
|                 | Y26A                   | 0.047 ± 0.04         | 0.016                                    | 730 ± 70                                 | 2.1                                      |
|                 | K29A                   | 4.4 ± 0.6            | 1.5                                      | 300 ± 40                                 | 0.90                                     |
|                 | Y33A                   | 0.44 ± 0.02          | 0.15                                     | 510 ± 40                                 | 1.5                                      |
|                 | K59A                   | 1.5 ± 0.1            | 0.50                                     | 760 ± 70                                 | 2.2                                      |
|                 | E61A                   | 1.1 ± 0.1            | 0.37                                     | 410 ± 60                                 | 1.2                                      |
|                 | E62A                   | 1.1 ± 0.1            | 0.37                                     | 590 ± 70                                 | 1.7                                      |
|                 | H72A                   | 2.4 ± 0.3            | 0.80                                     | 600 ± 70                                 | 1.8                                      |
|                 | R80A                   | 0.85 ± 0.04          | 0.28                                     | 770 ± 70                                 | 2.3                                      |
|                 | H11A                   | 2.0 ± 0.2            | 0.67                                     | 330 ± 30                                 | 0.97                                     |
|                 | R113A                  | 0.19 ± 0.02          | 0.063                                    | 380 ± 30                                 | 1.1                                      |
|                 | Y139A                  | NA                   |                                         | >2,000                                  | >5.9                                     |
|                 | K29A Y33A              | 0.61 ± 0.07          | 0.20                                     | 340 ± 80                                 | 1.0                                      |
|                 | K59A E61A              | 1.7 ± 0.1            | 0.57                                     | 520 ± 60                                 | 1.5                                      |
|                 | K59A E62A              | 0.73 ± 0.5           | 0.24                                     | 410 ± 30                                 | 1.2                                      |
| NatA-Naa15p mutations | R448A$^a$         | 0.40 ± 0.04          | 0.13                                     | 350 ± 30                                 | 1.0                                      |
|                 | F449A$^a$              | 0.38 ± 0.05          | 0.29                                     | 540 ± 60                                 | 1.6                                      |
|                 | F474A                  | 0.90 ± 0.08          | 0.30                                     | 330 ± 40                                 | 0.97                                     |
|                 | F536A                  | 1.8 ± 0.6            | 0.60                                     | 450 ± 60                                 | 1.3                                      |
|                 | F533A F536A            | 3.0 ± 0.3            | 1.0                                      | 500 ± 50                                 | 1.5                                      |
| Monomeric Naa10p WT | Ser1                | ND                   |                                         |                                         |                                         |
|                 | Met1                   | ND                   |                                         |                                         |                                         |
|                 | Glu1                   | 0.19 ± 0.02          | 0.063                                    | 1,720 ± 250                              | 5.1                                      |

$^a$ $K_m$ values are for the substrate peptide described in the Substrate column. $^b$ WT, wild type. $^c$ All normalizations are relative to wild-type NatA catalytic parameters generated with the Ser1 N-terminal substrate. $^d$ The acetyl CoA $K_m$ was calculated for these variants (WT = 59 ± 5 µM; L22A = 53 ± 6 µM). $^e$ These mutations disrupted stable complex formation. Where $K_{cat}$ is NA (not applicable), the $K_m$ is >2,000 µM, thus a rate could not be calculated from our assay. Where $K_{cat}$ is ND (not determined), activity could not be detected in our assay. Errors represent s.d. ($n = 3$).

relative to acetyl-CoA toward 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 N-terminal substrate binding by NatA.

We were able to cocrystallize NatA bound to CoA-SASEA and determine its X-ray crystal structure to 2.60-Å resolution, using the refined NatA-acetyl CoA 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. We observed only modest rearrangements in Naa10p upon CoA-SASEA binding as compared to acetyl-CoA 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). We performed these assays in the presence of a substrate peptide with an N-terminal sequence identical to that of the inhibitor, a Met1 N-terminal substrate peptide that corresponds to the NAA50 in vivo substrate and a substrate peptide with an N-terminal 1-Glu-Glu-Glu-3 sequence that matches the sequence for γ-actin, the only known NAA10 in vivo substrate. Although monomeric Naa10p is inactive with the Met1 and Ser1 N-terminal substrates, we were able to generate catalytic parameters with the Glu1 N-terminal substrate ( kat = 0.19 ± 0.02 s$^{-1}$; K m = 1.72 ± 0.25 µM) (Table 2). Whereas previous studies were able to detect acetylation of the N-terminal Glu1 substrate by the NatA complex$^{26}$, we were unable to detect activity toward this substrate. In contrast, we were able to determine kinetic parameters for the Ser1 N-terminal substrate in the presence of a saturating amount of acetyl CoA ( kat = 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 ( kat K m$^{-1}$), thus suggesting that steric alterations to the N-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 kat of the NatA-Naa10p E24Q mutant was reduced by 99%, and the K m was unaffected, thus emphasizing 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, whereas we could not calculate the catalytic efficiency for the NatA-Naa10p Y139A mutant, owing to the immeasurably high K m (Table 2), which is probably caused by the involvement of Tyr139 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 toward the N-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 is 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 N-terminal substrate appear to be out of place for catalysis by the Naa10p monomer, the subunit is able to catalyze acetylation of the Glu1 N-terminal substrate, thus suggesting that the acetylation reaction by monomeric Naa10p is catalyzed through an alternate mechanism to that of the NatA complex. To investigate the possibility that the Glu1 N-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 N-terminal Ser1 or Met1 N-terminal substrates used in this study, is able to acetylate the Glu1 N-terminal substrate with a rate of 0.65 ± 0.08 s⁻¹, which is ~22% of the rate of the wild-type NatA complex with its cognate Ser1 N-terminal substrate peptide (Table 2). We suspect that the more open N-terminal residue-binding pocket is able to accommodate an N-terminal Glu1 substrate peptide and that the carboxylate group of the glutamate residue–binding pocket is able to accommodate an N-terminal Glu1 residue.

To determine the molecular basis for the inability of NatA to acetylate N-terminal methionine and glutamate residues, we modeled an N-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, thus suggesting that Naa10p Glu24 is particularly important for excluding these N-terminal substrates. This is consistent with the importance of the corresponding human NAA50 Val29 residue for N-terminal methionine recognition and acetylation. We also modeled substrate N-terminal threonine and valine residues, the most bulky NatA substrates, into the active site to show that these residues' side chains, similarly to serine's, 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. 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 and H111A revealed that Naa10p His72 and His111 (Fig. 5a). However, point mutations of H72A and H111A resulted in a 10-fold decrease in catalytic activity (kcat = 2.4 ± 0.3 s⁻¹) and 10100-fold decrease in catalytic activity (kcat = 2.0 ± 0.2 s⁻¹), respectively.

**Figure 2** Structure of the Naa10p monomer bound to acetyl CoA. (a) Structural alignment of the uncomplexed Naa10p monomer (silver cartoon) bound to acetyl CoA (gray stick format) with the active Naa10p from the NatA structure (teal cartoon). (b) 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) 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 repositioned residues 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) Structural alignment of the complexed and uncomplexed forms of Naa10p with hNAA50 (pink) bound to CoA (gray stick format).

**Figure 3** Inhibitor structures and IC50 curves. (A) A ChemDraw representation of CoA, shown along with R groups that correspond to the structure of acetyl CoA, acetylonyl CoA and the bisubstrate inhibitor (CoA-SASEA). (B) Dose-response curves corresponding to the titration of CoA-SASEA and acetylonyl CoA into wild-type NatA (Naa10p–Naa15p) and hNAA50 acetyltransferase reactions. IC50 values for each inhibitor against each NAT are also indicated. Reactions were performed in triplicate; error bars, s.d. of each measurement.
results arguing against the importance of these residues in catalysis by Naa10p (Table 2). Notably, NAA50 also has an active site water molecule that probably has a catalytic role. This water molecule is not observed in the Naa10p active site and further supports the idea that Nata uses 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 proximity to the N-terminal residue and could be playing a part 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 $K_m$, thus suggesting that the carboxylic acid moiety is essential for catalysis but is not required for substrate binding (Table 2). In both Gcn5 and ESA1, mutation of the general base glutamate to a glutamine also results in a decrease in $k_{cat}$ by ~2 orders of magnitude, and this 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 causes an increase in side chain $pK_a$ that allows it to directly deprotonate the

---

**Figure 4** Structure of the NatA complex bound to a bisubstrate inhibitor. (a) Structure of the bisubstrate inhibitor (orange) as it appears when bound to Naa10p in 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) 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 dashed lines. N-terminal peptide residues are numbered 1–4 sequentially from the N-terminus. (c) Bar graph showing the catalytic efficiencies of mutants. NA, not applicable, as this mutant has a $K_m$ $>$ 2,000 μM, and catalytic efficiency cannot be calculated. WT, wild type. (d) Model of the inhibitor docked into the uncomplexed form of Naa10p. Residues that mediate substrate interactions in the complexed form of the enzyme are shown in stick format. (e) 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 N-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. N-terminal peptide residues are numbered 1–4 sequentially from the N-terminus. (b) Sequence alignment of key features from Naa10 catalytic subunits from *S. pombe* (Sp), *Arabidopsis thaliana* (At), *Caenorhabditis elegans* (Ce), *Candida albicans* (Ca), *Drosophila melanogaster* (Dm) and *Homo sapiens* (Hs). Black circles (•) mark proposed key catalytic residues, and plus signs (+) mark proposed substrate-binding residues. A line has been placed over numbered residues. (c) Alignment of the NatA active site with the corresponding active sites of other protein and the small-molecule GNAT enzymes ESA1, MOF1 and αTAT. Key catalytic residues for each enzyme are highlighted. His120/His122 signifies two residues that contribute to general base catalysis. (d) Sequence alignment of catalytic subunits of NAT complexes from NatA, NatB and NatC from *S. pombe* (Sp) and *H. sapiens* (Hs). Black circles (•) mark proposed key catalytic residues, and plus signs (+) mark proposed substrate-binding residues.
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 it could be involved in proper positioning of the α-amino group within the active site. Although 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, owing 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, by altering the pK 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 (Fig. 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, and this reinforces the importance of these residues in catalysis (Fig. 5b).

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

**DISCUSSION**

The data presented here reveal that Naa15p allosterically reconfigures the Naa10p active site for sequence-specific N-terminal acetylation. Specifically, this study highlights a unique mechanism in which NatA substrate binding specificity is coupled to the catalytic strategy. This finding is supported by evidence that an N-terminal residue of an incompatible substrate, such as a methionine, would have to displace Glu24, an essential catalytic residue, for binding to occur (Fig. 4c). Furthermore, upon modeling all bulky, β-branched NatA cognate substrate N-terminal residues into the binding pocket, we find that none of them clash with the Glu24 residue, and this further supports the importance of this residue and its role in substrate-specific acetylation (Fig. 4c). We also note that in the absence of Glu24, the Glu1 N-terminal substrate can undergo acetylation, probably because 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.

Comparison of the Naa10p active site with the corresponding region of NAA50 and other protein or small molecule acetyltransferases shows prominent differences that appear to underlie the requirement of Naa15p binding for Naa10p acetyltransferase activity toward conventional substrates. In particular, three catalytically essential Naa10p residues in the α1–α2 loop N terminal to the catalytic core region—Leu22, Glu24 and Tyr26—are not correctly positioned for catalysis in the uncomplexed form of the enzyme. Notably, 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 acetyltransferases 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's key residues for catalysis are on the opposite side of the substrate, relative to those of NAA50 and the other acetyltransferases, 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, and this reinforces the importance of these residues in catalysis (Fig. 5b). In addition, on the basis of 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), and it also reinforces the importance of this loop for enzymatic activity of NAT enzymes. However, NatB and NatC both acetylate methionine-containing N-terminal substrates, so these complexes probably use other strategies to accommodate their respective substrates. Many protein lysine side chain acetyltransferases, including HAT1 (ref. 30), HBO1 (ref. 29) and Sas2 (ref. 28), also require auxiliary subunits for activated acetylation, and it is of strong interest to determine whether 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 molecules for cell-based studies, owing 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, whereas acetyl CoA analogs will probably 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.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** The coordinates of the structures have been deposited in the Protein Data Bank under accession codes 4KVX (Naa10p–acytel CoA), 4KVO (Naa10p–Naa15p–acytel CoA) and 4KVM (Naa10p–Naa15p–bisubstrate inhibitor).

**Note:** Any Supplementary Information and Source Data files are available in the online version of the paper.

**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 STC 0332341. T.A. was supported by the Research Council of Norway and the Norwegian Cancer Society. We also acknowledge Marmorstein laboratory members and E. Skordalakes for helpful discussions.

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 videos. 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.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

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. USA 106, 8157–8162 (2009).
2. Forte, G.M., Poole, M.R. & Stirling, C.J. N-terminal acetylation inhibits protein targeting to the endoplasmic reticulum. PLoS Biol. 9, e1001073 (2011).
3. Hwang, C.S., Shemorry, A. & Varshavsky, A. N-terminal acetylation of cellular proteins creates specific degradation signals. Science 327, 973–977 (2010).
4. Scott, D.C., Monda, J.K., Bennett, E.J., Harper, J.W. & Schulman, B.A. N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. Science 334, 674–678 (2011).
5. Starheim, K.K., Gevaert, K. & Arnesen, T. Protein N-terminal acetyltransferases: when the start matters. Trends Biochem. Sci. 37, 152–161 (2012).
6. Yi, C.H. et al. Metabolic regulation of protein N-alpha-acetylation by Bcl-xL promotes cell survival. Cell 146, 607–620 (2011).
7. Arnesen, T. et al. Identification and characterization of the human ARD1-NAT3 protein acetyltransferase complex. Biochem. J. 386, 433–443 (2005).
8. Gauthsi, M. et al. The yeast Nα-acetyltransferase NAtA is quantitatively anchored to the ribosome and interacts with nascent polypeptides. Mol. Cell. Biol. 23, 7403–7414 (2003).
9. Mullen, J.R. et al. Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. EMBO J. 8, 2067–2075 (1989).
10. Park, E.C. & Szostak, J.W. ARD1 and NAT1 proteins form a complex that has N-terminal acetyltransferase activity. EMBO J. 11, 2087–2093 (1992).
11. Polevoda, B., Hoskins, J. & Sherman, F. Properties of Nat4, an N-terminal acetyltransferase subunit. Mol. Cell. Proteomics 10, M110.004580 (2011).
12. Van Damme, P. et al. Proteome-derived peptide libraries allow detailed analysis of the substrate specificities of Nα-acetyltransferases and point to hNaa10p as the post-translational actin Nα-acetyltransferase. Mol. Cell Proteomics 10, 2132–2143 (2011).
13. Yu, M. et al. Correlation of expression of human arrest-defective-1 (hARD1) protein with breast cancer. Cancer Invest. 27, 978–983 (2009).
14. Fluge, Ø., Bruland, O., Akslen, L.A., Varhaug, J.E. & Lillehaug, J.R. NATH, a novel gene overexpressed in papillary thyroid carcinomas. Oncogene 21, 5056–5068 (2002).
15. Yu, M. et al. Immunohistochemical analysis of human arrest-defective-1 expressed in cancers in vivo. Oncol. Rep. 21, 909–915 (2009).
16. Polevoda, B. & Sherman, F. Composition and function of the eukaryotic N-terminal acetyltransferase subunits. Biochem. Biophys. Res. Commun. 308, 1–11 (2003).
17. Polevoda, B. & Sherman, F. N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. J. Mol. Biol. 325, 595–622 (2003).
18. Yu, M., Gevaert, K., Gronryko, D., Velde, R., Varhaug, J.E. & Arnesen, T. Composition and biological significance of the human Nε-terminal acetyltransferases. BMC Proc. 3 (suppl. 6), S3 (2009).
19. Polevoda, B., Norbeck, J., Takakura, H., Blomberg, A. & Sherman, F. Identification and characterization of genes and mutants for N-terminal acetyltransferases from Saccharomyces cerevisiae. EMBO J. 18, 6155–6168 (1999).
20. Van Damme, P. et al. Metabolic regulation of protein Nα-acetylation by Bcl-xL promotes cell survival. Cancer Res. 66, 10677–10682 (2006).
21. Ren, T. et al. Generation of novel monoclonal antibodies and their application for detecting ARD1 expression in colorectal cancer. Cancer Lett. 264, 83–92 (2008).
22. Kalvik, T.V. & Arnesen, T. Protein N-terminal acetyltransferases in cancer. Oncogene 32, 269–276 (2013).
ONLINE METHODS and purification. The full-length Naa10p gene (encoding residues 1–177) was cloned out of the S. pombe genome (ATCC), and several C-terminal truncation constructs were engineered into a modified PETDUET vector containing a TEV protease–cleavable His6 tag. All constructs were transformed into Rosetta (DE3)pLysS competent E. coli cells, which were grown to an OD600 of 0.7 and induced with 0.5 mM IPTG at 16 °C for ~16 h. Cells were isolated by centrifugation and lysed by sonication in a buffer containing 25 mM Tris, pH 8.0, 500 mM NaCl, and 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-h dialysis into lysis buffer. The solution was passed through an additional nickel column to remove TEV protease as well as any uncleaved Naa10p. The resin was then washed with approximately seven column volumes of lysis buffer supplemented with 25 mM imidazole, which was pooled with the initial elution. The resin was then washed with approximately seven 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 25 mM sodium citrate monobasic, pH 5.5, 100 mM NaCl, and 2 mM DTT and loaded onto a buffer supplemented with 25 mM imidazole, which was pooled with the initial elution. The resin was then washed with approximately seven 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 25 mM 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 with a salt gradient (100–750 mM NaCl). Peak fractions were concentrated to 8 mg mL−1 (20-kDa concentrator; Amicon Ultra, Millipore), as measured by UV280 absorbance (extinction coefficient = 18,000), for crystallization trials. The selenomethionine derivative was prepared by expression of Naa10p (residues 1–156) in minimal medium (Molecular Dimensions) containing 50 μg L−1 of selenomethionine (Sigma), which was purified with the scheme described above.

Naa10p–Naa15p complex expression and purification. The full-length Naa15p gene (encoding residues 1–729) was cloned out of the S. pombe genome and engineered into MCSII of a modified pETDUET vector containing a TEV protease–cleavable His6 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 1 mL, 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 for crystallization trials. The selenomethionine-labeled complex was obtained as described above. Point-mutation proteins were generated with the Stratagene QuikChange protocol, and protein containing point mutations was expressed and purified with the protocols described above44.

Naa10p crystallization and structure determination. Naa10p at a concentration of 8 mg mL−1 was mixed with acetyl CoA at a 1:3 molar ratio. Initial crystallization hits were obtained from an Naa10p construct containing residues 1–161 (Naa10p residues 1–161) with 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 (residues 1–156), which did not crystalize in the absence of the seed crystals. Large crystals of Naa10p (residues 1–156) were obtained from this screen in a 1:1 mixture of protein and well solution (14% PEG 3350 and 0.1 M bis-Tris, pH 6.5) with 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 included a drop with a 2.5:0.5:2.0 ratio of protein (8 mg mL−1)/glycerol (50%)/well solution, mixed in that order. Selenium-methylated crystals of this construct were obtained with identical conditions with a lower protein concentration (6.5 mg mL−1). Diffraction-quality crystals required 1–3 days to grow to maximum dimensions.

Data sets were collected at beamlines X25 and X29A at the National Synchrotron Light Source (Brookhaven National Laboratory) and processed with 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 data sets were processed to 3.50-Å resolution. The two wavelengths MAD data were used by the Hybrid Substructure Search in Phenix to identify a total of eight heavy-atom sites in the asymmetric unit, and initial phases were obtained with SOLVE49. Density modification was performed in Resolve by application of four-fold NCS (corresponding to each heterodimer) and with histogram matching and solvent flattening23. Initially, the Naa15p helices were built as polyalanine chains, and the Naa10p molecules were manually placed into the density-modified map with Coot. The resulting structure was used as a molecular replacement model for the selenomethionine-derivative data set 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 of the map to 3.0σ and placement of 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 with the methionine residues as a guide. The final structure was used as a molecular replacement model for the high-resolution data sets (200° of data each) of the Naa10p–Naa15p–acetyl CoA (3.10-Å resolution) and the Naa10p–Naa15p–bisubstrate inhibitor (2.60-Å resolution) complex. Refinement of the binary-complex structures was carried out in Phenix, and four-fold NCS restraints were applied to the two separate groups: the full-length Naa10p molecules, and residues 100–731 of Naa15p. The positioning of the first five 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 with the open-access server). The backbone for the e20–021 loop from residues 356–377 can be traced; however, it is largely disordered (average B factor = 196 Å2), as it probably samples numerous conformations in the crystal. No density was observed for a loop from...
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 acetyl CoA (500 µM) was used in all enzymatic reactions, and the substrate peptide (NH₂–H2SASEAGVRWGPVGRRRRP–COOH; GenScript) concentration was varied to determine steady-state catalytic parameters. The first seven residues of this peptide correspond to the N-terminal sequence of a protein that has been shown to be highly N-terminally acetylated in yeast, whereas the C-terminal 12 residues form a basic tag that make the substrate peptide unique with our assay.1 The Glui N-terminal substrate peptide (NH₂–EEIAALLWGRWGPVGRRRRP–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 N-terminal substrate peptide (NH₂–MLGEPGRWGPVGRRRRP–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 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 981 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 acetyl CoA. The papers were then dried with acetone and added to 4 mL of scintillation fluid, and the signal was measured with 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- N-terminal peptide were carried out for 1 h with 500 µM acetyl CoA, substrate peptide (100–1,500 µM) and up to 1 µM untagged enzyme. No activity could be detected. Reactions with the Glui N-terminal peptide were performed at 25 °C for 30 min in the presence of 200 nM enzyme and 500 µM acetyl CoA at various peptide concentrations (30–1,000 µM). For the NatA complex, substrate peptide K₆₅ and Vₘ₅₆ parameters were derived by titration of the substrate peptide at eight different concentrations ranging from 100 to 1,500 µM in the presence of 500 µM acetyl CoA. To obtain kinetic data for the point mutants L22A, Y26A, and Y139A, 500 nM of enzyme was used in the reaction. Additionally, the acetyl CoA Kₘ values for the wild-type protein as well as the L22A mutant were determined by titration of the acetyl CoA at eight different concentrations (ranging from 20 to 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 with a standard curve created with known concentrations of radioactive acetyl CoA added to scintillation fluid. 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 with 100 nM wild-type protein for 3 min in the buffer described for the previous assays. Iₕ₅₀ values for Naa10p were calculated by titration of inhibitor (CoA–SASEA, 0.01–120 µM and acetyl CoA, 5–10,000 µM) into reaction mixtures containing fixed concentrations of substrate peptide (500 µM) and acetyl CoA (200 µM). Similarly, Iₕ₅₀ values for NAA50 (expressed and purified as previously described10) were calculated by titration of inhibitor (CoA–SASEA, 0.01–240 µM and acetyl CoA, 5–10,000 µM) into reaction mixtures containing fixed concentrations of substrate peptide (1,200 µM) and acetyl CoA (200 µM). To clarify, peptide concentrations were held constant at Kₚ. To calculate the Kₕ of the CoA–SASEA inhibitor versus acetyl CoA, substrate peptide concentration was held constant at 1,200 µM, and kinetic curves were generated with acetyl CoA concentrations ranging from 15 to 500 µM in the presence of inhibitor concentrations ranging from 0 to 15 µM. To calculate the Kₕ of the CoA–SASEA inhibitor versus substrate peptide, acetyl CoA concentration was held constant at 500 µM, and kinetic curves were generated with peptide concentrations ranging from 200 to 1,000 µM in the presence of inhibitor concentrations ranging from 0 to 30 µM. All reactions described here were performed in duplicate, and data were fit to a sigmoidal dose-response curve with GraphPad Prism, version 5.01.

Bisubstrate inhibitor synthesis. The peptide SASEA was synthesized on a 100-µmol scale with standard Fmoc solid-phase synthesis techniques. Rink amide resin (100–200 mesh), Fmoc-Ala-OH, Fmoc-Glu(Bu)-OH, Fmoc-Ser(Bu)-OH, and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Novabiochem (currently EMD Millipore). All other reagents, including Sigmacote, were purchased from Sigma-Aldrich. A clean, oven-dried glass reaction vessel (RV) was treated with Sigmacote and rinsed with dimethylformamide (DMF) before use. Rink amide resin (151 mg, substitution 0.66 mmol 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 with 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-acetylseraurea with carbodiimide chemistry. Bromoacetic acid (10 equiv.; 139 mg, 1 mmol) was preactivated with N,N′-diisopropylcarbodiimide (20 equiv.; 309 µL, 2 mmol) in 2 mL dry CH₂Cl₂ 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 CH₂Cl₂. Residual solvent was removed with vacuum suction, and the beads were dried under vacuum for 30 min. A cocktail of trifluoracetic acid (TFA)/trisopropylsilane/water (38:1:1 v/v/v) 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 N₂ flow after each incubation. The solutions were combined and concentrated under reduced pressure. The residue was dissolved in a H₂O-CH₂Cl₂ solution (3:1 v/v) and lyophilized. The crude lyophilized powder was dissolved in 3 mL triethylammonium bicarbonate buffer, pH 8.0. 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 reverse-phase high-performance liquid chromatography (HPLC) with a Varian ProStar HPLC instrument outfitted with a diode array detector (currently Agilent Technologies) and Vydac 218TP C18 analytical and prep columns (Grace-Vydac) with a linear solvent gradient. The aqueous phase was 0.1% TFA in H₂O, and the organic phase was 0.1% TFA in CH₃CN. 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 washout 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 LR MALDI-TOF mass spectrometer to confirm identity. MALDI-MS(−) m/z cladir for Ca₃H₂F₂N₁O₂P₃S₂[M–H]− 1268.31, found 1268.32.

44. Braman, J., Papworth, C. & Greener, A. Site-directed mutagenesis using double-stranded plasmid DNA templates. Methods Mol. Biol. 57, 31–44 (1996).
45. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 (1997).
46. McCoy, A.J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
47. Emstler, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 (2004).
48. Langer, G., Cohen, S.X., Lamzin, V.S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARPlWARP version 7. Nat. Protoc. 3, 1171–1179 (2008).

49. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010).

50. Painter, J. & Merritt, E.A. A molecular viewer for the analysis of TLS rigid-body motion in macromolecules. Acta Crystallogr. D Biol. Crystallogr. 61, 465–471 (2005).

51. Painter, J. & Merritt, E.A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr. D Biol. Crystallogr. 62, 439–450 (2006).

52. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. 55, 849–861 (1999).

53. Terwilliger, T.C. Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. 56, 965–972 (2000).