Supplementary Data

Structure of the mammalian adenine DNA glycosylase MUTYH: insights into the base excision repair pathway and cancer.

Teruya Nakamura[1,2,*], Kohtaro Okabe[1], Shogo Hirayama[1], Mami Chirifu[1], Shinji Ikemizu[1], Hiroshi Morioka[1], Yusaku Nakabeppu[3] and Yuriko Yamagata[1,4]

[1] Graduate School of Pharmaceutical Sciences, Kumamoto University, 5-1 Oehonmachi, Chuo-ku, Kumamoto, 862-0973 Kumamoto, Japan

[2] Priority Organization for Innovation and Excellence, Kumamoto University, 5-1 Oehonmachi, Chuo-ku, Kumamoto, 862-0973 Kumamoto, Japan

[3] Division of Neurofunctional Genomics, Department of Immunobiology and Neuroscience, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

[4] Shokei University and Shokei University Junior College, 2-6-78, Kuhonji, Chuo-ku, Kumamoto, 862-8678 Kumamoto, Japan

* To whom correspondence should be addressed. Tel: +81-96-371-4638; Fax: +81-96-371-4638; Email: tnaka@gpo.kumamoto-u.ac.jp

Table of contents

Table S1. Data collection and refinement statistics.

	MUTYH-DNA complex			CTD-PCNA
	Form I	**From II**	Phasing	**complex**
Data collection				
Wavelength (Å)	1.0	0.98	1.7	0.98
Space group	*I*222	*I*222	*I*222	$P6_3$
Unit-cell lengths (Å)	*a* = 74.3, *b* = 107.2, *c* = 156.4	*a* = 71.7, *b* = 108.5, *c* = 158.5	*a* = 72.1, *b* = 108.8, *c* = 158.5	*a* = *b* = 87.8, *c* = 124.4
Resolution range (Å)	46.89-2.45 (2.49-2.45)	42.53-1.97 (2.00-1.97)	42.53-2.75 (2.80-2.75)	36.40-2.70 (2.77-2.70)
No. of observed reflections	146,029	290,882	223,653	157,014
No. of unique reflections	23,331	43,985	30,921	14,970
Completeness (%)	99.8 (97.2)	99.4 (98.8)	98.8 (97.9)	99.9 (100)
R_{merge} (%)	9.4 (61.9)	5.3 (57.6)	7.7 (59.2)	4.3 (119.1)
$< I/\sigma I >$	23.6 (1.5)	33.1 (1.9)	54.1 (9.3)	29.9 (2.3)
Refinement statistics				
Resolution range (Å)	46.89-2.45	42.53-1.97		36.40-2.70
No. of reflections used	23,315	43,974		14,938
Completeness (%)	99.6	99.2		99.9
R_{work} / R_{free} (%)	19.8/22.1	17.8/19.7		23.5/28.6
R.m.s.d. in bonds (Å)	0.003	0.006		0.002
R.m.s.d. in angles (deg.)	0.585	0.832		0.498
Ramachandran plot				
Favoured (%)	97.0	97.8		91.9
Allowed (%)	3.0	2.2		8.1

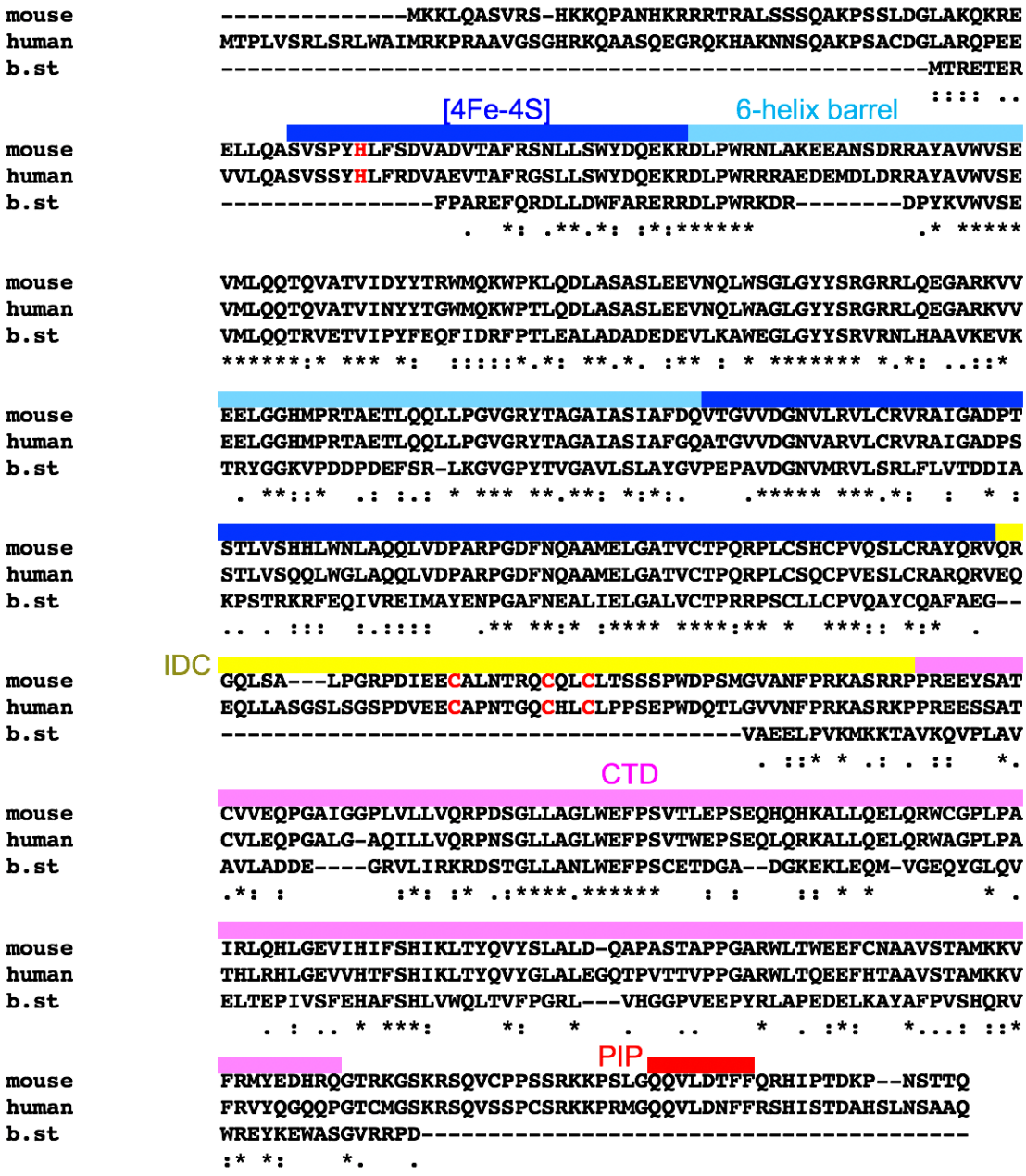

Figure S1. Sequence alignment of MUTYH and *B. stearothermophilus* MutY. Alignment was performed using Clustal W (1). The ligands of the Zn-binding motif are shown in red.

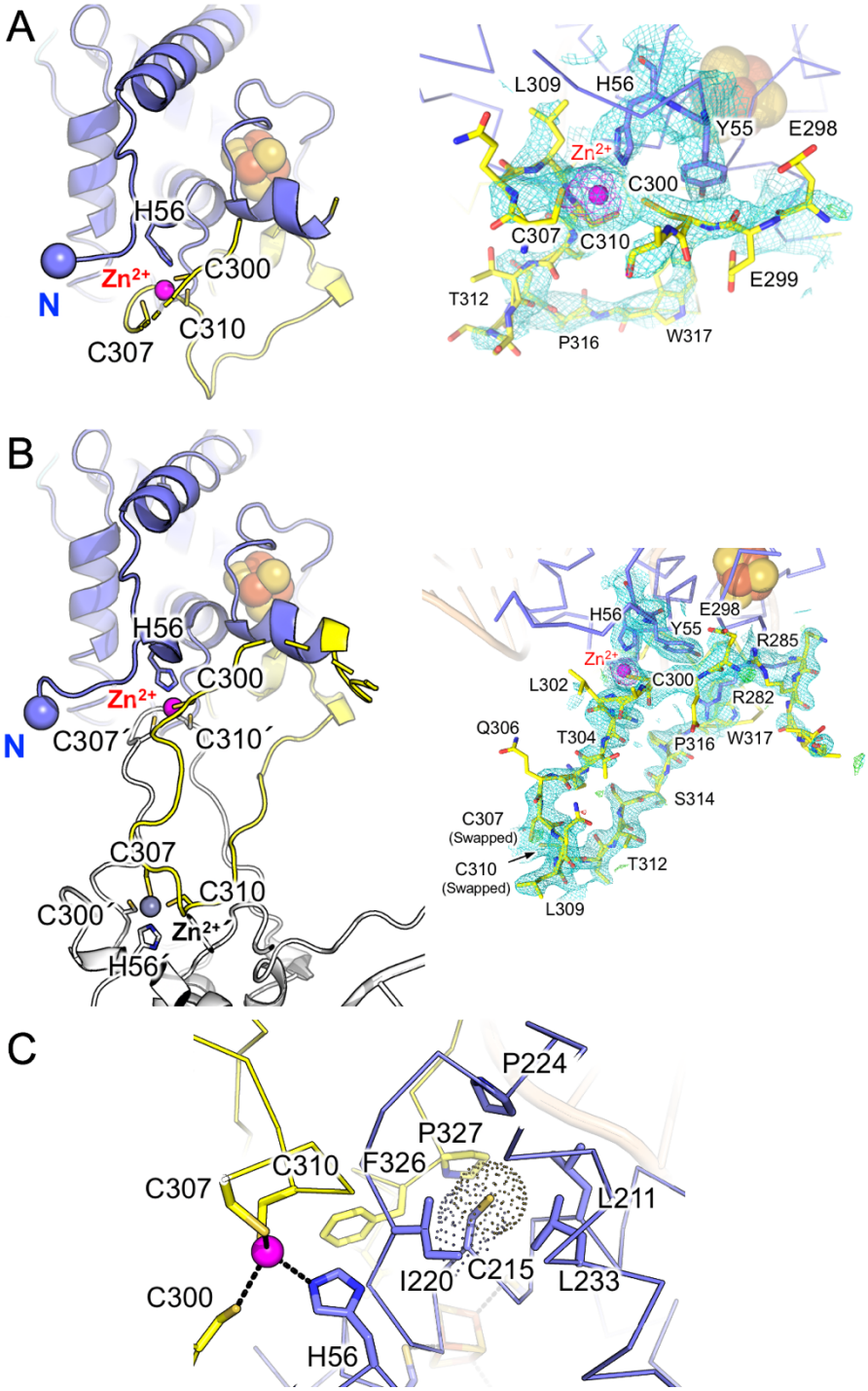

Figure S2. Coordination and electron densities of Zn-binding motif. (A) Form I. The $2F_o - F_c$ map (cyan, 1.0σ), the $F_o - F_c$ map (green and red, ±3.5σ), and the anomalous difference Fourier map calculated using X-ray with a wavelength of 1 Å (pink, 3.5σ) are shown as meshes. (B) Form II. A neighbouring symmetry mate in Form II is shown in white. (C) Cys215 forms a hydrophobic core close to the Zn-binding motif.

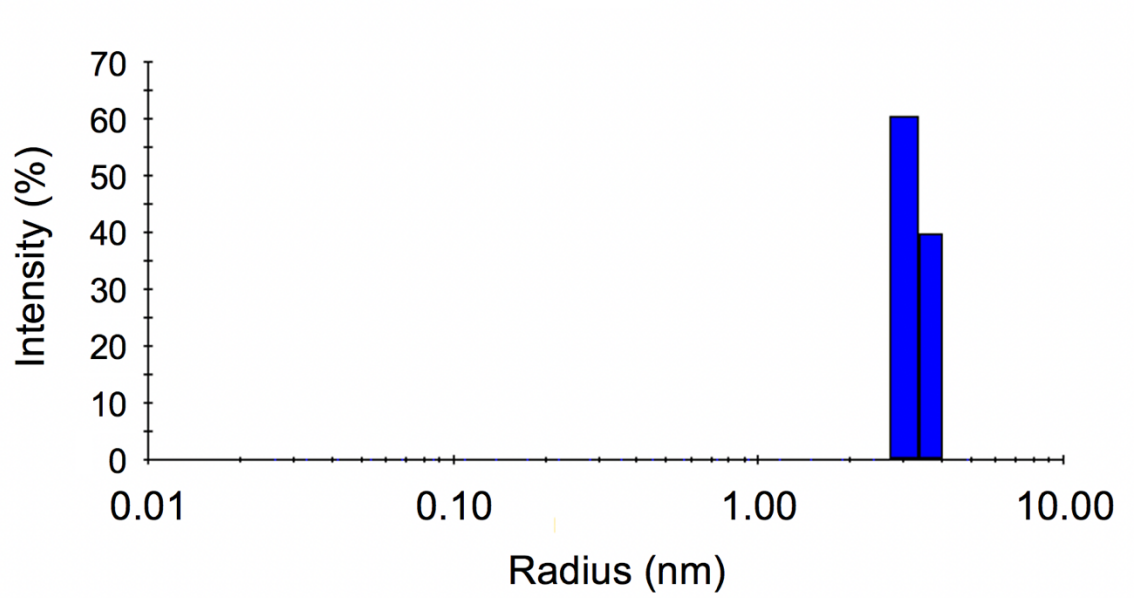

Figure S3. Particle size distribution of the MUTYH-DNA complex (1 mg/mL, at 298 K) by DLS analysis. The complex is monodisperse with an intensity of 100 %, a radius of 3.4 nm, and a polydispersity of 11.9%. The estimated molecular mass of the complex by DLS is 57 k.

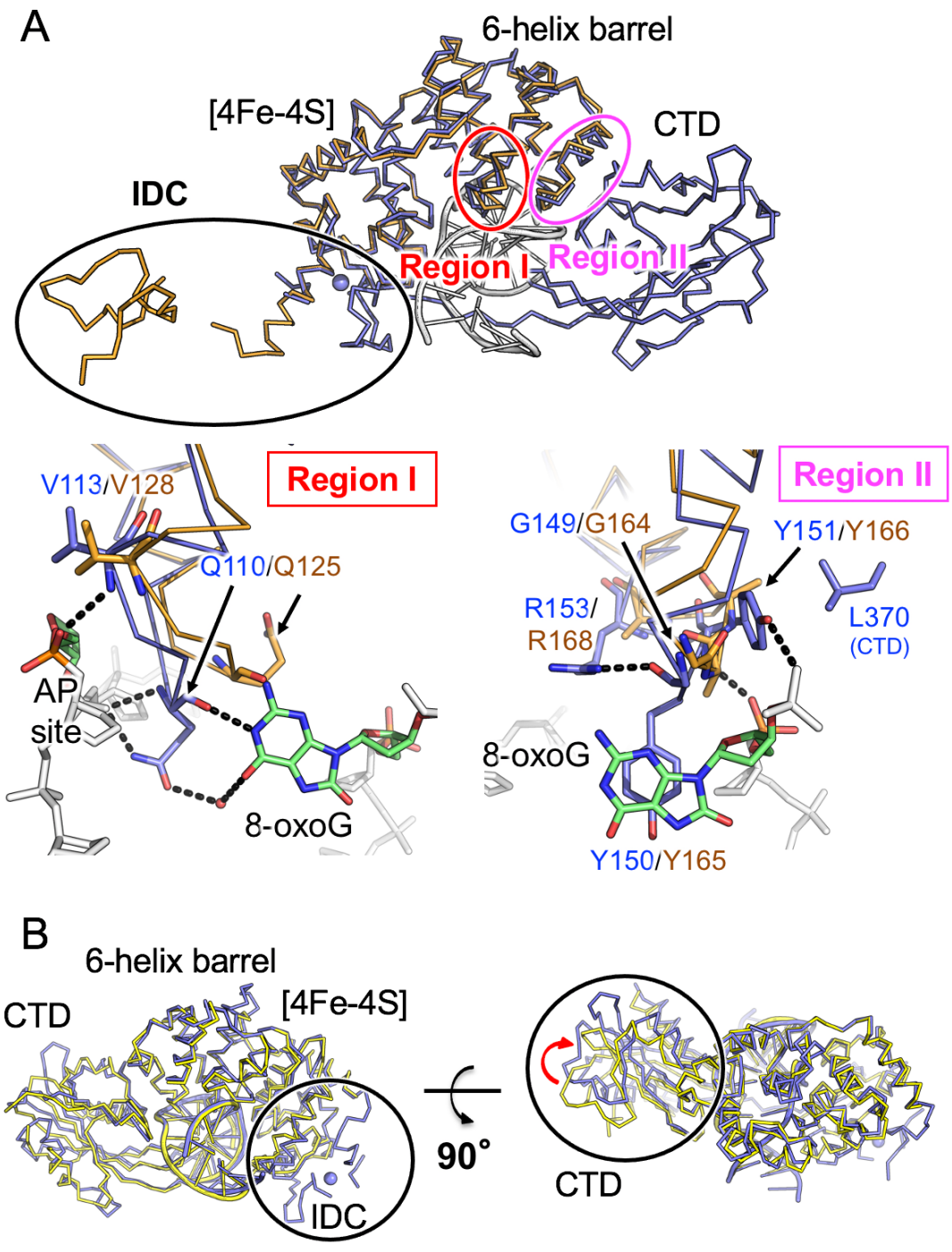

Figure S4. Structural comparison of MUTYH and MutY. (A) Structural comparison between the mouse MUTYH-DNA complex (slate and white) and the NTD of human MUTYH (orange, PDB ID: 3N5N). The IDC and DNA binding regions are indicated by circles. (B) Superposition of the MutY-DNA complex (yellow, PDB ID: 3G0Q) onto the mouse MUTYH-DNA complex (slate).

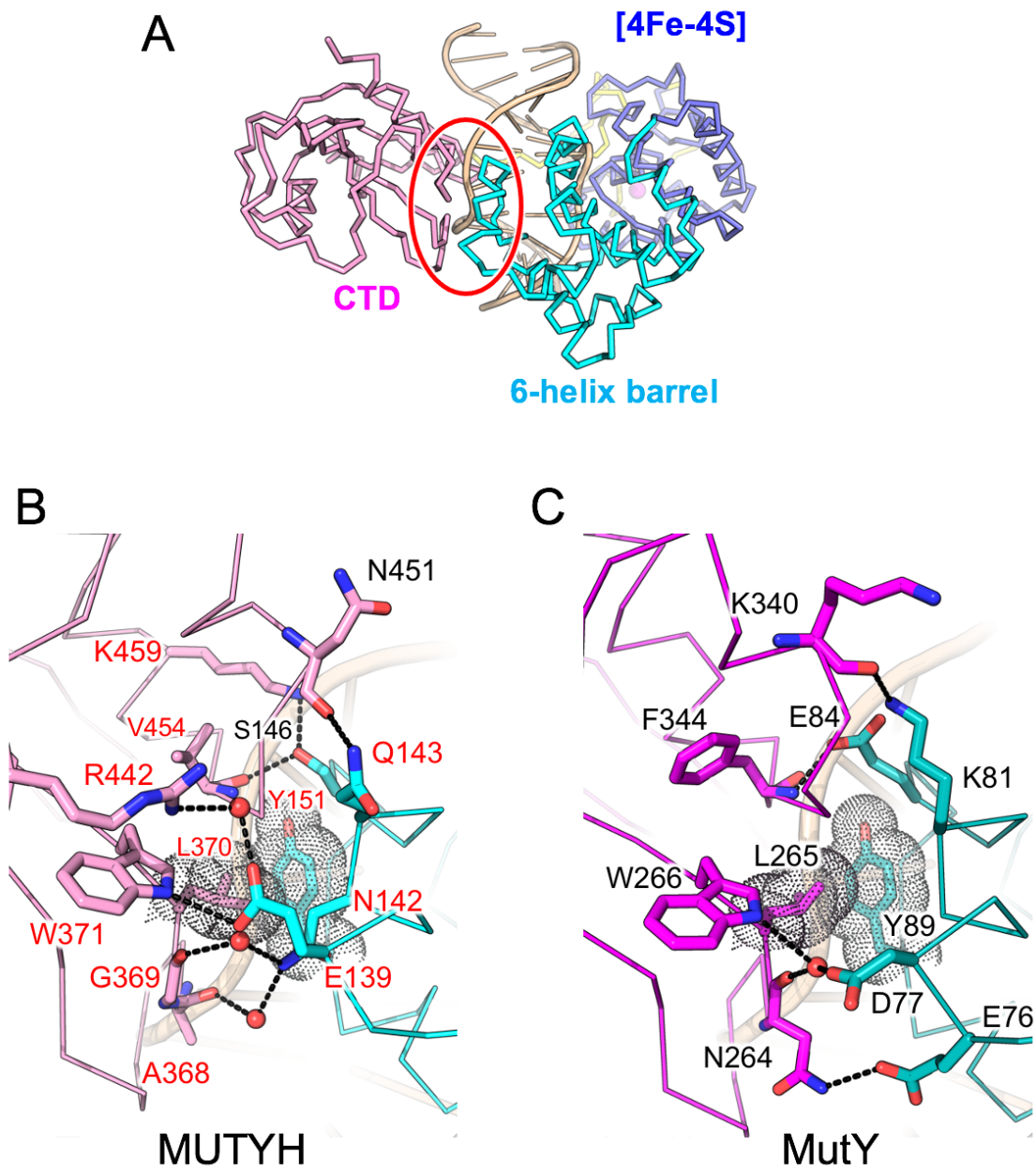

Figure S5. Interactions between the six-helix barrel domain and the CTD. (A) Overall view of the MUTYH-DNA complex. The interacting region is indicated by a red circle. (B) Interactions in MUTYH. The conserved residues between mice and humans are labelled in red. (C) Interactions in MutY (PDB ID: 3G0Q).

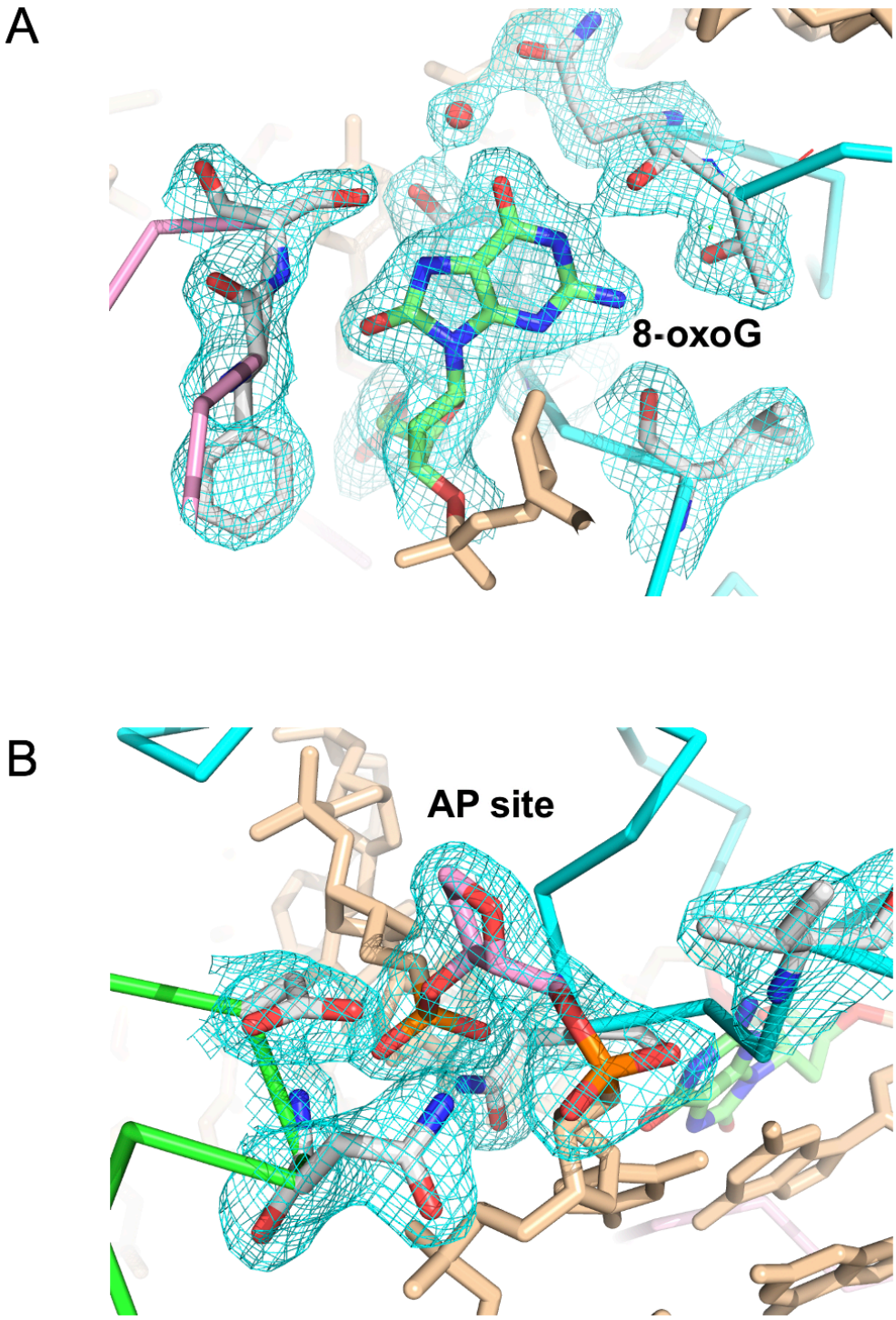

Figure S6. Electron densities of 8-oxoG and AP site. (A) 8-OxoG. The 2 F_o – F_c map (cyan, 1.0σ) and the F_o – F_c map (green and red, ±3.5σ) are shown as meshes. (B) AP site.

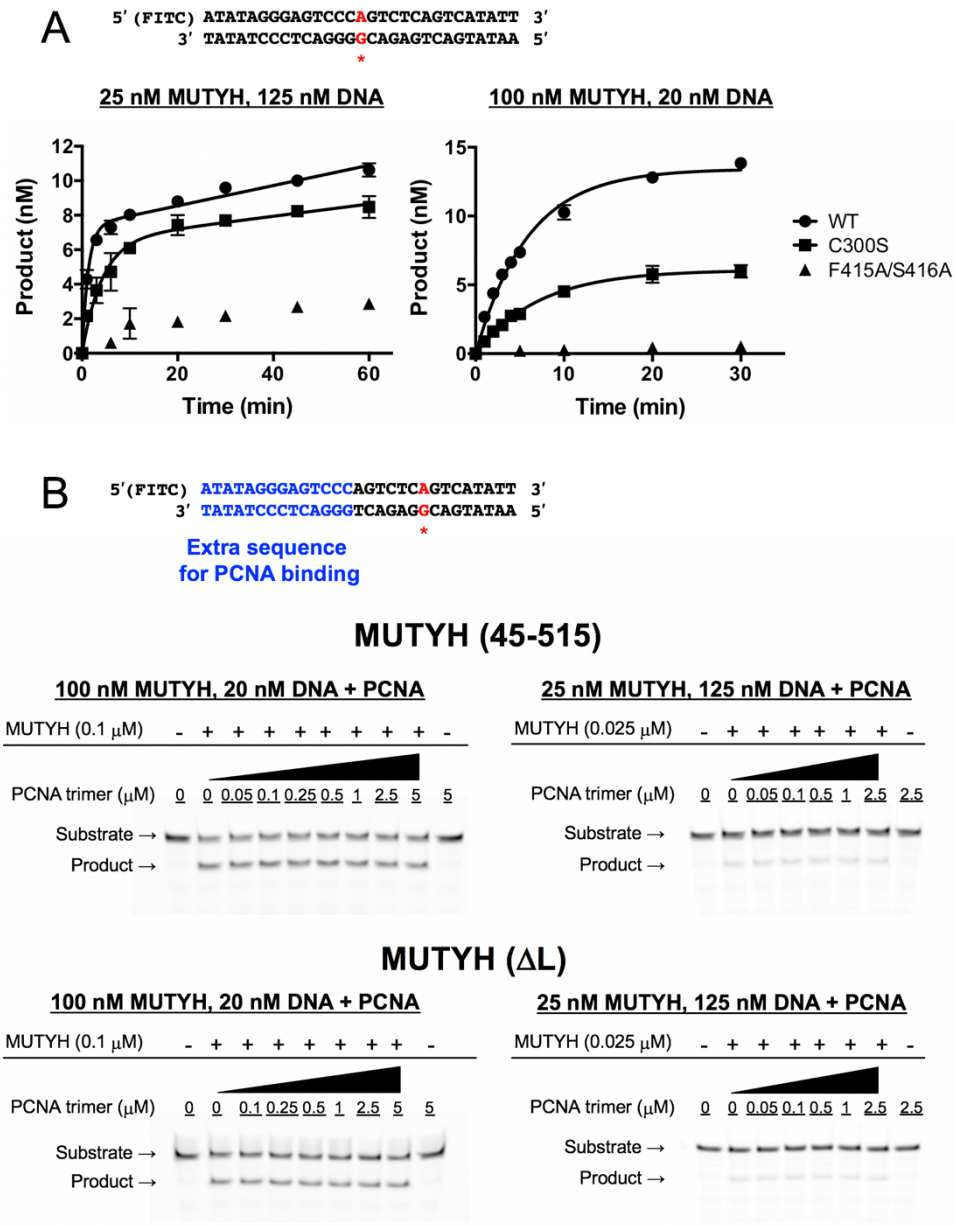

Figure S7. Activity measurement of MUTYH. (A) Adenine DNA glycosylase activity of MUTYH (45–487) (wild type, C300S, and F415A/S416A). G* indicates 8-oxoG. The data represent the mean ± SD of three independent experiments. The k_2 values (min^{-1}) of the wildtype and C300S were 0.17 ± 0.01 and 0.14 ± 0.01, respectively. The k_3 values (min^{-1}) of the wildtype and C300S were 0.008 ± 0.002 and 0.005 ± 0.002, respectively. The k_2 values of MUTYH (45–487) are lower than those of the full-length MUTYH in the previous report (2). The data of F415S/S416A could not be fitted due to its very weak activity. (B) Adenine DNA glycosylase activity of MUTYH (45–515) (wild type and ΔL) in the presence of PCNA.

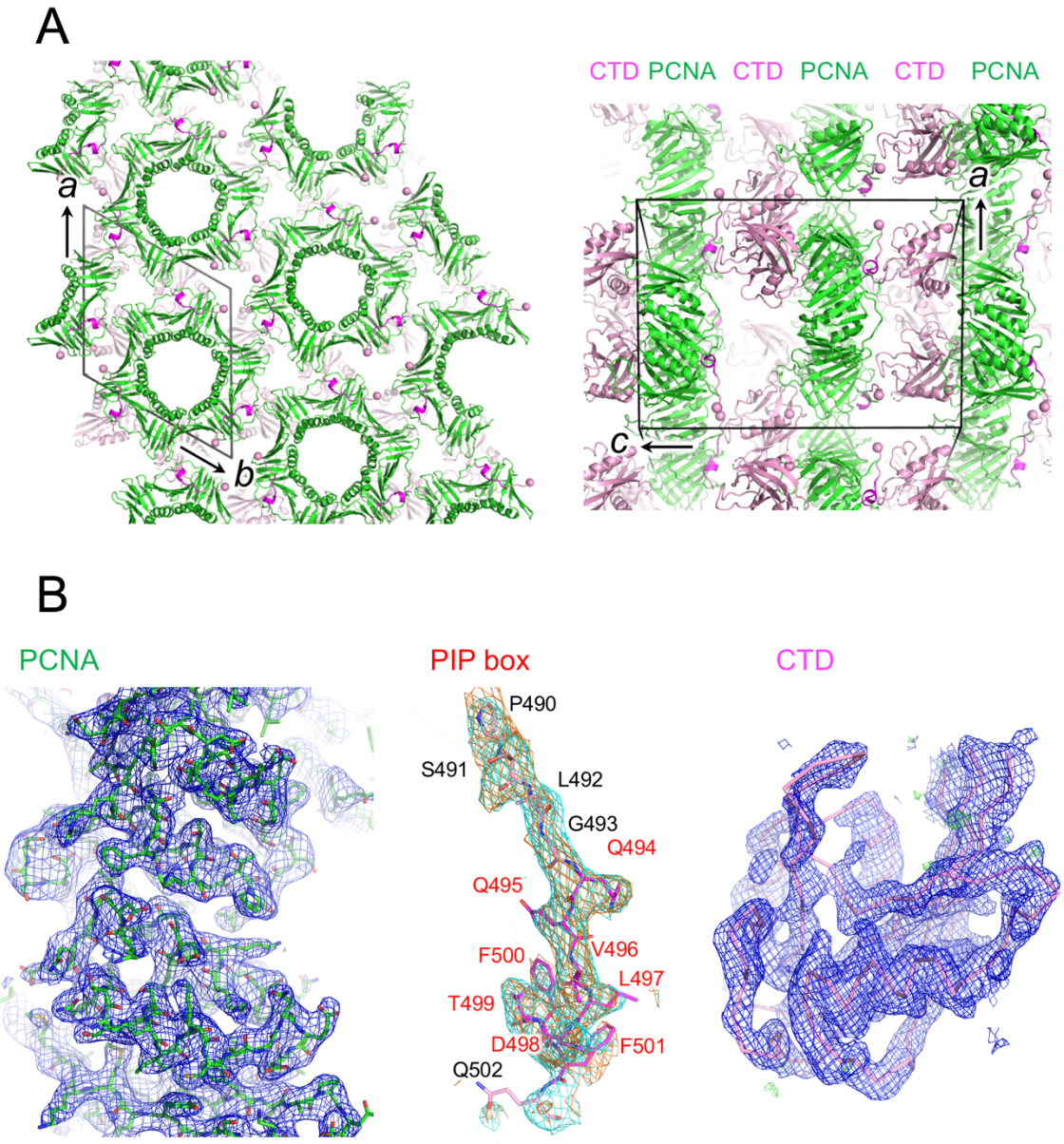

Figure S8. Crystal packing and electron densities of CTD-PCNA. (A) Crystal packing of the CTD-PCNA crystal. (B) Electron densities of PCNA, the PIP box, and the CTD. The 2 $F_o - F_c$ map (blue or cyan, 1.0σ), the $F_o - F_c$ map (green and red, ±3.5σ), and the composite omit map with the anneal method (orange, 1.0σ) are shown as meshes.

References

1. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R., *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
2. Engstrom,L.M., Brinkmeyer,M.K., Ha,Y., Raetz,A.G., Hedman,B., Hodgson,K.O., Solomon,E.I. and David,S.S. (2014) A zinc linchpin motif in the MUTYH glycosylase interdomain connector is required for efficient repair of DNA damage. *J. Am. Chem. Soc.*, **136**, 7829–7832.