Supplemental files to:

Crystal structure of the Golgi-associated human N-alpha acetyltransferase 60 reveals the molecular determinants for substrate-specific acetylation

Svein Isungset Støve[1,2], Robert S. Magin[3,4, 5], Håvard Foyn[1], Bengt Erik Haug[6], Ronen Marmorstein[3,4,5] and Thomas Arnesen[1,2]

[1]Department of Molecular Biology, University of Bergen, N-5020 Bergen, Norway; [2]Department of Surgery, Haukeland University Hospital, N-5021 Bergen, Norway; [3]Department of Biochemistry and Biophysics and Abramson Family Cancer Research Institute, and the [4]Graduate Group in Biochemistry and Molecular Biophysics, Perelman School of Medicine at the University of Pennsylvania, PA 19104, USA; [5]Program in Gene Expression and Regulation, Wistar Institute, PA 19104, USA; [6]Department of Chemistry, University of Bergen, N-5020 Bergen, Norway

Corresponding authors: Ronen Marmorstein, email: marmor@mail.med.upenn.edu;
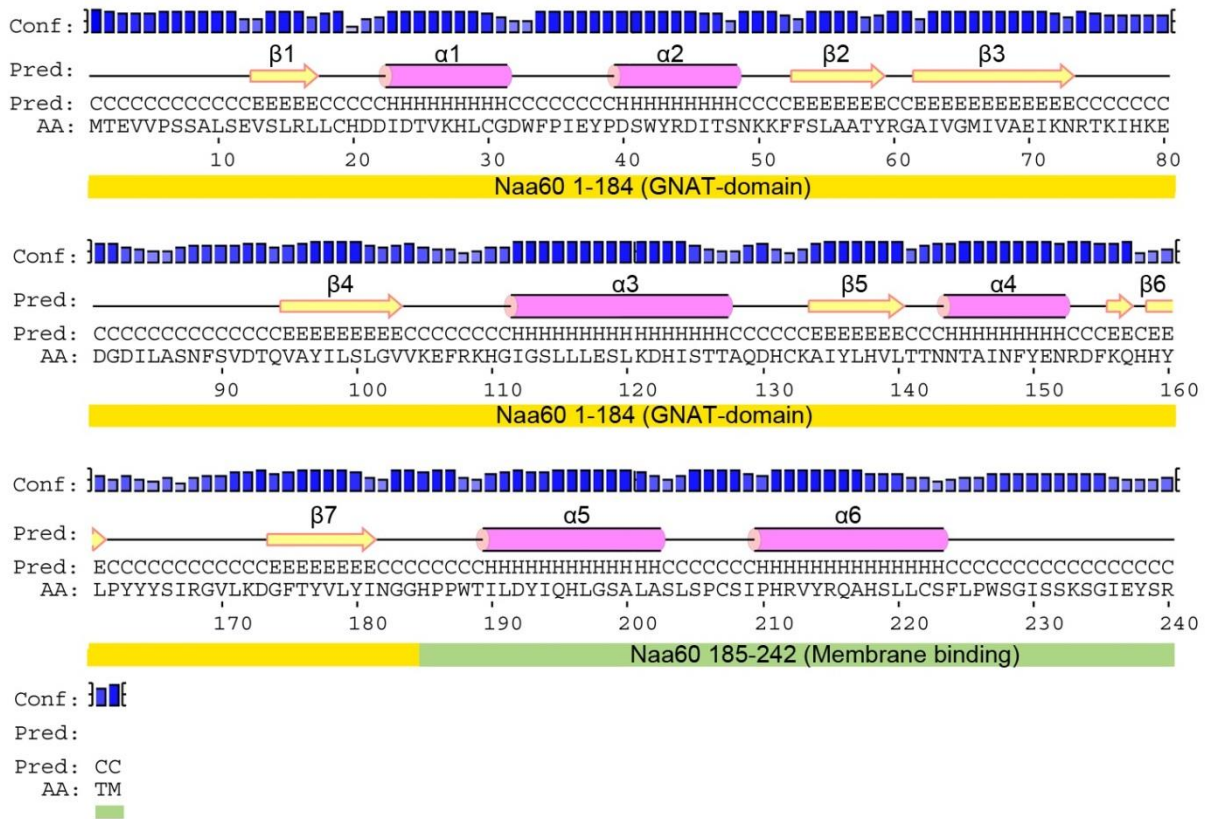
Thomas Arnesen, email: thomas.arnesen@uib.no

Figure S1, related to Figure 1: Secondary structure predictions of full-length hNaa60. Secondary structure prediction was created with the online secondary structure prediction tool PSIPRED (Buchan et al., 2013). Beta-strands are shown as yellow arrows and alpha helixes as pink cylinders. Secondary structure prediction of β1-7 and α1-4 in the GNAT domain of hNaa60 is consistent with the secondary structure elements seen in our X-ray structure. In addition, PSIPRED predicts two alpha helixes in the C-terminal part of the enzyme which we hypothesize to form a domain which interact with organellar membranes (Aksnes et al., 2015).
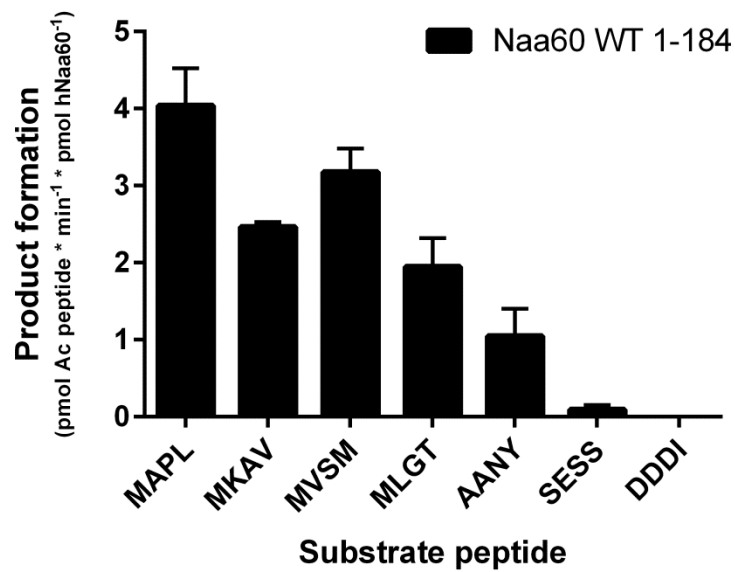
Figure S2, related to Figure 1. Substrate specificity profile for hNaa60 1-184. In order to confirm that hNaa60 1-184 was catalytically active, we tested the activity of this construct towards several different substrate polypeptides. The measured NAT activity of this variant closely resemble the substrate specificity profile of full length hNaa60 *in vivo* (Aksnes et al., 2015; Van Damme et al., 2011)
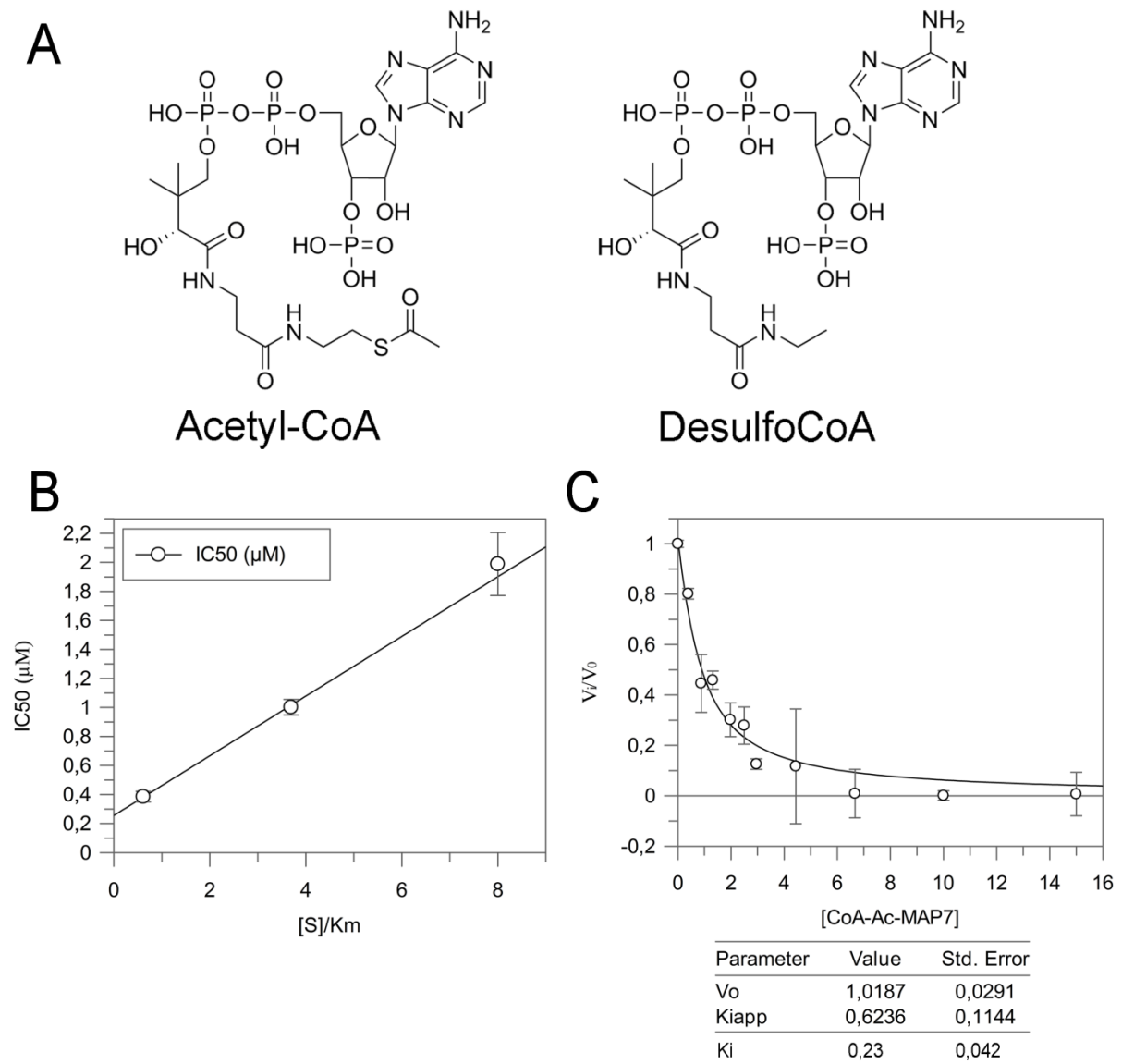
Figure S3, related to Figure 1: Characterization of the CoA-Ac-MAPL$_7$ inhibitor. A) Chemical representation of Acetyl-CoA and desulfoCoA structures. B) [S]/Km plot showing that the inhibitor is a competitive inhibitor for Ac-CoA binding. C) Determination of Kiapp by the Morrison equation. Product formation is shown as fractional activity. The determined Kiapp was inserted into appropriate Cheng-Prusoff equation and Ki was calculated to 0.23
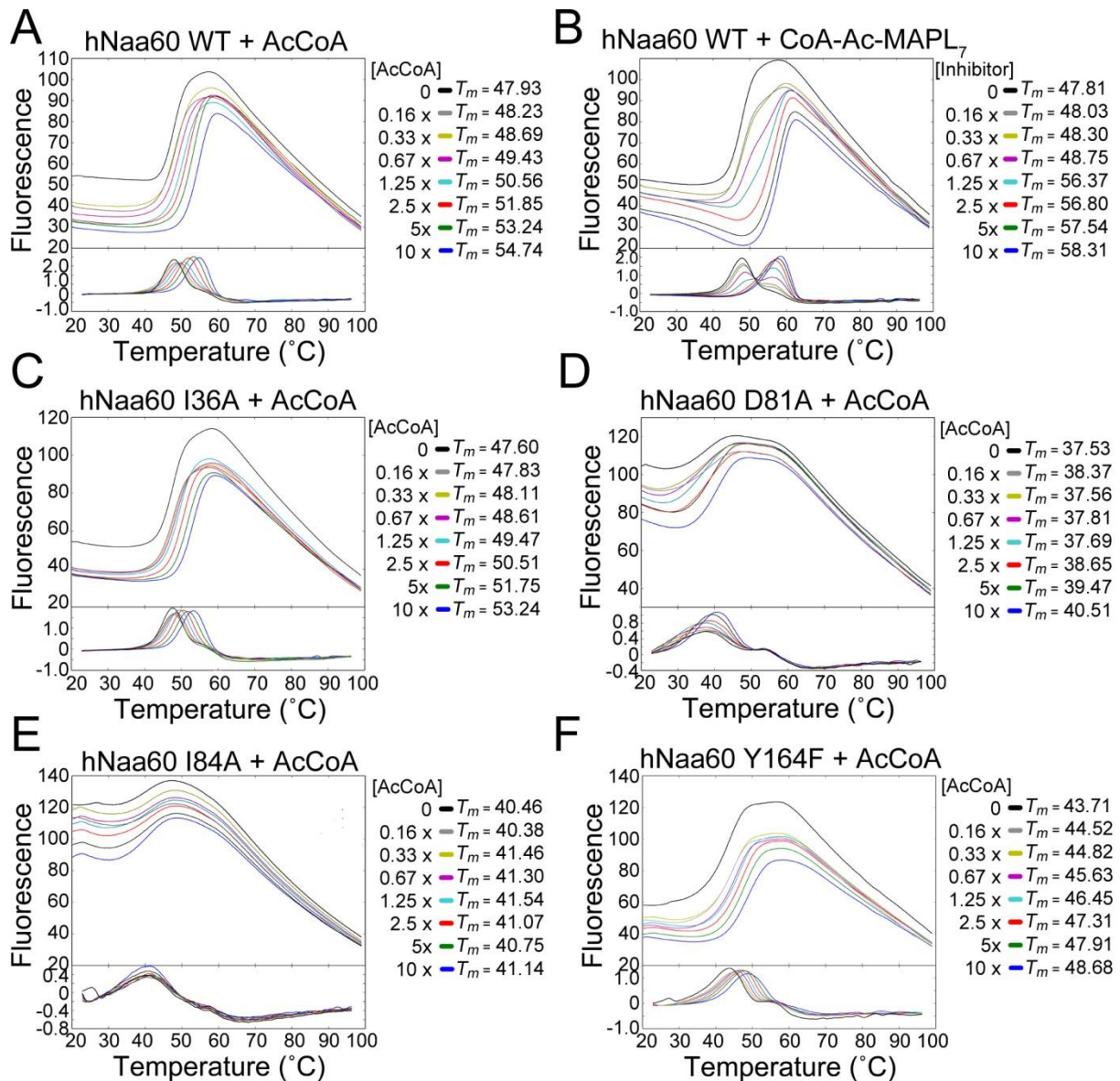
Figure S4, related to Figure 5: Thermal stability of wild-type hNaa60 and mutants. DSF-monitored thermal denaturation of wild-type hNaa60 and mutants. Enzymes were mixed with SYPRO Orange and either CoA or CoA-Ac-MAPL$_7$ at a molar ratio varying from 0.16:1 – 10:1, and the stability was measured at temperatures increasing from 22 to 95°C. All experiments were performed in triplicates, each figure show a representative result from these experiments. A) Thermal denaturation of Apo wild type hNaa60 and wild-type hNaa60 in the presence of increasing concentration of Ac-CoA. B) Thermal denaturation of Apo wild-type hNaa60 and wild-type hNaa60 in the presence of increasing concentration of the bisubstrate analogue CoA-Ac-MAPL$_7$. C) - F) Thermal denaturation of Apo hNaa60 mutants I36A, D81A, I84A and Y164F and hNaa60 mutants I36A, D81A, I84A and Y164F respectively in the presence of increasing concentration of Ac-CoA.
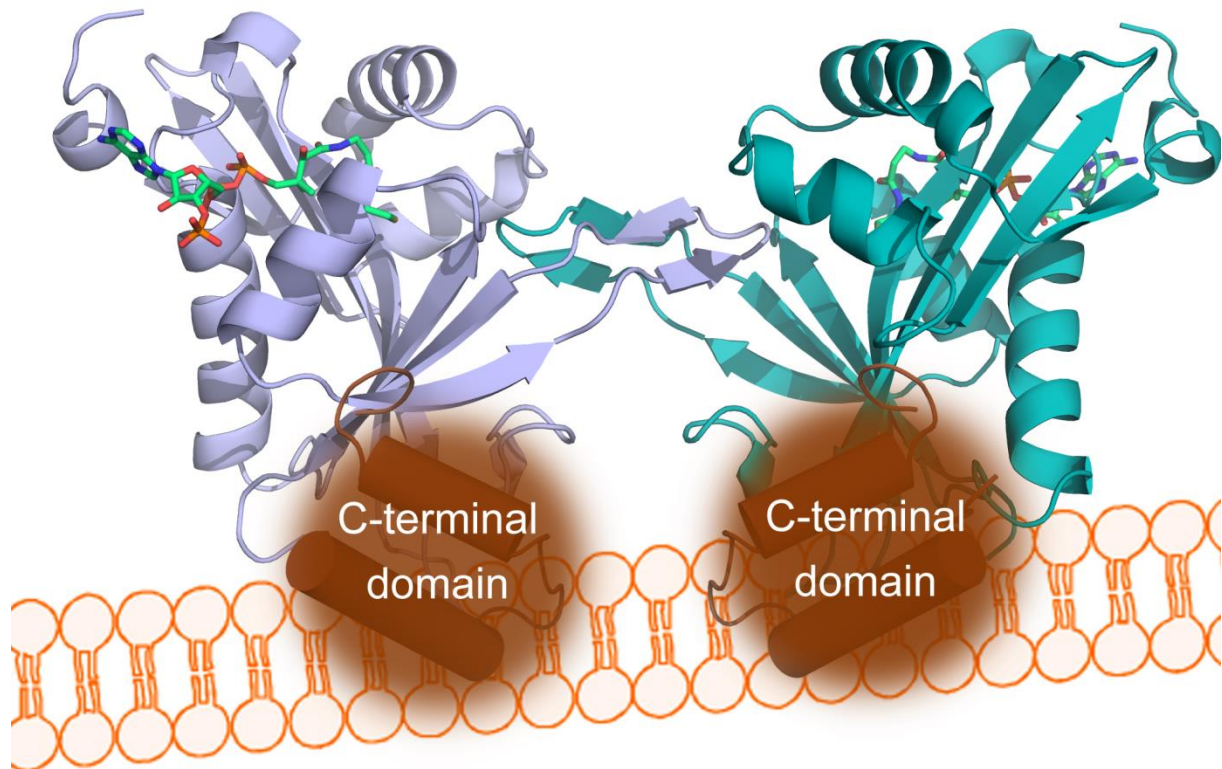
Figure S5, related to Figure 6: Schematic illustration of the potential positioning of the hNaa60 homodimer relative toorganellar membranes. hNaa60 is known to associate with Golgi membranes through a C-terminal domain. In the structure of the hNaa60 homodimer, both C-terminal domains are pointing in the same direction, suggesting that hNaa60 homodimerization also can occur *in vivo* even though hNaa60 is associated with Golgi membranes.
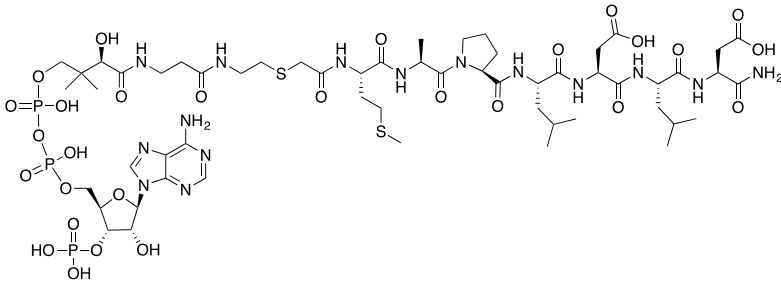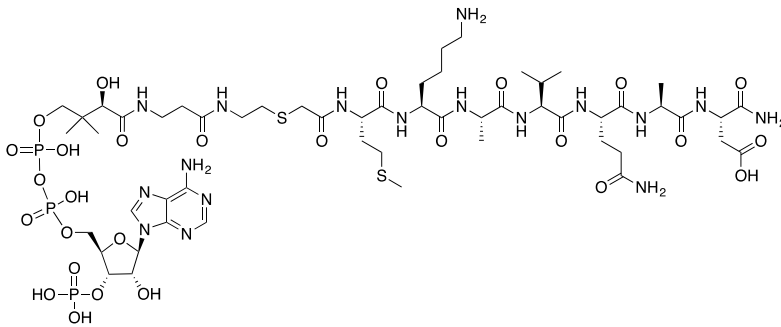
Table S1, related to Figure 4: Catalytic parameters hNaa60 WT and mutants

Enzyme	k_{cat}/K_m (μM^{-1} min^{-1})	k_{cat} (min^{-1})	k_{cat} (normilized to WT)	K_m (μM)	K_m (normilized to WT)
WT	2.27×10^{-2}	3.90 ± 0.20	1	171.40 ± 23.44	1
P35A	0.36×10^{-2}	0.83 ± 0.21	0,21	221.20 ± 42.00	1,29
I36A	0.36×10^{-2}	1.43 ± 0.22	0,37	378.20 ± 133.70	2,21
P35A/I36A/ I167A	ND	ND		ND	
Y38A	ND	ND		ND	
D81A	3.11×10^{-2}	18.84 ± 1.96	3,80	476.00 ± 128.90	2,78
I84A	4.55×10^{-2}	26.32 ± 0.82	6,75	577.00 ± 45.00	3,37
Y97F	ND	ND		ND	
L140A	1.26×10^{-2}	18.40 ± 1.22	4,72	1444.00 ± 144.90	8,42
Y164A	3.21×10^{-2}	15.82 ± 1.80	4,06	490.60 ± 113.30	2,86
Y164F	3.51×10^{-2}	13.00 ± 1.70	3,33	368.90 ± 105.9	2,15
Y165F	0.54×10^{-2}	0.84 ± 0.13	0,22	152.00 ± 66.68	0,89
I167A	0.74×10^{-2}	1.59 ± 0.20	0,41	210.60 ± 68.13	1,23

Table S2, related to Figure 4: Catalytic parameters hNaa50 and hNaa60 towards different substrate polypeptides.

Peptide	k_{cat}/K_m (μM^{-1} min^{-1}) hNaa60	k_{cat}/K_m (μM^{-1} min^{-1}) hNaa50	k_{cat} (min^{-1}) hNaa60	k_{cat} (min^{-1}) hNaa50	K_m (μM) hNaa60	K_m (μM) hNaa50
$MKAV_{24}$	2.27×10^{-2} $\pm 0.32 \times 10^{-2}$	1.40×10^{-2} $\pm 0.42 \times 10^{-2}$	3.90 ± 0.20	3.71 ± 0.33	171.4 ± 23.4	265.3 ± 78.3
$MVSM_{24}$	4.52×10^{-2} $\pm 1.71 \times 10^{-2}$	5.35×10^{-2} $\pm 0.79 \times 10^{-2}$	4.52 ± 0.52	36.32 ± 2.50	100.3 ± 36.2	677.3 ± 88.6
$MAPL_{24}$	4.75×10^{-2} $\pm 1.44 \times 10^{-2}$	2.43×10^{-2} $\pm 1.71 \times 10^{-2}$	4.64 ± 0.41	4.07 ± 0.41	97.5 ± 28.1	167.4 ± 62.3
$MKQY_{24}$	2.41×10^{-2} $\pm 0.88 \times 10^{-2}$	0.45×10^{-2} $\pm 0.25 \times 10^{-2}$	3.21 ± 0.32	1.01 ± 0.20	133.9 ± 46.9	225.6 ± 115.5
$MLGP_{24}$	1.24×10^{-2} $\pm 0.72 \times 10^{-2}$	9.51×10^{-2} $\pm 0.95 \times 10^{-2}$	4.77 ± 1.06	36.70 ± 6.14	383.0 ± 203.3	385.6 ± 151.1

Table S3, related to the experimental procedure of bisubstrate analogue synthesis: Structure and molecular weight of the synthesized bisubstrate analogues.

Structure	m/z Calcd[a] [M-H]$^-$	m/z Found
CoA-Ac-MAPL$_7$		
	1579.5	1579.2
CoA-Ac-MKAV$_7$		
	1567.5	1567.3

a) Based on average isotope distribution.

Supplemental experimantal procedures

hNaa60 cloning, expression and protein purification.

The full-length h*NAA60* gene (1-242) was subcloned into the MCSI of a modified petDUET

vector containing a SUMO protease cleavable N-terminal SUMO tag. By the use of site-

directed mutagenesis, several C-terminal truncation constructs was prepared by introducing

stop codons in the coding sequence of the gene, and constructs were transformed into Rosetta

(DE3)pLys competent cells. The truncated hNaa60 1-184 protein was stable through

purification, and this truncation construct was thus used for all further experiments, except

when noted. The transformed cells were grown to an OD_{600} of 0.6-0.8, cultures were cooled

down to 17 ˚C, protein expression was induced by the addition of isopropyl 1-thio-β-

galactopyranosidase (final concentration of 0.5 mM) and proteins were expressed overnight.

Cells were harvested by centrifugation and lysed by sonication in lysis buffer (25 mM Tris-

HCl (pH 8.0), 1 M NaCl, 10 mM β-mercaptoethanol and 10 μg/ml

phenylmethanesulfonylfluoride or cOmplete EDTA-free protease inhibitor tablets (1 tablet per

50 mL)). Cell lysate was subsequently centrifuged at 17000 x g for 30 minutes, and the

supernatant was collected and passed over nickel resin (Thermo Scientific). The resin was

washed with >25 column volumes of wash buffer (25 mM Tris-HCl (pH 8.0), 250 mM NaCl,

20 mM Imidazole, 10 mM β-mercaptoethanol) and bound proteins were eluted with elution

buffer (25 mM Tris-HCl (pH 8.0), 250 mM NaCl, 300 mM Imidazole, 10 mM β-

mercaptoethanol). The N-terminal SUMO tag was removed by mixing the eluted protein with

SUMO protease (Ulp-1) overnight in a dialysis bag while the protein was dialyzed into a

buffer without imidazole (25 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10 mM β-

mercaptoethanol). In order to isolate hNaa60 from uncut SUMO-hNaa60 fusion proteins, and

the free SUMO tag, the dialyzed sample was subjected to reverse nickel purification. The

flow through was collected, dialyzed into size exclusion buffer (25 mM HEPES (pH 7.5), 300 mM NaCl, 1 mM DTT) and concentrated to a final volume of 1 mL. The concentrated sample was then loaded on a HiLoad Superdex 200 size exclusion chromatography column, and fractions containing hNaa60 1-184 were collected and used for protein crystallization.

For enzymatic studies, SUMO-hNaa60 1-184 was purified by passing the cell supernatant over a 1 mL HisTrap column. The colum was washed with 20 column volumes of wash buffer (25 mM Tris-HCl (pH 8.0), 250 mM NaCl, 20 mM Imidazole, 10 mM β-mercaptoethanol) and bound proteins were eluted with elution buffer (25 mM Tris-HCl (pH 8.0), 300 mM NaCl, 300 mM Imidazole, 10 mM β-mercaptoethanol), upconcentrated and loaded on a HiLoad Superdex 200 size exclusion chromatography column.

Synthesis of bisubstrate analogues

Bisubstrate analogues were prepared by Fmoc-based solid phase peptide synthesis (SPPS) on a Rink amide MBHA resin (100-200 mesh, typically 0.78 mmole/g loading) using glass reaction vessels (Peptides International, Kentucky USA) following a previously described manual SPPS protocol (Narawane et al., 2014). Each amino acid (4 equiv.) was coupled using HCTU (3.9 equiv.) and DIPEA (8 equiv.) in dimethylformamide (DMF) at room temperature for 10 min. Following final Fmoc-deprotection of the N-terminal amino acid, the resin-bound peptides were treated with bromoacetic acid (8 equiv.) diisopropylcarbodiimide (DIC, 8 equiv.) in DMF for 1 h as previously described (Foyn et al., 2013). Final deprotection and cleavage from the solid support was facilitated by treatment of the resin with a mixture of trifluoroacetic acid (TFA), triisopropylsilane (TIS) and water (95:2.5:2.5 v/v, 12.5 mL/g of initial resin used) for 2 h. The resin was removed by filtration and washed with an additional portion of cleavage cocktail (12.5 mL/g of initial resin used). The combined TFA fractions

were evaporated under a stream of nitrogen until approximately 5 mL of the solution remained. Diethyl ether (45 mL) was added and the precipitate was collected by centrifugation (2000 x *g* for 5 min) and washed twice more with diethyl ether. The crude products were dried and purified by reverse-phase high performance liquid chromatography (RP-HPLC) before they were lyophilized. Coenzyme A (CoA) was attached by allowing bromoacetylated peptides to react with CoA (trilithium salt, 2 equiv.) in triethylammonium bicarbonate buffer (1 M, pH 8.0) for 16 h. The final bisubstrate analogs were purified by RP-HPLC and lyophilized before verification by ESI-MS. The bisubstrate analogues were purified to a typical purity of >85%. The structure, theoretical molecular weight and experimental molecular weight can be found in **Table S3**.

In vitro acetylation assays

In the quantitative acetylation assay the following oligopeptides (synthesized by Biogenes) were used:

MKAV$_{24}$ ([H]MKAVQAD RWGRPVGRRRRPVRVYP[OH]) representing Leucine-rich repeat-containing protein 59,

MAPL$_{24}$ ([H]MAPLDLD RWGRPVGRRRRPVRVYP[OH]) representing Protein phosphatase 6,

MKQY$_{24}$ ([H]MKQYQGS RWGRPVGRRRRPVRVYP[OH]) representing Transmembrane protein 222,

MVSM$_{24}$ ([H]MVSMSFK RWGRPVGRRRRPVRVYP[OH]) representing lysosomal-associated transmembrane protein 4A and

MLGP$_{24}$ [H] MLGPEGG RWGRPVGRRRRPVRVYP [OH] representing hnRNP F.

Determination of IC50 values: Purified SUMO-hNaa60 1-184 (500 nM) was mixed with

MAPL$_{24}$ substrate peptide (300 µM), Ac-CoA (300 µM) and inhibitor (0-15 µM) or

desulfoCoA (0-400 µM) in a total of 50 µL at 37°C for 45 min. IC50 values were determined

by fitting the data to Eq. 1 using Grafit 7,

$$\% \, Activity = \frac{100\%}{1 + ([I]/IC50)^s} \qquad \text{(Eq. 1)}$$

where [I] is the concentration of inhibitor and *s* is a slope factor (hill coefficient).

Determination of K_i^{app} for CoA-Ac-MAPL$_7$: Purified SUMO-hNaa60 1-184 (500 nM) was

mixed with MAPL$_{24}$ substrate peptide (300 µM), Ac-CoA (300 µM) and CoA-Ac-MAPL$_7$ (0-

15 µM) in a total of 50 µL at 37°C for 45 min. K_i^{app} was determined by fitting the data to the

Morrisons's equation (Eq. 2) using Grafit 7,

$$\frac{v_i}{v_0} = \frac{\left([E]_T + [I]_T + K_i^{app}\right) - \sqrt{\left([E]_T + [I]_T + K_i^{app}\right)^2 - 4[E]_T[I]_T}}{2[E]_T} \qquad \text{(Eq. 2)}$$

where v_i/v_0 is fractional activity, $[E]_T$ is total enzyme concentration and $[I]_T$ is total inhibitor

concentration.

Determination of modality of CoA-Ac-MAPL$_7$: IC50 values for CoA-Ac-MAPL$_7$ at different

concentrations of Ac-CoA (50, 300, 650 µM). IC50 values were then replotted against $[S]/K_m$.

Given the linear increase in IC$_{50}$ with increasing $[S]/K_m$ the inhibitor was interpreted as a

competitive inhibitor (Copeland, 2005)

Determination of K_i for CoA-Ac-MAPL$_7$: Knowing the inhibitor modality, K_i^{app} was inserted

into the appropriate Cheng-Prusoff equation for competitive inhibitors (Eq. 3)

$$K_i = \frac{K_i^{app}}{1 + ([S]/K_m)} \qquad \text{(Eq. 3)}$$

The standard deviation was determined after creating an approximate function of K_i using

taylor series (linearization) (Eq. 4) where a is the average value of K_i^{app} and b is the average

value of K_m:

$$K_i = \frac{K_i^{app}}{1 + ([S]/K_m)} = \frac{K_i^{app} * K_m}{[S] + K_m} \approx f(a,b) + \frac{b}{b + [S]}\left(K_i^{app} - a\right) + \frac{a[S]}{(b + [S])^2}(K_m - b) \qquad \text{(Eq. 4)}$$

f(a,b) is defined according to Eq. 5:

$$f(a,b) = \frac{ab}{[S] + b} \qquad \text{(Eq. 5)}$$

The approximation of the standard deviation was then calculated using Eq. 6:

$$STD(K_i) \approx \sqrt{\frac{K_m^{\,2}}{(K_m + [S])^2}\left(STD\left(K_i^{app}\right)\right)^2 + \frac{\left(K_i^{app}[S]\right)^2}{(K_m + [S])^4}\left(STD(K_m)\right)^2} \qquad \text{(Eq. 6)}$$

Differential scanning fluoremetry (DSF)

DSF experiments were performed in triplicates and as described previously (Bustad et al., 2012). Purified wild-type hNaa60 and mutants were diluted to a final concentration of 0.1 mg/mL in a 25 mM HEPES (pH 7.5) buffer containing 300 mM NaCl and 1 mM DTT. Enzymes were mixed with 5x SYPRO Orange and added to a 1:2 dilution series of either Ac-CoA or the CoA-Ac-MAPL$_7$ bisubstrate analogue. Thermal denaturation was studied in 8 different conditions; enzyme without any cofactor and enzyme in 7 different concentrations of CoA or CoA-Ac-MAPL$_7$. Samples were heated from 22 to 95°C at a scan rate of 2°C/min and fluorescence was quantified with a data pitch of 0.2°C on a LightCycler 480 Real-Time PCR System from Roche Applied Science (F. Hoffmann–La Roche Ltd., Basel, Switzerland).

References

Aksnes, H., Van Damme, P., Goris, M., Starheim, K.K., Marie, M., Stove, S.I., Hoel, C., Kalvik, T.V., Hole, K., Glomnes, N.*, et al.* (2015). An organellar nalpha-acetyltransferase, naa60, acetylates cytosolic N termini of transmembrane proteins and maintains Golgi integrity. Cell reports *10*, 1362-1374.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., and Jones, D.T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res *41*, W349-357.

Bustad, H.J., Skjaerven, L., Ying, M., Halskau, O., Baumann, A., Rodriguez-Larrea, D., Costas, M., Underhaug, J., Sanchez-Ruiz, J.M., and Martinez, A. (2012). The peripheral binding of 14-3-3gamma to membranes involves isoform-specific histidine residues. PloS one *7*, e49671.

Copeland, R.A. (2005). Evaluation of enzyme inhibitors in drug discovery : a guide for medicinal chemists and pharmacologists (Hoboken, N.J.: Wiley-Interscience).

Foyn, H., Jones, J.E., Lewallen, D., Narawane, R., Varhaug, J.E., Thompson, P.R., and Arnesen, T. (2013). Design, synthesis, and kinetic characterization of protein N-terminal acetyltransferase inhibitors. ACS chemical biology *8*, 1121-1127.

Narawane, S., Budnjo, A., Grauffel, C., Haug, B.E., and Reuter, N. (2014). In silico design, synthesis, and assays of specific substrates for proteinase 3: influence of fluorogenic and charged groups. Journal of medicinal chemistry *57*, 1111-1115.

Van Damme, P., Hole, K., Pimenta-Marques, A., Helsens, K., Vandekerckhove, J., Martinho, R.G., Gevaert, K., and Arnesen, T. (2011). NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. PLoS genetics *7*, e1002169.