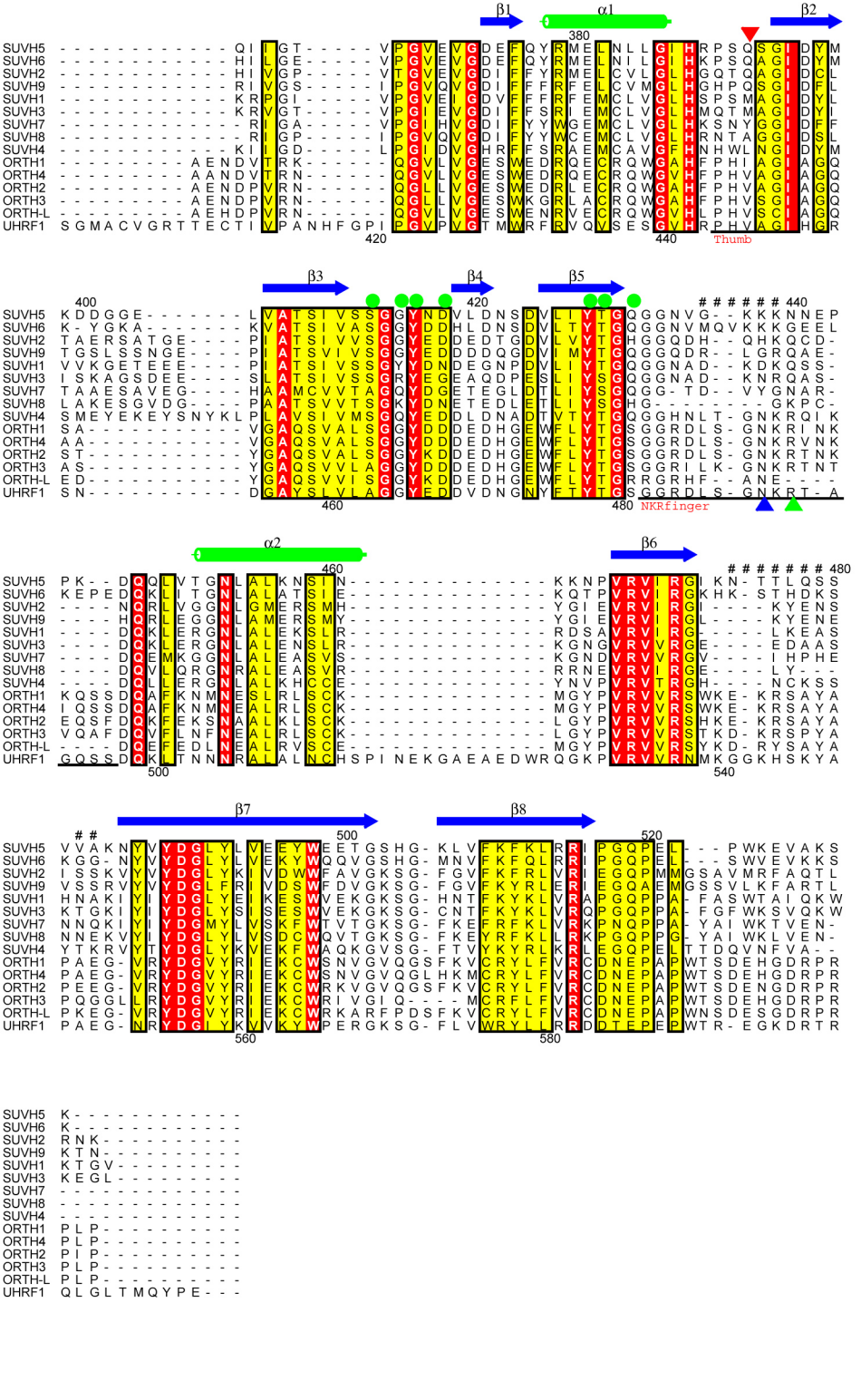
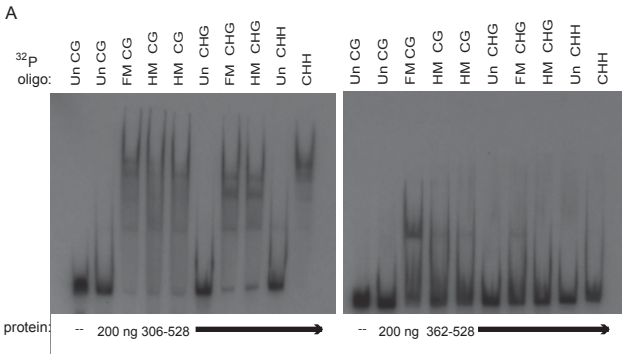
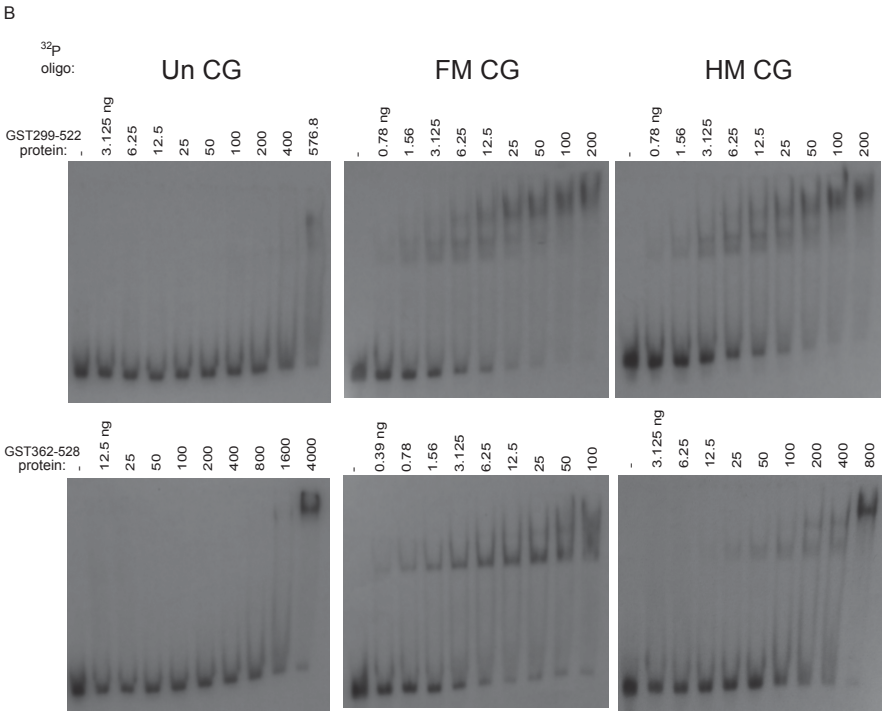
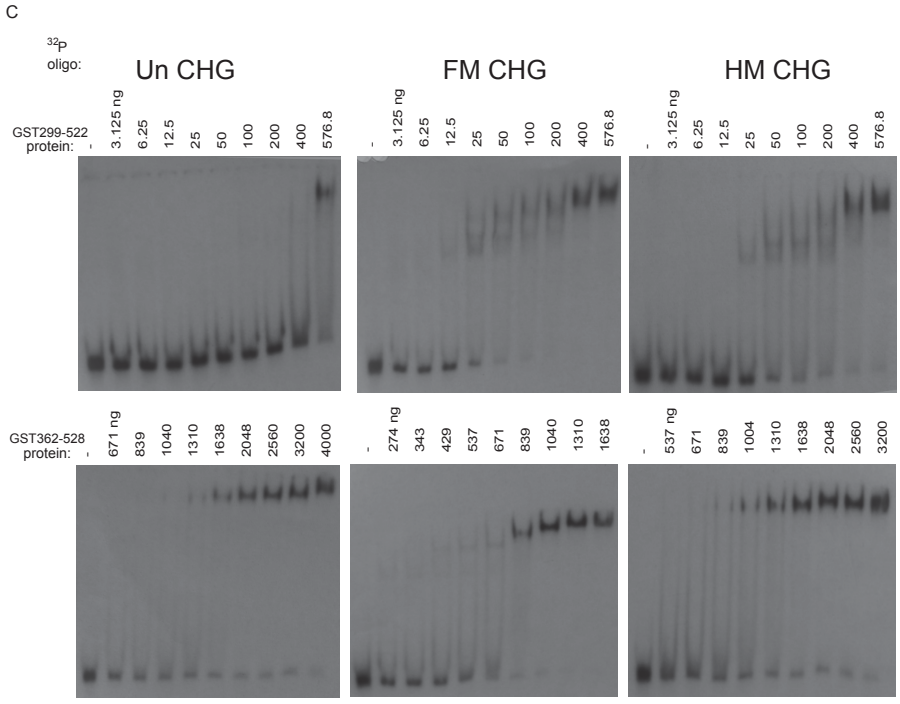
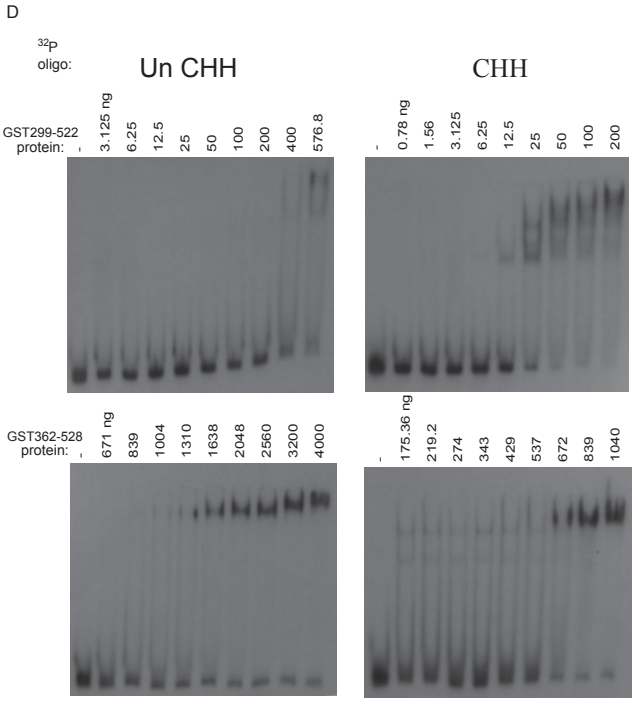

β1
α1
β2
β3
β4
β5
α2
β6
β7
β8
Thumb
NKRfinger
SUVH5
SUVH6
SUVH2
SUVH9
SUVH1
SUVH3
SUVH7
SUVH8
SUVH4
ORTH1
ORTH4
ORTH2
ORTH3
ORTH-L
UHRF1

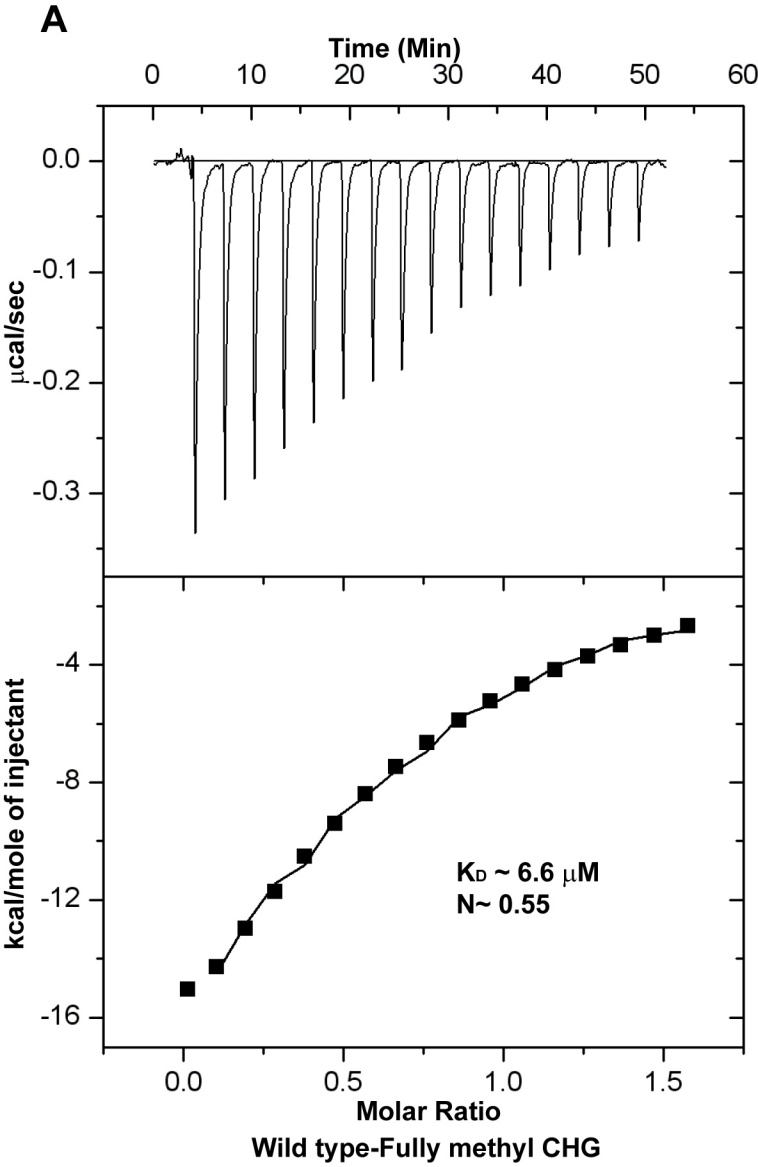
Rajakumara_Supplemental Fig. S2

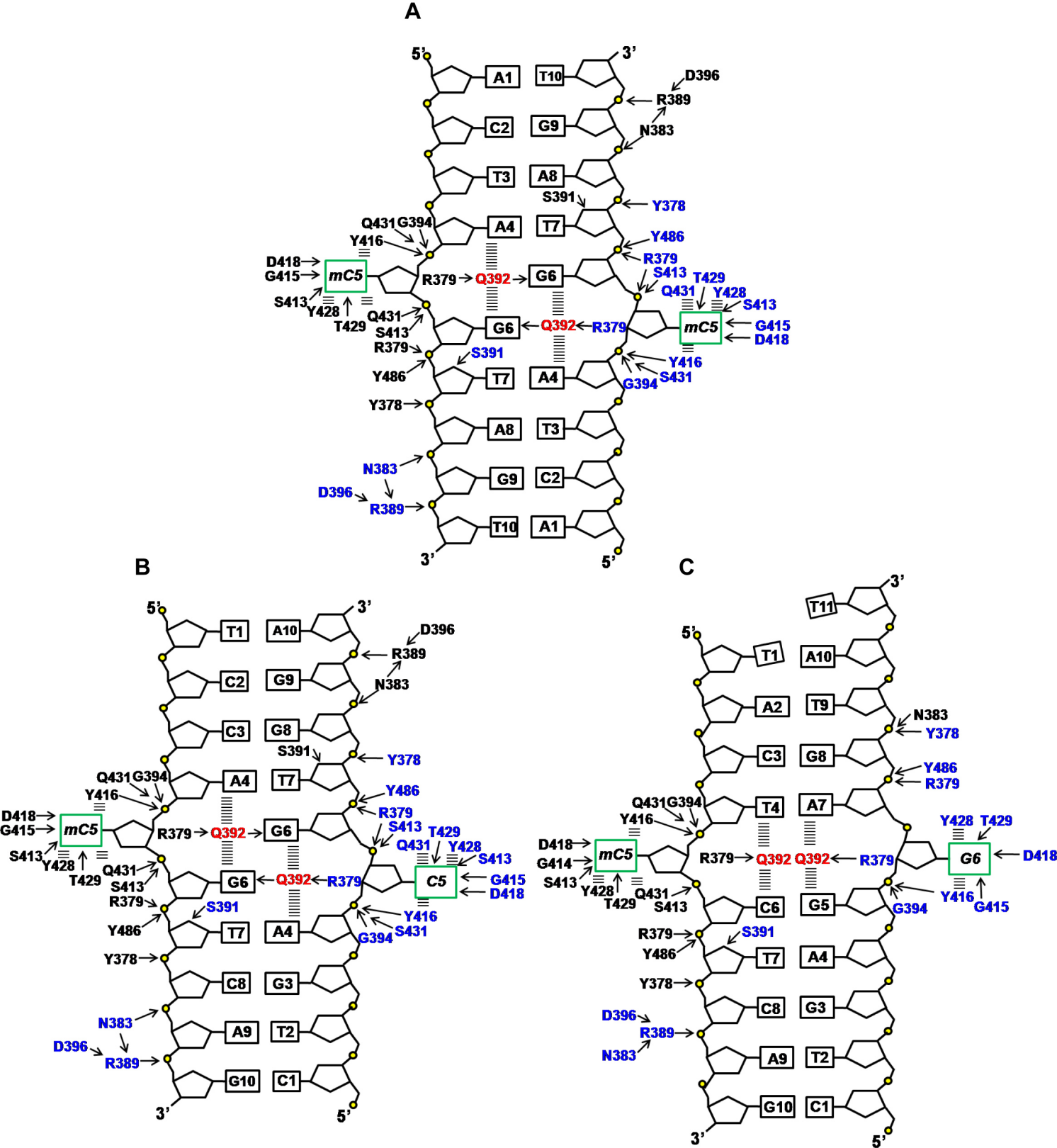


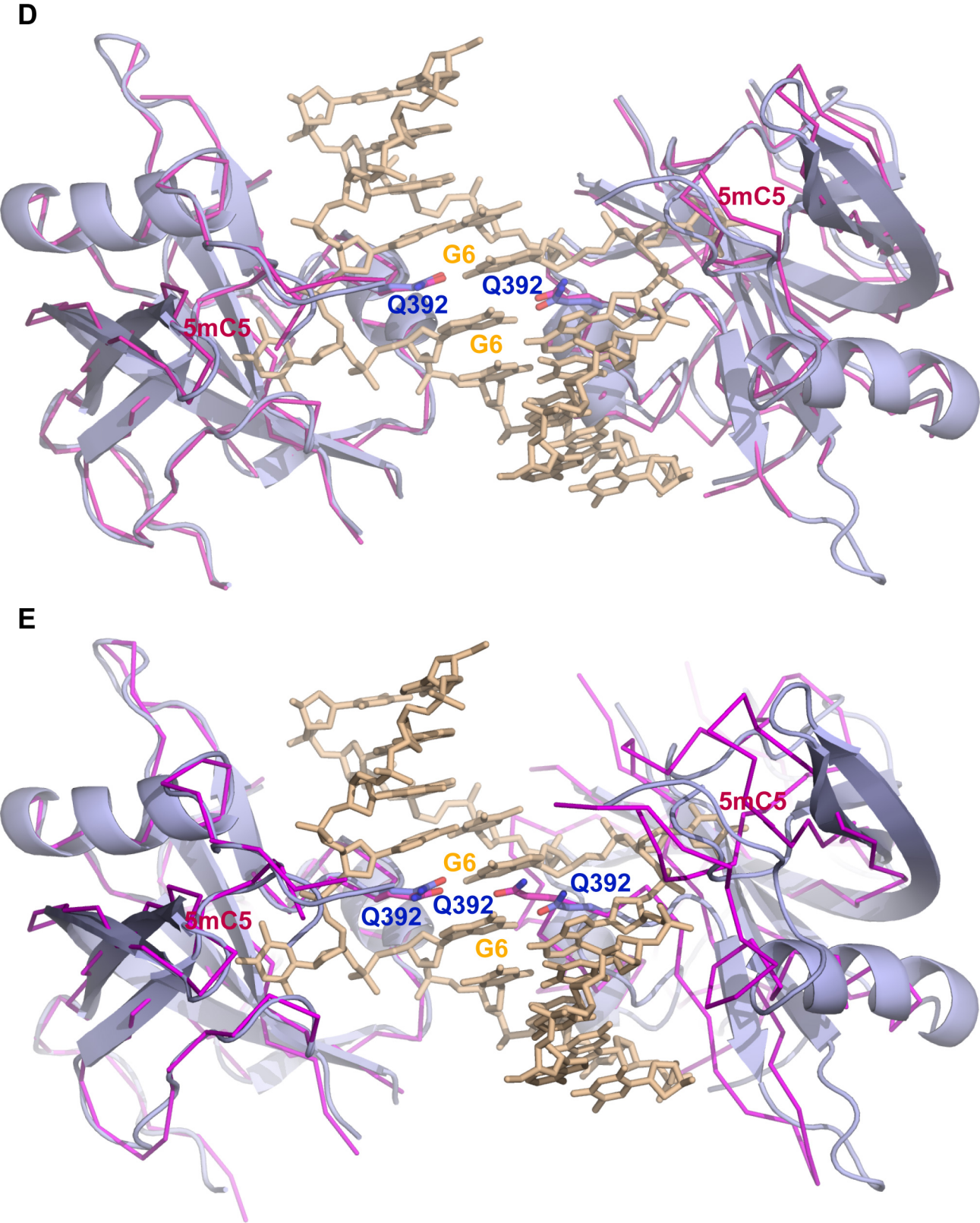


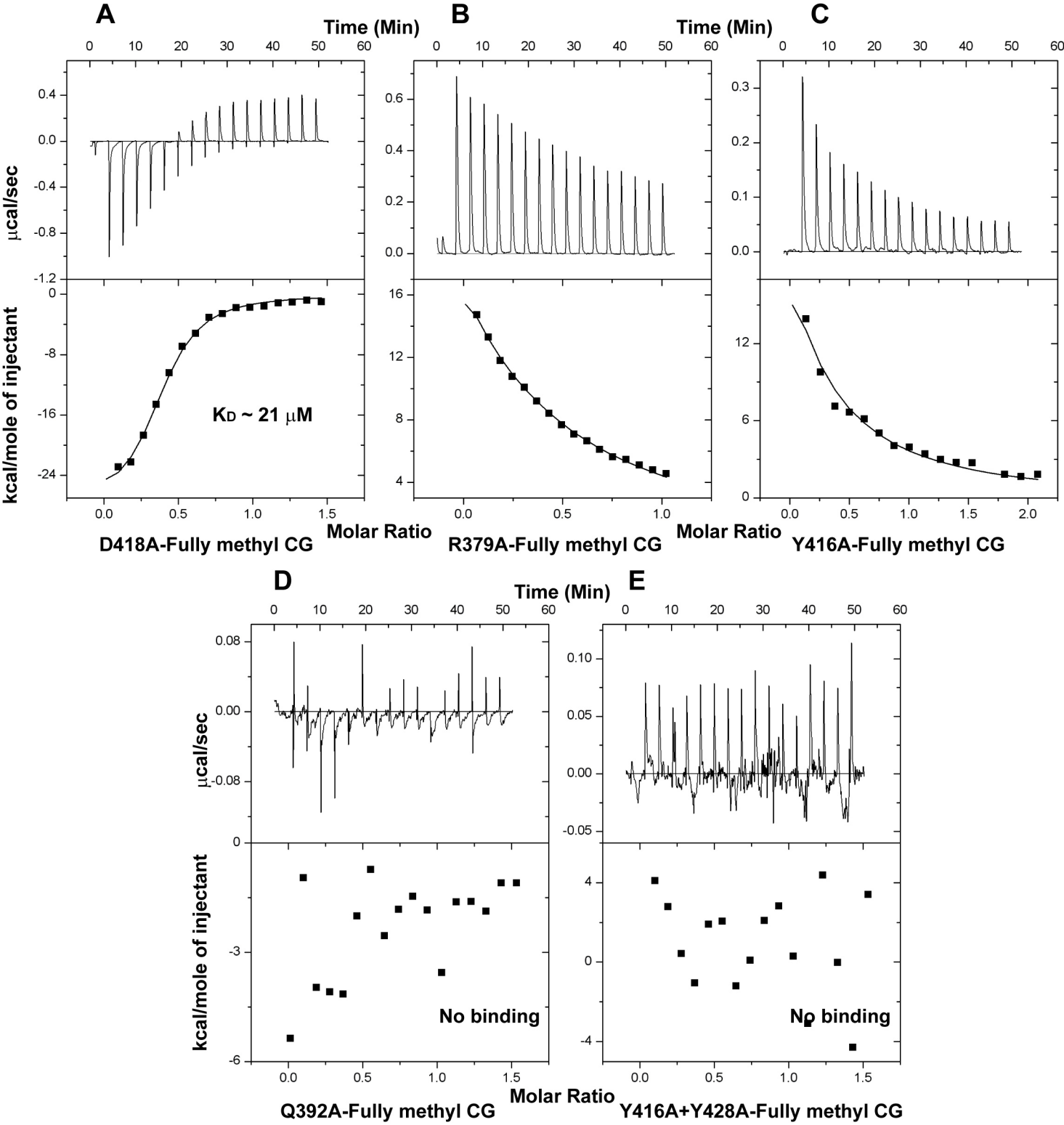


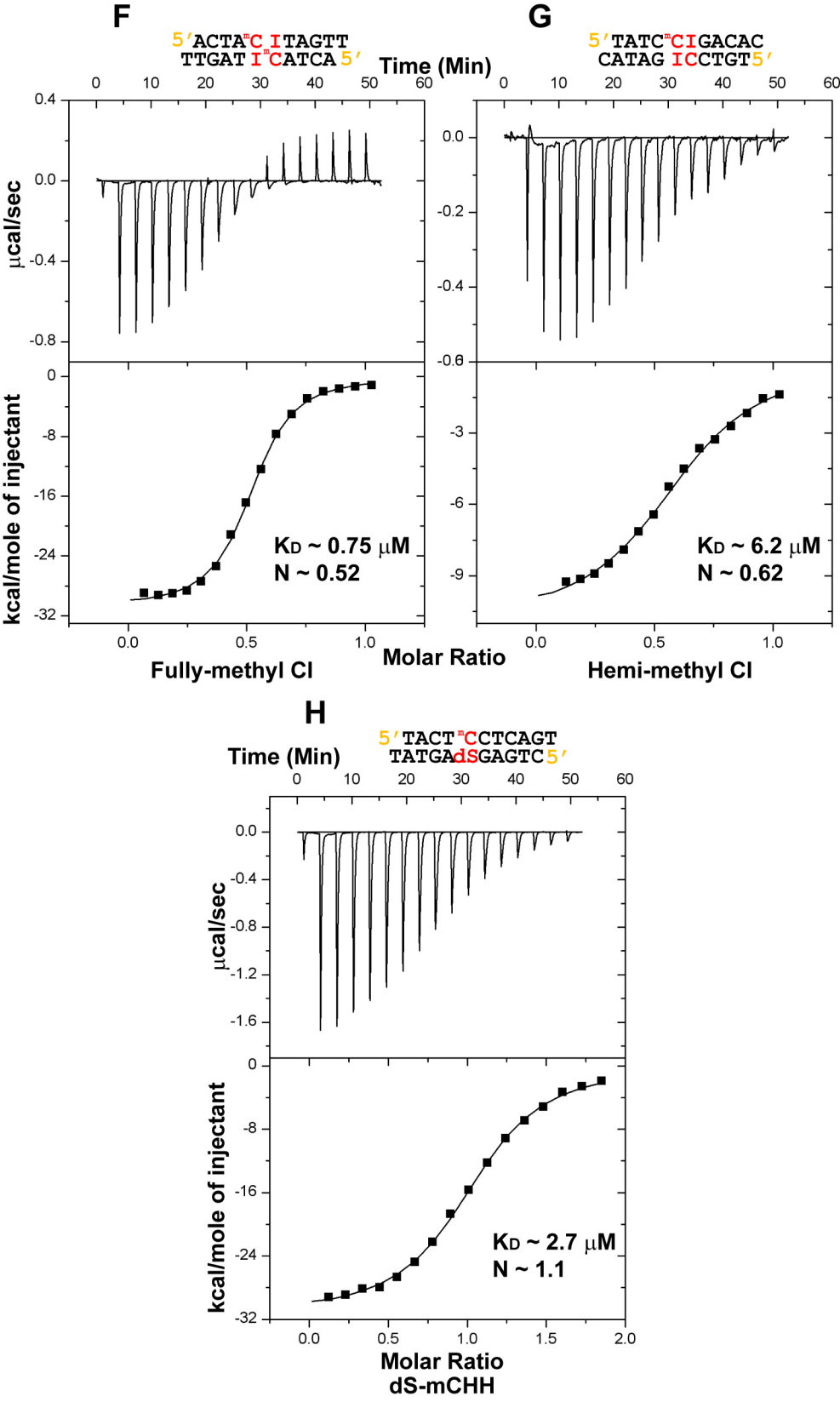



A
Time (Min)
0
10
20
30
40
50
60
0.0
-0.1
-0.2
-0.3
µcal/sec
-4
-8
-12
-16
kcal/mole of injectant
K_D ~ 6.6 µM
N~ 0.55
0.0
0.5
1.0
1.5
Molar Ratio
Wild type-Fully methyl CHG

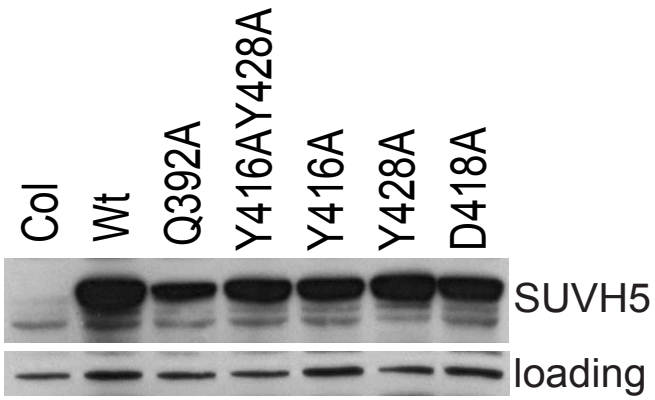

A
5'
3'
A1
T10
C2
G9
T3
A8
A4
T7
G6
G6
A4
T7
A8
T3
G9
C2
T10
A1
mC5
mC5
D396
R389
N383
S391
Y378
Y486
R379
S413
T429
Q431
Y428
S413
G415
D418
Y416
S431
G394
Q392
R379
Q431
G394
Y416
D418
G415
S413
Y428
T429
S413
Q431
R379
S391
Y486
Y378
N383
D396
R389
B
5'
3'
T1
A10
C2
G9
C3
G8
A4
T7
G6
G6
A4
T7
C8
G3
A9
T2
G10
C1
C5
mC5
D396
R389
N383
S391
Y378
Y486
R379
S413
T429
Q431
Y428
S413
G415
D418
Y416
S431
G394
Q392
R379
Q431
G394
Y416
D418
G415
S413
Y428
T429
S413
Q431
R379
S391
Y486
Y378
N383
D396
R389
C
5'
3'
T11
T1
A10
A2
T9
C3
G8
T4
A7
C6
G5
G6
T7
A4
C8
G3
A9
T2
G10
C1
mC5
D396
R389
N383
S391
Y378
Y486
R379
Y428
T429
Q431
G394
Y416
D418
G414
G415
S413
Y428
T429
Q431
S413
R379
Q392
Q392
R379
S391
Y486
Y378
N383
D396
R389
Y416
G415
G394


D
5mC5
G6
Q392
Q392
G6
5mC5
E
5mC5
G6
Q392
Q392
Q392
G6
5mC5

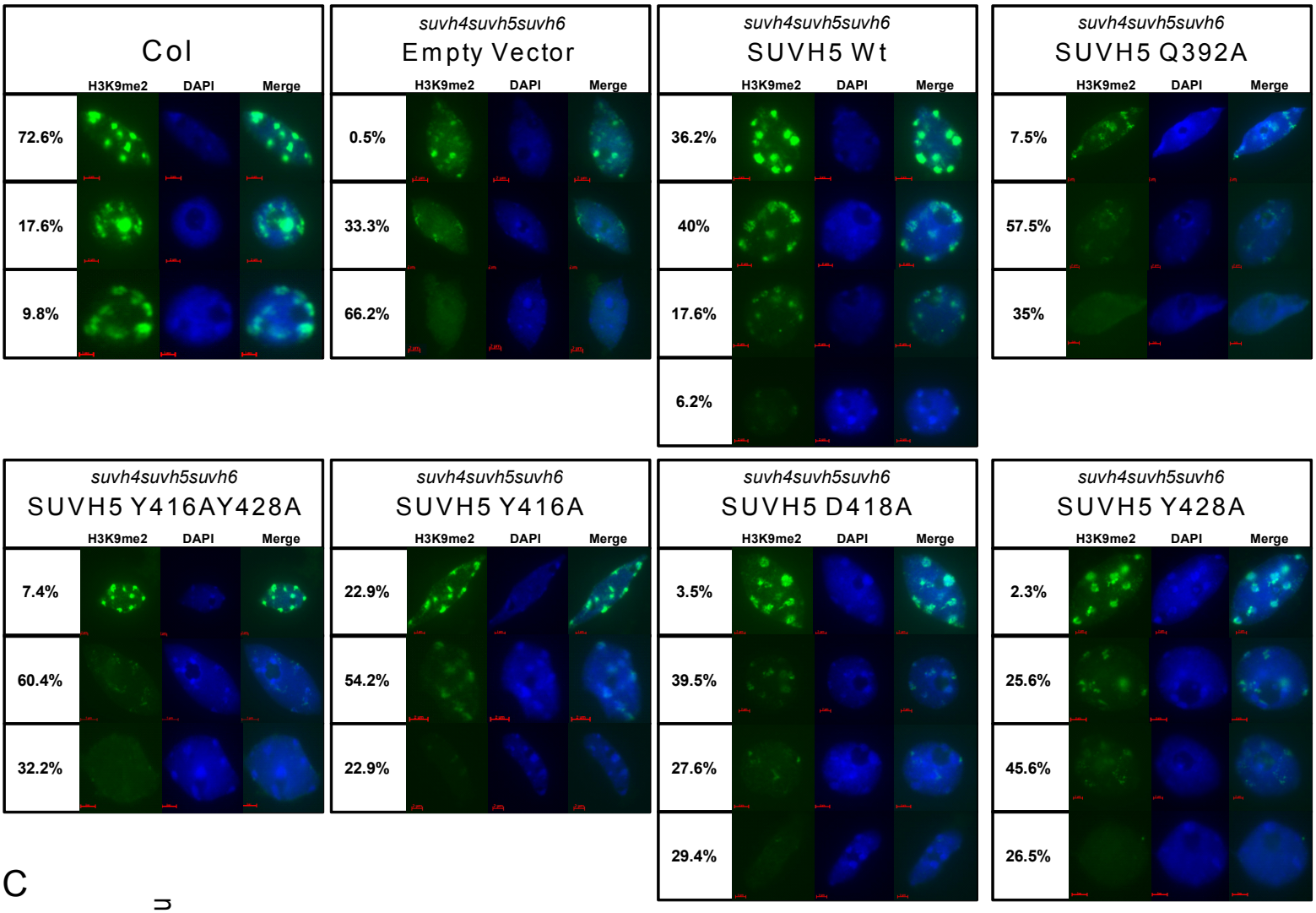

A
B
C
Time (Min)
μcal/sec
kcal/mole of injectant
KD ~ 21 μM
Molar Ratio
D418A-Fully methyl CG
R379A-Fully methyl CG
Y416A-Fully methyl CG
D
E
No binding
No binding
Q392A-Fully methyl CG
Y416A+Y428A-Fully methyl CG


F
5′ACTAmC ITAGTT
TTGAT ImCATCA 5′
Time (Min)
µcal/sec
kcal/mole of injectant
KD ~ 0.75 µM
N ~ 0.52
Fully-methyl CI
G
5′TATCmCIGACAC
CATAG ICCTGT5′
KD ~ 6.2 µM
N ~ 0.62
Hemi-methyl CI
Molar Ratio
H
5′TACTmCCTCAGT
TATGAdSGAGTC5′
Time (Min)
µcal/sec
kcal/mole of injectant
KD ~ 2.7 µM
N ~ 1.1
Molar Ratio
dS-mCHH

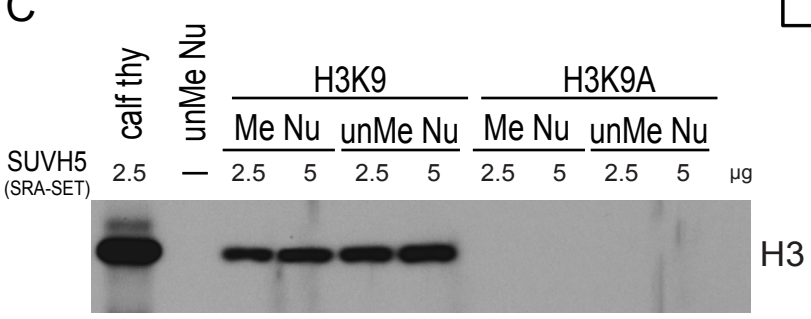
A

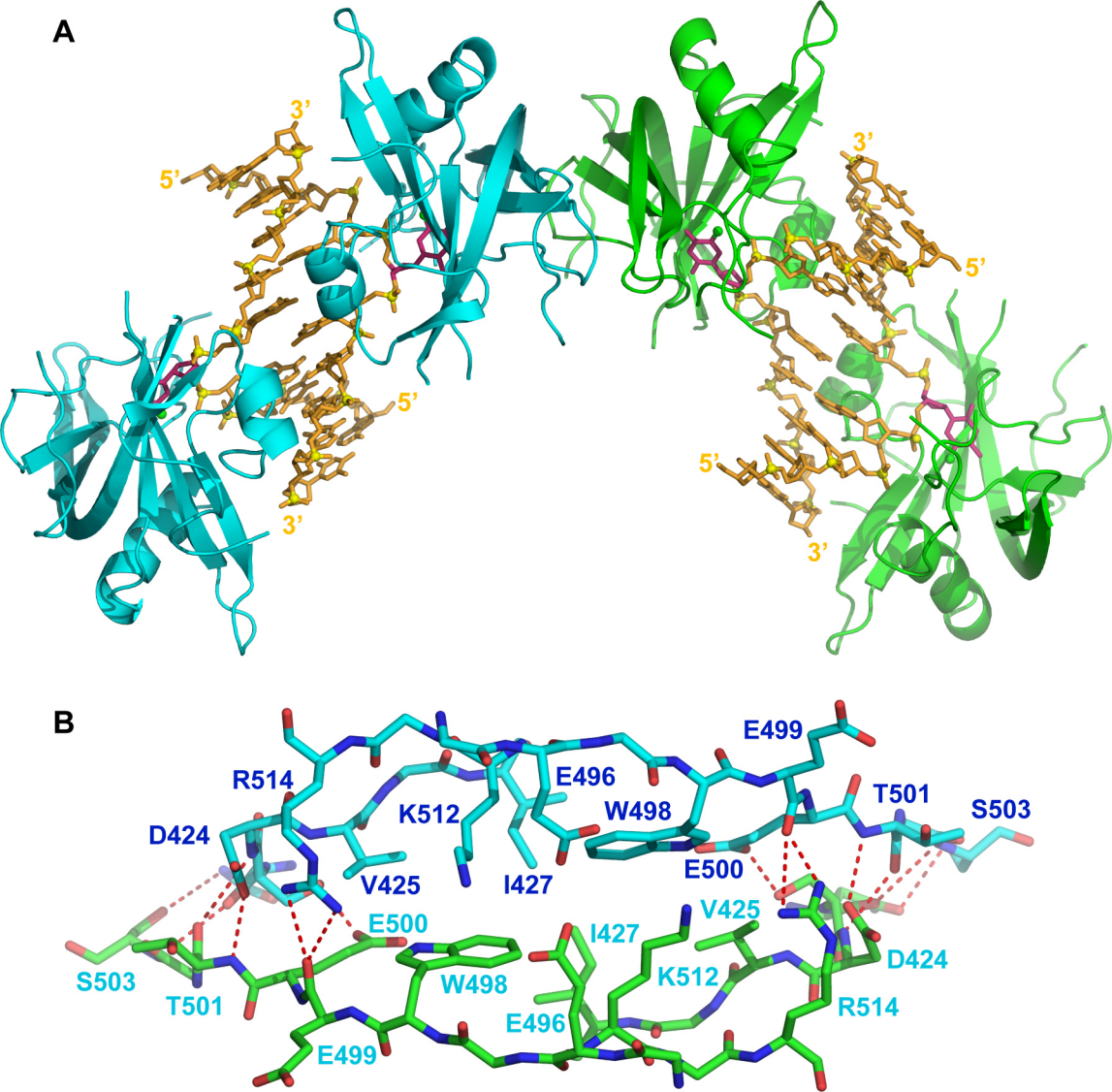

suvh4suvh5suvh6
pSUVH5::3xFlag-SUVH5
Col
Wt
Q392A
Y416AY428A
Y416A
Y428A
D418A
SUVH5
loading

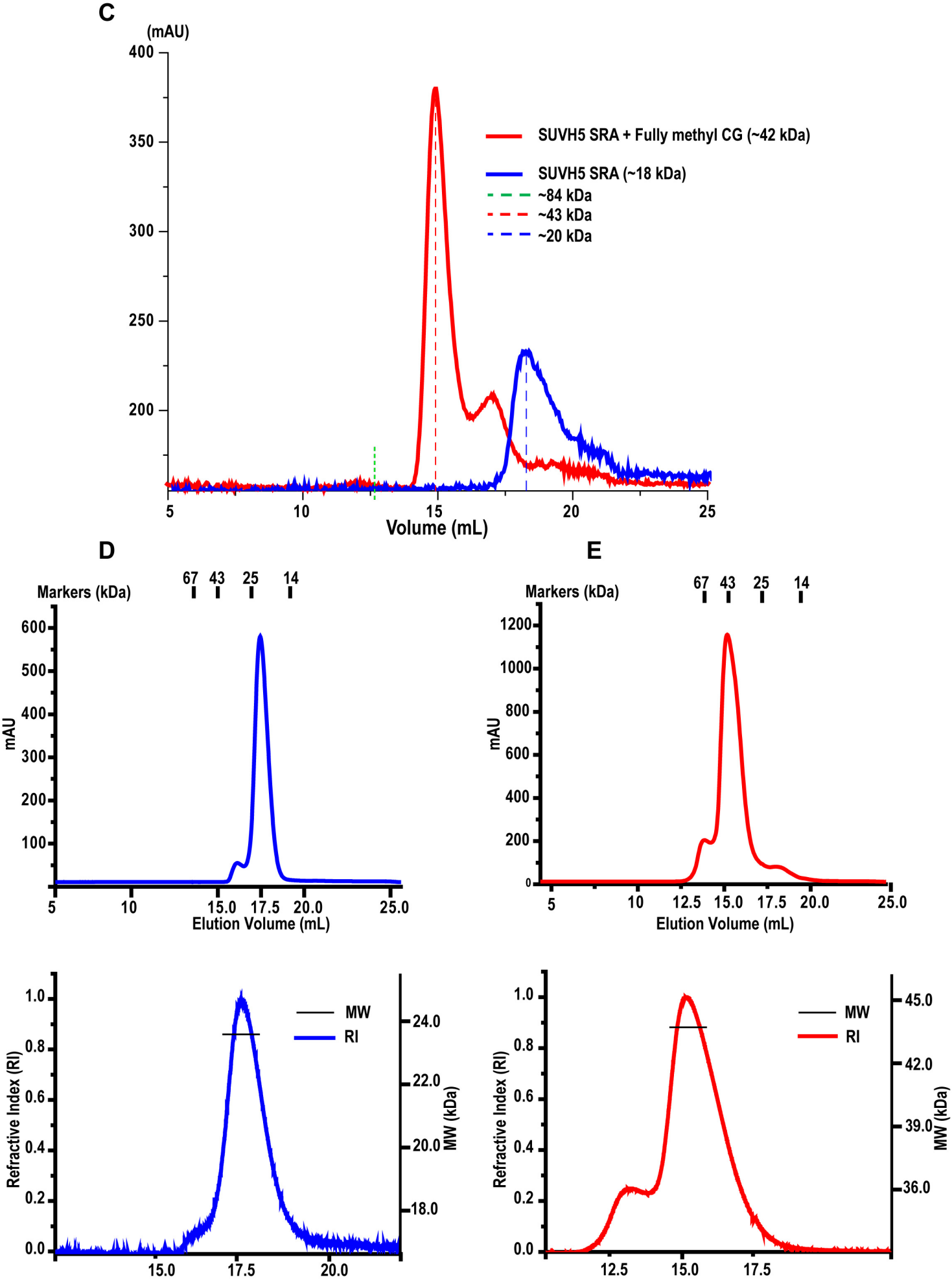
B

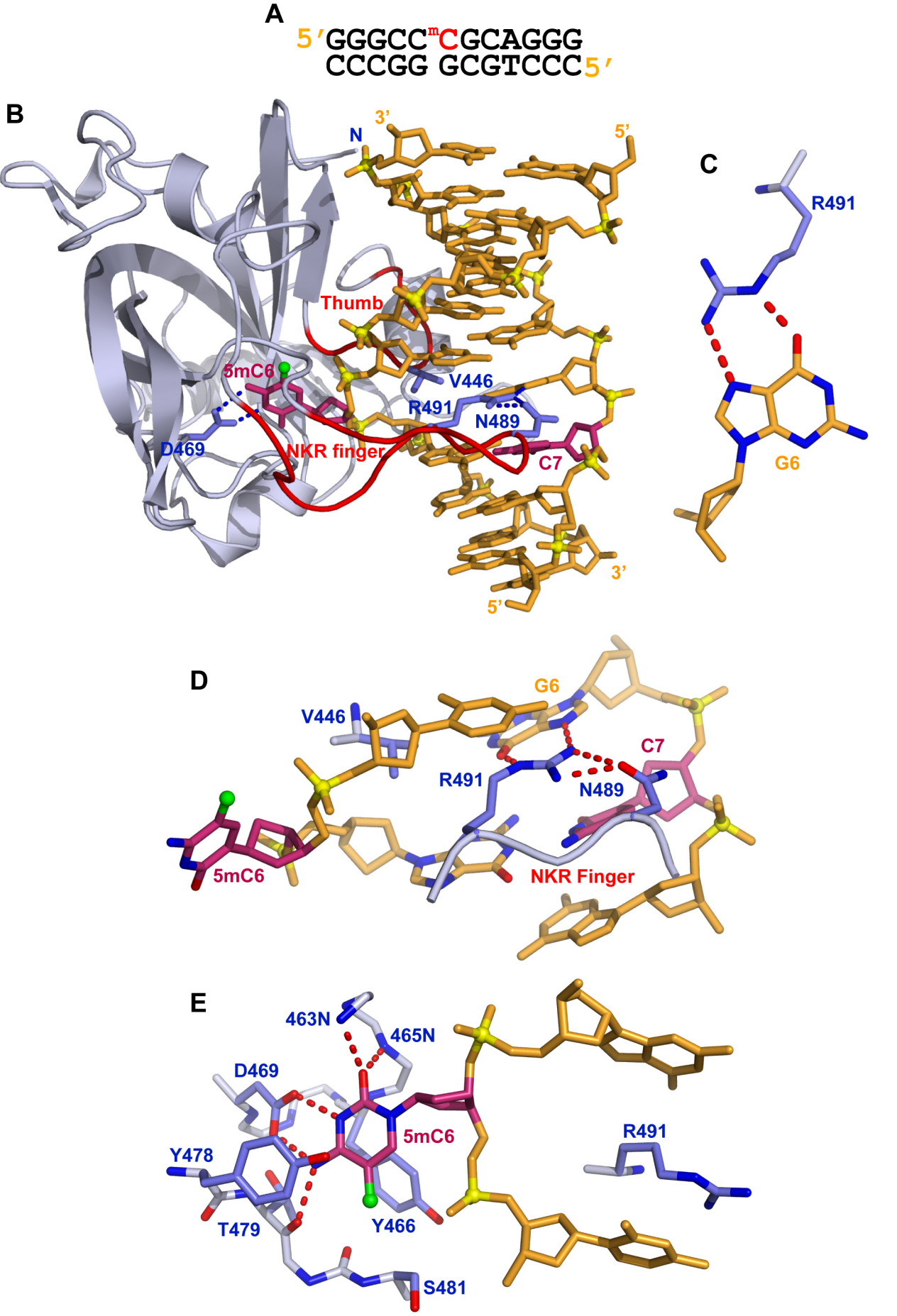

Col
H3K9me2
DAPI
Merge
72.6%
17.6%
9.8%
suvh4suvh5suvh6
Empty Vector
0.5%
33.3%
66.2%
SUVH5 Wt
36.2%
40%
17.6%
6.2%
SUVH5 Q392A
7.5%
57.5%
35%
SUVH5 Y416AY428A
7.4%
60.4%
32.2%
SUVH5 Y416A
22.9%
54.2%
22.9%
SUVH5 D418A
3.5%
39.5%
27.6%
29.4%
SUVH5 Y428A
2.3%
25.6%
45.6%
26.5%

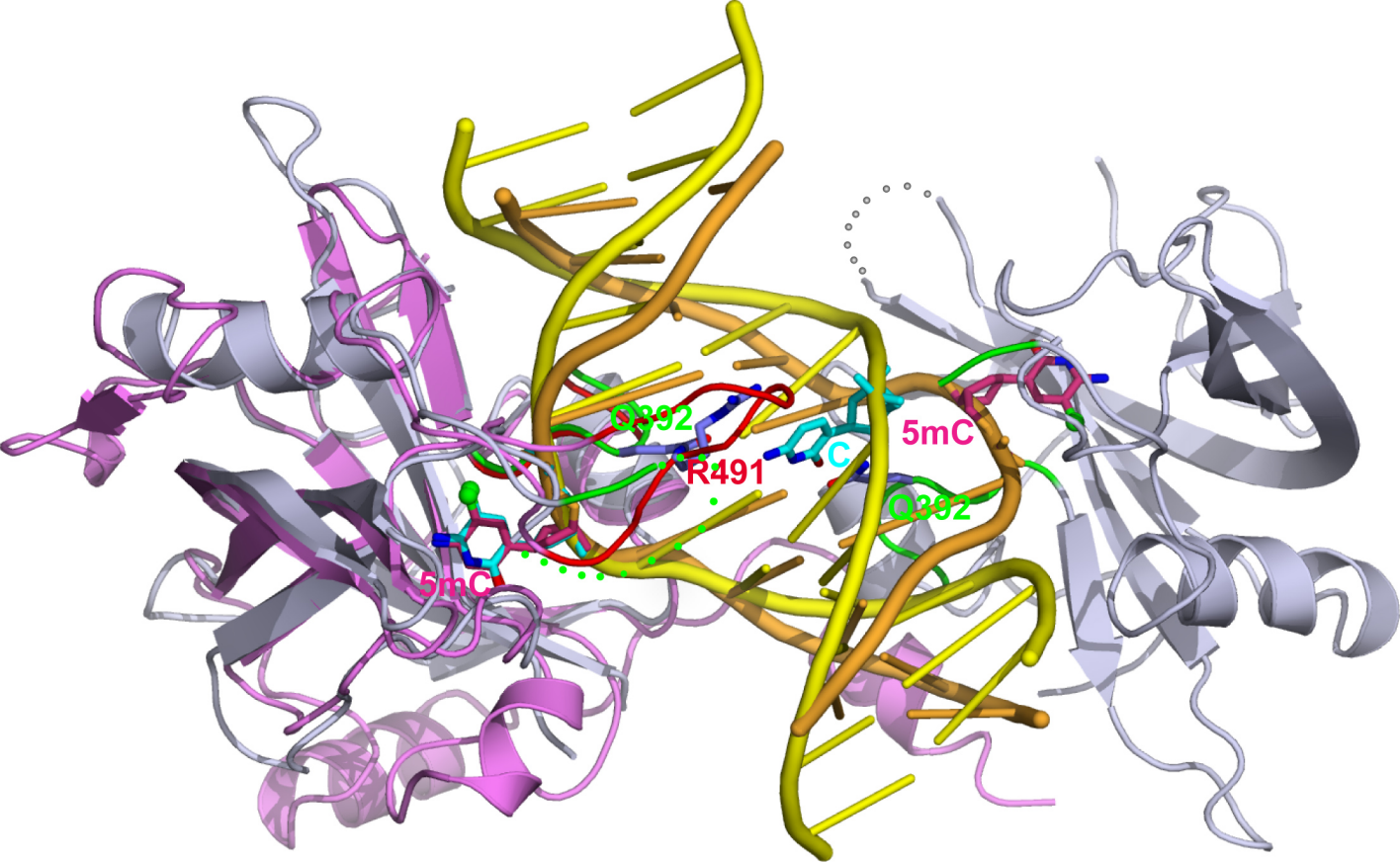
C

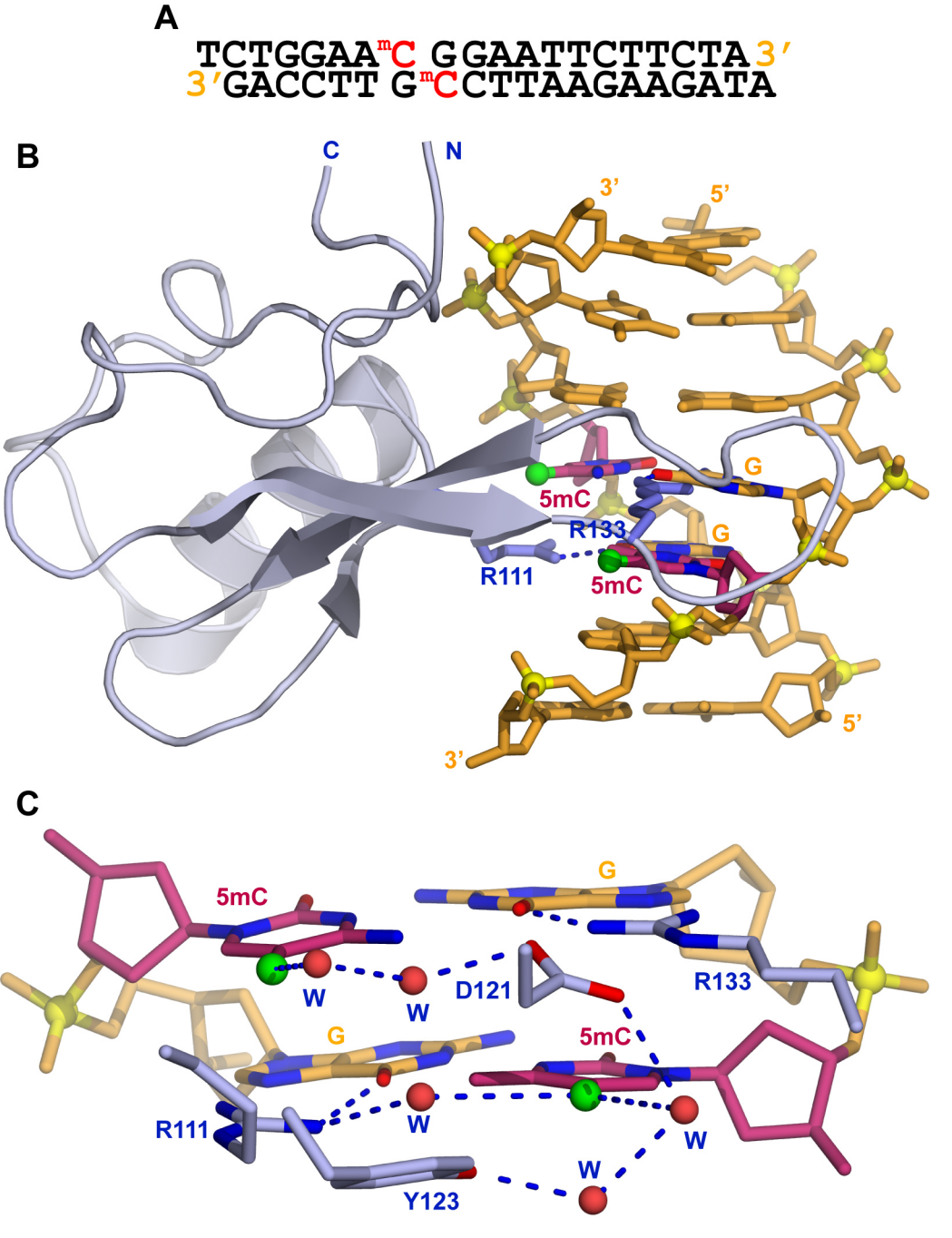

calf thy
unMe Nu
H3K9
H3K9A
Me Nu
unMe Nu
Me Nu
unMe Nu
SUVH5
(SRA-SET)
2.5
—
2.5
5
2.5
5
2.5
5
2.5
5
µg
H3


A
3'
5'
5'
3'
3'
5'
5'
3'
B
R514
D424
K512
V425
E496
I427
W498
E499
T501
S503
E500
S503
T501
E500
W498
E499
I427
V425
K512
E496
D424
R514


C
(mAU)
400
350
300
250
200
SUVH5 SRA + Fully methyl CG (~42 kDa)
SUVH5 SRA (~18 kDa)
~84 kDa
~43 kDa
~20 kDa
5
10
15
20
25
Volume (mL)
D
Markers (kDa)
67
43
25
14
600
500
400
300
200
100
mAU
5
10
15.0
17.5
20.0
25.0
Elution Volume (mL)
E
Markers (kDa)
67
43
25
14
1200
1000
800
600
400
200
0
mAU
5
10
12.5
15.0
17.5
20.0
25.0
Elution Volume (mL)
1.0
0.8
0.6
0.4
0.2
0.0
Refractive Index (RI)
MW
RI
24.0
22.0
20.0
18.0
MW (kDa)
15.0
17.5
20.0
1.0
0.8
0.6
0.4
0.2
0.0
Refractive Index (RI)
MW
RI
45.0
42.0
39.0
36.0
MW (kDa)
12.5
15.0
17.5


A
5′GGGCC mC GCAGGG
CCCGG GCGTCCC5′
B
N
3′
5′
Thumb
5mC6
V446
R491
N489
D469
NKR finger
C7
3′
5′
C
R491
G6
D
G6
V446
C7
R491
N489
5mC6
NKR Finger
E
463N
465N
D469
5mC6
R491
Y478
T479
Y466
S481


Q392
R491
5mC
C
5mC
Q392


A
TCTGGAA mC G GAATTCTTCTA 3′
3′ GACCTT G mC CTTAAGAAGATA
B
C
N
3’
5’
G
5mC
R133
G
R111
5mC
5’
3’
C
5mC
G
W
W
D121
R133
G
5mC
R111
W
W
W
Y123

SUPPLEMENTAL MATERIALS AND METHODS

Protein ex pression and purification. The cDNA encoding full-length SUVH5 was obtained from Prof. Judith L. Bender, Department of Molecular Biology, Cell Biology, and Biochemistry, Brown University. We generated a hexahistidine-sumo tagged construct containing SUVH5 residues 362-528 that corresponds to the SRA domain. Protein was expressed in *Escherichia coli* Rosetta2 DE3 (Novagen). Cells were grown at 37 ºC till OD_{660} reached 0.5-0.6, and then the temperature was decreased to 20 ºC and the culture was induced with 0.4 mM of isopropyl-1-thio-D-galactopyranoside. The cell culture was allowed to grow for 15 hrs, following which the cells were harvested and resuspended in lysis buffer (25 mM Tris-HCl, pH 7.5, 700 mM NaCl, 10 mM imidazole, 5% glycerol and 3 mM β-mercaptoethanol). Cells were lysed by three passes through an ice-cold French pressure cell press and then the lysate was clarified by centrifugation at 40,000g for 1 hr. The hexahistidine-sumo fusion protein was purified on a nickel-charged column (HisTrap HP, GE healthcare). After elution with a 600 mM imidazole containing buffer, the fusion protein was cleaved with Ulp1 protease at 15 U ml^{-1} during a 16 hr dialysis step at 4 ºC. The protein was further purified by cation-exchange (HiTrap Heparin HP) and gel filtration chromatography (HiLoad Superdex 200 26/60) and equilibrated with 15 mM Tris-HCl, pH 7.5, 150 mM NaCl, 3 mM DTT and 5% Glycerol. Purified protein was concentrated to 20 mg mL^{-1} at 4 ºC in Vivaspin 20 mL (Vivascience AG) 10,000 cut-off concentrator. The feedback inhibition method was used to obtain the seleno-methionyl derivative of the SUVH5 SRA domain; the cells were grown in M9 medium, and the protein was purified as above.

Generation of mutants. SRA mutants were generated by using QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene) on a plasmid carrying the cDNA of the SUVH5 protein and mutations were confirmed by sequencing. Mutated proteins were expressed and purified as above.

ITC measurements. The equilibrium dissociation constant of wild-type and mutant SUVH5 SRA domains (Gln392Ala, Arg379Ala, Tyr416Ala, Asp418Ala and Tyr416Ala/Tyr428Ala), as well as guanine to inosine substitutions in fully-methylated and hemi-methylated CG DNA and guanine to abasic nucleotide (dSpacer) in methylated CHH DNA were determined using a VP-ITC calorimeter (MicroCal, LLC) at 25 °C.Data were processed with MicroCal Origin software. The

SRA domain and DNA were dialyzed overnight against a buffer containing 40 mM Tris-HCl, 50 mM NaCl, and 1 mM β-mercaptoethanol, pH 7.5 at 4°C. Calorimetric titration was performed by injection of synthetically methylated DNA oligomers containing methylated cytosine bases in different sequence contexts into a cell containing the SUVH5 SRA protein. The wild-type protein and the different methylated DNA oligomer concentrations used were 50 μM to 100 μM and 300 μM to 600 μM, respectively. A fully-methylated CG DNA duplex (700 μM) was used for determining the binding affinity of mutant SRA domains (350 μM).

Crystallization. Each DNA oligomer was dissolved in 15 mM Tris-HCl, pH 7.5, 20 mM NaCl and 4 mM $MgCl_2$ prior to mixing with equimolar amounts of complementary strands to generate the hemi-methylated CG and methylated CHH duplexes. These DNA duplexes, along with the self-complementary fully-methylated CG DNA, were individually heated to 95 ºC for 5 min, and then annealed at 4 ºC for 12 hr.

Crystals of selenomethionyl-derivatized SUVH5 SRA bound to fully-methylated DNA were grown at 18 ºC using the sitting-drop method by mixing 150 nL of protein solution with 150 nL of well solution using the Mosquito crystallization robot. Crystallization of the SUVH5 SRA domain bound to fully-methylated CG DNA was undertaken under two different conditions and produced crystals in tetragonal ($P4_22_12$) and hexagonal ($P6_122$) space groups using 6 mg mL^{-1} of protein in a protein:DNA duplex molar ratio of 1:1.2. The tetragonal crystals were grown from a condition containing 19% PEG 4000, 100 mM Tris-HCl, pH 7.5 and 40 mM $CaCl_2$, whereas the hexagonal crystals were grown in condition containing 23% PEG 3350, 100 mM Bis-Tris-propane, pH 7.0 and 80 mM Mg-acetate.

Crystals of the SUVH5 SRA domain bound to hemi-methylated DNA were grown using 10 mg mL^{-1} of protein in a protein:DNA duplex molar ratio of 1:1.3 under two different conditions. The first condition contained 0.2 M K-Na tartrate, 20% PEG 3350 and the second condition contained 19% PEG4000, 100 mM MES, pH 6.5, 200 mM NaCl and 3 mM spermine.

Crystals of the SUVH5 SRA domain bound to methylated CHH DNA were grown using 10 mg mL^{-1} of protein in a protein:DNA duplex molar ratio of 1:1.5, from a condition containing 18% PEG4000, 100 mM Na-cacodylate, pH 6.5. All crystals were flash frozen at 100 K in cryoprotectant containing total of 37% ethylene glycol and PEG from the reservoir.

Data collection, structure determination and refinement. The X-ray diffraction data set for the $P4_22_12$ crystals of the selenomethionine-labeled SUVH5 SRA domain bound to fully-methylated CG DNA were collected at the Advanced Photon Source (APS) synchrotron beam line 24-IDC at three wavelengths and the structure was solved by the Multiple-Wavelength Anomalous Dispersion (MAD) technique using the SHARP program (de La Fortelle and Bricogne 1997). The PHENIX program (Adams et al. 2002) was used for automated model building and density modification. One strand of the self-complementary DNA was built into the density using the COOT program (Emsley and Cowtan 2004) and the structure of the complex was refined against 2.2 Å diffraction data with simulated annealing and restraint minimization methods using the PHENIX program. The final R-free and R-factor values of the model are 25.2% and 21.7%, respectively, for data from 30.0 Å to 2.2 Å. Missing residues include residues 437-441, 474-483 and four C-terminal amino acids, as well as the 3'-overhang residue (thymine) of the DNA duplex. The crystal of the complex in the $P6_122$ space group was solved by molecular replacement (MR) based on the apo-SRA domain from the $P4_22_12$ structure as the search model using the Molrep-Auto MR program implemented in CCP4 suite (Vagin and Teplyakov 2010). The COOT program was next used for building the self-complementary DNA duplex. The PHENIX program was used for energy minimization, non-crystallographic symmetry (NCS) averaging and restraints refinement. The final R-factor and R-free of the model are 25.0% and 30.3%, respectively, for data from 20.0 Å to 2.65 Å.

The Molrep-Auto MR program was used for solving the structure of the SRA domain bound to hemi-methylated CG DNA. The asymmetric unit contains two SRA molecules bound per DNA duplex. The non-symmetric hemi-methylated CG duplex is crossed by a non-crystallographic dyad axis in the crystal resulting in two opposite orientations of the hemi-methylated CG duplex in the complex. Therefore, the DNA was modeled in the COOT as a 10-bp duplex with half occupancy, with a second duplex of opposite orientation generated by non-crystallographic symmetry.

The PHENIX program was used for simulated annealing, NCS averaging, restraints and occupancy refinement. The structure of the complex was refined to an R-factor of 22.5% and R-free of 26.1%.

Diffraction data on the crystal of the SUVH5 SRA domain bound to methylated CHH DNA was collected at the APS synchrotron beam line 24-ID-E at single wavelength and the structure is solved by molecular replacement using the Molrep-Auto MR program. Manual

model building of the DNA duplex was carried out using the COOT program. Simulated annealing, restraint minimization and NCS averaging were performed using the PHENIX program. The final model had an R-factor of 22.9 and R-free of 27.3 for data from 30.0 Å to 2.75 Å resolution.

The crystallographic statistics for all structures presented above are listed in Supplementary Table 1.

Protein constructs and purification for EMSA assay s. GST fusions with two slightly different size SUVH5 SRA domains corresponding to amino acids 299-522 or amino acids 362-528 were generated using the Gateway cloning system (Invitrogen). These regions of the SUVH5 were amplified out of genomic DNA from Col plants using the following primers: 5'-CACCATGTGC TCA AGC AGT GGG GAT AGT GCT CGG TA-3' and 5'CTAGGGAAGCTCGGGTTG TCCAGGAATACGCCG-3' or 5'- CACCCAGATTATTGGAACTGTGCCT-3' and 5'-CGCAACCTCTTTCCAGGGAAG-3', respectively. PCR products were cloned into the pENTR or pENTR/TEV vector, respectively, and then recombined into the pDEST15 vector per manufacturer instructions. Proteins were expressed in BL21 AI cells and purified as previously described (Johnson et al. 2007; Johnson et al. 2008).

Electrophoretic Mobility Shift As says The DNA probes and conditions are largely as previously described (Johnson et al. 2007). Briefly, for testing binding on the entire battery of probes, 200 ng of each protein in binding buffer (50mM Tris, pH 6.8, 80% glycerol, 1mg/mL BSA, 14mM βme, 2 mM DTT) supplemented with NaCl to a final concentration of 50 mM, and an excess of unlabeled lambda DNA as a competitor, were incubated with ^{32}P end-labeled probe for 20 minutes at room temperature and then run on a 6% acrylamide gels with 5% glycerol in Tri-borate-EDTA buffer. Gels were then fixed, dried and exposed to film. For binding curves, the same conditions were used except the concentration of protein added was titrated. Autoradiographs were quantified using the ImageQuant TLv2005 software.

Epitope tagged SUVH5 constructs and mutagenesis.

Generation of Gateway entry clones. A genomic DNA fragment containing the upstream promoter region and open reading frame of SUVH5 were PCR amplified out of genomic DNA isolated from Col tissue using primers 5'-CACCGTCGACGCATGAATTAATTTATCACATTATAATGTC-3' and 5'-TTAGTAGAGCCTACCACTACACTCAGCGGAA-3'. The PCR product was cloned into the pDONR vector (Invitrogen) per manufacturer instructions. An amino-terminal 3xFlag-BLRP tag (5'GGTACCAACAAATGG CTGGTGGACTTAACGATATCTTCGAAGCTCAGAAGATTGAATGGCATGAGGATACTGGTGG ATCTAGCATCCCCGGACTTGAGGTTCTTTTCCAAGGACCTCTCGAGGACTACAAAGACGAT GACGACAAAGACTACAAAGACGATGACGACAAAGACTACAAAGACGATGACGACAAAAAG CTTGATATCGTGGTACC-3') was inserted into a Kpn I restriction site engineered into the SUVH5 genomic sequence at the start codon using the following primers 5'-GTAGGCGTTGGGCATCATGGTACCTTCAGAGTCATCAATA-3' and 5'TATTGATGACTCTGAAG GTACCATGATGCCCAACGCCTAC-3' and the quikchange site directed mutagenesis kit (Stratagene). Mutations within the SRA domain of this construct were also introduced by site directed mutagenesis using the primers listed (Supplementary Table 2).

Generation of gatew ay destinati on clones and *Arabidopsis* plant lines. The described pDONR constructs were digested with the Nhe I restriction enzyme and then recombined into a modified gateways destination vector, described previously (Johnson et al., 2008), except the BASTA gene was replace with the Hygromycin resistance gene, using Kpn I and Sac I restriction sites. DNA constructs were transformed into the *Agrobacterium* AGLO strain by electroporation. *Arabidopsis* plants carrying the *suvh4* (SALK_041474), *suvh5* (GABI kat 263C05) and *suvh6* (Garlic_1244_F04) mutant alleles were transformed with each SUVH5 construct, using the floral dip method described in (Clough and Bent 1998). Seeds from transformed plants were selected with hygromycin.

Western and southern blotting. The expression of eptitope-tagged lines were assessed by western blotting using protein extracted from T_1 flowers and the ANTI-FLAG M2 Monoclonal Antibody-Peroxidase Conjugate (Sigma A 8592) at a dilution of 1:5000. Complementation of the epitope-tagged wild-type and SRA mutant SUVH5 constructs were assessed by Southern blotting using Msp I digested genomic DNA extracted from the same T_1 plants as were used for the western blots and a probe specific of the *Ta3* locus as previously described (Johnson et al. 2008).

Nuclei isolation and H3K9 imm unofluorescence. Nuclei were isolated from seedlings as described previously (Jasencakova et al., 2003). A primary antibody to H3K9 dimethylation (Abcam 1220) was used at a dilution of 1:100 and anti-mouse Alexa Fluor 488 (Invitrogen) was used as a secondary antibody at a dilution of 1:150. Nuclei were visualized as described previously (Li et al., 2006).

Nucleosome preparati on and HMTase assa ys. Histone methylation assays were done as described previously (Johnson et al., 2008) using 2.5-5 µg of SRA-SET SUVH5 protein (aa 362-794). Hexahistidine-sumo tagged SET-SRA construct was expressed in *Escherichia coli* Rosetta2 DE3 (Novagen). The hexahistidine-sumo fusion protein was purified on a nickel- charged column and subsequently cleaved with Ulp1 protease at 15 U ml^{-1} during a 16 hr dialysis step at 4 ºC. Nucleic acid contaminants were removed by further purification with cation- exchange (HiTrap Heparin HP) column. Final purification involves gel filtration chromatography (HiLoad Superdex 200 26/60) using the buffer of 15 mM Tris-HCl, pH 7.5, 150 mM NaCl, 3 mM DTT and 5% Glycerol. Activity was assessed using mononucleosomes assesmbled by salt gradient dialysis from recombinant histone octamers and a 209-bp DNA construct containing the 601 nucleosome positioning sequence. Histone octamers were reconstituted from *Xenopus laevis* core histones expressed in *E. coli* essentially as described previously (Luger et al. 1999). The DNA construct encompassing the 601 nucleosome positioning sequence (Lowary and Widom 1998) was generated by PCR. For CG-methylated mononucleosomes, 601 DNA was treated with the CG-specific methyltransferase M.SssI (NEB) prior to nucleosome reconstitution. Methylation was confirmed by restriction analysis with Hpa II. Assembly of mononucleosome particles was verified by electrophoresis on 4.5% polyacrylamide gels in Tris-glycine-EDTA buffer.

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., .Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. 2002. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**(Pt 11): 1948-1954.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. 2008. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455**(7214): 818-821. .

Avvakumov, G.V., Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H., and Dhe-Paganon, S. 2008. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455**(7214): 822-825.

Clough, S.J. and Bent, A.F. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**(6): 735-743.

de La Fortelle, E. and Bricogne, G. 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods *Methods in Enzymology* **276**: 472-494

Emsley, P. and Cowtan, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2126-2132.

Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E., and Cheng, X. 2008. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455**(7214): 826-829.

Ho, K.L., McNae, I.W., Schmiedeberg, L., Klose, R.J., Bird, A.P., and Walkinshaw, M.D. 2008. MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol Cell* **29**(4): 525-531.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. 2007. The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr Biol* **17**(4): 379-384.

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. 2008. SRA-domain proteins required for DRM2-mediated de novo DNA methylation. *PLoS Genet* **4**(11): e1000280.

Lowary, P.T. and Widom, J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**(1): 19-42.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. 1999. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* **304**: 3-19.

Vagin, A. and Teplyakov, A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 1): 22-25.

Supplementary Table 1. Crystallographic data and statistics

Data collection and refinement statistics					
Crystal	SUVH5SRA-Fully-methylated CG DNA		SUVH5SRA-Fully-methylated CG DNA	SUVH5SRA-Hemi-methylated CG DNA	SUVH5SRA-Methylated CNN DNA
Beam line	APS-24ID-C		APS-24ID-E	APS-24ID-C	APS-24ID-E
	Peak	Inflection			
Wavelength	0.97918	0.97929	0.97918	0.97918	0.97918
Space group	$P4_22_12$		$P6_122$	$P4_2$	$P6_122$
Unit cell					
a, b, c (Å)	76.57, 76.57,73.91		104.46,104.46,176.69	76.38, 76.38, 74.25	102.97,102.97,170.49
A, β, γ (^{o})	90, 90, 90		90, 90, 120	90, 90, 90	90,90,120
Resolution (Å)	30-2.20 (2.28-2.20)[a]		20-2.65 (2.74-2.65)	20-2.37 (2.45-2.37)	30-2.75(2.85-2.75)
R_{sym}	0.051 (0.51)		0.134 (0.83)	0.068 (0.384)	0.088 (0.7)
I/σ (I)	49.4 (4.1)		27.1 (3.0)	30.1 (1.8)	26.0 (2.3)
Completeness (%)	99.7 (99.9)		100 (100)	98.8 (90.4)	99.8 (100)
Redundancy	11.4 (11.4)		28.1 (28.9)	7.3 (4.5)	8.5 (9.5)
Number of unique reflections	11,770		17,149	17,234	14,578
R_{work}/R_{free} (%)	21.7/25.2		25.0/30.3	22.5 /26.1	22.9 /27.3
Number of non-H atoms					
Protein	1136		2243	2173	2240
DNA	203		426	406	425
Water	63		36	51	18
Other			1	1	
Average B factors (Å^2)					
Protein	44.2		51.9	58.5	66.9
DNA	42.4		54.9	53.0	83.0
Water	47.6		38.2	57.5	67.3

R.m.s. deviations				
Bond lengths (Å)	0.007	0.006	0.007	0.007
Bond angles (°)	1.19	1.02	1.26	1.29

[a] Highest resolution shell (in Å) shown in parentheses.

Supplementary Table 2. Mutagenesis primer sets

Mutation	Orientation	Primer sequence 5' to 3'
Q392A	Forward	CTTCTTGGTATACATAGACCAAGTGCAAGTGGTATCGACTATATGAAAGA
Q392A	Reverse	TCTTTCATATAGTCGATACCACTTGCACTTGGTCTATGTATACCAAGAAG
Y416A	Forward	AGTATTGTATCCTCGGGAGGTGCTAATGACGTGCTTGATAACTC
Y416A	Reverse	GAGTTATCAAGCACGTCATTAGCACCTCCCGAGGATACAATACT
Y428A	Forward	TAACTCTGATGTCTTGATCGCCACGGGTCAAGGCGGAAAT
Y428A	Reverse	ATTTCCGCCTTGACCCGTGGCGATCAAGACATCAGAGTTA
Y416AD418A	Forward	GAGTATTGTATCCTCGGGAGGTGCTAATGCCGTGCTTGATAACTCTGATG
Y416AD418A	Reverse	CATCAGAGTTATCAAGCACGGCATTAGCACCTCCCGAGGATACAATACTC
D418A	Forward	CCTCGGGAGGTTATAATGCCGTGCTTGATAACTC
D418A	Reverse	GAGTTATCAAGCACGGCATTATAACCTCCCGAGG
N417A	Forward	CGAGTATTGTATCCTCGGGAGGTTATGCTGACGTGCTTGATAAC
N417A	Reverse	GTTATCAAGCACGTCAGCATAACCTCCCGAGGATACAATACTCG
N417K	Forward	AGTATTGTATCCTCGGGAGGTTATAAGGACGTGCTTGATAA
N417K	Reverse	TTATCAAGCACGTCCTTATAACCTCCCGAGGATACAATACT

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure S1: Multiple sequence alignment of the SRA domains of the SET domain associated SUVH proteins and the RING domain associated (ORTH 1-4, ORTH-L and UHRF1) family proteins. Sequence numbering of the SUVH family proteins is based on SUVH5 whereas for the RING domain associated proteins it is based on UHRF1. Secondary structural elements of the SUVH5 SRA are indicated above the sequence and disordered regions are shown by ## symbols. Residues highlighted in a background color-code correspond to their conservation level: fully-conserved, red; conservative substitutions, yellow. The thumb and NKR finger segments corresponding to the UHRF1 structure are underlined in black. The residue that inserts into the DNA duplex and displaces the 5mC residue in the SUVH5 SRA complex is indicated as an inverted red triangle. Green and blue upright triangles correspond to residues that replace the looped out 5mC and mask the unmodified C in the UHRF1 SRA complex, respectively. Filled green circles designate residues that interact with the 5mC in the binding pockets of the SUVH5 SRA and UHRF1 SRA domains.

Supplemental Figure S2. EMSA assays using untagged SRA domain proteins and binding curves using GST fusion proteins. (A) EMSA experiments using untagged 306-522 **(Left)** or 362-528 **(Right)** SUVH5 SRA domain proteins for comparison with the GST299-522 and GST362-528 constructs, respectively. The context and methylation state of each radiolabeled double stranded DNA oligonucelotide are indicated above. Un, unmethylated; FM, fully methylated; HM hemi-methylated;--, indicates no protein. **(B-D)** EMSA experiments using GST299-522 **(Upper)** or GST362-528 **(Lower)** SUVH5 SRA domain fusion proteins and radiolabeled double stranded DNA oligonucelotide in the **(B)** CG, **(C)** CHG, or **(D)** CHH sequence contexts. The concentration of SRA protein was titrated and the total protein used in each lane is indicated in ng.

Supplemental Figure S3. ITC measurements of binding of the SUVH5 SRA to methylated CHG DNA.

The measured binding parameters are K_D = 6.6 μM and N = 0.55.

Supplemental Figure S4. Schematic representation showing the intermolecular contacts between the SUVH5 SRA domains and the 5mC-containing DNA duplexes.

The stoichiometry of all three complexes was two SRA domains bound per DNA duplex.

(A) Intermolecular contacts in the SUVH5 SRA complex with fully-methylated CG DNA.

(B) Intermolecular contacts in the SUVH5 SRA complex with hemi-methylated CG DNA.

(C) Intermolecular contacts in the SUVH5 SRA complex with methylated CHH DNA.

Intermolecular hydrogen bond contacts are shown as arrows, while stacking interactions and van der Waals contacts are shown as series of short, parallel lines. The residues in blue are from one SRA and the residues in black are from the second SRA. The Gln392 residues from both SRA molecules are colored in red.

(D) Superposition of complexes contianing the SUVH5 SRA bound to fully-methylated CG (blue ribbon representation) and hemi-methylated CG (magneta line representation) DNA. The SRA domains on the left are superimposed to allow visualization of the relative displacement of the SRA domains on the right.

(E) Superposition of complexes containing the SUVH5 SRA bound to fully-methylated CG (blue ribbon representation) and methylated CHH (magneta line representation) DNA. The SRA domains on the left are superimposed to allow visualization of the relative displacement of the SRA domains on the right.

Supplemental Figure S5. ITC measurements of the binding of various SUVH5 SRA domain mutants to fully-methylated CG DNA and of the wild-type SUVH5 SRA to guanine-substituted CG and CHH DNAs.

(A) Asp418Ala mutant. The measured binding parameters are K_D = 21.0 μM.

(B) Arg379Ala mutant. The binding is weak and endothermic.

(C) Tyr416Ala mutant. The binding is weak and endothermic.

(D) Gln392Ala mutant. The binding is too weak to measure.

(E) Tyr416Ala/Tyr428Ala double mutant. The binding is too weak to measure.

(F) Guanine to inosine substitution in fully-methylated CG DNA. The measured binding parameters are K_D of 0.75 μM and N of 0.52.

(G) Guanine to inosine substitution in hemi-methylated CG DNA. The measured binding parameters are K_D of 6.2 μM and N of 0.62.

(H) Guanine to abasic site (designated dS) substitution in methylated CHH DNA. The measured binding parameters are K_D of 2.7 μM and N of 1.1.

Supplemental Figure S6. Western and Immunofluorescence analysis of various SUVH5 transgenes.

(A) Western blot using an antibody against the Flag epitope showing the expression the wild-type and mutant versions of the SUVH5 protein in the T_2 generation. Protein extracted from the non transgeneic colombia (Col) ecotype was used as a negative control. An unknown protein that cross reacts with the Flag antibody and serves as an internal loading control is shown in the lower panel.

(B) Immunofluorescence detection of H3K9 dimethylation in nuclei isolated from the indicated genotype (Top). Images representing each class of nuclei are shown and percentages out of 200 nuclei are indicated (left).

(C) Histone methyltransferase assays using a SUVH5 SRA-SET protein (aa 362-794) and either calf thymus histones (Calf thy), wild-type mono-nucleosomes (H3K9) assembled using methylated (Me Nu) or unmethylated (unMe Nu) DNA, or mutant mono-nucleosomes (H3K9A) assembled using methylated or unmethylated DNA as a substrate. ^{3}H radiolabeled SAM was supplied as the methyl donor. The concentration of SUVH5 protein used for each assay is indicated in μg and the position of histone 3 (H3) is indicated (right).

Supplemental Figure S7. Packing arrangement in the crystal of two molecules of the SUVH5 SRA domain bound to a fully-methylated CG DNA duplex and gel-filtration and tandem gel-filtration-MALS analyses of SUVH5 SRA in either the free state or bound to fully-methylated CG DNA.

(A) Crystal packing interactions between the SRA domains of adjacent complexes of the SUVH5 SRA domain bound to fully-methylated CG DNA (two SRAs bound per duplex). The SRA domains of one complex are colored in cyan and those of the other are colored in green.

(B) The interfacial interactions between the SRA domains of adjacent complexes of the SUVH5 SRA domain bound to fully-methylated CG DNA (two SRAs bound per duplex). The SRA domains of one complex are colored in green and those of the other are colored in cyan.

(C) Superposition of the gel-filtration chromatogram for the SUVH5 SRA in the free state (blue) and bound to fully-methylated CG DNA (red). Dashed blue and red lines indicate the peak-elution volume of the free SUVH5 SRA and two SUVH5 SRA domains bound to fully-methylated CG DNA complex, respectively. The dashed green line represents the elution volume equivalent to ~ 84 kDa corresponding to a complex consisting of a pair of SUVH5 SRA domains bound to fully-methylated CG DNA.

(D) Tandem gel-filtration-MALS analyses of SUVH5 SRA domain in the free state. Gel-filtration and MALS chromatogram are indicated in blue on top and bottom panels respectively. Horizontal solid black line represents a mean molecular weight estimated by MALS.

(E) Tandem gel-filtration-MALS analyses of the complex of two SUVH5 SRA domains bound to fully-methylated CG DNA. Gel-filtration and MALS chromatogram are indicated in red on top and bottom panels respectively. Horizontal solid black line represents a mean molecular weight estimated by MALS.

Supplemental Figure S8. Published structure of the UHRF1 SRA domain bound to hemi-methylated CG DNA (Arita et al. 2008; Avvakumov et al. 2008; Hashimoto et al. 2008)**.**

(A) Sequence of the hemi-methylated CG 12-mer DNA duplex containing a centrally located 5mC-G step on the top strand and a complementary unmethylated C-G step on the partner strand.

(B) Stick (DNA) and ribbon (protein) representation of the 2.2 Å crystal structure of the 1:1 UHRF1 SRA - hemi-methylated CG DNA duplex complex (PDB: 3CL2). The DNA is colored in

orange except for the 5mC5 base, which is colored in purple. The 5-methyl group is shown as a small green sphere. Backbone phosphorus atoms are shown as yellow balls and the 5'- and 3'-ends of the DNA are labeled. The SRA domain is colored in blue. The flipped-out 5mC6 residue is recognized by a single SRA domain of UHRF1 in the complex. Two loops, designated the NKR finger and thumb, are colored in red. Arg491 from the NKR finger partially fills the hole left by the flipped out 5mC6 base, where it functions by interacting with the orphaned G6.

(C) The side chain of Arg491, which inserts into the DNA duplex, forms intermolecular hydrogen bonds with the Hoogsteen edge of the G6 base.

(D) The side chain of Arg491 forms a network of intramolecular hydrogen bonds with the backbone of the NKR finger segment. The side chain of Asn489 prevents potential flipping out of the un-methylated C6 base on the partner strand. Val446, which corresponds in sequence to Gln392 in the SUVH5 SRA-DNA complex, interacts with the hydrophobic side chain of Arg491.

(E) Interaction of the looped out 5mC5 base with residues lining the binding pocket. The 5mC5 base is positioned between the aromatic rings of Tyr466 and Tyr478 with its Watson-Crick edge forming hydrogen bonds with the Asp469 side chain and with several other protein backbone atoms.

Supplemental Figure S9. Superimposition of the SUVH5 SRA domain bound to fully-methylated CG DNA (2 SRAs bound per duplex) and the UHRF1 SRA domain bound to hemi-methylated CG DNA (1 SRA bound per duplex). The two SUVH5 SRA domains and the fully-methylated CG DNA are colored in blue and orange, respectively. The UHRF1 SRA domain and the hemi-methylated CG DNA are purple and yellow, respectively. The thumb segments of the SUVH5 and UHRF1 SRA domains are colored in green and red, respectively. The NKR finger segment of the SUVH5 SRA domain is disordered and is shown by a dotted light green line, whereas its counterpart in the UHRF1 SRA domain is shown by a solid red line. Symmetrical 5mC bases in the SUVH5 SRA domains are colored in magenta, while the 5mC and the C in the UHRF1 SRA domain are shown in cyan. Green spheres represent the 5-methyl group of the 5mC bases in both structures. The SUVH5 SRA domain recognizes the minor groove of the DNA whereas the UHRF1 SRA domain mainly recognizes the major groove of the DNA. Note that the base that inserts into the DNA duplex and replaces the flipped out 5mC is

Gln392 of the thumb loop in the SUVH5 SRA-DNA complex and Arg491 of the NKR finger in the UHRF1 SRA-DNA complex.

Supplemental Figure S10. Structure of MeCP2 bound to full y-methylated CG DNA (Ho et al. 2008)**.**

(A) The crystal structure of the methyl C-G binding domain (MBD) of MeCP2 bound to fully-methylated CG DNA (PDB: 3C2I).

(B) The MBD domain recognizes the major groove centered on the 5mC-G steps through a distinct set of asymmetrical protein-DNA interactions.

(C) Expanded view of water-mediated intermolecular contacts within the major groove centered on the 5mC-G steps.

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. 2002. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**(Pt 11): 1948-1954.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. 2008. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455**(7214): 818-821.

Avvakumov, G.V., Walker, J.R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C.H., and Dhe-Paganon, S. 2008. Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455**(7214): 822-825.

Clough, S.J. and Bent, A.F. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**(6): 735-743.

de La Fortelle, E. and Bricogne, G. 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods *Methods in Enzymology* **276**: 472-494

Emsley, P. and Cowtan, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2126-2132.

Hashimoto, H., Horton, J.R., Zhang, X., Bostick, M., Jacobsen, S.E., and Cheng, X. 2008. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455**(7214): 826-829.

Ho, K.L., McNae, I.W., Schmiedeberg, L., Klose, R.J., Bird, A.P., and Walkinshaw, M.D. 2008. MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol Cell* **29**(4): 525-531.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. 2007. The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr Biol* **17**(4): 379-384.

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. 2008. SRA-domain proteins required for DRM2-mediated de novo DNA methylation. *PLoS Genet* **4**(11): e1000280.

Lowary, P.T. and Widom, J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**(1): 19-42.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. 1999. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* **304**: 3-19.

Vagin, A. and Teplyakov, A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 1): 22-25.