



**HAL**  
open science

## The plastidial protein acetyltransferase GNAT1 forms a complex with GNAT2, yet their interaction is dispensable for state transitions

Annika Brünje, Magdalena Füßl, Jürgen Eirich, Jean-Baptiste Boyer, Paulina Heinkow, Ulla Neumann, Minna Konert, Aiste Ivanauskaite, Julian Seidel, Shin-Ichiro Ozawa, et al.

### ► To cite this version:

Annika Brünje, Magdalena Füßl, Jürgen Eirich, Jean-Baptiste Boyer, Paulina Heinkow, et al.. The plastidial protein acetyltransferase GNAT1 forms a complex with GNAT2, yet their interaction is dispensable for state transitions. *Molecular and Cellular Proteomics*, 2024, pp.100850. 10.1016/j.mcpro.2024.100850 . hal-04744988

**HAL Id: hal-04744988**

**<https://hal.science/hal-04744988v1>**

Submitted on 19 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## The plastidial protein acetyltransferase GNAT1 forms a complex with GNAT2, yet their interaction is dispensable for state transitions

Annika Brünje<sup>a</sup>, Magdalena Füßl<sup>a</sup>, Jürgen Eirich<sup>a</sup>, Jean-Baptiste Boyer<sup>b</sup>, Paulina Heinkow<sup>a</sup>, Ulla Neumann<sup>e</sup>, Minna Konert<sup>c</sup>, Aiste Ivanauskaite<sup>c</sup>, Julian Seidel<sup>d</sup>, Shin-Ichiro Ozawa<sup>f</sup>, Wataru Sakamoto<sup>f</sup>, Thierry Meinzel<sup>b</sup>, Dirk Schwarzer<sup>d</sup>, Paula Mulo<sup>c</sup>, Carmela Giglione<sup>b</sup>, Iris Finkemeier<sup>a</sup>

<sup>a</sup>*Plant Physiology, Institute of Plant Biology and Biotechnology (IBBP), University of Münster, D-48149 Münster, Germany*

<sup>b</sup>*Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), F-91198 Gif-sur-Yvette, France*

<sup>c</sup>*Department of Life Technologies, Molecular Plant Biology, University of Turku, FI-20520 Turku, Finland*

<sup>d</sup>*Interfaculty Institute of Biochemistry, University of Tübingen, D-72076 Tübingen, Germany*

<sup>e</sup>*Central Microscopy, Max Planck Institute for Plant Breeding Research, D-50829 Köln, Germany*

<sup>f</sup>*Institute of Plant Science and Resources (IPSR) Okayama University, Kurashiki, Okayama, 710-0046, Japan*

### **Corresponding author:**

Prof. Dr. Iris Finkemeier

Tel.: +49 (251)8323805

Email: [iris.finkemeier@uni-muenster.de](mailto:iris.finkemeier@uni-muenster.de)

**Running Head:** Plastidial GNAT1 N-acetyltransferase interacts with GNAT2

## Highlights

- GNAT1 is a plastidial acetyltransferase with dual functionality, displaying robust *in vitro* and *in planta* activities, particularly on N-terminal substrate sites.
- The *in planta* set of N-terminal substrates comprises a subset of the substrate range previously identified for GNAT2.
- Structural models of GNAT1 and GNAT2 reveal a high degree of architectural similarity beyond their significant amino acid sequence homology.
- In contrast to GNAT2, the absence of GNAT1 does not impact state transitions in *Arabidopsis*, nor does it render susceptibility to high light conditions.
- Robust interactions were observed between GNAT1 and GNAT2, as well as between GNAT2 and GNAT3, whereas no interaction was detected between GNAT1 and GNAT3.

## Abstract

Protein *N*-acetylation is one of the most abundant co- and post-translational modifications in eukaryotes, extending its occurrence to chloroplasts within vascular plants. Recently, a novel plastidial enzyme family comprising eight acetyltransferases that exhibit dual lysine and N-terminus acetylation activities was unveiled in *Arabidopsis*. Among these, GNAT1, GNAT2, and GNAT3 reveal notable phylogenetic proximity, forming a subgroup termed NAA90. Our study focused on characterizing GNAT1, closely related to the state transition acetyltransferase GNAT2. In contrast to GNAT2, GNAT1 did not prove essential for state transitions and displayed no discernible phenotypic difference compared to the wild type under high light conditions, while *gnat2* mutants were severely affected. However, *gnat1* mutants exhibited a tighter packing of the thylakoid membranes akin to *gnat2* mutants. *In vitro* studies with recombinant GNAT1 demonstrated robust N-terminus acetylation activity on synthetic substrate peptides. This activity

was confirmed *in vivo* through N-terminal acetylome profiling in two independent *gnat1* knockout lines. This attributed several acetylation sites on plastidial proteins to GNAT1, reflecting a subset of GNAT2's substrate spectrum. Moreover, co-immunoprecipitation coupled to mass spectrometry revealed a robust interaction between GNAT1 and GNAT2, as well as a significant association of GNAT2 with GNAT3 - the third acetyltransferase within the NAA90 subfamily. This study unveils the existence of at least two acetyltransferase complexes within chloroplasts, whereby complex formation might have a critical effect on the fine-tuning of the overall acetyltransferase activities. These findings introduce a novel layer of regulation in acetylation-dependent adjustments in plastidial metabolism.

### **In Brief**

Protein acetylation represents a prevalent co- and post-translational modification within chloroplasts of higher plants. Acetyltransferases facilitate the addition of acetyl groups to the amino group of the N-terminus and lysine side chains. Notably, eight enzymes have recently been identified within plant cell chloroplasts. In this context, we focused on characterizing GNAT1, closely related to the state transition acetyltransferase GNAT2, which is recognized for its regulatory function in balancing photosynthetic light harvesting. Our findings demonstrate a shared set of target sites between GNAT1 and GNAT2 *in vivo*. Both enzymes establish a stable and robust protein interaction, yet this interaction is not essential for state transitions.

### **Keywords**

acetyltransferase, GNAT, chloroplasts, *N*-acetylation, protein interaction, cross-linking, mass spectrometry, Arabidopsis

## Introduction

The acetylation of amino groups within proteins is one of the major post-translational modifications (PTMs) in all organisms. N-terminal acetyltransferases (NATs) or lysine acetyltransferases (KATs) catalyze the transfer of acetyl groups from the donor substrate acetyl coenzyme A (acetyl-CoA) to the targeted amino groups, which are either the  $\alpha$ -amino group of protein N-termini or the  $\epsilon$ -amino group of internal lysine residues [1]. In contrast to N-terminal acetylation (NTA), the acetylation of lysine residues (KA) is reversible, whereby lysine deacetylases (KDACs) are responsible for the removal of acetyl groups from proteins. Furthermore, KA occurs post-translationally, whereas NTA is described mainly as a co-translational modification with the exception of N-terminally processed proteins, such as those imported to chloroplasts and mitochondria [2–5].

With the help of state-of-the-art mass spectrometry (MS), the regulation and abundance of NTA as well as KA in proteomes of diverse organisms was investigated in detail in recent years. Since KA was discovered in the mid-1960s on histones [6], research on this modification was for several decades predominantly centered on histone acetylation and its effects on chromatin remodeling and gene regulation. Around 40 years later, significant technical progress and newly introduced immunoprecipitation-based enrichment protocols allowed an immense increase in the number of KA-sites reported to date [7–10]. Recent studies revealed, for instance, about 7000 KA-sites on more than 2500 proteins of human HeLa cells or, in the case of *Arabidopsis thaliana*, more than 5000 sites identified on about 2600 seedling or plant leaf proteins [11–14]. Acetylation of terminal amino acids in eukaryotes is known as one of the most widespread protein modifications. For example, 85 % of the human proteome and about 81% of the *Arabidopsis* proteome undergo NTA [15–17].

Since plants have a sessile lifestyle, they depend on fast and effective acclimation mechanisms to regulate their cellular metabolism. A variety of specific and tightly regulated signaling processes allow plants to react within seconds, when environmental conditions are rapidly changing, for

instance, in the case of fluctuating light intensities or qualities as well as changing temperatures and nutrient or water availabilities [18, 19]. Furthermore, as photosynthesizing organisms, plants harbor chloroplasts in their green tissues. These organelles are able to convert light energy into chemical energy to produce organic compounds, which feed into a multitude of metabolic pathways [20, 21]. Several studies demonstrated that KA as well as NTA are widespread modifications in chloroplasts, whereby the fraction of plastidial proteins undergoing post-translational NTA is rather high in comparison to its proportion in the cytosol [13, 22–25].

Eight plastid-localized acetyltransferases were recently identified exhibiting dual KA and NTA activities [23, 26, 27]. These enzymes form a novel family of chloroplast acetyltransferases, which is subordinated to the superfamily of general control non-repressible 5 (GCN5)-related *N*-acetyltransferases (GNATs). Recent studies revealed the dual specificity of these plastidial GNATs, a property that has been debated and rather rejected for the cytosolic NAT enzymes for some time [28–30]. However, the investigation of their physiological roles is still in its beginnings, and a significant role in photosynthesis was recently discovered for GNAT2 [27]. Arabidopsis mutant lines deficient in GNAT2 lost their ability to perform state transitions and did not form the state transition complex consisting of photosystem I and light harvesting complex II proteins (PSI-LHCII) [27]. Furthermore, *gnat2* lines exhibited downregulated KA of specific photosynthetic proteins including subunits of PSI, PSII, and LHCII as well as a retarded growth phenotype under fluctuating light conditions [27, 31]. On the level of structural thylakoid membrane organization, *gnat2* mutant plants resemble the phenotype of plants deficient in the kinase STN7 concerning grana packing and defective response to the shift from darkness to light [32, 27, 31]. In the state transition process, a pool of LHCII trimers (L-LHCII) is able to change its association from either PSII to PSI or vice versa, thereby serving as an additional antenna. Depending on the prevalent light qualities and quantities, the excitation energy harvested by LHCII proteins can be dynamically distributed between PSII and PSI to enable a balanced flow of electrons within the photosynthetic electron transport chain [33–35]. For more than twenty years, it is well known that changes in the

L-LHCII association are regulated by reversible phosphorylation of the LHCII subunits LHCB1 and LHCB2 catalyzed by the kinase STN7 and the phosphatase PPH1/TAP38 [36, 37]. PSII-favoring light conditions trigger the phosphorylation of LHCB1 and 2, whereas PSI-favoring illumination leads to dephosphorylation of both subunits [38, 39]. Intriguingly, *gnat2* plants, just like STN7 deficient mutants, are not able to adapt to PSII-favoring light by directing the LHCII trimers to associate with PSI, even if the phosphorylation state of LHCB1 and 2 is not impaired in *gnat2* [27, 32]. These findings clearly highlight lysine acetylation, next to phosphorylation, as an additional layer of post-translational regulation in the network of chloroplast light responses. Nonetheless, it also raises numerous questions, such as about the precise mechanism that mediates the acetylation-dependent regulation of state transitions and which other metabolic pathways within the chloroplast require KA or NTA of proteins.

In this study, we investigated the function of GNAT1, a yet uncharacterized member of the new chloroplast-localized GNAT family. The acetyltransferase domain of GNAT1 reveals an especially high degree of similarity to the corresponding domain of GNAT2 [26], leading to speculations regarding a potentially similar function *in planta*. Furthermore, this study reports on an acetyltransferase complex within the chloroplast, whereby complex formation might have a critical effect on the fine-tuning of the overall acetyltransferase activities and substrate specificities. These findings open an additional promising perspective in the study of KA as well as of NTA within chloroplasts.

## Experimental Procedures

### *Heterologous expression and purification of recombinant GNAT1 and GNAT2 proteins*

The plasmid for heterologous expression of His<sub>6</sub>-GNAT1 was cloned as described previously for GNAT2/NSI [27], using the vector system pQE-30 (Qiagen). The open reading frame of GNAT1 (AT1G26220.1) was amplified from Arabidopsis cDNA with the Phusion High-Fidelity DNA

Polymerase (Thermo), whereby the GNAT1 coding sequence was trimmed by the region coding for the predicted transit peptide (36 N-terminal amino acids). The following primers were used: GNAT1\_pQE-30\_for, 5'-CTAGGATCCGCTGCAATGCAAC and GNAT1\_pQE-30\_rev, 5'-CAACCCGGGTTATTTCTTGTCTCTGT. Recombinant His<sub>6</sub>-GNAT2 protein was overexpressed and purified as reported by Koskela and coworkers, while the procedure was slightly modified in the case of GNAT1 [27]. In brief, recombinant His<sub>6</sub>-GNAT1 was expressed in the *E. coli* strain SoluBL21 (Genlantis) induced with 0.5 mM IPTG (Roth) for 15 h at 21°C. Cells were harvested, resuspended in cold buffer (100 mM K-phosphate, pH 8.0, 300 mM NaCl, 10% [v/v] glycerol, protease inhibitor cocktail [Sigma-Aldrich]), and disrupted with a French Press. Afterwards, 5 mM DTT and 10 units of lysozyme were added and His<sub>6</sub>-GNAT1 was purified from the soluble phase by Ni-NTA affinity chromatography (HisTrap FF, Cytiva) using the FPLC-system ÄKTA go (Cytiva).

#### *HPLC-based N<sub>ε</sub>-Lys- and N<sub>α</sub>-N-terminus acetylation assay*

His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2 were analyzed concerning their N<sub>ε</sub>-Lys- and N<sub>α</sub>-N-terminus acetylating activities by using artificial peptides, which function as KAT or NAT substrates, respectively [40, 26]. The sequences of the substrate probes generally comprise a modified lysine residue at position nine that carries a dinitroaniline group to allow the specific detection of the peptides by recording their absorbance at 360 nm. A lysine side chain at position five is either free to allow KA by a KAT or is blocked by an acetyl group (NTA peptide substrate). In the case of the NTA peptide probes, substrates with varying N-terminal amino acids were analyzed (Ala, Gly, Ser, Thr, Val, Met, or Leu), whereas the KA-specific probe generally provides an N-terminal Ala residue with a blocked N<sub>α</sub>-amino group. In analogy to the substrates reported by Seidel and coworkers, the KA peptide substrate contains an additional modified lysine residue at position three that is linked to an amino aminobenzoic acid moiety *via* its N<sub>ε</sub>-amino group [40]. In the NTA substrate probe, position three corresponds to an unmodified Gln residue. In summary, the KA substrate



peptide sequence is, here in one-letter code: ac-A-A-K(oAba)-G-A-K-A-A-K(Dnp)-Ahx-r-r-r and the sequence of the NTA substrate X-A-Q-G-A-K(ac)-A-A-K(Dnp)-Ahx-r-r-r, whereby oAba refers to ortho-aminobenzoic acid, Dnp to dinitroaniline, Ahx to 6-aminohexanoic acid, and *r* to D-arginine. For measuring catalytic activities of the recombinant His<sub>6</sub>-GNAT proteins, His<sub>6</sub>-GNAT1, His<sub>6</sub>-GNAT2, or both together, were incubated with the substrate peptide (100 μM) in reaction buffer (150 mM K-phosphate, pH 8.0, 50 mM NaCl, protease inhibitor cocktail [Sigma-Aldrich]) and the reaction was started by addition of acetyl-CoA (200 μM). From time points between 0 and 4 h, 11 μl-aliquots were collected and the reaction was stopped with 99 μl trifluoroacetic acid (TFA; final concentration 2 % [v/v]). Reaction products were analyzed by using a reversed-phase HPLC chromatograph (Shimadzu), as previously described [27]. Reaction rates were calculated from the peak areas of the chromatogram corresponding to the respective peptide probe.

#### *Computational analysis of AlphaFold 2-predicted GNAT structure models*

The GNAT structure models presented here were obtained from the AlphaFill database (www.alphafill.eu; [41]), which integrates missing ligands, cofactors, and ions into the protein structures provided by the AlphaFold 2 protein structure database [42]. AlphaFold 2 Multimer was used to model structures of protein complexes (version 2.3.1; [43]). Further processing steps, such as computational removal of transit peptide sequences and superimpositions of GNAT structures, were performed by using the software PyMOL (The PyMOL Molecular Graphics System, Version 4.5 Schrödinger, LLC). The protein sequence alignment of GNAT1 and GNAT2 (without transit peptide) was derived by using the PRALINE Multiple Sequence Alignment tool [44, 45].

#### *Plant Material*

*Arabidopsis thaliana* plant lines were cultivated in an 8 h light/ 16 h darkness regime at a PPFD of 120 μmol · m<sup>-2</sup> · s<sup>-1</sup>. All lines presented in this study are based on the ecotype Col-0. The mutant lines overexpressing GNAT1-, GNAT2- or GNAT3-GFP, respectively, were described in [26], while

the control line overexpressing chloroplast-targeted GFP was provided by Markus Schwarzländer, University of Münster, Germany [46]. Seeds of the T-DNA insertion mutants SALK\_062388 (*gnat1-1*) and SALK\_150736 (*gnat1-2*) were ordered from NASC and the T-DNA line SALK\_033944 (*gnat2-1, nsi-1*) was characterized previously [27].

High light and darkness treatment were applied on seedlings, which were grown on sterile agar plates containing ½ MS medium and cultivated for eight days under long-day conditions (16 h light/ 8 h darkness, PPFD of 120  $\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ). Photosynthetic performance was determined by detecting PSII-specific chlorophyll fluorescence signals via the IMAGING-PAM fluorometer before the treatment and after the indicated number of days under darkness or highlight (PPFD = 800  $\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) regime. In total, three plates were analyzed per condition, each with 15 individual seedlings per plant line that were selected for fluorescence detection.

#### *Genotyping of T-DNA lines*

The two Arabidopsis lines *gnat1-1* and *gnat1-2* were screened by PCR to amplify a DNA fragment of the intact gene locus as well as of a region covering the locus of the T-DNA insertion. In the first case, the primer pairs *gnat1-1\_LP*, 5'-GCAAGAAAGAATGCAGCAAAC and *gnat1-1\_RP*, 5'-AACCATTCTTTGATCCCATC, as well as *gnat1-2\_LP*, 5'-CATTCTGAATTGCAGGAGAG and *gnat1-2\_RP*, 5'-TGTAATCCTCAATCAACCGC were used. To confirm the insertion of T-DNA, each RP-primer was combined with the T-DNA specific oligonucleotide SALK\_LB, 5'-GATTTAGTGCTTTACGGCACCTC to examine the T-DNA orientation suggested by the Salk Institute Genomic Analysis Laboratory. To reveal “inverse” T-DNA insertion, the primers *gnat1-1\_LP* and *gnat1-2\_LP* were used in combination with SALK\_LB, respectively. The absence of mRNA coding for GNAT1 was verified by endpoint RT-PCRs proving both, the absence/presence of the full-length transcript and of an mRNA fragment, which covers the region downstream of the T-DNA insertion site (*gnat1\_P1\_for*, 5'-CATGTTTCTCGGAGGCACAATCTC; *gnat1\_P2\_for*, 5'-GATCTCGAATCCAGAGGCTTTC; *gnat1\_P3\_rev*, 5'-

CAAAC TTTT TATTTCTTGT TTTCTCTG TTTTG). As a reference, additional RT-PCRs were performed allowing the amplification of a fragment of the *ubc21*-transcript (Ubiquitin-conjugating enzyme 21; *ubc21\_for*, 5'-CTTTGCAACCTCCTCAAGTTC; *ubc21\_rev*, 5'-CATCATCATCCTTTCTTAGGCATAG). For RT-PCR, RNA was isolated from leaf material by phenol-chloroform-isoamylalcohol (25:24:1) extraction, and subsequent precipitation of RNA was achieved by using 50 % isopropanol. Afterwards, the RNA preparation was further purified by precipitation with 2 M LiCl and washing with 70% ethanol. For cDNA synthesis, genomic DNA was digested with the RNase-free DNase RQ1 (Promega), and cDNA was synthesized with the help of the FastGene Scriptase II cDNA kit (NIPPON Genetics).

#### *Cloning of GNAT1, 2, and 3 for Bimolecular Fluorescence Complementation (BiFC) assays*

The open reading frames without STOP codon of GNAT1 (AT1G26220.1), GNAT2 (AT1G32070.1) and GNAT3 (AT4G19985.1) were recombined from entry clones into the Gateway vectors pBatTL-B-sYFPc and pBatTL-B-sYFPn [47, 48] by performing LR reactions. These vector systems allow C-terminal fusions of either the C-terminal or the N-terminal domain of a split-YFP to the protein of interest. In the case of GNAT1 and 2, coding DNA sequences amplified from *Arabidopsis* (Col-0) cDNA were first cloned into the pGWR8 vector [49] and then transferred to the pENTR/D-TOPO (Thermo Fisher) gateway entry vector system by type II endonuclease restriction and subsequent DNA ligation. The entry vector containing GNAT3 was generated by using the pENTR/D-TOPO kit directly in combination with the PCR product amplified from *Arabidopsis* (Col-0) cDNA. The following primers were used to obtain GNAT1, 2, and 3 coding sequences, respectively: *gnat1\_pENTR\_for*, 5'-TATACCCGGGATGTTTCTCGGAGG; *gnat1\_pENTR\_rev*, 5'-TATAGGATCCTTTCTTGT TTTCTCTG TTTG C; *gnat2\_pENTR\_for*, 5'-TATACCCGGGATGCTACTAATCCCA; *gnat2\_pENTR\_rev*, 5'-TATAGGATCCCTTTGGGTACCAAACATG; *gnat3\_pENTR\_for*, 5'-CACCTTAAGGAAATGGGTTTG; *gnat3\_pENTR\_rev*, 5'-TGCCTCCAAGCTCTTTGTGA. Vector

constructs were confirmed by sequencing and used for the transient transformation of *Arabidopsis* (Col-0) protoplasts.

*Protoplast isolation and transformation followed by Bimolecular Fluorescence Complementation (BiFC) assays via confocal laser scanning microscopy (CLSM)*

Protoplasts of wild type *Arabidopsis* Col-0 plants grown for six weeks in 8 h light/ 16 h darkness conditions were isolated with the tape-sandwich method [50]. Suspensions of extracted protoplasts were then processed according to the polyethylene glycol method [51, 52]. Transfected protoplasts were resuspended in buffer W1 (4 mM MES-KOH pH 5.7, 0.5 M mannitol, 20 mM KCl) and incubated for 8-24 h under constant agitation and application of low light intensity (PPFD of  $25 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ). The YFP signal was recorded by using a Leica SP5 imaging system (Leica Microsystems) in combination with the water immersion objective lens HCX PL APO lambda blue 63.0 x 1.20 WATER UV. YFP emission was measured at 520-550 nm by applying an excitation wavelength of 514 nm. As a reference, chlorophyll auto-fluorescence was recorded above 700 nm.

*Extraction of thylakoid proteins*

Fresh *Arabidopsis* leaves were harvested and ground in cold buffer (300 mM sucrose, 50 mM HEPES-KOH, pH 7.6, 5 mM  $\text{MgCl}_2$ , 1 mM Na-EDTA, 1.25% [w/v] BSA, 22 mM ascorbate, 10 mM NaF, plant-specific protease inhibitor P9599 [Sigma]) by using a ball mill (Retsch) for 20 s at the highest frequency. The homogenate was filtered through one layer of miracloth (Millipore) and the collected filtrate was centrifuged for 4 min, 4000g, 4°C to sediment thylakoid membranes and chloroplasts. Chloroplasts were disrupted by resuspending the pellet in hypotonic buffer (5 mM sucrose, 10 mM HEPES-KOH, pH 7.6, 5 mM  $\text{MgCl}_2$ , 10 mM NaF, plant-specific protease inhibitor P9599). The remaining thylakoid membranes were pelleted by centrifugation for 5 min at 18000g,

4°C, resuspended in storage buffer (100 mM sucrose, 10 mM HEPES-KOH, pH 7.6, 10 mM MgCl<sub>2</sub>, 10 mM NaF) and flash-frozen in liquid N<sub>2</sub>.

#### *Solubilization of thylakoid membrane proteins and native gel electrophoresis*

Membrane complexes of the extracted thylakoids were solubilized by using either 1 % (w/v) digitonin or 1 % (w/v) dodecyl maltoside [53]. Further steps of sample preparation as well as casting of large-pore blue native gels followed by gel electrophoresis under non-denaturing conditions were performed as previously described [53].

#### *77K chlorophyll emission spectra measurements*

Freshly harvested Arabidopsis leaves were ground in cold buffer (300 mM sorbitol, 20 mM HEPES-KOH, pH 7.9, 2.5 mM Na-EDTA, 10 mM ammonium hydrogencarbonate, 2 mM ascorbate, 0.1% [w/v] BSA, 10 mM NaF, plant-specific protease inhibitor P9599 [Sigma]) with a ball mill (Retsch) for 20 s at highest frequency and the homogenate was then filtered through one layer of miracloth (Millipore). The chlorophyll concentration of the extract was determined according to the protocol of Porra and coworkers [54], and adjusted to 10 µg/ml chlorophyll before it was transferred to glass tubes and flash-frozen in liquid N<sub>2</sub>. Spectra were recorded with a Jasco FP-6500 spectrometer by selecting an excitation wavelength of 435 nm and detecting the chlorophyll fluorescence in an emission wavelength range from 600 to 800 nm. The data pitch was set to 1 nm and the scanning speed to 200 nm/min. Five emission spectra per biological replicate were measured and averaged.

#### *Chlorophyll fluorescence measurements*

PSII-specific chlorophyll a fluorescence was analyzed by using the pulse-amplitude-modulation (PAM) fluorometry with either Dual-PAM-100 or IMAGING-PAM (Walz) instruments. For the characterization of Arabidopsis seedlings grown on agar plates, photosynthetic performance was

determined with the IMAGING-PAM system, which allows the simultaneous measurement of several selected spots per run. To record PSI-specific absorption parameters in addition to PSII-specific fluorescence signals, the Dual-PAM-100 equipped with DUAL-E and DUAL-DR measuring heads was used. Light response curves were measured to characterize the photosynthetic performances of the different Arabidopsis plant lines during the application of increasing light intensities, whereby one intact leaf per plant was used as one biological replicate. Prior to each measurement, plants were dark incubated for 20 min followed by the fixation of a selected leaf between the measuring heads and determination of the initial basic fluorescence  $F_0$  and the maximum fluorescence  $F_m$ . The maximum P700 reduction was measured after a 10 s-phase of FR illumination *via* the application of a saturating light flash [55]. Afterwards, the leaf was treated with actinic light, which was increased in intensity in 2 min-intervals. Each illumination phase included the application of a saturating light flash at the end of the interval allowing the determination of the PSII- and PSI-specific, light-dependent parameters, such as the effective quantum yield of PSII,  $Y(II)$  [56], the nonphotochemical quenching NPQ [57], the photochemical quenching parameters  $qP$  and  $qL$  [58, 59], the quantum yields of regulated and nonregulated nonphotochemical quenching,  $Y(NPQ)$  and  $Y(NO)$  [59], as well as the quantum yield of PSI,  $Y(I)$ , and the donor and acceptor side limitation of PSI,  $Y(ND)$  and  $Y(NA)$  [55].

For inducing and analyzing state transitions in Arabidopsis, the same Dual-PAM-100 setup was used. After an initial dark incubation of 20 min, the  $F_0$ - and  $F_m$ -parameters of a single leaf were measured by applying a saturating light flash. Photosynthesis was then activated by illumination with far-red and weak red light ( $23 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) for a period of 5 min, followed by a phase of 20 min with only red-light incubation to induce state 2. Induction of state 1 was afterwards achieved by switching the far-red light source on again for an additional phase of 20 min. At the end of each 20 min-interval, a saturating light flash was given to determine the maximum fluorescence level in state 1 ( $F_{m1}$ ) and state 2 ( $F_{m2}$ ), respectively. The state transition parameters  $qT$ ,  $qS$ , and  $IB$  were calculated according to the protocols of Yeates and coworkers as well as of Ruban and Johnson,

whereby the measure of energy imbalance, IB, was used as an internal control to optimize the light settings of the measuring device [60, 61].

#### *Determination of CO<sub>2</sub> assimilation rates*

Six weeks old plants, which were adapted to GL (PPFD = 120  $\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ), were placed into the measuring chamber of a LI-6800 (LI-COR) gas exchange system and incubated in darkness with a constant CO<sub>2</sub> concentration of 400 ppm. As soon as a stable CO<sub>2</sub> level was reached, the actual assimilation rate was recorded and averaged over a period of 20-30 sec. In a stepwise approach, the light intensity was increased and the CO<sub>2</sub> assimilation rate was determined. The leaf area of each plant was calculated by dissecting the corresponding rosette and calculating its overall surface with the help of the image analysis software ImageJ [62]. For correction of CO<sub>2</sub> assimilation rates, each plant pot containing only the leftover soil was measured in the same experimental setup [63]. Six plants per genotype were analyzed and differences between the plant lines tested for significance by a One-Way ANOVA approach and a Fisher LSD test with a significance level of  $p \leq 0.05$ .

#### *Sample preparation for Transmission Electron Microscopy*

Wild type and mutant plants were grown for six weeks under short-day conditions (8 h light/ 16 h darkness, PPFD of 120  $\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) prior to the preparation of leaf discs. The subsequent steps of processing were then performed as described previously [27]. In total, 160 grana stacks of 16 TEM images, which were derived from at least five biological replicates, were analyzed per genotype.

#### *Lysine Acetylome profiling*

Arabidopsis leaves were harvested in the middle of the light phase from plants, which have been grown for 6 weeks in an 8 h light/ 16 h darkness regime. The leaves were flash-frozen in liquid N<sub>2</sub>

and ground to a fine powder before proteins were extracted using a filter-assisted sample preparation protocol [64, 65]. The proteins were trypsin-digested and afterwards dimethyl-labeled on C<sub>18</sub> Sep-Paks plus short columns (Waters) as previously reported [66]. A label swap was introduced. Equal amounts of light- and medium-labeled peptides were combined for each replicate and dried in a vacuum centrifuge. Before the enrichment of lysine-acetylated peptides, samples were resuspended in TBS buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl). A 10 µg-peptide aliquot per sample was stored for whole proteome analysis, while the remaining peptide was treated with an anti-acetyllysine antibody coupled to agarose beads (ICP0388, Immunechem, www.immunechem.com) as described previously [64]. Peptides were desalted, fractionated in three steps using SDB stage tips, and dried by vacuum centrifugation [67].

*Non-denaturing extraction of the leaf proteome and enrichment of GNAT-GFP interacting proteins*

Arabidopsis leaf proteins were extracted in a native buffer and enriched as described in detail by Née and coworkers [68]. In brief, frozen leaves were ground in extraction buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 10% [v/v] glycerol, 2 mM EDTA, 5 mM DTT, 0.5% [v/v] Triton X-100, plant-specific protease inhibitor P9599 [Sigma]) and the lysate was incubated for 1 h at 4°C. After centrifugation, an aliquot of 20 µg protein per replicate of the soluble phase was stored for whole proteome analysis. For the pulldown analysis, 2 mg protein were mixed with 20 µl GFP-Trap A bead slurry (ChromoTek) and incubated for 2 h at 4°C under constant agitation. The beads were washed with cold washing buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 10% [v/v] glycerol, 2 mM EDTA, 5 mM DTT) to remove non-binding proteins. At this point, the chemical cross-linker disuccinimidyl sulfoxide (DSSO) was introduced where indicated to generate covalent intra- and inter-protein bonds between neighboring lysine residues that can be cleaved later during the MS run. DSSO was added in a concentration of 12 mM and incubated for 1 h at 8 °C. Excess of DSSO was then quenched by the addition of Tris buffer in a final concentration of 20 mM. Afterwards,



proteins were denatured, reduced, and alkylated on-bead as previously described [13]. Proteins were digested with LysC overnight at room temperature, followed by a tryptic digest for additional 4 h at 37°C. Peptides were then desalted using C<sub>18</sub> stage tips and dried by vacuum centrifugation. Proteins of the 20 µg-total-proteome aliquots were purified from the extraction buffer and digested into peptides *via* the sample preparation protocol described by Hughes and coworkers using paramagnetic beads (Cytiva) [69].

#### *LC-MS/MS data acquisition*

The dried peptide pellets were resuspended in 2% [v/v] ACN, and 0.1% [v/v] TFA for analysis. Samples were measured using an EASY-nLC 1200 coupled to a Q Exactive HF or an Orbitrap Exploris 480 mass spectrometer. In the case of cross-linking experiments samples were measured in a Vanquish Neo coupled to an Orbitrap Eclipse equipped with a FAIMS pro duo (all instruments Thermo Fisher Scientific, see also Suppl. Data 7). During liquid chromatography, peptides were separated on 20-cm-frit-less pulled emitters (CoAnn Technologies Inc., 0.75-µm inner diameter), packed in-house with reversed-phase ReproSil-Pur C<sub>18</sub> AQ 1.9 µm resin (Dr. Maisch), whereby the column was kept at 50°C in a column oven throughout the run. The following parameters were applied for total-proteome analysis, while values stated in brackets indicate parameters used for lysine acetylome and co-immunoprecipitation analysis; if not stated separately, the parameters were identical. Peptides were eluted for 115 (68) minutes applying a segmented linear gradient of 0 to 98% solvent B (solvent A, 0% [v/v] ACN, 0.1% [v/v] FA; solvent B, 80% [v/v] ACN, 0.1% [v/v] FA) with a flow rate of 300 (250) nL/min. Mass spectra were acquired in a data-dependent acquisition mode with a Top15 method. The mass range was set to 300 to 1759 *m/z* at a resolution of 60000 (120000) FWHM, maximum IT of 55 ms, and a target value of 3 x 10<sup>6</sup> ions. Precursors were selected within an isolation window of 1.3 (1.2) *m/z* and HCD fragmentation was performed with a normalized collision energy of 25. MS/MS spectra were acquired with a target value of 10<sup>5</sup> (5 x 10<sup>4</sup>) ions at a resolution of 15000 FWHM, maximum IT of 55 (150) ms, and a fixed first mass

of  $m/z$  100. Peptides revealing a charge of +1, >6, or with an unassigned charge state were excluded from fragmentation for MS<sup>2</sup>, and dynamic exclusion for 30 s prevented repeated selection of precursors. Cross-linked peptides were measured twice using both the “XLMS Cleavable MS2 MS3” and “XLMS Cleavable MS2 MS2 MS3” method templates with 2 CVs (-50 V and -75 V) per run enabling the triggering of cross-link specific fragmentation spectra based on the specific mass shift of DSSO (31.9721 Da). Raw data (other than cross-link) were processed using MaxQuant software version 2.0.3.0 [70, 71], searching against the latest version of the Araport reference proteome for *Arabidopsis thaliana* (Araport11 containing 48,266 entries; [72]). A list of common contaminants and a reverse decoy database were enabled for the search. Carbamidomethylation was used as a fixed modification and oxidation of methionine and N-terminal acetylation were set as variable modifications. In addition, lysine acetylation was enabled as variable modification in the case of lysine acetylome profiling for enriched samples. In the case of dimethyl-labeled peptides, the multiplicity of MS<sup>1</sup>-based quantification was set to two selecting the light and intermediate label. For processing of data derived from label-free GNAT-GFP co-enrichment approaches, the options of Label-Free Quantification (LFQ) and intensity-Based Absolute Quantification (iBAQ) were enabled. Trypsin and, if required, LysC, were selected as proteases, and maximal two missed cleavages (total-proteome, co-immunoprecipitation assay) or four missed cleavages (acetylome) were accepted. The minimum length of valid peptides was set to seven amino acids. Re-quantify and “match-between runs” was enabled. Other settings were kept as default: PSM and protein FDR was 0.01. A tolerance of peptide of 20 / 4.5 ppm was allowed for first / main search, while the MS/MS peak tolerance was 20 ppm. Data for cross-link detection was processed using ProteomeDiscoverer 3.0.1.27 and XlinkX [73]. The Araport reference proteome for *Arabidopsis thaliana* was filtered for proteins localizing to the plastid according to SUBA4 reducing the search space to 3,318 entries [74]. Trypsin was set as a peptidase allowing up to 2 missed cleavages. Carbamidomethylation was used as a fixed modification and oxidation of methionine and N-terminal acetylation were set as variable

modifications, DSSO was used as a cross-linker. PSM, CSM, and protein FDR were 0.01. A tolerance of precursor ions of 10 ppm, FTMS Fragment Mass Tolerance of 20 ppm, and ITMS Fragment Tolerance of 0.5 Da were set. MS raw data were uploaded to the JPOST repository (identifier: JPST002262). In addition, spectra were also available *via* MS-viewer under sc2f0nxgso (for the co-immunoprecipitation profiles) and sjczk1ixdg (for the lysine acetylation profiles).

### *Data analysis*

Downstream data analyses were performed using Perseus version 1.6.15.0 [75] and R version 4.1.1 in combination with the additional LIMMA statistics package [76, 77]. Prior to differential expression or enrichment analyses, reverse hits and potential contaminants were removed. For the co-immunoprecipitation experiments, valid quantifications in at least three of four biological replicates per experiment (genotype) were required and in addition, significantly enriched protein groups had to possess a  $\log_2$  fold change of  $\geq 1$  and a false discovery rate-corrected p-value  $\leq 0.05$  (LIMMA). Figures were prepared with the help of the software tool OriginPro version 2022b (OriginLab Corporation, Northampton, USA).

### *MS-based characterization of protein N-termini and NTA profiling*

N-terminome analysis was performed on Arabidopsis leaves from the same plants that were harvested for the lysine acetylome profiling. These leaves were frozen in liquid nitrogen and ground before being processed following the SILProNAQ protocol [78]. The powder was solubilized in a specific buffer for protein extraction (50 mM HEPES-NaOH pH 7.2, 1.5 mM  $\text{MgCl}_2$ , 1 mM EGTA, 10% [v/v] glycerol, 1% [v/v] Triton X-100, 2 mM PMSF, 150 mM NaCl, protease inhibitor cocktail tablet). Protein amounts were determined by Bradford assays and 1 mg of protein was used for each sample. After denaturation of the proteins by reduction and subsequent alkylation of cysteine residues, proteins were N-terminally labeled with d3-N-Acetoxy succinimide, before being digested with trypsin. Peptides were desalted using  $\text{C}_{18}$  SepPak cartridges and

vacuum dried. Dried peptides were resuspended in 5 mM  $\text{KH}_2\text{PO}_4$ , 30% [v/v] ACN, 0.05% [v/v] FA (SCX buffer A) and loaded onto a Polysulfoethyl A column (polyLC, 2.1 x 200 mm, 5  $\mu\text{m}$ , 200  $\text{\AA}$ ). The elution was performed using a multi-step gradient of 5 mM  $\text{KH}_2\text{PO}_4$ , 350 mM KCl, 30% [v/v] ACN, 0.05% [v/v] FA (SCX buffer B) as follows: 0% B for 5 min; 0-26% B in 10 min; 26-35% B in 25 min; 35-60% B in 5 min; 100% B for 10 min; 0% B for 20 min. Fractions were collected every 2 min, for 30 min, at a constant flow rate of 0.2 ml/min. The fractions 2-11 were kept for analysis and vacuum-dried. The dried SCX fractions were resuspended in 30  $\mu\text{l}$  of 2% [v/v] ACN, 0.1% [v/v] TFA each, before being loaded individually on an EASY-nLC 1000 coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific). The peptide separation was done using a 15 cm  $\text{C}_{18}$  capillary column (NTCC-360/100-5-153, Nikkyo Technos Co., Ltd.) and a 55 min method that allows for fragmentation of singly charged ions. The MS full scans were set for an m/z range of 400-2000 Th, with a 60,000 FWHM resolving power. The 20 most intense ions were selected each time for CID fragmentation. With the chromatography buffers being 2% [v/v] ACN, 0.1% [v/v] FA (buffer A) and 100% [v/v] ACN, 0.1% [v/v] FA (buffer B), the separation was done as follows, with a constant flow rate of 300 nl/min: 5-35% B in 42 min; 35-80% B in 1 min; 80% B for 12 min. Raw data files from a sample's fractions were processed together using Mascot Distiller version 2.6.2 (Matrix Science) to search against the Araport11 database (48,266 entries [72] with the mass tolerances respectively set to 10 ppm and 0.5 Da for parent and fragment ions, and cysteine carbamidomethylation and lysine d3-acetylation as fixed modifications, methionine oxidation as a variable modification. The enzyme was defined as Semi-Trypsin/P and up to 5 miss cleavages were allowed to account for incomplete trypsin digestion that can occur with acetylated lysine residues. The NTA was quantified by selecting the "Acetylation" quantitation method. Identification and quantitation results were exported from Mascot and parsed through the eNcounter script [78, 26]. Thereby, the peptide ion score was set to 25 and the significance threshold to 0.2. These values were adjusted to a score of 30 and an e-value of 0.05 by the

eNcounter tool [79]. Further data processing was done in Microsoft Excel using the built-in statistics tools.

Journal Pre-proof

### *Experimental design and Statistical Rationale*

We sampled material from four individual plants for MS analysis and processed them with no additional technical replicates. Detailed statistical and bioinformatic analyses of proteomics data are described in the respective sections above. The respective tests as well as p-value- and FDR-cutoffs are indicated within the respective figure legends. Supplemental Data 7 outlines the respective instruments used to generate the data.

### **Results**

#### *Recombinant His<sub>6</sub>-GNAT1 shows a relaxed substrate specificity in vitro.*

Based on their phylogenetic relationship, GNAT1, 2, and 3 were previously assigned to a distinct subgroup of plastidial NATs called NAA90 [25, 26]. Since GNAT1 and GNAT2 showed the highest degree of proximity within the phylogenetic tree, we first compared the *in vitro* substrate preferences of the recombinant GNAT proteins in acetyltransferase enzyme assays. For these *in vitro* assays, synthetic peptides were used, which are specifically designed to measure dual acetyltransferase activities in an established HPLC-based assay (Fig. 1 A; [40, 26, 27]). The peptides share a nearly identical amino acid sequence and contain either a free N-terminus with varying amino acid residues (Ala, Gly, Ser, Thr, Val, Met, or Leu) or a free lysine side chain as a substrate for acetylation. In addition, the peptides carry a dinitroaniline side chain close to the C-terminus to specifically record their absorbance at 360 nm during HPLC analysis. The recombinant N-terminally His<sub>6</sub>-tagged GNAT1 and GNAT2 proteins lacking the predicted plastid transit peptides were expressed and purified from *E. coli* (Suppl. Fig. 1). For the comparison of their catalytic activities, we used fixed substrate (100  $\mu$ M peptide, 200  $\mu$ M acetyl-CoA) and protein concentrations (2  $\mu$ M) in an end-point enzyme assay. GNAT1 exhibited almost no activity using the KA substrate peptide even after 4 h of incubation, while GNAT2 showed a considerable lysine acetylation activity as previously reported (Fig. 1 B; [26, 27]). The substrate preference of GNAT1

for different peptide N-termini was slightly different from that of GNAT2. While GNAT2 showed the highest activity on N-terminal Val residues, followed by Thr and Leu, GNAT1 exhibited its highest activity on N-terminal non-polar amino acids, such as Met, Ala, Gly, or Val (Fig. 1 C). Still, both enzymes were able to acetylate all provided NTA substrates. To compare their enzymatic properties, we determined the acetyl-CoA-dependent Michaelis-Menten kinetics for both GNATs on the lysine peptide as well as on one of their preferred N-terminus peptides. The Michaelis-Menten kinetics revealed a low KA activity for GNAT1, which was enhanced with increasing concentrations of acetyl-CoA and resulted in a maximum velocity  $V_{max}/[E]$  of  $1.7 \pm 0.3$  nmol product  $\cdot$  min $^{-1}$   $\cdot$   $\mu$ mol enzyme $^{-1}$ . GNAT2 in contrast, exhibited a more than 10-fold higher activity with a  $V_{max}/[E]$  of  $19.3 \pm 0.4$  nmol product  $\cdot$  min $^{-1}$   $\cdot$   $\mu$ mol enzyme $^{-1}$  (Fig. 1 D, Table 1). Moreover, GNAT2 had a lower  $K_m$ -value ( $K_m = 16.9 \pm 1.9$   $\mu$ M) for acetyl-CoA in combination with the KA peptide than GNAT1 ( $K_m = 302.1 \pm 89.1$   $\mu$ M) indicating a higher affinity for acetyl-CoA in GNAT2. When using one of the preferred peptide substrates for NTA, such as N $_{\alpha}$ -Ala, Michaelis-Menten parameters were measured for GNAT1 with a  $V_{max}/[E]$  of  $515.7 \pm 14.8$  nmol product  $\cdot$  min $^{-1}$   $\cdot$   $\mu$ mol enzyme $^{-1}$  and a  $K_m$  of  $137.6 \pm 9.8$   $\mu$ M (Fig. 1 E). Also, GNAT2 exhibited a distinct saturation curve profile on an NTA substrate peptide, here consisting of an N-terminal valine residue, with an almost three-fold higher  $V_{max}/[E]$  of  $1400.7 \pm 3.0$  nmol product  $\cdot$  min $^{-1}$   $\cdot$   $\mu$ mol enzyme $^{-1}$  than GNAT1 and a  $K_m$  of  $23.4 \pm 0.3$   $\mu$ M. The  $V_{max}/[E]$ -values for both enzymes underline the preference of both enzymes for the NTA peptide probes and underpin that GNAT2 has a higher catalytic KA and NTA activity than GNAT1 in the course of specific enzyme kinetics measurements (Table 1).

*Predicted GNAT1 and GNAT2 protein structures exhibit a high degree of similarity in the architecture of their functional acetyltransferase domains.*

GNAT1 and 2 show the highest phylogenetic relationship among the plastidial GNATs and reveal a significant amino acid sequence conservation (41% sequence identity as assigned with PRALINE Multiple Sequence Alignment tool, Suppl. Fig. 2). To compare the catalytic sites of both

enzymes we used the Pfam (Protein families) acetyltransferase domains as a basis [80]. Most notably, regions of the  $\beta$ 3- and the  $\alpha$ 3-motif, as well as two patches in between, are consistently characterized by a majority of non-polar amino acid residues. These residues form mainly groups of two or three amino acids flanking the highly conserved sites of polar residues (Fig. 2 A). As there is no experimentally determined 3D structure available, AlphaFold 2 predicted models of mature GNAT1 and GNAT2 amino acid sequences were generated for tertiary structure comparisons. These predictions reveal an identical organization of the secondary structure motives belonging to the respective acetyltransferase domain (Fig. 2 A, B). The state-of-art computational method provided by AlphaFold 2 allows highly accurate protein structure modeling without necessarily requiring experimentally validated reference structures and serves as an excellent basis for additional tools such as AlphaFill, which even enables the *in silico* incorporation of ligand molecules into the predicted structure models [42, 41]. By taking advantage of both methods and extending the computed protein structures of GNAT1 and GNAT2 by the structure model of their cofactor coenzyme A, respectively, a superimposition of both GNAT models was achieved centered on the cofactor. The identical architecture of the catalytic domains of GNAT1 and GNAT2 including the complete conformity of the primary acetyl-CoA binding sites can be derived as an obvious characteristic (Fig. 2 C). These models elucidate the observed similarities in the catalytic activities of both enzymes but do not offer insights into their differences. Those might be due to differences in their structures outside of the catalytic domain.

*GNAT1 regulates a variety of acetylation sites on plastidial proteins as identified by in vivo lysine and N-terminus acetylome profiling.*

To investigate the *in vivo* substrates of GNAT1 and to compare them to the previously identified substrates of GNAT2, we isolated two independent homozygous T-DNA insertion lines from the Arabidopsis stock centers (*gnat1-1*, *gnat1-2*; Suppl. Fig. 3, 4). Both *gnat1* lines showed no obvious growth phenotype compared to the wild type (WT), while the *gnat2-1* mutant exhibited delayed



growth as reported previously (Fig. 3 A; Table 2 [27, 31, 81]. Furthermore, each of the *gnat1* lines carries a single or a double T-DNA insertion within the first 100 bp of the exon region of the *GNAT1* gene locus (AT1G26220), as determined by sequencing of the T-DNA borders (Suppl. Fig. 3). Despite the verification of the homozygous T-DNA incorporation into the *GNAT1* gene, a trimmed *GNAT1* mRNA transcript was identified by Reverse Transcriptase-PCR for both mutant lines (Suppl. Fig. 4). This truncated transcript would lack an intact plastidial targeting sequence in case it would be translated in frame. Hence, the identified *gnat1* T-DNA lines are both treated as KO lines for the plastidial *GNAT1* protein. For proteome as well as KA and NTA analyses, we used whole rosettes of both *gnat1* T-DNA lines and of WT plants harvested 4 h after the onset of light. All plants were grown for 6-7 weeks under short-day conditions before four to eight biological replicates were collected (*gnat1-1* and *gnat1-2* in four replicates, each, and WT in eight replicates). For proteome and KA profiling, the proteins were extracted, trypsinated, and the peptides dimethyl-labeled with light and medium labels. A label swap of light and medium dimethyl-labels was introduced for two of the four *gnat1* sample replicates, respectively, and for four of the eight WT sample replicates. Differentially labeled WT and mutant peptide samples from each biological replicate were pooled for subsequent analyses. In the samples used for the quantitative proteome profiling, 5302 protein groups were identified with 2625 protein groups that were quantified in at least three biological replicates of eight replicates in total. To reveal significant differences, which are consistent in both *gnat1* mutant lines, both lines were treated as one sample group in the LIMMA statistical analyses. No significant differences in protein abundances were observed between WT and the *gnat1* mutants (Fig. 3 B, Suppl. Data 2). In the course of the KA analyses, the pooled peptides were enriched using an acetyllysine antibody coupled to agarose beads before LC-MS/MS. From 912 identified KA sites, 713 were quantified in at least three from eight biological replicates and used for further downstream analyses. In the end, only one KA site was identified as significantly downregulated in the absence of *GNAT1* in both mutants in comparison to the WT plants, but this site revealed a remarkable decrease in its acetylation abundance by a

factor of more than a hundred-fold (Fig. 3 C, Suppl. Table 1, Suppl. Data 1). The affected protein is the chloroplast-encoded PSII subunit PSBD (D2) that harbors the identified KA residue in its N-terminal domain (K7). Together with cytochrome  $b_{559}$ , PSBD forms the first subcomplex unit in the process of *de novo* PSII assembly, thereby providing the basic structure for binding and arranging additional subunits as well as redox-active cofactors [82]. In the mature PSII complex, a heterodimer composed of PSBD and PSBA (D1) represents the central reaction center core, which coordinates the reactive chlorophyll a pair of PSII [83]. Interestingly, the KA-site that has been previously identified as the most downregulated KA site in the *gnat2* acetylome profile was also found in the acetylome of the *gnat1* mutants [27]. It can be assigned to the plastid inner envelope potassium transporters KEA1 and KEA2 (K168/K170) and in *gnat1*, it showed a slight but significant upregulation in KA in comparison to the WT (Fig. 3 C, Suppl. Table 1, Suppl. Data 1). This finding might hint at a compensatory GNAT2 activity in the absence of GNAT1. The overall protein abundances of the PSBD subunit as well as of KEA1 and KEA2 were not altered between mutant and WT plants (see Fig. 3 B, Suppl. Table 1, Suppl. Data 2).

To identify the *in vivo* NTA substrates of GNAT1, samples of three biological replicates per genotype were processed according to the SILProNAQ protocol [78]. Analyses of WT and *gnat1* resulted in the identification of 14,359 N-termini corresponding to 1482 non-redundant proteoforms (Suppl. Data 3). Of these proteoforms, 633 were quantified with respect to their N-terminal acetylation yield (489 in WT as well as 498 and 493 in *gnat1-1* and *gnat1-2*, respectively). In contrast to KA, the N-terminus acetylome of both mutant plant lines revealed a higher number of peptides with de-regulated NTA compared to WT. Since the two *gnat1* KO mutants exhibited similar overall profiles, we combined the data from both mutants (Fig. 3 D, Suppl. Table 2, Suppl. Data 3). Moreover, as no major differences to the previously published WT data sets were found, we also decided to combine the WT data for a more robust and statistically relevant interpretation (Suppl. Data 8; [26, 84]). Out of 377 quantified N-termini, nine proteins were significantly decreased in their NTA yield in the *gnat1* background (i.e. > 20% decrease; Fig. 3 D, Suppl. Table 2).

Interestingly, one of these proteins, the Fructose-1,6-bisphosphatase 1 (F16P1), revealed acetylation of two contiguous neo-N-termini starting with different amino acids. Altogether, these nine proteins were predicted to be plastid-localized and their function was linked to processes such as the Calvin-Benson cycle, chlorophyll, vitamin or amino acid metabolisms as well as protein biosynthesis (Suppl. Table 2, Suppl. Data 3). The strongest decrease of NTA was found for the Phosphomethylpyrimidine synthase THIC (from a yield of 88.1% in the WT to 35.8% in the *gnat1* mutant resulting in an NTA yield difference of -52.3%), but acetylation was still detectable in the mutant background. The most remarkable decrease was observed for the N-terminus of a plastidial forkhead-domain protein (ZHD10; Suppl. Table 2, Suppl. Data 3), whose level of NTA was nearly abolished (from a yield of 38.9% in the WT to 1.6% in the *gnat1* mutant resulting in an NTA yield difference of -37.3%). Other affected proteins included the Asparagine tRNA ligase SYNO, the Fructose-1,6-bisphosphatase 1 (F16P1), the Protochlorophyllide reductase C (PORC), or the Zeaxanthin epoxidase (ZEP), which were downregulated in NTA in a range from -48.5 to -21.7% (NTA yield differences in %, *gnat1*-WT).

Next, we compared the N-termini that were mostly affected in *gnat1* with those identified in *gnat2* [26]. The sites with significantly decreased NTA in *gnat1* were also found in *gnat2* to be less acetylated, whereby the effect was at least similar or often even stronger in *gnat2* (Suppl. Fig. 5 A, B). Interestingly, a distinct number of additional proteins, such as High Chlorophyll Fluorescence Phenotype 244 (HC244), the chloroplast Translocon protein TIC55, the NDH-subunit PNSB2, the Argininosuccinate synthase ASSY, and the Light harvesting complex protein LHCA5 showed differentially acetylated N-termini only in the *gnat2* plants and revealed a severely reduced acetylation yield or even a full loss of NTA in comparison to the WT (Suppl. Fig. 5 A, B, Suppl. Tables 2, 3). We therefore concluded that the spectrum of target sites attributed to GNAT1 is mainly a subset of target sites subjected to NTA by GNAT2. However, the NTA analyses also show that GNAT2 is not able to fully compensate for the loss of GNAT1 in the *gnat1* plants (Fig. 3 D). Moreover, the data indicate that GNAT1 supports the overall GNAT2 activity.

*gnat1* mutants reveal a more compact thylakoid stacking but no defects in state transitions.

Since several proteins of the photosynthetic machinery were affected in their acetylation status in *gnat1* mutants compared to WT, we performed further physiological characterizations. No significant differences were observed concerning dry weight per plant, chlorophyll (Chl) content, and Chl *a/b* ratio, as well as basic parameters linked to photosynthesis efficiency, such as the maximum quantum yield of PSII ( $F_v/F_m$ ), and the amplitude of PSI oxidation ( $P_m$ ) (Table 2, Suppl. Table 5). On average, *gnat1* rosettes showed a slightly higher biomass compared to WT. However, this difference was not significant due to the high variance among individual plants ( $n = 6$ ), but similar observations have been made before on the analysis of a single *gnat1-2* line in a previous study [81]. On the contrary, *gnat2-1* exhibited a significantly lower dry weight compared to WT and *gnat1* plants, a higher amplitude of  $P_m$ , and a slightly lower Chl *a/b* ratio, whereby the latter was already previously reported [27, 31, 81]. Since *gnat2* plants are not capable of performing photosynthetic state transitions as previously reported and GNAT1 and GNAT2 share similar protein substrates (Fig.1 B, 1 C), we were interested to see whether *gnat1* mutants also show some deficiency in state transitioning [27, 81, 31, 32]. No defect in the formation of the PSI-LHCII state transition complex was observed in *gnat1* and WT thylakoids that were treated with red light to induce state 2 and solubilized with 1% digitonin prior to IpBN-PAGE analysis (Fig. 4). In contrast, the absence of the PSI-LHCII complex in thylakoids extracted from *gnat2-1* could be clearly recognized. The state-specific association of the LHCII antennae was additionally analyzed by recording 77K chlorophyll emission spectra of crude leaf extracts (Fig. 5 A). After red light treatment, the ratio between PSI and PSII chlorophyll fluorescence (emission maximum at 735 and 690 nm, respectively) was increased in *gnat1* extracts, mirroring the emission characteristics of WT plant extracts. As expected, no detectable PSI-LHCII state transition complex formation upon red light illumination was detected in *gnat2-1* (Fig. 5 A). Furthermore, the method of Pulse Amplitude Modulated (PAM) fluorometry was applied to track LHCII antennae adjustment in a real-

time manner depending on the predominance of red or far-red light conditions. The maximum fluorescence yields of PSII and PSII-associated LHCII chlorophyll molecules were determined under both conditions and the offset value was monitored as qT parameter for state transitions. Additionally, the balancing of energy flow (qS) upon state transitions from state 1 to state 2 can be evaluated (Table 2, Suppl. Fig. 6). Again, *gnat1-1* and *gnat1-2* did not differ from WT plants, neither in qT nor in qS. The deficiency of state transition performance of *gnat2-1* mutants, however, becomes clearly apparent, since qT and qS were notably reduced as it was also previously observed by Koskela and coworkers [27].

Since an increased association of RBCL and RCA to the thylakoid membranes has been reported for *gnat1-2* and *gnat2-1* in previous studies, we monitored the rate of CO<sub>2</sub> assimilation in the mutant lines compared to WT plants [81]. In relation to increasing light intensities applied in the experimental setup, the mutant plants exhibited a slightly higher rate of CO<sub>2</sub> fixation than WT, especially in the range of high light intensities (Suppl. Fig. 7). Starting from a PPFD of 400  $\mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  and higher, most notably *gnat1-2* showed a significantly enhanced CO<sub>2</sub> assimilation with a rate about 10-15% higher than WT. When investigating the ultrastructural thylakoid membrane organization, a similar phenotype of *gnat1* and *gnat2* lines compared to WT could be observed (Suppl. Fig. 8, 9). A more compact packing of grana stacks in *gnat2* was previously reported and could be again identified for *gnat2-1* in this study [27]. By analyzing transmission electron microscopy images obtained from mutant and WT chloroplasts, more compact grana stacks were also detected in *gnat1-1* and *gnat1-2* (Suppl. Fig. 8 A). However, *gnat1* mutant lines showed more grana stacks with a lower number of membranes compared to WT and *gnat2* (Suppl. Fig. 8 B, C). In both *gnat1* mutant lines, only 1.0-1.3% of all grana stacks consist of 10 membrane layers or more, whereas in chloroplasts of WT or *gnat2-1* mutants around five times as many can be found.

*High light treatment strongly affects the photosynthetic performance of gnat2 but not of gnat1 seedlings.*

To reveal whether the significant downregulation of PSBD-K7 acetylation in the *gnat1* mutant plants could affect the integrity of PSII upon high light treatment, we exposed eight days old seedlings grown on sterile media plates of both *gnat1* lines as well as WT and *gnat2-1* plants to continuous high light (HL) conditions or continuous darkness (D). A notable drop in the maximum quantum yield  $F_v/F_m$  was detected for *gnat2-1* treated with HL (Fig. 5 B, C, Suppl. Table 6). With a constant decrease starting from day 1 of HL, *gnat2-1* mutant seedlings exhibited a depletion of photosynthetic activity by nearly 50 % on day 3 of HL in comparison to WT and *gnat1* mutants (Fig. 5 B). The same trend was observed for the effective quantum yield  $Y(II)$  determined by analyzing the light response of the seedlings *via* Imaging-PAM measurements (Suppl. Table 6). The difference between *gnat2-1* and WT or *gnat1* plants was especially striking in the range of low- to medium-range light intensities applied in the course of the light response analyses. Interestingly, the decrease in the yield of photochemical quenching was compensated in *gnat2* on day 1 of HL mainly by an increase in the yield of non-regulated energy dissipation  $Y(NO)$ . While  $Y(NO)$  decreased back to WT or *gnat1* level on days 2 and 3, the yield of non-photochemical quenching  $Y(NPQ)$  was then increased, indicating an upregulation of light-induced energy quenching mechanisms in *gnat2-1*. This increase in  $Y(NPQ)$ , also reflected by the coefficients NPQ and  $q_N$ , required an acclimation phase, since it was only observable after two days of HL. Hence, we can conclude that the defect in lysine acetylation of PSBD in *gnat1* does not result in an altered high light sensitivity in these lines, but it appears conceivable that the more severe defect of KA and NTA of photosynthetic proteins in *gnat2-1* contributes to the high light sensitivity of this mutant.

*Pulldown and BiFC analyses reveal a strong interaction of GNAT1 with GNAT2.*

To identify proteins, which could have a regulatory impact on GNAT1, we investigated its interactions with other proteins with the help of Arabidopsis lines overexpressing GNAT1 C-terminally linked to GFP. For a quantitative comparison, a plant line overexpressing chloroplast-targeted GFP (cp-GFP) served as a background control to identify non-specifically GFP-interacting proteins [68]. A co-immunoprecipitation assay using GFP-trap beads was performed on plant leaf extracts from 6-7-weeks-old plants followed by the identification and quantification of co-isolated proteins by mass spectrometry (IP-MS) from four biological replicates each. In the GNAT1-GFP pulldown, a strong enrichment of GNAT2 (around 2000-fold) next to the pulled GNAT1-GFP was observed (Fig. 6 A). To confirm the interaction between these two GNATs, the same experiment was also conducted using a GNAT2-GFP Arabidopsis overexpression line. Also, by introducing GNAT2 as the bait for interacting proteins, the strong GNAT1-GNAT2 interaction could be verified (Fig. 6 A). Intriguingly, next to GNAT1 the third GNAT of the NAA90 subfamily, GNAT3, appeared as one of the most abundant prey proteins in the GNAT2 pulldown. In a third IP-MS experiment using GNAT3-GFP as bait for interacting proteins, the association with GNAT2 could be confirmed, whereas GNAT1 was not detected among the co-enriched proteins (Fig. 6 A). In the GNAT1-GFP pulldown, 23 additional interacting proteins were identified with an enrichment factor of at least 2-fold (Suppl. Data 4, 6). Among those proteins, Seedling Plastid Development 1 (SPD1), a member of the nucleoside-triphosphate hydrolase superfamily, was identified as 60-fold enriched. In addition to SPD1, the co-eluted proteins Accumulation and Replication of Chloroplasts 11 (ARC11), FTSZ1-1, FTSZ2-1, FTSZ2-2, CPN60A, RABE1b, and HCF136, indicate an enrichment of proteins involved in the process of plastid organization. Within this cluster, an enrichment of proteins involved in chloroplast fission (ARC11, FTSZ1-1, FTSZ2-1, FTSZ2-2) was noticed. Furthermore, the 6-Phosphogluconate Dehydrogenases PGD1 and PGD3 of the oxidative pentose phosphate pathway were co-precipitated. Enzymes from sulfate assimilation, such as the ATP sulfurylases APS1, APS2, and APS4, proteins involved in arginine

and proline synthesis, and proteins involved in photosynthesis (CRR41, LHCB4.3, HCF136, FNRL, CPN60A), plastid transcription (RNA Polymerase subunit Alpha, PTAC17) and protein synthesis (RABE1b, RPL16, SVR3) were among the co-enriched proteins. Interestingly, of these putative GNAT1 partners, no candidate was identified to undergo significant acetylation by GNAT1, neither in NTA nor in KA (Suppl. Data 1, 3).

The immunoprecipitation of GNAT2-GFP resulted in the co-enrichment of six unique plastidial proteins apart from GNAT1 and 3 (Suppl. Data 4, 6). The most co-enriched protein was the Enhancer of NSI Activity (ENA; more than 700-fold), which is a DnaJ-type chaperone with a cysteine-rich (CR) domain. An enrichment of nearly 180-fold was observed for a poorly characterized Tetratricopeptide repeat (TPR)-like superfamily protein (AT1G78915), followed by proteins enriched by about two-fold, such as two proteins involved in protein synthesis (RPS17, RPL9), the plastidial histone deacetylase 14 (HDA14), and Cold-Regulated 15B (COR15B), which is involved in the protection against abiotic stresses. Only three proteins were found co-enriched by about two-fold in both GNAT1- and GNAT2-GFP pulldowns. Among those, the Translocon at the outer Envelope Membrane of Chloroplasts 75-III (MAR1), Cold-Responsive 6.6 (KIN2), and a Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein (CYP26-2) were identified (Suppl. Data 4, 6). Our results indicate that GNAT1 and GNAT2 can form a stable complex, but both enzymes also interact with other proteins independently.

None of the plastidial proteins identified as enriched upon GNAT immunoprecipitation exhibited significant up- or downregulation in its total abundance, pointing towards a true interaction with the GNATs (Suppl. Fig. 10, Suppl. Data 5, 6). Interestingly, within the set of de-regulated protein groups in GNAT3-GFP overexpression lines, the level of quantified GNAT1 was reduced to less than 20% in comparison to the plastid GFP control line. In contrast, neither GNAT1- nor GNAT2-GFP overexpression revealed an impact on the other physiologically expressed GNATs (Suppl. Data 5, 6).



To confirm the interactions between the GNATs, we performed a bimolecular fluorescence complementation (BiFC) assay using Arabidopsis protoplasts. For this purpose, the GNATs were C-terminally fused to either an N- or a C-terminal domain of a split-YFP protein. Co-transformation of protoplasts with GNAT combinations resulted in the identification of interacting GNATs in a pairwise manner indicated by the fluorescence emission of YFP. YFP-dependent fluorescence signals that were detected by confocal laser scanning microscopy (CLSM) were reported for combinations of GNAT1-GNAT2 and GNAT2-GNAT3, whereby both GNAT variants, either coupled to the N- or to the C-terminal domain of YFP, enabled the complementation of the YFP protein (Fig. 6 B). No direct interaction between GNAT1 and GNAT3 was observed, which supports the results of the GNAT-GFP pulldowns.

*Cross-linking experiments reveal potential heterodimer formation and independent complexes between GNAT1-GNAT2 and GNAT2-GNAT3.*

To identify the interaction sites between the GNATs, the protein cross-linking reagent disuccinimidyl sulfoxide (DSSO) was used. With the help of this chemical cross-linker, lysine residues at a distance of around 10 Å can be coupled, and the linked fragments of proteins identified by mass spectrometry. Only two inter-molecular cross-linked GNAT fragments were detected in the GNAT-GFP pulldowns: one coupling GNAT1 and GNAT2, and the other one connecting GNAT2 and GNAT3 (Fig. 7 A, Suppl. Fig. 11, 12). Modeling results obtained by AlphaFold 2 Multimer structure prediction of the corresponding hetero-dimers further support a dimerization of these GNATs [43]. All lysine residues of both cross-links are located within the interface area of the corresponding dimer models with a predicted distance of 14.3 (GNAT1-GNAT2, Fig. 7 A) and 5.1 Å (GNAT2-GNAT3; Suppl. Fig. 12). Noteworthy, confidence score calculations by AlphaFold 2 strongly suggest dimerization as the most likely form of GNAT multimerization with hetero- as well as homo-dimer predictions reaching high ranks of confidence

(Table 3, Suppl. Fig. 13). Modeled oligomers consisting of more than two GNATs, however, fall substantially in their score ranking (Suppl. Table 4).

*In vitro* assays indicate that the interaction of GNAT1 and GNAT2 does not boost their acetyltransferase activities.

After discovering the mutual affinity between GNAT1 and GNAT2 by co-immunoprecipitation from *Arabidopsis* leaves, recombinant His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2 were combined in the *in vitro* activity assay to analyze the effects of their interactions on their acetyltransferase activities. The N<sub>ε</sub>-Lys as well as four different N<sub>α</sub>-N-terminus substrate peptides, which have been demonstrated to correlate with the substrate preferences of GNAT1 and/or GNAT2 (Fig. 1 C), were tested either with a single GNAT or a combination of both. Interestingly, the mixing of GNAT1 (1 μM) and GNAT2 (1 μM) did not result in a burst, not even an addition of the individually determined acetyltransferase activities (Fig. 7 B, C). The highest difference can be observed throughout the acetylation of the N<sub>α</sub>-Ala peptide. While in the case of additive total GNAT activity a substrate conversion rate of around 450 nmol · min<sup>-1</sup> would be expected, the effective catalytic activity turned out to be only about 250 nmol · min<sup>-1</sup>, far less than the individual activity of GNAT1. In a similar range, with an effective activity of only about 55% of the hypothetical activity, the N<sub>α</sub>-Met peptide substrate was acetylated by GNAT1 and GNAT2. For the substrates N<sub>α</sub>-Val and N<sub>α</sub>-Thr, a smaller difference between effective and hypothetical conversion rates was observed with a percentage of around 76% and 73% of effective activity, respectively. The smallest difference can be determined in the case of the lysine acetylation rate of the corresponding N<sub>ε</sub>-Lys peptide, which enables an effective acetylation activity of 93% for the combination of GNAT1 and GNAT2. However, since GNAT1 shows negligible rates in acetylation of the N<sub>ε</sub>-Lys substrate in the single GNAT assay, GNAT2 is probably mainly responsible for the resulting enzymatic activity (Fig. 7 B). The lower catalytic rates in the combined assay might be explained by the competition of both enzymes for their substrates.

## Discussion

Protein acetylation is a prevalent modification in chloroplasts of plants [13, 15, 85, 86]. Eight novel enzymes that possess dual lysine and N-terminus acetylation activities have been recently discovered in Arabidopsis, which are conserved in the plant kingdom [26, 25]. GNAT1, GNAT2, and GNAT3 show high phylogenetic proximity, as they cluster on one distinct branch of the derived phylogenetic tree of Arabidopsis acetyltransferases, thereby defining a specific subtype of the GNAT-related plastidial acetyltransferases [26]. This subgroup was later named NAA90, with at least one ortholog being present in all representative members of the plant kingdom [25]. While most of the GNAT mutants have no apparent growth phenotype under standard growth conditions, *gnat2* was shown to be essential for photosynthetic state transitions and has a delayed growth phenotype, especially under fluctuating light conditions [27, 31, 81]. Here, we have expanded our knowledge concerning the role of the closely related GNAT1 enzyme.

*Structural characteristics of the Arabidopsis NAA90 enzymes in comparison to other GNAT acetyltransferases.*

A structurally conserved fold of  $\beta$ -sheets and  $\alpha$ -helices, which is shared by the majority of the GNAT acetyltransferases, was previously identified [4, 1]. One of its central elements is formed by the acetyl-CoA binding site located in the turn between strand  $\beta$ 4 and helix  $\alpha$ 3. In its architecture, it is arranged by additionally recruiting helix  $\alpha$ 4 and strand  $\beta$ 4 (Fig. 8 A; [29]). Moreover, the strands  $\beta$ 1- $\beta$ 4 and  $\beta$ 5- $\beta$ 6 exhibit an almost even surface of antiparallel organization resulting in a characteristic V-shaped configuration. Accessory structural features can be typically found in flanking N- and C-terminal regions contributing to the different functions of distinct GNATs [4]. It is assumed that individual substrate preferences are mainly defined by the configuration of the  $\beta$ 6- $\beta$ 7 and the  $\alpha$ 1- $\alpha$ 2 loop, which appear to form a narrow groove or even a tunnel in case of

evident NATs, thereby impeding an interaction with lysine residues of KAT-specific substrates (Fig. 8 A; [28, 29]). The substrate binding site of a typical KAT, such as the histone acetyltransferase GCN5, exhibits a rather open architecture and a less pronounced hairpin-like loop between the sheets  $\beta 6$  and  $\beta 7$  (Fig. 8 A; [28, 87]). In the case of the new plastidial GNAT family in *Arabidopsis*, no experimentally evaluated structure information is available. An acetyltransferase with a high degree of identity to GNAT2 of *Arabidopsis* is the serotonin *N*-acetyltransferase SNAT1 of *Oryza sativa*, which was mainly characterized by its function in melatonin biosynthesis [88]. As its crystal structure was recently resolved [89], SNAT1 might allow a first substantiation of structural interpretations drawn for its closest homologs GNAT2, GNAT1, and GNAT3 in *Arabidopsis*. To reveal the characteristic features of GNAT1, novel machine learning modeling approaches such as AlphaFold 2 were used as a valuable tool helping to classify this GNAT, which provides a new layer of conclusive information that facilitates the investigation of still poorly understood functional domains of proteins [90]. A remarkable congruence between the experimentally derived SNAT1 structure and a corresponding AlphaFold 2 model can be noted (Suppl. Fig. 15 A). In a comparable range, the modeled GNAT2 and the experimentally derived SNAT1 structure resemble one another, especially in the core region flanking the acetyl-CoA binding site (Suppl. Fig. 15 C). An interesting feature of GNAT2 that differs from SNAT1 is the extended loop between the  $\beta$ -sheets 2 and 3, which shows an accumulation of serine residues directly downstream to  $\beta$ -sheet 2 followed by several residues of negative charge, such as aspartate and glutamate (Suppl. Fig. 15 C, 16). Interestingly, the predicted models of GNAT1, 2, and 3 as well as the crystal structure of SNAT1 indicate the absence of the  $\beta$ -strands 6 and 7, so that the region defining the substrate binding surface area appears smaller and more accessible than the substrate binding sites of “classical” KATs and NATs (Fig. 8 B, Suppl. Fig. 14). Since the region between sheet  $\beta 5$  and helix  $\alpha 4$  is arranged in a position directly facing the  $\alpha 1$ - $\alpha 2$  loop instead, it might contribute to the shaping of the substrate binding site as well. In GNAT1 and 2, both, the  $\alpha 1$ - $\alpha 2$  and the  $\beta 5$ - $\alpha 4$  loop exhibit a very high confidence (pLDDT > 90) in

the context of architectural organization, possibly indicating a more or less rigid conformation. In GNAT3 in contrast, the  $\alpha$ 1- $\alpha$ 2 loop appears to be much more enlarged and covers the substrate binding surface completely (Fig. 8 B), though the confidence index is low to very low (pLDDT < 70 or even < 50) and the arrangement therefore rather dynamic (Suppl. Fig. 14, 17). In addition to the structural resolution of monomeric SNAT1, the study of Liao and coworkers also provides models presenting the homodimerization of SNAT1 as its preferred configuration, thereby revealing again a high degree of congruence between the experimentally derived SNAT1 homodimer and a GNAT2 homodimer model generated by AlphaFold 2 Multimer (Suppl. Fig. 15 E; [43, 89]). Experimentally derived structure information is necessary for a robust interpretation of functional elements and a deeper understanding of structural dynamics, for instance in the process of substrate interaction or complex formation.

*NAA90 enzymes form complexes in vivo.*

The first experimental evidence of GNAT oligomerization within the chloroplast was revealed in this study by four different approaches (coIP, cross-link MS, BiFC, and peptide assays) in *in planta*, *in vivo*, and *in vitro* setups. Throughout the profiling of proteins interacting with GNAT1, we found GNAT2 to be highly co-enriched during the GNAT1-GFP pulldown experiment and could confirm this interaction in a reversed set-up (Fig. 6 A). Furthermore, GNAT3 turned out to form a stable complex with GNAT2 as shown by GNAT2-GFP as well as GNAT3-GFP immunoprecipitation. These interactions between GNAT1 and GNAT2 and GNAT2 and GNAT3 were also traceable in transient split-YFP assays in Arabidopsis protoplasts in combination with CLS microscopy (Fig. 6 B). Including a chemical protein cross-linker in the procedure of GNAT-GFP immune-precipitation further underpinned the detected interactions and uncovered the concrete lysine residues of the respective GNATs that are in close proximity (Fig. 7 A, Suppl. Fig. 12). When mapping the lysine residues involved in the GNAT1-GNAT2 and the GNAT2-GNAT3 interaction in AlphaFold 2 Multimer-predicted dimer structures, both, the experimentally as well as

computationally derived information can be consistently brought together. Notably, both cross-links comprise lysine 258 (K258) of GNAT2, the terminal amino acid residue of its C-terminus, which appears to stretch out onto the cleft-like structure of GNAT1 or GNAT3 that forms the substrate binding site in both proteins, respectively. The dimer models of GNAT1-GNAT2 and GNAT2-GNAT3 indicate that the same surface-exposed amino acid residues of GNAT2 are involved in both interactions making the formation of multimers containing more than two GNATs rather unlikely to occur (Fig. 7 A, Suppl. Fig. 12, 18). This conclusion is also supported by the AlphaFold 2 multimer-confidence score index (Table 3, Suppl. Table 4), according to which all multimeric structures of three or more GNATs drop substantially in their confidence ranking [43]. Unraveling the functional impact of the GNAT dimer formation will be a promising subject of future research.

*In planta substrate preference of GNAT1 indicates its major function as N-terminal acetyltransferase.*

In the case of the cytosolic NAT machinery, the composition of distinct complexes in conjunction with the impact of its individual subunits is already well understood for many eukaryotic organisms [15, 91, 92]. Our analyses on the interplay between recombinant GNAT1 and GNAT2 revealed that a combination of both acetyltransferases did not result in a burst of activity, not even in a summation of the individual catalytic activities that were determined in the same setup for both enzymes separately (Fig. 7 B). The highest difference between the effective substrate conversion rate measured for a GNAT1-GNAT2 combination and a hypothetical enzymatic activity was thereby observed for the N<sub>α</sub>-Ala peptide. In the presence of both GNATs, only 53% of turnover was achieved, if 100% stands for the hypothetical sum of both individual activities. N<sub>α</sub>-peptide substrates that were predominantly favored by GNAT2 showed a smaller difference between effective and hypothetical GNAT activity. In general, the *in vitro* preference of recombinant GNAT2 for Val and Thr as N-terminal amino acids for acetylation is consistent with the previously reported

activity data [26]. The remarkably higher level of enzymatic activity observed in our study might be due to the different structural properties of the recombinant GNAT2 used here and the one introduced by Bienvenut and coworkers since the previously characterized variant was N-terminally coupled to a maltose-binding protein (MBP). In the comprehensive characterization performed in this study, all GNATs have been also analyzed for their NTA-substrate preferences in a cellular context during heterologous overexpression in *E. coli* [26]. On bacterial proteins, recombinant GNAT2 revealed only a relatively small subset of NTA-targets starting with a valine residue, whereas *in vitro*, the purified GNAT2 showed a clear preference for the N $_{\alpha}$ -Val peptide. However, the major subgroup of *E. coli*-proteins targeted by GNAT2 comprised a methionine residue at their N-terminus. GNAT1 exhibited only a very low yield of NTA during heterologous overexpression on *E. coli* proteins so no clear substrate preference could be derived. Though, the limited set of substrate N-termini already indicated a specificity for GNAT1 in accordance with NatA-type acetyltransferases favoring Ala, Val, Ser, and Thr [26]. This is also reflected by our NTA profile derived from *gnat1* mutant plants in comparison to WT. In general, the divergence in the substrate spectra, which can be observed in *in vitro* test systems, *in vivo* in *E. coli*, or *in planta*, might be explained by the very different testing conditions affecting the enzymatic activity. Specific factors might modulate activities and substrate preferences of GNAT1 and GNAT2 within their physiological cellular environment resulting in altered specificities when these factors are missing.

*The interaction of GNAT1 and GNAT2 is not required for state transitions or acclimation to high light*

Very recently, a subset of the plastidial GNATs (GNAT1, GNAT2, GNAT4, GNAT6, GNAT7, and GNAT10) was analyzed in a reverse genetics approach with a focus on changes in metabolite profiles and photosynthesis [81]. There, it was reported that these GNATs have individual impacts on a variety of metabolic pathways in Arabidopsis. In comparison to the other *gnat* mutants, *gnat1* only showed a few specific features on metabolome or phenotype level, such as the slight

accumulation of free serine. Ivanauskaite and co-workers identified *gnat2* as the only GNAT-deficient plant line with a visible phenotype exhibiting lower biomass and slightly differing photosynthetic properties in comparison to WT [81]. Interestingly, the retarded growth phenotype of *gnat2* was only apparent under low light cultivation and the effects on fresh weight as well as on rosette size diminished when higher light intensities were applied. A light-dependent phenotype of *gnat2* was also described in previous studies, for instance within a comparative analysis of growth and photosynthetic performance under fluctuating light cultivation or manifested as delayed flowering and retarded developmental stage phenotype; in the latter case again especially apparent under low light conditions [31, 93]. No phenotypic difference between WT and two independent *gnat1* mutant plant lines could be revealed under the tested light conditions, as well as no impact on the plants in their performance of photosynthetic state transitions. To test, if the adaption to high light could be affected in *gnat2* or *gnat1* mutant plants, we exposed seedlings of both genotypes to continuous high light illumination and evaluated their photosynthetic performance. A substantial drop in the maximum and effective quantum efficiency was observed in *gnat2* after one day of high light, along with a notable increase in the part of non-regulated energy dissipation Y(NO) (Fig. 5 B, C, Suppl. Table 6). This finding points towards a regulatory impact of GNAT2 in the process of light acclimation, since there was no sign of recovery for the maximum and the effective quantum yields detectable in the following days of high light treatment. The part of non-photochemical quenching of excessive energy Y(NPQ) showed a notable increase from the second day after the start of high light treatment, coincidentally with a beginning decrease of Y(NO) (Suppl. Table 6). This phenotype of a delayed NPQ upregulation, plausibly in response to the high level of potentially damaging, non-regulated energy dissipation, supports an additional function of GNAT2 in light acclimation beyond its role in the regulation of state transitions. For adaption to high light conditions, the precise adjustment of both types of photosynthetic electron transport, the linear and the cyclic electron flow, is well-known to be a significant factor. The cyclic electron flow (CEF) is characterized by a cycling of electrons around PSI, thereby comprising the



cytochrome *b<sub>6</sub>f* complex, PSI, the small electron carrier Ferredoxin, and specific proteins, which are able to re-inject the shuttled electrons into the plastoquinone pool, such as the proteins PGRL1 and PGR5 or the NADH dehydrogenase-like complex NDH-1 [94–96]. Due to this, the CEF has an essential function in preserving the photosynthetic electron flow, when no final acceptors downstream from the primary light reactions, such as NADP<sup>+</sup>, are available. Moreover, it is assumed that by increasing the proton influx into the lumen the CEF may play an important role in the high light-dependent downregulation of PSII and PSI activity by non-photochemical quenching and photosynthetic control, respectively [97–99]. The identification of proteins related to CEF among the NTA substrates of GNAT2 might hint at a defect in CEF in *gnat2*. The identified proteins comprise, for instance, the NDH-1 subunit PSNB2, the Ferredoxin-NADP<sup>+</sup> reductase FNR2, as well as the plastidial translocon subunit TIC55 that is described to form a redox-regulon involved in membrane-anchoring of the FNR (Suppl. Data 3; [100–102]). Intriguingly, some of the other proteins targeted by GNAT2-mediated NTA reveal a link to the repair and/or *de novo* synthesis of PSII, namely, the PSII assembly factor HC244 or the protein LPA3, a thylakoid membrane protein, which also appears to be involved in the maturation PSII [103, 104]. In addition, the NTA site with the strongest downregulation in the NTA yield identified in *gnat2* mutants (from a yield of 99.1% in the WT to 1.8% in *gnat2*, NTA yield difference of -97.2%), can be assigned to the plastidial S-sulfocysteine synthase CYSK4. Plants deficient in CYSK4 were described to reveal an accumulation of ROS, a remarkably reduced maximum quantum yield, and an upregulation of NPQ [105, 106]. Altogether, it can be summarized that the NTA profile of *gnat2* gives a certain hint for the involvement of the CEF and/or PSII assembly machinery in the mediation of the *gnat2*-specific phenotype. However, the precise mechanisms through which GNAT2 operates need to be deciphered in future work. No apparent phenotype could be detected for seedlings of the two GNAT1-deficient plant lines upon high light treatment, even though the KA profile pointed towards a potential function of GNAT1 in photosynthetic adaptation by revealing a remarkable down-regulation of the acetylation state of the PSII core protein PSBD (D2). For the acetylation site K7

of the PSBD protein, a remaining KA abundance of less than 1 % in *gnat1* compared to the WT was detected. Together with the PSII subunit PSBA (D1), PSBD builds the reaction center of PSII in the form of a heterodimer and coordinates its central chlorophyll *a* dimer P<sub>680</sub>. Additional inner antennae proteins and several membrane-intrinsic low-molecular-mass subunits surround the reaction center, thereby forming the full PSII complex. Interestingly, PSBD is involved in the first steps of PSII *de novo* assembly by binding the cytochrome b<sub>559</sub> and subsequently initiating the PSII assembly process [107, 82]. Recently, it has been reported that a substantial increase in PSBD transcription, which was obtained by inserting a strong promoter construct upstream to the *psbd* gene, did not result in enhanced protein abundance in tobacco plants [108]. In contrast, a reduction of the PSBD transcript level led to a considerable decrease, not only in PSBD abundance but also in PSII accumulation. These findings highlight the limiting function of PSBD synthesis on PSII assembly, but also indicate a strict control of PSBD protein abundance, presumably by regulatory mechanisms that target protein stability and degradation [108]. An important factor that can modulate the lifetime of a protein is the post-translational protein modification, such as the acetylation of lysine residues. Whether the acetylation of K7 mediated by GNAT1 has a regulatory impact on the stability of the PSBD subunit, is a question to be answered in future research. Under the light conditions, which were applied in this study for the proteome analyses, no change in PSBD protein abundance could be observed in *gnat1* mutants. Intriguingly, a remarkable upregulation in the acetylation of K7 and a neighboring lysine residue, K10, was discovered previously in thylakoid preparations of a mutant line deficient in the plastidial histone deacetylase 14 (HDA14), underpinning the putatively regulatory importance of these sites [13]. Also here, the altered levels in KA of the D2 protein were not accompanied by changes in protein abundance. The decrease in NTA abundance of several protein N-termini, for instance of the neo-N-terminus V61 of the zeaxanthin epoxidase (ZEP), led to the suggestion that GNAT1 could be involved in photosynthetic acclimation. ZEP catalyzes the conversion of zeaxanthin to violaxanthin *via* a two-step epoxidation reaction, whereby the formation of zeaxanthin from

violaxanthin, in turn, is induced under high light conditions as part of the energy-dependent, rapidly reversible component qE of the NPQ [109, 110]. Moreover, when zeaxanthin accumulates, the zeaxanthin-dependent quenching component qZ is activated, which forms a proportion of the enduring, slowly reversible part of the NPQ [111]. However, when comparing the NTA yield differences determined for *gnat1* plants with the NTA profile defined for GNAT2, it becomes apparent that the target sites attributed to GNAT1 form mainly a subset of the GNAT2 NTA substrate spectrum (Suppl. Fig. 5, Suppl. Table 2, 3, Suppl. Data 3). In line with this, seedlings deficient in GNAT1 did not reveal an NPQ phenotype upon HL treatment indicating a compensatory activity of GNAT2 in these plants (Suppl. Table 6).

In summary, our study provides important first insights into the specific features of GNAT1 and the interactions of GNAT1, 2, and 3 in the NAA90 subfamily. In future studies, it will be a promising task to examine the exact role of GNAT1 in complex with GNAT2. Further stress conditions that affect PSII assembly or electron flow balancing between linear and cyclic electron transfer should be tested on *gnat1* and *gnat2* mutant plants since our KA and NTA profiles indicate a potential involvement of GNATs in the regulation of these processes. On the other hand, higher-order mutant lines deficient in GNAT1 and GNAT2, or even GNAT1, 2, and 3, will be crucial to decipher the function of these complexes for photosynthesis regulation. Acetylome data derived from knockout plants as well as from heterologous overexpression of GNATs point towards a putative mutual compensation in the acetylation of specific sites. Furthermore, by additional *in vitro* activity assays the interplay between GNAT1, 2, and 3 could be further examined to decipher the preferred modes of multimerization.

**Data Availability**

Lysine acetylome, pull-downs and full proteome data can be found at JPOST (identifier JPST002262): <https://repository.jpostdb.org/entry/JPST002262>

In addition to that, spectra are also accessible via MSviewer. The data of the coimmunoprecipitation profiles can be accessed using the following URL:

[https://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report\\_title=MS-Viewer&search\\_key=sc2f0nxgso&search\\_name=msviewer](https://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=sc2f0nxgso&search_name=msviewer)

and the data of the acK profiling here: [https://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report\\_title=MS-Viewer&search\\_key=sjczk1ixdg&search\\_name=msviewer](https://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=sjczk1ixdg&search_name=msviewer)

N-terminomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) repository via the PRIDE submission tool: Project accession: PXD046152proteomecen. For every sample, search results (.xml) and raw files (.raw) have been made available, as well as Mascot Distiller project files (.rov) to allow the viewing of annotated spectra.

**Conflict of interest**

The authors declare no competing interests.

## Acknowledgments

We would like to thank Mark Aurel Schöttler for critical readings on the photosynthetic measurements.

## Funding

This work was supported by the Professorinnenprogramm of the University of Muenster and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummern FI1655/4-1, 445970965, 469950637, 507704013 to I.F.. Further support was provided by Deutscher Akademischer Austauschdienst (DAAD) within the Programme for Project-Related Personal Exchange, the European Research Area-Coordinated Action in Plant Sciences (ERA-CAPS) Research Programme 'KatNat', the Research Council of Finland 321616, 355787 and 330083 to P.M., A.I. and M.K.. Further support was provided by the French National Research Agency (ANR) by KatNat (ERA-CAPS, ANR-17-CAPS-0001-01) and CanMore (France-Germany PRCI, ANR-20 CE92-0040) to C.G., T.M., and J.B.B. This work has benefited from the support from a French State grant (Saclay Plant Sciences, reference n° ANR-17-EUR-0007, EUR SPS-GSR) under a France 2030 program (reference n° ANR-11-IDEX-0003), and from the facilities and expertise of the I2BC proteomic platform (Proteomic-Gif, SiCaPS) supported by IBiSA, Ile de France Region, Plan Cancer, CNRS and Paris-Saclay University and from COST Action CA20113, supported by COST (European Cooperation in Science and Technology). Further support was provided by the Grants-in-Aid for Scientific Research from Japan Society for the Promotion of Science KAKENHI 23H04959 to WS and by the RECTOR program in Okayama University, Japan, to S.O. and W.S.

## Author contributions

A.B., M.F., P.H., J.E., J.B.B., A.I., M.K, U.N., S.O. performed experiments; A.B., J.E., J.B.B., T.M., C.G., I.F. analyzed data; J.E. and A.B. performed computational analysis of protein structures;

A.B. and I.F. designed research; D.S. and J.S. provided reagents; A.I., M.K. and P.M. provided plant material; I.F., C.G., P.M., D.S. and W.S. provided supervision, A.B. and I.F. drafted the manuscript with inputs from T.M., C.G., J.B.B., M.K, P.M. and J.S., D.S..

## References

1. Vetting MW, S de Carvalho LP, Yu M et al. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys* 433:212–226. <https://doi.org/10.1016/j.abb.2004.09.003>
2. Drazic A, Myklebust LM, Ree R et al. (2016) The world of protein acetylation. *Biochim Biophys Acta* 1864:1372–1401. <https://doi.org/10.1016/j.bbapap.2016.06.007>
3. Aksnes H, Ree R, Arnesen T (2019) Co-translational, Post-translational, and Non-catalytic Roles of N-Terminal Acetyltransferases. *Mol Cell* 73:1097–1114. <https://doi.org/10.1016/j.molcel.2019.02.007>
4. Deng S, Marmorstein R (2021) Protein N-Terminal Acetylation: Structural Basis, Mechanism, Versatility, and Regulation. *Trends Biochem Sci* 46:15–27. <https://doi.org/10.1016/j.tibs.2020.08.005>
5. Marmorstein R, Zhou M-M (2014) Writers and readers of histone acetylation: structure, mechanism, and inhibition. *Cold Spring Harb Perspect Biol* 6:a018762. <https://doi.org/10.1101/cshperspect.a018762>
6. Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci U S A* 51:786–794. <https://doi.org/10.1073/pnas.51.5.786>
7. Choudhary C, Kumar C, Gnad F et al. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325:834–840. <https://doi.org/10.1126/science.1175371>
8. Zhao S, Xu W, Jiang W et al. (2010) Regulation of cellular metabolism by protein lysine acetylation. *Science* 327:1000–1004. <https://doi.org/10.1126/science.1179689>
9. Finkemeier I, Laxa M, Miguet L et al. (2011) Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. *Plant Physiol* 155:1779–1790. <https://doi.org/10.1104/pp.110.171595>

10. Wu X, Oh M-H, Schwarz EM et al. (2011) Lysine acetylation is a widespread protein modification for diverse proteins in Arabidopsis. *Plant Physiol* 155:1769–1778. <https://doi.org/10.1104/pp.110.165852>
11. Hansen BK, Gupta R, Baldus L et al. (2019) Analysis of human acetylation stoichiometry defines mechanistic constraints on protein regulation. *Nat Commun* 10:1055. <https://doi.org/10.1038/s41467-019-09024-0>
12. Liu S, Yu F, Yang Z et al. (2018) Establishment of Dimethyl Labeling-based Quantitative Acetylproteomics in Arabidopsis. *Mol Cell Proteomics* 17:1010–1027. <https://doi.org/10.1074/mcp.RA117.000530>
13. Hartl M, Füll M, Boersema PJ et al. (2017) Lysine acetylome profiling uncovers novel histone deacetylase substrate proteins in Arabidopsis. *Mol Syst Biol* 13:949. <https://doi.org/10.15252/msb.20177819>
14. Uhrig RG, Schläpfer P, Roschitzki B et al. (2019) Diurnal changes in concerted plant protein phosphorylation and acetylation in Arabidopsis organs and seedlings. *Plant J* 99:176–194. <https://doi.org/10.1111/tpj.14315>
15. Linster E, Wirtz M (2018) N-terminal acetylation: an essential protein modification emerges as an important regulator of stress responses. *J Exp Bot* 69:4555–4568. <https://doi.org/10.1093/jxb/ery241>
16. Aksnes H, Drazic A, Marie M et al. (2016) First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. *Trends Biochem Sci* 41:746–760. <https://doi.org/10.1016/j.tibs.2016.07.005>
17. Linster E, Stephan I, Bienvenut WV et al. (2015) Downregulation of N-terminal acetylation triggers ABA-mediated drought responses in Arabidopsis. *Nat Commun* 6:7640. <https://doi.org/10.1038/ncomms8640>
18. Dietz K-J, Turkan I, Krieger-Liszkay A (2016) Redox- and Reactive Oxygen Species-Dependent Signaling into and out of the Photosynthesizing Chloroplast. *Plant Physiol* 171:1541–1550. <https://doi.org/10.1104/pp.16.00375>
19. Unal D, García-Caparrós P, Kumar V et al. (2020) Chloroplast-associated molecular patterns as concept for fine-tuned operational retrograde signalling. *Philos Trans R Soc Lond B Biol Sci* 375:20190443. <https://doi.org/10.1098/rstb.2019.0443>
20. Kirchhoff H (2019) Chloroplast ultrastructure in plants. *New Phytol* 223:565–574. <https://doi.org/10.1111/nph.15730>
21. Maliga P, Bock R (2011) Plastid biotechnology: food, fuel, and medicine for the 21st century. *Plant Physiol* 155:1501–1510. <https://doi.org/10.1104/pp.110.170969>

22. Rowland E, Kim J, Bhuiyan NH et al. (2015) The Arabidopsis Chloroplast Stromal N-Terminome: Complexities of Amino-Terminal Protein Maturation and Stability. *Plant Physiol* 169:1881–1896. <https://doi.org/10.1104/pp.15.01214>
23. Dinh TV, Bienvenut WV, Linster E et al. (2015) Molecular identification and functional characterization of the first N $\alpha$ -acetyltransferase in plastids by global acetylome profiling. *Proteomics* 15:2426–2435. <https://doi.org/10.1002/pmic.201500025>
24. Giglione C, Fieulaine S, Meinnel T (2015) N-terminal protein modifications: Bringing back into play the ribosome. *Biochimie* 114:134–146. <https://doi.org/10.1016/j.biochi.2014.11.008>
25. Giglione C, Meinnel T (2021) Evolution-Driven Versatility of N Terminal Acetylation in Photoautotrophs. *Trends Plant Sci* 26:375–391. <https://doi.org/10.1016/j.tplants.2020.11.012>
26. Bienvenut WV, Brünje A, Boyer J-B et al. (2020) Dual lysine and N-terminal acetyltransferases reveal the complexity underpinning protein acetylation. *Mol Syst Biol* 16:e9464. <https://doi.org/10.15252/msb.20209464>
27. Koskela MM, Brünje A, Ivanauskaite A et al. (2018) Chloroplast Acetyltransferase NSI Is Required for State Transitions in *Arabidopsis thaliana*. *Plant Cell* 30:1695–1709. <https://doi.org/10.1105/tpc.18.00155>
28. Magin RS, March ZM, Marmorstein R (2016) The N-terminal Acetyltransferase Naa10/ARD1 Does Not Acetylate Lysine Residues. *J Biol Chem* 291:5270–5277. <https://doi.org/10.1074/jbc.M115.709428>
29. Abboud A, Bédoucha P, Byška J et al. (2020) Dynamics-function relationship in the catalytic domains of N-terminal acetyltransferases. *Comput Struct Biotechnol J* 18:532–547. <https://doi.org/10.1016/j.csbj.2020.02.017>
30. Liszczak G, Arnesen T, Marmorstein R (2011) Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. *J Biol Chem* 286:37002–37010. <https://doi.org/10.1074/jbc.M111.282863>
31. Koskela MM, Brünje A, Ivanauskaite A et al. (2020) Comparative analysis of thylakoid protein complexes in state transition mutants *nsi* and *stn7*: focus on PSI and LHCII. *Photosynth Res* 145:15–30. <https://doi.org/10.1007/s11120-020-00711-4>
32. Rantala M, Ivanauskaite A, Laihonon L et al. (2022) Chloroplast Acetyltransferase GNAT2 Is Involved in the Organization and Dynamics of Thylakoid Structure. *Plant Cell Physiol*. <https://doi.org/10.1093/pcp/pcac096>
33. Galka P, Santabarbara S, Khuong TTH et al. (2012) Functional analyses of the plant photosystem I-light-harvesting complex II supercomplex reveal that light-harvesting complex II loosely bound to photosystem II is a very efficient antenna for photosystem I in state II. *Plant Cell* 24:2963–2978. <https://doi.org/10.1105/tpc.112.100339>



34. Rochaix J-D (2014) Regulation and dynamics of the light-harvesting system. *Annu Rev Plant Biol* 65:287–309. <https://doi.org/10.1146/annurev-arplant-050213-040226>
35. Minagawa J, Tokutsu R (2015) Dynamic regulation of photosynthesis in *Chlamydomonas reinhardtii*. *Plant J* 82:413–428. <https://doi.org/10.1111/tpj.12805>
36. Bellaïflore S, Barneche F, Peltier G et al. (2005) State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature* 433:892–895. <https://doi.org/10.1038/nature03286>
37. Depège N, Bellaïflore S, Rochaix J-D (2003) Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas*. *Science* 299:1572–1575. <https://doi.org/10.1126/science.1081397>
38. Pesaresi P, Hertle A, Pribil M et al. (2009) Arabidopsis STN7 kinase provides a link between short- and long-term photosynthetic acclimation. *Plant Cell* 21:2402–2423. <https://doi.org/10.1105/tpc.108.064964>
39. Pribil M, Pesaresi P, Hertle A et al. (2010) Role of plastid protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron flow. *PLoS Biol* 8:e1000288. <https://doi.org/10.1371/journal.pbio.1000288>
40. Seidel J, Klockenbusch C, Schwarzer D (2016) Investigating Deformylase and Deacylase Activity of Mammalian and Bacterial Sirtuins. *Chembiochem* 17:398–402. <https://doi.org/10.1002/cbic.201500611>
41. Hekkelman ML, Vries I de, Joosten RP et al. (2021) AlphaFill: enriching the AlphaFold models with ligands and co-factors. *bioRxiv*:2021.11.26.470110. <https://doi.org/10.1101/2021.11.26.470110>
42. Jumper J, Evans R, Pritzel A et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–589. <https://doi.org/10.1038/s41586-021-03819-2>
43. Evans R, O'Neill M, Pritzel A et al. (2021) Protein complex prediction with AlphaFold-Multimer
44. Simossis VA, Kleinjung J, Heringa J (2005) Homology-extended sequence alignment. *Nucleic Acids Res* 33:816–824. <https://doi.org/10.1093/nar/gki233>
45. Heringa J (1999) Two strategies for sequence comparison: profile-preprocessed and secondary structure-induced multiple alignment. *Comput Chem* 23:341–364. [https://doi.org/10.1016/S0097-8485\(99\)00012-1](https://doi.org/10.1016/S0097-8485(99)00012-1)
46. Schwarzländer M, Fricker MD, Müller C et al. (2008) Confocal imaging of glutathione redox potential in living plant cells. *J Microsc* 231:299–316. <https://doi.org/10.1111/j.1365-2818.2008.02030.x>

47. Née G, Kramer K, Nakabayashi K et al. (2017) DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling pathway to control seed dormancy. *Nat Commun* 8:72. <https://doi.org/10.1038/s41467-017-00113-6>
48. Uhrig JF, Mutondo M, Zimmermann I et al. (2007) The role of Arabidopsis SCAR genes in ARP2-ARP3-dependent cell morphogenesis. *Development* 134:967–977. <https://doi.org/10.1242/dev.02792>
49. Rozhon W, Mayerhofer J, Petutschnig E et al. (2010) ASKtheta, a group-III Arabidopsis GSK3, functions in the brassinosteroid signalling pathway. *Plant J* 62:215–223. <https://doi.org/10.1111/j.1365-313X.2010.04145.x>
50. Wu F-H, Shen S-C, Lee L-Y et al. (2009) Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant Methods* 5:16. <https://doi.org/10.1186/1746-4811-5-16>
51. Damm B, Schmidt R, Willmitzer L (1989) Efficient transformation of Arabidopsis thaliana using direct gene transfer to protoplasts. *Mol Gen Genet* 217:6–12. <https://doi.org/10.1007/BF00330935>
52. Frank J, Kaulfürst-Soboll H, Rips S et al. (2008) Comparative analyses of Arabidopsis complex glycan1 mutants and genetic interaction with staurosporin and temperature sensitive3a. *Plant Physiol* 148:1354–1367. <https://doi.org/10.1104/pp.108.127027>
53. Järvi S, Suorsa M, Paakkarinen V et al. (2011) Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem J* 439:207–214. <https://doi.org/10.1042/BJ20102155>
54. Porra RJ, Thompson WA, Kriedemann PE (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 975:384–394. [https://doi.org/10.1016/S0005-2728\(89\)80347-0](https://doi.org/10.1016/S0005-2728(89)80347-0)
55. Klughammer C, Schreiber U (1994) An improved method, using saturating light pulses, for the determination of photosystem I quantum yield via P700+-absorbance changes at 830 nm. *Planta* 192:261–268. <https://doi.org/10.1007/BF01089043>
56. Genty B, Briantais J-M, Baker NR (1989) The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA) - General Subjects* 990:87–92. [https://doi.org/10.1016/s0304-4165\(89\)80016-9](https://doi.org/10.1016/s0304-4165(89)80016-9)

57. Demmig-Adams B (1990) Carotenoids and photoprotection in plants: A role for the xanthophyll zeaxanthin. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1020:1–24. [https://doi.org/10.1016/0005-2728\(90\)90088-L](https://doi.org/10.1016/0005-2728(90)90088-L)
58. Schreiber U, Schliwa U, Bilger W (1986) Continuous recording of photochemical and non-photochemical chlorophyll fluorescence quenching with a new type of modulation fluorometer. *Photosynth Res* 10:51–62. <https://doi.org/10.1007/BF00024185>
59. Kramer DM, Johnson G, Kiirats O et al. (2004) New Fluorescence Parameters for the Determination of QA Redox State and Excitation Energy Fluxes. *Photosynth Res* 79:209. <https://doi.org/10.1023/B:PRES.0000015391.99477.0d>
60. Yeates AM, Zubko MK, Ruban AV (2019) Absence of photosynthetic state transitions in alien chloroplasts. *Planta* 250:589–601. <https://doi.org/10.1007/s00425-019-03187-2>
61. Ruban AV, Johnson MP (2009) Dynamics of higher plant photosystem cross-section associated with state transitions. *Photosynth Res* 99:173–183. <https://doi.org/10.1007/s11120-008-9387-x>
62. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9:671–675. <https://doi.org/10.1038/nmeth.2089>
63. Correa Galvis V, Strand DD, Messer M et al. (2020) H<sup>+</sup> Transport by K<sup>+</sup> EXCHANGE ANTIporter3 Promotes Photosynthesis and Growth in Chloroplast ATP Synthase Mutants. *Plant Physiol* 182:2126–2142. <https://doi.org/10.1104/pp.19.01561>
64. Lassowskat I, Hartl M, Hosp F et al. (2017) Dimethyl-Labeling-Based Quantification of the Lysine Acetylome and Proteome of Plants. *Methods Mol Biol* 1653:65–81. [https://doi.org/10.1007/978-1-4939-7225-8\\_5](https://doi.org/10.1007/978-1-4939-7225-8_5)
65. Giese J, Lassowskat I, Finkemeier I (2020) High-Resolution Lysine Acetylome Profiling by Offline Fractionation and Immunoprecipitation. *Methods Mol Biol* 2139:241–256. [https://doi.org/10.1007/978-1-0716-0528-8\\_18](https://doi.org/10.1007/978-1-0716-0528-8_18)
66. Boersema PJ, Raijmakers R, Lemeer S et al. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* 4:484–494. <https://doi.org/10.1038/nprot.2009.21>
67. Kulak NA, Pichler G, Paron I et al. (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* 11:319–324. <https://doi.org/10.1038/nmeth.2834>
68. Née G, Tilak P, Finkemeier I (2020) A Versatile Workflow for the Identification of Protein-Protein Interactions Using GFP-Trap Beads and Mass Spectrometry-Based Label-Free Quantification. *Methods Mol Biol* 2139:257–271. [https://doi.org/10.1007/978-1-0716-0528-8\\_19](https://doi.org/10.1007/978-1-0716-0528-8_19)

69. Hughes CS, Moggridge S, Müller T et al. (2019) Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* 14:68–85. <https://doi.org/10.1038/s41596-018-0082-x>
70. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26:1367–1372. <https://doi.org/10.1038/nbt.1511>
71. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11:2301–2319. <https://doi.org/10.1038/nprot.2016.136>
72. Cheng C-Y, Krishnakumar V, Chan AP et al. (2017) Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J* 89:789–804. <https://doi.org/10.1111/tpj.13415>
73. Liu F, Rijkers DTS, Post H et al. (2015) Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat Methods* 12:1179–1184. <https://doi.org/10.1038/nmeth.3603>
74. Hooper CM, Castleden IR, Tanz SK et al. (2017) SUBA4: the interactive data analysis centre for *Arabidopsis* subcellular protein locations. *Nucleic Acids Res* 45:D1064-D1074. <https://doi.org/10.1093/nar/gkw1041>
75. Tyanova S, Temu T, Sinitcyn P et al. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731–740. <https://doi.org/10.1038/nmeth.3901>
76. R Core Team (2021) R: A Language and Environment for Statistical Computing. <https://www.R-project.org>
77. Ritchie ME, Phipson B, Di Wu et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47. <https://doi.org/10.1093/nar/gkv007>.
78. Bienvenut WV, Giglione C, Meinel T (2017) SILProNAQ: A Convenient Approach for Proteome-Wide Analysis of Protein N-Termini and N-Terminal Acetylation Quantitation. *Methods Mol Biol* 1574:17–34. [https://doi.org/10.1007/978-1-4939-6850-3\\_3](https://doi.org/10.1007/978-1-4939-6850-3_3)
79. Bienvenut WV, Scarpelli J-P, Dumestier J et al. (2017) EnCOUNTER: a parsing tool to uncover the mature N-terminus of organelle-targeted proteins in complex samples. *BMC Bioinformatics* 18:182. <https://doi.org/10.1186/s12859-017-1595-y>
80. Finn RD, Mistry J, Tate J et al. (2010) The Pfam protein families database. *Nucleic Acids Res* 38:D211-22. <https://doi.org/10.1093/nar/gkp985>

81. Ivanauskaite A, Rantala M, Laihonon L et al. (2023) Loss of Chloroplast GNAT Acetyltransferases Results in Distinct Metabolic Phenotypes in Arabidopsis. *Plant Cell Physiol* 64:549–563. <https://doi.org/10.1093/pcp/pcad017>
82. Komenda J, Reisinger V, Müller BC et al. (2004) Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in *Synechocystis* PCC 6803. *J Biol Chem* 279:48620–48629. <https://doi.org/10.1074/jbc.M405725200>
83. Dekker JP, Boekema EJ (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim Biophys Acta* 1706:12–39. <https://doi.org/10.1016/j.bbabi.2004.09.009>
84. Eirich J, Boyer J-B, Armbruster L et al. (2024) Light changes promote distinct responses of plastid protein acetylation marks \_submitted
85. Xue C, Liu S, Chen C et al. (2018) Global Proteome Analysis Links Lysine Acetylation to Diverse Functions in *Oryza Sativa*. *Proteomics* 18. <https://doi.org/10.1002/pmic.201700036>
86. Tan D-X, Reiter RJ (2020) An evolutionary view of melatonin synthesis and metabolism related to its biological functions in plants. *J Exp Bot* 71:4677–4689. <https://doi.org/10.1093/jxb/eraa235>
87. Rojas JR, Trievel RC, Zhou J et al. (1999) Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* 401:93–98. <https://doi.org/10.1038/43487>
88. Byeon Y, Back K (2016) Low melatonin production by suppression of either serotonin N-acetyltransferase or N-acetylserotonin methyltransferase in rice causes seedling growth retardation with yield penalty, abiotic stress susceptibility, and enhanced coleoptile growth under anoxic conditions. *J Pineal Res* 60:348–359. <https://doi.org/10.1111/jpi.12317>
89. Liao L, Zhou Y, Xu Y et al. (2021) Structural and Molecular Dynamics Analysis of Plant Serotonin N-Acetyltransferase Reveal an Acid/Base-Assisted Catalysis in Melatonin Biosynthesis. *Angew Chem Int Ed Engl* 60:12020–12026. <https://doi.org/10.1002/anie.202100992>
90. Gomes PSFC, Gomes DEB, Bernardi RC (2022) Protein structure prediction in the era of AI: Challenges and limitations when applying to in silico force spectroscopy. *Front Bioinform* 2:983306. <https://doi.org/10.3389/fbinf.2022.983306>
91. Arnesen T, van Damme P, Polevoda B et al. (2009) Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc Natl Acad Sci U S A* 106:8157–8162. <https://doi.org/10.1073/pnas.0901931106>
92. Deng S, McTiernan N, Wei X et al. (2020) Molecular basis for N-terminal acetylation by human NatE and its modulation by HYPK. *Nat Commun* 11:818. <https://doi.org/10.1038/s41467-020-14584-7>

93. Lee HY, Back K (2021) Melatonin Regulates Chloroplast Protein Quality Control via a Mitogen-Activated Protein Kinase Signaling Pathway. *Antioxidants* (Basel) 10. <https://doi.org/10.3390/antiox10040511>
94. Alboresi A, Storti M, Morosinotto T (2019) Balancing protection and efficiency in the regulation of photosynthetic electron transport across plant evolution. *New Phytologist* 221:105–109. <https://doi.org/10.1111/nph.15372>
95. DalCorso G, Pesaresi P, Masiero S et al. (2008) A complex containing PGRL1 and PGR5 is involved in the switch between linear and cyclic electron flow in Arabidopsis. *Cell* 132:273–285. <https://doi.org/10.1016/j.cell.2007.12.028>
96. Shikanai T, Endo T, Hashimoto T et al. (1998) Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. *Proc Natl Acad Sci U S A* 95:9705–9709. <https://doi.org/10.1073/pnas.95.16.9705>
97. Sato R, Ohta H, Masuda S (2014) Prediction of respective contribution of linear electron flow and PGR5-dependent cyclic electron flow to non-photochemical quenching induction. *Plant Physiol Biochem* 81:190–196. <https://doi.org/10.1016/j.plaphy.2014.03.017>
98. Ruban AV (2016) Nonphotochemical Chlorophyll Fluorescence Quenching: Mechanism and Effectiveness in Protecting Plants from Photodamage. *Plant Physiol* 170:1903–1916. <https://doi.org/10.1104/pp.15.01935>
99. Yamamoto H, Shikanai T (2019) PGR5-Dependent Cyclic Electron Flow Protects Photosystem I under Fluctuating Light at Donor and Acceptor Sides. *Plant Physiol* 179:588–600. <https://doi.org/10.1104/pp.18.01343>
100. Sirpiö S, Allahverdiyeva Y, Holmström M et al. (2009) Novel nuclear-encoded subunits of the chloroplast NAD(P)H dehydrogenase complex. *J Biol Chem* 284:905–912. <https://doi.org/10.1074/jbc.M805404200>
101. Lintala M, Schuck N, Thormählen I et al. (2014) Arabidopsis *tic62* *trp* mutant lacking thylakoid-bound ferredoxin-NADP<sup>+</sup> oxidoreductase shows distinct metabolic phenotype. *Mol Plant* 7:45–57. <https://doi.org/10.1093/mp/sst129>
102. Kramer M, Rodriguez-Heredia M, Saccon F et al. (2021) Regulation of photosynthetic electron flow on dark to light transition by ferredoxin:NADP(H) oxidoreductase interactions. *Elife* 10. <https://doi.org/10.7554/eLife.56088>
103. Hey D, Grimm B (2018) ONE-HELIX PROTEIN2 (OHP2) Is Required for the Stability of OHP1 and Assembly Factor HCF244 and Is Functionally Linked to PSII Biogenesis. *Plant Physiol* 177:1453–1472. <https://doi.org/10.1104/pp.18.00540>

104. Schneider A, Steinberger I, Strissel H et al. (2014) The Arabidopsis Tellurite resistance C protein together with ALB3 is involved in photosystem II protein synthesis. *Plant J* 78:344–356. <https://doi.org/10.1111/tpj.12474>
105. Bermúdez MA, Páez-Ochoa MA, Gotor C et al. (2010) Arabidopsis S-sulfocysteine synthase activity is essential for chloroplast function and long-day light-dependent redox control. *Plant Cell* 22:403–416. <https://doi.org/10.1105/tpc.109.071985>
106. Bermúdez MÁ, Galmés J, Moreno I et al. (2012) Photosynthetic adaptation to length of day is dependent on S-sulfocysteine synthase activity in the thylakoid lumen. *Plant Physiol* 160:274–288. <https://doi.org/10.1104/pp.112.201491>
107. Müller B, Eichacker LA (1999) Assembly of the D1 Precursor in Monomeric Photosystem II Reaction Center Precomplexes Precedes Chlorophyll a –Triggered Accumulation of Reaction Center II in Barley Etioplasts. *Plant Cell* 11:2365–2377. <https://doi.org/10.1105/tpc.11.12.2365>
108. Fu H-Y, Ghandour R, Ruf S et al. (2021) The availability of neither D2 nor CP43 limits the biogenesis of photosystem II in tobacco. *Plant Physiol* 185:1111–1130. <https://doi.org/10.1093/plphys/kiab052>
109. Kaiser E, Correa Galvis V, Armbruster U (2019) Efficient photosynthesis in dynamic light environments: a chloroplast's perspective. *Biochem J* 476:2725–2741. <https://doi.org/10.1042/BCJ20190134>
110. Niyogi KK, Grossman AR, Björkman O (1998) Arabidopsis mutants define a central role for the xanthophyll cycle in the regulation of photosynthetic energy conversion. *Plant Cell* 10:1121–1134. <https://doi.org/10.1105/tpc.10.7.1121>
111. Nilkens M, Kress E, Lambrev P et al. (2010) Identification of a slowly inducible zeaxanthin-dependent component of non-photochemical quenching of chlorophyll fluorescence generated under steady-state conditions in Arabidopsis. *Biochim Biophys Acta* 1797:466–475. <https://doi.org/10.1016/j.bbabi.2010.01.001>

## Figure and Table legends

### Figure 1. Biochemical characterization of recombinant His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2 reveals dual lysine and N-terminal acetylation activities. (A)

Synthetic substrate peptides used for N<sub>ε</sub>-Lysine (KA) and N<sub>α</sub>-N-terminus acetylation (NTA) assays in an HPLC-based set-up. The introduced dinitroaniline group allows a specific detection of the peptides at a wavelength of 360 nm. The amino acid residues were labeled in one-letter code. Ahx refers to 6-aminohexanoic acid and *r* to D-arginine. (B) KA and (C) NTA (Met, Ala, Gly, Val, Leu, Ser, Thr) activities of recombinant His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2. The activities were determined by end-point analysis of the acetylated fraction that was generated with the different synthetic substrate peptides after 4h (B) and 1h (C), respectively. The recombinant GNATs (2 μM) were incubated with the synthetic peptide substrate (100 μM) and the reaction was started by the addition of 200 μM acetyl-CoA at 30 °C. A fraction corresponding to 1000 pmol of peptide was analyzed by HPLC. The bar graphs (B, C) represent the final fraction of acetylated peptide (product) generated by 20 pmol of the recombinant GNAT (n = 3, ± SD). Grey rectangles show the individual activities of each replicate. Michaelis-Menten kinetics of the KA (D) and the NTA activity (E) performed with His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2, respectively. The catalytic activities were determined with increasing concentrations of acetyl-CoA (0, 25, 50, 100, 200, 500 μM) and the resulting curves were fitted according to the Michaelis-Menten function  $f(x) = V_{max} \cdot x \cdot (K_m + x)^{-1}$ . A fixed peptide substrate concentration of 100 μM was used (n = 3, ± SD).

### Figure 2. The predicted structures of GNAT1 and GNAT2 acetyltransferase domains are highly similar. (A)

Schematic overview of primary and secondary structures of GNAT1 and GNAT2 centered on the region annotated as acetyltransferase Pfam domain (PF00583, aligned region; Finn *et al.*, 2010). The colors, which highlight different amino acid residues displayed as one-letter code, represent the side chain features as follows: orange, non-polar; yellow, aromatic; green, polar (neutral); blue, positively charged, and red, negatively charged. The regions of acetyl-



CoA binding are displayed below the primary amino acid sequence with the boxed region representing the main binding site (Bienvenut *et al.*, 2020). In the third row, secondary structure elements determined by AlphaFold 2 structure prediction were marked by red ( $\alpha$ -helix) or blue ( $\beta$ -sheet) boxes (Jumper *et al.*, 2021). The consensus sequence combining the primary structure features of GNAT1 and GNAT2 is shown below. Amino acid residues, which are identical between both acetyltransferases, are shown in one-letter code, while differing amino acids in corresponding positions were replaced by an 'x'. **(B)** AlphaFold 2 structure predictions for the acetyltransferase domains of GNAT1 and GNAT2. Protein structure data were obtained from the AlphaFold 2 database and visualized using PyMOL (The PyMOL Molecular Graphics System, Version 4.5 Schrödinger, LLC). **(C)** Superimposition of the predicted acetyltransferase domain models presented in (B). Protein structure predictions of the AlphaFold 2 platform were extended by integration of the cofactor coenzyme A (CoA) as a stick model with the help of the AlphaFill tool and a superimposition of both GNATs was performed using PyMOL (Hekkelman *et al.*, 2021). The dashed yellow line highlights the main acetyl-CoA binding site of both GNATs.

**Figure 3. GNAT1 regulates a variety of lysine and N-terminal acetylation sites but shows no obvious growth phenotype. (A)** Phenotype of *gnat1* mutant lines in comparison to the wild type (WT) and the GNAT2-deficient knockout line *gnat2-1*. Plants were cultivated under short-day conditions (8 h light/ 16 h dark) with PPDF =  $120 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ , humidity = 50% and temperature = 22°C for six to seven weeks. Protein abundance **(B)**, lysine acetylation KA **(C)**, and N-terminal acetylation NTA **(D)** proteomic profiles displaying up- or downregulation in GNAT1 deficient plants in comparison to the WT. Sums indicate the total numbers of quantified protein groups or acetylation sites. For statistical evaluation of (B) and (C), the lines *gnat1-1* and *gnat1-2* were treated as one group and per protein group or protein site, a quantifiable identification in at least three out of eight biological replicates was required to be considered as valid. Altogether, four biological replicates of *gnat1-1* and of *gnat1-2*, respectively, and eight replicates of WT plants were processed, amounting to a number of four combined *gnat1-1*/WT and four combined

*gnat1-2*/WT samples. Green circles highlight plastid localized proteins, which show significant up- or downregulation in abundance or acetylation of a specific site ( $\log_2$  fold change  $\leq -0.5$  or  $\geq 0.5$  and false discovery rate-corrected p-value  $\leq 0.05$  [LIMMA]). For the investigation of the NTA profile of *gnat1* (D), the lines *gnat1-1* and *gnat1-2* were treated as one group, and the WT profile was derived by combining the data obtained in this work and previously published data set (Bienvenut *et al.*, 2020; Eirich *et al.*, 2024). For the experimental work that was performed in the course of the present study, three biological replicates of *gnat1-1*, of *gnat1-2*, and of WT plants, respectively, were processed. Significantly up- or downregulated N-termini that originate from proteins with plastid localization are marked by green color. The NTA differences between *gnat1* and WT were analyzed with GraphPad 10 and plotted as yield difference in % (*gnat1* - WT). An unpaired t-test analysis with FDR 1% was performed and N-termini displaying >20% difference in their NTA yield were regarded as significant.

1. THIC (T55, AT2G29630) 2. SYNO (T60, AT4G17300) 3a. F16P1 (V61, AT3G54050) 3b. F16P1 (A60, T3G54050) 4. ZHD10 (A72, AT2G21530) 5. ZEP (A61, AT5G67030) 6. PORC (T69, AT1G03630) 7. NAGK (T51, AT3G57560) 8. AT2G14880 (T46) 9. MORF9 (T61, AT1G11430)

**Figure 4. *gnat1* mutant plants are able to perform state transitions.** Large-pore blue native gel (lpBN-PAGE) images of thylakoid protein complexes isolated from Arabidopsis *gnat1-1*, *gnat1-2*, *gnat2-1*, and wild type (WT) plants. GL (PPFD =  $120 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) adapted plants were either treated with red (PPFD =  $53 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) or far-red light (PPFD =  $30 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) for at least one hour before thylakoid membranes were extracted and subsequently solubilized with 1% digitonin or  $\beta$ -dodecylmaltoside (DM). The solubilized protein complexes were separated by lpBN-PAGEs under native conditions. mc, megacomplex; sc, supercomplex; t, trimer; d, dimer; m, monomer.

**Figure 5. Phenotyping of *gnat1* and *gnat2* mutants after high light treatment. (A)** 77K chlorophyll fluorescence emission spectra displaying the chlorophyll emission signal of crude leaf extracts after red (R, state 2) and far-red light treatment (FR, state 1) of *gnat1-1*, *gnat1-2*, *gnat2-1*, and wild type (WT) plants. Spectra were normalized to the fluorescence peak at 685 nm and represent the average of three replicates per genotype. The fluorescence maximum around 685 nm originates from chlorophyll molecules associated with PSII, whereas the second maximum at 735 nm is the result of PSI-specific chlorophyll fluorescence emission. **(B)** Maximum quantum yields ( $F_v/F_m$ ) determined for *gnat1-1*, *gnat1-2*, *gnat2-1*, and WT seedlings treated with standard growth light (GL, PPDF =  $120 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ), continuous high light (HL, PPDF =  $800 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) or darkness for the indicated number of days. Plants were grown on plates containing sterile media for eight days under long-day conditions (16 h light/ 8 h darkness, PPDF =  $120 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ) before the treatment. Photosynthetic performances of the seedlings were analyzed by using the Imaging-PAM system (Walz) and differences between genotypes were tested for significance by One-Way ANOVA applying a Fisher LSD test with a significance level of  $p \leq 0.05$ . **(C)** Exemplary images of the analyzed seedlings grown on sterile media that were colored in relation to the maximum quantum yields ( $F_v/F_m$ ). These were determined after the indicated treatments. The correlation between colors and numerical values is predefined by the Imaging-PAM software and ranges from minimal  $F_v/F_m$  (0, black) to maximal  $F_v/F_m$  (1, magenta).

**Figure 6. GNAT1 exhibits a strong interaction with GNAT2 as revealed by co-immunoprecipitation and BiFC analyses.** Volcano plots of co-immunoprecipitation assays performed on leaf material of GNAT1-, GNAT2- or GNAT3-GFP overexpressing plant lines **(A)**. Sums represent the total number of quantified protein groups per experiment by Label-Free-Quantification (LFQ) *via* MaxQuant. For each protein group, the  $\log_2$  fold change was calculated by comparing LFQ intensities of the respective GNAT-GFP experiment with those of the control experiment comprising the co-immunoprecipitation of chloroplast-targeted GFP (cpGFP). Per genotype, four biological replicates were harvested and processed. To be recognized as

significant, valid quantifications in at least three of four biological replicates per experiment were required, and in addition, enriched protein groups had to possess a  $\log_2$  fold change of  $\geq 1$  and a false discovery rate-corrected p-value  $\leq 0.05$  (LIMMA). Green circles highlight protein groups with assigned chloroplast localization according to Suba4 [74]. Yellow circles represent the GFP protein that was found in the control experiment as well as in samples of the GNAT-GFP overexpressors as part of the respective fusion proteins. **(B)** Confocal Laser Scanning Microscopy (CLSM) analysis of protein interactions between GNAT1, GNAT2, and GNAT3 by Bimolecular Fluorescence Complementation (BiFC). Arabidopsis wild type (WT) protoplasts were transiently transformed with GNAT plasmid constructs enabling the expression of recombinant GNATs C-terminally fused to a split-YFP domain. Per assays, combinations of two GNATs were tested for interaction, which can be detected by YFP fluorescence emission as a result of sYFPN- and sYFPC-domain interplay. The main panels represent protoplasts in the channel of YFP detection showing either a fluorescence signal or in case of missing interaction a weak background emission obtained under the same conditions of signal detection. The three small images on the right side of each main panel display the chlorophyll autofluorescence (top), a merged view of YFP and autofluorescence signals (middle), and a bright-field image (bottom) of the respective protoplast. The scale bar corresponds to a distance of 20  $\mu\text{m}$ .

**Figure 7. Chemical cross-linking of GNAT1 and 2 during co-immunoprecipitation and combined *in vitro* activity assays.** **(A)** Structure predictions for the heterodimer GNAT1-GNAT2 derived by AlphaFold 2 Multimer and visualized using PyMOL (Evans *et al.*, 2021). Sequence data were obtained from the Araport 11 database and N-terminally truncated by the amino acid sequence corresponding to predicted chloroplast transit peptides. The lysine residues involved in the formation of the identified DSSO protein cross-link were shown as red stick models covered by the red, partially transparent mesh structure of the computed protein surface. The dotted yellow line marks the distance between the two amino groups of the lysine pair revealing a value of 14.3 Å. The position of the CoA molecule presented as a stick model was derived from a structure

prediction provided by the AlphaFill database (Hekkelman *et al.*, 2021). N<sub>ε</sub>-Lysine **(B)** and N<sub>α</sub>-N-terminus acetylation **(C)** activity defined for single GNATs and a combination of His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2. Both proteins were used in a concentration of 1 μM (single) or 1+1 μM (GNAT1 + GNAT2) and mixed with a synthetic peptide substrate in a concentration of 100 μM. Subsequently, the reaction was started by adding 200 μM of the acetyl-group donor acetyl-CoA, and an aliquot corresponding to 1000 pmol was analyzed by HPLC to detect a retention-time shift indicating the addition of an acetyl-group. The bar graph represents the fractions of acetylated peptide (product) generated per minute within the linear phase of the acetylation reaction (n = 3, ± SD).

**Figure 8. AlphaFold 2 structure predictions indicate the presence of accessory features flanking the more conserved ‘core’ domain. (A)** Crystal structures of NAA10 from *Schizosaccharomyces pombe* (PDB code 4KVM) and GCN5 from *Tetrahymena thermophila* (PDB code 1QSN) as well as structure models of GNAT1, 2, and 3 **(B)** provided by AlphaFold 2. The region of the α1-α2 loop is colored in red and the β6-β7 loop or the corresponding β5-α4-loop in GNAT1, 2, and 3, respectively, is colored in magenta. Each GNAT model was trimmed by the N-terminal region representing the plastid transit peptide. Furthermore, the extensive N-terminal loop of GNAT2 was removed from the corresponding images as indicated by the dashed black lines.

**Table 1. Enzymatic properties of recombinant His<sub>6</sub>-GNAT1 and His<sub>6</sub>-GNAT2.** The catalytic activities were measured with increasing concentrations of acetyl-CoA (0, 25, 50, 100, 200, 500 μM) and a fixed peptide substrate concentration of 100 μM (n = 3). The resulting curves were fitted according to the Michaelis-Menten function  $f(x) = V_{\max} \cdot x \cdot (K_m + x)^{-1}$  and the parameters  $V_{\max}$ ,  $K_m$ , and  $k_{\text{cat}}$  are shown as averages ± SD.

**Table 2: Physiological and photosynthesis-related parameters investigated for the wild type and the mutant lines *gnat1-1*, *gnat1-2*, and *gnat2*.** Differences between genotypes were

tested with ANOVA and for multiple comparisons, Fisher's LSD test was applied ( $n = 6$ ).  $F_v/F_m = (F_m - F_0)/F_m$ ,  $qT = (F_{m1} - F_{m2})/F_{m1}$ , and  $qS = (F_{S1'} - F_{S2'})/(F_{S1'} - F_{S2'})$ .

**Table 3: Confidence score values calculated by AlphaFold 2 Multimer for all homo- and hetero-dimer predictions including GNAT1, 2, and 3.** The average quality of each complex is evaluated by the intrinsic model accuracy estimates pTM and ipTM (predicted Template Modeling Score and interface predicted Template Modeling Score) and represented as a value between 0-1.

### Supplemental Figures

Suppl. Fig. 1\_Sequence alignment - GNAT1 and GNAT2.pdf

Suppl. Fig. 2\_Preparations of recombinant GNAT1 and GNAT2.pdf

Suppl. Fig. 3\_Confirmation of T-DNA insertions in *gnat1*.pdf

Suppl. Fig. 4\_Analysis of *gnat1* transcript abundance.pdf

Suppl. Fig. 5\_NTA yield difference – *gnat1* vs. *gnat2*.pdf

Suppl. Fig. 6\_Principle of qT determination.pdf

Suppl. Fig. 7\_CO<sub>2</sub> assimilation rates.pdf

Suppl. Fig. 8\_TEM graphs – Grana thylakoid organization.pdf

Suppl. Fig. 9\_TEM images.pdf

Suppl. Fig. 10\_CoIP experiment – total proteomes.pdf

Suppl. Fig. 11\_MS-specific properties of cross-linked peptides.pdf

Suppl. Fig. 12\_AF2-M model – GNAT2-GNAT3 cross-link

Suppl. Fig. 13\_AF2-M models of unconfirmed GNAT homo- and heterodimer.pdf

Suppl. Fig. 14\_GNAT monomers modeled by AF2 - per-residue confidences.pdf

Suppl. Fig. 15\_Structure superimpositions - GNAT models and SNAT1 crystal structures.pdf

Suppl. Fig. 16\_Sequence alignments - SNAT1 and GNAT1-3.pdf

Suppl. Fig. 17\_GNAT3 AT-domain and structural features of GNAT1-3.pdf

Suppl. Fig. 18\_Structural features of GNAT1-3 in the dimer constellations GNAT1-GNAT2 and GNAT2-GNAT3.pdf

### **Supplemental Tables**

Suppl. Table 1+2+3\_Overview of plastid KA+NTA sites.pdf

Suppl. Table 4\_Confidence of GNAT multimer predictions by AF2-Multimer.pdf

Suppl. Table 5\_Light response curve\_Dual-PAM100.pdf

Suppl. Table 6\_Seedlings under HL+D\_Imaging-PAM.pdf

### **Supplemental Data**

Suppl. Data 1+2\_ Suppl. Data 1+2\_Complete - KA profile+total proteome\_*gnat1*.xlsx

Suppl. Data 3\_Complete - NTA profiles *gnat1* and *gnat2*.xlsx

Suppl. Data 4+5\_Overview - Significant co-enriched protein groups+total proteome changes.xlsx

Suppl. Data 6\_Complete - GFP-CoIP profiles and total proteomes.xlsx

Suppl. Data 7\_Instrumentation used for the respective experiments

Suppl. Data 8\_ Arabidopsis WT total N-term eNcounter dataset containing all the identified N-termini of the 3 independent WT samples used in this report. Graphs were made from datasets #1, showing the protein NTA yields in the "2019", "Jul .2020" and "Oct. 2020" WT samples which were compared one to another.

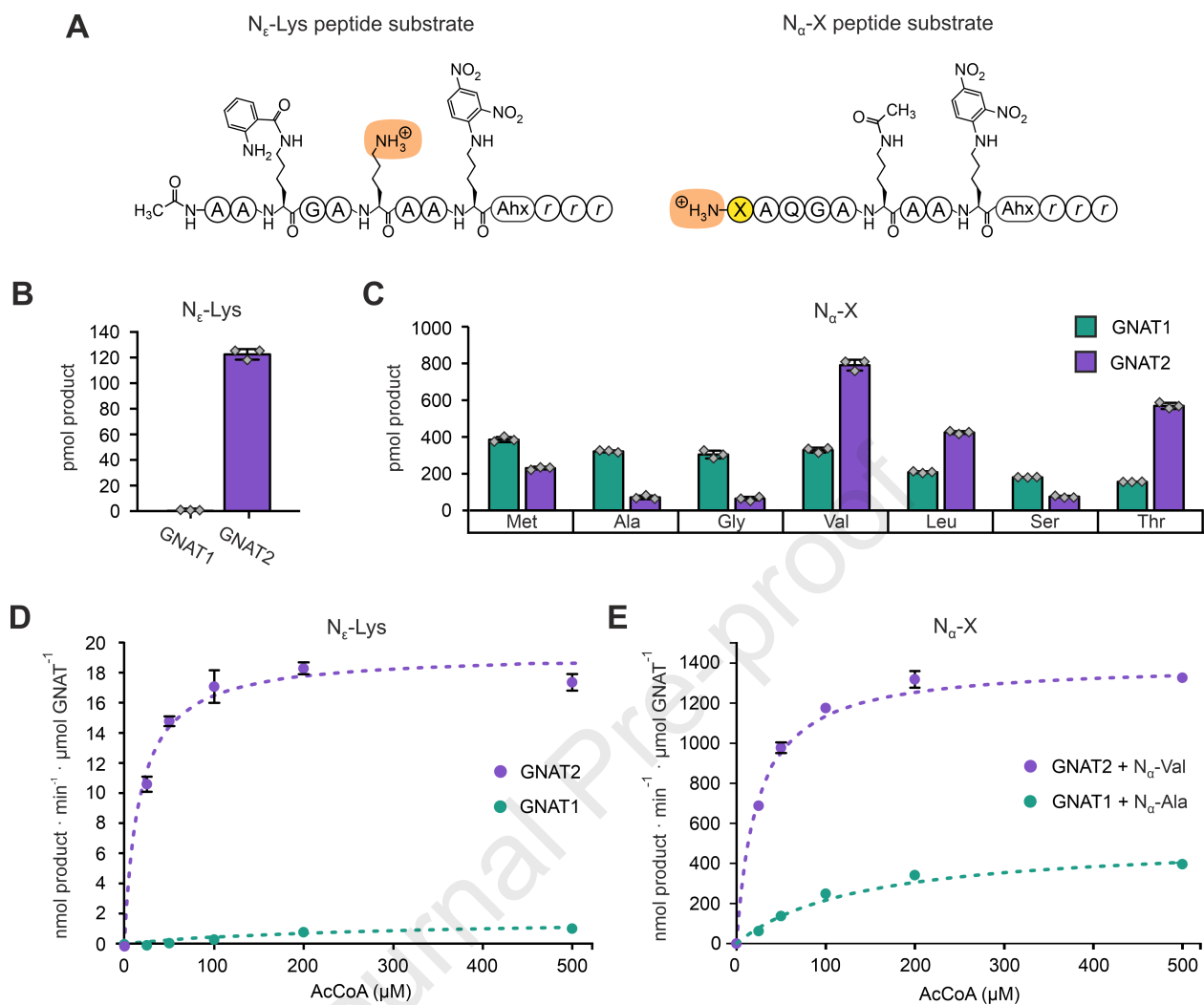
Journal Pre-proof



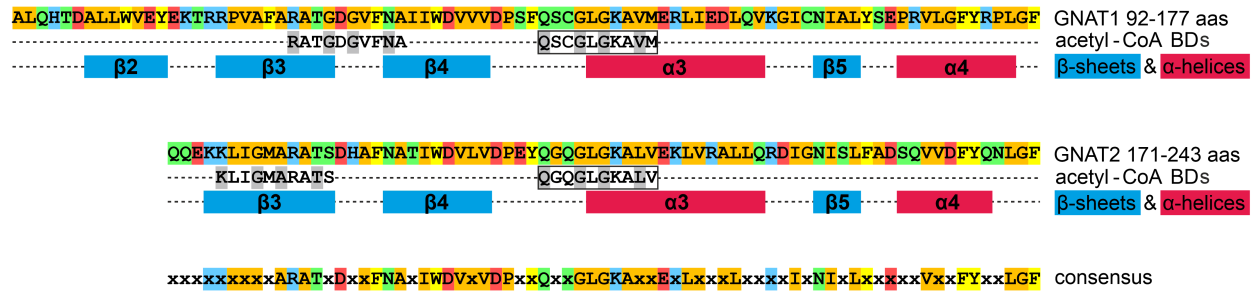
|  | KA activity<br>N $\epsilon$ -Lys peptide substrate |                    | NTA activity<br>N $\alpha$ -X peptide substrate |                            |
|--|--|--------------------|---|----------------------------|
|  | GNAT1  | GNAT2              | GNAT1<br>(N $\alpha$ -Ala)                      | GNAT2<br>(N $\alpha$ -Val) |
| $V_{\max}$<br>(nmol·min $^{-1}$ · $\mu$ mol GNAT $^{-1}$ ) | 1.74 $\pm$ 0.28                                    | 19.26 $\pm$ 0.43   | 515.7 $\pm$ 14.8                                | 1400.7 $\pm$ 3.0           |
| $K_m$<br>( $\mu$ M)  | 302.1 $\pm$ 89.1                                   | 16.9 $\pm$ 1.9     | 137.6 $\pm$ 9.8                                 | 23.44 $\pm$ 0.32           |
| $k_{\text{cat}}$<br>(min $^{-1}$ )                         | 0.002 $\pm$ 0.0003                                 | 0.019 $\pm$ 0.0004 | 0.515 $\pm$ 0.015                               | 1.401 $\pm$ 0.003          |

| Plant line     | Dry weight (mg)              | Chl a+b ( $\mu\text{g}/\text{mg}$ ) | Chl a/b                       | F <sub>v</sub> /F <sub>m</sub> | P <sub>m</sub>                | qT                           | qS                           |
|----------------|------------------------------|-------------------------------------|-------------------------------|--------------------------------|-------------------------------|------------------------------|------------------------------|
| Wild type      | 59.9 $\pm$ 14.2 <sup>a</sup> | 1.02 $\pm$ 0.11 <sup>a</sup>        | 3.64 $\pm$ 0.14 <sup>a</sup>  | 0.81 $\pm$ 0.01 <sup>a</sup>   | 0.84 $\pm$ 0.10 <sup>a</sup>  | 0.11 $\pm$ 0.01 <sup>a</sup> | 0.91 $\pm$ 0.01 <sup>a</sup> |
| <i>gnat1-1</i> | 66.2 $\pm$ 14.6 <sup>a</sup> | 1.02 $\pm$ 0.06 <sup>a</sup>        | 3.59 $\pm$ 0.15 <sup>a</sup>  | 0.81 $\pm$ 0.01 <sup>a</sup>   | 0.94 $\pm$ 0.09 <sup>bc</sup> | 0.10 $\pm$ 0.01 <sup>a</sup> | 0.91 $\pm$ 0.01 <sup>a</sup> |
| <i>gnat1-2</i> | 64.9 $\pm$ 14.7 <sup>a</sup> | 1.02 $\pm$ 0.09 <sup>a</sup>        | 3.55 $\pm$ 0.12 <sup>ab</sup> | 0.81 $\pm$ 0.02 <sup>a</sup>   | 0.91 $\pm$ 0.09 <sup>ab</sup> | 0.11 $\pm$ 0.01 <sup>a</sup> | 0.91 $\pm$ 0.02 <sup>a</sup> |
| <i>gnat2-1</i> | 45.3 $\pm$ 13.2 <sup>b</sup> | 1.01 $\pm$ 0.11 <sup>a</sup>        | 3.37 $\pm$ 0.13 <sup>b</sup>  | 0.81 $\pm$ 0.01 <sup>a</sup>   | 1.02 $\pm$ 0.06 <sup>c</sup>  | 0.00 $\pm$ 0.01 <sup>b</sup> | 0.28 $\pm$ 0.02 <sup>b</sup> |

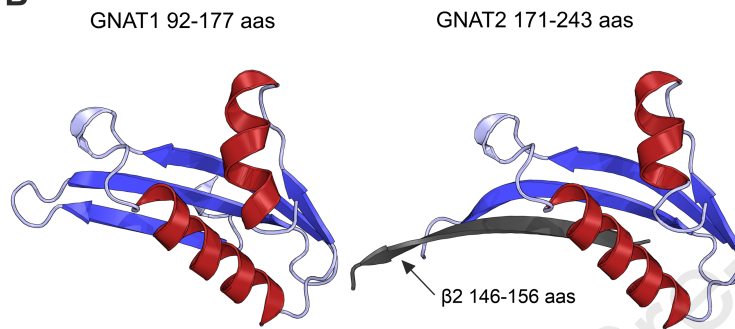
| Modeled GNAT dimer | AF2 confidence<br>(iptm+ptm) |
|--------------------|------------------------------|
| GNAT1-GNAT1        | 0.882                        |
| GNAT2-GNAT2        | 0.817                        |
| GNAT3-GNAT3        | 0.802                        |
| GNAT1-GNAT2        | 0.882                        |
| GNAT1-GNAT3        | 0.881                        |
| GNAT2-GNAT3        | 0.815                        |



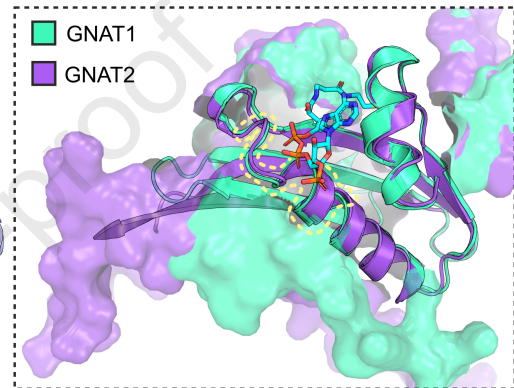
A

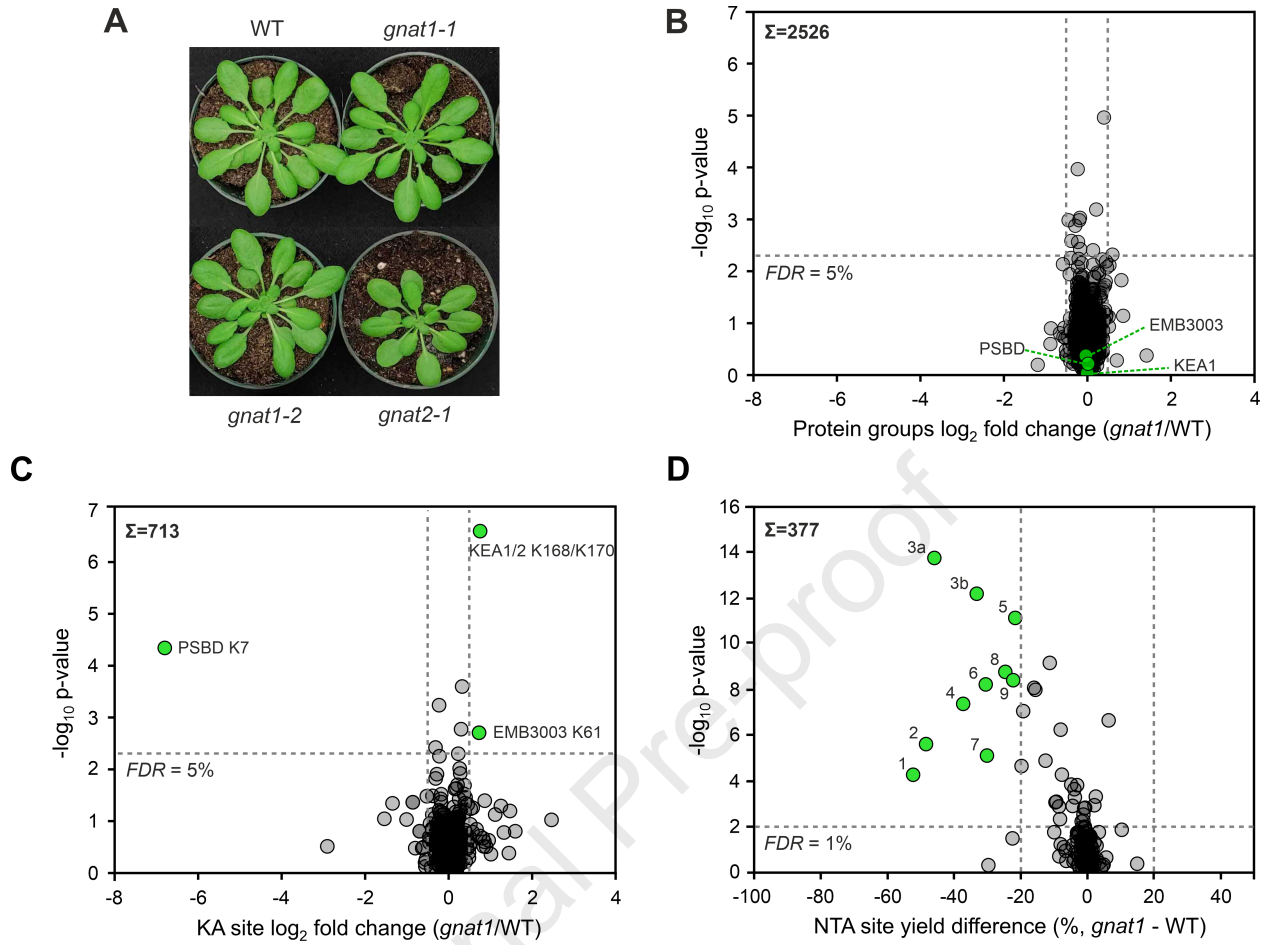


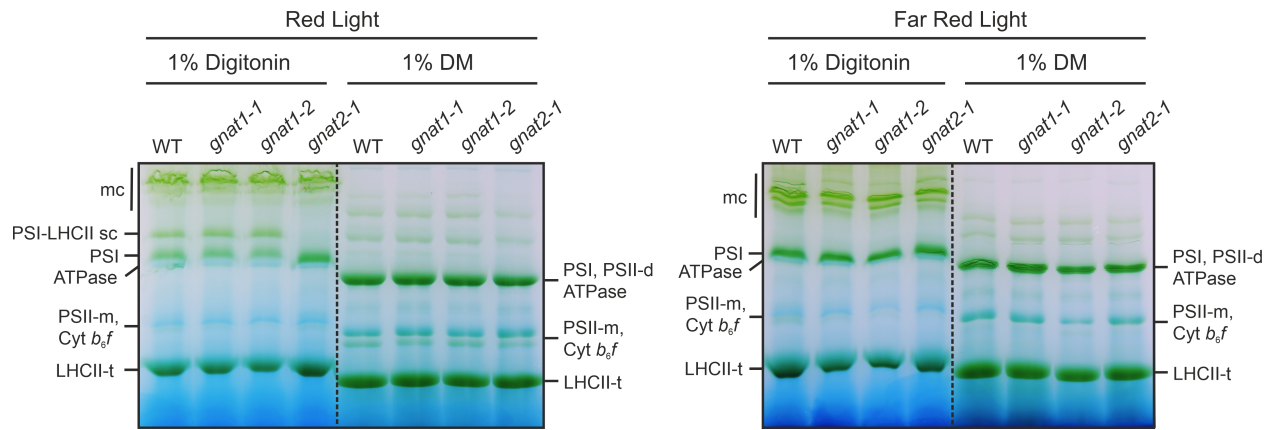
B

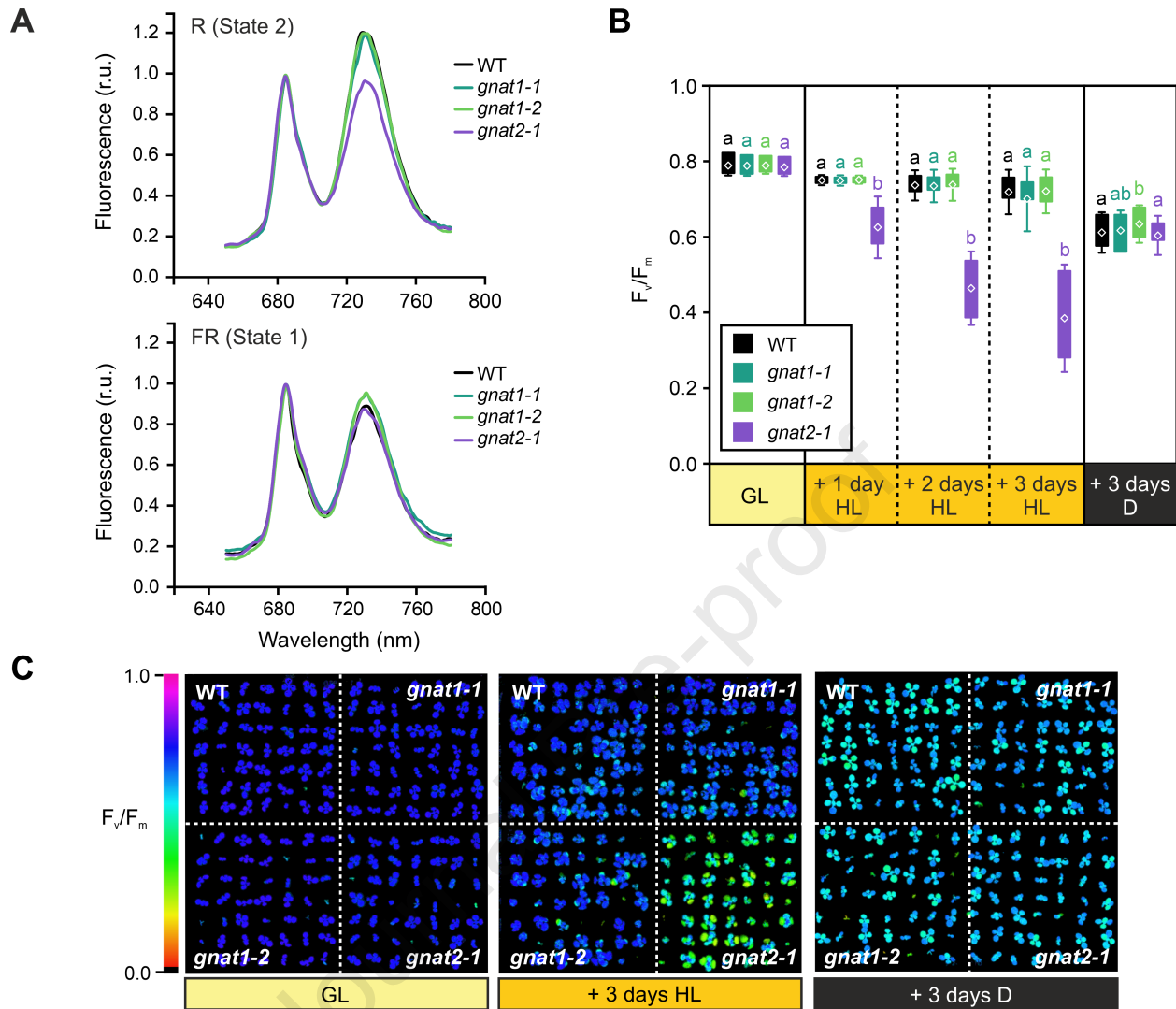


C

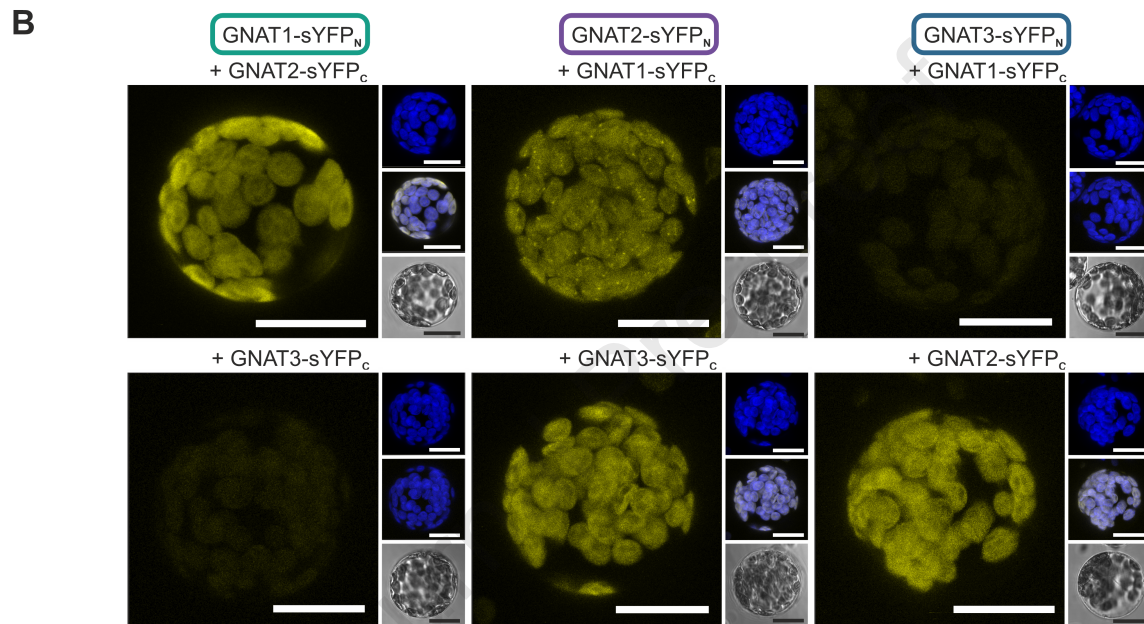
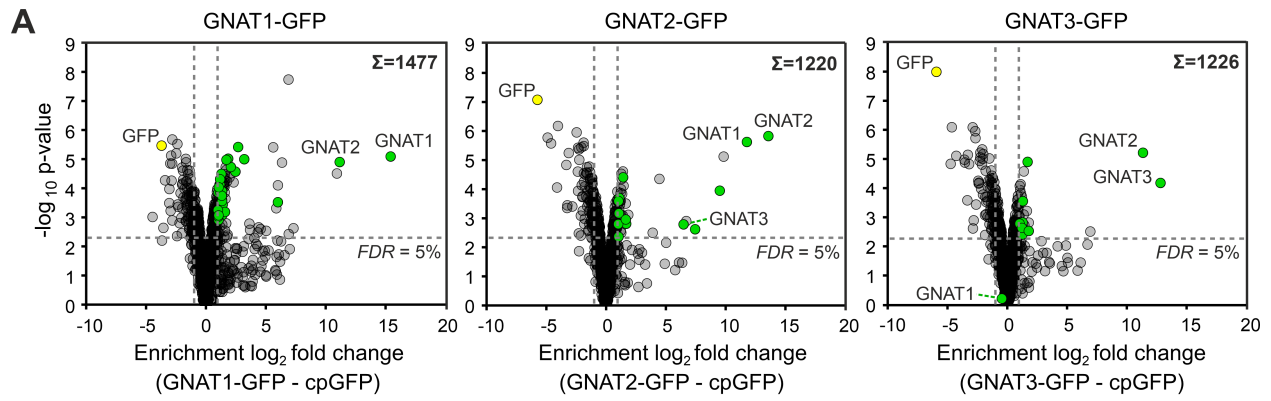




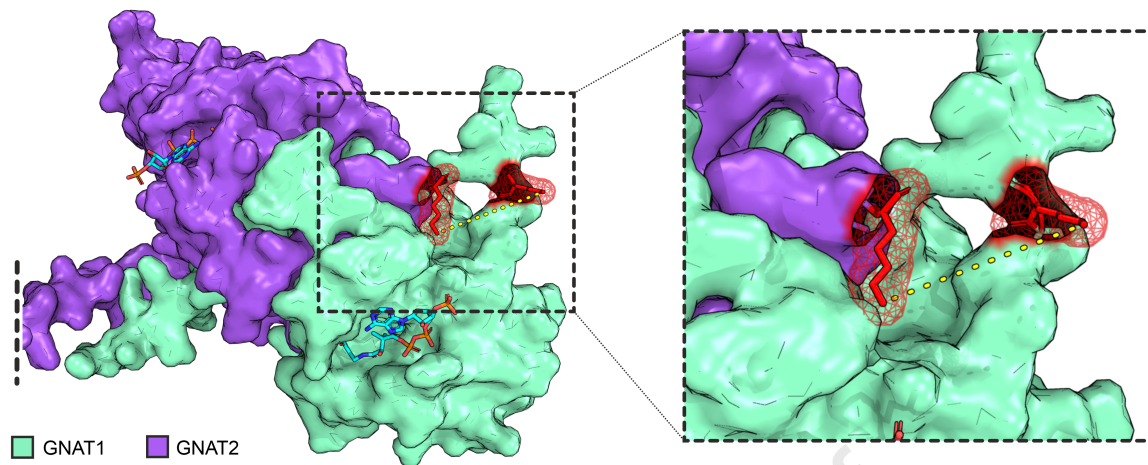




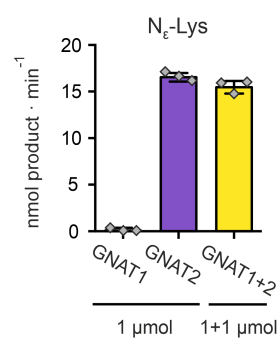




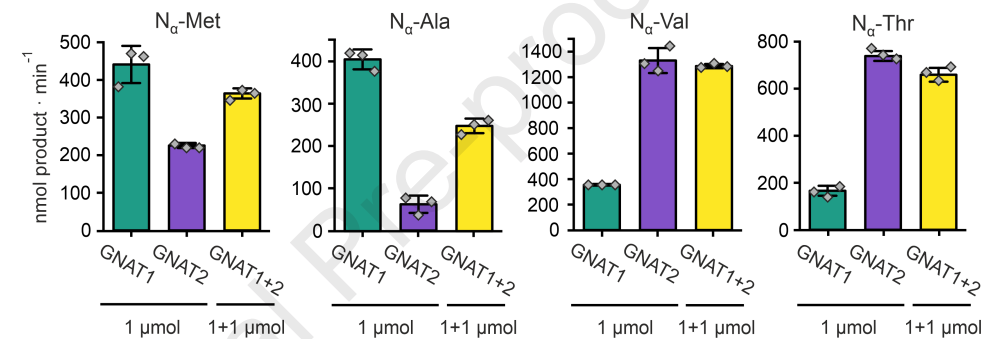
A



B



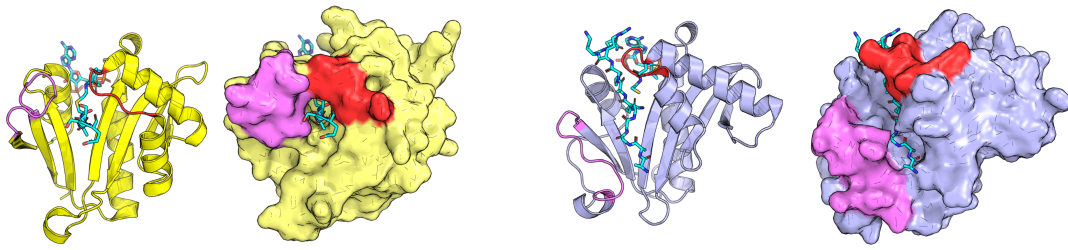
C



**A**

NAT (yNAA10)

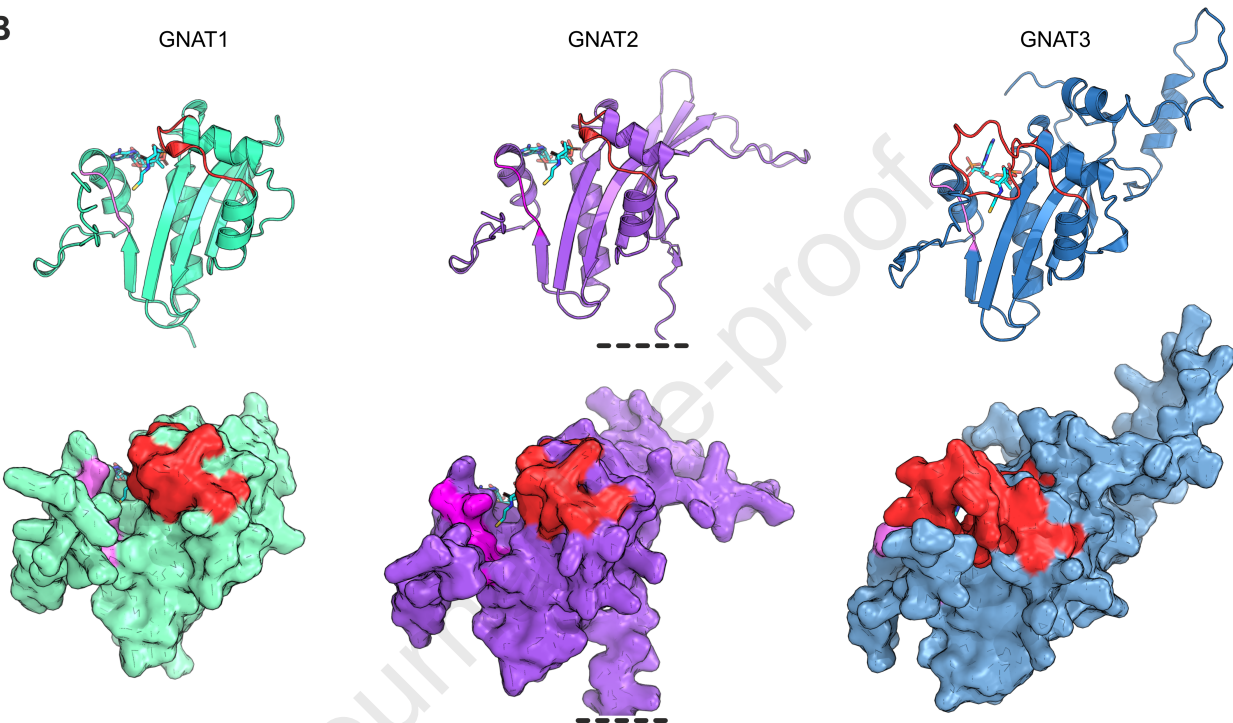
KAT (tGCN5)

**B**

GNAT1

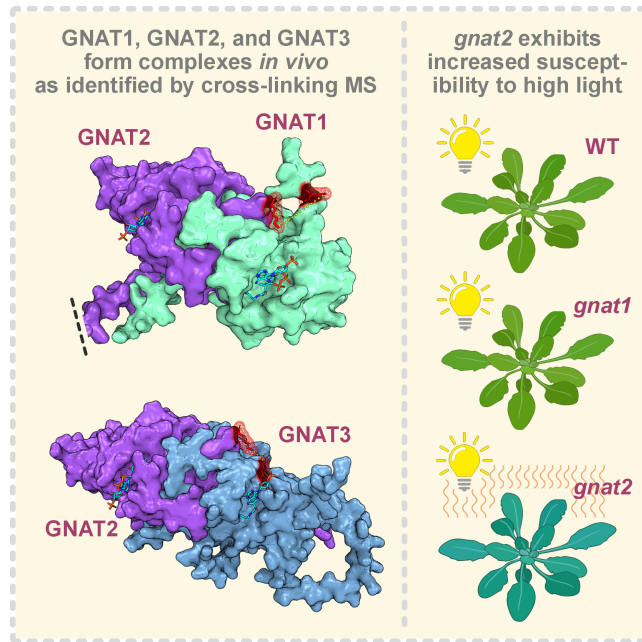
GNAT2

GNAT3



## Highlights

- GNAT1 is a plastidial acetyltransferase with dual activity in Arabidopsis.
- The N-terminal substrates of GNAT1 form a subset of the substrate set of GNAT2.
- Predicted structure models of GNAT1 and GNAT2 show a high architectural similarity.
- Absence of GNAT1 neither impacts state transitions nor acclimation to high light.
- Stable interactions were observed between GNAT1 & GNAT2 and GNAT2 & GNAT3.



**In Brief**

Protein acetylation is a common co- and post-translational modification in chloroplasts. Here, we characterize GNAT1, an acetyltransferase that adds acetyl groups to N-termini and lysine side-chains. GNAT1 is closely related to GNAT2, which regulates state transitions of the photosynthetic antenna proteins. Our findings show shared N-terminal substrate sites between GNAT1 and GNAT2 *in vivo*. Both enzymes form a complex, but this interaction is not essential for state transitions.

Journal Pre-proof

**Declaration of interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for *[Journal name]* and was not involved in the editorial review or the decision to publish this article.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof