Electronic Supplementary Material (ESI) for RSC Advances.
This journal is © The Royal Society of Chemistry 2017

SUPPLEMENTAL DATA

Structural Basis for Substrate Discrimination by *E. coli* Repair Enzyme, AlkB

Namrata Jayanth[1], Nirmala Ogirala[1], Anil Yadav[2] and Mrinalini Puranik[1]*

[1] National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bellary Road, Bangalore 560065, India
[2] Indian Institute of Science Education and Research (IISER), Pune, Maharashtra 411008, India
* To whom correspondence should be addressed. Mrinalini Puranik, Hindustan Unilever Research Centre, No. 62, Main Road, Whitefield, Bangalore 560066, India, Phone: 91-9980633836
E-mail: puranik.mrinalini@gmail.com, mrinalini.puranik@unilever.com

SUPPLEMENTAL RESULTS

AlkB does not bind the neutral, imino form of 1-me-dAMP

Raman excitation wavelength of 225 nm is specific for enhancement of the modes of the imino form[1]. The most intense bands in the 1-me-dAMP spectrum at 225 nm excitation correspond to the imino form at 1442 cm^{-1} (N6–H bending and N1–C6 stretching modes) and 1422 cm^{-1} (C2–H stretching and CH_3 bending vibrational modes)[1] as shown in (**Supplementary Fig. 4B**). The intensities of these bands decrease in the presence of AlkB. While no change in band position is observed, the reduction in intensities seen as negative bands in the difference spectrum (**Supplementary Fig. 4D**) implies a decrease in the population of the imino form.

We interpret this decrease in the imino population as a shifting of the equilibrium between the amino-imino forms of 1-me-dAMP because of preferential binding of the enzyme to the amino form. This observation further strengthens our hypothesis of higher affinity for the protonated amino form of 1-me-dAMP by AlkB in preference to the neutral imino form.

Interaction of AlkB with nucleotides flanking methylated substrates

The crystal structure of AlkB with the 1-me-dAMP trimer reveals likely interactions between the two thymine residues flanking 1-me-dAMP. The C2-O group of the dT residue at the 3'-end of 1-me-dAMP as seen from the crystal structure[2] interacts with Tyr 55 and the methyl group of the dT residue 5' to 1-me-dAMP forms contacts with Ser 129. UVRR spectroscopy depicts the changes in the thymine residues adjacent to the methylated substrates. This is particularly evident in the case of 3-me-dCMP containing trimer (**Fig. 5**) rather than the 1-me-dAMP containing trimer as most of the dT bands are masked by the 1-me-dAMP bands occurring in similar positions. The interactions are manifested as downshifts of bands at 1483 $cm^{-1}$and 1203 cm^{-1} corresponding to the methyl group bending mode coupled with pyrimidine ring vibrations in water. In D_2O, wavenumber shifts are observed for 1479 cm^{-1}, 1243 cm^{-1} and 1155 cm^{-1} bands, all corresponding to methyl group bending vibrations (**Fig. 5**).

Crystal structures of AlkB with methylated DNA show that AlkB establishes contacts with the phosphate backbone but not with the bases further than those immediately flanking the methylated base[3, 4]. Vibrational spectrum of AlkB with a longer substrate, a pentamer, 5'-dC-dT-(1-me-dAMP)-dT-dC-3' (**Supplementary Fig. 9**), shows no observable perturbation in bands corresponding to dC located at the two ends of the pentamer. Hence the difference in rates of catalysis reported for longer substrates with AlkB[4] is probably due to the phosphate backbone interactions of the enzyme with oligomers.

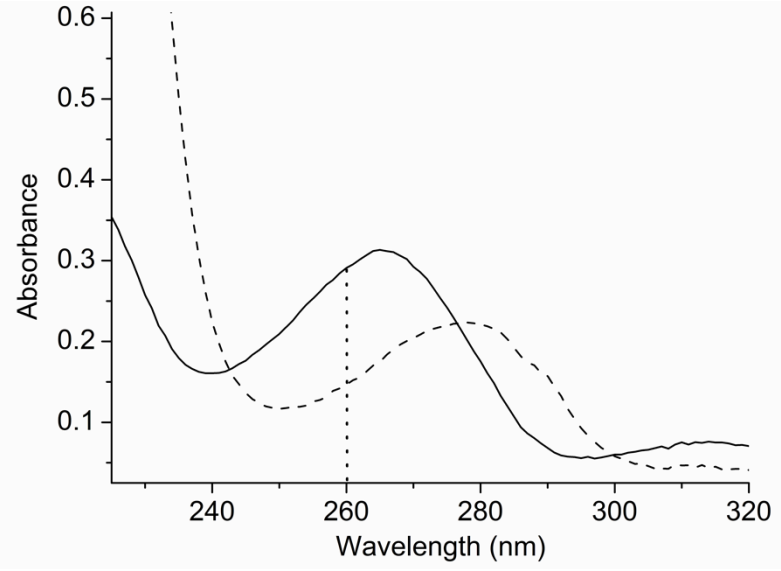
Co-factors modulate the conformation of AlkB suitable for substrate binding

Fe^{2+} and 2-oxoglutarate are the co-factors that regulate the catalytic mechanisms of the dioxygenase enzyme family. AlkB undergoes significant conformational dynamics throughout the process of binding to the enzyme and the release of product as seen from the solution-state NMR spectroscopic studies. AlkB appears to modulate its conformation for suitable substrate binding on addition of co-factors. The significance of co-factors in

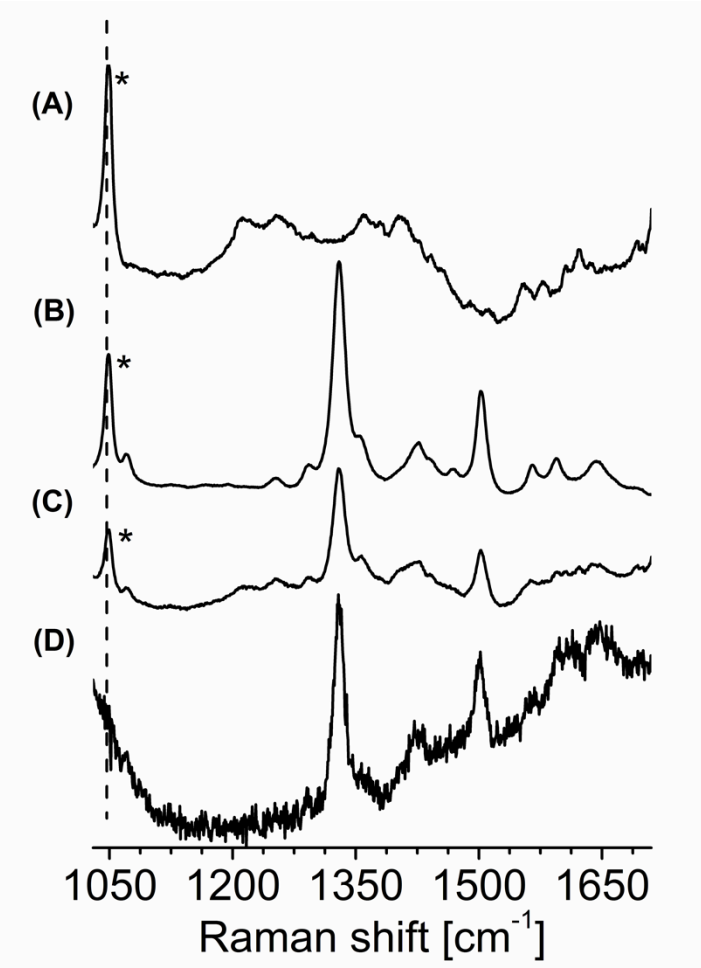
regulating substrate binding is seen from our experiments which demonstrate that co-factors are required for AlkB to bind methylated nucleotides.

As seen from **Supplementary Fig. 5B**, no changes are observed for the amino group of 1-me-dAMP upon addition of AlkB. There is also no reduction in intensity of bands or wavenumber shifts of the ring modes in the absence of co-factors in the difference spectrum obtained. This indicates that 1-me-dAMP does not have any strong interaction with AlkB in the absence of cofactors.

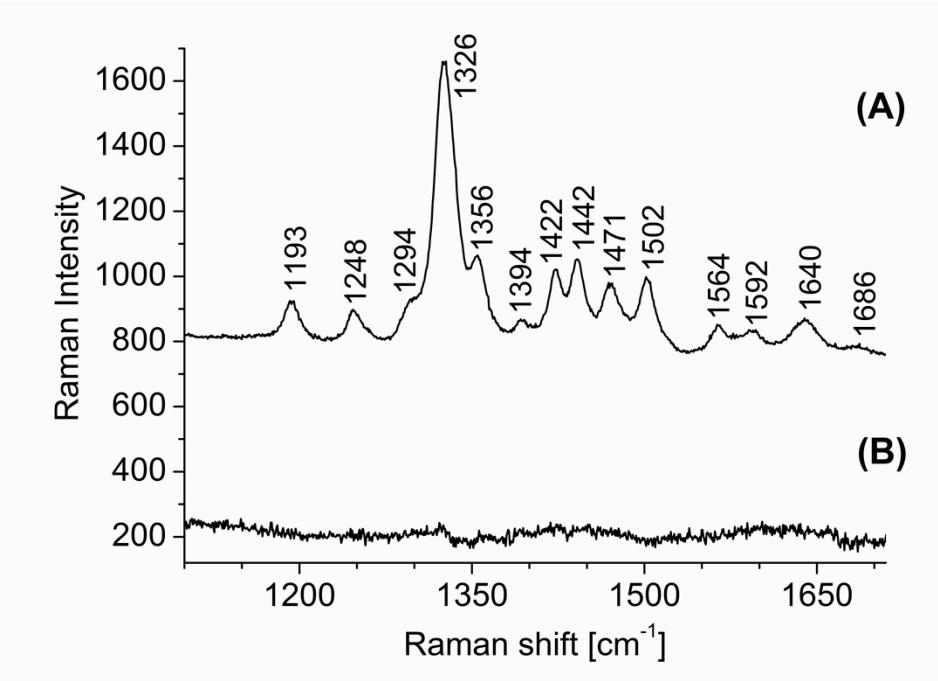
SUPPLEMENTAL FIGURES



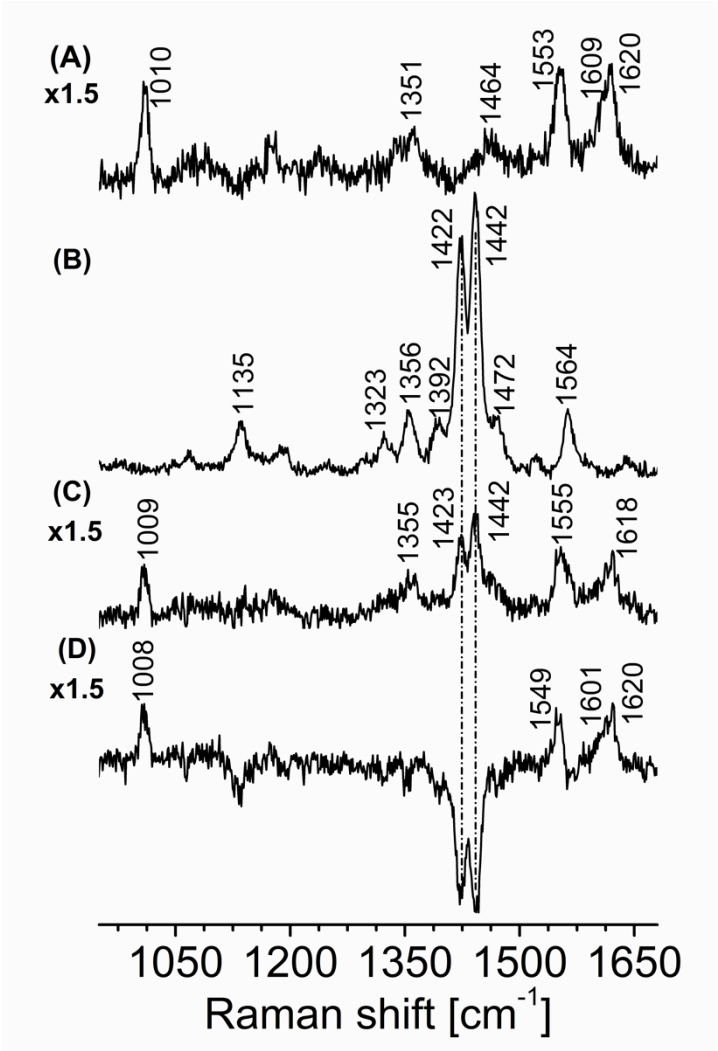
Supplementary Fig. S1. UV absorption spectra of 10 μM AlkB (dashed line) and 1 mM 1-me-dAMP (solid line) in 50 mM Hepes buffer, pH 8.0. The laser excitation wavelength used (λ_{exc}= 260 nm) is depicted as a dotted line.



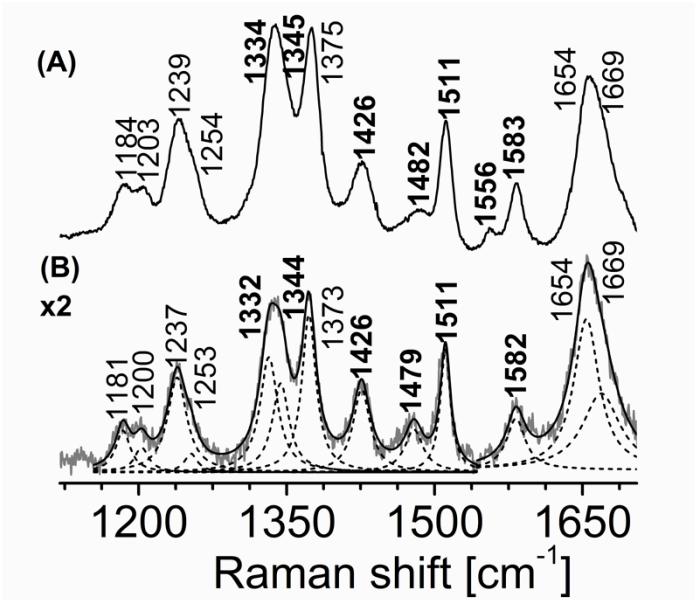
Supplementary Fig. S2. Subtraction protocol employed to remove contributions of unbound 1-me-dAMP and AlkB from AlkB•1-me-dAMP complex. Shown above are the spectra of **(A)** 200 μM AlkB in reaction buffer with 30 mM $NaNO_3$, **(B)** 1 mM 1-me-dAMP in reaction buffer containing 30 mM $NaNO_3$, **(C)** AlkB•1-me-dAMP complex in reaction buffer containing 30 mM $NaNO_3$ and **(D)** Difference spectrum of AlkB•1-me-dAMP complex with contributions from AlkB and 1-me-dAMP removed. The 1049 cm^{-1} band of $NaNO_3$ is marked with an asterisk.



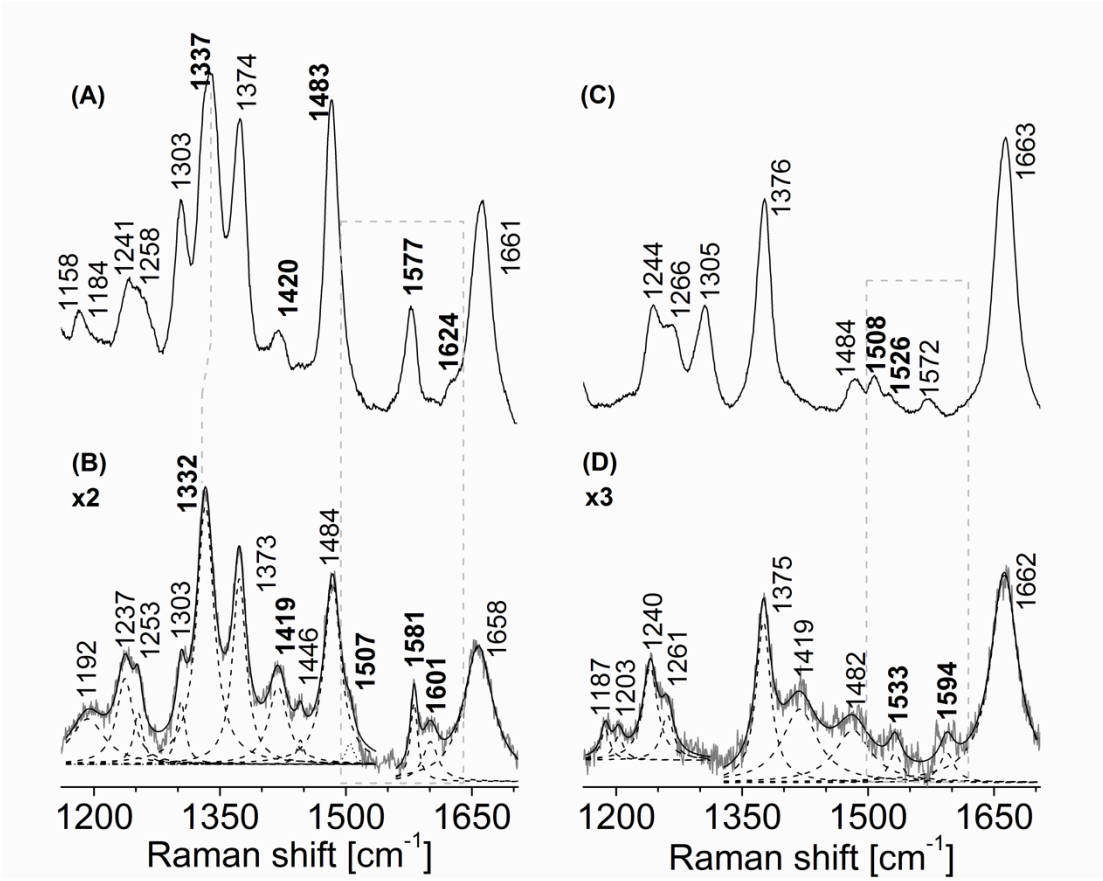
Supplementary Fig. S3. Resonance Raman spectra of 1-me-dAMP at pH 8.0 (imino form) with AlkB. **(A)** 1-me-dAMP in Hepes buffer, pH 8.0 (H_2O) and **(B)** Difference spectrum of AlkB•1-me-dAMP complex with contributions from unbound AlkB and unbound 1-me-dAMP removed. The spectra were obtained using laser excitation of 260 nm.



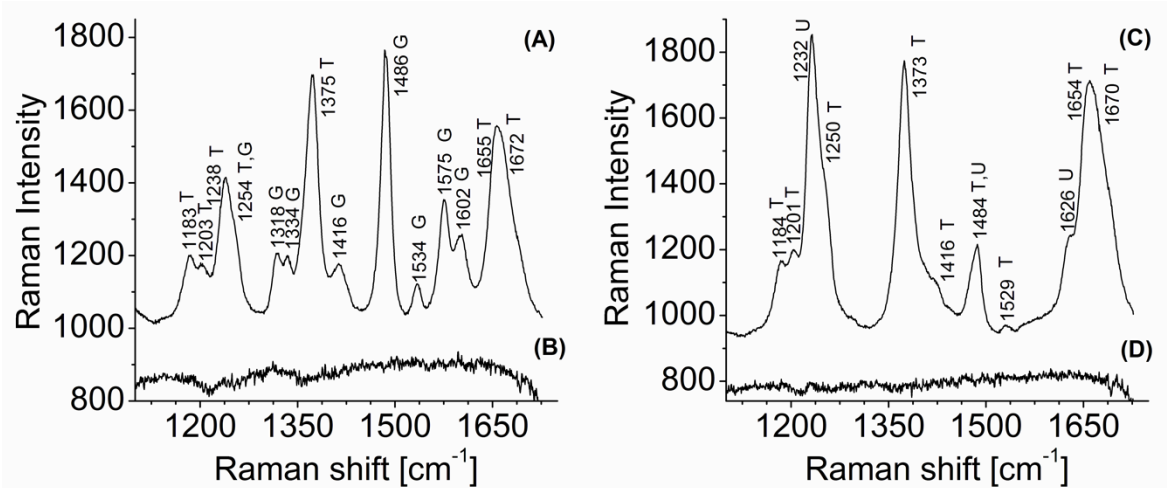
Supplementary Fig. S4. Resonance Raman spectra of AlkB and 1-me-dAMP, pH 8.0, in Hepes buffer (H_2O). **(A)** AlkB in Hepes buffer (H_2O), **(B**) 1-me-dAMP in Hepes buffer (H_2O), **(C)** Complex of AlkB•1-me-dAMP and **(D)** Difference spectrum of AlkB•1-me-dAMP complex with contributions from unbound AlkB and unbound 1-me-dAMP removed. The spectra were obtained using 225 nm laser excitation. Lorentzian band fits are shown as dashed lines.



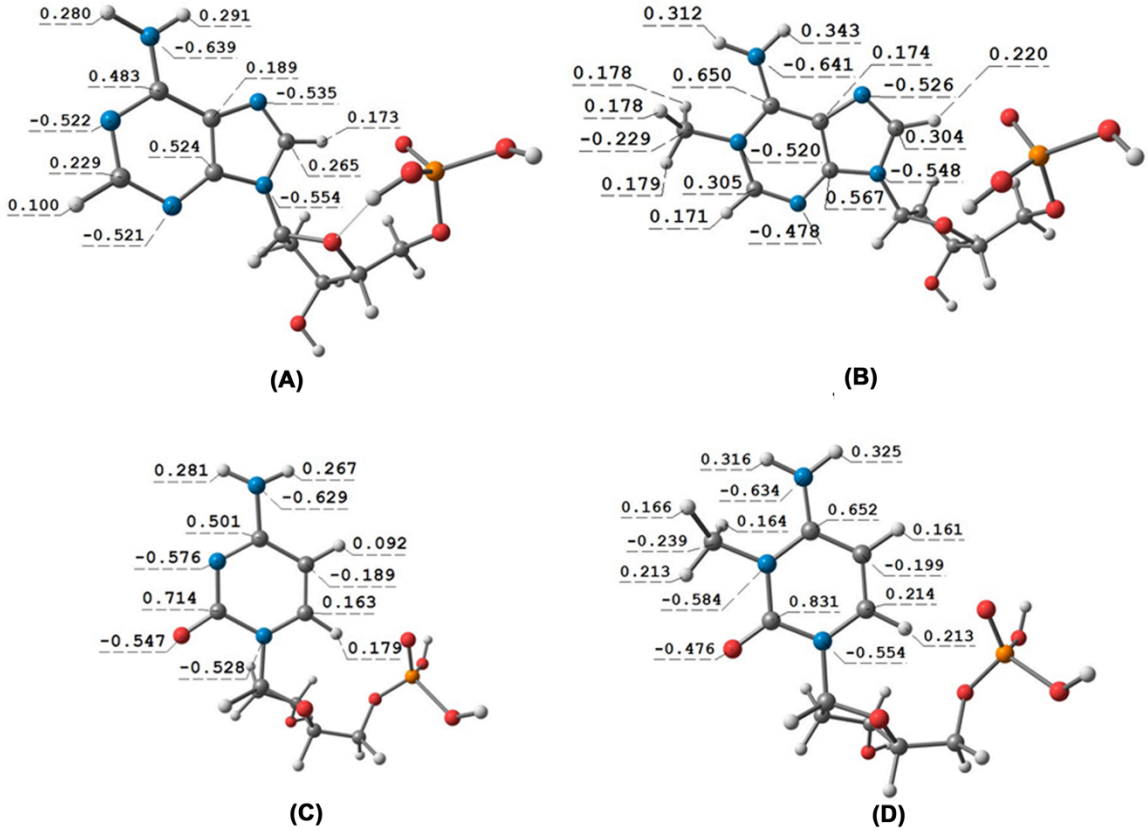
Supplementary Fig. S5. Resonance Raman spectra of AlkB with to 5'-dT-(1-me-dA)-dT-3' trinucleotide, pH 8.0., in the absence of co-factors (Fe^{2+} and 2-OG). **(A)** 5'-dT-(1-me-dA)-dT-3' trinucleotide in Hepes buffer (H_2O), **(B)** Difference spectrum of AlkB•dT-(1-me-dA)-dT complex with contributions from unbound AlkB and unbound dT-(1-me-dA)-dT (H_2O) removed. The result was multiplied by a factor of 2. Lorentzian band fits are shown as dashed lines. The spectra were obtained using laser excitation of 260 nm.



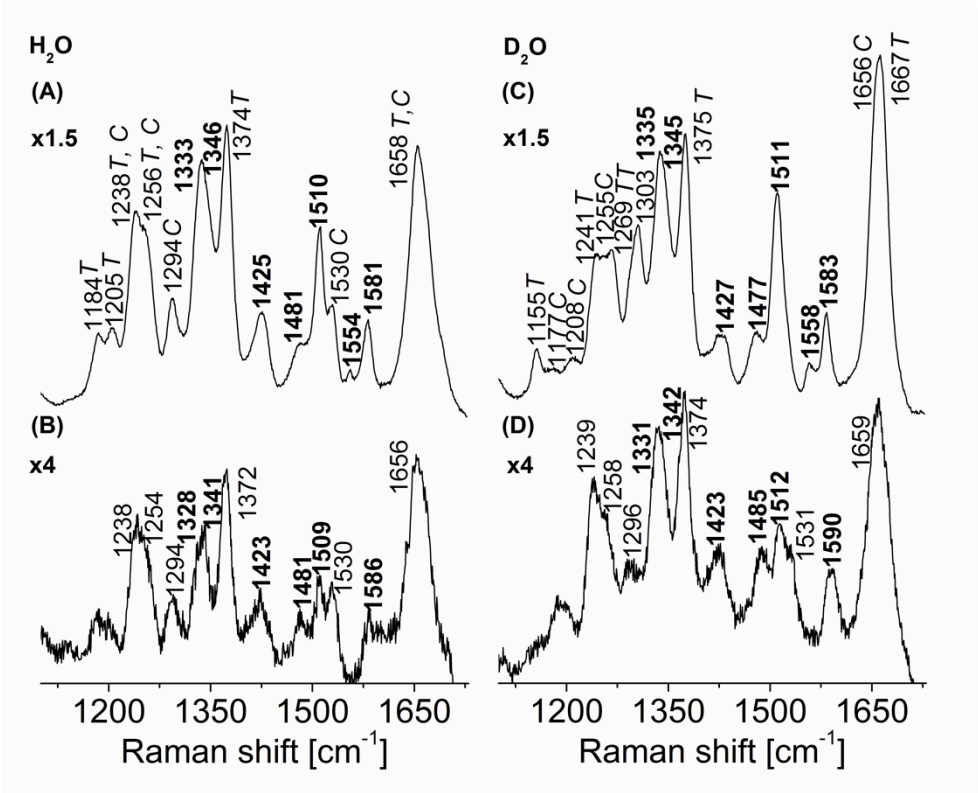
Supplementary Fig. S6. Resonance Raman spectra of AlkB bound to 5'-dT-dA-dT-3' trinucleotide and AlkB bound to 5'-dT-dC-dT-3' trinucleotide, pH 8.0. in D_2O **(A)** 5'-dT-dA-dT-3' trinucleotide in Hepes buffer (D_2O) and **(B)** Difference spectrum of AlkB•dT-dA-dT complex with contributions from unbound AkB and unbound dT-dA-dT removed. The result was multiplied by a factor of 2. **(C)** 5'-dT-dC-dT-3' trinucleotide in Hepes buffer (D_2O) and **(D)** Difference spectrum of AlkB•dT-dC-dT complex with contributions from unbound AlkB and unbound dT-dC-dT removed. The result was multiplied by a factor of 3. The spectra were obtained using laser excitation of 260 nm. The wavenumbers corresponding to dAMP and dCMP are depicted in bold. Lorentzian band fits are shown as dashed lines.



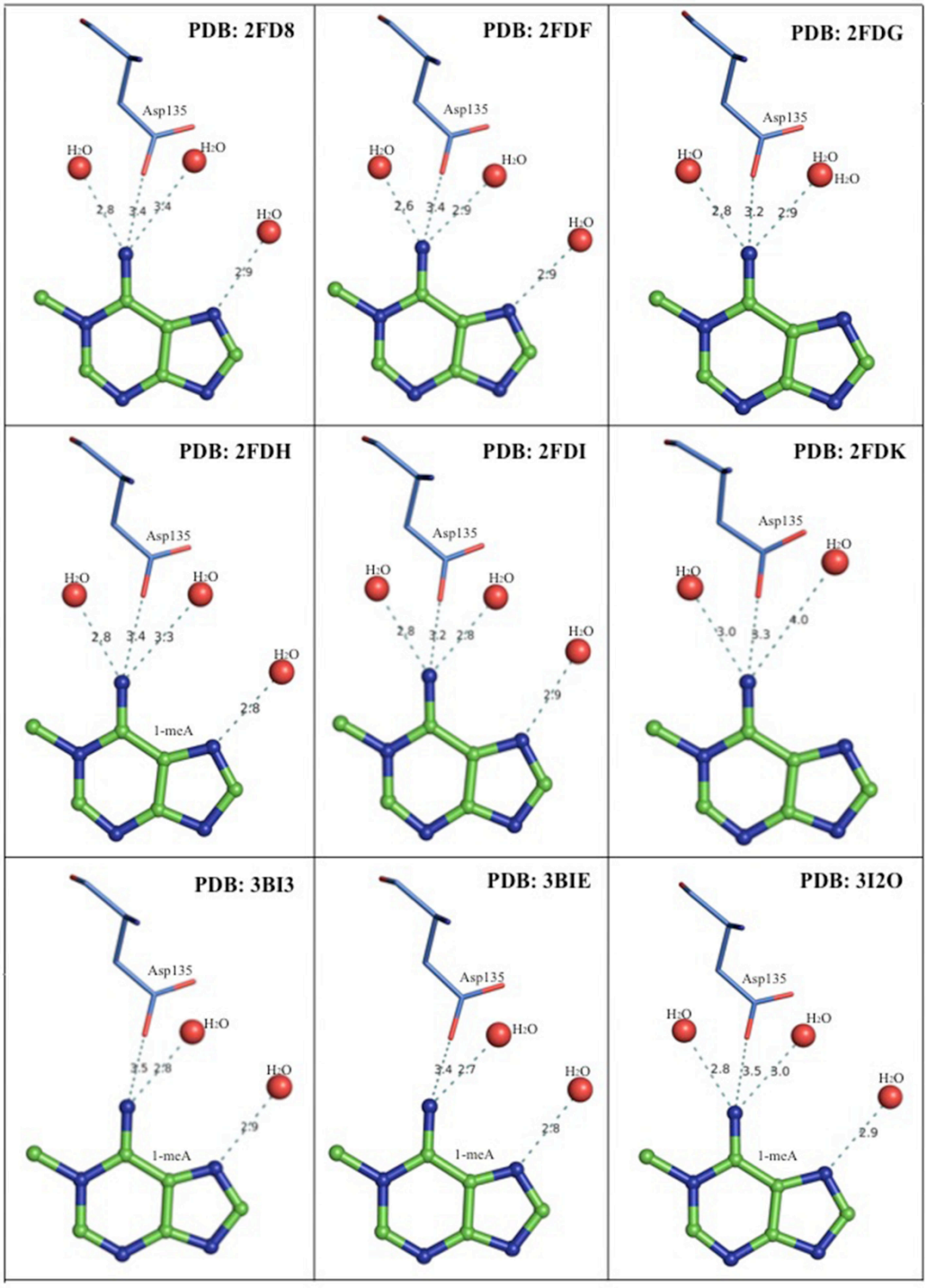
Supplementary Fig. S7. Resonance Raman spectra of **(A)** 5'-dT-dG-dT-3' trinucleotide in Hepes buffer, pH 8.0 (H_2O), **(B)** Difference spectrum of AlkB•dT-dG-dT complex with contributions from unbound AlkB and unbound dT-dG-dT removed. **(C)** 5'-dT-dU-dT-3' trinucleotide in Hepes buffer, pH 8.0 (H_2O), **(D)** Difference spectrum of AlkB•dT-dU-dT complex with contributions from unbound AlkB and unbound dT-dU-dT removed. The spectra were obtained using laser excitation of 260 nm.



Supplementary Fig. S8. Computed (B3LYP/6-31G** level of theory) charge distribution for **(A)** dAMP, **(B)** 1-me-dAMP, **(C)** dCMP and **(D)** 3-me-dCMP.



Supplementary Fig. S9. Resonance Raman spectra of AlkB bound to 5'-dC-dT-(1-me-dA)-dT-dC-3' pentamer, pH 8.0. **(A)** 5'-dC-dT-(1-me-dA)-dT-dC-3' pentamer in Hepes buffer (H_2O), **(B)** Difference spectrum of AlkB•dC-dT-(1-me-dA)-dT-dC complex with contributions from unbound AlkB and unbound dC-dT-(1-me-dA)-dT-dC (H_2O) removed. **(C)** 5'-dC-dT-(1-me-dA)-dT-dC-3'pentamer in Hepes buffer (D_2O) and **(D)** Difference spectrum of AlkB•dC-dT-(1-me-dA)-dT-dC complex with contributions from unbound AlkB and unbound dC-dT-(1-me-dA)-dT-dC (H_2O) removed. The spectra of pentamer in H_2O and D_2O were multiplied by a factor of 1.5 and the difference spectra were multiplied by a factor of 4. The spectra were obtained using laser excitation of 260 nm. The wavenumbers corresponding to 1-me-dAMP are depicted in bold.



Supplementary Fig. S10. Comparison of active-site interactions of AlkB with 1-me-dAMP from various crystallographic data[2, 4]. Hydrogen bonds are indicated by dashed lines and distances are in Å.

SUPPLEMENTARY TABLES

Supplementary Table S1. Experimental (λ_{exc}=260 nm) wavenumbers (cm^{-1}) of 1-me-dAMP (pH 6.0) and 5'-dT-(1-me-dA)-dT-3' (pH 8.0), free in solution and when bound to AlkB, in water and D_2O.

1-me-dAMP (H_2O)	**AlkB•1-me-dAMP[a] (H_2O)**	**dT-(1-me-dA)-dT (H_2O)**	**AlkB• dT-(1-me-dA)-dT[a] (H_2O)**	**Mode Assignments[c,d]**	**1-me-dAMP (D_2O)[b]**	**AlkB•1-me-dAMP[a] (D_2O)[b]**	**dT-(1-me-dA)-dT (D_2O)[b]**	**AlkB • dT-(1-me-dA)-dT[a] (D_2O)[b]**	**Mode Assignments[c,d]**
1643	1645 (+2)			N6H6 be, C2N3 str, C2H be, N1C6 str, C5C6 str, N7C8 str, C4N9 str, C8H be					
1594	1597 (+3)	1582	1584 (+2)	N6H6 be, N3C4 str, C8H be, C4C5 str, N9H be, C1H1 be	1584	1595 (+11)	1584	1590 (+7)	N3C4 str, N1C6 str, N9H be, C8H be, C1H1 be, N7C8 str
1565	1562 (-3)	1555		C2H be, N6H6 be, C6N6 str, C1H1be, C2N3 str	1556	1568 (+12)	1558		C2H be, C1H1 be, N9H be, C6N6 str, N3C4 str
1469	–	1479	1480 (0)	C1H1 be, N6H6 be, C8H be	1474	1491 (+15)	1480	1485 (+6)	C1H1 be
1503	1501 (-2)	1511	1510 (-1)	C1H1 be, C2N3 str, C2H be, N6H6 be, C4N9 str, C5N7 str, C5C6 str, N1C6 str	1500	1508 (+8)	1514	1510 (-4)	C1H1 be, C2H be
1426	1424 (-2)	1427	1423 (-3)	C2H be, N9H be, C8H be, C5N7 str, C1H1 be, N1C6 str	1432	1425 (-7)	1428	1424 (-3)	C1H1 be, C2H be, C5C6 str, N7C8 str, N9H be, C8H be, N6H6 be, N1C6 str

				C1H1 be, C2H be, C4N9 str, N7C8 str, C8H be	1415				C1H1 be, C2H be, C4N9 str, N7C8 str, C8H be
				C2H be, N9H be, C8H be, C1H1 be	1370	1360 (-10)			C2H be, N9H be, C8H be, C1H1 be
1358	–	1345	1341 (-4) (Decreases in intensity)	C2H be, N9H be, C8H be, C8N9 str	1355		1348	1343 (-5) (Decreases in intensity)	C2H be, C8H be, N9H be
1330	1328 (-2)	1333	1326 (-7) (Decreases in intensity)	N1C6 str, C5N7 str, N1C2 str, C8H be, N9H be, C1H1 be, N6H6 be	1333	1329 (-4)	1337	1329 (-7) (Decreases in intensity)	C1H1 be, N9H be, C8H be, N1C2 str, N1C6 str, C5N7 str, N6H6 be
1291	–			C1H1 be, N1C2 str, N9H be, C5N7 str	1272	1281 (+7)			C1H1 be, N1C2 str, N9H be, C5N7 str

[a]Values in parentheses represent the average shifts obtained from 3 data sets. [b]Exchangeable hydrogen atoms upon isotope labelling are shown in our previous publication[1]. [c]Mode assignments were obtained from vibrational frequencies computed for the nucleobase, 1-methyladenine, using B3LYP/6-31G** level of theory. A detailed table of assignments is listed in our previous publication[1]. [d]str: stretch; be: bend.

Supplementary Table S2. Experimental (λ_{exc}=260 nm) wavenumbers (cm^{-1}) of 3-me-dCMP and 5'-dT-(3-me-dC)-dT-3' free in solution and when bound to AlkB, in water and D_2O.

3-me-dCMP (H_2O)	**dT-3-me-dC-dT (H_2O)**	**AlkB*dT-3-me-dC-dT[a] (H_2O)**	**Mode Assignments[c,d]**	**3-me-dCMP (D_2O)[b]**	**dT-3-me-dC-dT (D_2O)[b]**	**AlkB*dT-3-me-dC-dT[a] (D_2O)[b]**	**Mode Assignments[c,d]**
1549	1549	1539 (-10)	N4H4 be, N3C4 str, C3H3 be, C4N4 str, N1H be, C6H be	1537	1537 (-12)	1552 (+15)	N1H be, N3C4 str, N1C2 str, C3H3 be
1457	1455	1447 (-8)	C3H3 be, N1H be, C5H be, N4H4 be				
1427			C3H3 be, N1H be, N4H4 be	1435	1439	–	N1H be, C3H3 be, C5H be, C4C5 str
			C3H3 be, N4H4 be, C5H be, C6H be	1407	1403	–	C3H3 be, N1H be, C6H be
1268	1270	1267 (-3) (Decreases in intensity)	C3H3 be, C2N3 str, N1H be, N4H4 be, C4N4 str	1275	1270 (0)	1263 (-7) (Decreases in intensity)	C3H3 be, C2N3 str, N4H4 be, C6H be, C4N4 str
			C5H be, C6H be, N4H4 be, N1C2 str, N1H be, C3H3 be	1179	1180	1184 (+4)	C5H be, N1H be, N3C3 str, N4H4 be
			C3H3 be	1151	1155	1151 (-3)	C3H3 be

[a]Values in parentheses represent the average shifts obtained from 3 data sets. [b]Exchangeable hydrogen atoms upon isotope labelling are depicted in our previous publication[1]. [c]Mode assignments were obtained from vibrational frequencies computed for the nucleobase, 3-methyl-cytosine, using B3LYP/6-31G** level of theory. A detailed table of assignments is listed in our previous publication[1]. [d]str: stretch; be: bend.

Supplementary Table S3. Computed (B3LYP/6-31G**) wavenumbers (cm^{-1}) and experimental (λ_{exc}=260 nm) wavenumbers (cm^{-1}) of and 5'-dT-dA-dT-3' free in solution and when bound to AlkB, in water and D_2O.

DFT	dT-dA-dT (H_2O)	AlkB*dT-dA-dT[a] (H_2O)	Mode Assignments[b,c]	DFT	dT-dA-dT (D_2O)	AlkB*dT-dA-dT[a] (D_2O)	Mode Assignments[b,c]
1657		-	N6H6 be, C5C6 str, C6N6 str, C2H be	1637 (-20)	1624		C6N6 str, C4C5 str, N1C2 str, C2H be, N9H be
1633	1581	1581 (0)	N3C4 str, C5C6 str, C5N7 str, N9H be, C8H be, N6H6 be, N1C6 str	1629 (-4)	1577 (-4)	1581 (+3)	N3C4 str, C5C6 str, N7C8 str, C8H be, N1C6 str
1601	1604	1599 (-5)	N6H6 be, C2H be, C4C5 str, C5C6 str, N1C2 str	1531 (-71)		1601	C2H be, N1C6 str, C8H be, N7C8 str, C4C5 str, N1C2 str
1517			C8H be, N7C8 str, N6H6 be, C5C6 str, C4C5 str				
1502	1509	1506 (-3)	C2H be, N6H6 be, N1C6 str, N3C2 str	1508 (+6)		1507	C2H be, C8H be, N7C8 str, C2N3 str, N1C6 str, N9H be
1355	1337	1336 (-1)	C2H be, C5N7 str, N1C2 str, C5C6 str, N6H6 be, C6N6 str	1359 (+4)	1337 (0)	1332 (-5)	C8H be, N1C6 str, C5N7 str, C4C5 str, N1C2 str, N9H be

[a]Values in parentheses represent the average shifts obtained from 2data sets. [b]Mode assignments were obtained from vibrational frequencies computed for the nucleobase, adenine, using B3LYP/6-31G** level of theory. [c]str: stretch; be: bend.

Supplementary Table S4. Computed (B3LYP/6-31G**) wavenumbers (cm^{-1}) and experimental (λ_{exc}=260 nm) wavenumbers (cm^{-1}) of 5'-dT-dC-dT-3' free in solution and when bound to AlkB, in water and D_2O.

DFT	dT-dC-dT (H_2O)	AlkB*dT-dC-dT[a] (H_2O)	Mode Assignments[b,c]	DFT	dT-dC-dT (D_2O)	AlkB*dT-dC-dT[a] (D_2O)	Mode Assignments[b,c]
1783		-	C2O str, N1H be	1782 (-1)			C2O str, N1H be
1671			C5C6 str, C5H be, C6H be, N3C4 str, N4H4 be, N4C4 str	1667 (-4)			C6H, C5C6 str, C5H be, C4N4 str, N1H be
1608	1598	1598 (0)	N4H4 be, C6H be			1594	
1545	1531	1531 (0)	C4C5 str, N1H be, N1C2 str, C5H be, N4H4 be, C5C6 str	1540 (-5)	1526 (-5)	1533 (+7)	C4C5 str, N1H be, N1C2 str, C5H be, N4H4 be, C5C6 str
1487			C5H be, N1C6 str , N3C4 str, N4H4 be, C6H be, C4N4' str	1506 (+19)	1508		C6H be, C5H be, N4C4 str, N4H4 be, C5C6 str

[a]Values in parentheses represent the average shifts obtained from 2 data sets. [b]Mode assignments were obtained from vibrational frequencies computed for the nucleobase, cytosine, using B3LYP/6-31G** level of theory. [c]str: stretch; be: bend.

REFERENCES

1. N. Jayanth and M. Puranik, *Journal of Physical Chemistry B*, 2011, **115**, 6234-6242.
2. B. Yu, W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney and J. F. Hunt, *Nature*, 2006, **439**, 879-884.
3. C. G. Yang, C. Q. Yi, E. M. Duguid, C. T. Sullivan, X. Jian, P. A. Rice and C. He, *Nature*, 2008, **452**, 961-964.
4. B. Yu and J. F. Hunt, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 14315-14320.