Published online 19 June 2023

Nucleic Acids Research, 2023, Vol. 51, No. 14 ***7451–7464***
https://doi.org/10.1093/nar/gkad529

Pyrimidine catabolism is required to prevent the accumulation of 5-methyluridine in RNA

Shangyu Gao[1,†], Yu Sun[1,†], Xiaoguang Chen[2], Changhua Zhu[1], Xiaoye Liu[3], Wenlei Wang[1], Lijun Gan[1], Yanwu Lu[1], Frank Schaarschmidt[4], Marco Herde[2], Claus-Peter Witte[2,*] and Mingjia Chen[1,*]

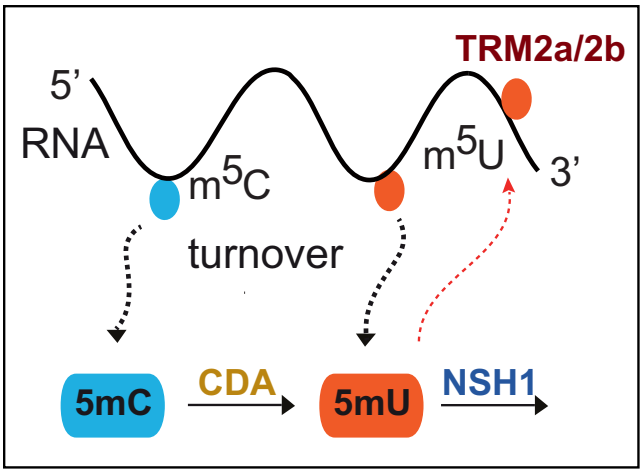
[1]College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, China, [2]Department of Molecular Nutrition and Biochemistry of Plants, Institute of Plant Nutrition, Leibniz University Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany, [3]Department of Criminal Science and Technology, Nanjing Forest Police College, Nanjing 210023, China and [4]Department of Biostatistics, Institute of Cell Biology and Biophysics, Leibniz University Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany

Received September 03, 2022; Revised May 31, 2023; Editorial Decision June 01, 2023; Accepted June 08, 2023

ABSTRACT

5-Methylated cytosine is a frequent modification in eukaryotic RNA and DNA influencing mRNA stability and gene expression. Here we show that free 5-methylcytidine (5mC) and 5-methyl-2′-deoxycytidine are generated from nucleic acid turnover in *Arabidopsis thaliana*, and elucidate how these cytidines are degraded, which is unclear in eukaryotes. First CYTIDINE DEAMINASE produces 5-methyluridine (5mU) and thymidine which are subsequently hydrolyzed by NUCLEOSIDE HYDROLASE 1 (NSH1) to thymine and ribose or deoxyribose. Interestingly, far more thymine is generated from RNA than from DNA turnover, and most 5mU is directly released from RNA without a 5mC intermediate, since 5-methylated uridine (m^5U) is an abundant RNA modification (m^5U/U ∼1%) in Arabidopsis. We show that m^5U is introduced mainly by tRNA-SPECIFIC METHYLTRANSFERASE 2A and 2B. Genetic disruption of 5mU degradation in the *NSH1* mutant causes m^5U to occur in mRNA and results in reduced seedling growth, which is aggravated by external 5mU supplementation, also leading to more m^5U in all RNA species. Given the similarities between pyrimidine catabolism in plants, mammals and other eukaryotes, we hypothesize that the removal of 5mU is an important function of pyrimidine degradation in many organisms, which in plants serves to protect RNA from stochastic m^5U modification.

GRAPHICAL ABSTRACT



INTRODUCTION

Of the more than 100 chemical modifications in various eukaryotic RNAs, the greatest diversity is found in tRNAs. In absolute quantity, most modifications are present in rRNAs, which are the most abundant RNA species. These modifications promote the stability and translation functions of these RNAs (1,2). Recent technical advances have shown that mRNA also carries modifications such as N^6-methyladenosine (m^6A) (3,4), N^1-methyladenosine (5), 5-methylcytidine (m^5C) (6), N^4-acetylcytidine (7) and pseudouridine (Ψ) (8). Specific enzymes, named 'writer' and 'eraser' proteins, are required for the introduction and removal of these modifications at the mRNA level, respectively. Such mRNA modifications are decoded by different 'reader' proteins and influence the stability and translation

*To whom correspondence should be addressed. Tel: +86 25 84396069; Fax: +86 25 84395869; Email: mjchen@njau.edu.cn
Correspondence may also be addressed to Claus-Peter Witte. Tel: +49 511 7624578; Fax: +49 511 7623611; Email: cpwitte@pflern.uni-hannover.de
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

efficiency of their targeted transcripts (9,10). Upon RNA turnover, modified nucleotides such as N^6-methyladenosine monophosphate (N^6-mAMP), pseudouridine monophosphate (ΨMP) and 5-methylcytdine monophosphate, are released together with other canonical nucleotides. A conserved N^6-mAMP deaminase, recently characterized in Arabidopsis (*Arabidopsis thaliana*) and human (*Homo sapiens*) cells, catalyzes the hydrolysis of N^6-mAMP to inosine monophosphate (11), an intermediate of general purine metabolism (12). The catabolism of Ψ, the most abundant modified nucleoside in non-coding RNA (ncRNA), has also been described in Arabidopsis. Free Ψ must be transported into peroxisomes, where it is phosphorylated by a pseudouridine kinase to ΨMP, which is hydrolyzed by a peroxisomal ΨMP glycosylase into uracil and ribose-5-phosphate (13,14), both intermediates of canonical pyrimidine nucleoside catabolism. Mammals lack these enzymes and excrete Ψ with the urine (15), whereas other animals, such as birds, fish and insects, have this catabolic pathway (13). Thus, in plants and many other eukaryotes, specialized enzymes are required for the conversion of the modified nucleotides N^6-mAMP and ΨMP into intermediates of general nucleotide metabolism, whereas in mammals some modified nucleosides can also be disposed of by secretion.

Unlike m^6A and Ψ, m^5C is widely distributed not only in ncRNAs but also in mRNAs of eukaryotic organisms (16–18). This C methylation is introduced by NOP2/SUN RNA METHYLTRANSFERASE FAMILY MEMBER 2 (NSUN2) in mammals (16) and by tRNA-SPECIFIC METHYLTRANSFERASE 4B (TRM4B) in Arabidopsis (6). It was observed that m^5C is enriched in 3′-untranslated regions and regions immediately downstream of translation initiation sites, where it is recognized by the reader protein ALY/REF EXPORT FACTOR (ALYREF), which mediates nuclear export of m^5C-methylated transcripts (16,19). In Arabidopsis, mutants of *TRM4B* lacking RNA m^5C methyltransferase activity undergo fewer cell divisions in the root apical meristem, resulting in shorter primary roots and increased sensitivity to oxidative stress (17). 5-Methyl-2′-deoxycytidine (5mdC) is a major epigenetic mark of DNA (^{5}mC). Upon RNA turnover, 5-methylcytidine (5mC) is released while DNA turnover or repair lead to the release of 5mdC.

In plant pyrimidine metabolism, CYTIDINE DEAMINASE (CDA) and NUCLEOSIDE HYDROLASE 1 (NSH1) are two important enzymes involved in nucleoside and nucleotide homeostasis (12,20). Cytidine degradation begins with its deamination by CDA to uridine (21), which is then hydrolyzed by NSH1 to uracil and ribose (22,23) (Supplementary Figure S1). NSH1 also hydrolyzes C5-methylated uridine (5mU) *in vitro* (24), but it is unknown whether 5mC is a substrate of CDA. If that were the case, the NSH1 substrate 5mU could originate from 5mC deamination. Alternatively or additionally, this methylated uridine may also be directly derived from m^5U-modified RNA species. Although m^5U has been detected in plant tRNAs (25), little is generally known about m^5U in plant RNAs. In mammals and yeast (*Saccharomyces cerevisiae*), m^5U has been found in rRNA and tRNA (26) as well as in mRNA (27,28), but its metabolic fate upon degradation of these RNAs has not yet been investigated. There are many similarities between plant and mammalian pyrimidine nucleoside catabolism (Supplementary Figure S1). Mammals possess a CDA that can deaminate C, dC and 5mC *in vitro* (29). The degradation of uridine (U), deoxyuridine (dU) and thymidine (dT) is catalyzed by phosphorylases (30–32) and not by hydrolases as in plants, but in both organisms the bases uracil or thymine (T) are released by these enzymes which are further metabolized in similar ways (33,34) (Supplementary Figure S1).

In this study, we investigated the origin of methylated (deoxy-)cytidines and (deoxy-)uridines in Arabidopsis and their metabolic fate, showing that CDA and NSH1 are required for the catabolism of most but not all of these compounds. Moreover, we present experimental evidence that plant RNAs contain m^5U and that TRM2A and TRM2B are the main writer enzymes for m^5U. We also investigated whether NSH1 and CDA are required to protect RNA from random methylation events and can show that NSH1 is indeed involved. Phenotypic consequences of defective pyrimidine nucleoside degradation were also investigated.

MATERIALS AND METHODS

Plant material and cultivation

Arabidopsis (*Arabidopsis thaliana*) ecotype Col-0 was chosen as the wild type. T-DNA insertion mutants of Arabidopsis from the SALK collection (35), the SAIL collection (36) and the GABI-Kat collection (37) were obtained from the Nottingham Arabidopsis Stock Centre. Complementation lines for *cda-1* and *nsh1* were described previously (21,38). Double mutants were obtained by crossing *cda-1* and *trm4b-1*, *cda-1* and *cmt3-11*, *cda-1* and *ros1-4*, *pyd1* and *cda-1*, *pyd1* and *nsh1*, and *trm2a* and *nsh1*; the pollen donor is named last in each case. The T-DNA insertions were confirmed using the primer pairs 448/N0262, N261/N61, N61/2616, N61/1426, 1197/Cp99, N61/Cp108, N61/Cp111, N61/Cp242 and N61/Cp237 for *cda-1*, *cda-2*, *nsh1*, *pyd1*, *trm4b-1*, *cmt3-11*, *ros1-4*, *trm2a* and *trm2b* lines, respectively, and the primer pairs N0261/N0262, N0261/N0262, 1905/2616, 1236/1426, Cp99/Cp100, Cp108/Cp109, Cp110/Cp111, Cp241/Cp242 and Cp237/Cp238 for the wild-type alleles of *cda-1*, *cda-2*, *nsh1*, *pyd1*, *trm4b-1*, *cmt3-11*, *ros1-4*, *trm2a* and *trm2b* lines, respectively (Supplementary Table S1).

Arabidopsis and *Nicotiana benthamiana* plants were cultivated in a controlled growth chamber (16 h light of 75 μmol/m/s at 22°C and 8 h dark at 20°C, 75% relative humidity). For metabolic profiling and RNA extraction, plants were grown on full nutrient soil or on Petri dishes containing half-strength Murashige and Skoog (MS) agar. For the 5mU and U treatments, the plants were grown on half-strength MS medium supplied with 500 ng/ml 5mU or 5000 ng/ml U, respectively. The leaf area of seedlings was quantified from photographs employing the ImageJ (v 1.8.0) software.

RNA extraction, reverse transcription and mRNA preparation

RNA from plants was isolated using Trizol reagent (Monad) and treated with DNase (4× gDNA wiper Mix;

Vazyme Biotech Co., Ltd), and cDNA was prepared using HiScript II Q Select RT Super Mix (Vazyme Biotech Co., Ltd). Oligo (dT) was used for reverse transcription. To measure gene-specific mRNA levels in the mutants, cDNAs were prepared from seedlings of each genotype. The polymerase chain reaction (PCR) used primer pairs Cp345/Cp346 and Cp237/Cp238, giving rise to a product of 2496 bp and 1743 bp from the wild-type alleles of *trm2a* and *trm2b*, respectively. The primers flanked the T-DNA insertions in each case. The sequences of primers used in analyses are listed in Supplementary Table S1. Non-polyadenylated RNA and mRNA were isolated from total RNA for digestion employing the PolyATract mRNA Isolation Systems III (Promega). For mRNA preparation, two rounds of purification were performed. The RNA was quantified by a NanoDrop photometer (Thermo Scientific), and the purity and the integrity were assessed by applying 100 ng of RNA per sample on a bioanalyzer (Agilent).

Protein purification and determination of kinetic constants

Recombinant Arabidopsis CDA and NSH1 were affinity purified after transient expression in *N. benthamiana* as described before (21,38). Purified protein was quantified using the Bradford reagent from Tiangen with bovine serum albumin (BSA) as a standard.

The enzymatic activity of CDA was assessed as described (21), with minor modifications. For individual reactions, recombinant CDA solutions were adjusted to 5 μg/ml using elution buffer of the purification procedure. A 150 μl aliquot of the substrate solution was incubated at 30°C for 5 min. The reaction was started by adding 50 μl of enzyme solution. In a time course, 40 μl aliquots were added to 160 μl of water, followed by the addition of 50 μl of phenol nitroprusside reagent and 100 μl of hypochlorite reagent for colorimetric ammonia quantification (39). The absorbance was determined by photometric measurement at 636 nm. Ammonium standard curves were generated by adding elution buffer instead of enzyme solution to the reaction mix and placing 40 μl aliquots in 160 μl of NH_4Cl solutions of different concentrations prior to detection. The kinetic constants were determined for C concentrations of 0.1, 0.25, 0.5, 1, 3, 5 and 10 mM, for dC concentrations of 0.05, 0.1, 0.25, 0.5, 1, 2 and 5 mM, for 5mC concentrations of 1, 3, 5, 10, 25, 50, 100 and 200 mM, for 5mdC concentrations of 1, 5, 10, 25, 50 and 100 mM, and for 2′-*O*-methylcytidine (2OmC) concentrations of 1, 5, 10, 20, 50, 100 and 200 mM. Kinetic curves were recorded in four independent repeats, and kinetic constants were determined by fitting the data to the Michaelis–Menten equation using the GraphPad Prism software.

Recombinant NSH1 solutions were adjusted to 57.5 μg/ml using 100 mM HEPES buffer (pH 8). For individual reactions, 200 μl of 20 mM uridine or 2′-*O*-methyluridine (2OmU) substrate solution was incubated at 30°C for 5 min. The reaction was started by adding 40 μl of NSH1 enzyme solution. In a time course, 60 μl aliquots were taken and the reaction was terminated by heating the mixture to 98°C for 10 min. The activity was determined by quantifying uracil production over time via high performance liquid chromatograpy (HPLC; Agilent Infinity 1260).

Metabolite analyses

Nucleotide metabolite analyses were performed using an Agilent 1290 HPLC system coupled to an Agilent 6460 triple-quadrupole mass spectrometer, and other metabolites were analyzed by an Agilent 1290 HPLC system coupled to a Sciex 6500 Qtrap mass spectrometer. Sample extraction was performed as described (20,38), with minor modifications. The plant material was frozen in liquid nitrogen and ground in 2 ml centrifuge vials with steel beads in a tissue grinder (Tissuelyser-48L, Shanghai Jingxin) at a frequency of 60 Hz for 5 min. For nucleoside and nucleobase measurement, 30 mg of homogenized plant material was extracted with 200 μl of 0.1% formic acid and then immediately incubated at 95°C for 10 min. After centrifugation at 20 000 *g* for 20 min, the suspension was passed through a SyringeFilter (Nylon/PES, 0.2 μm; Tianjin Branch billion Lung Experimental Equipment Co., Ltd) to remove any particles. For nucleotide analyses, plant material was extracted and samples prepared as previously described (20).

For the analysis of nucleosides and nucleobases, 2 μl samples were separated on a WATERS T3 column (2.1 × 100 mm; particle size 1.8 μm) at a flow rate of 0.3 ml/min. Solvent A was 0.1% formic acid and Solvent B was acetonitrile. The gradient was 0 min, 5% B; 0.5 min, 5% B; 3 min, 90% B; 4 min, 90% B; 4.1 min, 5% B; and 7 min, 5% B. The following mass transitions were monitored: m/z 244.3 to 111.9 (C), m/z 258.4 to 125.9 (5mC), m/z 258.2 to 111.9 (2OmC), m/z 228.3 to 111.9 (dC), m/z 242.3 to 125.9 (5mdC), m/z 245.3 to 133 (U), m/z 259.2 to 126.9 (5mU), m/z 259.1 to 127.1 (3-methyluridine; 3mU), m/z 259.3 to 112.9 (2OmU), m/z 243 to 127 (dT), m/z 111.1 to 111.1 (U) and m/z 127.9 to 110 (T). All compounds were quantified using external standards added to wild-type extracts from the corresponding growth conditions as matrix. C: 0, 100, 500, 1000, 5000, 20 000 and 50 000 ng/ml; 5mC: 0, 10, 50, 100, 500 and 2000 ng/ml; 2OmC: 0, 20, 100, 200, 1000, 4000 and 10 000 ng/ml; dC: 0, 5, 25, 50, 250, 1000, 2500 and 5000 ng/ml; 5mdC: 0, 0.125, 0.5, 1.25, 5 and 25 ng/ml; 5mU: 0, 10, 50, 100, 500 and 2000 ng/ml; 2OmU: 0, 20, 100, 200, 1000, 4000 and 10 000 ng/ml; dT: 0, 1.25, 5, 25, 50 and 100 ng/ml; uracil: 0, 2.5, 12.5, 25, 125, 250 and 500 ng/ml; and thymine: 0, 2.5, 12.5, 25, 125, 250 and 500 ng/ml were used as external standards.

For nucleotide analyses, 10 μl samples were separated on a Hypercarb column (50 × 4.6 mm, particle size 5 μm, Thermo Scientific) at a flow rate of 0.6 ml/min. Solvent A was 5mM NH_4Ac (pH 9.5) and Solvent B was acetonitrile. The gradient was 0 min, 4% B; 10 min, 30% B; 10.1 min, 100% B; 11.5 min, 100% B; 11.6 min, 4% B; and 20 min, 4% B. The following mass transitions were monitored: m/z 499 to 97 (5-methyl-UTP; 5mUTP) and m/z 485 to 97 (UTP). UTP was quantified against its internal standard, while 5mUTP was quantified using external standards (25, 50, 125, 250, 500 and 2500 ng/ml) added to the wild-type extracts from the corresponding growth condition as matrix.

Quantitative analyses of the nucleic acid modification by LC-MS/MS

RNA and DNA samples were fully digested into single nucleosides as described previously (11) with minor modifications. A 800 ng aliquot of total RNA, mRNA, non-polyadenylated RNA or DNA was individually digested by benzonase (0.4 U; Sigma-Aldrich) and phosphodiesterase I (0.004 U; Sigma-Aldrich), and then dephosphorylated by shrimp alkaline phosphatase (0.04 U; NEB) in 50 μl of buffer containing 10 mM Tris–HCl, pH 7.9, 1 mM $MgCl_2$ and 0.1 mg/ml BSA. After incubation at 37°C for 10 h, the samples were filtered in ultrafiltration tubes (3 kDa cut-off; Pall), and 2 μl aliquots were analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS). The nucleosides were separated and identified as described above. Standard solutions of C: 1, 5, 25 50, 100, 200, 400, 2000 and 10 000 ng/ml; U: 1, 5,25 50, 100, 200, 400, 2000 and 10 000 ng/ml; 5mC: 0.1, 0.5, 0.5, 5, 10, 20, 40, 200 and 1000 ng/ml; 5mU: 0.1, 0.5, 2.5, 5 and 10 ng/ml; 2OmC: 0.1, 0.5, 2.5, 5 and 10 ng/ml; 2′-*O*-methyl-deoxycytidine (2OmdC): 0.5, 2.5, 12.5, 62.5, 125, 250, 500, 1000 and 2000 ng/ml; or 5mdC: 0.5, 2.5, 12.5, 25, 50, 100, 200 and 400 ng/ml were used for quantification. The ratio of 5mC to C, 2OmC to C, 5mdC to dC, and 5mU to U were calculated based on the calibrated concentrations.

The ability of eukaryotic RNA polymerase II to use 5mUTP as substrate was assessed using the HeLa Scribe nuclear extract *in vitro* transcription system (Promega) as described previously (11). 5mUTP and UTP mixed in different ratios (0:100, 1:99, 2:98, 5:95, 10:90 and 20:80) were applied. RNA was isolated after a reaction time of 1 h and then subjected to digestion. The 5mU to U ratio was determined by LC-MS/MS as described above.

Quantification and statistical analysis

Data for the leaf area of seedlings from different genotypes grown under various conditions were analyzed and plotted with R (version 4.2.2) in combination with RStudio (version 2022.07.02). The final figure was created with GraphPad (version 9.0). Linear modeling of log-transformed data (package 'nlme' version 3.1-162) and subsequent testing for normality by visual inspection of normal plots with simulation envelopes (package 'hnp' version 1.2-6) showed normality for the log-transformed data generated for Figure 7 which were chosen for further processing. We tested whether the analysis would benefit from an extension with generalized least square models that account for heterogeneous variances. However, by calculating the lowest Akaike information criterion (AICc; package 'MuMIn' version 1.47.5), we found the lowest value for the model assuming equal variances which was chosen for further calculations. Confidence intervals (at 95%) and significance tests were calculated on the model-based means with the package 'emmeans' (version 1.8.5) considering the ratio of genotypes and mock-treated versus untreated seedlings. For other figures, if applicable, analysis of variance (ANOVA) followed by Tukey's honestly significant difference test was performed in R (version 4.2.2) for statistical evaluation. Different letters represent differences at the $P < 0.05$ significance level.

RESULTS

Exploration of the catabolism of methylated cytidines

To investigate whether CDA, encoded by At2g19570, can use 5mC and other modified cytidines as substrate, we affinity-purified recombinant CDA fused to a C-terminal Strep tag (CDA-Strep) transiently produced in *N. benthamiana* leaves (Supplementary Figure S2A) (21). The substrate range of the purified CDA was assessed using C and dC as well as eight modified (deoxy)cytidine nucleosides naturally found in plant RNA or DNA. Each nucleoside was provided at a concentration of 2 mM (Figure 1A). Five nucleosides were good substrates for CDA: C, 5mC, 2′-*O*-methylcytidine (2OmC), dC and 5mdC. For human CDA also C, 5mC and dC were good substrates *in vitro* (29), while other possible *in vivo* substrates have not yet been tested. For Arabidopsis CDA, it has already been shown that dC is a substrate *in vitro* (21), and it has also been noted that a mutant of *CDA* accumulates dC, suggesting that it is a substrate of CDA *in vivo* (20). CDA was inactive with the possible *in vivo* substrates N^3-methylcytidine, N^4-acetylcytidine, 5-hydroxymethylcytidine, 5-formylcytidine and 5-hydroxymethyl-deoxycytidine (Figure 1A). The catalytic efficiency (k_{cat}/K_M) of CDA was 35.9/mM/s for C and 64.8/mM/s for dC (Supplementary Figure S2B, C; Table 1), similar to previous results (21). For 5mC, 2OmC and 5mdC, the catalytic efficiency of CDA was 4.4, 4.3 and 13.8/mM/s, respectively, markedly lower than that for the two non-methylated substrates (Figure 1B, C, D; Table 1). This difference resulted mainly from higher K_M values for the methylated cytidines, indicating that CDA binds these substrates less efficiently than (d)C but catalyzes them well once they are bound.

CDA and NSH1 are required for methylated pyrimidine degradation *in vivo*

To investigate the *in vivo* activity of CDA with these methylated cytidine nucleosides, we used two independent homozygous T-DNA insertion mutants, *cda-1* and *cda-2*, and a complementation line expressing C-terminal hemagglutinin (HA) and Strep-tagged CDA in the *cda-1* background (called AtCDA + *cda-1*) reported previously (21). We quantified 5mC, 2OmC, 5mdC, C and dC contents by LC-MS/MS in extracts from 20-day-old seedlings of these lines and the wild type (Col-0). Contents of 5mC, 2OmC and 5mdC were ∼20, 32 and 10 times higher, respectively, in both *cda* lines compared with the wild type (Figure 2). Contents of cytidine and dC were ∼20 and 5.8 times higher, respectively, in *cda* than in the wild type (Supplementary Figure S3). The accumulation of these nucleosides was prevented by the *AtCDA* transgene in the *cda* background (Figure 2; Supplementary Figure S3). These data demonstrate that CDA is required *in vivo* for the catabolism of 5mC, 2OmC and 5mdC. Although mammalian CDA is very probably also involved in the degradation of methylated cytidines *in vivo*, its role and the physiological consequences of altered CDA activity in this context have so far not been investigated (40–42).

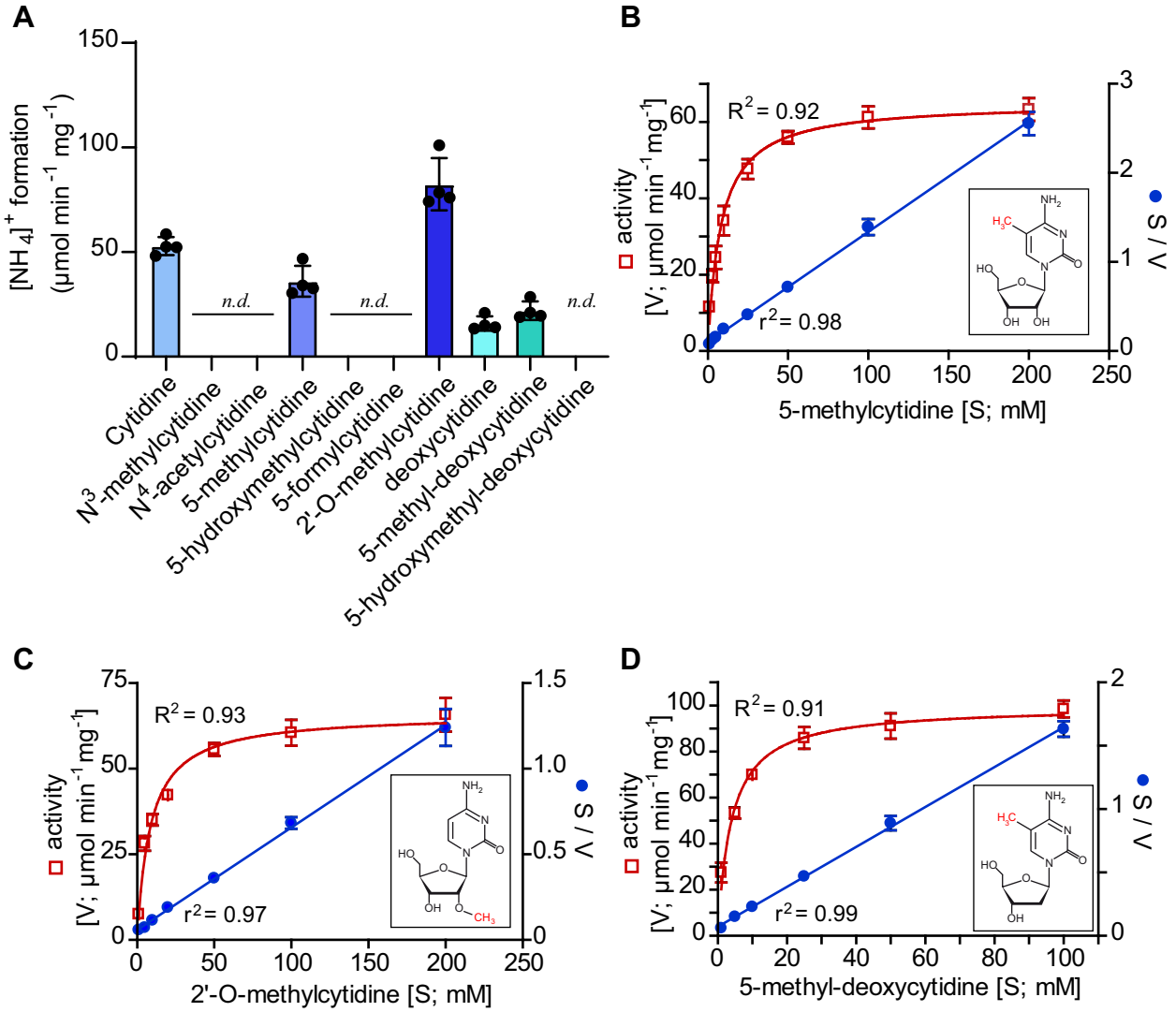


Figure 1. Biochemical analyses of CDA. (**A**) Enzymatic activity of CDA with different nucleosides as substrates, each at a concentration of 2 mM. Error bars are the SD ($n = 4$ independent reactions). *n.d.*, not detectable. (**B**) Determination of kinetic constants for CDA with 5-methylcytidine (5mC). S, substrate concentration; V, enzymatic velocity. The kinetic data were fitted with the Michaelis–Menten equation (red curve), or plotted as S/V over S (Hanes plot) and fitted by linear regression (blue line). Error bars are the SD ($n = 4$ independent reactions). (**C**) Same as in (B) but with 2′-*O*-methylcytidine (2OmC) as substrate. (**D**) Same as in (B) but with 5-methyl-2′-deoxycytidine (5mdC) as substrate.

Table 1. Kinetic parameters of recombinant Arabidopsis CDA for different substrates

Substrate	AtCDA		
	K_M (mM)	k_{cat} (/s)	k_{cat}/K_M (/mM/s)
Cytidine	0.82 ± 0.15	29.43 ± 1.65	35.9
5mC	8.05 ± 0.96	35.36 ± 1.04	4.4
2OmC	8.34 ± 1.12	35.46 ± 1.12	4.3
dC	0.28 ± 0.05	18.15 ± 0.66	64.8
5mdC	3.92 ± 0.61	54.1 ± 1.68	13.8

The deamination of cytidine by CDA produces uridine, which is an *in vivo* substrate of NSH1 (22). *In vitro*, NSH1 also hydrolyzes 5mU (24), prompting us to hypothesize that NSH1 might be involved in the catabolism of the other deamination products of CDA, i.e. 5mU, 2′-*O*-methyluridine (2OmU) and dT. To test this idea, we measured contents of these three compounds in 20-day-old seedlings of Col-0, a null mutant in *NSH1* (*nsh1*) and a complementation line (AtNSH1 + *nsh1*), which had been characterized previously (38). Contents of 5mU and dT were ∼275 and 3.3 times higher, respectively, in the mutant seedlings than in the wild type, and the *NSH1* transgene complemented this molecular phenotype (Figure 3A, B). An accumulation of dT in *nsh1* had been noted before (20). In contrast, 2OmU did not accumulate in *nsh1* seedlings and similar amounts of this compound were measured in all lines (Figure 3C). Consistent with this result, NSH1 lacked activity with 2OmU *in vitro* when this compound was incubated with the affinity-purified enzyme (Supplementary Figure S4). These data show that NSH1 is required for the

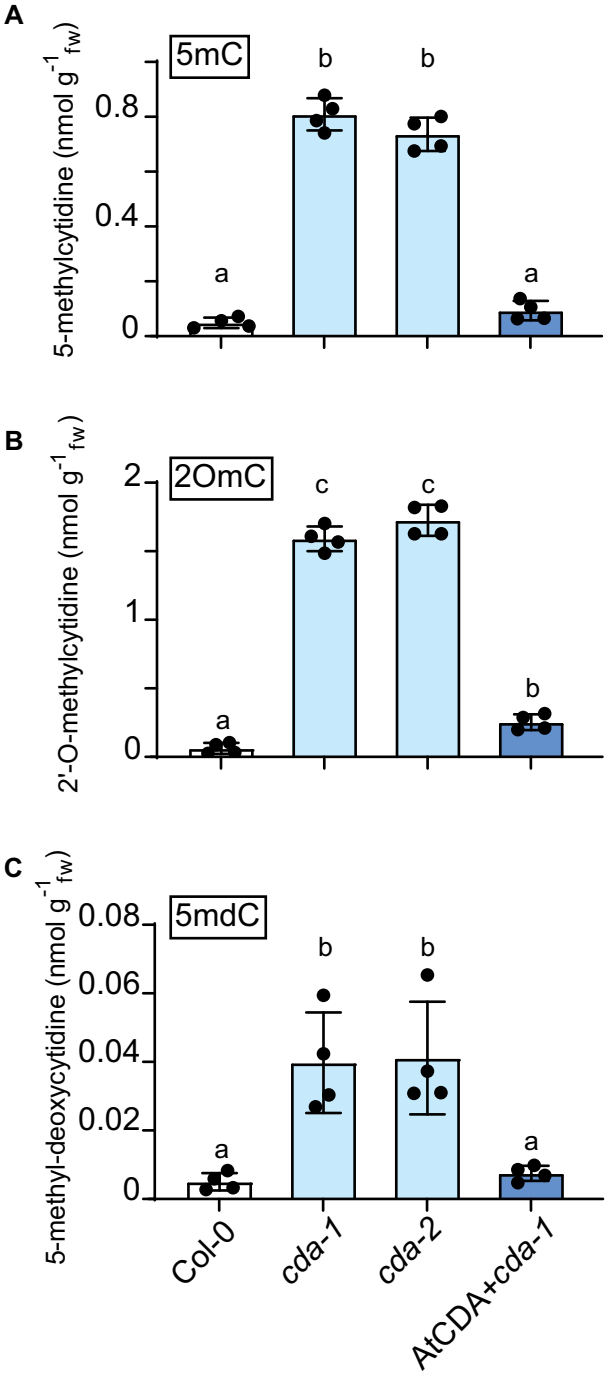


Figure 2. Metabolite analyses of seedlings varying in *CDA* expression. (**A**) 5mC contents in 20-day-old seedlings of the wild type (Col-0, white), the two *cda* lines (light blue) and the complementation line (dark blue). Error bars are the SD ($n = 4$ biological replicates). Different letters label data that are significantly different at $P < 0.05$. fw, fresh weight. (**B**) Same as in (A) but showing 2OmC contents. (**C**) Same as in (A) but showing 5mdC contents.

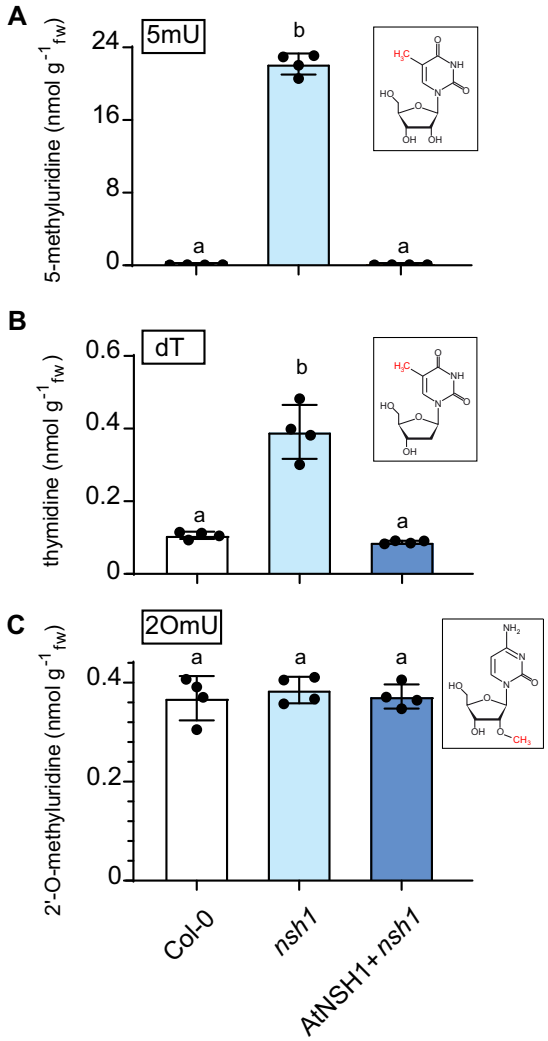


Figure 3. Metabolite analyses of seedlings varying in *NSH1* expression. (**A**) 5-Methyluridine (5mU) contents in 20-day-old seedlings of the wild type (Col-0, white), the *nsh1* line (light blue) and the complementation line (dark blue). Error bars are the SD ($n = 4$ biological replicates). Different letters label data that are significantly different at $P < 0.05$. fw, fresh weight. (**B**) Same as in (A) but showing thymidine (dT) contents. (**C**) Same as in (A) but showing 2′-*O*-methyluridine (2OmU) contents.

hydrolysis of 5mU and dT *in vivo*, whereas 2OmU is not an NSH1 substrate and will probably be degraded by another so far unknown process. Although mammalian RNA also contains m^5U (43), its metabolic fate after RNA degradation has not yet been investigated.

Pyrimidine ring catabolism degrades thymine which is mainly derived from RNA

NSH1 hydrolyzes 5mU and dT in the cytosol producing thymine (5-methyluracil) and either ribose or deoxyribose. It has been shown that the first enzyme of pyrimidine ring catabolism DIHYDROURACIL DEHYDROGENASE (DPYD), also called PYRIMIDINE 1 (PYD1), is required for uracil and thymine degradation, because mutants lacking this enzyme were strongly compromised in producing CO_2 from exogenously supplied radiolabeled substrates (34,44). It was also shown in these studies that mutants of *PYD1* accumulate uracil *in vivo*, whereas the content of thymine was not reported, probably because thymine was thought to be derived from DNA turnover or repair and therefore quantitatively of minor importance. Our data show that there is almost two orders of

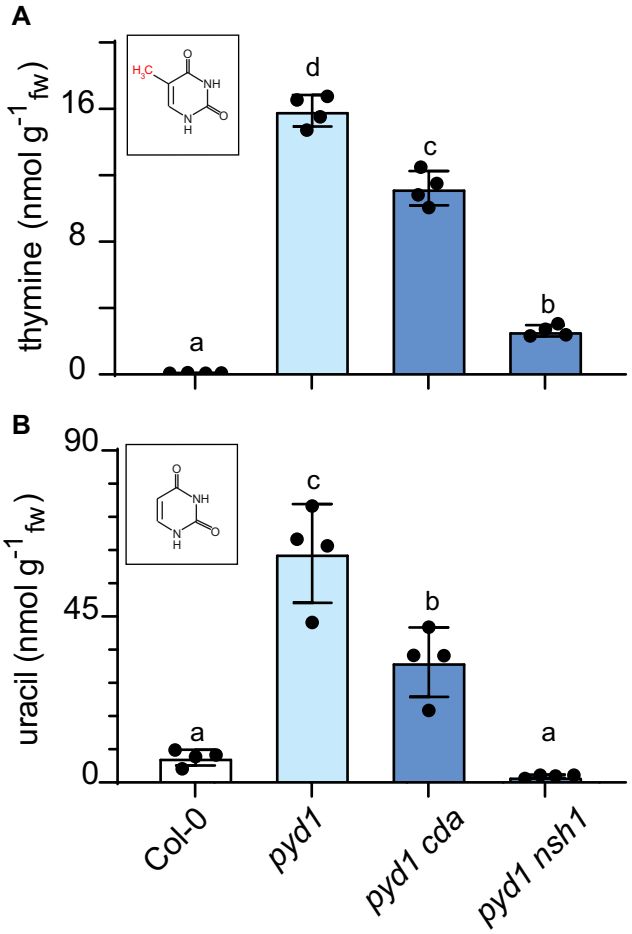


Figure 4. Metabolite analyses of *pyd1 cda* and *pyd1 nsh1* seedlings in comparison with *pyd1* seedlings and the wild type. (**A**) Thymine contents in 20-day-old seedlings of the wild type (Col-0, white), the *pyd1* line (light blue) and the *pyd1 cda-1* and *pyd1 cda-1* double mutant lines (dark blue). Error bars are the SD ($n = 4$ biological replicates). Different letters label data that are significantly different at $P < 0.05$. fw, fresh weight. (**B**) Same as in (A) but showing uracil contents.

magnitude more 5mU in *nsh1* seedlings than the corresponding deoxynucleoside dT, strongly suggesting that the thymine generated by NSH1 is derived mainly from RNA rather than from DNA turnover (Figure 3). Therefore, we hypothesized that a mutant of *PYD1* would accumulate substantial amounts of thymine, far more than would be expected from DNA turnover and repair. We used the *pyd1-1* allele (34) for this experiment and additionally crossed this mutant to mutants of *CDA* (*cda-1*) and *NSH1* (*nsh1-1*) (24). In 20-day-old seedlings, uracil but also thymine accumulated strongly in the *PYD1* mutant (Figure 4). There was about the same amount of thymine in *pyd1* seedlings (~16 nmol/g fw) as 5mU in the *nsh1* background (~22 nmol/g fw; Figure 3), supporting the idea that most thymine that enters catabolism is derived from RNA and not from DNA. The thymine accumulation was suppressed by ~30% and 85% in the double mutants *pyd1 cda* and *pyd1 nsh1*, respectively (Figure 4A). The weaker suppression in the *cda* background shows that much of the 5mU hydrolyzed by NSH1 does not stem from 5mC deamination but is probably directly derived from m^5U-containing RNA. So far, m^5U has only been detected in plant tRNAs (25), but the high accumulation of 5mU in *nsh1* indicates that other plant RNAs, especially the abundant rRNAs, also contain m^5U.

DNA turnover and repair contribute a comparatively small amount of thymine, as the *CDA* and *NSH1* mutants accumulate far fewer 5-methylated deoxynucleosides (5mdC and dT) than methylated nucleosides (5mC and 5mU, Figures 2 and 3). Nonetheless, in the *pyd1 nsh1* line, there is more thymine than in the wild type (Figure 4A), indicating that not all thymine is generated via nucleoside hydrolysis by NSH1. Some will be released by nucleobase excision repair directly from DNA without a nucleoside intermediate (45).

Origin of 5mC, 5mdC and 5mU

In plants, TRM4B, encoded by At2g22400, introduces a methyl group into cytosine residues in RNA molecules (6). A mutation in *TRM4B* (*trm4b-1*) (17) led to an ~2-fold decrease in m^5C abundance in total RNA relative to its abundance in the wild type (Figure 5A). To investigate whether the CDA substrate 5mC (Figures 1 and 2) stems from the degradation of m^5C-containing RNA, we crossed the *cda-1* mutant to the *trm4b-1* mutant and analyzed the resulting lines. In 20-day-old *cda trm4b* seedlings, ~41% less 5mC was detected compared with the *cda* single mutant (Figure 5B), demonstrating that a substantial amount of 5mC originates from m^5C RNA breakdown.

CHROMOMETHYLASE3 (CMT3; AT1G69770) and REPRESSOR OF SILENCING1 (ROS1; AT2G36490) are required for the addition and removal, respectively, of cytosine methylation in plant DNA (46). Mutations in *CMT3* (*cmt3-11*) (47) or *ROS1* (*ros1-4*) (48) resulted in lower or higher ^{5}mC abundance in DNA, respectively (Figure 5C). To test whether free 5mdC is released from the degradation of ^{5}mC-containing DNA, we crossed *cda-1* to *cmt3-11* and *ros1-4*. Compared with the wild type, which contains little free 5mdC, *cda* and *cmt cda* as well as *ros cda* seedlings all accumulated significantly more of this methylated deoxynucleoside *in vivo* (Figure 5D). However, compared with *cda*, levels were lower by ~26% in *cmt cda* and higher by ~19% in *ros cda* plants (Figure 5D), demonstrating that 5mdC accumulating in the *cda* background is derived from DNA degradation.

CDA can generate 5mU by deamination of 5mC. However, the concentration of 5mU in the *nsh1* line (Figure 3A) was 25 times higher than that of 5mC in the *cda* background (Figure 2A), probably because 5mU was also directly released from RNA degradation of m^5U-containing RNA. In plants, TRM2A and TRM2B have been postulated to be m^5U writer enzymes due to sequence similarity to TRM2 from yeast (25), but it has not yet been experimentally addressed if TRM2A and TRM2B are actually involved in RNA modification. We observed that m^5U abundance in total RNA was 36% and 42% lower in 20-day-old *trm2a* and *trm2b* seedlings (*trm2a*, SALK_085796; *trm2b*, SALK_106689; Supplementary Figure S5) while the double mutant even had a reduction by 93% compared with the wild type (Figure 5E), showing that indeed TRM2A and TRM2B play an important role in uridine methylation of

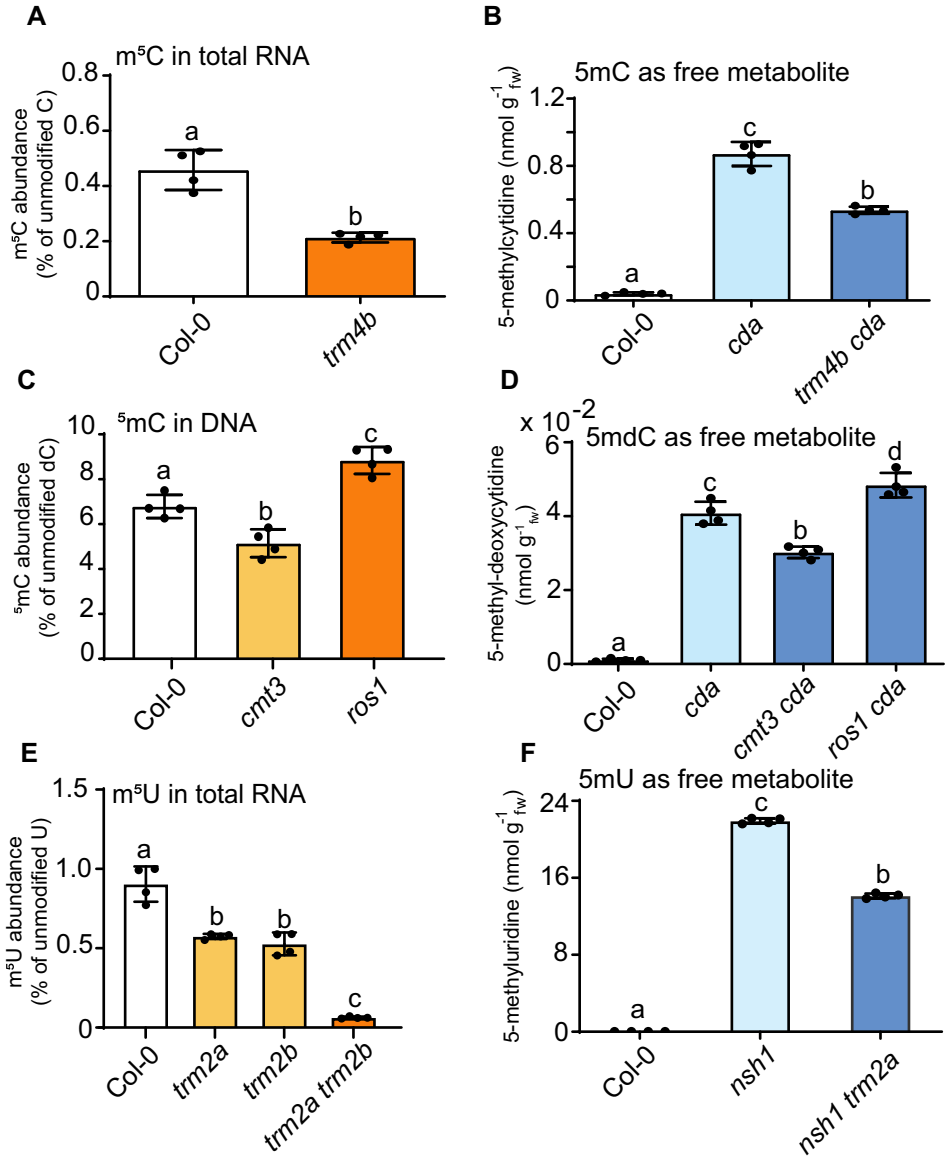


Figure 5. Determination of the origin of 5mC, 5mdC and 5mU in Arabidopsis. In all panels, 20-day-old seedlings were used. Error bars are the SD ($n = 4$ biological replicates). Different letters label data that are significantly different at $P < 0.05$. fw, fresh weight. (**A**) m^5C frequency relative to C in total RNA of the wild type (Col-0; white) and the *trm4b-1* line (dark orange). (**B**) 5mC contents of the wild type (Col-0, white), the *cda* line (light blue) and the *trm4b cda* double mutant line (dark blue). (**C**) ^{5}mC frequency relative to dC in DNA of the wild type (Col-0; white), the *cmt3-11* line (light orange) and the *ros1-4* line (dark orange). (**D**) 5mdC contents of the wild type (Col-0, white), the *cda-1* line (light blue) and *cmt3-11 cda* and *trm4b cda* double mutant lines (dark blue). (**E**) m^5U frequency relative to U in total RNA of the wild type (Col-0; white), the *trm2a* and *trm2b* lines (light orange) and the double mutant (dark orange). (**F**) 5mU content of the wild type (Col-0, white), the *nsh1-1* line (light blue) and *nsh1-1 trm2a* double mutant line (dark blue).

RNA. Then we raised the question of whether the NSH1 substrate 5mU is directly derived from methylated RNA. For this, we quantified 5mU in Col-0, *nsh1* and *nsh1 trm2a* seedlings and observed a 36% decrease of this compound in *nsh1 trm2a* plants compared with the single mutant of *NSH1* (Figure 5F). In line with the 25-fold stronger accumulation of 5mU in *nsh1* compared with 5mC in *cda* background, the data demonstrate that 5mU is derived not only from the deamination of 5mC but to a great extent directly from the turnover of m^5U-containing RNA *in vivo*. About 1% of uridines are 5-methylated in RNA (Figure 5E), thus we show that m^5U is a fairly abundant RNA modification in plants.

NSH1 protects RNA from m^5U misincorporation

As demonstrated above, several modified nucleosides accumulate in the *cda* or *nsh1* background (Figures 2 and 3). The highest amounts were observed for 5mC and 2OmC in *cda* and 5mU in *nsh1*. If such relatively abundant methylated pyrimidine nucleosides become phosphorylated to their corresponding triphosphates, they might represent substrates for RNA polymerases, leading to random methylation in newly synthesized RNA. Because RNA methylations are normally introduced at specific sites by writer proteins, such random methylation could disturb RNA function. Therefore, the catabolism of modified nucleosides by

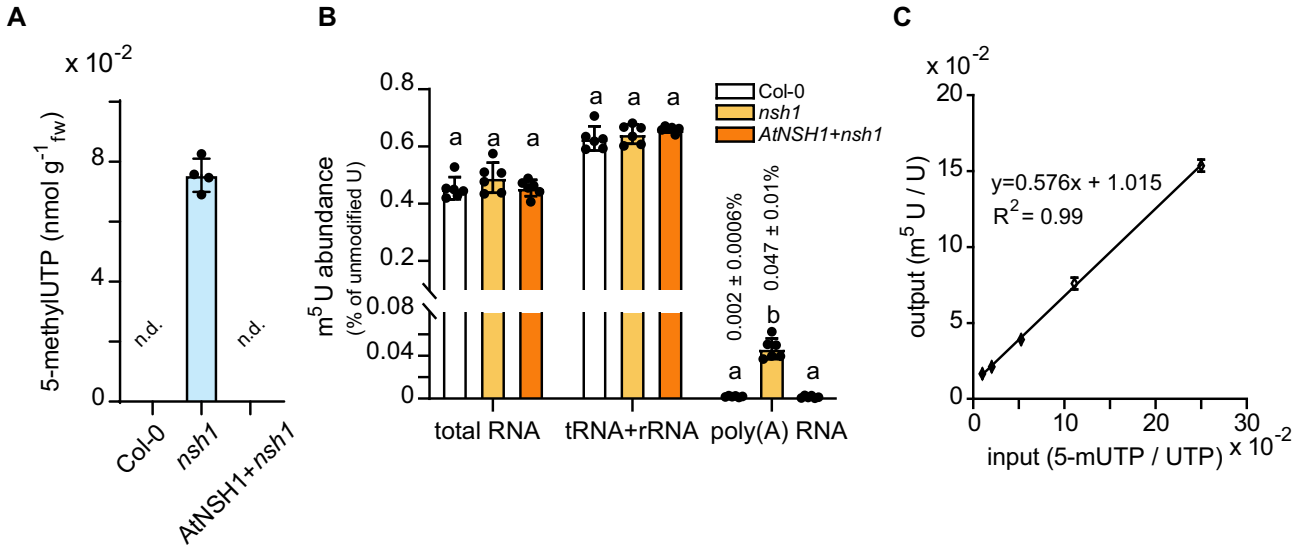


Figure 6. 5mUTP accumulation and m^5U frequency in different RNA species from Arabidopsis leaves varying in *NSH1* expression and selectivity of RNA polymerase II for UTP versus 5mUTP. (**A**) 5mUTP content in 20-day-old seedlings of the wild type (Col-0), the *nsh1* line (light blue) and the complementation line. Error bars are the SD ($n = 4$ biological replicates). The detection limit of 5mUTP was 50 pg on column. n.d., not detectable. fw, fresh weight. (**B**) m^5U frequency relative to U in total RNA, non-polyadenylated RNA (rRNA and tRNA) and mRNA of 20-day-old seedlings of the wild type (Col-0; white), the *nsh1* line (light orange) and the complementation line (dark orange). Error bars are the SD ($n = 6$ biological replicates). Different letters label data that are significantly different at $P < 0.05$. (**C**) Relative incorporation of m^5U referred to U into mRNA (output) synthesized *in vitro* by RNA polymerase II using a commercial HeLa cell transcription assay at varying input ratios of 5mUTP/UTP. Error bars are the SD ($n = 4$ independent transcription reactions).

CDA and NSH1 might be important to protect RNA from misincorporation of methylated nucleotides. To assess if 5mU can be phosphorylated *in vivo*, we quantified the 5-methyl-UTP (5mUTP) content in 20-day-old seedlings of the wild type, the *nsh1* line and the complementation line. 5mUTP was detected in *nsh1* seedlings but not in the wild type or the complementation line (Figure 6A). This shows that nucleoside and nucleotide kinases are able to phosphorylate 5mU to 5mUTP *in vivo*. Corresponding phosphorylation products of 5mC or 2OmC were not detected maybe because these nucleosides are at least 10-fold less abundant than 5mU in the respective mutants.

To test whether 5mUTP can be a substrate of RNA polymerases *in vivo*, the m^5U/U ratios were measured in mRNA, rRNA + tRNA, and total RNA from the wild type, *nsh1* and the complementation line. In total RNA or in tRNA + rRNA, the m^5U/U ratios were the same in all genotypes (Figure 6B). However, in highly purified mRNA, the m^5U/U ratio was significantly elevated in *nsh1* (mean: 0.047%) compared with the wild type or the complementation line (mean: 0.002%) (Figure 6B). That the m^5U content of mRNA was increased in the *nsh1* background indicated that 5mUTP is a substrate for RNA polymerase II. We investigated the selectivity of a eukaryotic RNA polymerase II for UTP versus 5mUTP, using an *in vitro* transcription assay employing HeLa cell extracts. A pre-set substrate ratio of 5mUTP to UTP was reflected in almost the same ratio in the RNA product (Figure 6C), showing that RNA polymerase II uses UTP and 5mUTP almost equally well. Thus, human RNA polymerase II must be protected from 5mUTP because it is not selective against this modified nucleotide. The same might be the case for RNA polymerase II from Arabidopsis since we measured more m^5U in mRNA of *nsh1*, which accumulates some 5mUTP. However, the selectivity of uridine kinases and uridine monophosphate kinases (uridylate kinases) for non-methylated substrates appears to suppress the production of 5mUTP, because in the *nsh1* background there is far less 5mUTP (0.076 nmol/g; Figure 6A) than 5mU (21.5 nmol/g; Figure 3A). Nonetheless, by degrading 5mU, NSH1 prevents the accumulation of this aberrant nucleoside and of 5mUTP. Together with the kinases, NSH1 thereby protects the mRNA from random incorporation of m^5U. A similar scenario has been described for protecting RNA and DNA from random accumulation of m^6A in Arabidopsis and mammalian cells (11,49,50).

We also measured the m^5C/C and C_m/C ratios in total RNA, rRNA + tRNA and highly purified mRNA, and the $^5mC/dC$ ratio in DNA from the wild type, the two *cda* mutants and the complementation line. However, we did not detect significant differences in these ratios for any of the genotypes (Supplementary Figure S6). Unlike NSH1, CDA does not seem to play a role in protecting the nucleic acids.

Seedling growth is reduced in the *NSH1* mutant

To investigate the biological consequences of compromised uridine and 5mU degradation, we assessed the physiological performance of the wild type, *nsh1* and the complementation line throughout their life cycles. While there were no apparent differences between the genotypes at later growth stages, many *nsh1* seedlings appeared smaller than those of the wild type or the complementation line early in development (Figure 7A). We selected 7-day-old seedlings (days after imbibition) grown on half-strength MS medium to quantify this phenotype. The mean seedling size was

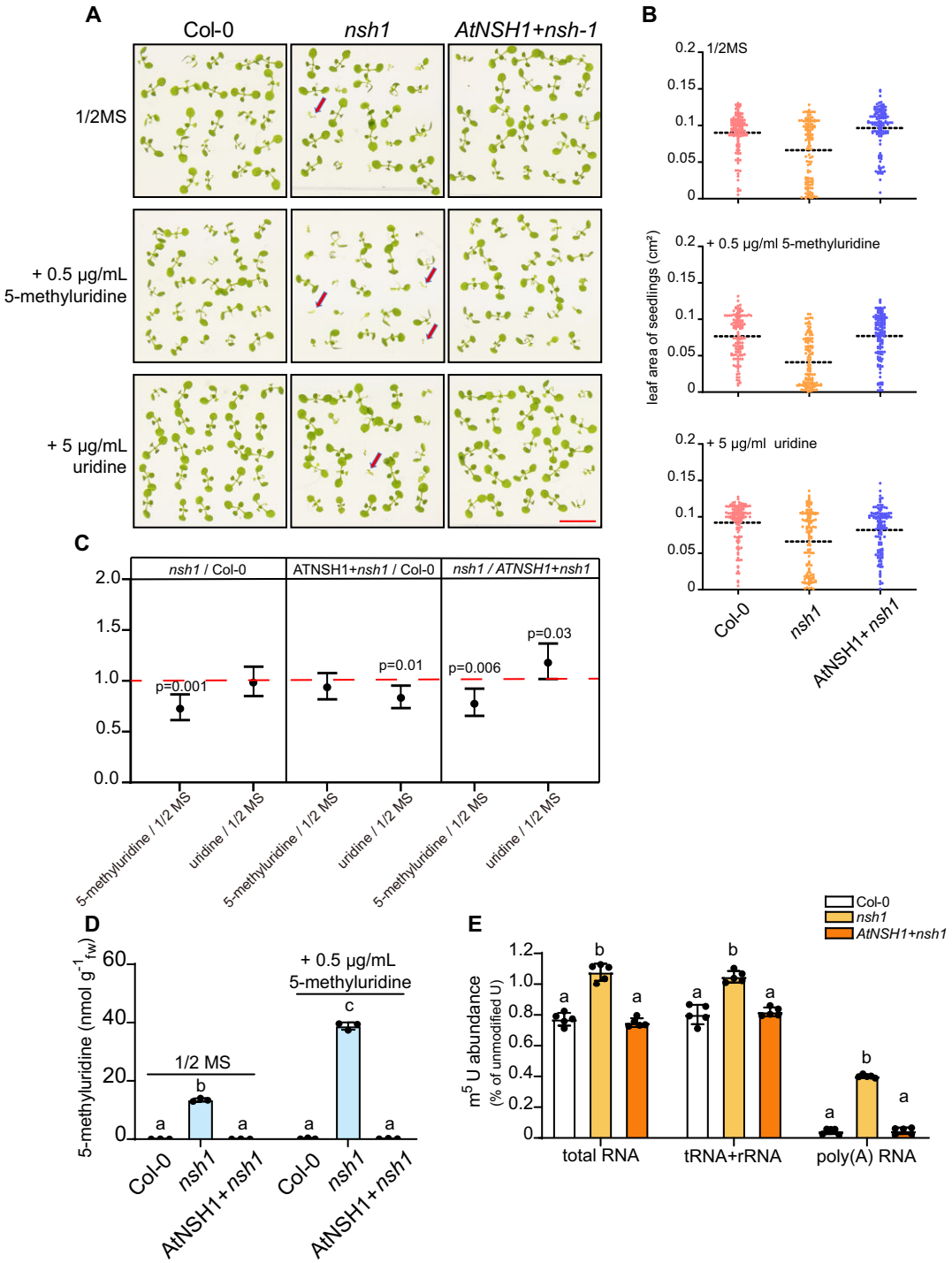


Figure 7. Growth phenotypes, metabolite profiles and m^5U frequencies in 7-day-old seedlings varying in *NSH1* expression grown in the presence or absence of 5mU or U. (**A**) Col-0, *nsh1* and the complementation line after 7 days of growth on standard medium (top panel) or in the presence of 0.5 µg/ml 5mU (middle panel) or 5 µg/ml U (bottom panel) under long-day conditions (16 h of light). Scale bar = 1 cm. (**B**) Leaf area quantification of Col-0, *nsh1* and the complementation line grown as shown in (A). $n = 125$ plants for each genotype from five replicates with 25 plants each germinated from seeds derived from different mother plants. (**C**) Statistical evaluation of data from (B). Ratios of ratios and confidence intervals at $P = 95\%$ as well as P-values are displayed. The ratio of the leaf area of a genotype on supplemented medium versus the leaf area on control medium was put in relation to the same ratio obtained from a second genotype. This allows statistical evaluation and quantification of whether a genotype is more strongly affected by 5mU or U treatment than another genotype. (**D**) 5mU contents in 7-day-old seedlings of the wild type (Col-0, white), the *nsh1* line (light blue) and the complementation line (dark blue) grown on standard medium or in presence of 0.5 µg/ml 5mU. Error bars are the SD ($n = 3$ biological replicates). Different letters label data that are significantly different at $P < 0.05$. fw, fresh weight. (**E**) m^5U frequencies relative to U in total RNA, non-polyadenylated RNA (rRNA and tRNA) and mRNA of 7-day-old seedlings of the wild type (Col-0; white), the *nsh1* line (light orange) and the complementation line (dark orange) grown in the presence of 0.5 µg/ml 5mU. Error bars are the SD ($n = 5$ biological replicates). Different letters label data that are significantly different at $P < 0.05$.

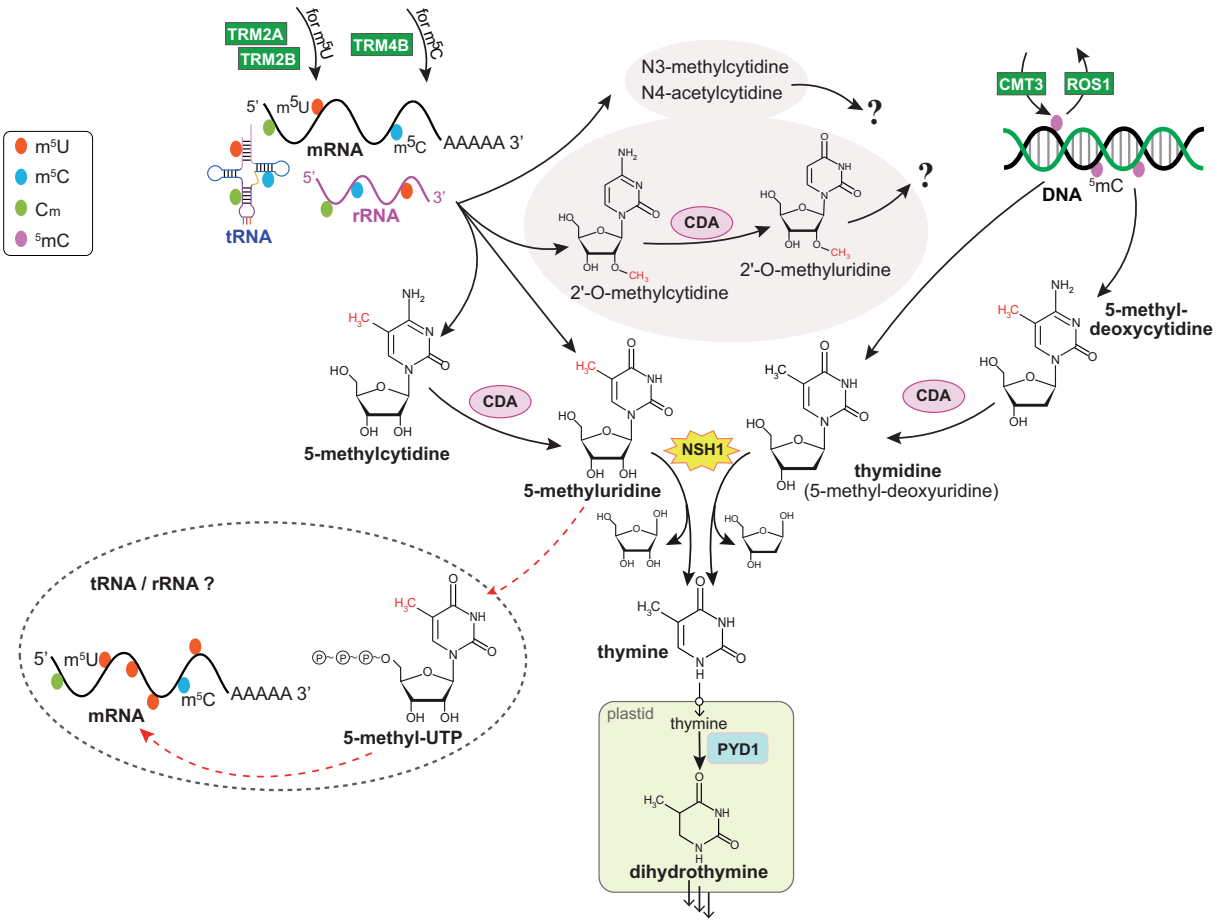


Figure 8. Pyrimidine methylation of RNA and DNA in Arabidopsis produces methylated nucleosides upon nucleic acid catabolism that enter different degradation pathways. Enzymatic reactions promoted in the mutant of *NSH1* are shown by red dashed arrows. Question marks indicate unresolved issues regarding the metabolic fate.

reduced in *nsh1* compared with the wild type or the complementation line (Figure 7B, top panel). We suspected that the partially hampered seedling development in *nsh1* might be caused by the misincorporation of m^5U into RNA (Figure 6B). To test this hypothesis, seedlings were grown on half-strength MS medium supplemented with 0.5 μg/ml 5mU or 5 μg/ml uridine. Supplementation with 5mU suppressed the growth of *nsh1* ∼27% more than it suppressed the growth of the wild type each relative to the control without 5mU in the medium (Figure 7C), whereas the 10-fold quantity of uridine in the medium did not lead to stronger growth depression of *nsh1* compared with the other genotypes relative to controls on media without supplements. Consistent with a negative effect on growth of 5mU, the wild type and the complementation line also appeared slightly smaller when grown on 5mU-containing media (Figure 7B), but not when grown in the presence of uridine. The 5mU content rose exclusively in *nsh1* seedlings by the external supply of this methylated nucleoside (Figure 7D). The elevated 5mU concentration was reflected in an ∼10-fold higher m^5U/U ratio in mRNA (mean: 0.40%) for seedlings grown in the presence of 5mU relative to seedlings grown in the absence of this compound (compare Figures 6B and 7E). In addition, the m^5U/U ratios in the total RNA and rRNA + tRNA pools isolated from *nsh1* seedlings grown in the presence of 5mU were also elevated relative to the ratios seen in the wild type and the complementation line. This indicates that RNA polymerases catalyzing tRNA/rRNA biosynthesis are also capable of using 5mUTP. Furthermore, all RNA species from all genotypes grown on medium containing 5mU had higher m^5U levels than those grown on half-strength MS medium alone (compare Figures 6B and 7E). Taken together, 5mU degradation by NSH1 appears to be a necessary house-cleaning reaction to prevent random RNA methylation at C5 of uridine. The data suggest that such RNA modifications negatively affect seedling establishment, although other reasons, for example metabolic alterations in *nsh1*, cannot be excluded as causes for the phenotypic alterations in this mutant.

DISCUSSION

Modified nucleosides occur naturally in RNA and DNA. We asked the question of how methylated (deoxy) pyrimidine nucleotides re-enter canonical nucleotide metabolism when nucleic acids are degraded in *A. thaliana* and have summarized our findings in a model (Figure 8). The data show that plant RNAs contain m^5U, which has also been reported for other organisms (2). Additionally, we show that TRM2A and TRM2B are required for m^5U modifications

of RNA. In *S. cerevisiae* there is only one *TRM2* gene and the corresponding enzyme is responsible for the methylation of uridine 54 (U54) in tRNAs (51). The function of TRM2A and TRM2B of plants has been predicted using the homology to the yeast enzyme (25), but so far their function in plants has not been investigated. There are also two *TRM2* genes in humans, with TRM2A recently shown to be responsible for the majority of m^5U modifications in RNA targeting mainly U54 of cytosolic tRNAs (43), while TRM2B is located in mitochondria targeting U54 of mitochondrial tRNAs as well as 12S rRNA (26). Our results show that in Arabidopsis, TRM2A and TRM2B contribute equally to m^5U methylation of RNA (Figure 5E). Together they are responsible for most such modifications which we speculate might also involve methylation of U54 of tRNAs but will probably also include rRNA targets. Upon RNA degradation, 5mU is released and the corresponding deoxynucleoside (dT) is derived from DNA turnover and repair (Figure 8). Both are substrates of NSH1, which generates thymine, a canonical nucleobase that can enter pyrimidine ring catabolism (34), and ribose or deoxyribose from these substrates. Notably, thymine is mostly derived from 5mU and, therefore, from RNA and not from DNA degradation (Figures 3A, B and 4A). In mammals and other eukaryotes, it has not yet been investigated how 5mC and 5mU from RNA degradation are metabolized, although the requirement to dispose of these modified nucleosides will probably be the same as in plants. The enzymatic makeup of pyrimidine nucleoside and nucleobase catabolism in plants and mammals is very similar (Supplementary Figure S1). Mammalian uridine phosphorylases release the base from the corresponding pyrimidine nucleoside, which in plants is the function of NSH1. Surprisingly, it has not yet been investigated if 5mU is a relevant substrate of eukaryotic uridine phosphorylases—either *in vitro* or *in vivo*. It is thus possible that compromised activity of this enzyme will lead to accumulation of 5mU, which may be of significant relevance to understand the physiological consequences of uridine phosphorylase deficiency in mammals.

Recently, it has been reported that m^5C is also present in plant RNAs (6), and it has long been known that DNA contains methylated dC (52). The methylated cytidines, 5mC and 2OmC from RNA degradation and 5mdC from DNA breakdown, are all *in vivo* substrates of CDA and are converted to the respective uridines (Figures 1 and 2). Two of the products, 5mU and dT, are then substrates of NSH1 and are hydrolyzed to thymine and ribose or 2-deoxyribose, respectively. Our data suggest that only 4% of 5mU is derived from 5mC because in *nsh1* seedlings there is 25-fold more 5mU than there is 5mC in the *cda* background (Figures 2A and 3A). Although the quantitative aspects of these data should not be overinterpreted, they do show that a large proportion of 5mU is derived directly from RNA without a 5mC intermediate. Interestingly, 2OmC is an *in vivo* substrate of CDA but the resulting 2OmU, which can also be directly derived from RNA, is not degraded by NSH1 either *in vitro* or *in vivo*. This suggests that plants possess a so far unknown enzyme for the demethylation of 2OmU. The methylation of (deoxy)ribose at the OH-group at carbon 2 is widespread and found in prokaryotes and eukaryotes (53,54), but enzymes that remove this methylation have not yet been described in any organism.

For canonical (deoxy)nucleosides, it has been shown that their concentrations impact on the amounts of (deoxy)nucleotides (20,55,56). However, the kinases that connect these metabolite pools are probably selective against modified nucleosides and nucleotides, which has been observed for adenylate kinases that phosphorylate AMP far better than N^6-mAMP (11). Here we show that 5mU can be converted to 5mUTP, when the methylated nucleoside accumulates in the *nsh1* background, and that 5mUTP represents a good substrate for RNA polymerase II *in vitro*. We assume that this is also the case *in vivo*, resulting in significantly elevated m^5U levels in mRNA (Figure 6; Figure 8, reactions occurring only in the *nsh1* mutant background are depicted by red dashed arrows). That RNA polymerases can use 5-substituted UTPs is well known and can be exploited for biotechnological applications (57). Interestingly, *in vivo* labeling of RNA in HeLa cells was only possible with 5-azide-substituted UTP but not with the corresponding nucleosides, suggesting that the aberrant nucleosides are not well recognized by the kinases that convert uridine to UTP, whereas the RNA polymerase is able to accept the modified UTP (58).

It is known that the mutant of *NSH1* has an increased content of several nucleosides and nucleotides (20,22–24,55). We observed a reduced growth performance of the mutant at the seedling stage, which could be due to the metabolic disturbance in this genetic background. However, the defect in growth was specifically aggravated by cultivating the *nsh1* plants on medium containing 5mU but not by growing them on medium containing 10-fold more uridine. Therefore, the *nsh1* growth phenotype seems rather to be associated with random m^5U incorporation into RNA, which became more prominent upon 5mU feeding (Figure 7), than with other metabolic alterations. By supplementing 5mU from outside to the *nsh1* mutant, we could detect elevated amounts of m^5U in rRNA and tRNA, showing that in principle this modified nucleoside can find its way into these RNAs *in vivo* (Figure 7D, E). However, given the stochastic nature of aberrant m^5U modification in RNA species in *nsh1*, it is at this stage impossible to demonstrate that the growth phenotype is caused by RNA alterations.

In summary we show that the function of NSH1 is not merely to recycle nutrients but also to protect RNA from random methylation. Such a role cannot be postulated for CDA because *CDA* mutants were unchanged in terms of RNA methylation and growth. NSH1 and CDA are required for the catabolism of several methylated pyrimidine nucleosides, but not for all modified pyrimidines that occur *in vivo*. For example, 2′-O-methylated uridine is not a substrate of NSH1, and N^4-acetylcytidine is not catabolized by CDA, suggesting that there must be other currently unknown catabolic enzymes for these compounds. Many other modified nucleosides, such as N^1-methyladenosine and 7-methylguanosine, also occur in RNA. Elucidating the metabolism of these nucleosides will be an interesting future challenge. For plants, our data highlight that the metabolism of RNA degradation products can be critical beyond preventing (deoxy)nucleotide imbalance (20,42) or compromising nucleobase and nucleoside salvage. Since

there are many similarities in pyrimidine catabolism between eukaryotes, this issue needs to be explored for other organisms, in particular humans.

DATA AVAILABILITY

All experimental data are available in the main text or in the Supplementary data. Expression plasmids and mutant lines are available on reasonable request from the authors. Sequence data from this article can be found under the following accession/locus numbers: At2g19570 (*CDA*), At2g36310 (*NSH1*), At3g17810 (*PYD1*), At2g22400 (*TRM4B*), At3g21300 (*TRM2A*), At2g28450 (*TRM2B*), At1g69770 (*CMT3*) and At2g36490 (*ROS1*).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Professor Honggui La for kindly sharing *cmt3-11* and *ros1-4* mutant seeds, Jialei Wang for help with mass spectrometry analyses, and Peng Liu for assistance with seedling phenotype analyses.
Author contributions: M.C. and C.-P.W. conceived the research. S.G., Y.S., X.C., C.Z., X.L., W.W., L.G. and Y.L. performed the experiments. M.C., M.H., F.S. and C.-P.W. analyzed the data, and M.C. and C.-P.W. wrote the manuscript. All authors proofread and agreed with the manuscript.

FUNDING

The Fundamental Research Funds for the Central Universities [YDZX2023013]; the Deutsche Forschungsgemeinschaft (DFG) [CH2292/1-1 to M.C. and C.-P.W., WI3411/8-1 and INST 187/741-1 FUGG to C.-P. W., and HE-5949/4-1 to M.H.]; the International Centre for Genetic Engineering and Biotechnology [CRP/CHN20-04_EC]; the National Natural Science Foundation of China [31900907]; the Natural Science Foundation of Jiangsu Province, China [BK20190528]; and Nanjing Agricultural University [start-up fund for advanced talents to M.C.]. Funding for open access charge: Nanjing Agricultural University, China.
Conflict of interest statement. None declared.

REFERENCES

1. Chou,H.-J., Donnard,E., Gustafsson,H.T., Garber,M. and Rando,O.J. (2017) Transcriptome-wide analysis of roles for tRNA modifications in translational regulation. *Mol. Cell*, **68**, 978–992
2. McCown,P.J., Ruszkowska,A., Kunkler,C.N., Breger,K., Hulewicz,J.P., Wang,M.C., Springer,N.A. and Brown,J.A. (2020) Naturally occurring modified ribonucleosides. *Wiley Interdiscip. Rev. RNA*, **11**, e1595.
3. Zhong,S., Li,H., Bodi,Z., Button,J., Vespa,L., Herzog,M. and Fray,R.G. (2008) MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell*, **20**, 1278–1288.
4. Dominissini,D., Moshitch-Moshkovitz,S., Schwartz,S., Salmon-Divon,M., Ungar,L., Osenberg,S., Cesarkas,K., Jacob-Hirsch,J., Amariglio,N., Kupiec,M. *et al.* (2012) Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. *Nature*, **485**, 201–206.
5. Dominissini,D., Nachtergaele,S., Moshitch-Moshkovitz,S., Peer,E., Kol,N., Ben-Haim,M.S., Dai,Q., Di Segni,A., Salmon-Divon,M., Clark,W.C. *et al.* (2016) The dynamic N^1-methyladenosine methylome in eukaryotic messenger RNA. *Nature*, **530**, 441–446.
6. Cui,X., Liang,Z., Shen,L., Zhang,Q., Bao,S., Geng,Y., Zhang,B., Leo,V., Vardy,L.A., Lu,T. *et al.* (2017) 5-Methylcytosine RNA methylation in *Arabidopsis thaliana*. *Mol. Plant*, **10**, 1387–1399.
7. Arango,D., Sturgill,D., Alhusaini,N., Dillman,A.A., Sweet,T.J., Hanson,G., Hosogane,M., Sinclair,W.R., Nanan,K.K., Mandler,M.D. *et al.* (2018) Acetylation of cytidine in mRNA promotes translation efficiency. *Cell*, **175**, 1872–1886.
8. Carlile,T.M., Rojas-Duran,M.F., Zinshteyn,B., Shin,H., Bartoli,K.M. and Gilbert,W.V. (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*, **515**, 143–146.
9. Gilbert,W.V., Bell,T.A. and Schaening,C. (2016) Messenger RNA modifications. Form, distribution, and function. *Science*, **352**, 1408–1412.
10. Frye,M., Harada,B.T., Behm,M. and He,C. (2018) RNA modifications modulate gene expression during development. *Science*, **361**, 1346–1349.
11. Chen,M., Urs,M.J., Sánchez-González,I., Olayioye,M.A., Herde,M. and Witte,C.-P. (2018) m^6A RNA degradation products are catabolized by an evolutionarily conserved N^6-methyl-AMP deaminase in plant and mammalian cells. *Plant Cell*, **30**, 1511–1522.
12. Witte,C.-P. and Herde,M. (2020) Nucleotide metabolism in plants. *Plant Physiol.*, **182**, 63–78.
13. Chen,M. and Witte,C.-P. (2020) A kinase and a glycosylase catabolize pseudouridine in the peroxisome to prevent toxic pseudouridine monophosphate accumulation. *Plant Cell*, **32**, 722–739.
14. Kim,S.-H., Witte,C.-P. and Rhee,S. (2021) Structural basis for the substrate specificity and catalytic features of pseudouridine kinase from *Arabidopsis thaliana*. *Nucleic Acids Res.*, **49**, 491–503.
15. Sekula,P., Dettmer,K., Vogl,F.C., Gronwald,W., Ellmann,L., Mohney,R.P., Eckardt,K.-U., Suhre,K., Kastenmüller,G., Oefner,P.J. *et al.* (2017) From discovery to translation: characterization of C-mannosyltryptophan and pseudouridine as markers of kidney function. *Sci. Rep.*, **7**, 17400.
16. Yang,X., Yang,Y., Sun,B.-F., Chen,Y.-S., Xu,J.-W., Lai,W.-Y., Li,A., Wang,X., Bhattarai,D.P., Xiao,W. *et al.* (2017) 5-methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m^5C reader. *Cell Res.*, **27**, 606–625.
17. David,R., Burgess,A., Parker,B., Li,J., Pulsford,K., Sibbritt,T., Preiss,T. and Searle,I.R. (2017) Transcriptome-wide mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and noncoding RNAs. *Plant Cell*, **29**, 445–460.
18. Tang,Y., Gao,C.-C., Gao,Y., Yang,Y., Shi,B., Yu,J.-L., Lyu,C., Sun,B.-F., Wang,H.-L., Xu,Y. *et al.* (2020) OsNSUN2-mediated 5-methylcytosine mRNA modification enhances rice adaptation to high temperature. *Dev. Cell*, **53**, 272–286.
19. Pfaff,C., Ehrnsberger,H.F., Flores-Tornero,M., Sørensen,B.B., Schubert,T., Längst,G., Griesenbeck,J., Sprunck,S., Grasser,M. and Grasser,K.D. (2018) ALY RNA-binding proteins are required for nucleocytosolic mRNA transport and modulate plant growth and development. *Plant Physiol.*, **177**, 226–240.
20. Straube,H., Niehaus,M., Zwittian,S., Witte,C.-P. and Herde,M. (2021) Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. *Plant Cell*, **33**, 270–289.
21. Chen,M., Herde,M. and Witte,C.-P. (2016) Of the nine cytidine deaminase-like genes in Arabidopsis, eight are pseudogenes and only one is required to maintain pyrimidine homeostasis in vivo. *Plant Physiol.*, **171**, 799–809.
22. Jung,B., Floerchinger,M., Kunz,H.-H., Traub,M., Wartenberg,R., Jeblick,W., Neuhaus,H.E. and Moehlmann,T. (2009) Uridine-ribohydrolase is a key regulator in the uridine degradation pathway of Arabidopsis. *Plant Cell*, **21**, 876–891.

23. Riegler,H., Geserick,C. and Zrenner,R. (2011) *Arabidopsis thaliana* nucleosidase mutants provide new insights into nucleoside degradation. *New Phytol.*, **191**, 349–359.
24. Jung,B., Hoffmann,C. and Moehlmann,T. (2011) Arabidopsis nucleoside hydrolases involved in intracellular and extracellular degradation of purines. *Plant J.*, **65**, 703–711.
25. Wang,Y., Pang,C., Li,X., Hu,Z., Lv,Z., Zheng,B. and Chen,P. (2017) Identification of tRNA nucleoside modification genes critical for stress response and development in rice and Arabidopsis. *BMC Plant Biol.*, **17**, 261.
26. Powell,C.A. and Minczuk,M. (2020) TRMT2B is responsible for both tRNA and rRNA m^5U-methylation in human mitochondria. *RNA Biol.*, **17**, 451–462.
27. Tardu,M., Jones,J.D., Kennedy,R.T., Lin,Q. and Koutmou,K.S. (2019) Identification and quantification of modified nucleosides in *Saccharomyces cerevisiae* mRNAs. *ACS Chem. Biol.*, **14**, 1403–1409.
28. Cheng,Q.-Y., Xiong,J., Ma,C.-J., Dai,Y., Ding,J.-H., Liu,F.-L., Yuan,B.-F. and Feng,Y.-Q. (2020) Chemical tagging for sensitive determination of uridine modifications in RNA. *Chem. Sci.*, **11**, 1878–1891.
29. Vincenzetti,S., Cambi,A., Neuhard,J., Garattini,E. and Vita,A. (1996) Recombinant human cytidine deaminase: expression, purification, and characterization. *Protein Expr. Purif.*, **8**, 247–253.
30. Cao,D., Leffert,J.J., McCabe,J., Kim,B. and Pizzorno,G. (2005) Abnormalities in uridine homeostatic regulation and pyrimidine nucleotide metabolism as a consequence of the deletion of the uridine phosphorylase gene. *J. Biol. Chem.*, **280**, 21169–21175.
31. Le,T.T., Ziemba,A., Urasaki,Y., Hayes,E., Brotman,S. and Pizzorno,G. (2013) Disruption of uridine homeostasis links liver pyrimidine metabolism to lipid accumulation. *J. Lipid Res.*, **54**, 1044–1057.
32. López,L.C., Akman,H.O., García-Cazorla,A., Dorado,B., Martí,R., Nishino,I., Tadesse,S., Pizzorno,G., Shungu,D., Bonilla,E. *et al.* (2009) Unbalanced deoxynucleotide pools cause mitochondrial DNA instability in thymidine phosphorylase-deficient mice. *Hum. Mol. Genet.*, **18**, 714–722.
33. van Kuilenburg,A.B.P., Meinsma,R., Beke,E., Assmann,B., Ribes,A., Lorente,I., Busch,R., Mayatepek,E., Abeling,N.G.G.M., van Cruchten,A. *et al.* (2004) Beta-ureidopropionase deficiency: an inborn error of pyrimidine degradation associated with neurological abnormalities. *Hum. Mol. Genet.*, **13**, 2793–2801.
34. Zrenner,R., Riegler,H., Marquard,C.R., Lange,P.R., Geserick,C., Bartosz,C.E., Chen,C.T. and Slocum,R.D. (2009) A functional analysis of the pyrimidine catabolic pathway in Arabidopsis. *New Phytol.*, **183**, 117–132.
35. Alonso,J.M., Stepanova,A.N., Leisse,T.J., Kim,C.J., Chen,H.M., Shinn,P., Stevenson,D.K., Zimmerman,J., Barajas,P., Cheuk,R. *et al.* (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, **301**, 653–657.
36. Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C. *et al.* (2002) A high-throughput Arabidopsis reverse genetics system. *Plant Cell*, **14**, 2985–2994.
37. Kleinboelting,N., Huep,G., Kloetgen,A., Viehoever,P. and Weisshaar,B. (2012) GABI-Kat SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database. *Nucleic Acids Res.*, **40**, D1211–D1215.
38. Baccolini,C. and Witte,C.-P. (2019) AMP and GMP catabolism in Arabidopsis converge on xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. *Plant Cell*, **31**, 734–751.
39. Witte,C.P. and Medina-Escobar,N. (2001) In-gel detection of urease with nitroblue tetrazolium and quantification of the enzyme from different crop plants using the indophenol reaction. *Anal Biochem.*, **290**, 102–107.
40. Frances,A. and Cordelier,P. (2020) The emerging role of cytidine deaminase in human diseases: a new opportunity for therapy. *Mol. Ther.*, **28**, 357–366.
41. Gemble,S., Buhagiar-Labarchède,G., Onclercq-Delic,R., Biard,D., Lambert,S. and Amor-Guéret,M. (2016) A balanced pyrimidine pool is required for optimal Chk1 activation to prevent ultrafine anaphase bridge formation. *J. Cell. Sci.*, **129**, 3167–3177.
42. Gemble,S., Ahuja,A., Buhagiar-Labarchède,G., Onclercq-Delic,R., Dairou,J., Biard,D.S.F., Lambert,S., Lopes,M. and Amor-Guéret,M. (2015) Pyrimidine pool disequilibrium induced by a cytidine deaminase deficiency inhibits PARP-1 activity, leading to the under replication of DNA. *PLoS Genet.*, **11**, e1005384.
43. Carter,J.-M., Emmett,W., Mozos,I.R., Kotter,A., Helm,M., Ule,J. and Hussain,S. (2019) FICC-Seq: a method for enzyme-specified profiling of methyl-5-uridine in cellular RNA. *Nucleic Acids Res.*, **47**, e113.
44. Cornelius,S., Witz,S., Rolletschek,H. and Möhlmann,T. (2011) Pyrimidine degradation influences germination seedling growth and production of Arabidopsis seeds. *J. Exp. Bot.*, **62**, 5623–5632.
45. Roldán-Arjona,T., Ariza,R.R. and Córdoba-Cañero,D. (2019) DNA base excision repair in plants: an unfolding story with familiar and novel characters. *Front. Plant Sci.*, **10**, 1055.
46. Kong,W., Li,B., Wang,Q., Wang,B., Duan,X., Ding,L., Lu,Y., Liu,L.-W. and La,H. (2018) Analysis of the DNA methylation patterns and transcriptional regulation of the NB-LRR-encoding gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.*, **96**, 563–575.
47. Forgione,I., Wołoszyńska,M., Pacenza,M., Chiappetta,A., Greco,M., Araniti,F., Abenavoli,M.R., van Lijsebettens,M., Bitonti,M.B. and Bruno,L. (2019) Hypomethylated drm1 drm2 cmt3 mutant phenotype of *Arabidopsis thaliana* is related to auxin pathway impairment. *Plant Sci*, **280**, 383–396.
48. Li,J., Yang,D.-L., Huang,H., Zhang,G., He,L., Pang,J., Lozano-Durán,R., Lang,Z. and Zhu,J.-K. (2020) Epigenetic memory marks determine epiallele stability at loci targeted by de novo DNA methylation. *Nat. Plants*, **6**, 661–674.
49. Musheev,M.U., Baumgärtner,A., Krebs,L. and Niehrs,C. (2020) The origin of genomic N^6-methyl-deoxyadenosine in mammalian cells. *Nat. Chem. Biol.*, **16**, 630–634.
50. Liu,X., Lai,W., Li,Y., Chen,S., Liu,B., Zhang,N., Mo,J., Lyu,C., Zheng,J., Du,Y.-R. *et al.* (2021) N^6-methyladenine is incorporated into mammalian genome by DNA polymerase. *Cell Res*, **31**, 94–97.
51. Nordlund,M.E., Johansson,J.O., von Pawel-Rammingen,U. and Byström,A.S. (2000) Identification of the TRM2 gene encoding the tRNA(m^5U54) methyltransferase of *Saccharomyces cerevisiae*. *RNA*, **6**, 844–860.
52. Zhang,H., Lang,Z. and Zhu,J.-K. (2018) Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.*, **19**, 489–506.
53. Azevedo-Favory,J., Gaspin,C., Ayadi,L., Montacié,C., Marchand,V., Jobet,E., Rompais,M., Carapito,C., Motorin,Y. and Sáez-Vásquez,J. (2021) Mapping rRNA 2'-O-methylations and identification of C/D snoRNAs in *Arabidopsis thaliana* plants. *RNA Biol.*, **18**, 1760–1777.
54. Ramakrishnan,M., Rajan,K.S., Mullasseri,S., Palakkal,S., Kalpana,K., Sharma,A., Zhou,M., Vinod,K.K., Ramasamy,S. and Wei,Q. (2022) The plant epitranscriptome: revisiting pseudouridine and 2'-O-methyl RNA modifications. *Plant Biotechnol. J.*, **20**, 1241–1256.
55. Heinemann,K.J., Yang,S.-Y., Straube,H., Medina-Escobar,N., Varbanova-Herde,M., Herde,M., Rhee,S. and Witte,C.-P. (2021) Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase. *Nat. Commun.*, **12**, 6846.
56. Chen,X., Kim,S.-H., Rhee,S. and Witte,C.-P. (2023) A plastid nucleoside kinase is involved in inosine salvage and control of purine nucleotide biosynthesis. *Plant Cell*, **35**, 510–528.
57. Milisavljevič,N., Perlíková,P., Pohl,R. and Hocek,M. (2018) Enzymatic synthesis of base-modified RNA by T7 RNA polymerase. A systematic study and comparison of 5-substituted pyrimidine and 7-substituted 7-deazapurine nucleoside triphosphates as substrates. *Org. Biomol. Chem.*, **16**, 5800–5807.
58. Sawant,A.A., Tanpure,A.A., Mukherjee,P.P., Athavale,S., Kelkar,A., Galande,S. and Srivatsan,S.G. (2016) A versatile toolbox for posttranscriptional chemical labeling and imaging of RNA. *Nucleic Acids Res.*, **44**, e16.