

Supplementary Information

Divergent architecture of the heterotrimeric NatC complex explains N-terminal acetylation of cognate substrates

Grunwald et al.

Content

Supplementary Table 1: Primers sequences for vector construction and site-directed mutagenesis

Supplementary Fig. 1: Sequence alignment of Naa30 and Naa38 orthologs

Supplementary Fig. 2: Sequence alignment of Naa35 orthologs

Supplementary Fig. 3: MALDI-MS analysis of Naa38 and purification of NatC

Supplementary Fig. 4: Structural comparison of NatC subunits Naa30 and Naa38

Supplementary Fig. 5: NatC subunit interfaces

Supplementary Fig. 6: Michaelis-Menten kinetics of NatC for various substrate peptides

Supplementary Fig. 7: Michaelis-Menten kinetics for NatC mutants

Supplementary Fig. 8: Electron density maps for the active site of NatC

Supplementary Fig. 9: Comparisons of the peptide-binding surfaces in NAT orthologs

Supplementary Fig. 10: Peptide ligand-induced conformational changes

Supplementary Fig. 11: Replicates of the NatC/ribosome co-sedimentation assay

Supplementary Fig. 12: Replicates of the yeast dilution spot assay

Supplementary Fig. 13: Models of NatC on the ribosomal surface

Supplementary References

Supplementary Table 1 Primers sequences for vector construction and site-directed mutagenesis

Primer name (cloning strategy)	Sequence (5' to 3')
For construction of NatC bacterial expression vectors	
RBS-Naa30 Fw (PvuI)	GGCCGATCGAAGGAGATATACCATGGAAATAGTGTACAAGCCATTG
Naa30ΔC17 Rv (XhoI)	GGCCTCGAGTCATTAAGTCAACGGCAATATCAGTTTAAAG
HRV3C-Naa35 Fw (BamHI)	GGCGGATCCGCTGGAAGTTCTGTTCCAGGGGCCCCATGGAAGTAGACAGT ATATTAG
HRV3C-Naa35ΔN17 Fw (BamHI)	GGCGGATCCGCTGGAAGTTCTGTTCCAGGGGCCCCAATTGGTTGACGTC ACAAG
HRV3C-Naa35ΔN44 Fw (BamHI)	GGCGGATCCGCTGGAAGTTCTGTTCCAGGGGCCCCGATTGTTGAGGGT ACAC
HRV3C-FLAG-Naa35 Fw (BamHI)	GGCGGATCCGCTGGAAGTTCTGTTCCAGGGGCCCCGATTATAAAGATGAT GATGATAAAATGGAAGTAGACAGTATATTAG
Naa35 Rv (NotI)	GGCGCGGCCGCTTATTTATAGCGGTCTTGC
Naa38 Fw (NdeI)	GGCCATATGGACATCTTGAAACTGTC
Naa38 Rv (MfeI)	GGCCAATTGCTAAACAATATTAGCCATCAATTC
Naa38ΔC11 Fw (site-directed mutagenesis)	CTGCAGGAGCTTACTTAGCAATTGGATATC
Naa38ΔC11 Rv (site-directed mutagenesis)	GATATCCAATTGCTAAGTAAGCTCCTGCAG
For construction of FLAG-Naa35 yeast expression vectors	
pRS416_Naa35 Fw (Gibson ¹)	TCGAATTCCTGCAGCCCGGGGCCAGCAAATACTATGTTTCACTGC
pRS416_Naa35 Rv (Gibson)	GCGGCCGCTCTAGAAGTAGTGCGGGATTTTCCCTTCGGTTAC
FLAG-Naa35 Fw (Gibson)	ATGGATTATAAAGATGATGATGATAAAATGGAAGTAGACAGTATATTAGG
FLAG-Naa35 Rv (Gibson)	TTTATCATCATCATCTTTATAATCCATTCTCGGAAACGTACCAGC
For generation of NatC-Naa30 mutations by site-directed mutagenesis	
L27A_Fw	GTATTAATAAACTAATAGATGCGGATGCATCAGAGCCGTA CTCTCGATATAC
L27A_Rv	GTATATCGAGTACGGCTCTGATGCATCCGCATCTATTAGTTTTTAATAC
S28A_Fw	CTAATAGATGCGGATCTAGCAGAGCCGTA CTCTCGATATAC
S28A_Rv	GTATATCGAGTACGGCTCTGCTAGATCCGCATCTATTAG
E29A_Fw	GGATCTATCAGCGCCGTA CTCTCG
E29A_Rv	CGAGTACGGCGCTGATAGATCC
E29Q_Fw	GGATCTATCACAGCCGTA CTCTCG
E29Q_Rv	CGAGTACGGCTGTGATAGATCC
Y31F_Fw	GGATCTATCAGAGCCGTTCTCGATATACGTATATAGGTATTTTC
Y31F_Rv	GAAATACCTATATACGTATATCGAGAACGGCTCTGATAGATCC
Y80A_Fw	CGTGAGACTGAGAGGAGCTATAGGAATGCTAGCCGTAG
Y80A_Rv	CTACGGCTAGCATTCCCTATAGCTCCTCTCAGTCTCACG
Y80F_Fw	GACTGAGAGGATTTATAGGAATGC
Y80F_Rv	GCATTCCCTATAAATCCTCTCAGTC
E118A_Fw	GATCATGTTAGCAACAGAGGTGG
E118A_Rv	CCACCTCTGTTGCTAACATGATC
E118Q_Fw	GATCATGTTACAAACAGAGGTGG
E118Q_Rv	CCACCTCTGTTTGTAACATGATC
E120A_Fw	GTTAGAAACAGCGGTGGAAAATTC
E120A_Rv	GAATTTTCCACCGCTGTTTCTAAC
E120Q_Fw	GTTAGAAACACAGGTGGAAAATTC
E120Q_Rv	GAATTTTCCACCTGTGTTTCTAAC
Y130A_Fw	GCGGCTCTAACTTAGCTGAAGGAATGGGTTTCATC
Y130A_Rv	GATGAAACCCATTCTTCAGCTAAGTTTAGAGCCGC
Y130F_Fw	GCGGCTCTAACTTATTTGAAGGAATGGGTTTCATC
Y130F_Rv	GATGAAACCCATTCTTCAAATAAGTTTAGAGCCGC
Y145F_Fw	GAAACGGATGTTCCGCTACTTCTTGAATGAAGGGGACG
Y145F_Rv	CGTCCCTTCATTCAAGAAGTAGCGGAACATCCGTTTC

For generation of NatC-Naa35 mutations by site-directed mutagenesis

F47A_Fw	GATCCCAGGTTTGATTTGGCCGAGGGTACACATTCTTTAG
F47A_Rv	CTAAAGAATGTGTACCCTCGGCCAAATCAAACCTGGGATC
K59A_Fw	CTTTAGAGGTCAACAACTCCGCATTAGACTCCAGTCTGATAG
K59A_Rv	CTATCAGACTGGAGTCTAATGCGGAGTTGTTGACCTCTAAAG
Naa35-EPR1 ^a Fw	GCTGCAAACGTGGCACTGGCAGGATATATAGGAATGCTAGC
Naa35-EPR1 Rv	TGCCAGTGCCACGTTTGCAGCCGGGTCCATCTTACACACTAT
Naa35-EPR2-Tip1 ^b Fw	GAACGCAGCACTTGCAGCATTGAAAGCTGGTGAAAAG
Naa35-EPR2-Tip1 Rv	GCTGCAAGTGCTGCGTTCATGGAATGAATGGC
Naa35-EPR2-Tip2 ^c Fw	GACCAATTGGCGCTAAAGTACGCGTTTGCTATGGATAATGAAATG
Naa35-EPR2-Tip2 Rv	CTTTAGCGCCAATTGGTCGGCCTTTTACCAGCTTTCAAT
Naa35-EPR3 ^d -a Fw	CTTTATCGAAACTGAAGTGGCCAATGTCGTTAGTTCTATTG
Naa35-EPR3-a Rv	CAATAGAACTAACGACATTGGCCACTTCAGTTTCGATAAAG
Naa35-EPR3-b Fw	GCATTATCAGTCAACTCAGCAGTTATCGTCAACACTTTG
Naa35-EPR3-b_Rv	CTGAGTTGACTGATAATGCTGCAGAAGTATGTTTTAGTTTACAATAG
Naa35-EPR4 ^e Fw	GCGAACATTGCAAACGCGGACTCCACGAATACAAG
Naa35-EPR4 Rv	CGCGTTTGCAATGTTGCTTTCAAAGTGTGACGATAAC

RBS, ribosome-binding site HRV3C, Human Rhinovirus (HRV) 3C protease cleavage site (LEVLFQGP); FLAG, FLAG-epitope (DYKDDDDK); EPR, Electropositive region

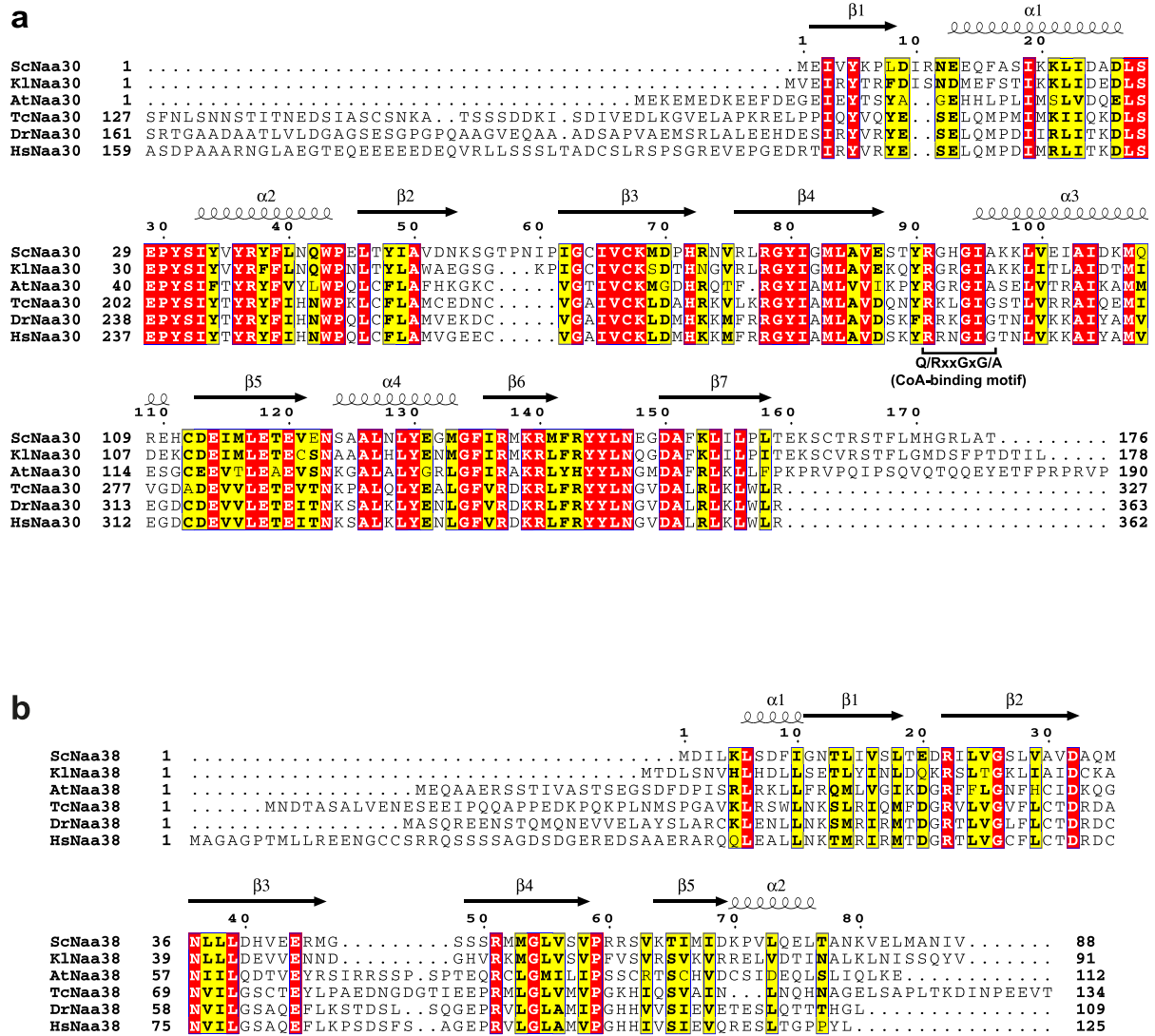
^aThe Naa35-EPR1 primer pair generates Naa30 point mutations H72A, R73A, R76A and R78A.

^bThe Naa35-EPR2-Tip1 primer pair generates Naa35 point mutations K500A, K501A, K503A and K504A.

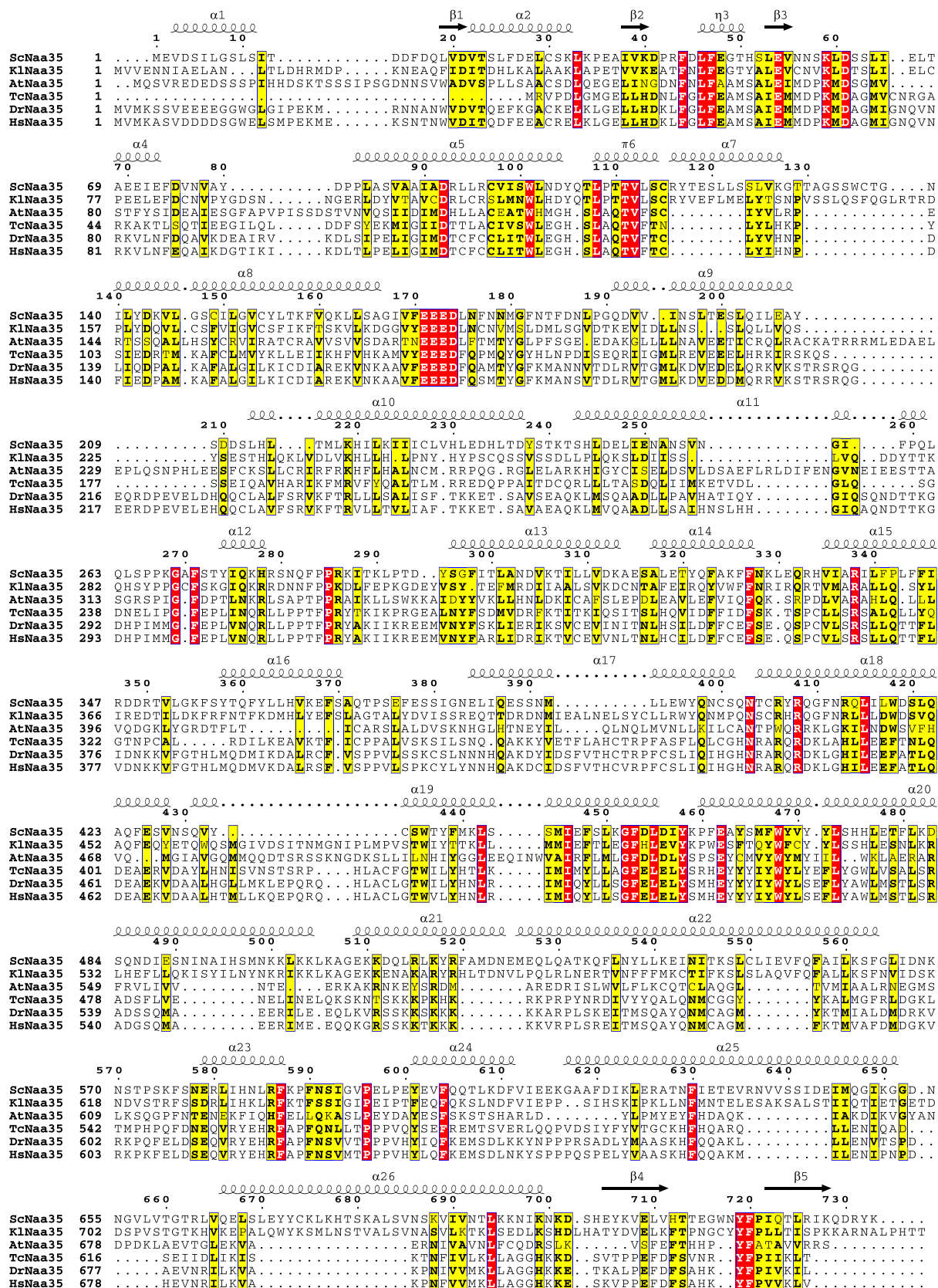
^cThe Naa35-EPR2-Tip2 primer pair generates Naa35 point mutations K511A, R515A, and R519A.

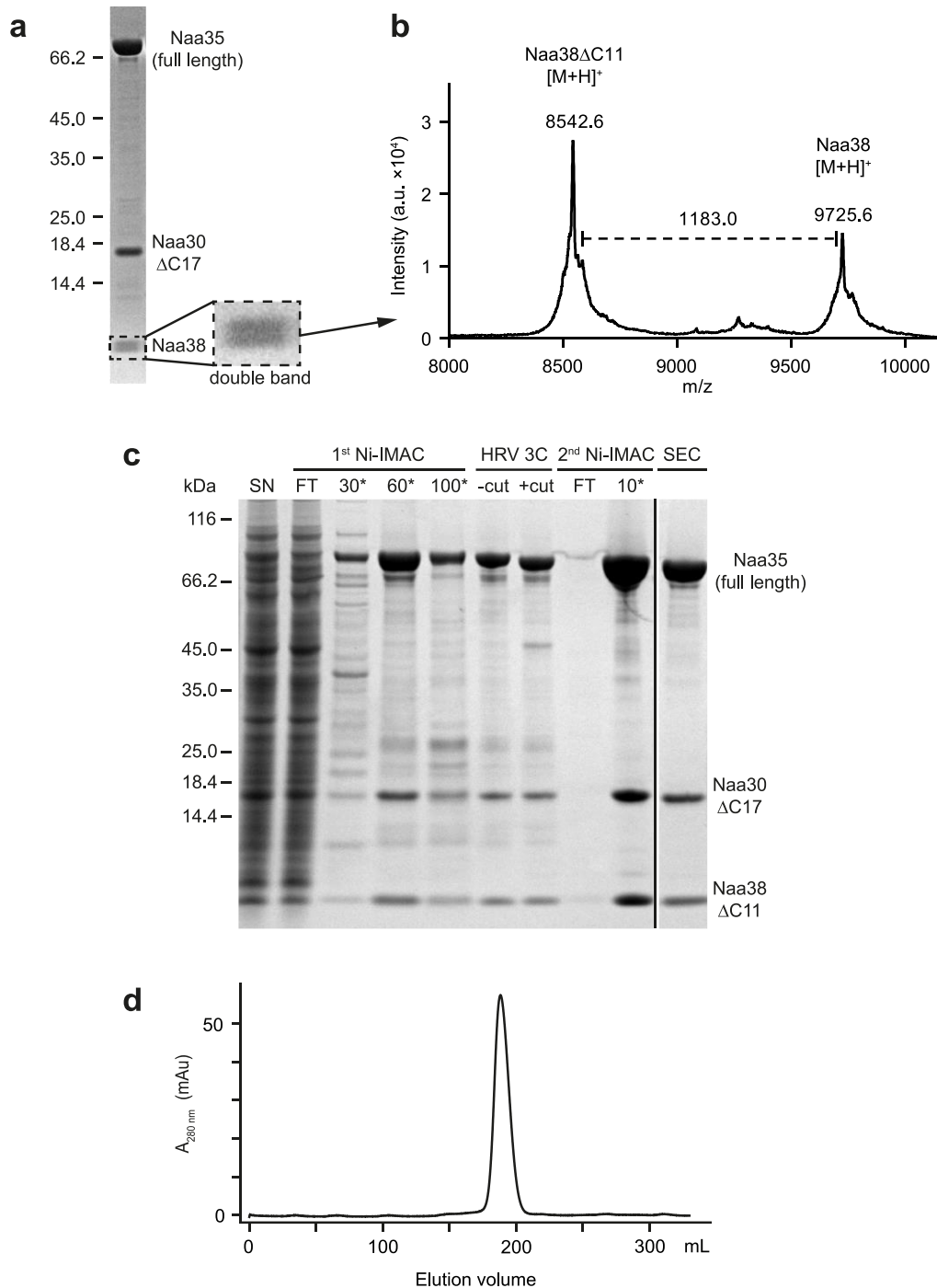
^dThe two Naa35-EPR3 primer pairs were used successively to generates Naa35 point mutations R636A, K681A, and K688A.

^eThe Naa34-EPR4 primer pair generates Naa35 point mutations K696A, K699A and K701A.

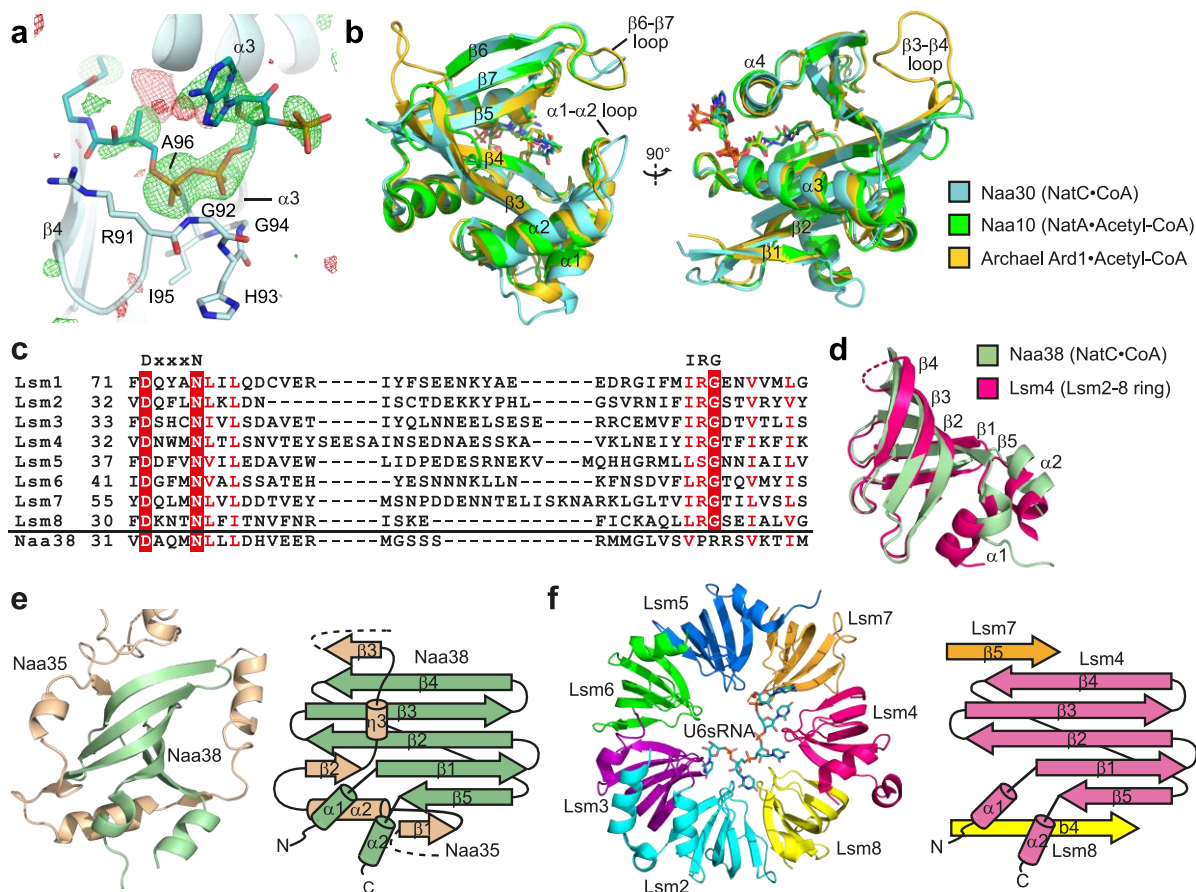


Supplementary Fig. 1 Sequence alignment of Naa30 and Naa38 orthologs. a, b Completely conserved residues are highlighted with red boxes, highly conserved with yellow boxes. Secondary structures assignments are shown on top. The following sequences of Naa30 and Naa38 were aligned: Sc, *Saccharomyces cerevisiae* (accession numbers NP_015376.1 and NP_009948.1); Kl, *Kluyveromyces lactis* (XP_452068.1 and XP_451488.1), At, *Arabidopsis thaliana* (NP_181348.1 and NP_680719.1); Tc, *Tribolium castaneum* (XP_975323.1 and XP_008192183.1); Dr, *Danio rerio* (NP_001129721.2 and NP_001153763.1) and Hs, *Homo sapiens* (NP_001011713.2 and NP_001307854.1).

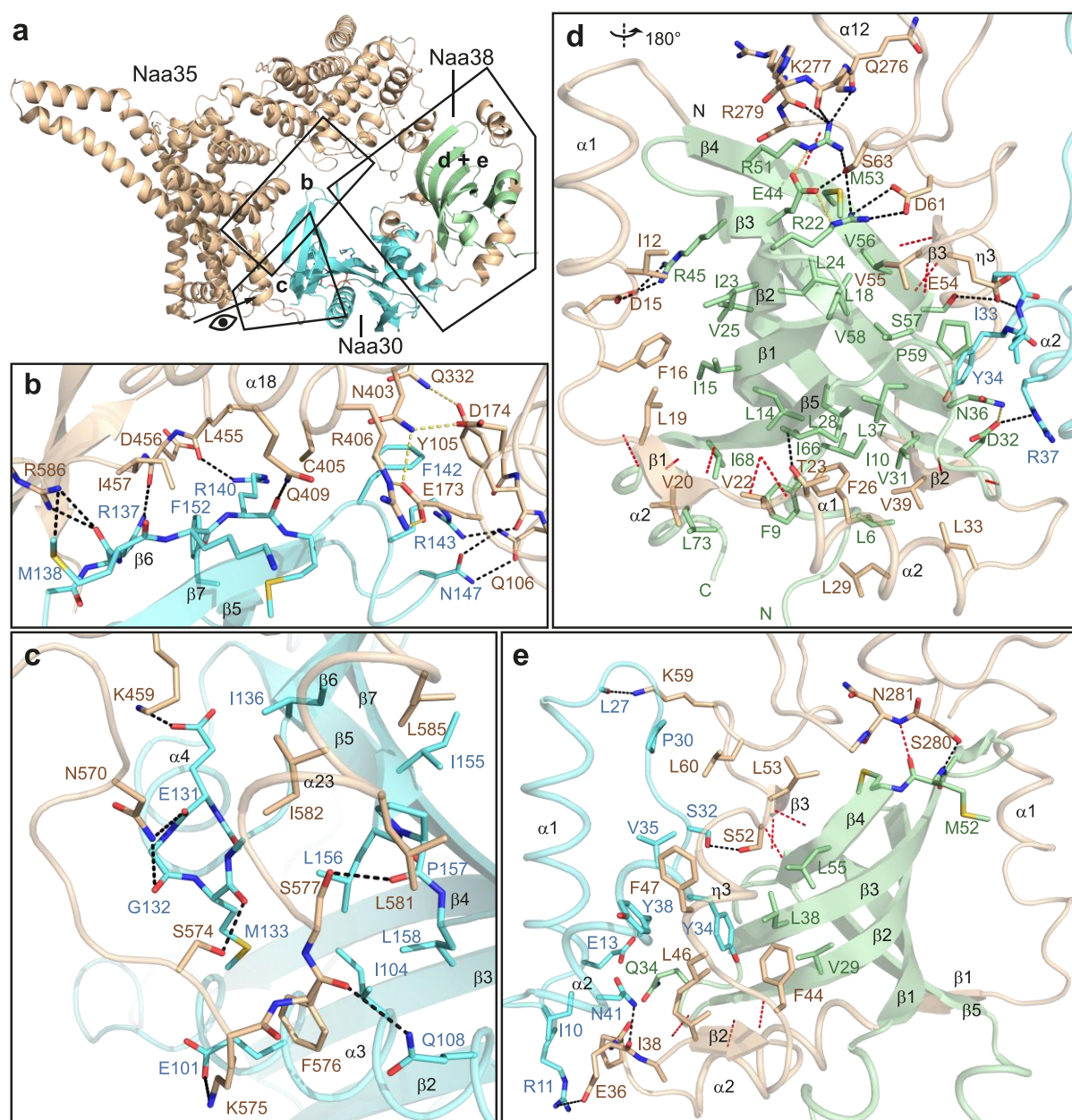




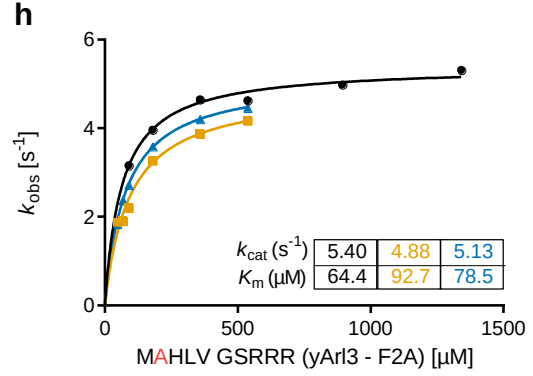
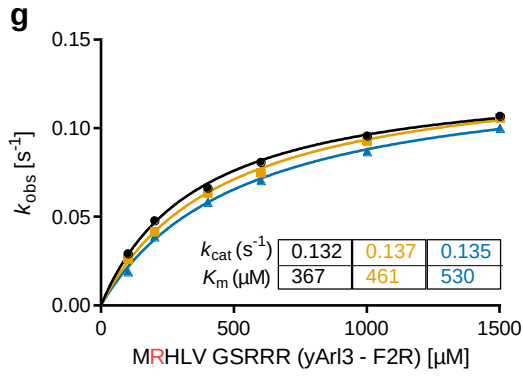
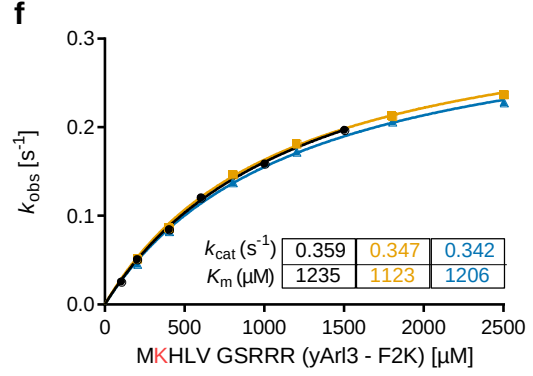
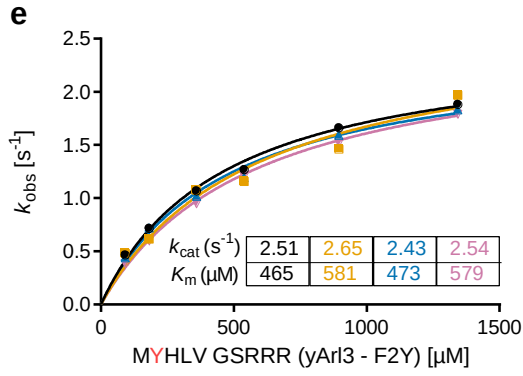
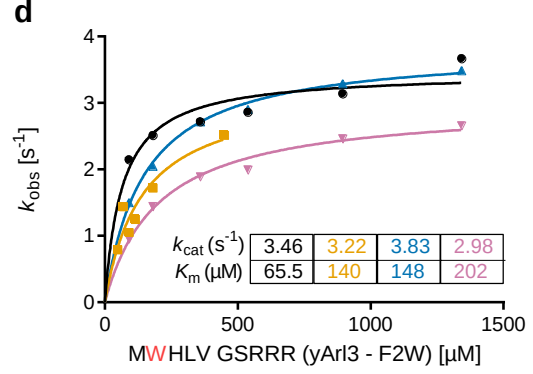
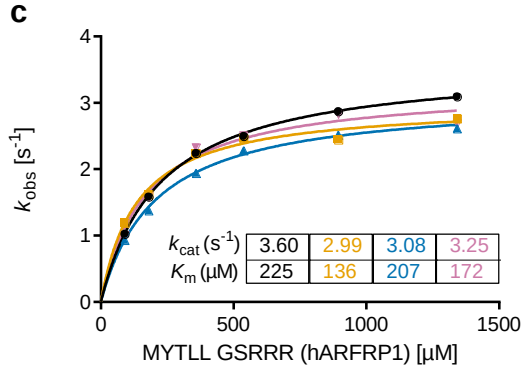
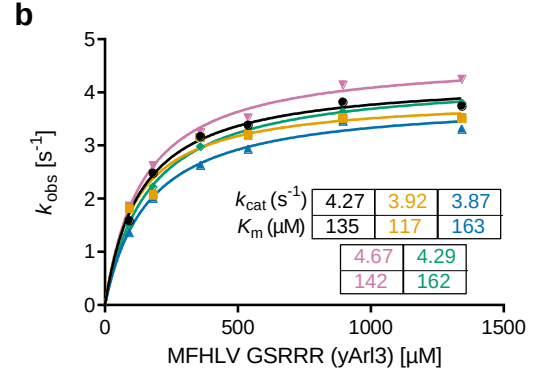
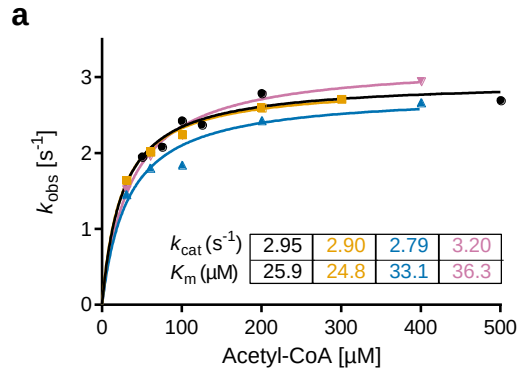
Supplementary Fig. 3 MALDI-MS analysis of Naa38 and purification of NatC. **a** During an initial purification of a NatC complex containing subunits Naa30ΔC17 (residues 1–159), full-length Naa35 (residues 1–733) and full-length Naa38 (residues 1–88), a partial proteolytic degradation of the small subunit Naa38 was observed. The purification of this NatC construct was performed only once. **b** MALDI-MS analysis of the protein purification shown in subfigure a. The right peak corresponds to full-length Naa38 (theoretical MW = 9724.6 Da) and the left peak was attributed to a partial proteolytic degradation product of Naa38, which lacks eleven C-terminal residues (Naa38ΔC11, theoretical MW = 8541.1 Da). **c** Samples collected from a purification of NatC WT containing subunits Naa30ΔC17, Naa38ΔC11 and full-length Naa35. The latter is preceded by an N-terminal 6×His-tag, followed by a human rhinovirus 3C cleavage site. Purifications of NatC WT were repeated three times with identical results. Ni-IMAC, Ni²⁺-immobilized metal ion affinity chromatography; SN, supernatant; FT, flow-through; *, Imidazole concentrations during wash and elution steps; HRV 3C, human rhinovirus 3C protease cleavage with fractions before (-cut) and after (+cut) digest; SEC, size exclusion chromatography. The juxtaposed SEC lane is from the same gel. **d** Size exclusion chromatography profile for the NatC complex, run on a S200, 26/600 column. Uncropped images for panels a and c are provided in the Source Data file.

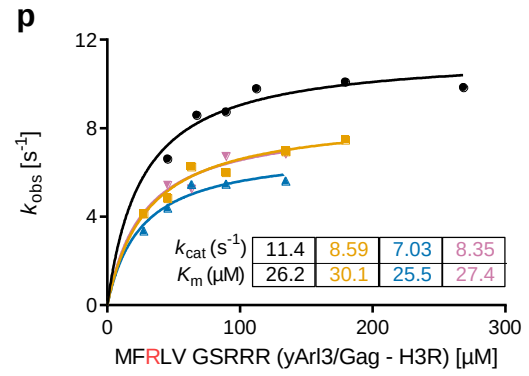
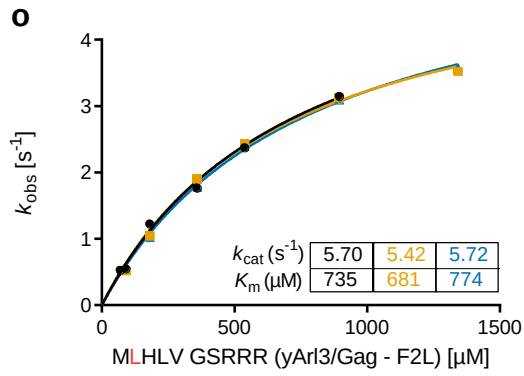
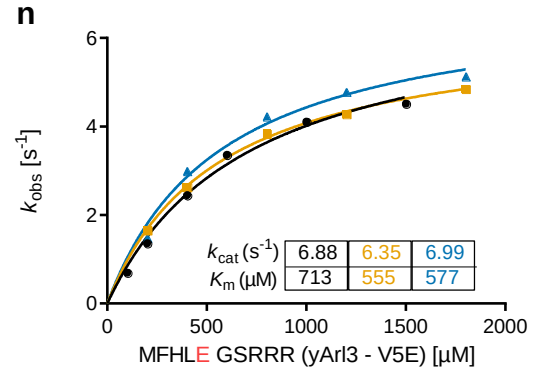
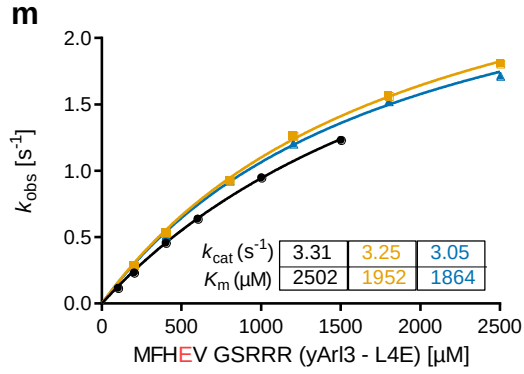
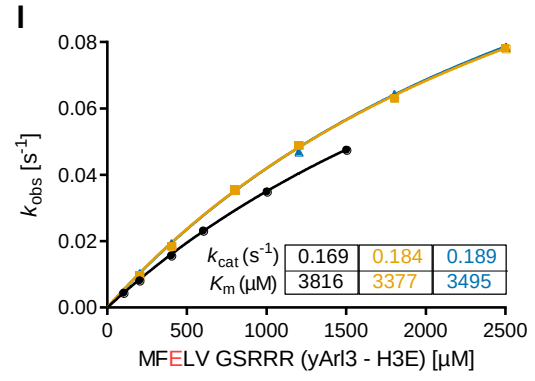
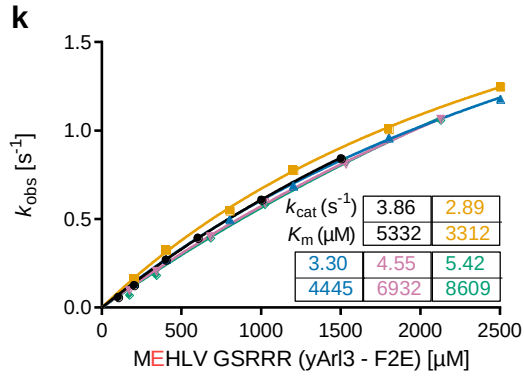
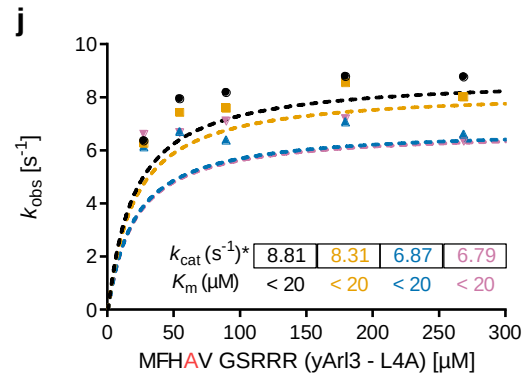
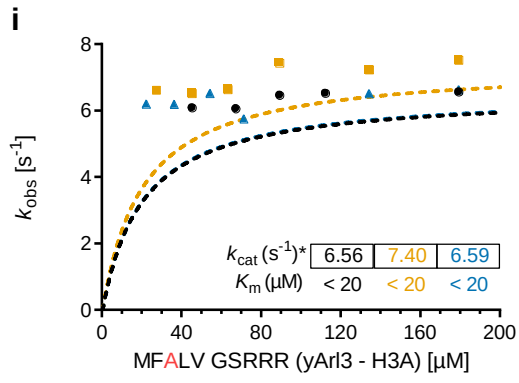


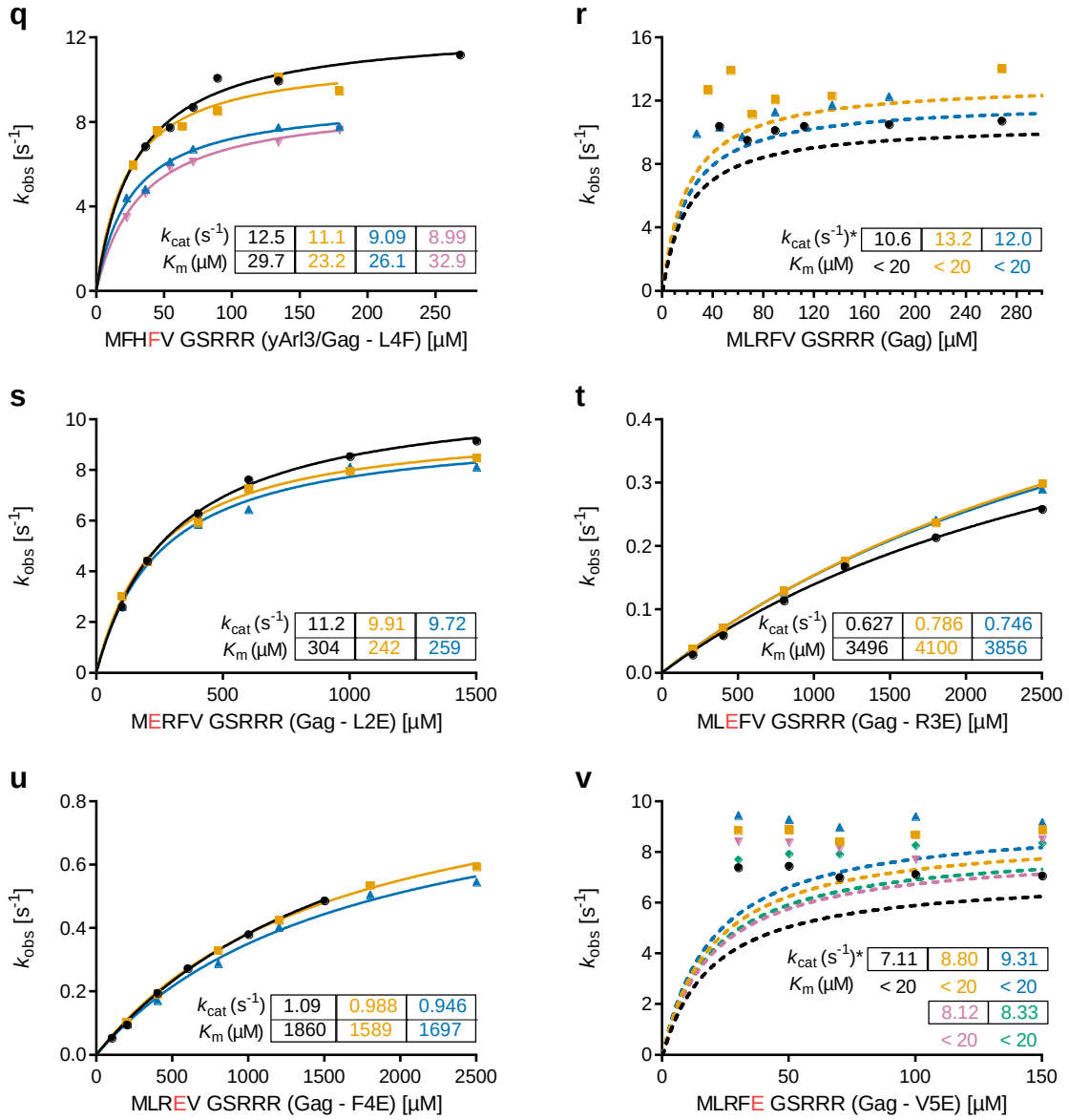
Supplementary Fig. 4 Structural comparison of NatC subunits Naa30 and Naa38. **a** CoA (teal stick format) bound to the Naa30 subunit of the NatC•CoA structure. A simulated annealing omit map contoured to 3.0 σ shows positive (green) and negative (red) difference electron density. Residues within the conserved CoA-binding motif RxxGxG/A are shown as sticks. **b** Superposition of Naa30 with its closest structural homologs. The RMSD between Naa30 and the archaeal Ard1 NAT ortholog from *Sulfolobus solfataricus* (4LX9) is 1.6 Å (over 149 C α atoms) and between Naa30 and the Naa10 subunit of the *Schizosaccharomyces pombe* NatA complex (4KVO) complex 1.7 Å (over 150 C α atoms). **c** Sequence alignment of the *S. cerevisiae* Naa38 and Lsm1-8 proteins showing the region between the conserved DxxxN and IRG motifs. Completely conserved residues are highlighted with red boxes, highly conserved residues are colored in red. **d** Superposition of Naa38 with its closest structural homolog Lsm4 (RMSD of 1.8 Å, over 69 C α atoms) of the heptameric Lsm2-8 ring (4M7A). **e, f** Comparison of the β -sheet topology of the Naa35-Naa38 β -sheet (**e**) with the β -sheet topology of Lsm4 in contact with its neighboring Lsm units (Lsm7 and Lsm8) within the Lsm2-8 ring (**f**).



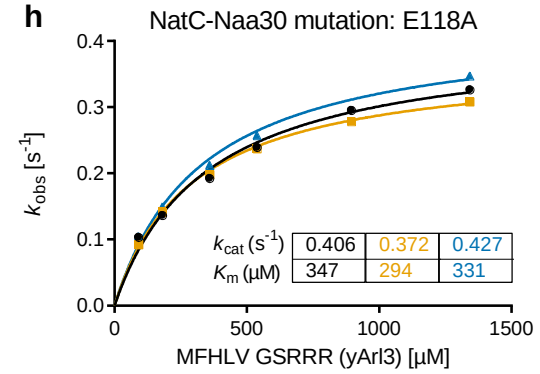
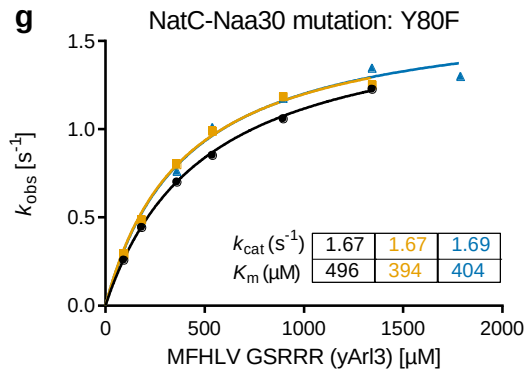
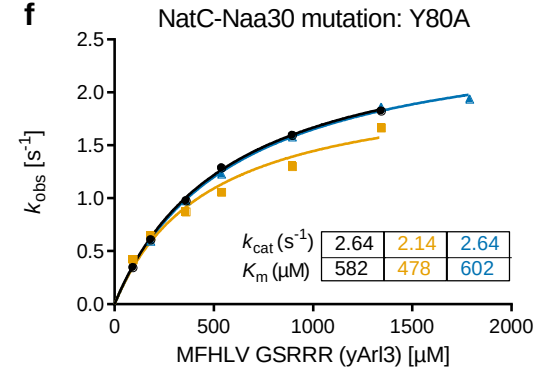
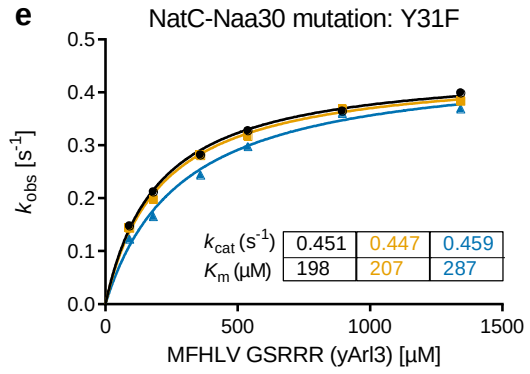
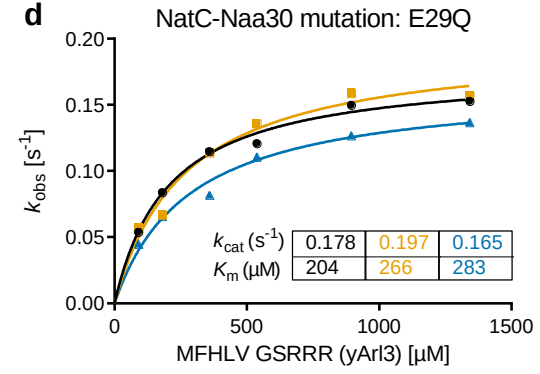
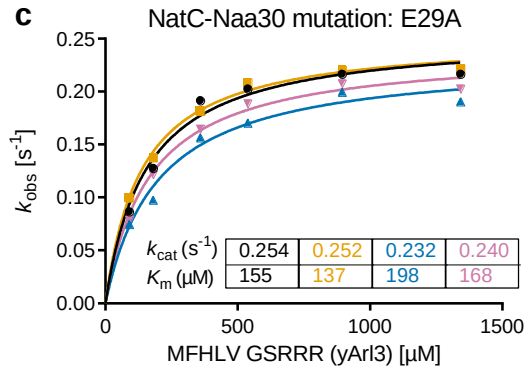
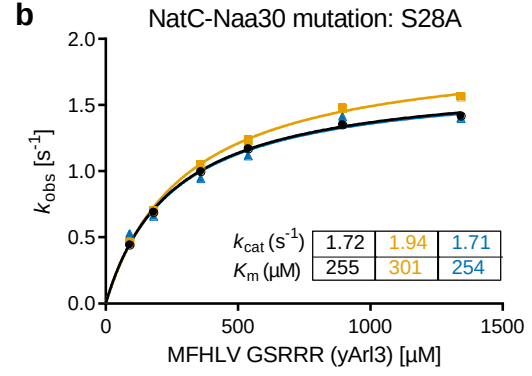
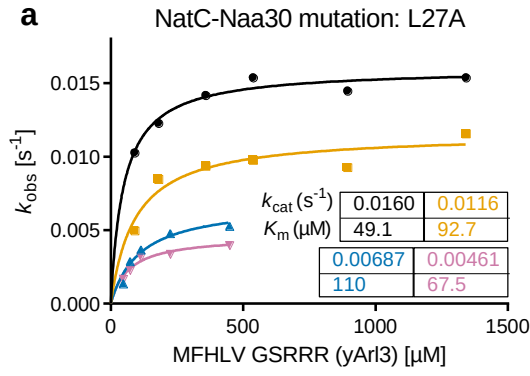
Supplementary Fig. 5 NatC subunit interfaces. **a** Cartoon representation of the NatC•CoA structure. **b**, **c** Magnified views of the Naa30-Naa35 interface, which is dominated by hydrogen-bonding interactions (dashed lines, colored in black) and involves the C-terminal half (residues 101–158) of the Naa30 subunit. Key interface residues are shown in stick format. **d**, **e** Magnified views of the interfaces between the Naa38 subunit, the Naa35 N-terminus (residues 1–64), a short stretch around Naa35-helix α 12 (residues 276–282) and the N-terminal half of the Naa30 subunit (residues 10–41). The interface between Naa38 and Naa35 is dominated by hydrophobic side chain interactions. Additionally, the β -sheet of Naa38 is extended by three short β -strands (β 1– β 3) from the Naa35 N-terminus (backbone-backbone interactions are colored in red).

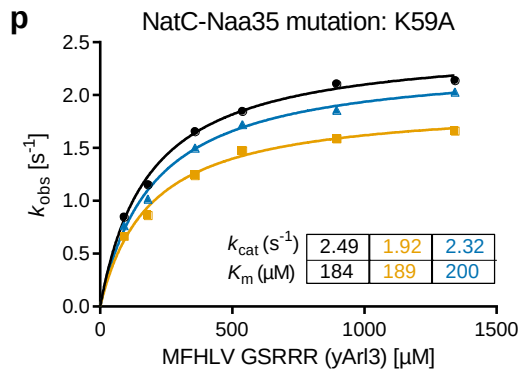
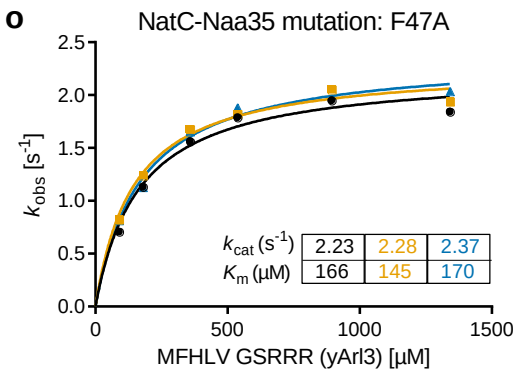
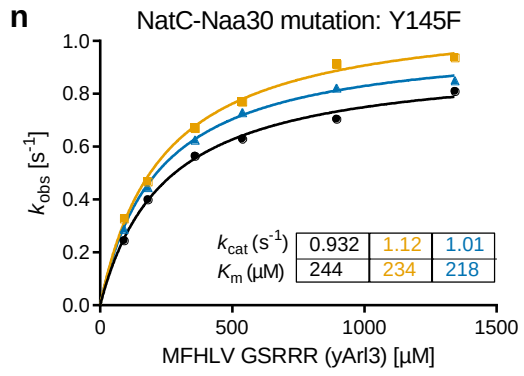
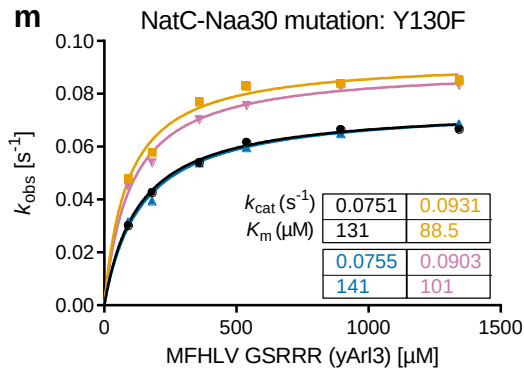
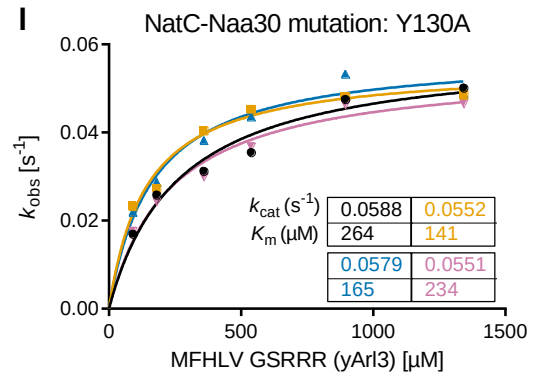
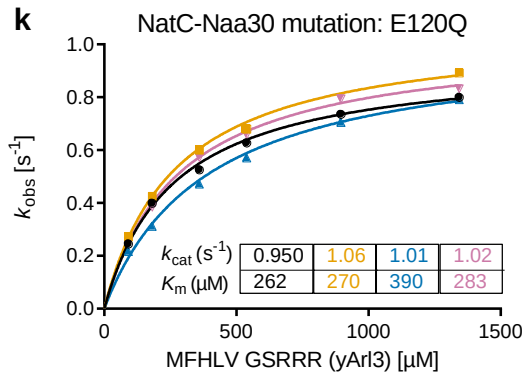
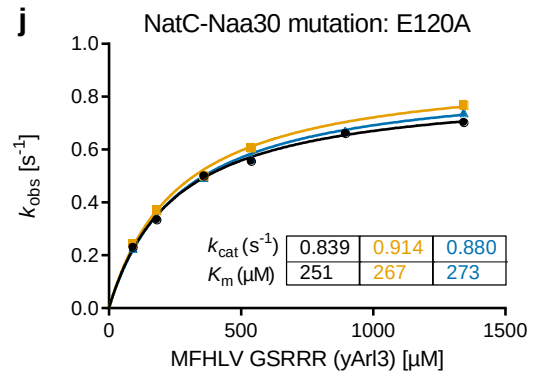
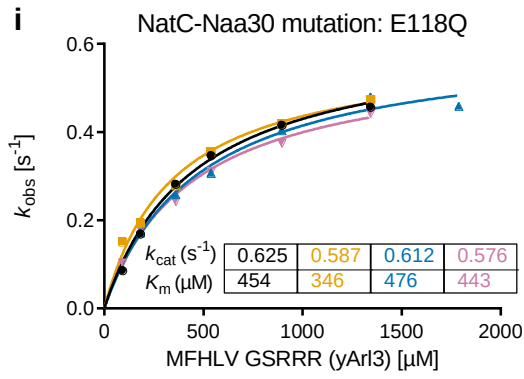


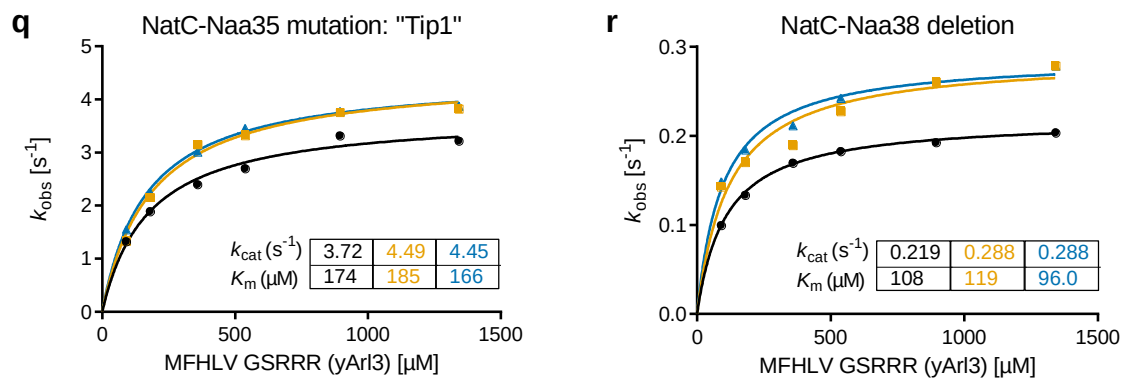




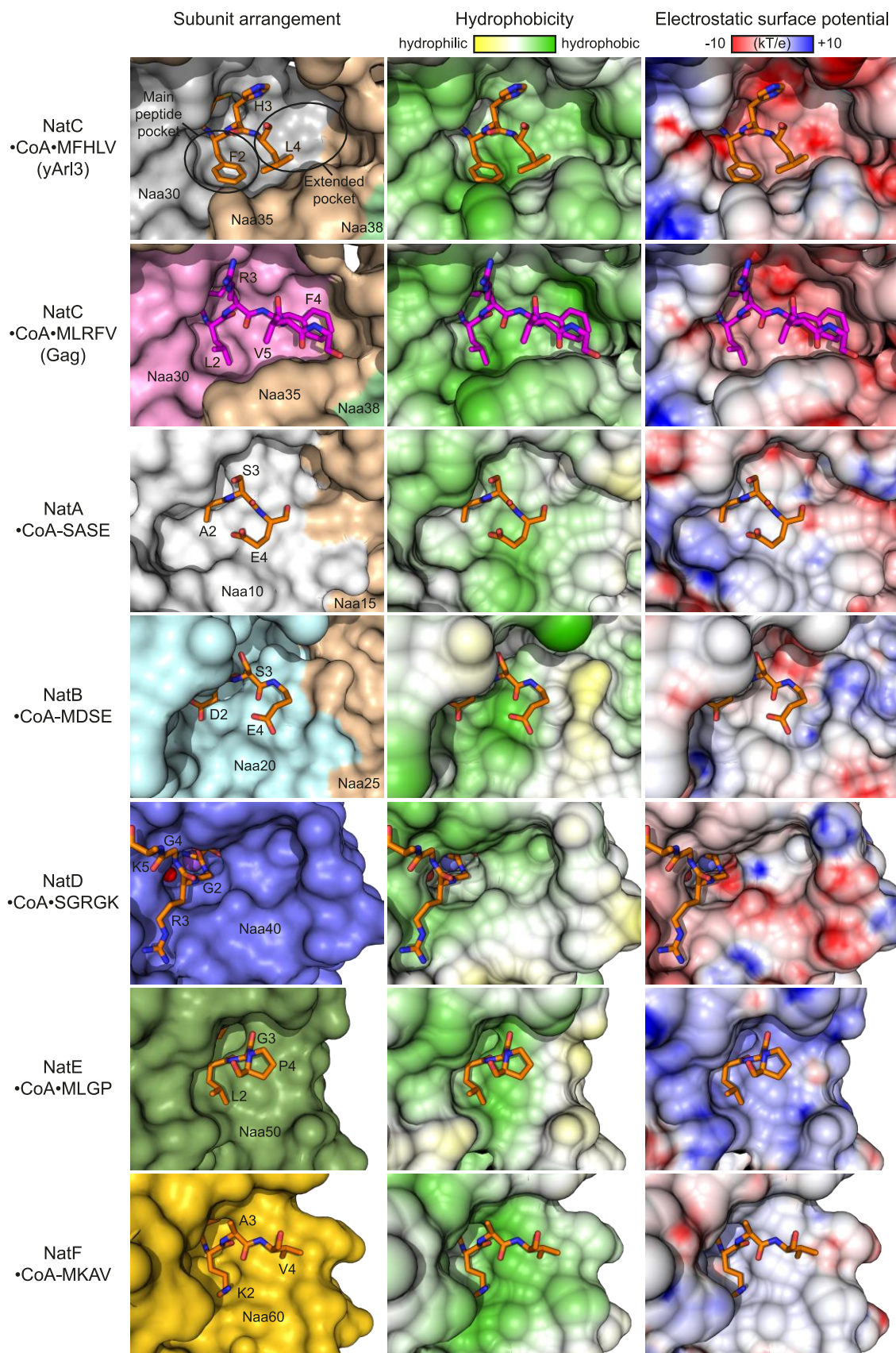
Supplementary Fig. 6 Michaelis-Menten kinetics of NatC for various substrate peptides. All reactions were performed with NatC WT (Naa30 Δ C17, full-length Naa35, Naa38 Δ C11). k_{obs} values represent initial reaction velocities (v_0) divided by the NatC concentration. Each graph (colored in black, orange, blue, purple and green) represents an independent experiment. Catalytic parameters (K_m , k_{cat}) for each replicate are shown below the plots in the corresponding color. **a** Kinetics for acetyl-CoA (variable substrate) with the yArl3 peptide (MFHLVGSRRR) at a concentration of 1340 μM ($\sim 10 \times K_m$). **b–v** Kinetics for different substrate peptides (variable substrate), as indicated below each X-axis, with acetyl-CoA at a concentration of 500 μM ($\sim 10 \times K_m$). K_m values for kinetics in subfigures **i**, **j**, **r** and **v** were below the smallest measured substrate concentrations and below the detection limit of the colorimetric assay. For these kinetics, the k_{cat} of each replicate was determined from average k_{obs} values for substrate concentrations above 100 μM . Dashed curves represent theoretical Michaelis-Menten curves for each replicate, assuming a K_m of 20 μM . In all four cases (subfigures **i**, **j**, **r** and **v**), the graphs clearly indicate that the real K_m value for these peptides is below 20 μM .



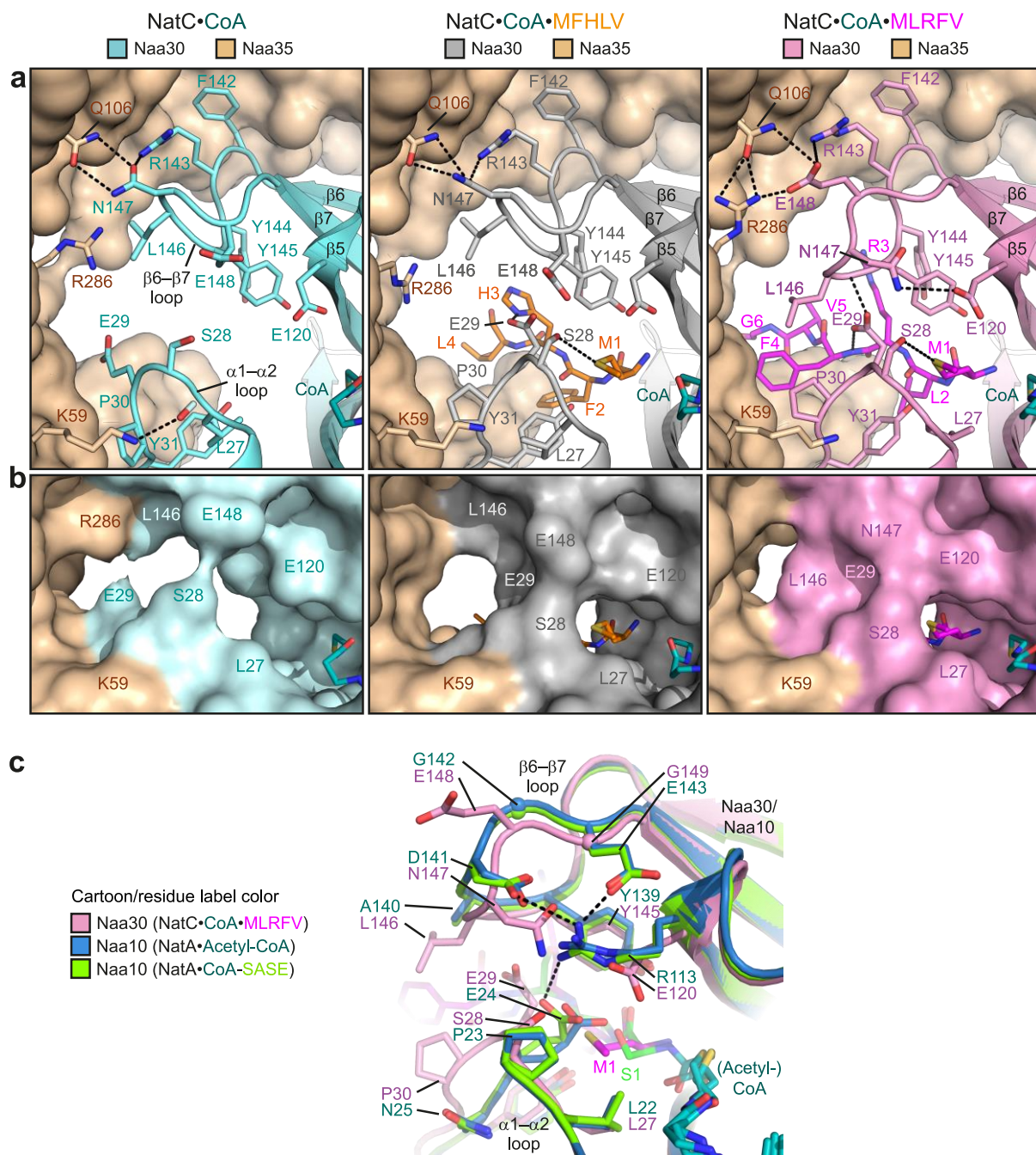




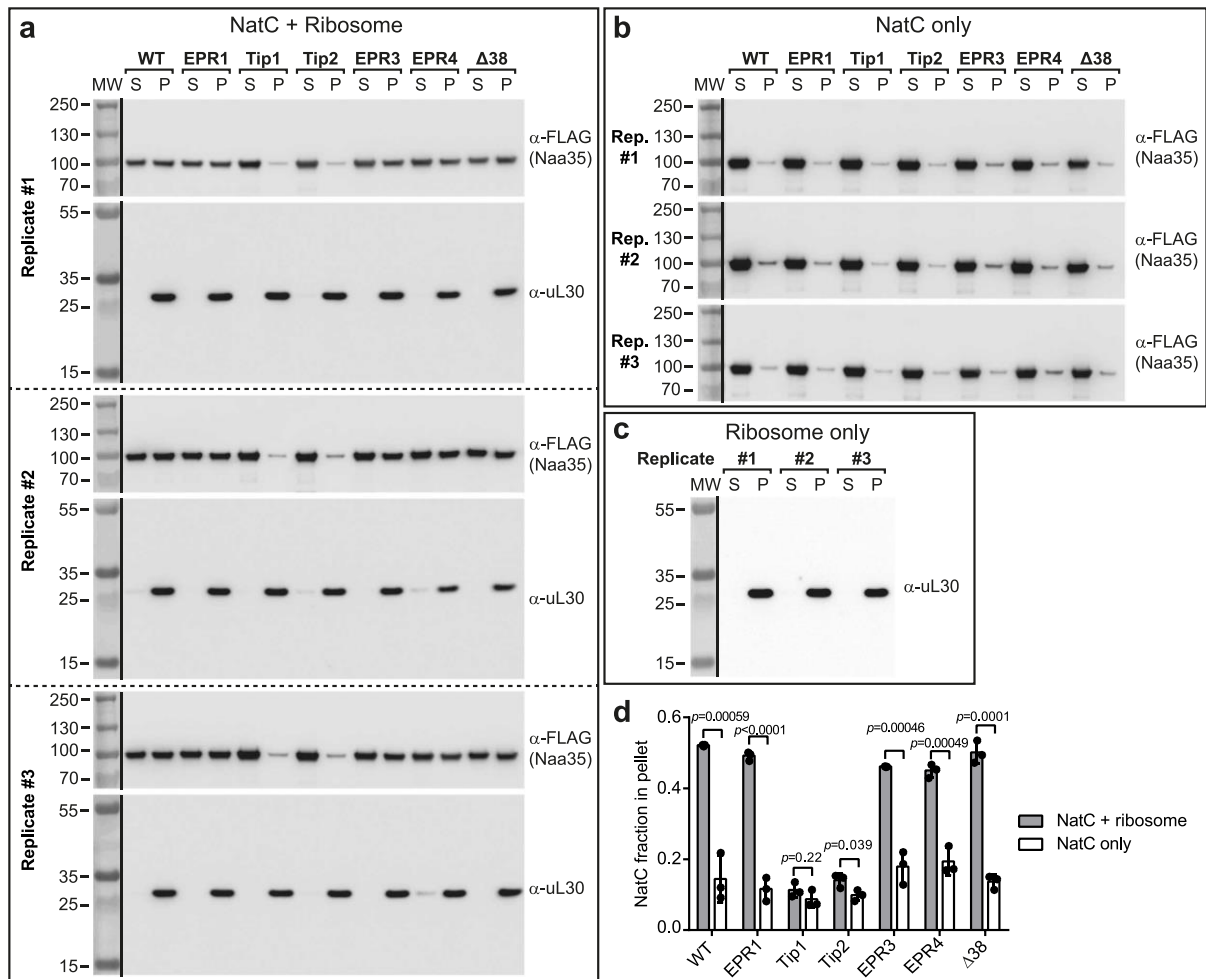
Supplementary Fig. 7 Michaelis-Menten kinetics for NatC mutants. All NatC mutants are based on the NatC WT (Naa30 Δ C17, full-length Naa35, Naa38 Δ C11) construct. Point mutations and the affected subunit are specified above each subfigure. All reactions were performed with the yArl3 peptide (MFHLVGSRRR) at a concentration of 1340 μM ($\sim 10 \times K_m$). k_{obs} values represent initial reaction velocities (v_0) divided by the NatC concentration. Each graph (colored in black, orange, blue, purple and green) represents an independent experiment. Catalytic parameters (K_m , k_{cat}) for each replicate are shown below the plots in the corresponding color. The Tip1 mutant contains Naa35 point mutations: K500A, K501A, K503A and K504A.



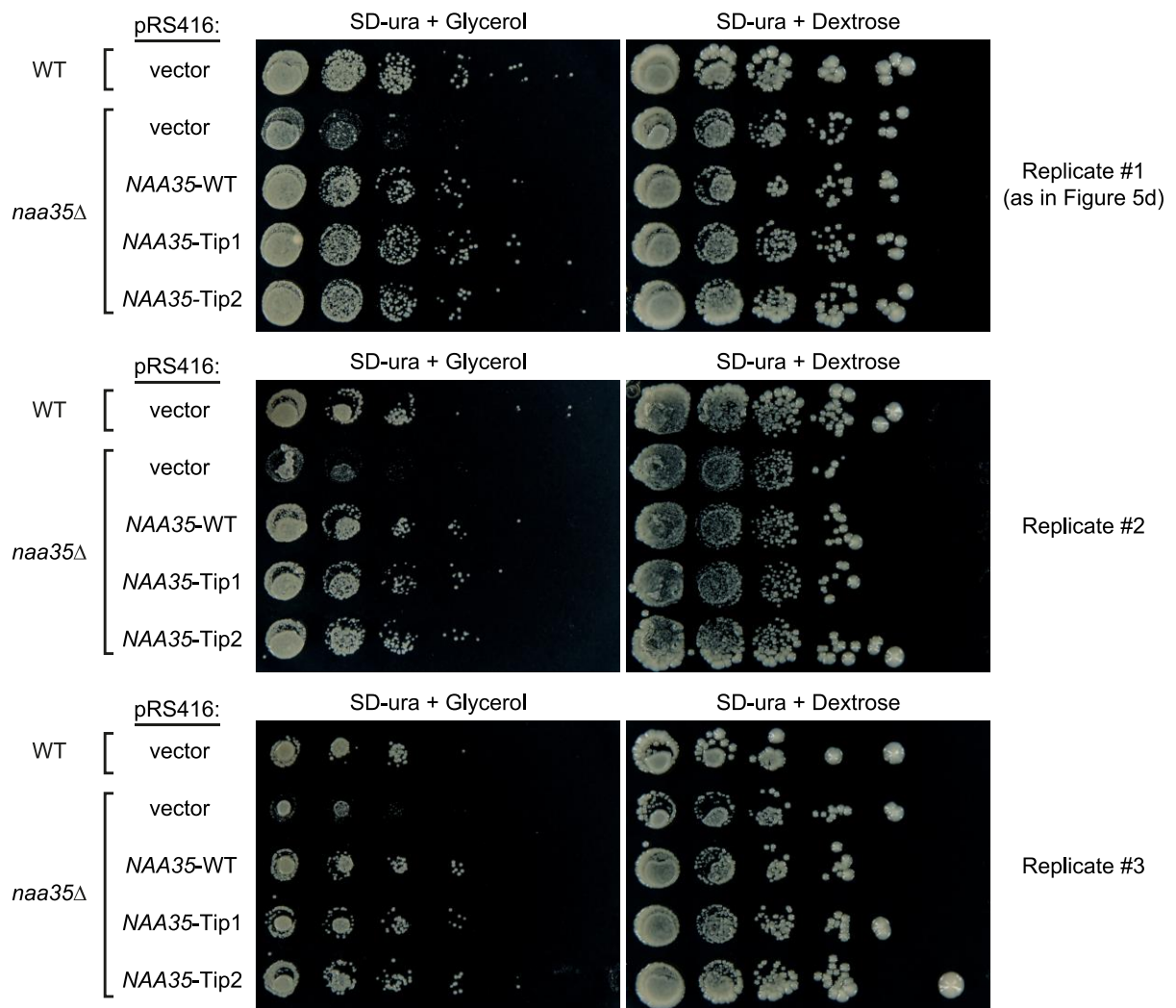
Supplementary Fig. 9 Comparisons of the peptide-binding surfaces in NAT orthologs. Substrate peptide recognition in the peptide-binding pocket of NatC•CoA•MFHLV, NatC•CoA•MLRFV, NatA•CoA-SASE (pdb 4KVM), NatB•CoA-MDSE (5K18), NatD•CoA-SGRGK (4U9W), NatE•CoA-MLGP (3TFY) and NatF•CoA-MKAV (5ICV). NAT complex subunits are colored individually to show their contribution to the peptide binding pocket (left). Additionally, NAT surfaces are colored according to hydrophobicity (middle) and electrostatic potential (right).



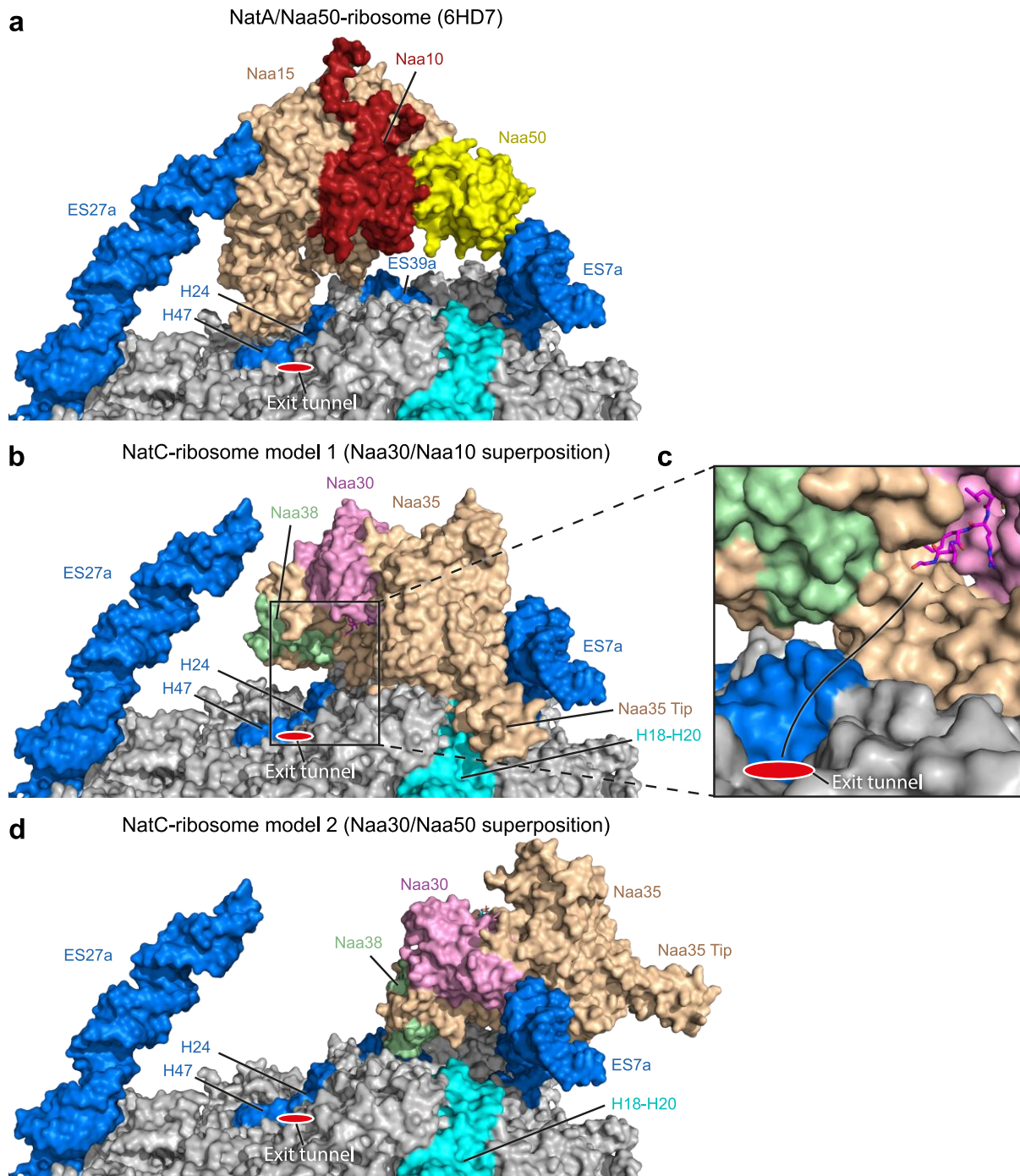
Supplementary Fig. 10 Peptide ligand-induced conformational changes. **a** Mixed surface (Naa35) and cartoon/stick (Naa30) representation, showing the structural rearrangements of the Naa30 $\alpha 1-\alpha 2$ and $\beta 6-\beta 7$ loops upon ligand binding. **b** Surface projection of NatC, showing the effect of ligand binding on the shape and diameter of the central tunnel. **c** Superposition of Naa30 (NatC•CoA•MLRFV) with Naa10 of the *Schizosaccharomyces pombe* NatA complex in the presence of Acetyl-CoA (pdb 4KVO) or a bisubstrate conjugate CoA-SASE (4KVM).



Supplementary Fig. 11 Replicates of the NatC/ribosome co-sedimentation assay. **a–c** Western blot of three individual replicates of the NatC/ribosome co-sedimentation assay (replicate #1 of Fig.11a,b is additionally shown in Fig. 5b). NatC constructs in the supernatant (S) and pellet (P) fractions were immunodetected via a FLAG tag at the Naa35 N-terminus; yeast ribosomes via the ribosomal protein uL30. EPR, electropositive region. **b** NatC sedimentation control. **c** Ribosome sedimentation control. All images in a–c show the molecular weight marker (MW) in lane 1 as an incident white light image, while all remaining lanes are chemiluminescence images from the same blot. **d** Quantification (chemoluminescence band intensities) of the NatC fraction in the pellet ($P/(S+P)$) of subfigures **a** and **b**. All data represent mean $\pm$ SD ($n=3$ independent experiments). NatC/ribosome co-sedimentation experiments were compared to the NatC sedimentation control using unpaired two-tailed t-tests with the Holm-Sidak correction. Uncropped images for panels a–c are provided in the Source Data file.



Supplementary Fig. 12 Replicates of the yeast dilution spot assay. Three individual replicates of the yeast dilution spot assay obtained for individually transformed yeast cells (replicate #1 is additionally shown in Fig. 5d). SD-ura, synthetic defined dropout medium without uracil. Serial tenfold dilutions of *S. cerevisiae* WT (BY4741) and *naa35Δ* (Y00294) strains, transformed with a pRS416 yeast centromere vector, carrying no insert (vector), *NAA35*-WT or the *Naa35*-Tip1 (K500A, K501A, K503A, K504A) or *Naa35*-Tip2 mutant (K511A, R515A, R519A). Cells were grown at 37 °C for 5 days on SD-ura agar plates, supplemented with 3% glycerol or 2% dextrose, respectively.



Supplementary Fig. 13 Models of NatC on the ribosomal surface. **a** Molecular model of the NatA/Naa50-ribosome complex (6HD7). NatA/Naa50 interacts with ribosomal RNA expansion segments (ES) and RNA helices (H), both colored in blue, near the ribosomal exit tunnel. The catalytic subunit Naa30 of the NatC•CoA•MLRFV structure was used for a superposition with Naa10 (RMSD = 1.10 Å over C α 127 atoms) or Naa50 (RMSD = 2.70 Å over C α 124 atoms) of the NatA/Naa50-ribosome complex: **b** Model 1 of the NatC-ribosome complex, generated by a Naa30/Naa10 superposition. In this model, the auxiliary subunit Naa35 is in close contact with the surface of the ribosome. Moreover, the Naa35 tip region would contact ribosomal helices H18–H20 (cyan). **c**, Magnified view of model 1, showing the close proximity between peptide-binding site and the ribosomal exit tunnel. This model would thus be consistent with a conserved arrangement of the acetylation centers in NatA and NatC relative to nascent polypeptide chain emerging from the ribosome. **d** Model 2 of the NatC-ribosome complex, generated by a superposition of Naa30 with Naa50. Hardly any interaction between NatC and the ribosome was observed and the Naa35 tip is positioned far away from the ribosome surface, arguing against model 2 representing a native conformation.

Supplementary References

1. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345 (2009).