Supplementary Information

A New Protein Architecture for Processing Alkylation Damaged DNA: The Crystal Structure of DNA Glycosylase AlkD

Emily H. Rubinson, Audrey H. Metz, Jami O'Quin, and Brandt F. Eichman

Figure Legends

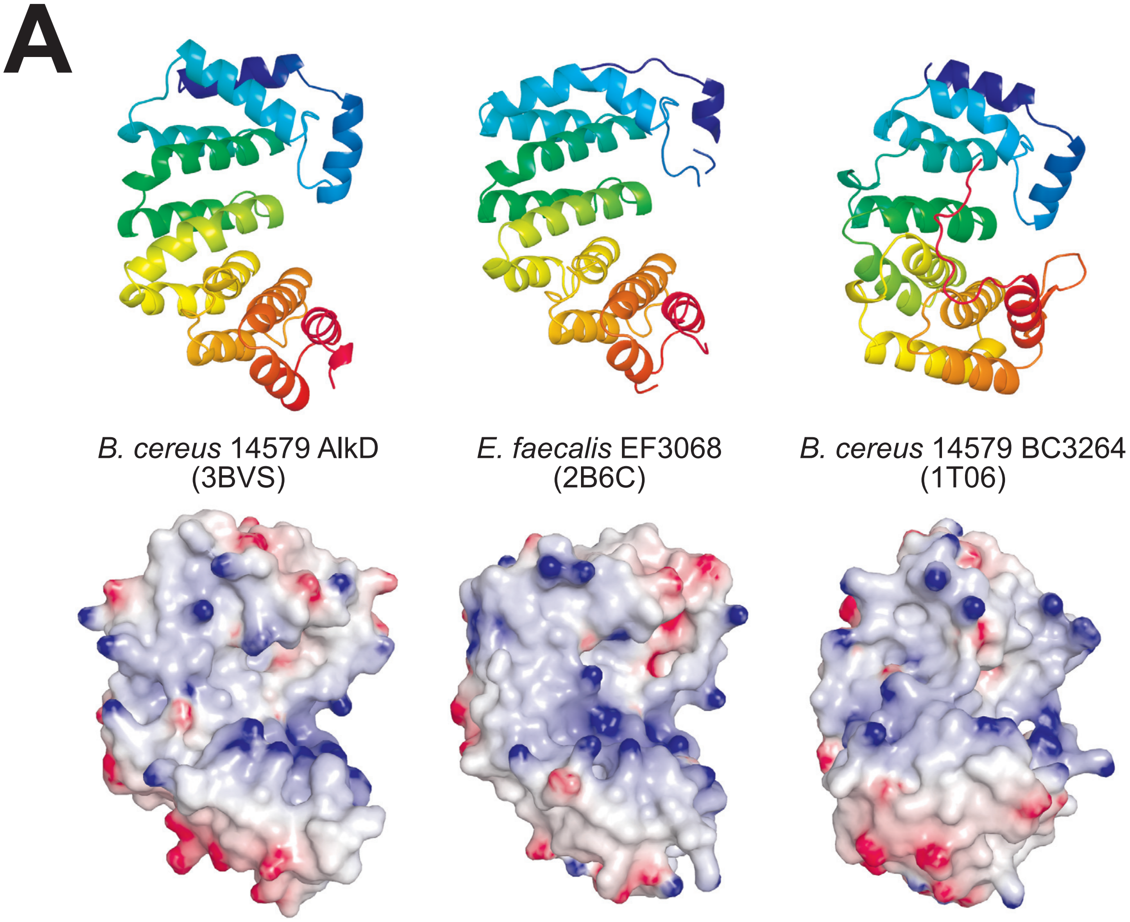
Figure S1. Comparison of the closest AlkD structural orthologs. (A) Ribbon and electrostatic potential surface representations of *B. cereus* AlkD, *E. faecalis* EF3068 (PDB code 2B6C) and *B. cereus* BC3264 (PDB code 1T06) are shown. (B) Superposition of AlkD and EF3068.

Figure S2. Sequence alignment of five putative AlkD homologs. The sequences of *Bacillus cereus* AlkD, *Bacillus anthracis* (NP_847331), *Listeria monocytogenes* (NP_465770), *Enterococcus faecalis* (NP_816680), *Dictyostelium discoideum* (XP_635637), and *Entamoeba histolytica* (XP_654596) are aligned. The secondary structure identified in the crystal structure is shown schematically above the sequence. Aromatic residues inside the putative active site are marked with a yellow triangle, and Asp113 and Arg148 residues important for 7mG excision and DNA binding are marked with a red asterisk. Positively charged residues lining the concave surface are boxed.

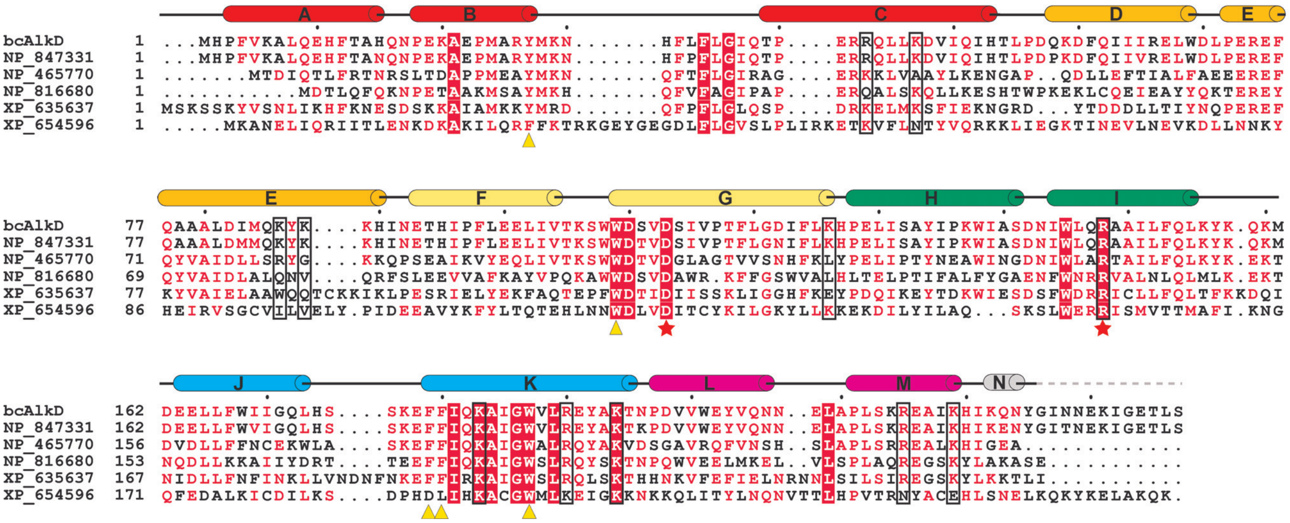
Figure S3. Helical repeat proteins identified by the DALI server as being closely related in structure to AlkD (see Table S1). Shown are HEAT proteins—protein phosphatase 2A (PP2A) 56kD regulatory subunit (2JAK, 2NPP), SCF ubiquitin ligase regulatory subunit Cand1 (1U6G), PP2A 65kD scaffolding subunit (1B3U)—and Armadillo repeat proteins Importin-α (1WA5) and Importin–β (1QBK). Proteins are rendered as ribbon (top) and electrostatic potential surface representations (bottom). Electrostatic potentials (red negative, blue positive, -7 to +7 k_BT) were calculated with the program DelPhi [1]. Note that AlkD is the only structure to contain a positively charged concave surface.

Figure S4. AlkD-DNA binding. Shown is the binding curve for AlkD against tetrahydrofuran (THF)-DNA. Binding of AlkD to abasic DNA was measured by the change in fluorescence anisotropy as AlkD was titrated into a solution containing 25mer oligonucleotide containing an internal THF residue and a 6-carboxyfluorescein moiety at the 3′-end.

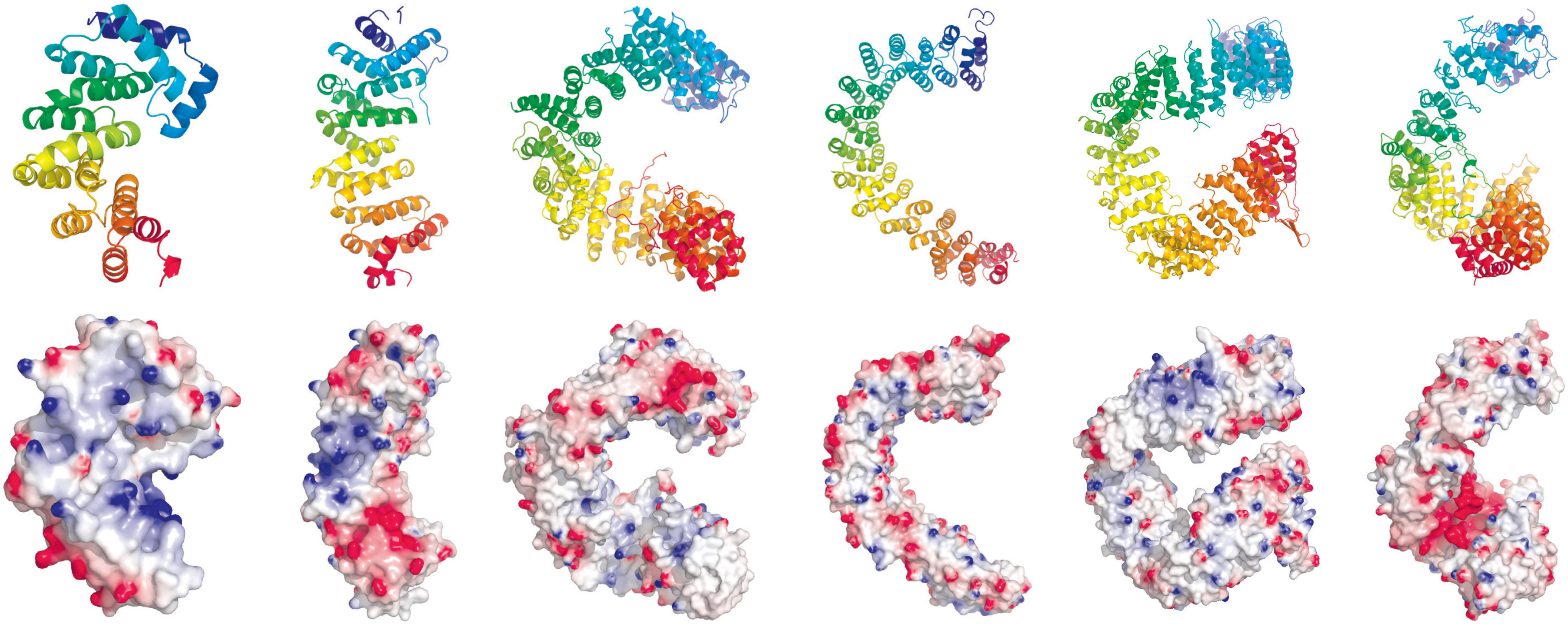
Sup. Figure S1



Sup. Figure S2



Sup. Figure S3



AlkD
3BVS

phosphatase 2A
B56G subunit
2JAK

Cand1
(Cul1/Roc1)
1U6G

phosphatase 2A
PR65/A subunit
1B3U

Importin-α
1WA5

Importin-β
1QBK

Sup. Figure S4

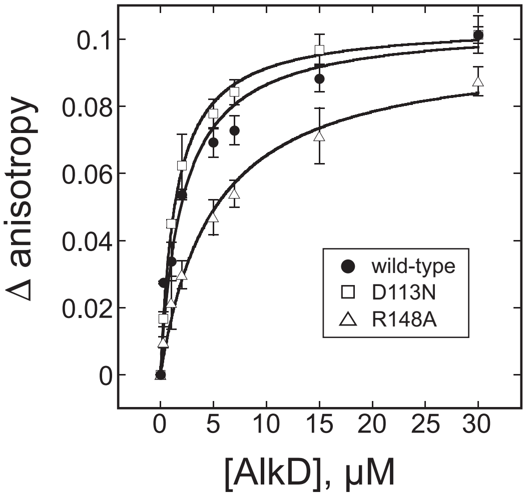


Table S1. Top 20 DALI hits for BcAlkD.

Rank	PDB ID	Chain	Z Score	RMSD	Name	Reference
1	2b6c	A	23.9	2.2	Hypothetical protein, unknown function	[2]
2	1t06	A	13.0	3.8	Hypothetical protein, unknown function	[3]
3	2jak	A	8.6	6.3	Protein phosphatase 2A 56kD regulatory subunit	[4]
4	1u6g	C	8.5	3.9	Cullin interacting protein, Cand1	[5]
5	1b3u	A	8.1	3.5	Protein phosphatase 2A 65kD scaffolding subunit	[6]
6	2db0	A	7.8	2.9	253aa long hypothetical protein	[7]
7	2npp	B	7.6	6.2	Protein phosphatase 2A 56KDa subunit (holoenzyme)	[8]
8	1wa5	C	7.5	4.1	Importin-α	[9]
9	1qbk	B	7.2	4.2	Importin-β2	[10]
10	2vgl	A,B	7.0	8.9	Adaptor-related protein complex 2 (AP2) α2 subunit	[11]
11	1qgr	A	6.4	4.2	Importin-β subunit	[12]
12	1dvp	A	6.1	3.8	Membrane trafficking protein Hrs	[13]
13	3bct	A	6.0	4.5	β-catenin (*Mus musculus*)	[14]
14	1oyz	A	6.0	3.6	Hypothetical protein, unknown function	[15]
15	2fv2	A	5.9	7.3	Required for cell differentiation protein, Rcd-1	[16]
16	1ee4	A	5.9	4.7	Importin-α	[17]
17	2gl7	A	5.7	9.4	β-catenin	[18]
18	2ilr	A	5.5	7.9	Fanconi Anemia Group E protein, FANCE	[19]
19	1upk	A	5.3	5.6	Mouse protein 25α, MO25α	[20]
20	1ft1	A	5.3	3.3	Protein farnesyltransferase, FTase	[21]

References to Supplementary Information

1. Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A. & Honig, B. (2002). Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J Comput Chem* **23**, 128-37.
2. Osipiuk, J., Hatzos, C., Moy, S., Collart, F. & Joachimiak, A. X-ray structure of predicted DNA alkylation repair enzyme from *Enterococcus faecalis*. *PDB ID: 2B6C-A*.
3. Zhang, R., Wu, R., Moy, S. & Joachimiak, A. 1.9 A crystal structure of a hypothetical protein from *Bacillus cereus* ATCC 14579. *PDB ID: 1T06-A*.
4. Magnusdottir, A., Stenmark, P., Flodin, S., Nyman, T., Kotenyova, T., Nilsson, P., Ogg, D. & Nordlund, P. Crystal Structure of the Human Pp2A Regulatory Subunit, B56G. *PDB ID: 2JAK-A*.
5. Goldenberg, S. J., Cascio, T. C., Shumway, S. D., Garbutt, K. C., Liu, J., Xiong, Y. & Zheng, N. (2004). Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. *Cell* **119**, 517-28.
6. Groves, M. R., Hanlon, N., Turowski, P., Hemmings, B. A. & Barford, D. (1999). The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* **96**, 99-110.
7. Handa, N., Nishino, A., Kishishita, S., Murayama, K., Shirouzu, M. & Yokoyama, S. Crystal structure of PH0542. *PDB ID: 2DB0-A*.
8. Xu, Y., Xing, Y., Chen, Y., Chao, Y., Lin, Z., Fan, E., Yu, J. W., Strack, S., Jeffrey, P. D. & Shi, Y. (2006). Structure of the protein phosphatase 2A holoenzyme. *Cell* **127**, 1239-51.
9. Matsuura, Y. & Stewart, M. (2004). Structural basis for the assembly of a nuclear export complex. *Nature* **432**, 872-7.
10. Chook, Y. M. & Blobel, G. (1999). Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. *Nature* **399**, 230-7.
11. Collins, B. M., McCoy, A. J., Kent, H. M., Evans, P. R. & Owen, D. J. (2002). Molecular architecture and functional model of the endocytic AP2 complex. *Cell* **109**, 523-35.
12. Cingolani, G., Petosa, C., Weis, K. & Muller, C. W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* **399**, 221-9.
13. Mao, Y., Nickitenko, A., Duan, X., Lloyd, T. E., Wu, M. N., Bellen, H. & Quiocho, F. A. (2000). Crystal structure of the VHS and FYVE tandem domains of Hrs, a protein involved in membrane trafficking and signal transduction. *Cell* **100**, 447-56.
14. Huber, A. H., Nelson, W. J. & Weis, W. I. (1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**, 871-82.
15. Kuzin, A., Semesi, A., Shkarina, T., Savchenko, A., Arrowsmith, C. & Tong, L. X-ray structure of YIBA *E. coli* Northeast Structural Genomics Consotrium target ET31. *PDB ID: 1OYZ-A*.
16. Garces, R. G., Gillon, W. & Pai, E. F. (2007). Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16**, 176-88.
17. Conti, E. & Kuriyan, J. (2000). Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure* **8**, 329-38.

18. Sampietro, J., Dahlberg, C. L., Cho, U. S., Hinds, T. R., Kimelman, D. & Xu, W. (2006). Crystal structure of a beta-catenin/BCL9/Tcf4 complex. *Mol Cell* **24**, 293-300.
19. Nookala, R. K., Hussain, S. & Pellegrini, L. (2007). Insights into Fanconi Anaemia from the structure of human FANCE. *Nucleic Acids Res* **35**, 1638-48.
20. Milburn, C. C., Boudeau, J., Deak, M., Alessi, D. R. & van Aalten, D. M. (2004). Crystal structure of MO25 alpha in complex with the C terminus of the pseudo kinase STE20-related adaptor. *Nat Struct Mol Biol* **11**, 193-200.
21. Park, H. W., Boduluri, S. R., Moomaw, J. F., Casey, P. J. & Beese, L. S. (1997). Crystal structure of protein farnesyltransferase at 2.25 angstrom resolution. *Science* **275**, 1800-4.