Supplemental material for

Structural insights into methylated DNA recognition by the C-terminal zinc fingers of the DNA reader protein ZBTB38

Nicholas O. Hudson[1], Frank G. Whitby[2], Bethany A. Buck-Koehntop[1,3]

[1]Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

[2]Department of Biochemistry, University of Utah, Salt Lake City, Utah 84112

[3]To whom correspondence should be addressed: Phone: 801-581-3186; Fax: 801-581-8433; E-mail: koehntop@chem.utah.edu

Contains supplemental Tables S1-S2 and supplemental Figures S1-S4.

Table S1. Oligonucleotide sequences utilized for NMR, X-ray crystallography and EMSA

NMR	
mCZ38BS_18mer_F	5′-^{1}GCACTCATmCGGmCGCAGAC^{18}-3′
mCZ38BS_18mer_R	5′-^{19}GTCTGmCGCmCGATGAGTGC^{36}-3′
mCZ38BS_21mer_F	5′-^{1}GCACTCATmCGGmCGCAGATCAG^{21}-3′
mCZ38BS_21mer_R	5′-^{22}CTGATCTGmCGCmCGATGAGTGC^{42}-3′
X-ray Crystallography	
mCZ38BS_18mer_F	5′-^{1}GCACTCATmCGGmCGCAGAC^{18}-3′
mCZ38BS_18mer_R	5′-^{19}GTCTGmCGCmCGATGAGTGC^{36}-3′
EMSA	
mCZ38BS_27mer_F (ATMGGMG)	5′-^{1}GCACTCATmCGGmCGCAGATCAGCTAGCC^{27}-3′
mCZ38BS_27mer_R	5′-^{28}GGCTAGCTGATCTGmCGCmCGATGAGTGC^{54}-3′
mCZ38BS_ACMG_F (ACMGGMG)	5′-GCACTCACmCGGmCGCAGATCAGCTAGCC-3′
mCZ38BS_ACMG_R	5′-GGCTAGCTGATCTGmCGCmCGGTGAGTGC-3′

Table S2. Summary of protein:DNA contacts between ZBTB38 ZF 6-9 and ZBTB38 ZF 6-9_K1055R with mCZ38BS

ZBTB38 ZF 6-9 (1006-1124)		mCZ38BS		
*Residue	Atom	Base	Atom	Distance (Å)
Tyr-1010	Oη	C2	O1P	2.6/2.4
Tyr-1010	Oη	C2	P	3.7
Tyr-1010	Cζ	C2	O1P	3.4
Lys-1017	Nζ	T30	O1P	3.7/2.6
Ser-1021	Cβ	A3	O2P	3.4/3.3
Pro-1022	Cδ	A3	O2P	3.3
Pro-1022	Cδ	C2	O1P	3.4/3.5
Ser-1023	N	A3	O2P	2.8
Ser-1023	Oγ	A3	O2P	2.8
Ser-1023	Cβ	A3	O2P	3.1
Ser-1023	Cβ	A3	P	3.6
Ser-1023	N	A3	P	3.7
Ser-1023	Cβ	A3	C2’	4.2/4.1
Met-1027	Cγ	T30	C7	4.1
Met-1027	Sδ	T30	C7	4.2
His-1028	Nδ1	T30	O2P	2.8/2.7
His-1028	Nδ1	T30	P	3.5/3.6
Cys-1031	Sγ	A29	O5’	3.4/3.3
Cys-1031	Sγ	A29	C3’	3.8/3.6
Cys-1031	Sγ	A29	P	3.9/3.8
Cys-1031	Sγ	A29	C5’	4.0/3.9
Cys-1031	Sγ	A29	C2’	4.2
Arg-1045	Cζ	C27	O5’	3.3/3.2
Arg-1045	Nη1	C27	O5’	3.3
Arg-1045	Nη2	C27	O3’	3.5
Arg-1045	Nη1	C27	O1P	3.7/3.4
Arg-1045	Cζ	C27	C5’	3.9/4.0
Arg-1045	Nη2	C27	O5’	3.4
Arg-1045	Nη2	G28	O1P	4.2
Phe-1047	Cδ2	G28	O2P	3.3
Phe-1047	Cε2	C27	C5’	3.8/4.0
Phe-1047	Nζ	C27	C5’	4.2
Val-1049	Cγ2	T30	C7	3.9/4.0
Val-1049	Cγ2	A29	C8	4.2
Gly-1051	Cα	C6	C5	4.2
Asn-1052	Nδ2	A29	N7	2.9
Asn-1052	Oδ1	A29	N6	2.9/3.0
Asn-1052	Nδ2	G28	C8	3.6/3.5
Asn-1052	Cγ	G28	C8	4.2
Asn-1052	Nδ2	G28	N7	3.6
Gln-1054	Oε1	W103		2.5
W103		C6	O1P	3.0
Lys-1055	Nζ	T8	O4	2.6
Lys-1055	Nζ	T8	C7	3.8
Lys-1055	Nζ	W59		2.6
Lys-1055	Cε	T8	C7	4.2
W59		G28	C5	2.8
Arg-1055	Nη2	G28	O6	2.7
Arg-1055	Nη1	T8	O4	2.9

Arg-1055	Nη2	G28	N7	3.2
Arg-1055	Nη2	A29	N6	3.3
Arg-1055	Cζ	G28	O6	3.4
Arg-1055	Nη2	G28	C6	3.5
Arg-1055	Nη1	G28	O6	3.3
Arg-1055	Cζ	C27	C5	4.1
Arg-1055	Cζ	C27	5mC	3.8
Arg-1055	Nε	C27	5mC	3.8
Arg-1055	Nη1	C27	5mC	4.0
Arg-1055	Cδ	C27	5mC	4.0
His-1056	Nδ1	C27	O2P	2.7/2.8
His-1056	Nδ1	C27	P	3.8/3.7
Ile-1059	Cδ1	C26	C3’	3.8/3.6
Ile-1059	Cδ1	C26	C2’	4.2
Ile-1059	Cδ1	C26	C5’	4.1
K1064	Nζ	W98		2.9
W98		G25	O1P	2.5
K1064	Nζ	W57		2.8
W57		G25	O1P	2.6
Lys-1073	Nζ	W121		2.6
W121		G25	O1P	2.8
Phe-1075	Cδ2	G25	O2P	3.5
Leu-1077	Cδ1	C27	5mC	3.7/3.8
Asn-1078	Nδ2	A7	O1P	3.4
Glu-1079	Oε2	C27	N4	2.8
Glu-1079	Oε2	C9	N4	2.9
Glu-1079	Oε2	C9	5mC	3.7/3.5
Glu-1079	Oε1	T8	C7	3.6/3.7
Glu-1079	Cβ	C9	5mC	3.8
Glu-1079	Oε2	C27	5mC	4.0/3.9
Glu-1079	Cδ	C27	C7	4.2
Glu-1079	Cδ	C9	5mC	4.2/4.1
Glu-1079	Cβ	C9	5mC	3.8
Glu-1079	Cδ	T8	C7	4.1
Glu-1079	Cβ	T8	C7	4.2
Thr-1080	Cγ2	C24	C2’	4.1
Lys-1082	Nζ	T8	O2P	2.6/2.7
Ile-1083	Cγ2	C24	5mC	4.1/4.0
His-1084	Nδ1	C24	O2P	2.7/2.5
His-1084	Nδ1	C24	P	3.7/3.7
Ile-1087	Cδ1	G23	C3’	4.1/4.0
Arg-1093	Nη1	C9	O2P	2.7/2.8
Arg-1093	Nη1	C9	P	3.7
Arg-1093	Nη2	C9	O2P	3.9
Arg-1093	Nη1	C9	O1P	4.1/4.0
Tyr-1094	Oη	G10	O2P	3.1/2.5
Tyr-1105	Oη	G23	O2P	2.7/2.6
Tyr-1105	Cε2	T22	C2’	3.7
Tyr-1105	Oη	G23	P	3.7/3.6
Tyr-1105	Cε2	T22	C3’	3.8
Tyr-1105	Cε1	C24	5mC	3.8
Tyr-1105	Cζ	G23	C8	3.9/4.1
Tyr-1105	Cζ	T22	C2’	4.0/4.1
Tyr-1105	Cε1	G23	C8	4.1
Tyr-1105	Cζ	G23	O2P	3.4
Tyr-1105	Cε2	G23	O2P	3.4

Leu-1106	Cδ2	G10	C3’	4.1
Leu-1106	Cδ2	G10	C2’	4.2
Ser-1107	Oγ	T22	C7	3.7
Thr-1108	Oγ1	T22	O2P	2.7/2.6
Thr-1108	Cβ	T22	O2P	3.4
Thr-1108	Cγ2	C21	C3’	4.1/4.0
Arg-1110	Nη1	G11	O2P	3.5
Asn-1111	Nδ2	C21	O2P	2.7
Asn-1111	Cβ	C21	O2P	3.4
Arg-1115	Nη2	C21	O1P	3.0
Arg-1115	Nη2	C21	O2P	4.2
Arg-1115	Nη1	C21	O1P	3.1/3.2
Arg-1115	Nη2	T20	O3’	3.1
Arg-1115	Nη2	C21	P	3.6

*Contacts shared between the two structures are shown in black, while contacts and distances unique to the WT structure or the K1055R variant are shown in red and blue, respectively.

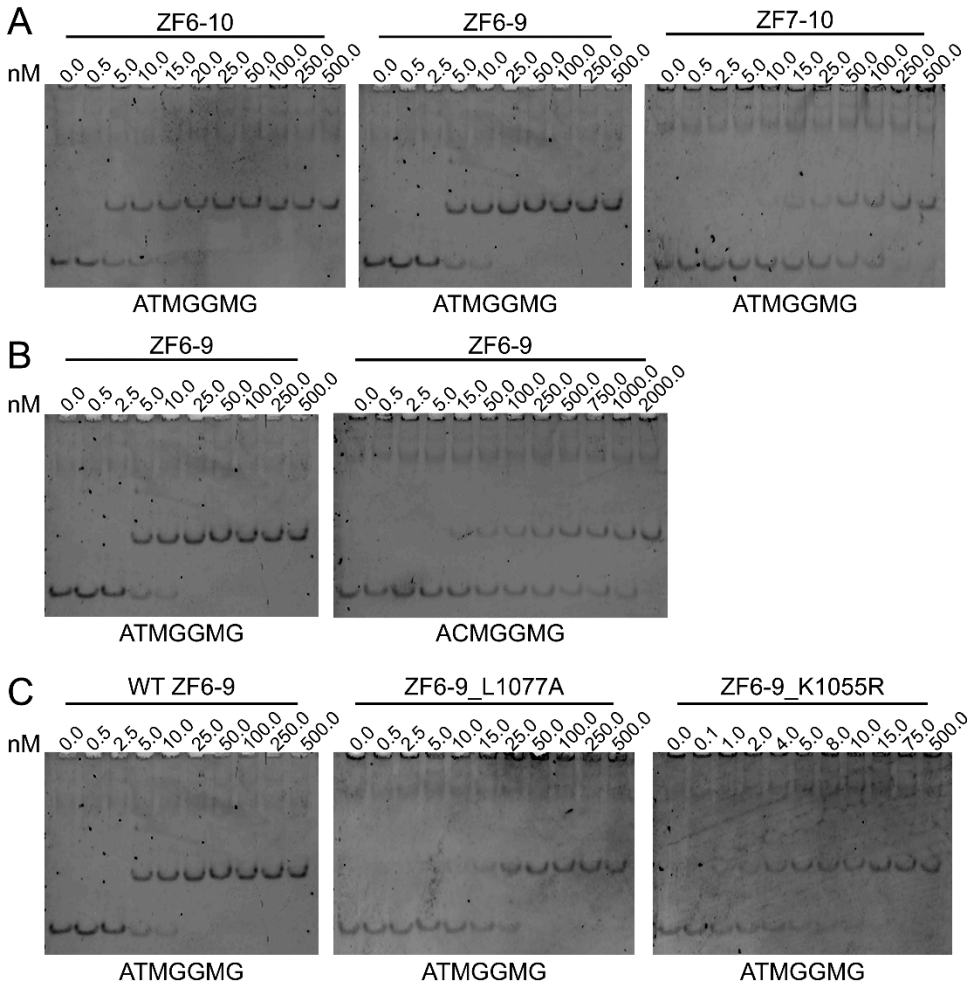

Figure S1. Representative EMSA gels for the binding isotherms depicted in Figures 1D (*A*), 3B (*B*), and 3D/4B (*C*). The wild-type (WT) ZF6-9 in complex with the mCZ38BS_27mer from (*A*) is reproduced in (*B*) and (*C*) for comparative purposes. M denotes a methylated cytosine.

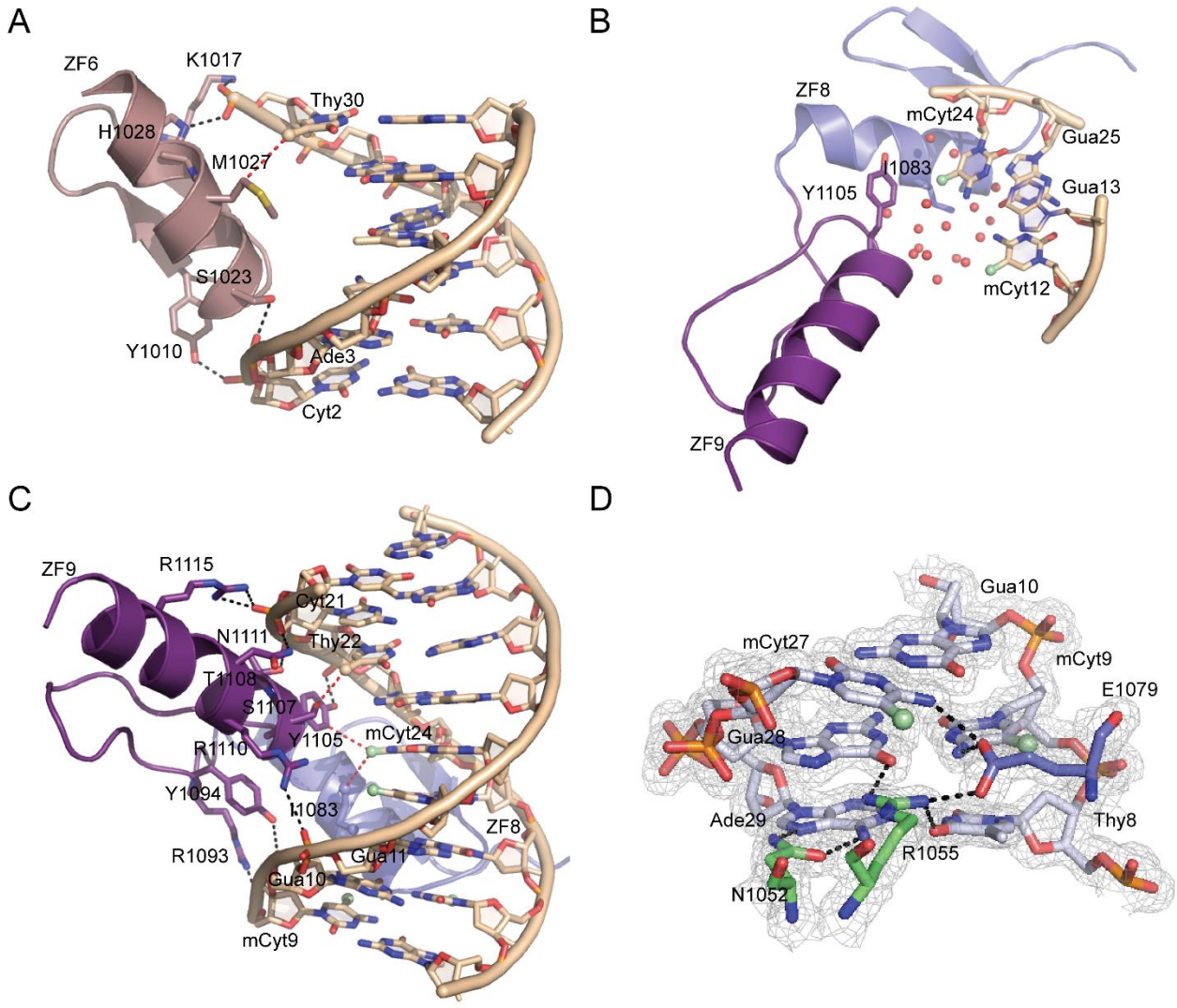

Figure S2. *A,* Summary of interactions for ZF6 with the mCZ38BS. *B,* Image depicting the solvation layer surrounding the 3′-mCpG site (red spheres). *C,* Summary of interactions for ZF9 with the mCZ38BS. A side chain interaction between I1083 in ZF8 and mC24 is also depicted. *D,* Interactions between Asn-1052, Arg-1055 and Glu-1079 with the core T8:A29, mC9:G28 and G10:mC27 base pairs. Electron densities (2Fo-Fc) contoured at 1σ are shown. For all panels, black dotted lines depict classical H-bond interactions, while red dotted lines represent van der Waals interactions.

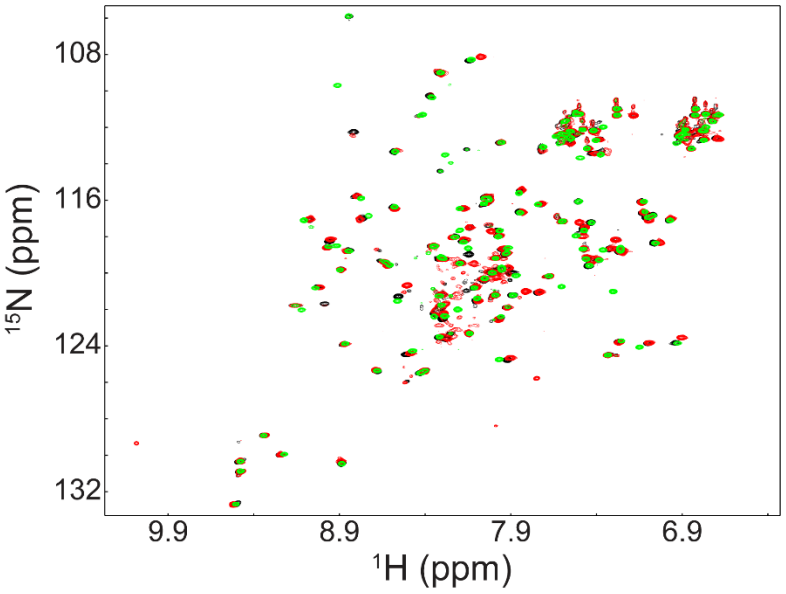

Figure S3. ^{1}H/^{15}N HSQC spectral overlay of WT ZBTB38 ZFs 6-9 (black), the K1055R (red) and L1077A (green) variants, indicating that the respective point mutations do not disrupt protein structural integrity.

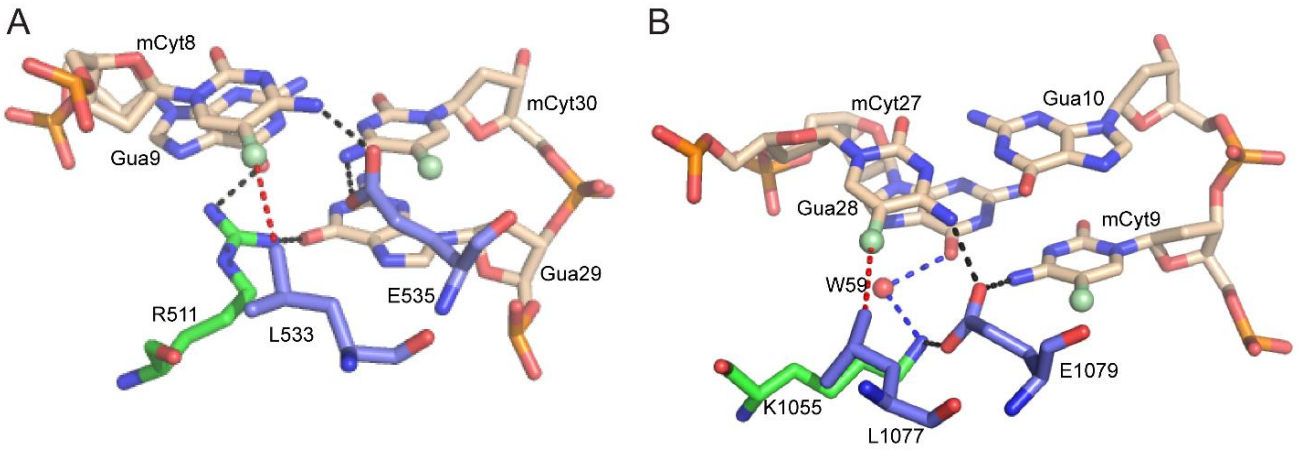

Figure S4. Comparison of the core mCpG recognition between the ZBTB33:MeECad complex (*A*, PDB ID: 4F6N) and the ZBTB38 ZF 6-9:mCZ38BS complex (*B*). Black dotted lines represent classical hydrogen bond interactions, blue dotted lines represent water mediated hydrogen bonds and red dotted lines represent van der Waals interactions.