농학박사학위논문

애기장대 유래 **Pseudouridine Kinase** 의 구조와 생화학적 기능 분석

Structural and Biochemical Analysis of Pseudouridine Kinase from *Arabidopsis thaliana*

2021 년 8 월

서울대학교 대학원

농생명공학부 응용생명화학전공

김 상 훈

A Dissertation for the Degree of Doctor of Philosophy

Structural and Biochemical Analysis of Pseudouridine Kinase from *Arabidopsis thaliana*

August 2021

Sang-Hoon Kim

Applied Life Chemistry Major

Department of Agricultural Biotechnology

Seoul National University

애기장대 유래 **Pseudouridine Kinase** 의

구조와 생화학적 기능 분석

Structural and Biochemical Analysis of

Pseudouridine Kinase from *Arabidopsis thaliana*

지도교수 이 상 기

이 논문을 농학박사학위논문으로 제출함
2021년 8월

서울대학교 대학원
농생명공학부 응용생명화학전공
김 상 훈

김상훈의 박사학위논문을 인준함
2021년 8월

위 원 장 _____

부 위 원 장 _____

위 원 _____

위 원 _____

위 원 _____

Structural and Biochemical Analysis of Pseudouridine Kinase from *Arabidopsis thaliana*

Advisor: Sangkee Rhee

A Dissertation Submitted in Partial Fulfillment
of the Requirement for the Degree of

DOCTOR OF PHILOSOPHY

to the Faculty of
Applied Life Chemistry Major,
Department of Agricultural Biotechnology

at

SEOUL NATIONAL UNIVERSITY

by

Sang-Hoon Kim

Date Approved

July, 2021

Ki-Bong Oh

Sangkee Rhee

Euiyoung Bae

Jeong-Young Suh

Sung-Gun Kim

Abstract

RNA modifications are chemically diverse site-specific events, achieved in most cases through an enzyme-dependent reaction. They regulate the stability of RNAs, mRNA–protein interactions, and translation efficiency. Compared to studies related to the biological role and biosynthesis process of RNA modifications, studies on the degradation of modified RNA are relatively incomplete. Recently, metabolic fate of pseudouridine, one of the most prevalent modified RNAs, was characterized in plant *Arabidopsis thaliana*. In the *A. thaliana*, two enzymes are responsible for pseudouridine degradation: PSEUDOURIDINE KINASE (PUKI) and PSEUDOURIDINE MONOPHOSPHATE GLYCOSYLASE (PUMY). PUKI and PUMY are involved in phosphorylating pseudouridine into pseudouridine monophosphate and hydrolyzing pseudouridine monophosphate into uracil and ribose 5'-phosphate, respectively. The resulting products can be subjected to a general pathway for pyrimidine catabolism or the salvage pathway. In this thesis, I conducted structural and biochemical analyses of PUKI from *A. thaliana* (AtPUKI), the enzyme catalyzing the first step in pseudouridine degradation. AtPUKI, a member of the phosphofructokinase B (PfkB) family of carbohydrate kinases, is a homodimeric α/β protein with a protruding small β-strand domain, which serves simultaneously as dimerization interface and dynamic substrate specificity determinant. AtPUKI has a unique nucleoside binding site specifying the binding of pseudourine, of which one is mediated by a loop from the small β-strand domain of the adjacent monomer. Conformational transition of the dimerized small β-strand domains containing active site residues is required for substrate specificity. This dynamic feature explains the higher catalytic efficiency for pseudouridine over uridine. Both

substrates bind well to the AtPUKI with similar K_m value, but only pseudouridine is turned over efficiently. These results provide an example for structural and functional divergence in the PfkB family and highlight how AtPUKI avoids futile uridine phosphorylation which *in vivo* would disturb pyrimidine homeostasis.

Key words: Pseudouridine catabolism, Non-canonical nucleotide catabolism, Pseudouridine kinase, Crystal structure, Substrate specificity, *Arabidopsis thaliana*

Student Number: 2013-23265

Contents

Abstract ..... **i**
Contents ..... **iii**
List of Figures ..... **v**
List of Tables ..... **vii**
List of Abbreviations ..... **viii**

Introduction **1**
RNA modification 2
Chemical property and role of pseudouridine 5
Catabolism of canonical pyrimidine nucleoside 6
Catabolism of non-canonical nucleotides 10
Materials and Methods **17**
Cloning and purification of AtPUKI 18
Crystallization and structure determination 27
Activity assays 33
Results and Discussion **37**
Structure of unliganded AtPUKI 38
Structural homologs of AtPUKI and the structure of the monovalent cation-binding site 47
The ternary complex of AtPUKI with pseudouridine and ADP 55
Functional features of the AtPUKI active site residues 67
High fidelity of AtPUKI for pseudouridine 88
The substrate pocket of AtPUKI in comparison to ribokinase and adenosine kinase 90
The Mg^{2+}-binding site and catalysis in AtPUKI 93

Reference.....**97**
Accession numbers.....**102**
Abstract in Korean.....**103**
Acknowledgement.....**105**

List of Figures

Introduction

Figure 1. Types and functions of RNA modification 3

Figure 2. Metabolism of uridine 8

Figure 3. Reaction scheme of pseudouridine catabolism 13

Figure 4. Sequence comparison of PUKI from different sources 15

Materials and Methods

Figure 5. scheme of the construct for AtPUKI purification 20

Figure 6. Purification profiles of AtPUKI 22

Figure 7. SDS-PAGE analysis of AtPUKI 24

Figure 8. Crystal of AtPUKI 29

Figure 9. Scheme of enzyme-coupled assay for steady-state kinetic assay of AtPUKI 35

Results and Discussion

Figure 10. Determination of the molecular weight of soluble AtPUKI 41

Figure 11. Monomeric and dimeric structures of AtPUKI 43

Figure 12. Structural comparison of the two subunits of the unliganded AtPUKI 45

Figure 13. Effects of metal ions, different nucleoside triphosphates, and phosphate on AtPUKI activity 49

Figure 14. The monovalent cation-binding site in AtPUKI 51

Figure 15. Metal analyses by inductively coupled plasma–atomic emission spectroscopy (ICP-AES) and crystallographic refinements 53

Figure 16. Structural features for the ternary complex of AtPUKI with pseudouridine and ADP 58

Figure 17. Conformational changes in the ternary complex 61

Figure 18. Pseudouridine in the AtPUKI ternary complex 63

Figure 19. Possible binding modes of pseudouridine and ATP in AtPUKI 65

Figure 20. Kinetic analyses of wild-type AtPUKI 70

Figure 21. Substrate specificity of AtPUKI and chemical structure of substrate candidates.....72

Figure 22. Activity comparisons using a variety of AtPUKI mutants.....74

Figure 23. Kinetic parameters of AtPUKI variants with pseudouridine and structurally and chemically similar nucleosides.....76

Figure 24. Kinetic analyses of AtPUKI mutants for pseudouridine.....78

Figure 25. Kinetic analyses of AtPUKI variants for uridine 81

Figure 26. Kinetic analyses of AtPUKI variants for cytidine 84

Figure 27. Kinetic analyses of AtPUKI variants for 5-methyl uridine.....86

Figure 28. The substrate pocket of AtPUKI, ribokinase, and adenosine kinase.....91

Figure 29. The Mg^{2+}-binding site of AtPUKI 95

List of Tables

Table 1. List of AtPUKI mutants used in this study 26
Table 2. Data collection and refinement statistics.....31

List of Abbreviation

PUKI	PSEUDOURIDINE KINASE
PUMY	PSEUDOURIDINE MONOPHOSPHATE GLYCOSYLASE
m^6A	N^6-methylated adenine
m^1A	N^1-methylated adenine
Ψ	pseudouridine
mRNA	messenger RNA
tRNA	transfer RNA
rRNA	ribosomal RNA
N^6-mAMP	N^6-methyl-adenosine monophosphate
ΨMP	pseudouridine monophosphate
PfkB	phosphofructokinase B
ATP	adenosine triphosphate
GTP	guanosine triphosphate
CTP	cytidine triphosphate
UTP	uridine triphosphate
ADP	adenosine diphosphate
TEV	tobacco etch virus
MBP	maltose binding protein
SeMet	seleno-L-methionine
IMAC	immobilized metal affinity chromatography
SEC	size exclusion chromatography
UDP	uridine diphosphate

UMP	uridine monophosphate
NC-β-alanine	N-carbamoyl-β-alanine
PRPP	5-phospho-α-D-ribose 1-diphosphate
NSH1	nucleoside hydrolase 1
PLUTO	plastidic nucleobase transporter
NDPK	nucleoside diphosphate kinase
UMK	UMP kinase
UPRT	uracil phosphoribosyltransferase
DPYD	dihydropyrimidine dehydrogenase
DPYH	dihydropyrimidine hydrolase
β-UP	β-ureidopropionase
BAAT	β-Ala aminotransferase

Introduction

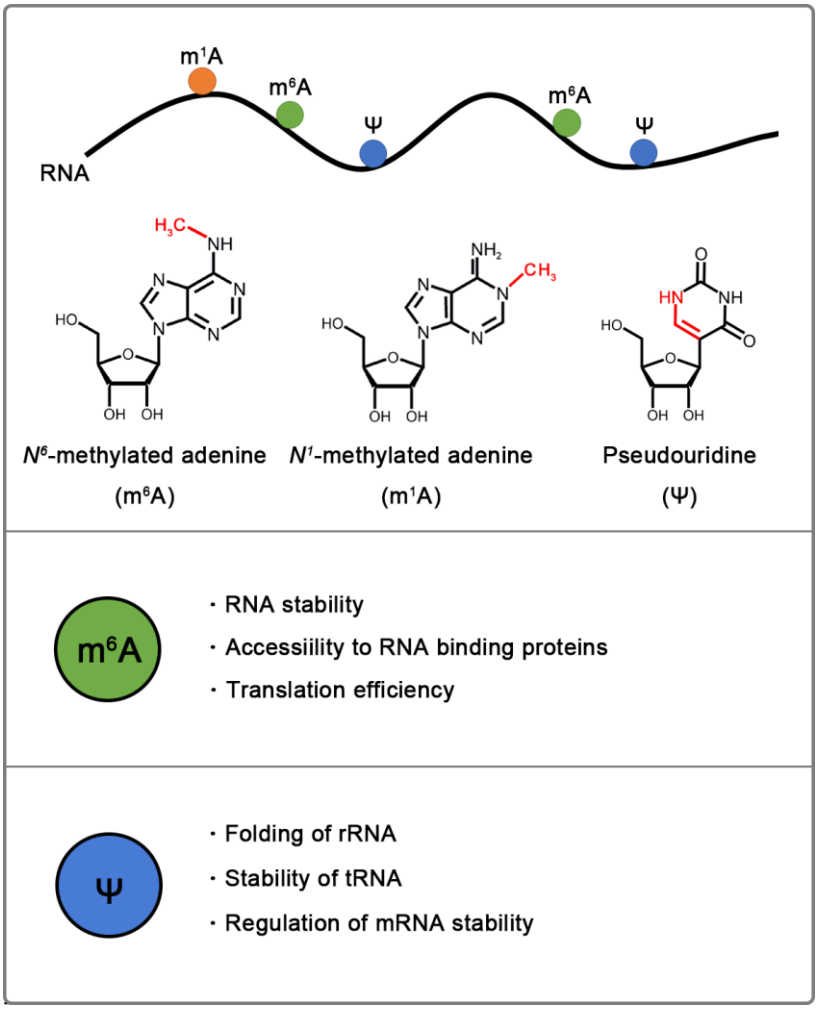
Contents of the thesis are published as Kim et al., Nucleic Acids Res. 2021 Jan 11;49(1):491-503.

RNA modification

RNA modifications play an important role in regulating the stability of various coding and non-coding RNAs, and influence gene expression (Roundtree et al., 2017; Frye et al., 2018). Across all organisms, there are presently more than 140 known post-transcriptional modifications in RNAs, including N^6-methylated adenine (m^6A), N^1-methylated adenine (m^1A) and pseudouridine (Ψ) (Fig. 1). Each RNA modification requires specific enzymes. Enzymes introducing and removing modifications are called writers and erasers, respectively, and the information stored in the modifications is interpreted via reader proteins (Roundtree et al., 2017; Frye et al., 2018; Zaccara et al., 2019). Of the RNA modifications, m^6A and Ψ, a C5-glycoside isomer of uridine (Fig. 1), are the most prevalent (Gilbert et al., 2016; Zaccara et al., 2019). m^6A occurs mainly in messenger RNAs (mRNAs) and regulates for example mRNAs stability, accessibility to RNA-binding proteins, and translation efficiency (Zaccara et al., 2019). By contrast, Ψ was first found in non-coding RNA, including transfer RNA (tRNA) and ribosomal RNA (rRNA) (Cohn, 1960), and only recently in mRNAs (Carlile et al., 2014; Schwartz et al., 2014). Recently, through the pseudouridine sequencing, more than 2000 Ψ sites were mapped in more than 1900 human transcripts (Li et al., 2015). In *Arabidopsis thaliana*, there are 187 Ψ sites in rRNAs, 232 Ψ sites in tRNAs, and more than 450 Ψ sites in mRNAs (Sun et al., 2019). Especially, 5' untranslated regions (UTRs) and coding sequences in *A. thaliana* mRNAs preferentially contain Ψs.

Figure 1. Types and functions of RNA modification

Representative RNA modifications and their chemical structures are shown. Currently, more than 140 distinct site-specific RNA modification have been discovered. Each modification affects a variety of molecular process.


m1A
m6A
Ψ
m6A
Ψ
RNA
N6-methylated adenine
(m6A)
N1-methylated adenine
(m1A)
Pseudouridine
(Ψ)
m6A
· RNA stability
· Accessiility to RNA binding proteins
· Translation efficiency
Ψ
· Folding of rRNA
· Stability of tRNA
· Regulation of mRNA stability

Chemical property and role of pseudouridine

Ψ was firstly discovered in RNAs in the 1950s and is considered as one of the most prevalent modification (~0.2-0.6% of uridine is modified into Ψ, Li et al., 2015). The Ψ is synthesized via base-specific isomerization by an enzyme called pseudouridine synthase. As a result of isomerization, the Ψ has chemical properties which differ from those of uridine (Adachi et al., 2019). In addition to the unusual carbon-carbon glycosidic bond between the uracil and ribosyl moiety, Ψ has an additional hydrogen bond donor (Fig1, marked in red). These chemical property impact on RNA stands, such as giving greater rigidity to the phosphodiester backbone of the RNA and/or stabilizing base pairs of Ψ-A (compared to those of U-A) (Penzo et al., 2017).

Ψ plays a role in the folding and translational accuracy of rRNAs, and it is also essential for the stabilization of tRNA structures (Charette and Gray, 2000; Helm, 2006). In mRNA from yeast and human cells, pseudouridylation is regulated in response to changing cellular environments (Carlile et al., 2014; Schwartz et al., 2014), and its presence alters translation (Eyler et al., 2019). In fact, an artificial conversion from U to Ψ in the termination codons of mRNA (UAA, UAG, and UGA) changes them into the missense codons (Karijolich and Yu., 2011). Especially, the rapid mRNA degradation induced by nonsense-meditated mRNA decay was perturbed by the U to Ψ conversion. In plant mRNA, the role of pseudouridylation is not yet clear.

Catabolism of canonical pyrimidine nucleoside

Canonical pyrimidine nucleosides, such as uridine and/or cytidine, are to some extent absorbed from the environment (Girke et al., 2014), but most are considered to be released from RNA degradation. And a majority of the RNA comes from the ribosomes (Floyd et al., 2015). In A. *thaliana*, most rRNA is degraded by RNase in the vacuole, and it is presumed that nucleotides produced by RNase are dephosphorylated by vacuolar phosphatases (Witte and Herde, 2020). The resulting product, nucleoside, is metabolized in two pathways (Fig. 2): salvage pathway and degradation pathway.

In the uridine catabolism, it undergoes a common pretreatment process before being metabolized by the above-mentioned pathways. The cytoplasmic uridine is hydrolyzed to uracil and ribose by nucleoside hydrolase 1 (NSH1) (Jung et al., 2009). Then, the uracil is reallocated by plastidic nucleobase transporter (PLUTO) from cytosol to plastid (Witz et al., 2012), where the fate of transported uracil is decided.

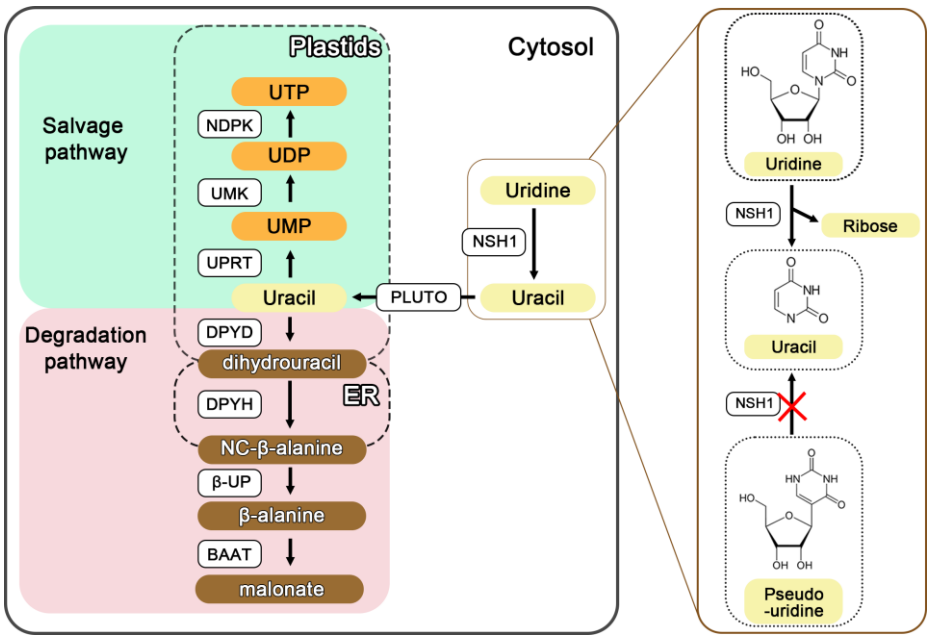
The uracil can be re-converted into nucleotide, which is named “salvage pathway” (Fig. 2, marked in green box). Uracil phosphoribosyltransferase (UPRT) catalyzes the anabolism of UMP using uracil and activated 5-phospho-α-D-ribose 1-diphosphate (PRPP) (Witte and Herde, 2020). the UMP is phosphorylated twice in succession by UMP kinase (UMK) and nucleoside diphosphate kinase (NDPK) to synthesize UTP in plastid (Zrenner et al., 2006).

Instead of being salvaged, the uracil can also be degraded, which is called “degradation pathway” (Fig. 2, marked in red box). In the first degradation reaction, dihydropyrimidine dehydrogenase (DPYD) catalyzes the reduction of uracil ring to dihydrouracil (Tintemann et al., 1985) and then the ring of dihydrouracil is opened by dihydropyrimidine hydrolase (DPYH),

generating N-carbamoyl-β-alanine (NC-β-alanine). Subsequently, the carbamino group of NC-β-alanine is released by β-ureidopropionase (β-UP) to yield β-alanine (β-Ala) (Cornelius et al., 2011). Recently, it has been suggested that β-Ala may be degraded by β-Ala aminotransferase (BAAT), but the enzyme has not yet been assigned (Wu et al., 2016).

Figure 2. metabolism of uridine

In *A. thaliana*, uridine is metabolized by two pathways; salvage and degradation. UTP, uridine triphosphate; UDP, uridine diphosphate; UMP, uridine monophosphate; NC-β-alanine, N-carbamoyl-β-alanine; NSH1, nucleoside hydrolase 1; PLUTO, plastidic nucleobase transporter; NDPK, nucleoside diphosphate kinase; UMK, UMP kinase; UPRT, uracil phosphoribosyltransferase; DPYD, dihydropyrimidine dehydrogenase; DPYH, dihydropyrimidine hydrolase; β-UP, β-ureidopropionase; BAAT, β-Ala aminotransferase.


Plastids
Cytosol
Salvage pathway
Degradation pathway
UTP
NDPK
UDP
UMK
UMP
UPRT
Uracil
PLUTO
Uridine
NSH1
Uracil
DPYD
dihydrouracil
DPYH
ER
NC-β-alanine
β-UP
β-alanine
BAAT
malonate
Uridine
NSH1
Ribose
Uracil
NSH1
Pseudo -uridine

Catabolism of non-canonical nucleotides

Compared to the knowledge regarding the biogenesis of RNA modifications and metabolism of canonical nucleotides, little is known about the metabolic fate of non-canonical nucleotides derived from the degradation of modified RNAs. Metabolic turnover of m^6A- or Ψ-containing RNAs produces the non-canonical nucleotides N^6-methyl-adenosine monophosphate (N^6-mAMP) or pseudouridine monophosphate (ΨMP) as products. Since the non-canonical nucleotide has a chemically different base from that of canonical nucleotide, it cannot be used as a substrate by enzymes to metabolize canonical nucleotide. Therefore, eukaryotes have pathways for the degradation of these modified nucleotides. Some of the pathways and enzymes that perform the role have recently been discovered. (Chen et al., 2018; Chen and Witte, 2020). In *Arabidopsis* plants and human cells, N^6-mAMP deaminase hydrolyzes N^6-mAMP, generating inosine monophosphate (IMP) (Chen et al., 2018), which is an intermediate of either purine nucleotide biosynthesis or catabolism (Baccolini and Witte, 2019; Witte and Herde, 2020). The N^6-mAMP deaminase reaction thus converts a non-canonical nucleotide into a nucleotide occurring in general nucleotide metabolism. The elucidation of the structure of N^6-mAMP deaminase from Arabidopsis showed that this enzyme undergoes a ligand-induced conformational change and identified the amino acids in the active site that mediate the substrate specificity (Jia and Xie, 2019; Wu et al., 2019). The discovery of the *in-vivo* role of N^6-mAMP deaminase has shown for the first time that eukaryotic cells can catabolize non-canonical nucleotides released from modified nucleic acids.

The catabolic pathway for ΨMP has recently been described in *Arabidopsis* (Chen and Witte, 2020). In this plant, a significant proportion of RNA degradation seems to occur in the vacuole (Floyd et al., 2015; Hickl et

al., 2019), where breakdown of Ψ-containing RNA generates ΨMP (presumably 3'-ΨMP), which is then dephosphorylated to pseudouridine and exported to the cytosol (Chen and Witte, 2020). For degradation, pseudouridine must reach the peroxisome. The catabolic pathway for pseudouridine in the peroxisome consists of two enzymes: PSEUDOURIDINE KINASE (PUKI), which phosphorylates pseudouridine to 5'-ΨMP, and PSEUDOURIDINE MONOPHOSPHATE GLYCOSYLASE (PUMY), which hydrolyzes 5'-ΨMP producing uracil and ribose 5-phosphate (Fig. 3) (Chen and Witte, 2020). Both reaction products are intermediates of main metabolic pathways. Uracil may be reincorporated into uridine monophosphate in a salvage pathway or may enter degradation pathway (Fig. 2) (Loh et al., 2006; Zrenner et al., 2009). Ribose 5-phosphate may be activated to PRPP needed for phosphoribosyl transfer reactions or may enter the pentose phosphate pathway. PUKI and PUMY are required for the removal of pseudouridine. Malfunction of these enzymes causes massive accumulation of pseudouridine, resulting in spurious formation of 5'-ΨMP catalyzed by cytosolic UMP kinases. Cytosolic 5'-ΨMP appears to be particularly toxic and causes delayed seed germination and growth inhibition (Chen and Witte, 2020).

Degradation of ΨMP was initially observed for pyrimidine auxotrophic *Escherichia coli* mutants (Breitman, 1970). Later, detailed investigations of these mutants led to the discovery of PUKI and PUMY, called YeiC and YeiN in *E. coli* (Preumont et al., 2008). Bioinformatic analyses suggested that PUKI and PUMY are present in many organisms including eukaryotes, but not in mammals (Preumont et al., 2008; Reumann, 2011; Chen and Witte, 2020). Interestingly, the kinase and the glycosidase reside on a single polypeptide in animals, but these putatively bifunctional enzymes have not

been characterized so far (Chen and Witte, 2020). PUKI is a member of the phosphofructokinase B (PfkB) family of carbohydrate kinases (Park and Gupta, 2008; Chen and Witte, 2020). PUKI from Arabidopsis (AtPUKI; encoded at the locus At1g49350) has putative orthologs of highly similar sequence in other plants (Fig. 4) (Chen and Witte, 2020). PUKI sequences formed a well-supported clade when all PfkB kinases from five evolutionary distant plant species were phylogenetically analyzed (Schroeder et al., 2018). This indicates that PUKI has conserved sequence elements of functional relevance that are distinct from other PfkB family kinases. One intriguing biochemical observation for AtPUKI is its high specificity towards the non-canonical nucleoside pseudouridine (Chen and Witte, 2020), leading to the question of how AtPUKI discriminates its authentic substrate from other chemically and/or structurally similar pyrimidine nucleosides.

In this thesis, I determined crystal structures of AtPUKI in the absence of a ligand, its binary complex with adenosine triphosphate (ATP), and a ternary complex with pseudouridine and adenosine diphosphate (ADP). I also carried out mutational and kinetic analyses to investigate the substrate specificity of the enzyme. The results provide a structural and catalytic rationale for the high preference of AtPUKI for the non-canonical nucleoside pseudouridine.

Figure 3. Reaction scheme of pseudouridine catabolism

Pseudouridine degradation requires two enzymatic reactions occurring in the peroxisome. PUKI phosphorylates pseudouridine with ATP as a phosphate donor to 5'-ΨMP, which is hydrolyzed into uracil and ribose 5-phosphate by PUMY.

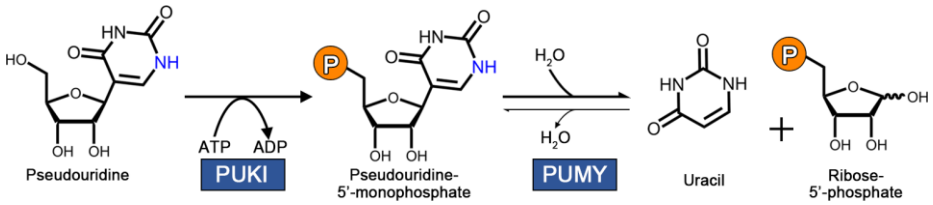
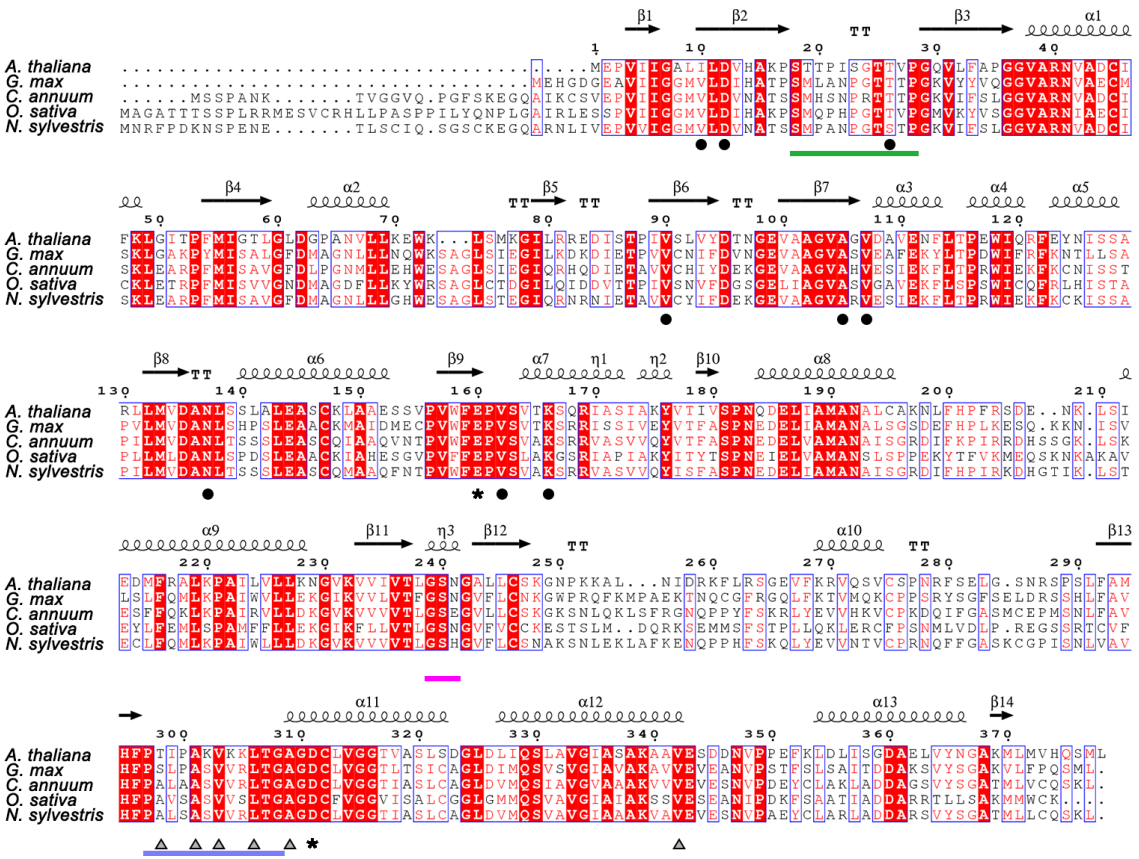

HO
HN
O
O
NH
O
OH OH
Pseudouridine
ATP ADP
PUKI
P
HN
O
O
NH
O
OH OH
Pseudouridine-
5'-monophosphate
H2O
H2O
PUMY
HN
O
O
NH
Uracil
P
O
OH
OH OH
Ribose-
5'-phosphate

Figure 4. Sequence comparison of PUKI from different sources

The amino acid sequence of AtPUKI (At1g49350) is compared to that of other PUKI homologs from soybean (Glycine max), pepper (Capsicum annuum) rice (Oryza sativa), and a tobacco species (Nicotiana sylvestris). Highly conserved residues are shown in red and are boxed in blue, while strictly conserved residues are shown with a red background. Secondary structural elements defined in unliganded AtPUKI are shown for the corresponding sequences. Residues or structural elements are indicated using the following color code: residues in the substrate pocket for pseudouridine, black circles; ADP-binding residues, gray triangles; putative catalytic residues, black asterisks; the nucleoside substrate-binding loop, green bar; the small ATP-binding loop (Gly239–Asn241), magenta bar; the large ATP-binding loop (Pro297–Gly308), light blue bar. This figure was prepared using ESPript (Robert and Gouet, 2014).


A. thaliana
G. max
C. annuum
O. sativa
N. sylvestris

Materials and Methods

Cloning and purification of AtPUKI

A synthetic cDNA of AtPUKI (At1g49350) (Bioneer, Daejeon, South Korea) codon-optimized for expression in *E. coli* was amplified by PCR. The resulting PCR product was cloned into a modified pET-28a vector (Merck, Kenilworth, NJ, USA) containing a tobacco etch virus (TEV) protease cleavage site at the junction between the coding sequence for an N-terminal $(His)_6$-maltose binding protein (MBP) and the multiple cloning site. Later, I found that the $(His)_6$-MBP tag could not be cleaved by the TEV protease during purification. For a facile cleavage of the tag, which is required for crystallization, additional linkers coding for Gly-Gly-Gly-Ser (GGGS) were inserted, resulting in the N-terminal region of the AtPUKI fusion protein containing the sequence $(H)_6$–MBP–GGGS–GGGS–ENLYFQS (TEV protease cleavage site)–GGGS (Fig. 5). *E. coli* BL21 (DE3) cells (Merck) transformed with the resulting construct were cultured at 37°C in Luria-Bertani medium until the absorbance at 600 nm reached 0.6. Expression of AtPUKI was induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside, followed by incubation at 12°C for 40 h. Cells were collected by centrifugation and sonicated in buffer A (50 mM Tris-HCl at pH 8.0, 300 mM NaCl, 2 mM dithiothreitol, and 5% [w/v] glycerol), and the supernatant was obtained by centrifugation at 30,000 × g for 1 h at 4°C. The fusion protein was purified using a BabyBio Ni-NTA column (Bio-Works Technologies AB, Uppsala, Sweden) equilibrated with buffer A, and was then eluted with buffer A containing an additional 0.5 M imidazole. The $(His)_6$–MBP tag of AtPUKI was subsequently removed by a TEV protease treatment overnight at 4°C using a 20:1 molar ratio of AtPUKI to protease. The N-terminal His-tagged fragment and the TEV protease were removed by a BabyBio Ni-NTA column, and the tag-free AtPUKI was further purified by size-exclusion

chromatography using a Superdex 200 column (GE Healthcare, Chicago, IL, USA) with buffer A (Fig. 6). The purify in each purification steps was confirmed by SDS-PAGE (Fig. 7).

For structural analyses of AtPUKI, seleno-L-methionine (SeMet)-substituted AtPUKI was expressed in *E. coli* BL21 (DE3) in M9 minimal medium containing the required amino acids according to a published protocol (Van Duyne et al., 1993). The purification of SeMet-substituted AtPUKI was carried out as described before for the native AtPUKI.

For further functional and structural analysis, various AtPUKI mutants were produced by site-directed mutagenesis using the QuikChange method (Agilent) (Table 1). To purify the various mutants of AtPUKI, I followed purification procedures identical to those described above.

Figure 5. scheme of the construct for AtPUKI purification

For a facile cleavage of the tag, additional linkers (i.e., Gly-Gly-Gly-Ser (GGGS)) were inserted around TEV protease cleavage site. The resulting construct contains $(His)_6$–MBP–GGGS–GGGS–ENLYFQS (TEV protease cleavage site)–GGGS in the N-terminal region of the AtPUKI recombinant protein.

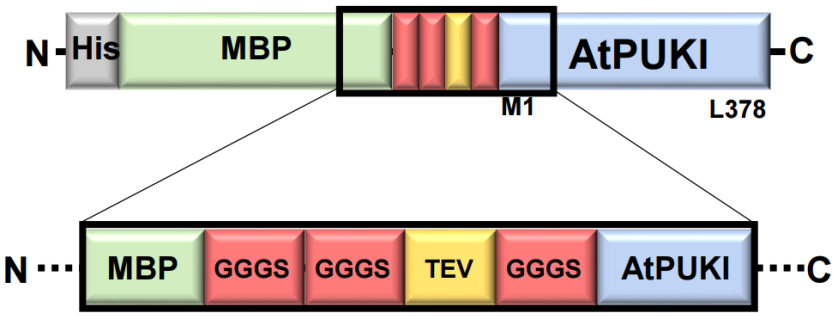

N
His
MBP
AtPUKI
C
M1
L378
N
MBP
GGGS
GGGS
TEV
GGGS
AtPUKI
C

Figure 6. Purification profiles of AtPUKI

AtPUKI was purified using immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC). The *blue line* and *green line* represent protein absorbance at 280nm and concentration of imidazole (%), respectively. (A) In the 1st IMAC, N-terminal His-tag containing AtPUKI was purified. (B) After TEV protease treatment for 14-16 h at 4℃, 2nd IMAC was performed. The N-terminal MBP containing His-tag fragment and the TEV protease were removed by a BabyBio Ni-NTA column. (C) The tag-free AtPUKI was further purified by SEC using a Superdex 200 column.

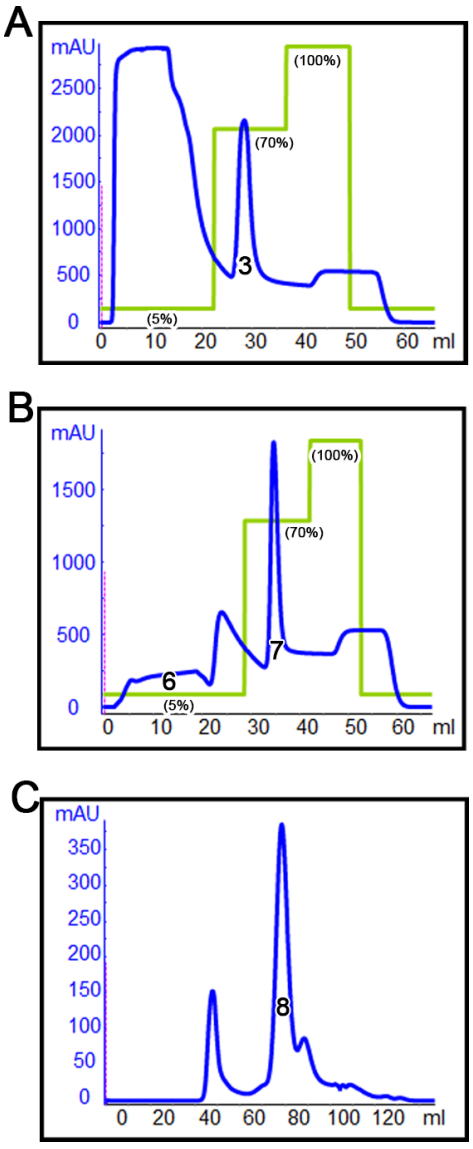
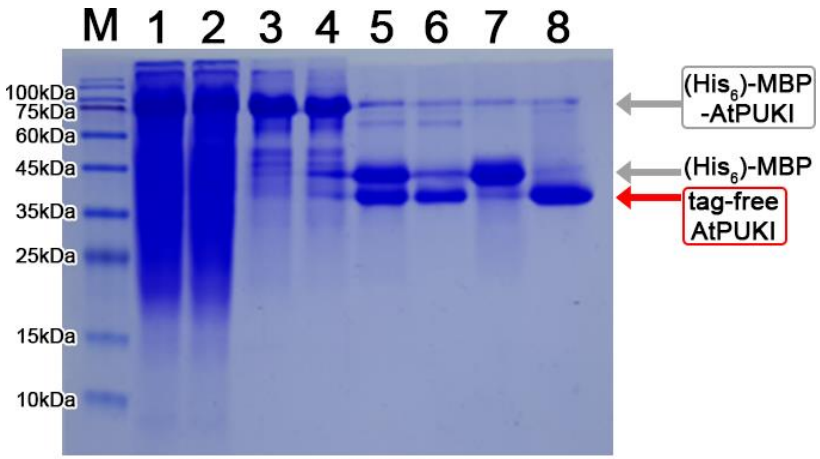

A
mAU
2500
2000
1500
1000
500
0
(100%)
(70%)
(5%)
3
0 10 20 30 40 50 60 ml
B
mAU
1500
1000
500
0
(100%)
(70%)
(5%)
6
7
0 10 20 30 40 50 60 ml
C
mAU
350
300
250
200
150
100
50
0
8
0 20 40 60 80 100 120 ml

Figure 7. SDS-PAGE analysis of AtPUKI

AtPUKI was purified with several purification steps. Each purified sample was analysed by 15% SDS-PAGE gel. On the SDS-PAGE gel, each lane indicates the purify of AtPUKI; lane M, protein marker; lane 1, total protein after cell lysis; lane 2, supernatant after centrifugation; lane 3, after 1st IMAC; lane 4, Immediately after TEV protease treatment; lane 5, 14-16h after TEV protease treatment; lane 6, unbound fraction in the 2nd IMAC; lane7, bound fraction in the 2nd IMAC; lane8, SEC.



4 : immediately after TEV protease treatment
5 : TEV protease treatment for 14–16h at 4℃

Table 1. List of AtPUKI mutants used in this study

	Residue number	Amino acid	Mutated amino acid	Nomen-clature	Role of the residue
1	10	Ile	Ala	I10A	Pseudouridine interaction
2 3	12	Asp	Ala Asn	D12A D12N	Pseudouridine interaction
4 5 6	26	Thr	Ala Val Ser	T26A T26V T26S	Pseudouridine interaction
7	137	Asn	Ala	N137A	Pseudouridine interaction
8 9	160	Glu	Ala Gln	E160A E160Q	Catalytic residue
10	166	Lys	Ala	K166A	Pseudouridine interaction
11 12	298	Thr	Ala Val	T298A T298V	ATP interaction
13	303	Val	Ala	V303A	ATP interaction
14	306	Leu	Ala	L306A	ATP interaction
15	311	Glu	Ala	D311A	Catalytic residue
16	342	Val	Ala	V342A	ATP interaction

Crystallization and structure determination

Initial crystallization was carried out by the sitting-drop vapor-diffusion method at 22°C using native AtPUKI (8 mg/mL) with The JCSG Core I–IV Suites, pH Clear Suite and Protein Complex Suite (Qiagen). In initial crystallization screening, crystals of AtPUKI were obtained with reservoir solution No. 8 of The JCSG Core I Suite consisting of 0.1 M tri-sodium citrate and 20% PEG3350. The crystallization conditions were optimized by slightly changing the precipitant concentration. Crystals of both native and SeMet-substituted AtPUKI were reproduced in crystallization conditions containing 0.1 M tri-sodium citrate and 22% PEG3350 (Fig. 8). Subsequently, co-crystallization of a binary com plex with 2 mM ATP and 4 mM $MgCl_2$, and a ternary complex with 4 mM ADP and 4 mM pseudouridine were also achieved under crystallization conditions identical to those for native AtPUKI. Prior to data collection, crystals for the binary and ternary complexes were further soaked in the crystallization solution plus 2 mM ATP and 4 mM $MgCl_2$ for the binary complex or 4 mM ADP and 4 mM pseudouridine for the ternary complex. In all cases, 20% glycerol was used as a cryoprotectant for data collection at 100 K.

Diffraction data were collected with a 0.5° oscillation angle at beamline 7A of the Pohang Accelerator Laboratory (Pohang, South Korea). The diffraction images were indexed, integrated, and scaled with HKL2000 (Otwinowski and Minor, 1997), and a $CC_{1/2}$ statistical value of approximately 0.5 was used for the high-resolution cut-off (Karplus and Diederichs 2012; Diederichs and Karplus, 2013). All crystals belong to the space group $P2_1$, with two monomers in the asymmetric unit. Specifically, sub-structure determination and the initial phase calculations for the structure of SeMet-substituted AtPUKI were carried out with the single-wavelength anomalous

dispersion method using the program PHENIX (Adams et al., 2010). The refined model of the SeMet-substituted AtPUKI structure was used as search model for a molecular replacement of its binary and ternary complexes, which was conducted using the PHENIX AutoMR program (Adams et al., 2010). The models were manually rebuilt by COOT (Emsley et al., 2010) and refined with PHENIX (Adams et al., 2010). After several cycles of refinement, most residues of AtPUKI were located from Met1 to Leu378, except for highly disordered residues, particularly between Phe200 and Pro300. Details on data collection and refinement statistics are shown in Table 2.

Figure 8. Crystal of AtPUKI

Crystals were obtained under a crystallization solution of 0.1 M tri-sodium citrate and 22% PEG3350

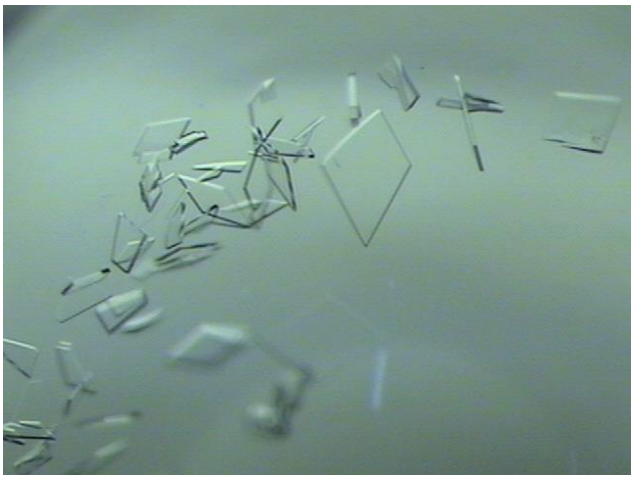


Table 2. Data collection and refinement statistics

Data set	Unliganded AtPUKI	Ternary complex pseudouridine-ADP	Binary complex ATP
PDB ID	**7C1X**	**7C1Y**	**7C1Z**
Data collection			
Crystal	SeMet	Native	Native
Wavelength (Å)	0.97935	0.97935	0.97935
Resolution (Å)	50.0-2.4 (2.49-2.40) [a]	50.0-2.1 (2.18-2.10) [a]	50.0-2.1 (2.18-2.10) [a]
Unique reflections	7,873	11,095	11,212
Multiplicity	3.4 (3.4)	3.6 (3.7)	3.5 (3.5)
Completeness (%)	99.2 (99.3)	99.6 (100.0)	99.5 (99.7)
Mean I/sigma(I)	13.4 (1.7)	13.1 (1.1)	15.3 (1.4)
Wilson *B*-factors (Å^2)	48.3	43.1	36.4
$CC_{1/2}$ [b]	0.97 (0.73)	0.98 (0.52)	0.99 (0.61)
Space group	*P2$_1$*	*P2$_1$*	*P2$_1$*
Unit cell *a*, *b*, *c* (Å)	80.7, 48.4, 92.8	79.3, 48.9, 91.8	80.6, 47.9, 91.9
α , β , γ (º)	90.0, 109.1, 90.0	90.0, 107.5, 90.0	90.0 109.0 90.0
Refinement			
R-work [c] (%)	20.8	19.1	18.6
R-free [d] (%)	24.2	23.5	23.7
No. of atoms	5225	5380	5465
Macromolecules	5198	5209	5290
Ligands	-	171	66
Water	27	83	109
RMS(bonds) (Å)	0.003	0.008	0.008
RMS(angles) (º)	0.89	1.23	1.25
Ramachandran			
favored (%)	97.4	98.0	96.7
outliers (%)	0.2	1.9	0.0
Average *B*-factor (Å^2)	57.8	54.7	45.2
Macromolecules	57.8	54.8	45.0
Ligands	-	49.5	60.9
Water	57.4	49.0	47.6

[a]Numbers in parentheses refer to data in the highest resolution shell.

[b]The $CC_{1/2}$ is the Pearson correlation coefficient (CC) calculated from each subset containing a random half of the measurements of unique reflection

[c]$R_{work} = \Sigma \, ||F_{obs}|-|F_{cal}|| / \Sigma |F_{obs}|$

[d]R_{free} is the same as R_{obs} for a selected subset (5%) of the reflections that was not included in prior refinement calculations.

Activity assays

I employed two different assay protocols for AtPUKI: an enzyme-coupled assay (Mallette and Kimber, 2018; Kang et al., 2019) and a direct assay (Andersson and Mowbray, 2002) not requiring coupling enzymes. The wild-type and various mutants of AtPUKI were purified as described above.

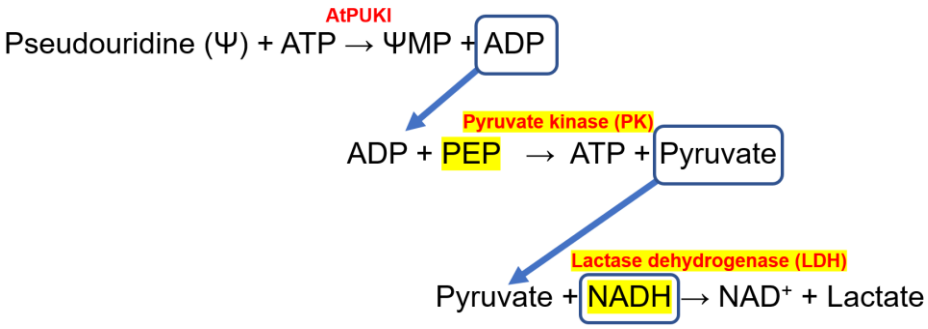
Given that the enzyme-coupled assay requires K^+ and Mg^{2+} ions as co-factors for one of the coupling enzymes (pyruvate kinase, see below), a direct assay protocol was also needed to determine the effects of different monovalent and divalent cations on AtPUKI activity, and to determine the substrate preferences of AtPUKI toward various nucleotide triphosphates including ATP, guanosine triphosphate (GTP), cytidine triphosphate (CTP), and uridine triphosphate (UTP). Typically, the assay mixture included 20 mM divalent cation, 50 mM monovalent cation, 0.003% phenol red pH 8.1, 4 mM ATP, and the wild-type AtPUKI. After incubation at 25°C for 2 min, 1.25 mM pseudouridine was added to initiate the enzyme reaction, and changes in the absorbance at 430 nm were monitored for 30 s. For the monovalent cation dependency of AtPUKI, its activity was measured in the absence or presence of 50 mM LiCl, NaCl, KCl, RbCl, or CsCl. For the effects of divalent cations, the AtPUKI activity was also assayed in the absence or presence of 20 mM $MgCl_2$, $MnCl_2$, or $CaCl_2$. For the substrate preference regarding the phosphate donor, the direct assay was performed in the presence of 1 mM ATP, GTP, CTP, or UTP. All assays were conducted in triplicate.

The enzyme-coupled assay was employed for steady-state kinetic measurements using K^+ and Mg^{2+} as mono- and di-valent cations, because these were the most stimulating ions for AtPUKI activity in the direct assay. Briefly, ADP released from an AtPUKI-dependent reaction was detected in the coupled assay, with concurrent consumption of NADH, via two sequential

reactions by pyruvate kinase and the NADH-dependent lactate dehydrogenase (Fig. 9). Specifically, the pre-reaction mixture consisted of 40 mM Tris HCl pH 7.5, 20 mM $MgCl_2$, 50 mM KCl, 0.2 mM NADH, 2 mM phosphoenolpyruvate, indicated concentrations of AtPUKI, 18–28 units of lactate dehydrogenase and 12–20 units of pyruvate kinase per mL, and one of the substrates (either pseudouridine or ATP). Following incubation at 25°C for 2 min, the other substrate was added to initiate the enzyme reaction. Then I monitored the decrease in absorbance at 340 nm using a UV-visible light spectrophotometer (Jasco, Tokyo, Japan). I verified that the concentrations of the coupling enzymes were saturating in assay and that the initial velocity was dependent on the concentration of AtPUKI. The initial velocity was determined between 15 and 45 s and expressed as the corresponding ADP concentration change per min, using an extinction coefficient of 6,220 $M^{-1}cm^{-1}$ at 340 nm for NADH. Sigmaplot (Systat Software, San Jose, CA, USA) was used to calculate the K_m and V_{max}.

Figure 9. Scheme of enzyme-coupled assay for steady-state kinetic assay of AtPUKI

In the coupled assay, ADP released from an AtPUKI-dependent reaction was detected, with concurrent consumption of NADH, via two sequential reactions by pyruvate kinase and the NADH-dependent lactate dehydrogenase. Additional components for enzyme-coupled assay are indicated by yellow boxes.


AtPUKI
Pseudouridine (Ψ) + ATP → ΨMP + ADP
Pyruvate kinase (PK)
ADP + PEP → ATP + Pyruvate
Lactase dehydrogenase (LDH)
Pyruvate + NADH → NAD+ + Lactate

RESULTS AND DISCUSSION

Structure of unliganded AtPUKI

The calculated molecular weight of monomeric AtPUKI is 40.3 kDa. The native protein had a molecular weight of approximately 80 kDa determined by size-exclusion chromatography during protein purification, indicating that it is a dimer (Fig. 10). Consistent with the chromatographic result for the enzyme in solution, crystalline AtPUKI was also a dimer in the asymmetric unit, with a buried surface area of 2420 $Å^2$ calculated from the PISA analyses (Krissinel and Henrick, 2007). Therefore, the dimer observed in the crystal represents the biologically relevant dimeric state.

Monomeric AtPUKI (Met1–Leu378) is folded into an α/β domain, with a protruding small β-stranded domain (Figs. 4 and 11A). The α/β domain is reminiscent of a Rossmann fold employing the β–α unit as a basic structural motif. In the α/β domain, juxtaposition of the β–α units results in a central β-sheet, with seven α-helices (α3–α9) on the convex side of the β-sheet and the remaining α-helices (α1, α2, and α10–α12) on the other side. These structural arrangements result in a crevice along the central β-sheet (Fig. 11B, area in magenta). Specifically, the central β-sheet contains 10 β-strands that are positioned in the order of β5-β4-β1-β8-β9-β10-β11-β12-β13-β14 with a parallel orientation, except for β12 and β14. Among the β–α units, the β1–α1 and β5–α3 units on the edge of the central β-sheet contain ~30-residue-long insertions in the loop region that are folded into β-strands: β2 and β3 in β1–α1, and β6 and β7 in β5–α3 (Figs. 4 and 11A). The two consecutive β-strands in each insertion are folded in an antiparallel manner and protrude from the main domain. In the small domain, the antiparallel β-strands from the two β–α units are packed in an edge-to-edge orientation, resulting in a continuous four β-stranded domain (Fig. 11A). The small domain is bent over the crevice on the top of the central β-sheet, leaving a cleft just underneath the small domain.

Dimerization of AtPUKI is mediated mainly by two elements: the small β-strand domain (Fig. 11B) and α2 on the edge of the central β-sheet (Fig. 11A). The face-to-face orientation with an almost orthogonal manner of the four β-strands in the small domains from both subunits results in a β-clasp motif for dimerization (Sigrell et al., 1998). In the β-clasp motif, there are two loops protruding from each subunit that are present in β2–β3 and β6–β7. These two loops are oriented in a two-fold symmetric way with the two equivalent loops from the adjacent subunit (Fig. 11B, zoom-in view). In detail, the loop present in the β2–β3 strand of one subunit crosses over the β-clasp motif and interacts with the loop in the β6*–β7* strand of the adjacent subunit (*indicates residues or elements from the adjacent monomer) and vice versa. Later I found that the loop in β2–β3 (i.e., Ser18–Pro28) is involved in the active site of the adjacent subunit (see below). Therefore, I refer to this loop as the substrate-binding loop in the following.

The two subunits of the dimer adopt different conformations. In one monomer, referred to as subunit-A, the area of the crevice along the central β-sheet is less exposed compared to the other monomer (i.e., subunit-B) (Fig. 11B). These features originate from differences in the bending angle of the small β-strand domains toward the crevice on the large α/β domain (Fig. 12). Structural superposition showed that the large α/β domains of the two subunits maintain an essentially identical conformation, with a root-mean-square-deviation (RMSD) of 0.63 Å for the 298 Cα atoms in the large domain. However, due to differences in orientation, the small domains of both subunits showed large positional deviations, up to 10 Å, for their corresponding Cα atoms (Fig. 12). In fact, the small domain in subunit-A, specifically the β6 and β7 regions, is bent by 15 degrees more compared to that in subunit-B, leaving the crevice in subunit-A less accessible from the surface. Accordingly, the

inter-atomic Cα distance between Ala101 in the small domain β7 region and Ser291, the residue most distant from the small domain, is different for the respective subunits: ~36 Å for subunit-A and ~43 Å for subunit-B. This asymmetry indicates that the small domains, particularly the β6 and β7 regions, have dynamic features in AtPUKI.

Figure 10. Determination of the molecular weight of soluble AtPUKI

Size-exclusion chromatographic analysis of AtPUKI after removal of the $(His)_6$–MBP tag was performed on a Superdex 200 column (GE Healthcare) with a buffer containing 50 mM phosphate (pH 7.4) and 150 mM NaCl. The peak of AtPUKI had an elution volume corresponding to a molecular weight of approximately 80 kDa, indicating that AtPUKI is a dimer in solution. Insert, molecular mass markers (GE Healthcare): aprotinin (6.5 kDa), ribonuclease (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumin (44 kDa), conalbumin (75 kDa) and aldolase (158 kDa).

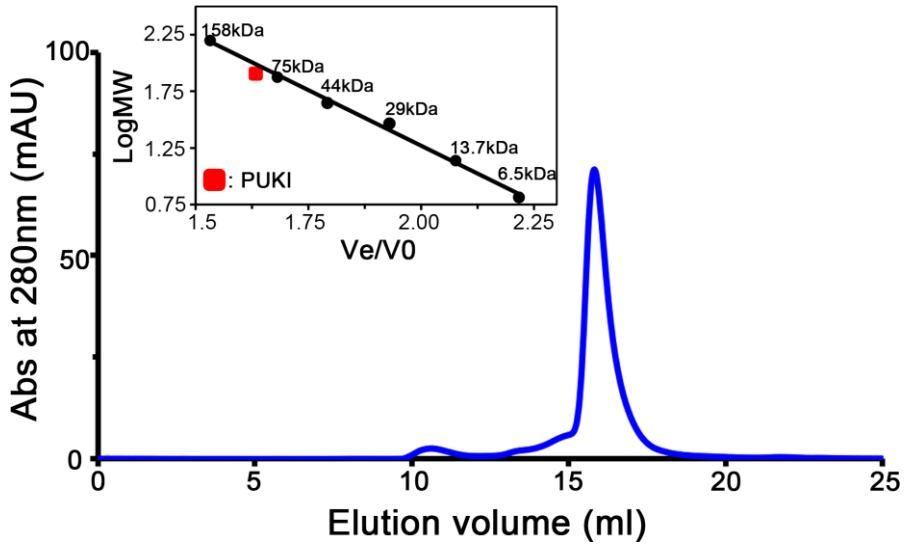

Abs at 280nm (mAU)
100
50
0
0
5
10
15
20
25
Elution volume (ml)
LogMW
2.25
1.75
1.25
0.75
1.5
1.75
2.00
2.25
Ve/V0
158kDa
75kDa
44kDa
29kDa
13.7kDa
6.5kDa
: PUKI

Figure 11. Monomeric and dimeric structures of AtPUKI

(A) The monomeric structure of unliganded AtPUKI is shown in a side view along the central β-sheet (yellow), with the protruding small β-stranded domain on the top (yellow). The monovalent ion binding site is indicated with a green circle containing the letter M. (B) Top view of dimeric AtPUKI is displayed in surface representation, with a zoom-in view of the β-clasp motif for dimerization. Rotation of the side view in (A) horizontally by 90° towards the reader would orient AtPUKI almost in the position of subunit-B shown in yellow. The surface in magenta forms a crevice comprising the binding sites for pseudouridine and ADP (see the characterization of the ternary complex). In subunit-A, the crevice surface is more extensive than in subunit-B. In a zoom-in view, the substrate-binding loops between β2 and β3 from both subunits are labeled and positioned almost in a two-fold symmetric manner around the perpendicular two-fold axis running through the center, as indicated by the dot.

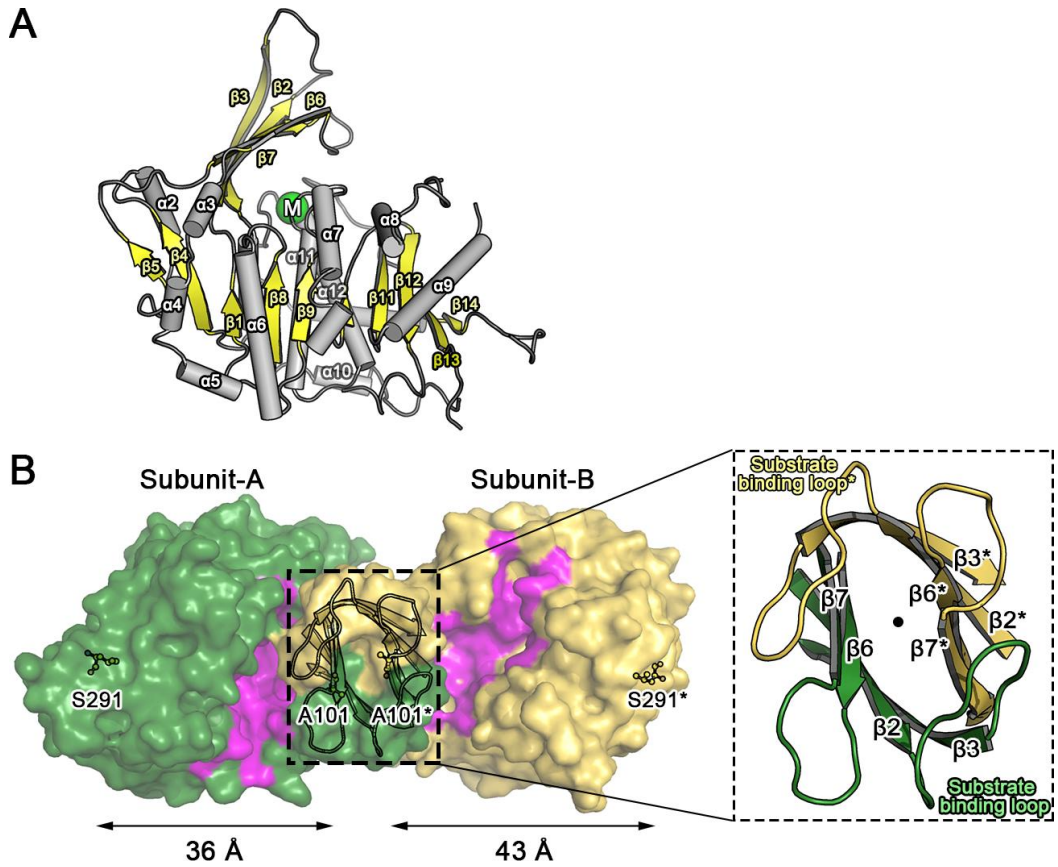
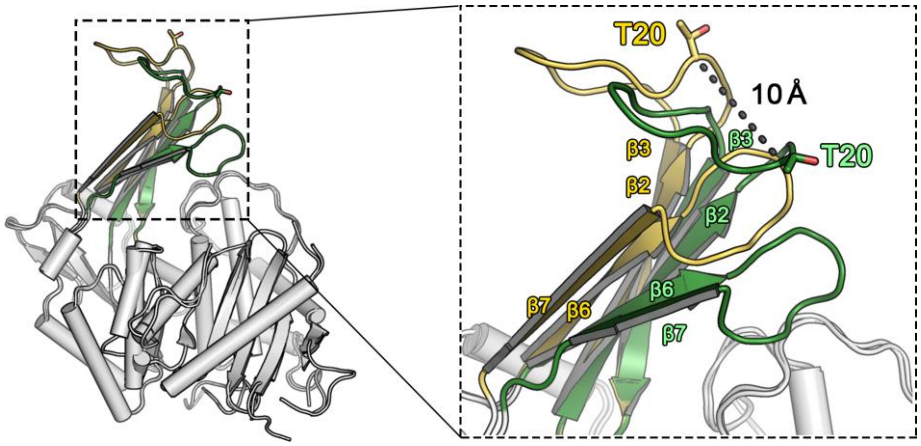

A
β3
β2
β6
β7
α2
α3
M
α7
α8
α11
β5
β4
α12
β12
α9
α4
β8
β11
β14
β1
α6
β9
β13
α5
α10
B
Subunit-A
Subunit-B
S291
A101
A101*
S291*
36 Å
43 Å
Substrate binding loop*
β3*
β7
β6*
β2*
β6
β7*
β2
β3
Substrate binding loop

Figure 12. Structural comparison of the two subunits of the unliganded AtPUKI

The two monomers of the dimeric AtPUKI in the unliganded form were superimposed based on the large α/β domain, with an RMSD of 0.63 Å for 298 Cα atoms, and are displayed in a side view with an orientation similar to Figure 11(A). In a zoom-in view, the small domains from both monomers show large positional deviations of as much as 10 Å for the Cα atoms of Thr20 in the substrate-binding loop. Specifically, the β6 and β7 regions of the small domain in subunit-A (green) exhibit a higher bending angle toward the cavity on the central β-sheet than the corresponding regions of subunit-B (yellow). The bending angle is calculated based on the relative position of β6 in each subunit.


T20
10Å
β3
β3
T20
β2
β2
β6
β7
β6
β7

Structural homologs of AtPUKI and the structure of the monovalent cation-binding site

Sequence analyses indicated that AtPUKI is a member of the PfkB family (Park and Gupta, 2008). Accordingly, structural features of monomeric and dimeric AtPUKI are also highly homologous to those of this protein family. A structure similarity search using the program DALI (Holm and Laakso, 2016) indicated that the AtPUKI monomer exhibits high structural homology with the PfkB family kinases catalyzing the phosphorylation of ribose, other carbohydrates, and adenosine, with a Z-score > 20. Among those structures, *E. coli* ribokinase (EcRBSK; Protein Data Bank [PDB] ID, 1RK2) (Sigrell et al., 1999) is the closest homolog with a Z-score of 37.7, along with other ribokinases from *A. thaliana* (AtRBSK; PDB ID, 6ILR; Z-score, 35.0) (Kang et al., 2019) and *Vibrio cholerae* (PDB ID, 4XDA; Z-score, 34.8) (Paul et al., 2015). Human adenosine kinase (PDB ID, 1BX4; Z-score, 30.6) is also homologous to the AtPUKI monomer. Due to these structural similarities, I adopted the structural nomenclature from ribokinase, particularly for the ATP binding features (Kang et al., 2019; Sigrell et al., 1998).

A direct (*i.e.* coupling enzyme-independent) activity assay showed that AtPUKI reaches maximum activity in the presence of the monovalent cation K^+, the divalent cation Mg^{2+}, and with ATP as substrate (Fig. 13A–C). These functional features are essentially identical to those of ribokinase. Inorganic phosphate is not necessary for AtPUKI activity (Fig. 13D). This is different from ribokinase of *E. coli* and from human adenosine kinase, which require inorganic phosphate as an activator (Maj and Gupta, 2001; Park et al., 2004). The binding sites for the cations and ligands in ribokinases from various sources are well characterized (Sigrell et al., 1998; Paul et al., 2015; Kang et al., 2019). In AtPUKI, the monovalent cation-binding site was present at the

location corresponding to that in ribokinases. It is formed by two adjacent loops: one between α12 and α13, and the other preceding α11, which was defined as the large ATP loop in ribokinase for the ATP-binding element (Fig. 14). The identity of the metal in a crystal of AtPUKI was proposed based on metal analyses including inductively coupled plasma atomic emission spectroscopy (ICP-AES), and crystallographic refinements (Fig. 15). ICP analyses showed that throughout the purification procedure, AtPUKI has not remained complexed with the monovalent cation candidates, such as Li^+, Na^+, K^+, Rb^+ (Fig. 15A). The results imply the metal in a crystal of AtPUKI may be from the purification buffer or the crystallization solution. Also, the identity of the metal ion was crystallographically analysed (Fig. 15B). Based on coordination geometries and putative metal-to-ligand distances, both Na^+ and K^+ are potential participants (see panels *a* and *b* for sodium and potassium ions, respectively). However, refined B-factors and the *Fo-Fc* map of the metal ions indicated that Na^+ is the most probable metal. Taken together, most likely a sodium ion occupies the position of the monovalent cation in the structure, although more direct evidence is required to fully ascertain the identity of the metal.

It is coordinated by the main chain carbonyl oxygens of Lys305, Thr307, Ala341, Ser344, and Asp346 (Fig. 14), and the coordination shell differs from the square bipyramidal shell in AtRBSK with six ligands (Kang et al., 2019). The binding environment of the divalent cation Mg^{2+} is unique to AtPUKI. It will be discussed below in the description of the binary complex with ATP (see the section "Mg^{2+}-binding site and catalysis in AtPUKI").

Figure 13. Effects of metal ions, different nucleoside triphosphates, and phosphate on AtPUKI activity

The relative activity depending on (A) monovalent cations (each at 50 mM), (B) divalent cations (each at 20 mM), and (C) different nucleoside triphosphates as phosphate donors (each at 1 mM) was assessed with a direct activity assay. (D) Using an enzyme-coupled assay, the effects of phosphate on AtPUKI activity were measured. The mean value of the highest measurement was set to 100% in each panel. Error bars are SD (n = 3). *n.d.*, not detected.

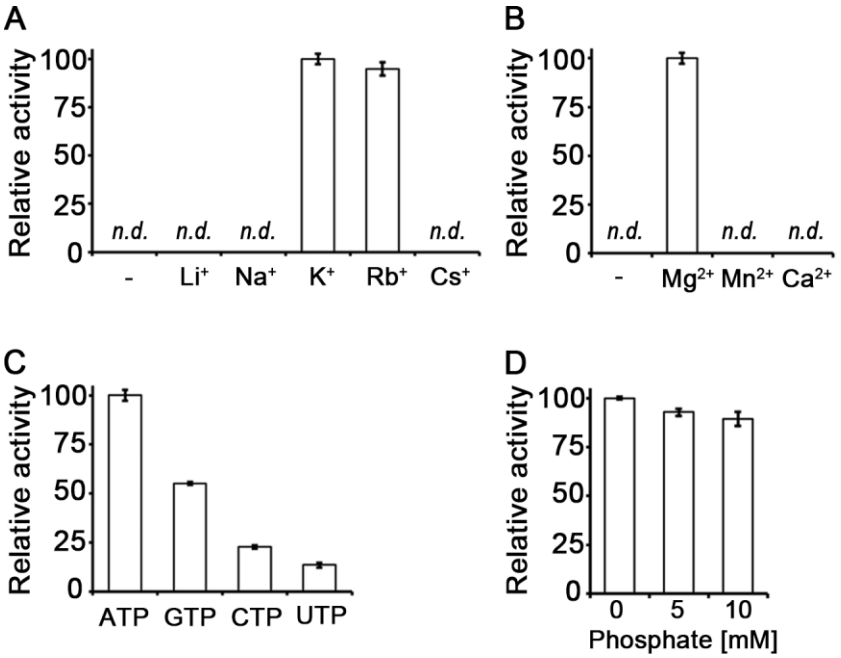

A
Relative activity
100
75
50
25
0
n.d.
n.d.
n.d.
n.d.
-
Li+
Na+
K+
Rb+
Cs+
B
Relative activity
100
75
50
25
0
n.d.
n.d.
n.d.
-
Mg2+
Mn2+
Ca2+
C
Relative activity
100
75
50
25
0
ATP
GTP
CTP
UTP
D
Relative activity
100
75
50
25
0
0
5
10
Phosphate [mM]

Figure 14. The monovalent cation-binding site in AtPUKI

The cation represented by the green sphere connects and stabilizes two loops: the loop between α12 and α13 (gray) and the large ATP-binding loop (light blue) preceding α11. The dashed lines represent coordination between the monovalent cation and the main chain carbonyl oxygen atoms of the indicated residues.

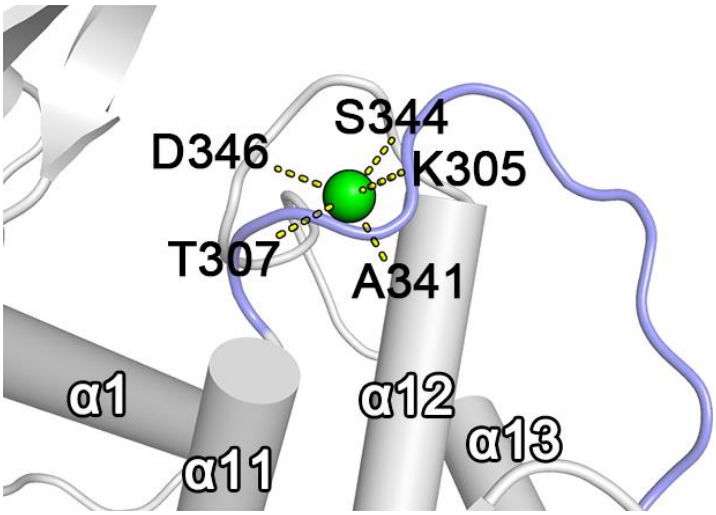

S344
D346
K305
T307
A341
α1
α12
α13
α11

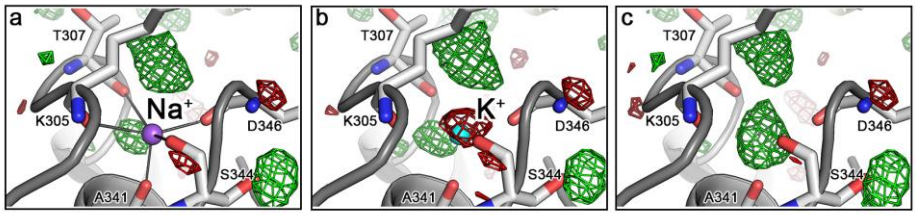
Figure 15. Metal analyses by inductively coupled plasma–atomic emission spectroscopy (ICP-AES) and crystallographic refinements

(A) Results from ICP-AES indicated that no monovalent cations from the four candidates, *i.e.*, Li^+, Na^+, K^+, Rb^+, remained complexed with AtPUKI throughout the purification procedure. For ICP-AES, AtPUKI was purified as described in the manuscript using buffer A (50 mM Tris-HCl at pH 8.0, 300 mM NaCl, 2 mM dithiothreitol, and 5% [w/v] glycerol). Then, AtPUKI in buffer A was dialyzed against a metal-free buffer (50 mM Tris-HCl at pH 8.0). After dialysis, AtPUKI was concentrated to ~1 mg/ml (1,000 ppm), and metals were analyzed in duplicate by ICP-AES (Agilent 5800 ICP-OES Spectrometer, Santa Clara, CA, USA). The sodium ion was the only metal detected in the dialysis buffer and AtPUKI, at nearly equal concentrations. (B) Crystallographic refinements in the ternary complex suggest that a sodium ion is located in the ATP-binding loop (Fig. 14). Green indicates a positive *Fo-Fc* map and red a negative *Fo-Fc* map; both maps are contoured at 3.0 sigma. In both cases, refinement of the proposed metal was performed with an occupancy of 1.0 for the metal ions when comparing refined B-factors.

A

	Li (ppm)	Na (ppm)	K (ppm)	Rb (ppm)
Buffer	*n.d.*	32.87	*n.d.*	*n.d.*
AtPUKI	*n.d.*	32.54	*n.d.*	*n.d.*

B



	Na^+	K^+	No metal
***R*-work**	18.6	18.6	18.7
***R*-free**	23.7	23.9	23.9
Average *B*-factors (Å^2) of the coordinating atoms			
AtPUKI	40.5	40.8	41.0
Metal	47.3	64.1	-

The ternary complex of AtPUKI with pseudouridine and ADP

I determined the crystal structure of AtPUKI complexed with pseudouridine and ADP. By co-crystallization and further soaking of the crystal with both ligands, I achieved unequivocal electron densities for pseudouridine and ADP (Fig. 16A). The structure of the ternary complex displays conformational differences compared to the unliganded AtPUKI (Fig. 17). However, there are virtually no changes in the tertiary structures of both subunits, with RMSD values of 0.67 Å for the 351 Cα atoms of subunit-A and of 0.48 Å for the 342 Cα atoms of subunit-B between the ternary complex and the unliganded AtPUKI. The changes occur mainly in the quaternary structure by a rigid-body rotation of one subunit. As a result, the inter-atomic Cα distances between reference points on both subunits are altered. For example, the distance between the two equivalent Pro202 residues in the two subunits is reduced by about 3 Å due to substrate binding (Fig. 17). In the ternary complex, the substrate pseudouridine and the product ADP are bound to both subunits.

In subunit-A, pseudouridine and ADP, which are separated by ~6 Å between the O5' of the pseudouridine ribosyl moiety and the β-phosphate atom of ADP (Fig. 16A), are bound in an extended mode along the crevice in the concave side of the central β-sheet. The small β-strand domain is located in an overhanging manner above the two ligands. Pseudouridine is located at the cleft underneath the small β-strand domain, with its planar nucleobase (hereafter referred to as uracil-Ψ) almost perpendicular to the ribose (Fig. 16A and B). Overall, pseudouridine is oriented such that the hydroxyl O5' of the ribosyl moiety is positioned toward the β-phosphate group of ADP, whereas its nucleobase points into the opposite direction and is closer to the dimerization interface. The uracil-Ψ ring is embedded into the pocket lined

with the hydrophobic residues Ile10, Val90, Ala105, Val107, and Val162, and the hydrophilic residues Thr26*, Asn137, and Lys166 (Fig. 16B). The extended side chain of Lys166 likely serves to limit the depth of the pocket. The hydrophobic residues in the pocket are in stacking orientations against the planar uracil-Ψ ring, with Val162 on one side of the ring and the remaining residues, mainly from β6 and β7 of the small domain, on the other side. The hydrophilic residues are located within 3.0 Å of uracil-Ψ ring. In particular, Asn137 mediates bidendate interactions with the N3 and O4 of the base (Fig. 16B), and Lys166 interacts with the O2 of the base. Another key residue for the pseudouridine specificity is Thr26*, a residue in the substrate-binding loop of the adjacent subunit-B involved in forming the binding pocket of subunit-A. Its side chain hydroxyl group is located within ~2.6 Å of the N1 of uracil-Ψ, the hallmark atom for differentiating uracil-Ψ from uracil in uridine.

However, the possible involvement of Thr26* in determining pseudouridine specificity was observed only in subunit-A. In subunit-B, the side chain hydroxyl group of Thr26 is more than ~7 Å away from the N1 of uracil-Ψ, mainly due to the different conformation of the substrate-binding loop in β2–β3 (Fig. 16C). Additionally, subunit-B also shows noticeable differences in the position of the small β-strand domain relative to that in subunit-A (Figs. 12 and 16C). The β6 and β7 region, containing Val90, Ala105, and Val107 for hydrophobic interactions with uracil-Ψ, is about 1.5 to 3.0 Å more distant from uracil-Ψ in subunit-B than in subunit-A. As a consequence, the relative positions of pseudouridine in the two subunits show some differences. In subunit-B, pseudouridine is positioned further away from ADP compared to subunit-A, but positional displacements are minor, ranging only from 0.3 to 1.3 Å depending on the respective atoms (Fig. 18). For example, the pseudouridine ribosyl moiety is found almost at the same

location in the two subunits with displacements of only 0.5 to 0.7 Å, but larger displacements of 0.3 to1.3 Å are observed for the planar uracil-Ψ ring located near the hydrophobic residues in β6 and β7 region of the small β-strand domain. Therefore, the hydrophobic interactions are likely to be weaker in the active site of subunit-B and the specificity for uracil-Ψ is partially lost. These observations suggest that the small β-strand domain, which contains the substrate-binding loop and the hydrophobic residues interacting with uracil-Ψ, is able to undergo structural transitions.

Unlike the nucleobase, the pseudouridine ribosyl moiety and ADP are bound almost at the same location in both subunits (Fig. 16 B and C). In particular, the planar ribosyl moiety sits equally on the bed of β2 and β3 in the small domain. Its O2' and O3' form bidendate interactions with Asp12, and the O5' is within ~2.8 Å of Asp311, a putative catalytic residue as suggested for ribokinase (Kang et al., 2019). ADP is bound in an extended configuration to the crevice along the central β-sheet (Fig. 16A), analogous to the binding of ADP in EcRBSK (Sigrell et al., 1998) and AtRBSK (Kang et al., 2019). Its adenine base is embedded into the cavity between the two loops, the large and small ATP loop defined as the ATP-binding elements in ribokinase (Sigrell et al., 1998; Kang et al., 2019), and is involved in extensive hydrophobic interactions (Fig. 16D). Its associated ribosyl and di-phosphate groups are extended towards the ribosyl moiety of pseudouridine (Fig. 16A), fulfilling the geometrical requirement for a phosphorylation reaction occurring on the ribosyl O5' of the substrate. When the ATP in subunit-A of the AtPUKI-ATP binary complex is overlaid on the ADP in the corresponding subunit of the ternary complex, the distance between the O5' of the ribose in pseudouridine and the γ-phosphate atom of ATP is ~4.4 Å (Fig. 19), a distance proximal enough for a kinase reaction.

Figure 16. Structural features for the ternary complex of AtPUKI with pseudouridine and ADP

(A) The binding sites of pseudouridine and ADP in subunit-A are displayed, with an overlaid omitted *Fo-Fc* electron density map contoured at 2.5 σ. This enlarged surface representation is a top view of the crevice (magenta) from subunit-A of the ternary complex. The orientation of the ternary complex is similar to the one shown in Figure 11B, illustrating that the small domain (green) partly covers pseudouridine and ADP in the crevice. The proposed catalytic base Asp311 indicated as blue surface is located in the vicinity of the O5' ribosyl moiety of pseudouridine. (B) The binding environment of pseudouridine is shown for subunit-A of the ternary complex. Residues are classified in the following color code: small domain residues, green; residues from subunit-B, yellow and asterisk; catalytic base Asp311, blue label; other residues, gray. The dashed lines indicate possible hydrogen bonds within 3.0 Å. Notably, the substrate-binding loop from subunit-B is part of the subunit-A active site, and Thr26* therefore interacts specifically with the N1 of the nucleobase of pseudouridine.

(C) The binding environment of pseudouridine in subunit-B is shown. For comparison with that of subunit-A, the orientation of the view in subunit-B is adjusted to be almost identical to that of (B). Residues in subunit-B are indicated with an asterisk and classified using the following color code: small domain residues, yellow; residues from subunit-A, green; catalytic base Asp311, blue label; other residues, gray. Notably, the hydrophobic residues in the β6* and β7* region of the small domain and the substrate-binding loop from subunit-A are farther distant from the nucleobase (see text for details) in subunit-B than in subunit-A. (D) The ADP-binding mode is shown in subunit-B of the ternary complex, with the small ATP-binding loop (Gly239–Asn241, magenta) and large ATP-binding loop (Pro297–Gly308, light blue) as defined in Figure 4. The adenine base is enclosed by hydrophobic residues and the monovalent cation-binding site is indicated as a green circle. These features are identical in both subunits.

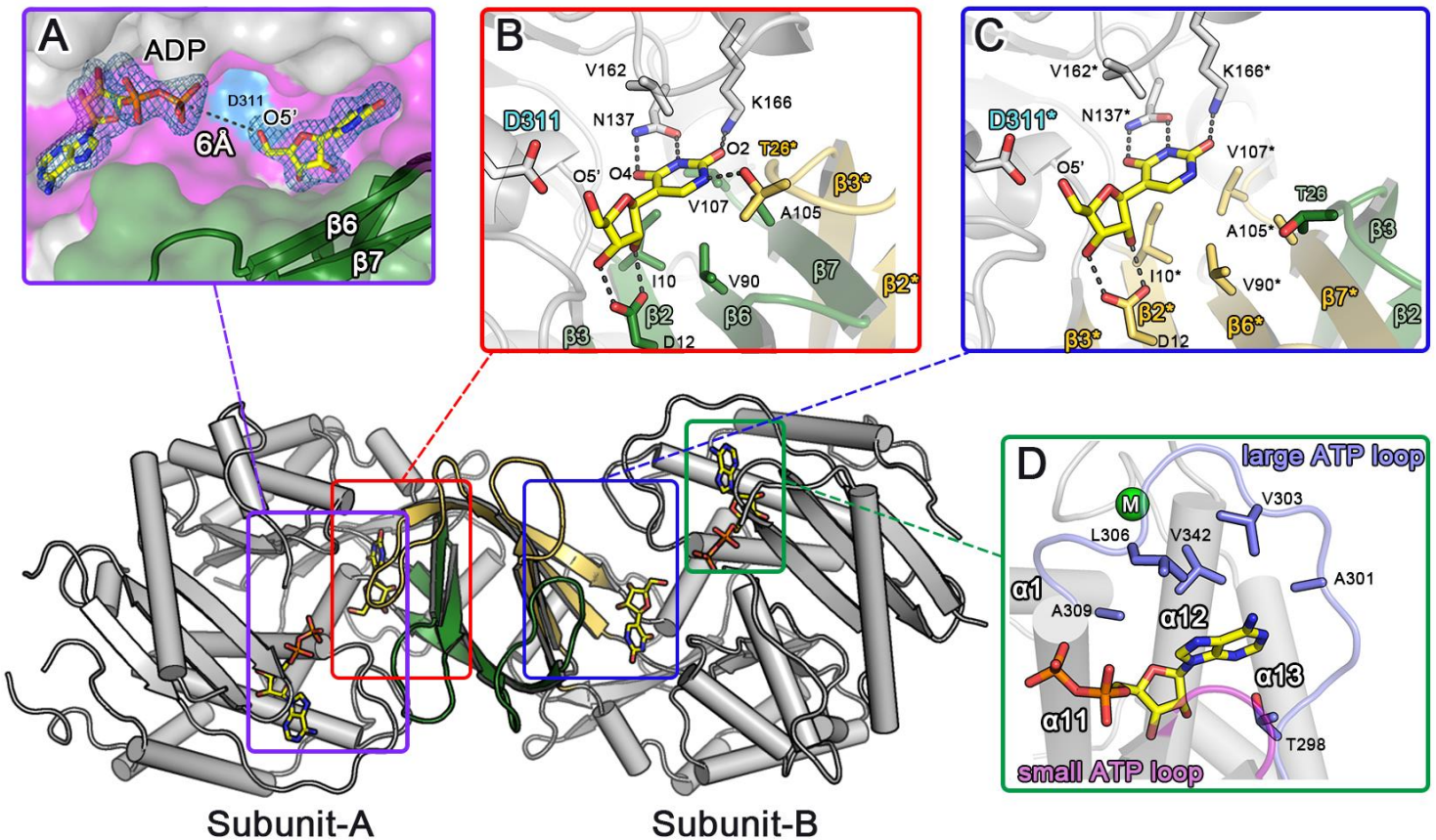

A
ADP
D311
O5'
6Å
β6
β7
B
V162
K166
D311
N137
O2
T26*
O5'
O4
β3*
V107
A105
I10
V90
β7
β2*
β2
β6
β3
D12
C
V162*
K166*
D311*
N137*
V107*
O5'
T26
A105*
β3
I10*
V90*
β7*
β2
β2*
β6*
β3*
D12
D
large ATP loop
M
V303
L306
V342
A301
α1
A309
α12
α13
α11
T298
small ATP loop
Subunit-A
Subunit-B

Figure 17. Conformational changes in the ternary complex

Conformational changes in the ternary complex. The ternary complex of AtPUKI was compared to the unliganded dimer by superimposition of the subunits-A (gray) from both structures. By substrate binding, subunit-B (magenta in the ternary complex versus green in the unliganded PUKI) undergoes a possible rigid-body rotation toward subunit-A. Due to this movement, the inter-atomic Cα distance between the two equivalent Pro202 residues on the edge of the large domain is changed from 47.6 Å to 44.7 Å.

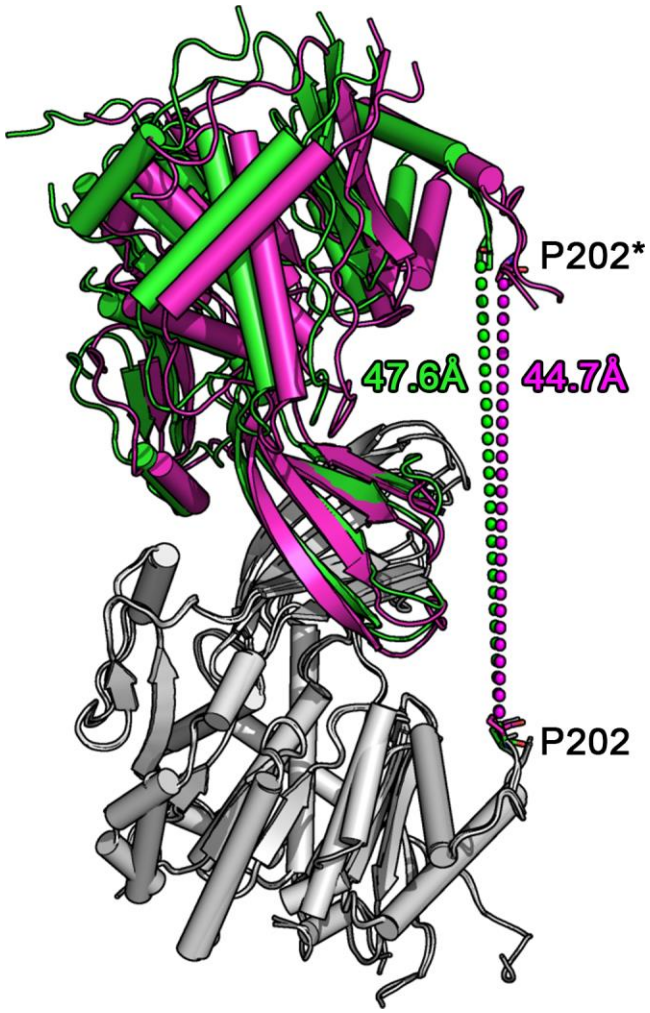

P202*
47.6Å
44.7Å
P202

Figure 18. Pseudouridine in the AtPUKI ternary complex

The two monomers of the dimeric AtPUKI in a ternary complex were superimposed based on the large α/β domain, with an RMSD of 0.53 Å for 278 Cα atoms. Color code: large α/β domains from both subunits (gray), the small domain and pseudouridine in subunit-A (green) and subunit-B (yellow). The zoom-in view shows pseudouridine and its nearby residues including the catalytic residue Asp311 from both subunits. Due to the changed relative location of the β6 and β7 regions of the small domain in subunit-A (green) versus subunit-B (yellow), pseudouridine shows positional differences for the corresponding atoms in a range of 0.3~1.3 Å.

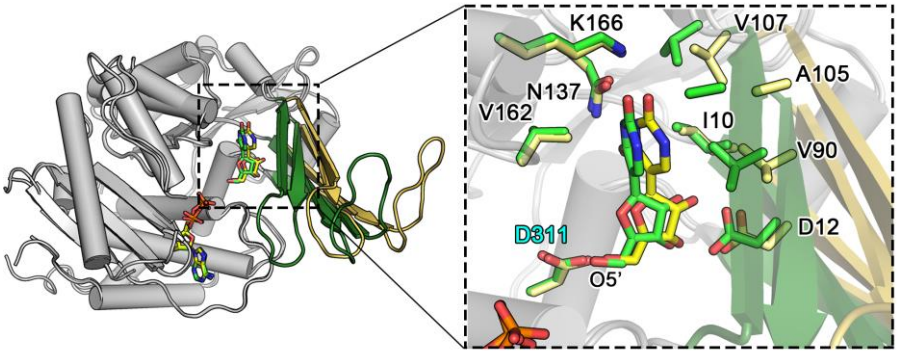
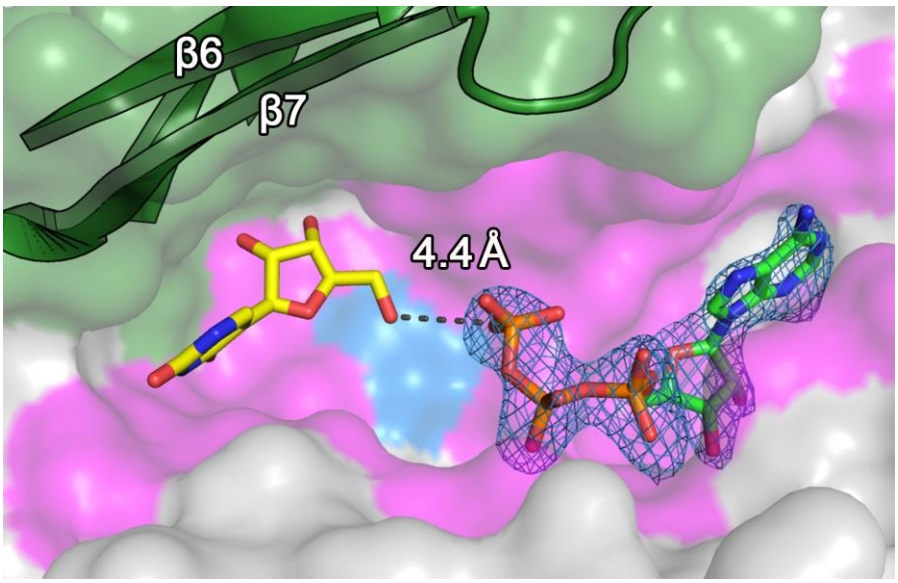

K166
V107
A105
N137
V162
I10
V90
D311
D12
O5'

Figure 19. Possible binding modes of pseudouridine and ATP in AtPUKI

Pseudouridine from the ternary complex was modeled into the AtPUKI-ATP binary complex. In the model, the binding site of pseudouridine is essentially identical to that in the ternary complex. The labels of this presentation are identical to those of Figure 16(A). ATP in the binary complex is overlain with an omitted *Fo-Fc* electron density map contoured at 2.5 σ.


β6
β7
4.4Å

Functional features of the AtPUKI active site residues

By employing an enzyme-coupled assay, I first carried out steady-state kinetic analyses of the wild-type AtPUKI (Fig. 20). Under my assay conditions, the K_m and k_{cat} values are 39.6 µM and 5.28 s^{-1} for pseudouridine, and 761 µM and 5.80 s^{-1} for ATP. Under saturation conditions for ATP (i.e., 4 mM ATP), I evaluated the substrate selectivity of AtPUKI towards various nucleosides at a concentration of 1.25 mM each (Fig. 21). Consistent with a previous study (Chen and Witte, 2020), AtPUKI was highly specific for pseudouridine, with relative activities towards cytidine and uridine of 15.4% and 1.6%, respectively, and negligible activities for other potential substrates. Thus, AtPUKI exhibits a narrow selectivity towards pyrimidine nucleosides and likely possesses structural determinants distinguishing pseudouridine from the chemically similar uridine or cytidine.

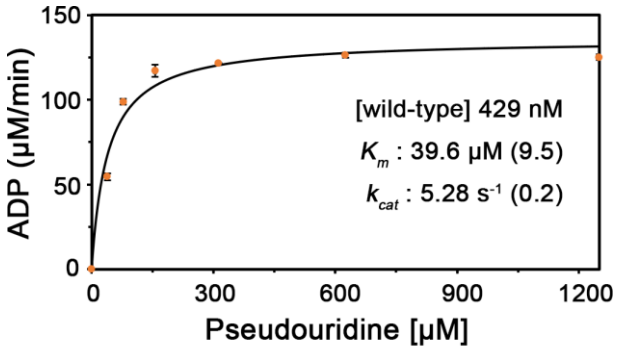
I further performed activity assays and kinetic analyses of AtPUKI mutants. Residues for site-directed mutagenesis were selected based on the binding site of pseudouridine and ADP in subunit-A (Fig. 16 B and D). ADP production was monitored for 30 s in these assays using 4 mM ATP and 1.25 mM pseudouridine as substrates. Mutations in the nucleoside binding pocket reduced the activity to a greater extent than those in the ADP-binding site (Fig. 22). Mutation of the three residues that interact likely via hydrogen bonds with the base of pseudouridine (T26A, T26V, N137A, K166A) led to activities in the range of 2.5 to 16.3% relative to the wild-type. The T26S mutant exhibited an activity of ~53%, suggesting a role of the identified hydrogen bond in catalysis. The D12A and D12N mutations that probably interfere with proper ribose binding abolished the activity almost completely, whereas a mutation of the hydrophobic residue proximal to uracil-Ψ (I10A) maintained ~50% activity.

To further validate the structure-based functional assignments of the active site residues, the kinetic parameters of the various mutants were determined. The analysis confirmed the proposed functional roles for active site residues (Figs. 23–27). For example, I10A increased the apparent K_m value by ~2-fold but not the k_{cat} value, as would be expected for a nucleobase-interacting residue (Figs. 23 and 24). Asn137 plays a central role in recognizing uracil-Ψ via a bidentate interaction with the base. Accordingly, a ~114-fold increase in the K_m value and a ~4.3-fold decrease in the k_{cat} value were observed for the N137A mutant. The K166A mutation resulted in a negative effect on K_m (2.5-fold higher) and k_{cat} (5.7-fold lower), because the long side chain of Lys166 is located at the bottom end of the binding pocket and its terminal NZ recognizes the O2 of uracil-Ψ (Fig. 16B). The most intriguing residue was Thr26 in the substrate-binding loop donated from the other subunit that recognizes the N1 of the nucleobase (Fig. 16B). Whereas the T26A and T26V mutants lose the ability to form the proposed hydrogen bond to N1 of the base, the T26S variant maintains this interaction. Accordingly, two groups can be discerned based on kinetic features: (1) the T26A and T26V mutants both with an about 4-fold higher K_m and a 6-fold lower k_{cat} value relative to the wild-type AtPUKI and (2) the T26S variant with an almost equal K_m to the wild-type and a k_{cat} value reaching 62% of that of the wild-type enzyme. These data clearly show that Thr26 plays an important role in recognizing uracil-Ψ and demonstrate that the structure of subunit-A in the ternary complex represents the catalytically active conformation of the substrate-binding loop in AtPUKI (Fig. 16B). Consistently, the active site residues described are highly conserved in PUKI orthologs of plants (Fig. 4). Especially, the hydrophilic residues Thr26, Asn137, and Lys166 interacting with the uracil-Ψ are also highly conserved in PUKI (YeiC) from bacteria and

in the metazoan bifunctional enzymes containing both PUKI and PUMY domains (Chen and Witte, 2020). In some enzymes, the equivalent position to Thr26 is replaced by serine (Fig. 4). Therefore, it is probable that the binding mode of uracil-Ψ to the active site of AtPUKI is representative for PUKI in general.

Figure 20. Kinetic analyses of wild-type AtPUKI

Steady-state kinetic analyses of wild-type AtPUKI with pseudouridine and ATP as substrates. Error bars are SD (n = 3). The measurements did not show any recognizable deviations among the measured values; thus, the error bars are not clearly visible. The K_m and k_{cat} values were obtained using SigmaPlot, with standard errors (SEs) noted in parentheses.


150
100
50
0
ADP (μM/min)
0
300
600
900
1200
Pseudouridine [μM]
[wild-type] 429 nM
K_m : 39.6 μM (9.5)
k_{cat} : 5.28 s^{-1} (0.2)

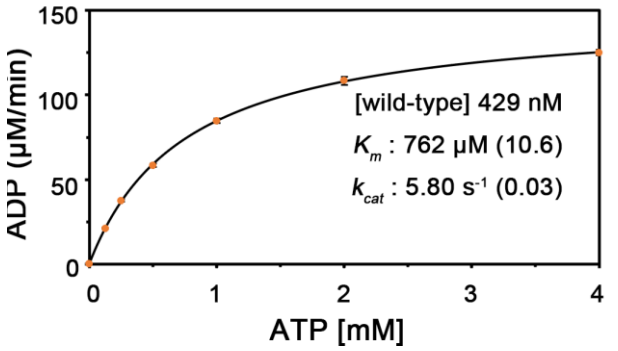

150
100
50
0
ADP (μM/min)
0
1
2
3
4
ATP [mM]
[wild-type] 429 nM
K_m : 762 μM (10.6)
k_{cat} : 5.80 s^{-1} (0.03)

Figure 21. Substrate specificity of AtPUKI and chemical structure of substrate candidates

(A) Specific activity of AtPUKI with different substrate candidates at 1.25 mM, including ribose (blue), pyrimidine nucleosides (green), and purine nucleosides (black). The resulting activity was compared to that with pseudouridine (red, 8.54 μmol min^{-1} mg^{-1} protein set to 100%). Specific activities are given in parentheses. Error bars are SD (n = 3). *n.d.*, not detected. (B) The nucleoside chemical structures used in the substrate specificity assay of AtPUKI are shown.

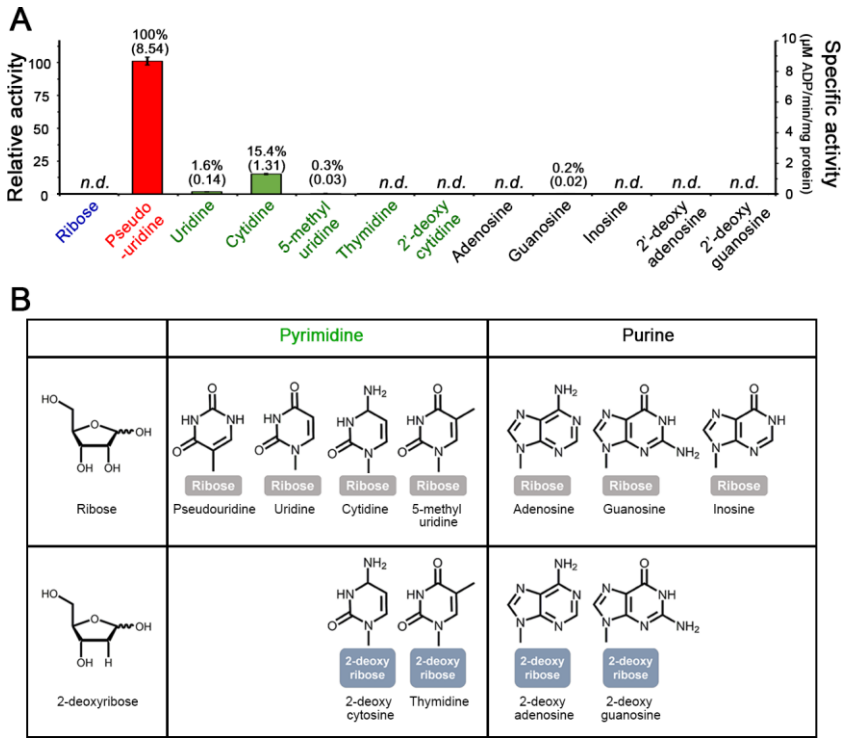

A
Relative activity
Specific activity
(μM ADP/min/mg protein)
100%
(8.54)
1.6%
(0.14)
15.4%
(1.31)
0.3%
(0.03)
0.2%
(0.02)
n.d.
Ribose
Pseudo-uridine
Uridine
Cytidine
5-methyl uridine
Thymidine
2'-deoxy cytidine
Adenosine
Guanosine
Inosine
2'-deoxy adenosine
2'-deoxy guanosine
B
Pyrimidine
Purine
Ribose
2-deoxyribose
Pseudouridine
Uridine
Cytidine
5-methyl uridine
Adenosine
Guanosine
Inosine
2-deoxy cytosine
Thymidine
2-deoxy adenosine
2-deoxy guanosine

Figure 22. Activity comparisons using a variety of mutants

Specific activity of AtPUKI variants with pseudouridine. Color code: mutants of the substrate pocket (blue), the ADP-binding site (black), and catalysis (yellow; see the section "Mg^{2+}-binding site and catalysis in AtPUKI"). The activities were compared to that of the wild-type AtPUKI (set to 100%); the wild-type enzyme data is identical to that shown in Figure 21. Specific activities are given in parentheses. Error bars are SD (n = 3). *n.d.*, not detected.

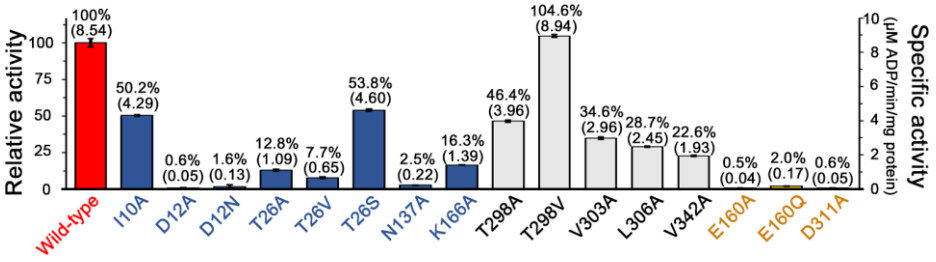
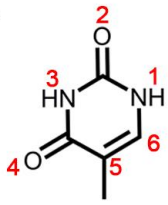

Relative activity
Specific activity
(μM ADP/min/mg protein)
100
75
50
25
0
10
8
6
4
2
0
100% (8.54)
50.2% (4.29)
0.6% (0.05)
1.6% (0.13)
12.8% (1.09)
7.7% (0.65)
53.8% (4.60)
2.5% (0.22)
16.3% (1.39)
46.4% (3.96)
104.6% (8.94)
34.6% (2.96)
28.7% (2.45)
22.6% (1.93)
0.5% (0.04)
2.0% (0.17)
0.6% (0.05)
Wild-type
I10A
D12A
D12N
T26A
T26V
T26S
N137A
K166A
T298A
T298V
V303A
L306A
V342A
E160A
E160Q
D311A

Figure 23. Kinetic parameters of AtPUKI variants with pseudouridine and structurally and chemically similar nucleosides

For clarity, only the nucleobases of the nucleosides are shown. In (A)–(D), steady-state kinetic analyses in the presence of 4 mM ATP as described in the section “Activity assays.” Atom numbers of the pyrimidine ring are given for each nucleoside; *n.d.*, not detected.

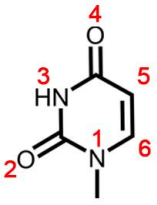
A



Pseudouridine

	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
Wild-type	39.6	5.28	1.33×10^{-1}
I10A	83.3	5.83	7.00×10^{-2}
T26A	167	8.21×10^{-1}	4.93×10^{-3}
T26V	161	7.46×10^{-1}	4.64×10^{-3}
T26S	31.6	3.29	1.04×10^{-1}
N137A	4500	1.24	2.76×10^{-4}
K166A	97.2	9.34×10^{-1}	9.61×10^{-3}

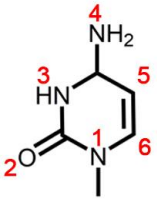
B



Uridine

	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
Wild-type	44.4	5.86×10^{-2}	1.32×10^{-3}
T26A	45.6	5.93×10^{-2}	1.30×10^{-3}
T26V	27.5	4.81×10^{-2}	1.75×10^{-3}
T26S	27.2	4.63×10^{-2}	1.70×10^{-3}
N137A	*n.d.*	*n.d.*	*n.d.*
K166A	132	4.53×10^{-2}	3.43×10^{-4}

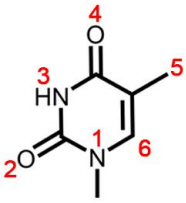
C



Cytidine

	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
Wild-type	422	1.08	2.57×10^{-3}
N137A	*n.d.*	*n.d.*	*n.d.*
K166A	228	7.81×10^{-1}	3.42×10^{-3}

D

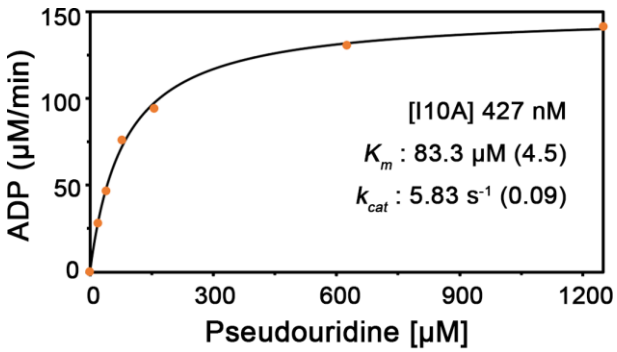


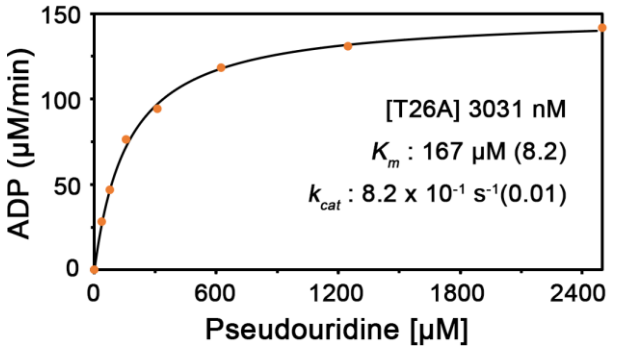
5-methyl uridine

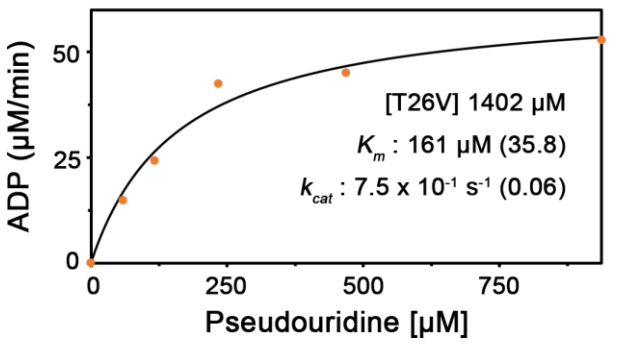
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
Wild-type	119	1.36×10^{-2}	1.14×10^{-4}
N137A	*n.d.*	*n.d.*	*n.d.*
K166A	204	3.30×10^{-2}	1.62×10^{-4}

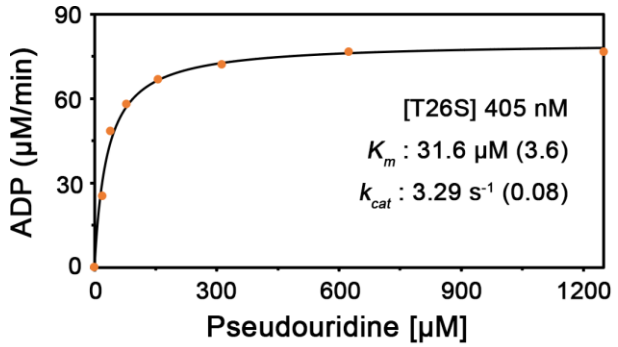
Figure 24. Kinetic analyses of AtPUKI mutants for pseudouridine

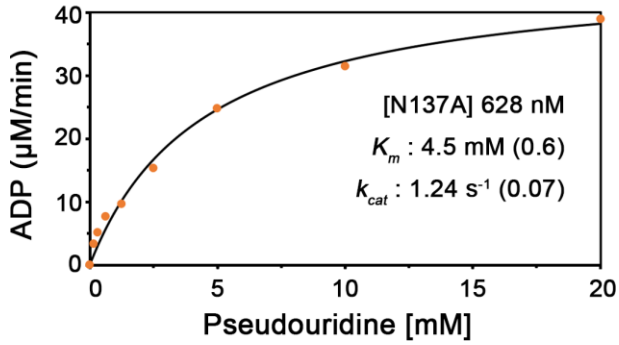
Steady-state kinetic analyses of AtPUKI mutants with pseudouridine and ATP as substrates. The K_m and k_{cat} values were obtained using SigmaPlot, with SEs noted in parentheses.


150
100
50
0
ADP (μM/min)
0
300
600
900
1200
Pseudouridine [μM]
[I10A] 427 nM
K_m : 83.3 μM (4.5)
k_{cat} : 5.83 s^{-1} (0.09)


150
100
50
0
ADP (μM/min)
0
600
1200
1800
2400
Pseudouridine [μM]
[T26A] 3031 nM
K_m : 167 μM (8.2)
k_{cat} : 8.2 x 10^{-1} s^{-1}(0.01)


50
25
0
ADP (μM/min)
0
250
500
750
Pseudouridine [μM]
[T26V] 1402 μM
K_m : 161 μM (35.8)
k_{cat} : 7.5 x 10^{-1} s^{-1} (0.06)


ADP (μM/min)
90
60
30
0
0
300
600
900
1200
Pseudouridine [μM]
[T26S] 405 nM
K_m : 31.6 μM (3.6)
k_{cat} : 3.29 s^{-1} (0.08)


ADP (μM/min)
40
30
20
10
0
0
5
10
15
20
Pseudouridine [mM]
[N137A] 628 nM
K_m : 4.5 mM (0.6)
k_{cat} : 1.24 s^{-1} (0.07)

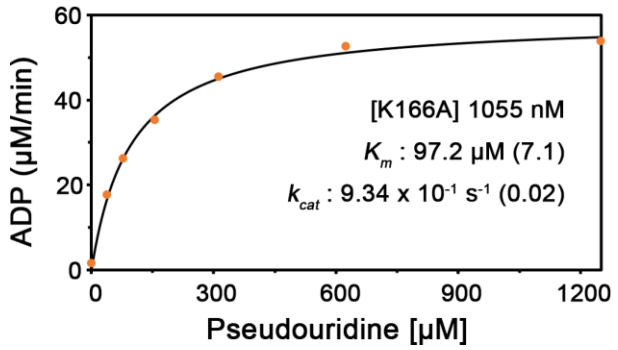
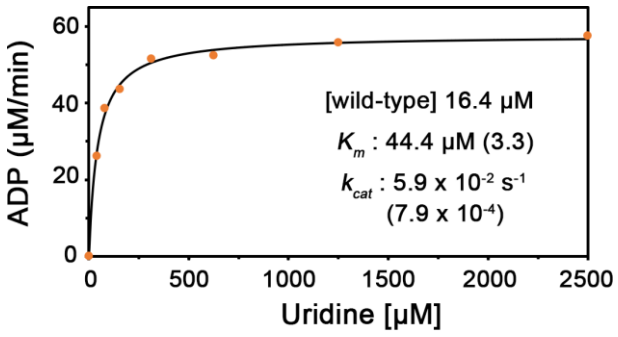
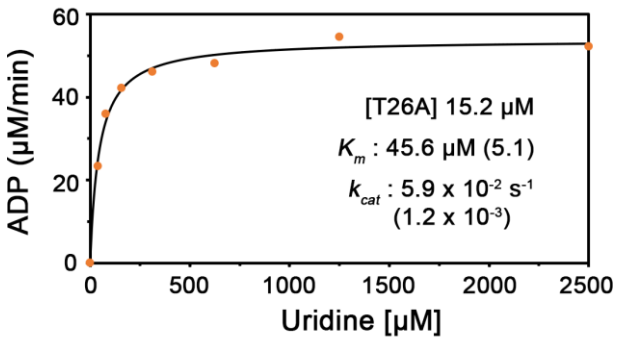
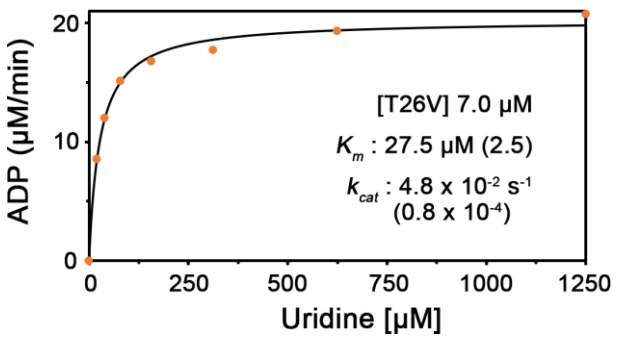

ADP (μM/min)
60
40
20
0
0
300
600
900
1200
Pseudouridine [μM]
[K166A] 1055 nM
K_m : 97.2 μM (7.1)
k_{cat} : 9.34 x 10^{-1} s^{-1} (0.02)

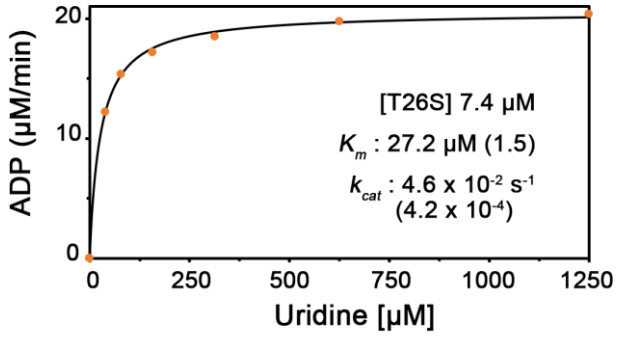
Figure 25. Kinetic analyses of AtPUKI variants for uridine

Steady-state kinetic analyses of AtPUKI mutants with uridine and ATP as substrates. The K_m and k_{cat} values were obtained using SigmaPlot, with SEs noted in parentheses.


ADP (μM/min)
60
40
20
0
0
500
1000
1500
2000
2500
Uridine [μM]
[wild-type] 16.4 μM
K_m : 44.4 μM (3.3)
k_{cat} : 5.9 x 10^{-2} s^{-1}
(7.9 x 10^{-4})


ADP (μM/min)
60
40
20
0
0
500
1000
1500
2000
2500
Uridine [μM]
[T26A] 15.2 μM
K_m : 45.6 μM (5.1)
k_{cat} : 5.9 x 10^{-2} s^{-1}
(1.2 x 10^{-3})


ADP (μM/min)
20
10
0
0
250
500
750
1000
1250
Uridine [μM]
[T26V] 7.0 μM
K_m : 27.5 μM (2.5)
k_{cat} : 4.8 x 10^{-2} s^{-1}
(0.8 x 10^{-4})


ADP (μM/min)
20
10
0
0
250
500
750
1000
1250
Uridine [μM]
[T26S] 7.4 μM
K_m : 27.2 μM (1.5)
k_{cat} : 4.6 x 10^{-2} s^{-1}
(4.2 x 10^{-4})

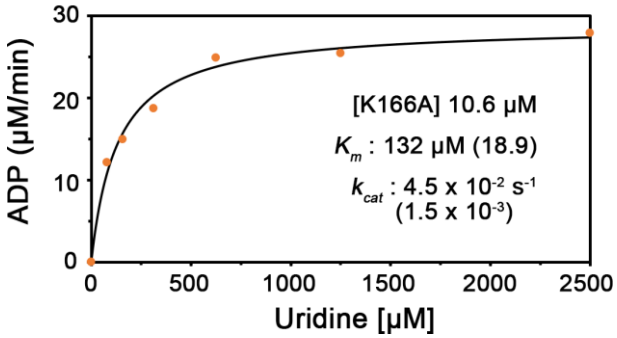
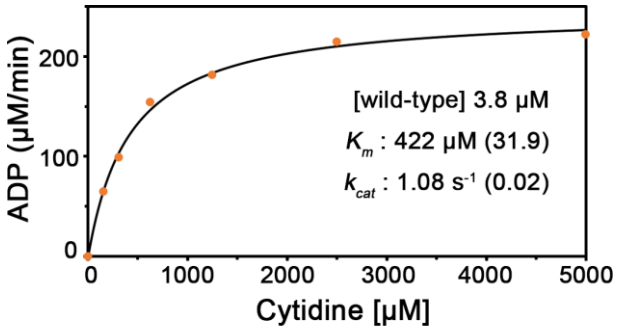

ADP (μM/min)
30
20
10
0
0
500
1000
1500
2000
2500
Uridine [μM]
[K166A] 10.6 μM
K_m : 132 μM (18.9)
k_{cat} : 4.5 x 10^{-2} s^{-1}
(1.5 x 10^{-3})

Figure 26. Kinetic analyses of AtPUKI variants for cytidine

Steady-state kinetic analyses of AtPUKI mutants with cytidine and ATP as substrates. The K_m and k_{cat} values were obtained using SigmaPlot, with SEs noted in parentheses.


ADP (μM/min)
200
100
0
0
1000
2000
3000
4000
5000
Cytidine [μM]
[wild-type] 3.8 μM
K_m : 422 μM (31.9)
k_{cat} : 1.08 s^{-1} (0.02)

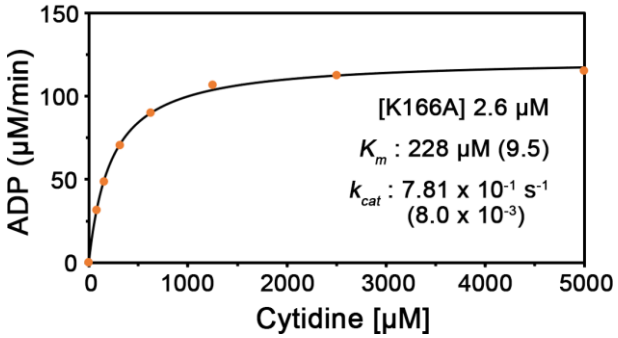
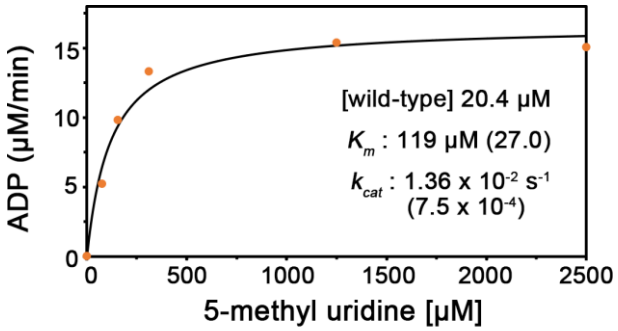
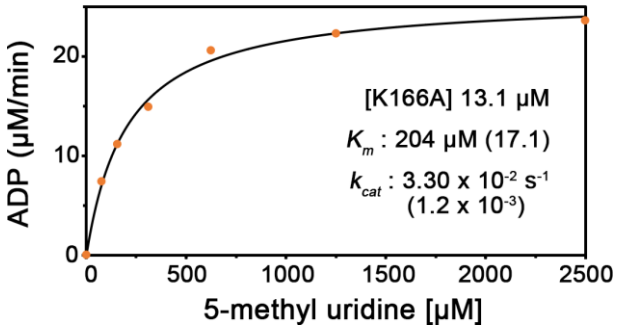

ADP (μM/min)
150
100
50
0
0
1000
2000
3000
4000
5000
Cytidine [μM]
[K166A] 2.6 μM
K_m : 228 μM (9.5)
k_{cat} : 7.81 x 10^{-1} s^{-1}
(8.0 x 10^{-3})

Figure 27. Kinetic analyses of AtPUKI variants for 5-methyl uridine

Steady-state kinetic analyses of AtPUKI mutants with 5-methyl uridine and ATP as substrates. The K_m and k_{cat} values were obtained using SigmaPlot, with SEs noted in parentheses.


ADP (μM/min)
15
10
5
0
0
500
1000
1500
2000
2500
5-methyl uridine [μM]
[wild-type] 20.4 μM
K_m : 119 μM (27.0)
k_{cat} : 1.36 x 10^{-2} s^{-1}
(7.5 x 10^{-4})


ADP (μM/min)
20
10
0
0
500
1000
1500
2000
2500
5-methyl uridine [μM]
[K166A] 13.1 μM
K_m : 204 μM (17.1)
k_{cat} : 3.30 x 10^{-2} s^{-1}
(1.2 x 10^{-3})

High fidelity of AtPUKI for pseudouridine

The question of how AtPUKI achieves its high fidelity, distinguishing pseudouridine from other pyrimidine nucleosides, was addressed by further kinetic analyses employing the substrates uridine, cytidine, and 5-methyl uridine carrying a methylated uracil base (Figs. 23–27).

Comparing the kinetic parameters of wild-type and mutant enzymes for chemically similar substrates confirmed the previous notion that the formation of several specific hydrogen bonds between AtPUKI and pseudouridine is the key factor for achieving this high substrate specificity. This is reminiscent to the so-called “direct readout” for sequence-specific interactions characterized in protein–DNA recognitions (Steitz, 1990; Pabo and Sauer. 1992). There are three crucial amino acids for the fidelity of AtPUKI: Thr26, Asn137, and Lys166. In this study, I noticed that Asn137 is essential for recognizing pseudouridine, but it is also absolutely required for binding the base of pyrimidine nucleosides in general, because the wild-type AtPUKI exhibited activity (although with large variation) to all substrates, whereas the N137A variant was essentially inactive. Specifically, the side chain carbonyl and amino groups of Asn137 fulfill the requirement for hydrogen bond formation by accepting and donating a hydrogen, respectively, from/to the hydrogen donor N3 and acceptor O4 of uracil-Ψ (or the equivalent N3 and O2 atoms of the other three nucleosides, Figs. 23 and 24). Lys166 likely recognizes the hydrogen acceptor O2 of uracil-Ψ that is equivalent to O4 in uridine but chemically differs from the hydrogen donor N4 in cytidine, belonging to the amino group distinguishing cytidine from uridine. Accordingly, the K_m values of AtPUKI for pseudouridine and uridine are almost equal, but that for cytidine is ~9.5-fold higher. However, the k_{cat} value for cytidine is higher than for uridine (Figs. 23, 25 and 26). Although AtPUKI is a better kinase for

cytidine than uridine at high substrate concentrations, this will be of little relevance *in vivo*, because the content of cytidine is generally lower than that of uridine in *Arabidopsis* (Chen et al., 2016) and uridine is not an *in vivo* substrate of AtPUKI (Chen and Witte, 2020). Thr26 in the substrate-binding loop probably forms a hydrogen bond to N1 of pseudouridine. The formation of this bond is not possible when 5-methyl uridine is the substrate (Figs. 23D and 27), resulting in an approximately 3-fold higher K_m value and a 390-fold reduction of the k_{cat} value compared to pseudouridine as substrate.

It is quite unusual that the K_m values for pseudouridine and uridine are almost identical for the wild-type AtPUKI, whereas there is an approximately 90-fold difference in the k_{cat} values (Figs. 23–25). Apparently, AtPUKI discriminates its substrates not by binding affinity but by differential turnover. AtPUKI and its mutant variants T26A, T26V, and T26S have similar K_m and k_{cat} values for uridine (Figs. 23B and 25). This is not the case for the authentic substrate, pseudouridine, for which the kinetic parameters of the T26S variant, in contrast to the T26A and T26V mutants, are almost like those of the wild-type. It appears that the interactions mediated via Thr26 in the substrate-binding loop facilitate catalysis. These observations together with the dynamic features of the β6 and β7 regions of the small domain and the substrate-binding loop, which were implied by comparing the structures of subunits-A and -B (Figs. 12 and 16 B and C), lead us to propose that AtPUKI undergoes a structural transition between a catalytically active form and an inactive form. The transition likely affects the catalytic efficiency, mainly through increasing the turnover. Probably only pseudouridine can induce the structural transition effectively recruiting the essential elements of the enzyme into the active site. Pseudouridine is therefore the by far best substrate for AtPUKI.

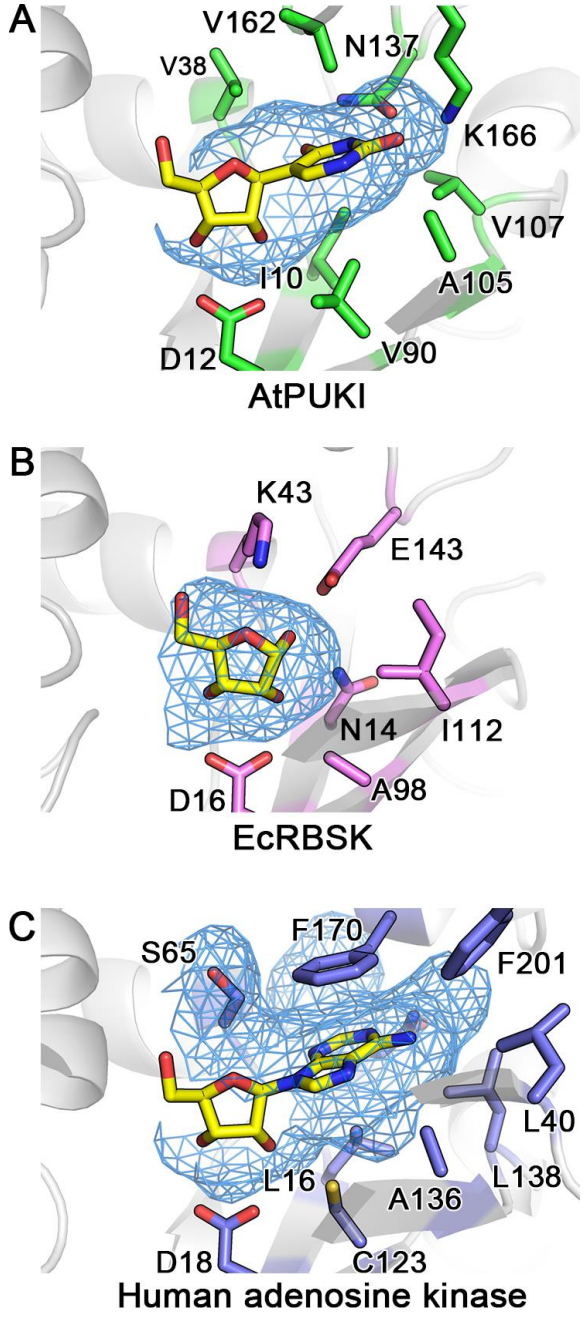
The substrate pocket of AtPUKI in comparison to ribokinase and adenosine kinase

Structure comparisons indicate that among structurally homologous proteins from the PfkB family, the nucleobase binding pocket of AtPUKI is unique in particular for the high number of specific interactions with the base involving also inter-subunit interactions and conformational changes (Fig. 28A).

In dimeric EcRBSK, each monomer has a vacancy suitable only for ribose, without any involvements of structural elements from the adjacent monomer. The region corresponding to the nucleobase pocket in AtPUKI is effectively occupied, mainly by Lys43 and Glu143, which are strictly conserved in the family of ribokinases (Fig. 28B). The monomeric human adenosine kinase contains a nucleoside binding pocket essentially at an equivalent location to where it is found in AtPUKI (Fig. 28C). However, the pocket is larger than that of AtPUKI to allow accommodation of a nucleoside with a purine base, and it consists mainly of hydrophobic residues which are also conserved in other adenosine kinases. In contrast to AtPUKI, the nucleobase binding part of the pocket lacks hydrophilic residues that could mediate specific hydrogen bonds to the adenine moiety. These structural features explain the more pronounced promiscuity of adenosine kinases regarding the substrate (Bennett and Hill, 1975; Romanello, 2013). It appears that the nucleoside binding pocket of PUKI is especially equipped for a thorough identity control of the nucleobase. This might be required to avoid the *in vivo* phosphorylation of similar nucleosides like uridine, which is more abundant than pseudouridine in *Arabidopsis* (Chen and Witte, 2020).

Figure 28. The substrate pocket of AtPUKI, ribokinase, and adenosine kinase

(A) The mesh in blue calculated by PyMol (Delano, 2002) represents the molecular surface of the protein in the absence of ligand(s), and therefore displays the pseudouridine binding pocket of AtPUKI. This view is almost identical to that of Figure 16(B). (B) Substrate pocket of *E. coli* ribokinase (PDB id, 1RK2) (Sigrell et al., 1998) showing a region of the kinase that is equivalent to that shown for AtPUKI in (A). Note that a nucleobase cannot be accommodated, mainly because Lys43 and Glu143 occupy the space. (C) The binding environment for adenosine in human adenosine kinase (PDB ID: 1BX4) shown in an identical manner as AtPUKI in (A).


A
V162
N137
V38
K166
V107
I10
A105
D12
V90
AtPUKI
B
K43
E143
N14
I112
D16
A98
EcRBSK
C
S65
F170
F201
L40
L16
A136
L138
D18
C123
Human adenosine kinase

The Mg^{2+}-binding site and catalysis in AtPUKI

A recent study on AtRBSK suggested a possible reaction mechanism by which ribose is converted into ribose-5'-phosphate involving the catalytic base Asp325 and catalytic acid Lys291 (Kang et al., 2019). In detail, $Asp325^{AtRBSK}$ (i.e., residue Asp325 of AtRBSK) in the vicinity of the ribose catalyzes the de-protonation of the ribose 5'-OH group, and $Lys291^{AtRBSK}$ near the β-phosphate group of ATP is responsible for the protonation of the ADP leaving group after a nucleophilic attack of the deprotonated O5' of ribose on the γ-phosphate group of ATP. Structural comparisons show that the binding sites for ribose, ADP, and ATP are almost identical between AtPUKI and members of the ribokinase family (Sigrell et al., 1998; Paul et al., 2015; Kang et al., 2019). From a structural perspective, the catalytic base $Asp325^{AtRBSK}$ corresponds to $Asp311^{AtPUKI}$ (Fig. 29A), but the catalytic acid $Lys291^{AtRBSK}$ does not have a functional counterpart in AtPUKI. Using the binary complex of AtPUKI with ATP, I observed that the Cα position of $Lys291^{AtRBSK}$ is superimposable to that of $Thr237^{AtPUKI}$, but the side chain of $Thr237^{AtPUKI}$ is more than 6 Å away from the β-phosphate group of ATP and unable to carry out a role as a catalytic acid (Fig. 29A).

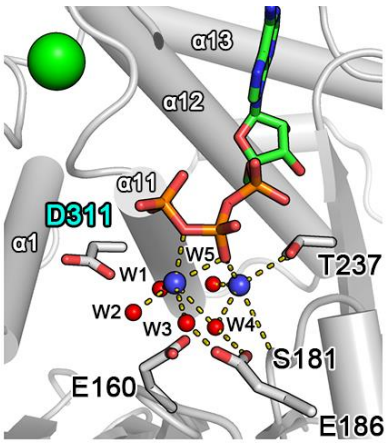
In the binary complex, I identified two Mg^{2+}-binding sites in the vicinity of the β- and γ-phosphate groups of ATP (Fig. 29A), based on the number of ligands and the coordination geometry. These features are identical in both monomers. One magnesium ion (Mg_A) between the β- and γ-phosphate groups of ATP, has six ligands: two from the oxygen atoms of ATP and the remaining four with water molecules (W1 to W4). The second magnesium ion (Mg_B) near the β-phosphate group of ATP, has five coordinating ligands, including two water molecules (W4 and W5), and three oxygen atoms: one from the β-phosphate group of ATP, one from the side chain hydroxyl group

of Thr237, and one from the main chain carbonyl group of Ser181. Therefore, there are no candidates for the catalytic acid in close vicinity of the β-phosphate group of ATP. In the Mg^{2+} shell of the binary complex, water molecules are coordinated within 3.0 Å with the nearby residues Glu160 and Glu186 (Fig. 29A): Glu160 to W4 and W5 and Glu186 to W3 and W4. The two glutamates are located more than 4 Å distant from the two magnesium ions, making a direct coordination with the magnesium ions impossible. Mutagenesis of Glu160 as well as a proposed catalytic base Asp311 essentially abolishes enzyme activity (Fig. 22). Mutants E160A, E160Q, and D311A only had 0.5%, 2.0%, and 0.6% of the specific activity of wild-type AtPUKI, respectively. Kinetic measurements varying the ATP concentration at a saturating concentration of pseudouridine (1.25 mM) indicated that the mutations caused minor effects on the K_m value for ATP but greatly reduced the enzyme efficiencies (k_{cat}/ K_m) to less than 3% compared to wild-type AtPUKI (Fig. 29B). These data and the structural environment around Glu160 suggest that a water molecule in the Mg^{2+}-coordinating shell likely plays a role in the formation of the ADP leaving group during the AtPUKI kinase reaction. Consistent with such a pivotal function, Glu160 and Glu186 are strictly conserved in the PUKI sequences from plants (Fig. 4) and also in the bifunctional enzymes from animals. In bacterial PUKI (YeiC) enzymes, a Glu or an Asp are found at the position equivalent to Glu160, and Glu186 is also here highly conserved (Chen and Witte, 2020).

Figure 29. The Mg^{2+}-binding site of AtPUKI.

(A) The Mg^{2+}-binding site in the vicinity of ADP in the ternary complex of AtPUKI. The following symbols are used: monovalent cation, green circle; magnesium ions, blue circles; water molecule, red circles. The Mg^{2+}-coordinating shell is indicated by the dashed lines, and for clarity, possible hydrogen bonds are not shown between Glu160 and water molecules including W4 and W5. (B) Steady-state kinetic analyses employing 1.25 mM pseudouridine and different concentrations of ATP, as described in the section "Activity assays." The D311A variant was produced to analyze a mutant of the catalytic base Asp311.

A



B

	K_m (μM)	k_{cat} (s^{-1})	k_{cat} / K_m ($\mu M^{-1} \cdot s^{-1}$)
Wild-type	762	5.80	7.62×10^{-3}
E160A	634	3.69×10^{-2}	5.82×10^{-5}
E160Q	564	1.34×10^{-1}	2.38×10^{-4}
D311A	997	4.14×10^{-2}	4.15×10^{-5}

Reference

Adach,I.H. and De Zoysa,M.D. (2019) Yu,Y.T. Post-transcriptional pseudouridylation in mRNA as well as in some major types of noncoding RNAs. Biochim Biophys Acta Gene Regul Mech. 1862(3):230-239.

Adams,P.D., Afonine,P.V., Bunk´oczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J. and Grossekunstleve,R.W. (2010) PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta. Crystallogr. D. Biol. Crystallogr. 66, 213–221.

Andersson,C.E. and Mowbray,S.L. (2002) Activation of ribokinase by monovalent cations. J. Mol. Biol. 315, 409-419.

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.

Bennett,L.L.Jr and Hill,D.L. (1975) Structural requirements for activity of nucleosides as substrates for adenosine kinase: orientation of substituents on the pentofuranosyl ring. Mol. Pharmacol. 11, 803–808.

Breitman,T.R. (1970) Pseudouridulate synthetase of *Escherichia coli*: Correlation of its activity with utilization of pseudouridine for growth. J. Bacteriol. 103, 263–264.

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.

Charette,M. and Gray,M.W. (2000) Pseudouridine in RNA: what, where, how, and why. IUBMB Life 49, 341–351.

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.

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.

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.

Cohn,W.E. (1960) Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. J. Biol. Chem. 235, 1488–1498.

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.

Delano,W.L. (2002) The PyMOL molecular graphics systems. DeLano Scientific. http://www.pymol.org

Diederichs,K. and Karplus,P.A. (2013) Better models by discarding data? Acta Crystallogr. D Biol. Crystallogr. 69, 1215-1222.

Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. Acta. Crystallogr. D. Biol. Crystallogr. 66, 486–501.
Eyler,D.E., Franco,M.K., Batool,Z., Wu,M.Z., Dubuke,M.L., Dobosz-Bartoszek,M., Jones,J.D., Polikanov,Y.S., Roy,B. and Koutmou,K.S. (2019) Pseudouridinylation of mRNA coding sequences alters translation. Proc. Natl. Acad. Sci. U.S.A. 116, 23068-23074.
Floyd,B.E., Morriss,S.C., MacIntosh,G.C. and Bassham,D.C. (2015) Evidence for autophagy-dependent pathways of rRNA turnover in *Arabidopsis*. Autophagy 11, 2199–2212.
Frye,M., Harada,B.T., Behm,M. and He,C. (2018) RNA modifications modulate gene expression during development. Science. 361, 1346-1349.
Girke,C., Daumann,M., Niopek-Witz,S., Möhlmann,T. Nucleobase and nucleoside transport and integration into plant metabolism. (2014) Front Plant Sci. 9;5:443.
Gilbert,W.V., Bell,T.A. and Schaening,C. (2016) Messenger RNA modifications: Form, distribution, and function. Science 352, 1408-1412.
Helm,M. (2006) Post-transcriptional nucleotide modification and alternative folding of RNA. Nucleic Acids Res. 34, 721-733.
Hickl,D., Drews,F., Girke,C., Zimmer,D., Mühlhaus,T., Hauth,J., Nordström,K., Trentmann,O., Neuhaus,H.E., Fehlmann,T., Keller,A., Simon,M. and Möhlmann,T. (2019) Differential degradation of RNA species by autophagy related pathways in plants. bioRxiv, 793950.
Holm,L. and Laakso,L.M. (2016) Dali server update. *Nucleic Acids Res*.44, W351–W355
Jia,Q. and Xie,W. (2019) Alternative conformation induced by substrate binding for *Arabidopsis thaliana* N^6-methyl-AMP deaminase. Nucleic Acids Res. 47, 3233-3243.
Jung,B., Flörchinger,M., Kunz,H.H., Traub,M., Wartenberg,R., Jeblick,W., Neuhaus,H.E. and Möhlmann,T. (2009) Uridine-ribohydrolase is a key regeulator in the uridine degradation pathway of Arabidopsis. Plant Cell. 21:876–891
Kang,P.-A., Oh,J., Lee,H., Witte,C.-P. and Rhee,S. (2019) Crystal structure and mutational analyses of ribokinase from *Arabidopsis thaliana*. J. Struct. Biol. 206, 110-118.
Karijolich,J. and Yu,Y.T. (2011) Converting nonsense codons into sense codons by targeted pseudouridylation. Nature. 15;474(7351):395-8.
Karplus,P.A. and Diederichs,K. (2012) Linking crystallographic model and data quality. Science 336, 1030-1033.
Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol. 372, 774–797
Li,X., Zhu,P., Ma,S., Song,J., Bai,J., Sun,F. and Yi,C.(2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol. 11(8):592-7.
Loh,K.D., Gyaneshwar,P., Markenscoff Papadimitriou,E., Fong,R., Kim,K.S.,

Parales,R., Zhou,Z., Inwood,W. and Kustu,S. (2006) A previously undescribed pathway for pyrimidine catabolism. Proc. Natl. Acad. Sci. U.S.A. 103, 5114-5119.
Maj,M.C. and Gupta,R.S. (2001) The effect of inorganic phosphate on the activity of bacterial ribokinase. J Protein Chem. 20, 139–144.
Mallette,E. and Kimber,M.S. (2018) Structural and kinetic characterization of (*S*)-1-amino-2-propanol kinase from the aminoacetone utilization microcompartment of *Mycobacterium smegmatis*. J. Biol. Chem. 293, 19909–19918.
Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307-326.
Pabo,C.O. and Sauer,R.T. (1992) Transcription factors: structural families and principles of DNA recognition. Annu. Rev. Biochem. 61, 1053-1095.
Park,J., Singh,B., Maj,M.C. and Gupta,R.S. (2004) Phosphorylated derivatives that activate or inhibit mammalian adenosine kinase provide insights into the role of pentavalent ions in AK catalysis. Protein J. 23, 167–177.
Park,J. and Gupta,R.S. (2008) Adenosine kinase and ribokinase: the RK family of proteins. Cell Mol. Life Sci. 65, 2875–2896.
Paul,R., Patra,M.D. and Sen,U. (2015) Crystal structure of apo and ligand bound *Vibrio cholerae* ribokinase (Vc-RK): role of monovalent cation induced activation and structural flexibility in sugar phosphorylation. Adv. Exp. Med. Biol. 842, 293–307.
Penzo,M., Guerrieri,A.N., Zacchini,F., Treré,D. and Montanaro,L. (2017) RNA Pseudouridylation in Physiology and Medicine: For Better and for Worse. Genes (Basel). 1;8(11):301.
Preumont,A., Snoussi,K., Stroobant,V., Collet,J.-F. and Van Schaftingen,E. (2008) Molecular identification of pseudouridine metabolizing enzymes. J. Biol. Chem. 283, 25238–25246.
Robert,X. and Gouet,P. (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324
Reumann,S. (2011) Toward a definition of the complete proteome of plant peroxisomes: Where experimental proteomics must be complemented by bioinformatics. Proteomics 11, 1764–1779.
Romanello,L., Bachega,J.F., Cassago,A., Brandão-Neto,J., DeMarco,R., Garratt,R.C. and Pereira,H.D. (2013) Adenosine kinase from *Schistosoma mansoni*: structural basis for the differential incorporation of nucleoside analogues. Acta. Crystallogr. D. Biol. Crystallogr. 69, 126–136.
Roundtree,I.A., Evans,M.E., Pan,T. and He,C. (2017) Dynamic RNA modifications in gene expression regulation. Cell 169, 1187-1200.
Schroeder,R.Y., Zhu,A., Eubel,H., Dahncke,K. and Witte,C.-P. (2018) The ribokinases of *Arabidopsis thaliana* and *Saccharomyces cerevisiae* are required for ribose recycling from nucleotide catabolism, which in plants is not essential to survive prolonged dark stress. New Phytol. 217, 233-244.
Schwartz,S., Bernstein,D.A., Mumbach,M.R., Jovanovic,M., Herbst,R.H.,

León-Ricardo,B.X., Engreitz,J.M., Guttman,M., Satija,R., Lander,E.S., Fink,G. and Regev,A. (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159, 148-162.

Sigrell,J.A., Cameron,A.D., Jones,T.A. and Mowbray,S.L. (1998) Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 å resolution: insights into a new family of kinase structures. Structure 6, 183–193.

Sigrell,J.A., Cameron,A.D. and Mowbray,S.L. (1999) Induced fit on sugar binding activates ribokinase. J. Mol. Biol. 290, 1009–1018.

Steitz,T.A. (1990) Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding. Q. Rev. Biophys. 23, 205-280.

Sun,L., Xu,Y., Bai,S., Bai,X., Zhu,H., Dong,H., Wang,W., Zhu,X., Hao,F. and Song,C.-P. (2019) Transcriptome-wide analysis of pseudouridylation of mRNA and non-coding RNAs in *Arabidopsis*. J. Exp. Bot. 70, 5089–5600.

Tintemann,H., Wasternack,C., Benndorf,R. and Reinbothe,H. (1985) The ratelimiting step of uracil degradation in tomato cell-suspension cultures and Euglena-gracilis invivo studies. Comp Biochem Physiol B Biochem Mol Biol 82: 787–792.

Van Duyne,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L. and Clardy,J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J. Mol. Biol. 229, 105-124.

Witte,C.-P. and Herde,M. (2020) Nucleotide metabolism in plants. Plant Physiol. 182, 63-78.

Witz,S., Jung,B., Fürst,S. and Möhlmann,T. (2012) De novo pyrimidine nucleotide synthesis mainly occurs outside of plastids, but a previously undiscovered nucleobase importer provides substrates for the essential salvage pathway in Arabidopsis. Plant Cell. 24(4):1549-59.

Wu,B., Zhang,D., Nie,H., Shen,S., Li,Y. and Li,S. (2019) Structure of *Arabidopsis thaliana* N^6-methyl-AMP deaminase ADAL with bound GMP and IMP and implications for N^6-methyl-AMP recognition and processing. RNA Biol. 16, 1504-1512.

Wu,S., Alseekh,S., Cuadros-Inostroza,Á., Fusari,C.M., Mutwil,M., Kooke,R., Keurentjes,J.B., Fernie,A.R., Willmitzer,L. and Brotman,Y. (2016) Combined use of genome-wide association data and correlation networks unravels key regulators of primary metabolism in Arabidopsis thaliana. PLoS Genet 12: e1006363.

Zaccara,S., Ries,R.J. and Jaffrey,S.R. (2019) Reading, writing and erasing mRNA methylation. Nat. Rev. Mol. Cell Biol. 20, 608-624.

Zrenner,R., Stitt,M., Sonnewald,U. and Boldt,R. (2006) Pyrimidine and purine biosynthesis and degradation in plants. Annu Rev Plant Biol. 57:805-36.

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.

Accession numbers

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data bank under accession number 7C1X, 7C1Y and 7C1Z.

Abstract in Korean

RNA 의 변형은 종류가 매우 다양하며 각각 RNA 들의 안정성, mRNA 와 단백질들간의 상호작용 그리고 단백질로의 번역과정에 효율성을 조절하는 등 다양한 역할을 담당한다. 대부분의 경우, RNA 의 변형은 효소들에 의해서 특정 RNA 에 부위 특이적으로 일어난다. 이들의 생물학적 역할, 생합성과정 관련 연구들과 비교해 변형된 RNA 의 분해 관련 연구는 비교적 미비하다. 가장 많이 발견되는 RNA 변형중 하나인 pseudouridine 의 대사 관련 효소들이 최근 애기장대에서 동정이 되었다. 애기장대에서는 두개의 효소에 의해 pseudouridine 의 분해를 촉매된다: PSEUDOURIDINE KINASE (PUKI) 와 PSEUDOURIDINE MONOPHOSPHATE GLYCOSYLASE (PUMY). PUKI 와 PUMY 는 pseudouridine 을 pseudouridine monophosphate 로 인산화 시키는 과정과 pseudouridine monophosphate 를 uracil 과 ribose 5'-phosphate 로 가수분해하는 과정을 각각 촉매 한다. 차례로 해당 산물들은 pyrimidine 의 일반적인 대사과정이나 재활용 과정에 사용된다. 본 학위논문에서는 pseudouridine 분해의 첫번째과정을 담당하는 효소인 애기장대 유래 PUKI 의 단백질 삼차 구조를 분석하고, 규명된 단백질의 구조를 기반으로 기질 특이성에 관련된 인자들을 밝히고자 한다. AtPUKI 는 PfkB family 에 속하는 carbohydrate kinase 로써, homodimer 가 생물학적 기능 단위이다. 구조적인 특징으로는 α/β domain 를 중심으로 가지고 있으며, 이로부터 유래한 β-strand domain 로 구성이 되어있다. 흥미롭게도, β-strand domain 는 dimerization interface 를 제공하는 동시에 기질 특이성을 결정짓는 역할을 수행한다. AtPUKI, pseudouridine 그리고 ATP 결합구조를 토대로 AtPUKI 에는 pseudouridine 이 결합할 수 있는 독특한 결합부위가 존재하며, 이는 여러 개의 친수성 아미노산들에 의해 매개가 되는 것을 확인하였다. 특히, 인접한 단량체의 β-strand domain 로부터 유래한 loop 또한 기질의 특이성을 결정짓는데 중요한

역할을 하며, 구조적인 변화의 수반이 요구된다는 것을 제시하였다. 이러한 동적인 특징은 AtPUKI 가 uridine 보다 pseudouridine 에 높은 촉매 효율을 보이는 이유를 잘 설명한다. 두 기질들은 모두 AtPUKI 와 비슷한 결합 친화도를 보이며 잘 결합하지만, 오직 pseudouridine 만이 AtPUKI 의 구조적인 변화를 야기하며 효소에 의해 효율적으로 분해된다(높은 turnover rate)는 것을 제시했다. 이러한 결과들은 어떻게 AtPUKI 가 uridine 을 포함하는 다른 pyrimidine nucleoside 의 항상성을 방해하지 않고, pseudouridine 의 인산화에만 관여하는지를 잘 설명하며, 더 나아가 PfkB family 에 속하는 다양한 효소들의 구조적 그리고 기능적 다양성의 예시로서 사용될 수 있다.

주요어 : Pseudouridine catabolism, RNA modification ,Non-canonical nucleotide catabolism, Pseudouridine kinase, Crystal structure, Substrate specificity, *Arabidopsis thaliana*

학번 : 2013-23265