

## Electronic Supplementary Material (ESI) for Chemical Communications.

### Supplementary Information

#### The C-terminal loop of *Arabidopsis thaliana* guanosine deaminase is essential to catalysis

Qian Jia,<sup>a</sup> Hui Zeng,<sup>a</sup> Huanxi Li,<sup>a</sup> Nan Xiao,<sup>a</sup> Jing Tang,<sup>a</sup> Shangfang Gao,<sup>a</sup> Jinbing Zhang,<sup>a</sup> and Wei Xie<sup>\*a</sup>

<sup>a</sup>MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory for Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou, Guangdong, 510006, People's Republic of China. E-mail: xiewei6@mail.sysu.edu.cn; Tel: 862039332943

#### Experimental Procedures

##### Materials and general method

##### Cloning, expression and protein purification

The full-length gene encoding GSDA (At5g28050, accession number NM\_122688.4) was amplified by PCR from the cDNA of *Arabidopsis thaliana* using the primers shown in Table S2. After the double digestion by the restriction enzymes NdeI and NotI, the PCR product was inserted into the pET-21b (+) vector. The proteins expressed would possess a C-terminal 6 × His tag. The full-length version of the enzyme, including wildtype (WT) and mutants, was employed for general biochemical studies (TLC, kinetic and TSA studies). For crystallization, the S29-Y185 fragment (SF) was subcloned into an engineered pET-28a (+) vector (MerckMillipore). All the mutations were created through Quikchange PCRs based on the full-length or the SF versions of the WT sequence and their primers were listed in Table S2.

The *Escherichia coli* strain BL21 (DE3) cells were transformed with the plasmids for overexpression. The transformed cells were cultured overnight in the Luria-Bertani broth containing 50 mg/L ampicillin or 30 mg/L kanamycin at 37°C. A 2-L fresh culture medium was inoculated with 20 mL of the overnight culture. When OD<sub>600</sub> reached 0.8, the expression of AtGSDA was induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After overnight growth at 18°C, the *E. coli* cells were then pelleted by centrifugation at 3,500 g for 20 min and resuspended in pre-chilled nickel-nitrilotriacetic acid (Ni-NTA) buffer A containing 40 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10 mM imidazole, 1 mM β-mercaptoethanol (β-ME)

and 1 mM phenylmethylsulfonyl fluoride (PMSF). The resuspended cells were disrupted by ultrasonication and the supernatant was obtained by centrifugation at 23,500 g for 1 h at 4°C. The supernatant was then applied onto Ni-NTA affinity resin (Qiagen) pre-equilibrated with Ni-NTA buffer A. The target protein was eluted with Ni-NTA buffer B containing 40 mM Tris-HCl (pH 8.0), 250 mM NaCl, 250 mM imidazole, 1 mM  $\beta$ -ME and 1 mM PMSF. The fractions were further subjected to anion exchange purification by a Q HP column (GE Healthcare) using a NaCl gradient, and protein was eluted at ~150 mM NaCl. The purified protein was pooled and dialyzed in a buffer consisting of 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM DTT. AtGSDA was stored at 3.0 mg/ml at -80°C after being flash-frozen by liquid nitrogen. The same protocol was adopted for the expression and purification of the WT-SF truncation or other mutants followed. For the crystallization purpose, the N-terminally fused His-tag was cleaved off by the PreScission protease (PSP) (GE Healthcare) treatment overnight at a molar ratio of 80:1 (AtGSDA/PSP), while dialyzing against a buffer containing 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM DTT. The target proteins without the tag were enriched by collecting the unbound fractions from a Histrap column (GE Healthcare), which employed the same buffers as those for Ni-NTA affinity chromatography. All the proteins were purified as described above and stored at -80°C after being flash frozen by liquid nitrogen.

### **Crystallization and structure determination**

Initial crystallization screens were set up using the sitting-drop vapor diffusion method, and 4.0 mg/mL SF proteins were mixed with an equal volume of the reservoir solution at 4°C respectively. Crystals of the apo-protein were obtained in a condition of 1.2 M Na-citrate and 0.1 M HEPES (pH 7.5). The cocrystals of the complexes with guanosine or xanthosine were obtained by mixing E82Q-SF or WT-SF with corresponding ligands at molar ratios of 1:5-10 respectively, prior to crystallization. All of the cocrystals were obtained under the identical condition to that of the apo-protein.

The crystals for the complexes were soaked for 1-3 min in a cryoprotective solution containing all the components of the reservoir solution supplemented with 20% glycerol (v/v). The soaked crystals were mounted on nylon loops and flash frozen in liquid nitrogen. The native data of apo-protein and guanosine-cocrystals were collected from frozen crystals at -173 °C on ADSC CCD detectors using Beamline 19U1 (BL19U1) at the Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, P.R. China).<sup>1</sup> The data were processed with the program *HKL3000*.<sup>2</sup> The data of xanthosine-cocrystals was collected using an Oxford Diffraction Xcalibur Nova diffractometer from the home X-ray source. The diffractometer was operated at 50 kV and 0.8

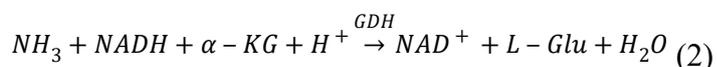
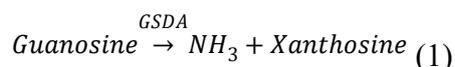
mA, with a rotation of 1° per frame at -120 °C. The data were recorded using a 65-mm Onyx CCD detector, and the exposure time was 30 s for each frame. All the space groups of the (co)crystals were  $P6_1$ , and there were two molecules in the asymmetric unit. To solve its structure, molecular replacement (MR) was first performed using Phaser<sup>3</sup> with the coordinates of tRNA-specific adenosine-34 deaminase from *Agrobacterium fabrum* (AtTadA, PDB 2A8N) as the search model.<sup>4</sup> The phase was further improved by density modification, and multiple cycles of refinement alternating with model rebuilding were carried out by PHENIX. Refine.<sup>5</sup> The dimer model was finished with 310 residues built into the density.<sup>6</sup> The final R-factor was 16.7% ( $R_{\text{free}} = 19.0\%$ ). The final model was validated by Molprobit.<sup>7</sup> The Ramachandran plot of the final model has 97.4%, 2.3%, and 0.3% of the residues in the most favorable, generously allowed and disallowed region. The structures of the ligand-bound complexes were solved in a similar fashion, but with the finished model of the apo-protein as the search probe (PDB 7DBF) for the following MR runs. The ligand-bound structures were very similar to that of the apo-form, except for the C-terminus, which was built into the density manually. After several rounds of refinement, the density of the ligands became evident in the difference map. The ligands were added in the final round of refinement when the structure refinement converges. All the data collection and structure refinement statistics were summarized in Table S3. The structural figures were produced with PyMOL (www.pymol.org).

### Deaminase assays by TLC

The assay mixture contained 20 mM Tris-HCl (pH 8.0), 150 mM NaCl and 5 mM guanosine or other ligands. 50  $\mu\text{M}$  full-length AtGSDA or mutants were added to initiate the reaction. Then the reactions were incubated for 30 min at 37°C. Aliquots of 0.5  $\mu\text{L}$  were taken and spotted onto a cellulose polyethyleneimine plate (Merck Millipore). The reactions were developed with a mixture of 5 mM  $(\text{NH}_4)_3\text{PO}_4$  and 50%  $\text{CH}_3\text{CN}$ .

### Kinetic studies

The enzymatic kinetics are investigated through the two stepwise reactions and characterized by the changes of the  $\text{NAD}^+/\text{NADH}$  absorbance at the wavelength of 340 nm.<sup>8,9</sup>



The reaction buffer consists of 20 mM HEPES (pH7.5), 150 mM NaCl and potential substrates of 50-600  $\mu$ M. AtGSDA was added to initiate the reaction. After 60 s, a mixture containing 0.4 mM NADH, 1 mM  $\alpha$ -ketoglutarate ( $\alpha$ -KG) and 4 units glutamate dehydrogenase (GDH) (Coolaber) was added to allow the formation of glutamate. The resulting reductions in OD<sub>340</sub> readings were monitored by a Synergy microplate reader (Biotek) after its stabilization and were compared to the standard curve established previously, to derive the changes in NADH concentrations. The corresponding  $K_M$  and  $k_{cat}$  were calculated by the Lineweaver-Burk plot.

## **TSA**

Test ligands were prepared in buffers consisting of 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT. A 20- $\mu$ l assay mixture containing 2.0  $\mu$ g GSDA wild type or mutants, 2- $\mu$ l 2 $\times$  SYPRO orange fluorescence dye (Sigma-Aldrich) and 3 mM ligand was mixed in 96-well PCR plates on ice. The plates were incubated at 25 °C for 10 min in a StepOnePlus Real-Time PCR system (Life Technologies) and were then gradually heated to 95 °C with a heating rate of 1 °C min<sup>-1</sup>. The fluorescence signals of the dye at 490/530 nm wavelengths (for excitation and emission, respectively) during the thermal denaturation were recorded every 30 s. The melting curves of the mutants were fitted by a Boltzmann model with Origin 8.0 software (OriginLab) to derive the melting temperature. The assays were conducted in triplicates for all the mutants and the control.

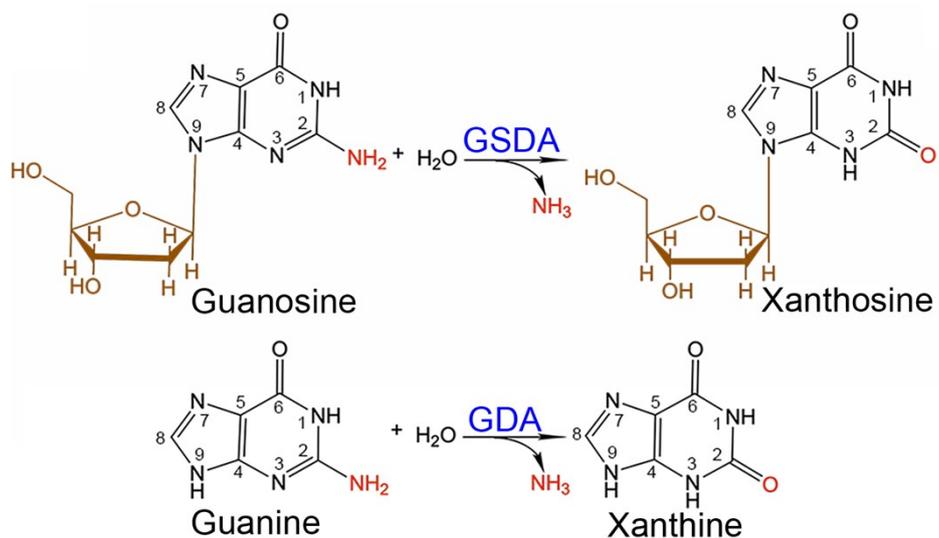
## **Mass spectrometry**

Samples were prepared by mixing 0.2 mM AtGSDA with the corresponding ligands at a molar ratio of 1:5 and were incubated for 30 min at 37°C. The samples were filtered by syringe filters, and were further diluted 100-fold before injection. The reaction products were either detected by an LTQ Orbitrap (Thermo) or a Tims TOF spectrometer (Bruker) with an electrospray ion source operating in the positive ion mode (ESI+). the reaction products

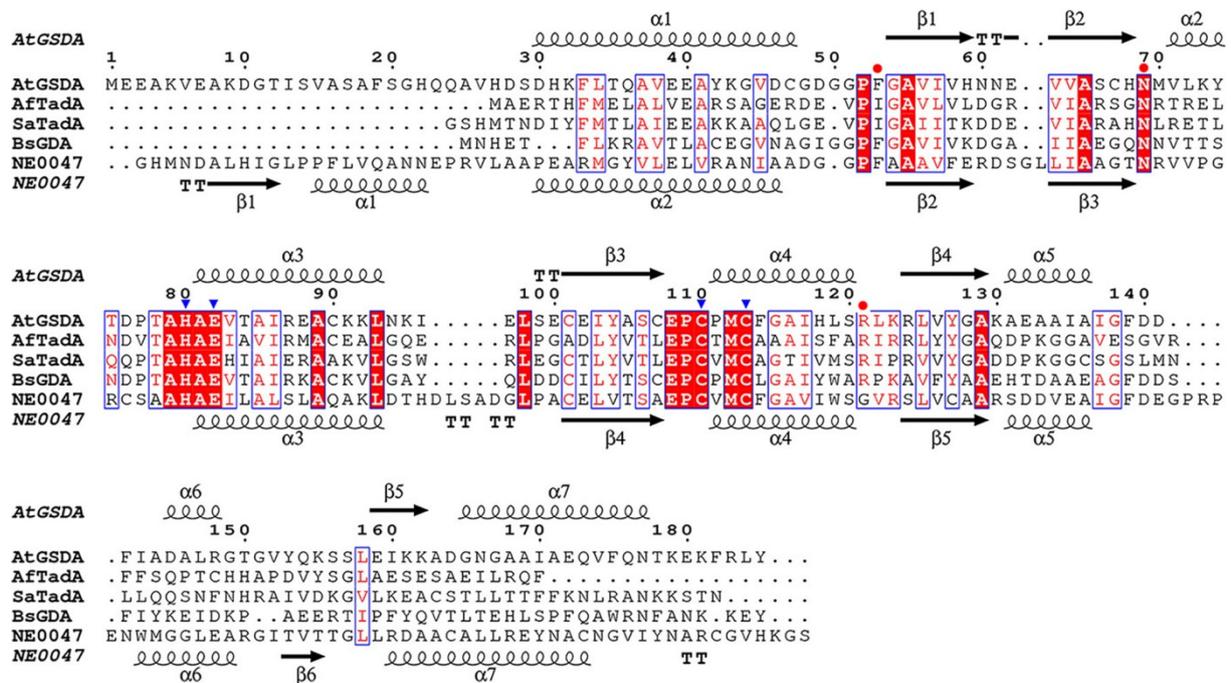
were separated into a C4-reverse column. The instrument parameters for the mass spectrometer were as follows: ion polarity = positive; scan range for full scan data collection = 50-1000 m/z; set capillary = 4500 V; set end plate offset = -500 V; set collision cell RF =500 Vpp; set nebulizer = 1.5 Bar; set dry heater = 220°C. Possible structures for the product ions were derived using Mass Frontier version 4.0 and the threshold for the correct molecular weight was limited to 5 ppm (Thermo).

## **Fluorescence spectroscopy**

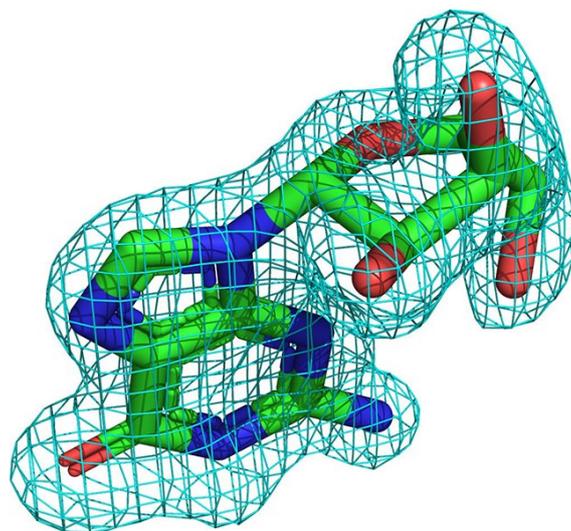
The assays were carried out with 5  $\mu\text{M}$  protein in the buffer of 150 mM NaCl, 20 mM Tris-HCl pH 8.0. The concentrations of guanosine or xanthosine used were between 5 and 320  $\mu\text{M}$ . The fluorescence spectra were conducted on a RF530R1PC fluorescence spectrophotometer (Shimadzu). The interactions were monitored by the quenching of the protein fluorescence. The excitation was set at 280 nm, and the emission was at 290-450 nm, with a slit width of 1 nm.



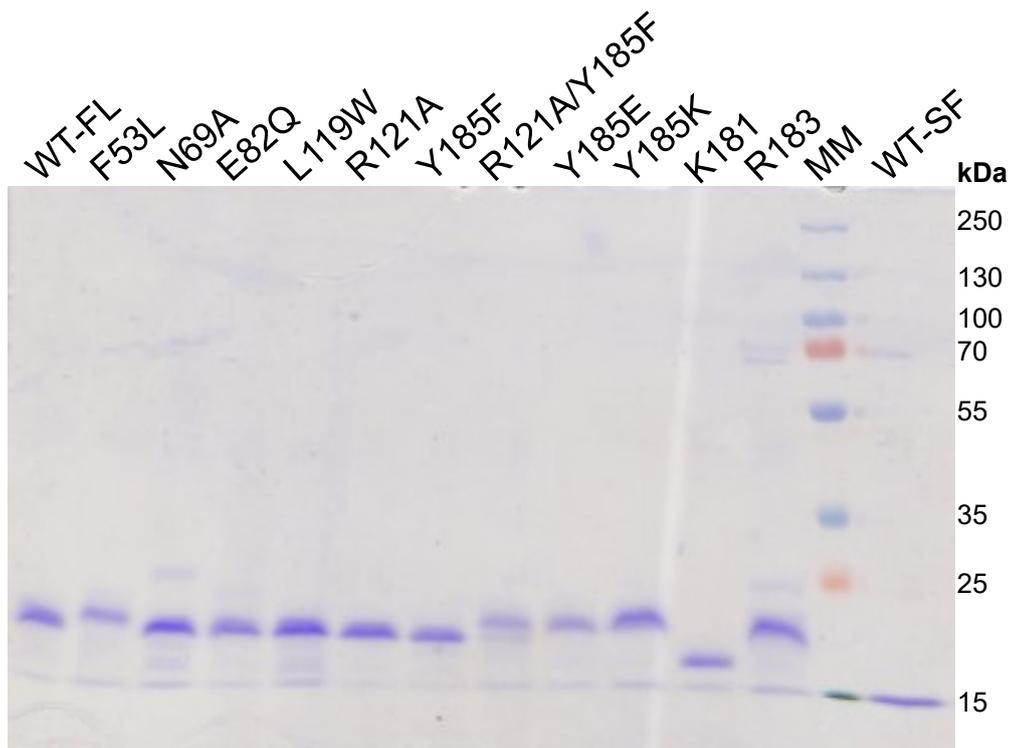
**Fig. S1** The deamination reactions catalyzed by the two types of enzymes. The chemical structural differences between the substrates and products were colored red and orange. The numbering of the purine rings was shown.



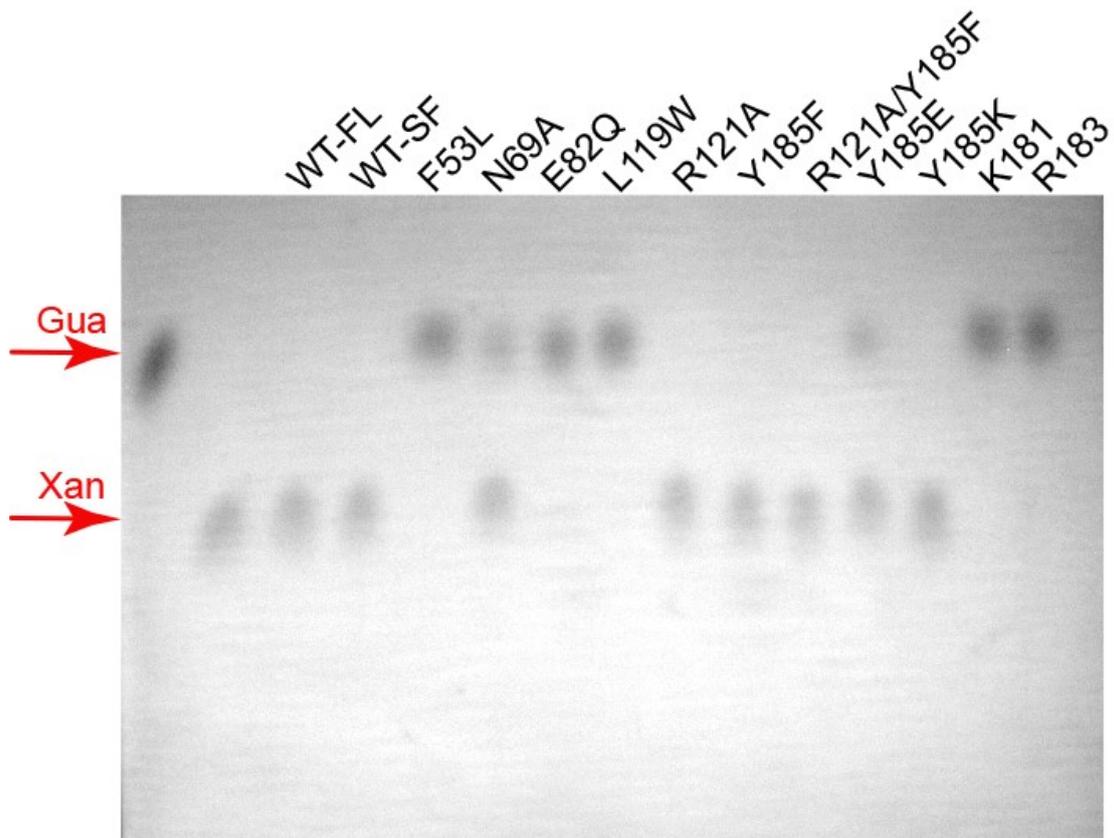
**Fig. S2** The multiple sequence alignment. The secondary structure elements were labeled above the sequences. The blue triangles dots indicated residues responsible for zinc binding and the general base Glu82, and the red solid dots indicated the essential residues for catalysis. The PDB codes for AtGSDA, AfTadA, SaTadA, BsGDA and NE0047 are 7DBF, 2A8N, 2B3J, 1WKQ and 4LCN, respectively.



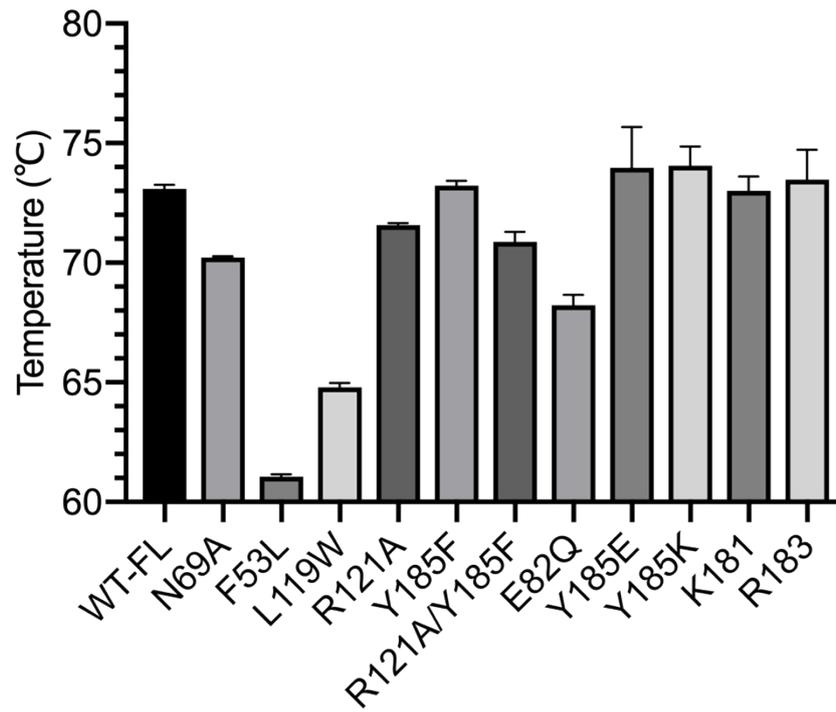
**Fig. S3** The substrate guanosine in sticks representation. The electron density of the composite omit map contoured at  $2\sigma$ .



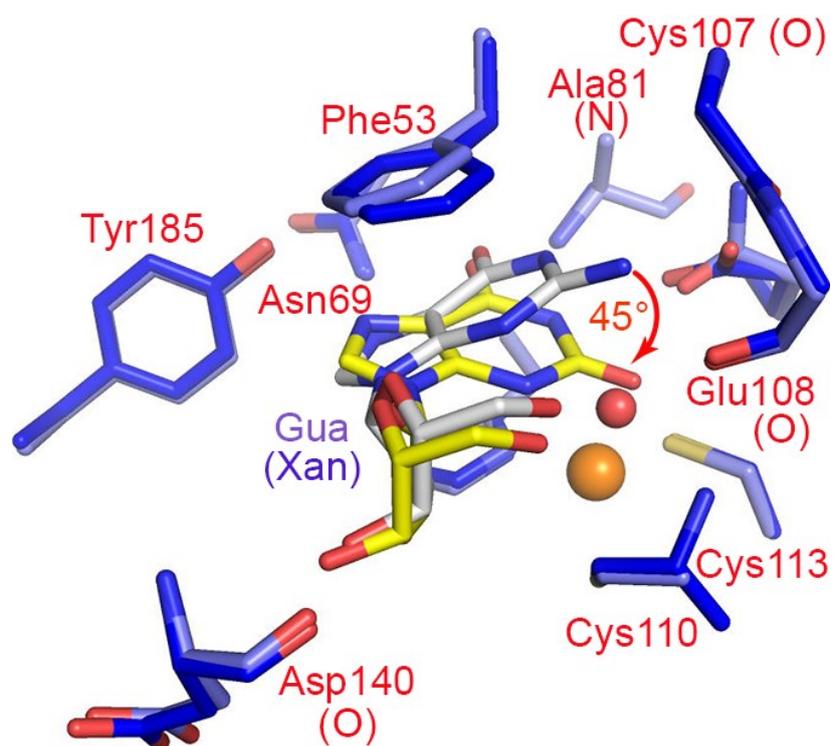
**Fig. S4** The purification results of the AtGSDA WT and mutants. MM: molecular marker.



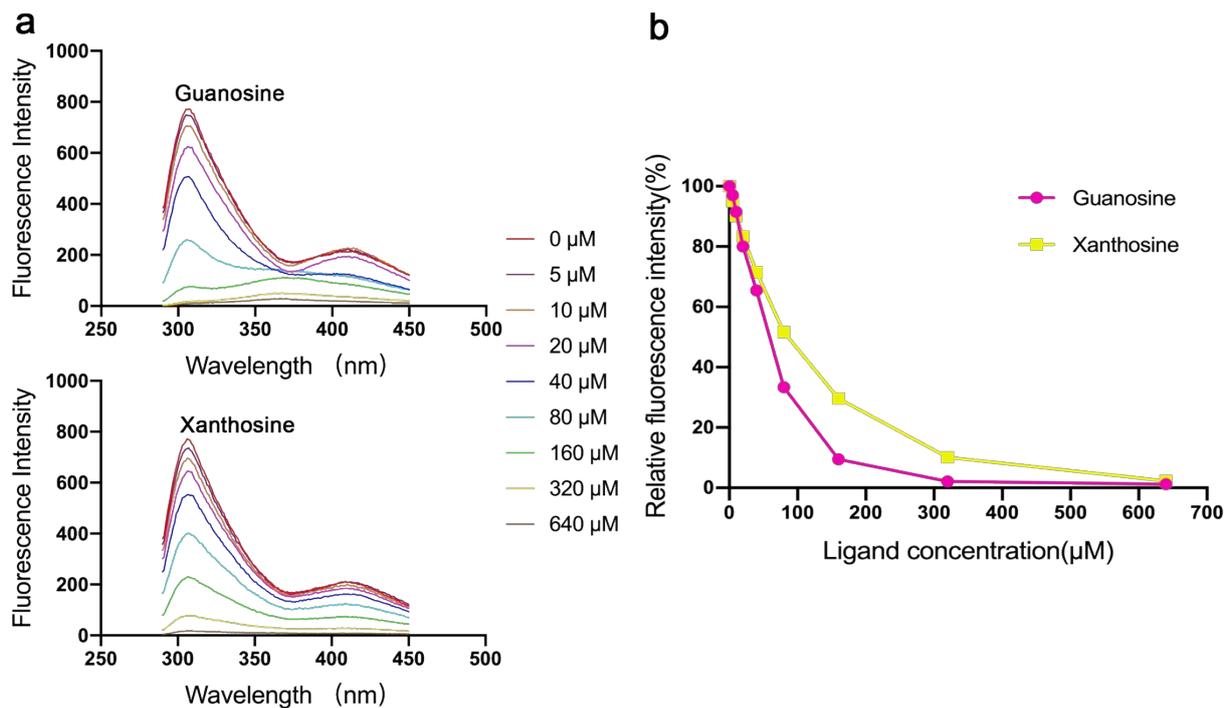
**Fig. S5** The deaminase activity assays for the WT and mutants as analyzed by TLC. The retention positions for guanosine and xanthosine were marked by the arrows on the left side. Gua: guanosine; Xan: xanthosine.



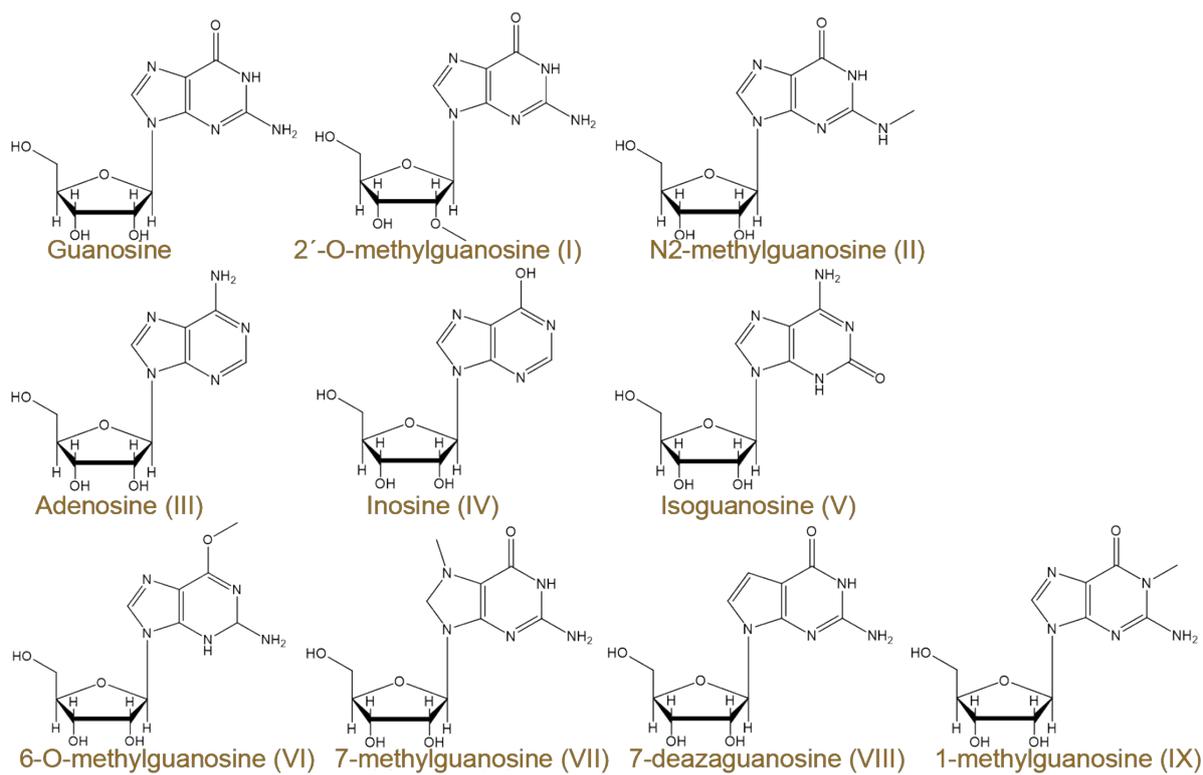
**Fig. S6** The TSA results of the AtGSDA WT and mutants. The vertical axis represented the T<sub>m</sub> values.



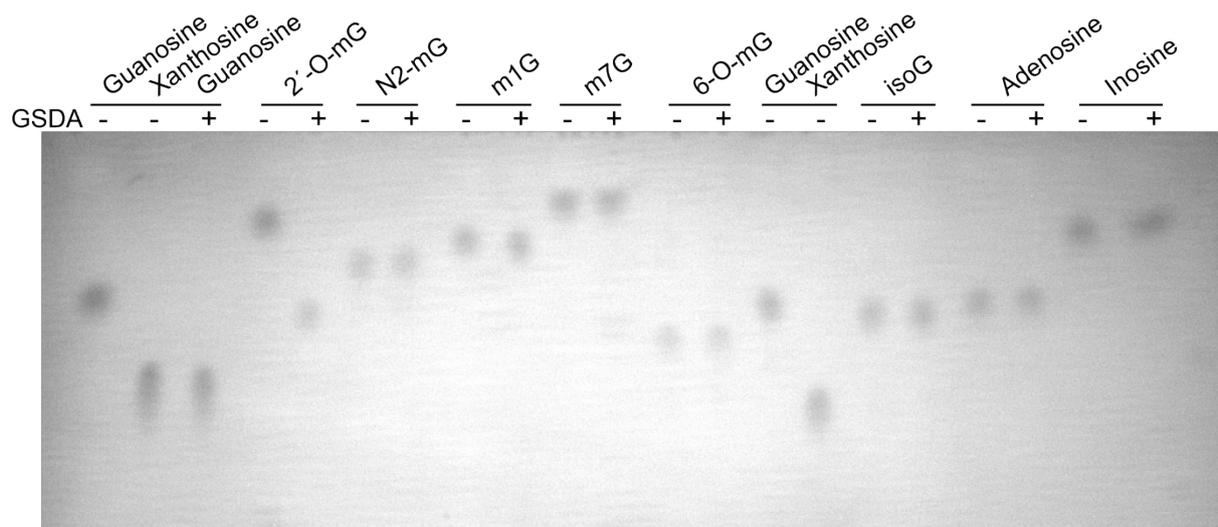
**Fig. S7** The conformational changes of the purine ring before (silver) and after the deamination reaction (yellow). The conserved residues critical to catalysis were shown in sticks and the red arrow signaled the rotation direction of the purine. Wat: the catalytic water; Gua: guanosine; Xan: xanthosine.



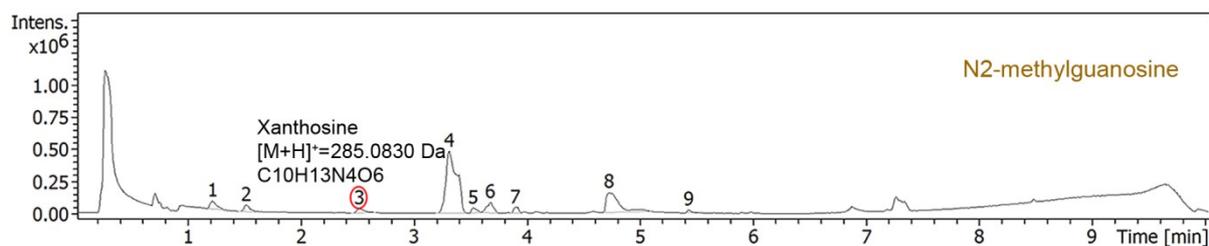
**Fig. S8** The fluorescence quenching of AtGSDA by the titration of guanosine/xanthosine. (a) The emission reduction of the fluorescence at 310 nm by the addition of guanosine (up) and xanthosine (down). Seven concentrations of the ligands were employed and their influences were indicated by the curves of different colors. (b) The relative fluorescence intensities of E82Q protein titrated by xanthosine and guanosine.



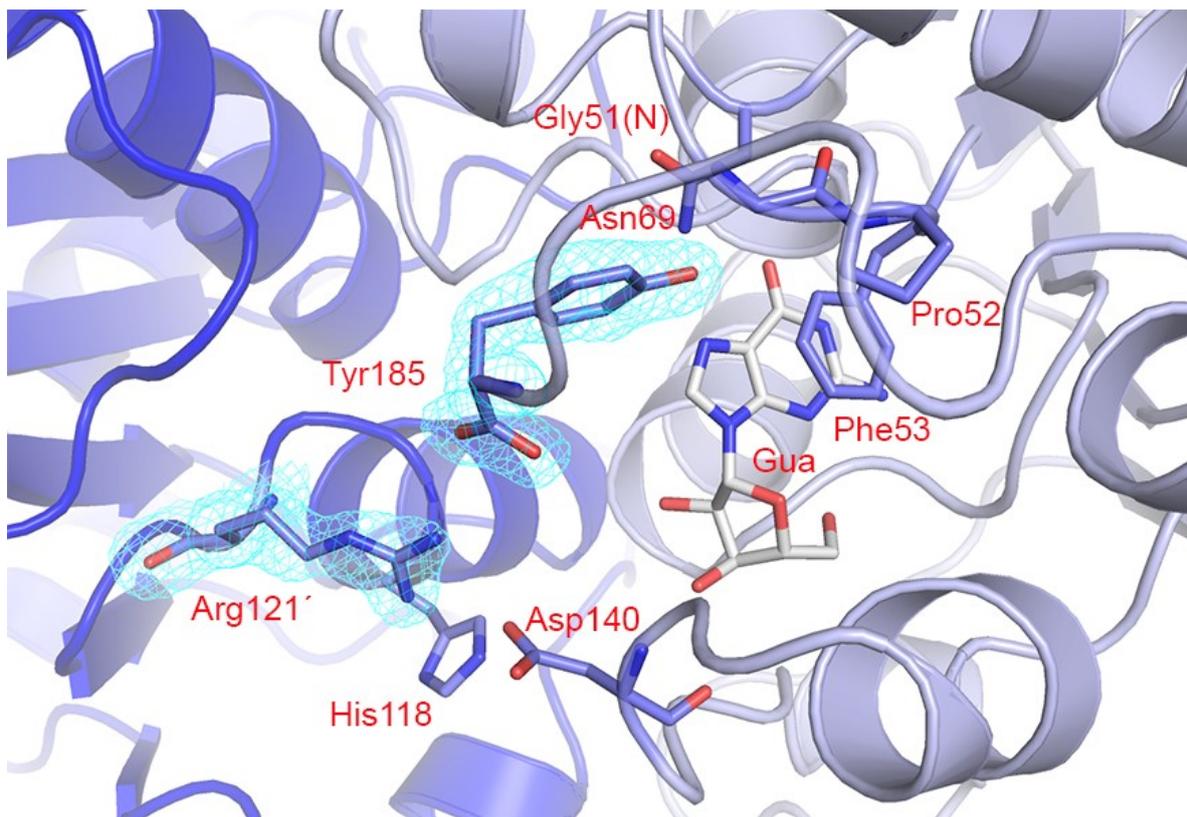
**Fig. S9** The guanosine derivatives tested for deamination possibilities.



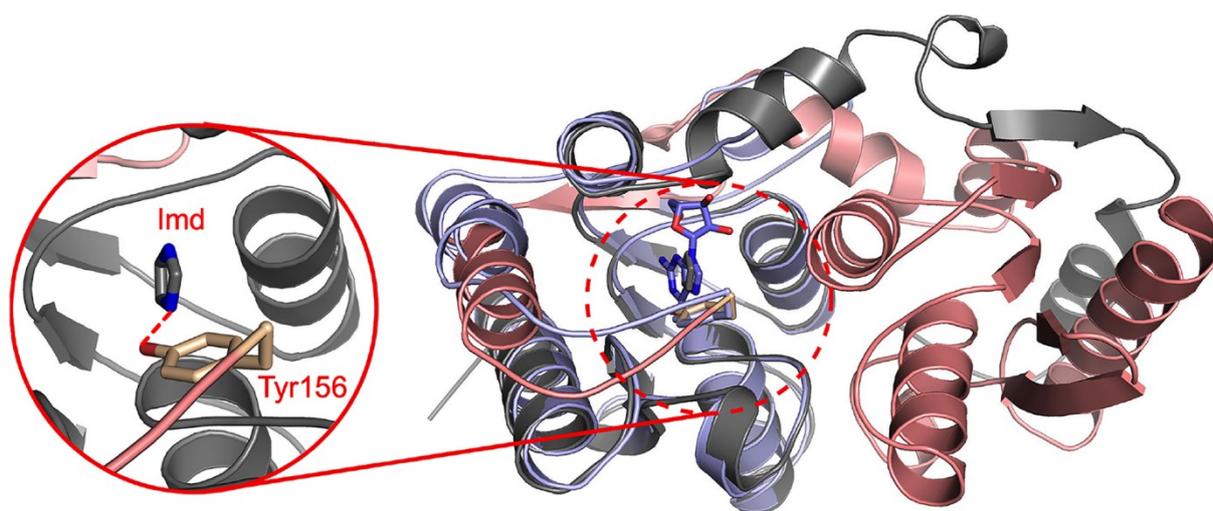
**Fig. S10** The reaction results of AtGSDA with various ligands as resolved by TLC. 2'-O-mG: 2-O'-methylguanosine; N2-mG: N2-methylguanosine; m1G: 1-methylguanosine; m7G: 7-methylguanosine; 6-O-mG: 6-O-methylguanosine; isoG: isoguanosine.



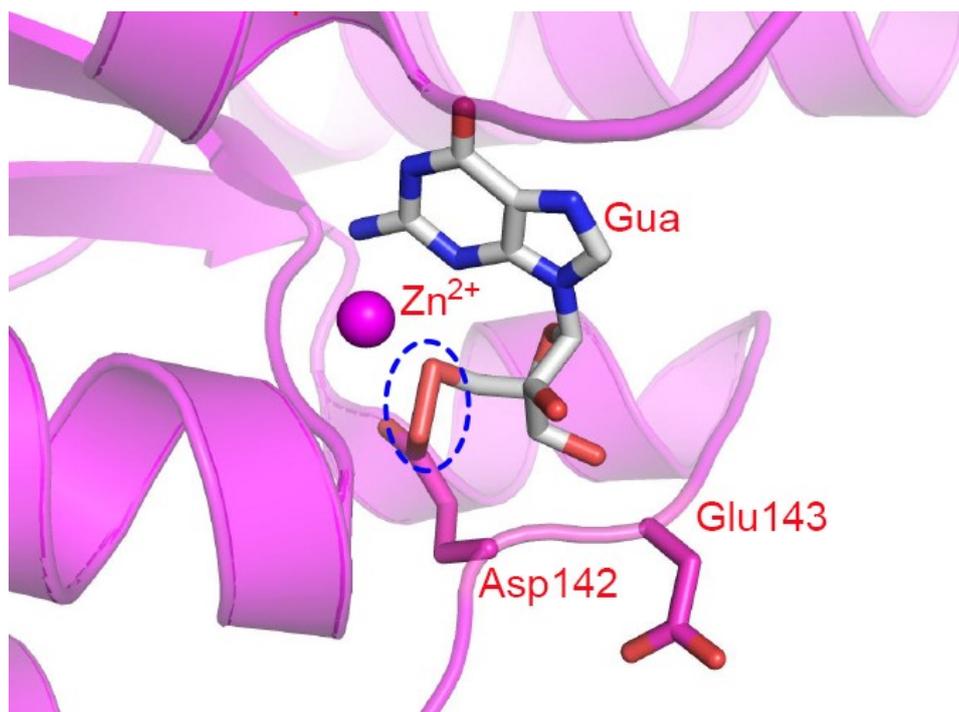
**Fig. S11** The chromatograms of the reaction results of AtGSDA with N2-methylguanosine. The possible peaks containing possible substrates or products (and corresponding molecular weights) were indicated. The red oval in the lower panel indicated the xanthosine peak generated from N2-methylguanosine.



**Fig. S12** The inter- and intra-molecular interaction network that Tyr185 participated in upon the binding of the ligands. The composite omit map was countered at  $2\sigma$  and indicated by the cyan mesh. Gua: guanosine.



**Fig. S13** The structural superimposition of the E82Q-guanosine complex (light blue, only one protomer was shown) to that of the BsGDA-imidazole complex (the two protomers were colored deep salmon and gray respectively). Imd: imidazole. The guanosine and imidazole were shown in sticks and the active site was circled. The key Tyr156-imidazole hydrogen bond was indicated by the red dashed lines in the close-up view and the guanosine molecule was not shown.



**Fig. S14** The structural differences in the substrate-binding pockets and the overall shapes between NE0047 and AtGSDA. The hypothetical model of the NE0047-guanosine complex. The coordinates of guanosine were from that of the superimposed E82Q-guanosine complex (PDB 7DC9). The Asp142Glu143 fragment were shown in sticks, and Asp142 was mistakenly connected to the ribose of guanosine by the pymol program due to their close distance, as indicated by the broken blue oval. Gua: guanosine.

**Table S1.** Kinetic constants for AtGSDA WT and mutants.

Proteins	$K_M$ (mM)	$k_{cat}$ (sec <sup>-1</sup> )	$k_{cat}/K_M$ (mM <sup>-1</sup> sec <sup>-1</sup> )	Relative $k_{cat}/K_M$
WT	0.040	0.336	8.400	1.000
F53L	ND			
N69A	0.045	0.021	0.467	0.056
E82Q	ND			
L119W	ND			
R121A	0.055	0.109	1.982	0.236
Y185F	0.107	0.161	1.505	0.179
R121A/Y185F	0.571	0.213	0.373	0.044
Y185E	0.053	0.083	1.567	0.187
Y185K	0.028	0.054	1.929	0.230
K181	0.068	0.006	0.088	0.010
R183	0.051	0.006	0.118	0.014

\*ND: not determined.

**Table S2.** The primers used in this study. The bases corresponding to the mutated residues were underlined.

Gene	Sense Primers	Antisense Primers
FL	5'- GCGGCAGCCATATGGAAGAAGCTAAAG TGGAAGCA-3'	5'- TTGCACTTGCGGCCGCGTATAAAC GGAACTTCTCCTTT-3'
S29- Y185	5'- GCGGCAGCCATATGAGCGACCATAAAT TCCTAACG-3'	5'- TTGCACTTGCGGCCGCGTATAAAC GGAACTTCTCCTTT-3'
L119W	5'- GGAGCCATCCATT <u>GGT</u> CGAGACTCAAG A-3'	5'- TCTTGAGTCTCGACCAATGGATGGC TCC-3'
E82Q	5'- AACTGCACATGCT <u>CAAGT</u> CACTGCCAT- 3'	5'- ATGGCAGTGACTIONGAGCATGTGCA GTT-3'
N69A	5'- CTAGCTGCCAC <u>GCT</u> ATGGTTTTGAAATA -3'	5'- TATTTCAAACCATAGCGTGGCAGC TAG-3'
F53L	5'- GTGATGGTGGCCCA <u>CTT</u> GGTGC GGTTGA TTG-3'	5'- CAATCACCGCACCAAGTGGGCCAC CATCAC-3'
R121A	5'- AGCCATCCATCTCTCG <u>GCA</u> CTCAAGAG GTTGGTTTA-3'	5'- TAAACCAACCTCTTGAGTGCCGAGA GATGGATGGCT-3'
Y185F	5'- GAGAAGTTCCGTTTTAT <u>TCG</u> CGGCCGCA CT-3'	5'- AGTGC GGCCGCGAATAAACGGAAC TTCTC-3'
Y185E	5'- GCGGCAGCCATATGGAAGAAGCTAAAG TGGAAGCA-3'	5'- TTGCACTTGTCGACTTATTCTAAAC GGAACTTCTCCTT-3'
Y185K	5'- GCGGCAGCCATATGGAAGAAGCTAAAG TGGAAGCA-3	5'- TTGCACTTGTCGACTTAGAATAAAC GGAACTTCTCCTT-3'
K181	5'- GCGGCAGCCATATGGAAGAAGCTAAAG TGGAAGCA-3'	5'- TTGCACTTGTCGACTCACTTCTCCT TTGTGTTCTGGAA-3'
R183	5'- GCGGCAGCCATATGGAAGAAGCTAAAG TGGAAGCA-3'	5'- TTGCACTTGTCGACTCAACGGAAC TCTCCTTTGTGTT-3'

**Table S3.** Data collection and refinement statistics.

	<b>Apo-WT</b>	<b>E82Q-guanosine</b>	<b>WT-xanthosine</b>
<b>PDB ID</b>	7DBF	7DC9	7DCA
Resolution (Å)	28.61-1.90 (1.97-1.90)	50-1.70 (1.76-1.70)	21.67-2.10 (2.21-2.10)
Space group	<i>P6<sub>1</sub></i>	<i>P6<sub>1</sub></i>	<i>P6<sub>1</sub></i>
Cell dimension(Å)			
<i>a, b, c</i> (Å)	119.12, 119.12, 39.45	119.81, 119.81, 39.88	119.35, 119.35, 39.79
$\alpha, \beta, \gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 120
$R_{\text{merge}}$	0.23 (0.78)	0.07 (0.91)	0.17 (0.81)
Redundancy	18.6 (13.8)	19.5 (17.0)	8.6 (5.6)
Completeness (%)	99.5 (97.5)	100 (99.8)	99.9 (100)
$CC_{1/2}$	0.97 (0.90)	1.00 (0.91)	0.99 (0.69)
$I/\sigma(I)$	19.2 (2.5)	34.3 (2.2)	12.1 (2.4)
<b>Refinement</b>			
Resolution range (Å)	28.61-1.90 (1.97-1.90)	28.78-1.70 (1.75-1.70)	21.67-2.10 (2.21-2.10)
No. reflections	25582	36244	19208
$R_{\text{work}}/R_{\text{free}}$ (%)	16.6/19.0	16.2/19.7	17.2/22.0
No. atoms			
Protein	2315	2394	2371
Ligand	2 (Zn <sup>2+</sup> )	2 (Zn <sup>2+</sup> ), 40 (GMP)	2 (Zn <sup>2+</sup> ), 40 (4UO)
Water	140	211	220
B-factor (Å <sup>2</sup> )			
Protein	35.6	31.4	20.5
Ligand	27.6 (Zn <sup>2+</sup> )	22.7 (Zn <sup>2+</sup> ), 28.5 (GMP)	14.0 (Zn <sup>2+</sup> ), 16.2 (4UO)
Water	41.4	40	27
R.m.s deviations			
Bonds (Å)	0.006	0.015	0.004
Angles (°)	0.73	1.31	0.71
Ramachandran favored (%)	97.4	96.8	96.8
Outliers (%)	0.33	0.00	0.00

<sup>a</sup>Values in parentheses are for the highest-resolution shell.

## References

- 1 W.Z. Zhang, J.C. Tang, S.S. Wang, Z.J. Wang, J.H. He, *Nucl. Sci. Tech.*, 2019, **30**, 170-180.
- 2 Z. Otwinowski, W. Minor, *Methods Enzymol*, 1997, **276**, 307-326.
- 3 A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, *J. Appl Crystallogr*, 2007, **40**, 658-674.
- 4 D.K. Wilson, F.B. Rudolph, F.A. Quioco, *Science*, 1991, **252**, 1278-1284.
- 5 P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 213-221.
- 6 P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 486-501.
- 7 A.A. Vaguine, J. Richelle, S.J. Wodak, *Acta Crystallogr D Biol Crystallogr*, 1999, **55**, 191-205.
- 8 H. Muratsubaki, K. Satake, K. Enomoto, *Analytical biochemistry*, 2006, **359**, 161-166.
- 9 J.B. French, Y. Cen, T.L. Vrablik, P. Xu, E. Allen, W. Hanna-Rose, A.A. Sauve, *Biochemistry*, 2010, **49**, 10421-10439.