Molecular Cell, Volume *43*

Supplemental Information

PHD Finger Recognition of Unmodified Histone H3R2 Links UHRF1 to Regulation of Euchromatic Gene Expression

Eerappa Rajakumara, Zhentian Wang, Honghui Ma, Lulu Hu, Hao Chen, Yan Lin, Rui Guo, Feizhen Wu, Haitao Li, Fei Lan, Yujiang Shi, Yanhui Xu, Dinshaw J. Patel, and Yang Shi

Figure S1

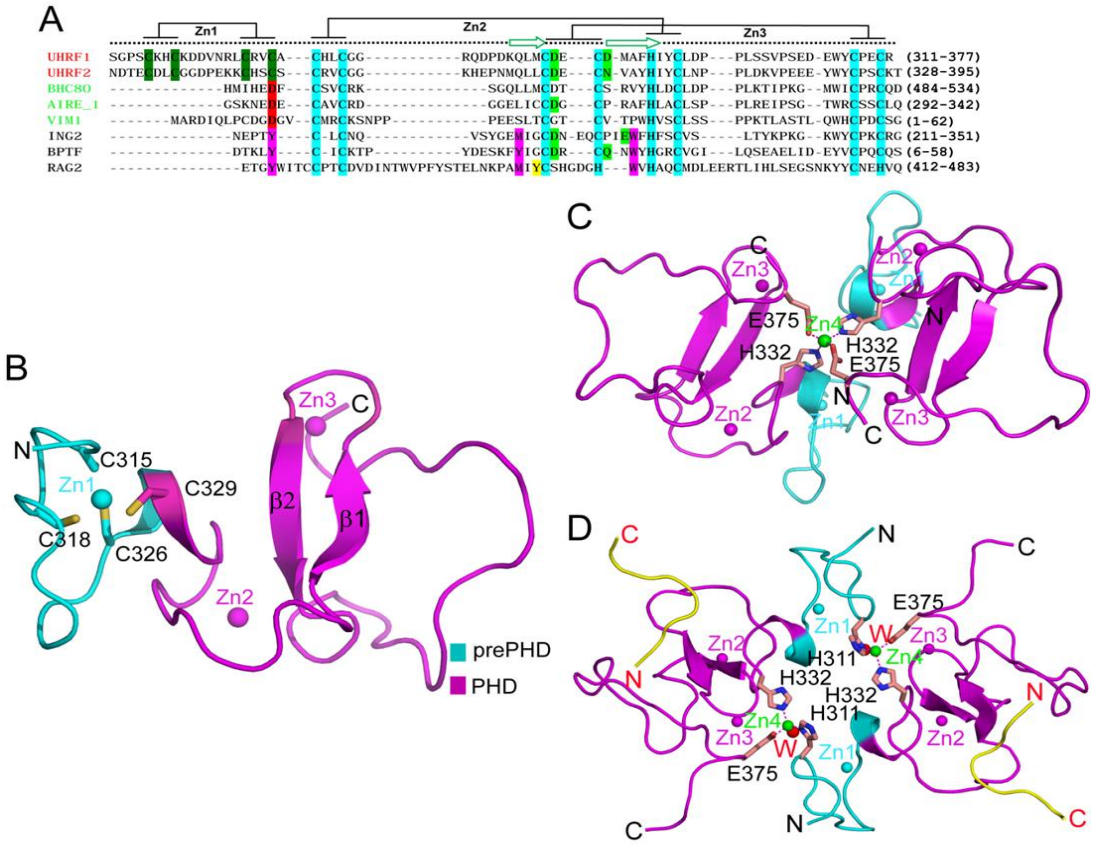
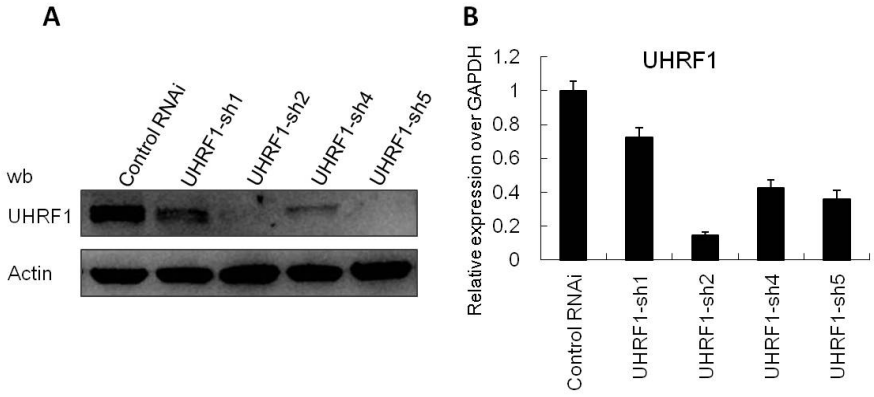


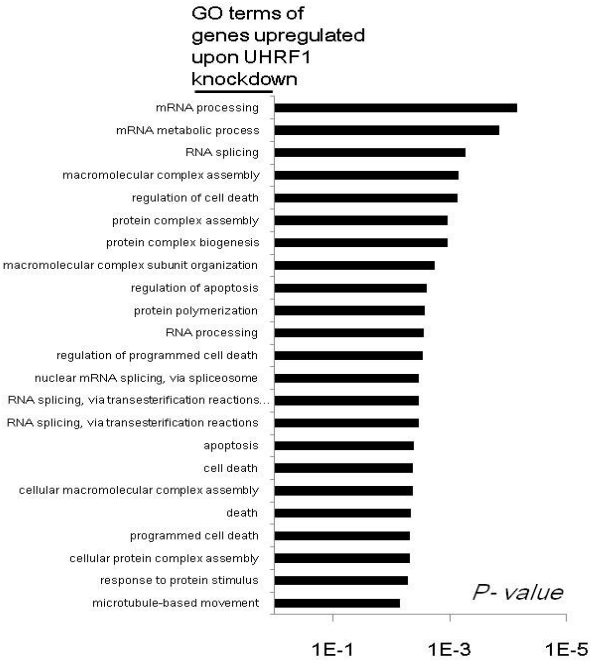
Figure S2

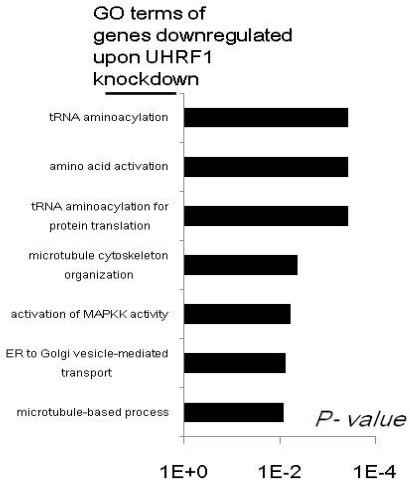


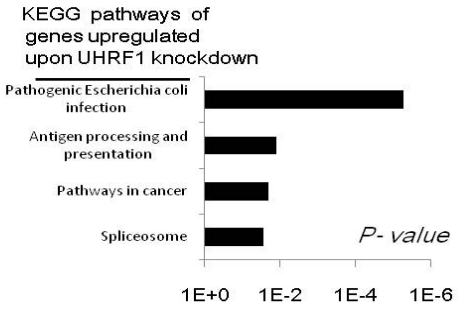
C

Knockdown UHRF1	Fold of Up-regulation	Fold of Down-regulation
	x≥1.4	x≥1.4
UHRF1-sh2	1953	1145
UHRF1-sh5	1459	2473

Figure S3


GO terms of genes upregulated upon UHRF1 knockdown
mRNA processing
mRNA metabolic process
RNA splicing
macromolecular complex assembly
regulation of cell death
protein complex assembly
protein complex biogenesis
macromolecular complex subunit organization
regulation of apoptosis
protein polymerization
RNA processing
regulation of programmed cell death
nuclear mRNA splicing, via spliceosome
RNA splicing, via transesterification reactions...
RNA splicing, via transesterification reactions
apoptosis
cell death
cellular macromolecular complex assembly
death
programmed cell death
cellular protein complex assembly
response to protein stimulus
microtubule-based movement
P- value
1E-1
1E-3
1E-5


GO terms of genes downregulated upon UHRF1 knockdown
tRNA aminoacylation
amino acid activation
tRNA aminoacylation for protein translation
microtubule cytoskeleton organization
activation of MAPKK activity
ER to Golgi vesicle-mediated transport
microtubule-based process
P- value
1E+0
1E-2
1E-4


KEGG pathways of genes upregulated upon UHRF1 knockdown
Pathogenic Escherichia coli infection
Antigen processing and presentation
Pathways in cancer
Spliceosome
P- value
1E+0
1E-2
1E-4
1E-6

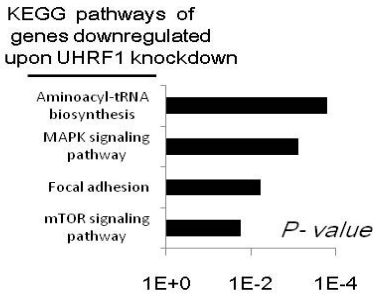

KEGG pathways of genes downregulated upon UHRF1 knockdown
Aminoacyl-tRNA biosynthesis
MAPK signaling pathway
Focal adhesion
mTOR signaling pathway
P- value
1E+0
1E-2
1E-4

Figure S4

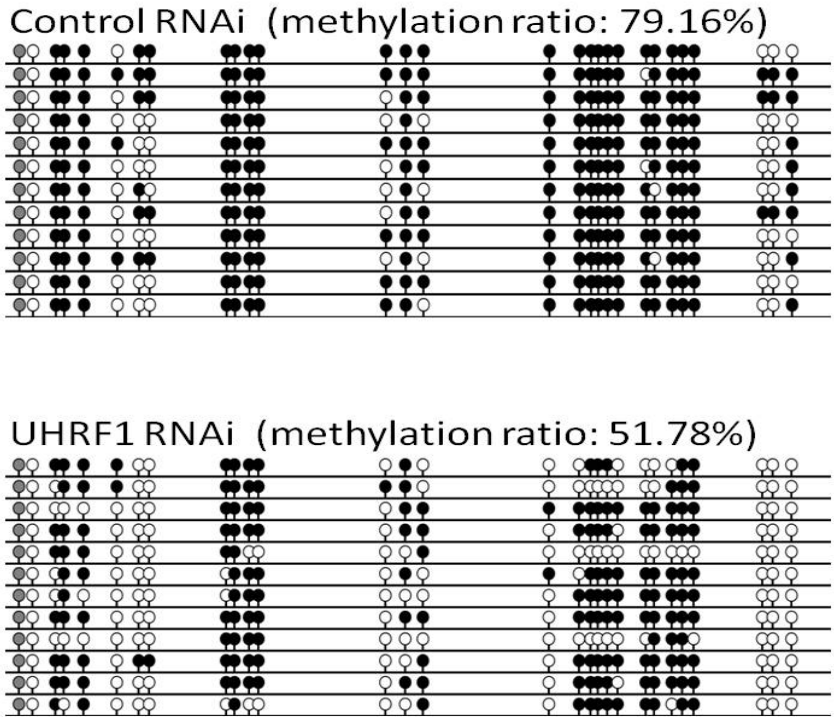


Figure S5

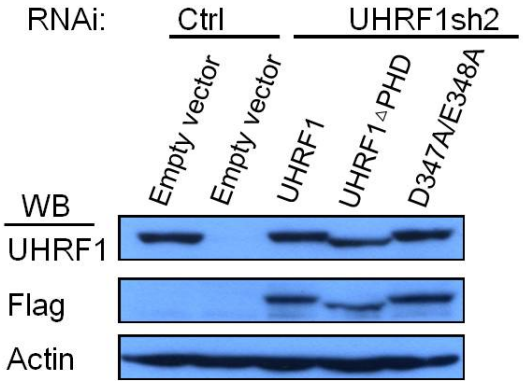
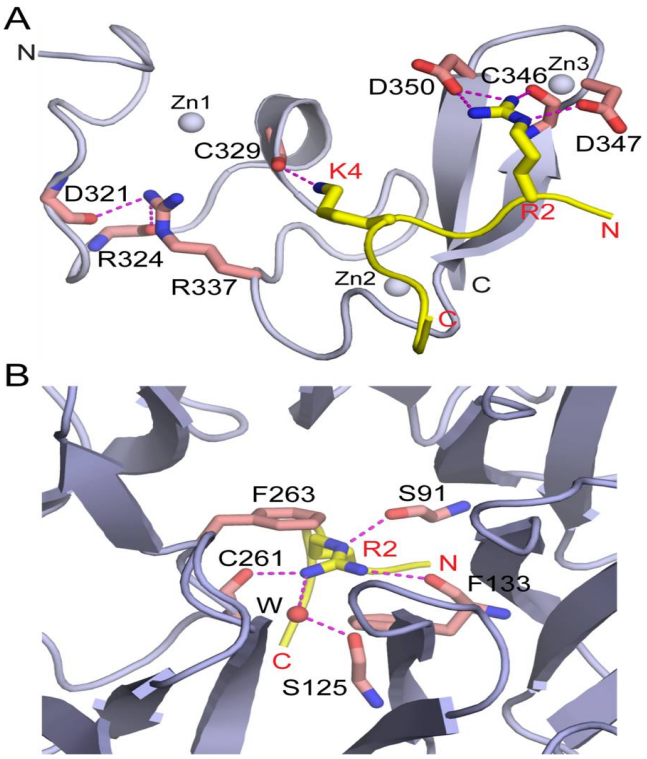


Figure S6



Supplemental figure legends

Figure S1. Sequence Alignment of PHD Fingers, the Crystal Structure of PHD_{UHRF1} in the Free State, and Details of Zn-mediated Dimerization of PHD_{UHRF1} in Free and H3(1-9) Bound States

(A) Multiple sequence alignment of PHD fingers involved in recognition of H3R2me0 (red color), H3K4me0 (green color) or H3K4me3 (black color). Three Zn-chelating residues that are found in the PHD_{UHRF1} monomer, are connected by solid lines. The secondary structural elements of PHD_{UHRF1} are indicated above the sequence (β-strands are in green arrows and loops are in dotted lines). Cys residues highlighted

in green background color coordinate the Zn1 atom of the pre-PHD in PHD_{UHRF1}, whereas residues that chelate Zn2 and Zn3 atoms in the canonical PHD finger fold are highlighted in cyan background color. Two Asp residues involved in the recognition of unmodified H3R2 in PHD_{UHRF1} are highlighted in a green color background. Residues highlighted in red and magenta color backgrounds are involved in recognition of H3K4me0 and H3K4me3, respectively.

(B) The crystal structure of PHD_{UHRF1} in the free state. The pre-PHD and canonical PHD fingers are colored in cyan and magenta, respectively. The side chains coordinated to Zn1 are labeled.

(C) A single Zn ion mediated dimerization of PHD_{UHRF1} (ribbon representation) in free state. All Zn atoms are represented as spheres. Canonical PHD finger with corresponding Zn-atoms are represented in magenta. Pre-PHD motif is represented in cyan. Zn atom coordination to residues involved in PHD finger dimerization are represented in green sphere(s). Red dashed lines indicate the interactions that bridge two PHD fingers through Zn-coordination.

(D) A pair of Zn ions mediating dimerization of PHD_{UHRF1} in H3(1-9) peptide bound state. The color coding and labeling is the same as in panel A. The H3 peptide in a ribbon representation is colored in yellow. Water molecules coordinate the Zn-atoms are represented as red spheres.

See also Figure 2.

Figure S2. **Gene Expression Microarray Analysis.**

UHRF1 stable knockdown and control HCT116 cell lines were established by infecting HCT116 cells with four PLKO.1-UHRF1 shRNAs (targeting 4 different regions of UHRF1 (see table S2) or plko.1-control shRNA virus for 16 hours, then selected with puromycin for 6 days until cells were stabilized.

(A) The knockdown efficiency was determined by Western blot with an anti-UHRF1 antibody.

(B) The knockdown efficiency was determined by qRT–PCR.

(C) Microarray analysis of genes affected by UHRF1 knockdown in HCT116. The numbers of the affected gene probes (out of a total of 54613) are shown with indicated fold of down- or up-regulation. UHRF1-sh2 and UHRF1-sh5 treated HCT116 cells were used for the microarray analyses.

See also Figure 4.

Figure S3. GO Term and KEGG Pathway Analysis of Genes Regulated by UHRF1.

GO term and KEGG pathway analysis were performed using genes (affected by both UHRF1sh2 and UHRF1sh5) from the Gene Expression Microarray data (fold≥1.4). The significantly over-represented GO terms ($P<0.01$) (A and B) and KEGG pathways ($P<0.05$) (C and D) in the up- and down-regulated group of genes were determined by GOTERM_BP_FAT and KEGG_PATHWAY, respectively (http://david.abcc.ncifcrf.gov/).

Figure S4. RARβ is methylated at its promoter region and knockdown UHRF1 can decrease the DNA methylation level.

BSP cloning-based sequencing analysis was performed using HCT116 genomic DNA. Black dot, methylated; white dot, unmethylated.

See also Figure 5.

Figure S5. **The Wild-type and Mutant Proteins were Comparably Expressed in the UHRF1 RNAi cells.**

Cells were lysed in Laemmli Sample Buffer with β-mercaptoethanol. Equal amounts of proteins were subjected to electrophoresis on an 8% SDS–PAGE gels, Blots were probed with anti-UHRF1, anti-flag or anti-actin antibodies.

See also Figure 6.

Figure S6. Structural Comparison between PHD_{UHRF1} and WDR5 WD-40 Motif Complexed to N-terminal H3 Peptides Containing Unmodified R2 and K4 Side Chains.

(A) Structure of the PHD_{UHRF1} with unmodified H3 peptide. Arg337 from the N-terminus of canonical PHD finger interacts with the main chain carbonyl oxygens from the pre-PHD. Main chain carbonyl oxygen from Cys329 of the N-terminus of canonical PHD finger interacts with the side chain of unmodified H3K4 residue. Two

Asp residues and the main chain carbonyl oxygen from the β-hairpin interact with unmodified H3R2 residue.

(B) Structure of WDR5 bound to H3(1-10)K4me2 peptide (PDB: 2H13). The unmodified R2 side chain is inserted into the central cavity of the torroidal WD-40 propeller fold, where it is oriented through direct and water-mediated hydrogen bonds and sandwiched between staggered side chains of Phe133 and Phe263.

See also Figure 3.

Supplemental Tables

Table S1. Genes up- and down-regulated in the presence of uhrf1 shRNAs
We have supplied the table as an Excel file.

Table S2. shRNA sequences for UHRF1.

UHRF1-sh1	ccgcaccaaggaatgtaccat
UHRF1-sh2	gcctttgattcgttccttctt
UHRF1-sh4	tgtgaaatactggcccgagaa
UHRF1-sh5	gcgctggctctcaactgcttt
Control RNAi(Scr)	aattctccgaacgtgtcacgt

Table S3. Primers for qPCR

APP.	gene	Forward primer (5'-3')	Backward primer (5'-3')
RT	UHRF1	GCAGAGGCTGTTCTACAGGG	GTGTCGGAGAGCTCGGAGT
RT	GAPDH	TGATGACATCAAGAAGGTGGTGAAG	TCCTTGGAGGCCATGTGGGCCAT
RT	SLC16A6	GGGACTCACATTCCACTGCTTG	ATGGCTCACTACCTTTGTTTCACC
RT	ACSL5	GAGGAACTGTGCCAAAACCAAG	GGATGTGCTCATACAGGCTGTCA
RT	AQP3	CAGTTCATAGGCACAGCCTCCC	GGTACACGAAGACACCCGCAAT
RT	DHRS2	CTGAGCCAGTTGCTGCCCTACA	CCTGGAACCACGCAGTTTACCC
RT	H19	CTCAGGAATCGGCTCTGGAAGG	GATGATGTGGTGGCTGGTGGTC
RT	SAMD9	CTAGCATCGTAGGAGGAAGAAA	ACCACTTCAAGACTGCTCCATT
RT	TIAM1	AAAACCGAGAGCCTTCCCTCAT	CACGGACTCACAGAACTCATCA
RT	ENC1	ATTCTCTGCCTGCTCTGCTTAT	GCAACTTCTTTTCCTTTTTTTC
RT	KRT23	CCCACTATGAGGAGGAACTGACG	TCGACTTTGATTCTTCCCGTGT
RT	VSNL1	TTCTTTCCTTATGGAGACGCCTCCA	GCTCATTTCTGGATGTCGCACT
RT	ISG15	GCAGCGAACTCATCTTTGCCAGTA	TCAGAGGTTCGTCGCATTTGTCC
RT	SUSD2	TCCTGGAGGAGCTGCCGGAC	TGGGTCTCCGAAGGCGGAGG
RT	ADAM19	GCATCGTTTCCCAGGACTTCTC	CCAGCCCTCTGTGATCTGTATTCT
RT	KLRC1	AAGATCCACACTGGGCTGATTT	ACTTGGGAAGAGAGTTTGCTGG
RT	MALAT1	CAGACCCTTCACCCCTCACCTC	CAAAAGCCCTCTCAGCCACTCA
RT	NXT2	CGGGCACTAACCAGGCTGTATC	ATGGACCAAATGAGAATGGACTTT
RT	ABCC2	CATCCCTCACAAACTGCCTCTT	GATTCATCCTCAGGCTTCCAGA
RT	SOX4	ACACTGGTGGCAGGTTAAGG	GGCGGCTGGTTAATATCTCA
RT	DDX58	CTGATTGCCACCTCAGTTGCTG	TGCTCTTCCTCTGCCTCTGGTT
RT	DDX60	TCTAGGCACAATCGGTGTCAAT	CTGGCAAGAAGCATAAGAATCAA
RT	HERC5	AATGCACGTTATGAACCAGGATA	TTAGGGAGGAAGAGGACACTGAA
RT	HIST1H2AC	CCGCTGGTTTTGGTGATTTTTGT	CGGGATGATGCGAGTCTTCTTGT
RT	MSI2	GCAAGAGGATCAGGCTCCAACC	CCACTCAGAAACTTCGCCCAGTC
RT	OASL	GGCAGAAGGGTACAGATGGGACA	CAGGAACCTGGAAGGACAGACG
RT	SMARCA1	GAGCAGAAAAGAAGAAACGGGC	ACCTTCTTGACATCCTTCTTTCC
RT	Sp110	ACTAAATGTGCCCGAAAGTCCA	GTTTCCAGTTCTTTGCGTTCCT
RT	ADAM19promoter	CATCAGGCTCTGGAGTTATACGG	CGCTAAGTGTCAAACGCTGTCT
RT	SUSD2promoter	GGCTGCCAACACCCTGTGGTTC	GGCCATCCGGTCTGTGATCTGC
RT	SAT2	ATCGAATGGAAATGAAAGGAGTCA	GACCATTGGATGATTGCAGTCA
RT	RPL30	CAAGGCAAAGCGAAATTGGT	GCCCGTTCAGTCTCTTCGATT
RT	RAR-β	TTCCAAGCTAAGCCGCCGCA	CCCAGCAGCCCTGCAAAAGC

Supplemental Experimental Procedures

Protein Expression and Purification

The cDNA encoding full-length UHRF1 was obtained from Open Biosystems. We generated a cDNA of PHD finger (311-380) in GST-tagged expression vector, as well as cDNAs of TUDOR (140-295) and SRA (427-630) domains in hexahistidine-Sumo tagged vectors. Plasmids were transformed into *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. The culture was then induced with 0.4 mM of isopropyl-1-thio-D-galactopyranoside, and allowed to grow for 15 hrs, following which they were harvested and resuspended in the lysis buffer (25 mM Tris-HCl, pH 7.5, 500 mM NaCl, 3 mM β-mercaptoethanol). Cells were lysed following three passes through an ice-cold French pressure cell press and then the lysate was cleared by centrifugation at 40,000g for 1 hr. The hexahistidine-Sumo fusion proteins were purified on a nickel-charged column (HisTrap HP, GE healthcare). After elution with a 750 mM imidazole-containing buffer, fusion proteins were cleaved with Ulp1 protease at 25 U /mL during a 16 hr dialysis step at 4 ºC. The GST tagged PHD_{UHRF1} was purified on a glutathione-sepharose column (GSTrap HP, GE healthcare). After elution with a 20 mM L-glutathione containing buffer, the fusion protein was cleaved with PreScission protease at 15 U/ml during a 16 hr dialysis step at 4 ºC. Proteins were further purified by gel filtration chromatography (HiLoad Superdex 75 26/60) and equilibrated with 15 mM Tris-HCl, pH 7.5, 50 mM NaCl, 3 mM DTT. Purified

PHD finger, TUDOR and SRA domains of UHFR1 were concentrated to 25 mg/mL, 10 mg/mL and 8 mg/mL at 4 ºC in Vivaspin 20 mL (Vivascience AG) 5,000 cut-off concentrator.

ITC Measurements

The equilibrium dissociation constants for PHD finger, TUDOR and SRA domains of UHRF1 bound to H3 peptides were determined using a VP-ITC calorimeter (MicroCal, LLC) at 25 °C and the data were processed with MicroCal Origin software. The proteins were dialyzed overnight against a buffer containing 40 mM Tris-HCl, 50 mM NaCl, and 2 mM β-mercaptoethanol, pH 7.5 at 4°C. Peptides dissolved in water were lyophilized over night. Calorimetric titration was performed by injection of synthetic peptides into a cell containing PHD, TUDOR or SRA domains of UHRF1. The protein and different peptides concentrations used were 100 µM to 150 µM and 1mM to 1.5 mM, respectively.

Crystallization

All crystals 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. Both PHD_{UHRF1}-H3(1-9) and PHD_{UHRF1}-H3(1-9)K4me3 complex crystals were grown from a condition containing 30% w/v polyethylene glycol 8,000, 0.2 M sodium acetate trihydrate and 0.1 M sodium cacodylate trihydrate, pH 6.5. Both complexes

are crystallized in tetragonal ($P4_32_12$) space group using 20 mg mL^{-1} of protein in a protein:peptide molar ratio of 1:1.1. Free-form crystals were grown using 15 mg mL^{-1} of protein under the condition contained 2 M ammonium sulfate and 0.1 M Tris, pH 8.5. All crystals were flash frozen at 100 K in a cryoprotectant containing a total of 37% ethylene glycol.

Data Collection, Structure Determination and Refinement

The X-ray diffraction data set for the $P4_32_12$ crystals of PHD_{UHRF1} bound to H3(1-9) peptide were collected on beamline X29 at the Brookhaven National Laboratory (BNL) synchrotron at Zn-peak wavelength and the structure was solved by the Single-Wavelength Anomalous Dispersion (SAD) technique using the PHENIX program (Zwart et al., 2008). The PHENIX (Adams et al., 2002) program was used for automated model building and density modification. Two H3 peptides bound to each molecule in the asymmetric unit were built into the density maps using the COOT program (Emsley et al., 2004) and the structure of the complex was refined against 1.80 Å diffraction data with simulated annealing and restraint minimization methods using the PHENIX program. The final R-factor and R-free values of the model are 21.5% and 24.3%, respectively, for data from 20.0 Å to 1.80 Å. Missing residues include a residue at N-terminus of the PHD finger and one peptide residue at C-terminus.

The data set of PHD_{UHRF1}-H3(1-9)K4me3 complex crystals were collected on 24-ID-C beam line at the Advanced Photo Source (APS), Argonne National Laboratory. The Molrep-Auto MR program (Vagin et al., 2010) was used for solving the structure of the PHD finger, using the PHD finger coordinates from the PHD_{UHRF1}-H3(1-9) complex structure. The asymmetric unit contains two PHD finger molecules, each bound by a peptide. The PHENIX program was used for simulated annealing, NCS averaging, restraints and occupancy refinement against 1.95 Å data. The structure of the complex was refined to an R-factor of 21.3% and R-free of 24.7%.

The 2.65 Å diffraction data set of the free form of PHD_{UHRF1} crystals was collected using in-house copper anode X-ray generator. The structure was solved by molecular replacement using the PHD finger model from PHD_{UHRF1}-H3(1-9) peptide complex structure. Simulated annealing, restraint minimization and NCS averaging were performed using the PHENIX program. The final model had an R-factor of 24.8% and R-free of 29.4% for data from 20.0 Å to 2.65 Å resolution.

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

Plasmid Construction and Antibodies

UHRF1 Point mutation and deletion constructs were generated by site-directed mutagenic PCR and cloned into vector plenti6.2 with an N-terminal Flag tag (plenti6.2-UHRF1, plenti6.2-UHRF1ΔPHD, plenti6.2-UHRF1-D347A/E348A), and

into vector DsRed-c1 (DsRed-c1-UHRF1, DsRed-c1-UHRF1ΔPHD, DsRed-c1-UHRF1-D347A/E348A). The following antibodies were obtained commercially: monoclonal antibody against 5-methylcytidine (5mC) (Eurogentec #BI-MECY-1000), anti-UHRF1 (santa cruz biotechnology sc-100606), anti-flag (Sigma F7425), anti-actin Ab (abcam ab34731). UHRF1 antibody for Chip was raised in house.

Cell Culture

HCT116 cells were obtained from American Type Culture Collection (ATCC) and maintained in McCoy's 5A medium containing 10% fetal bovine serum (FBS) (Hyclone) and 1% Pen-Strep. NIH-3T3 cells and 293T cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) (Hyclone) and 1% Pen-Strep.

Western Blotting

Cells were lysed in Laemmli Sample Buffer withβ-mercaptoethanol (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS, 0.01% Bromophenol Blue, β-mercaptoethanol, 710 mM). Equal amounts of proteins were subjected to electrophoresis on an 8% SDS–PAGE gels, and transferred to PVDF membranes (Bio-Rad). Blots were probed with anti-UHRF1 (santa cruz biotechnology sc-100606), anti-flag (Sigma F7425) or anti-actin Ab (abcam ab34731).

Quantification of UHRF1-Regulated Gene Expression and ChIP-qPCR Analyses

Quantification of endogenous genes expression regulated by UHRF1. Total cellular RNAs from 5 x 10^5 HCT116 cells treated with either UHRF1 or control shRNA were extracted using the TRIzol reagent (Invitrogen). RNA (1μg) of each sample was then reverse transcribed and used for real time quantitative PCR. ChIP-qPCR Analyses were performed to detect the presence of UHRF1 and H3K9me3 at ADAM19 and SUSD2 promoter regions. Conventional ChIP was performed as previously described (Shi et al., 2003). Primers for real-time PCR are listed in table S3.

Bisulfite sequencing

After treatment with either lentiviral UHRF1 or control shRNAs for 14 days, genomic DNA was purified from HCT116 cells. 50ng genomic DNAs were subjected to treatment with the EZ DNA Methylation Kit™ (Zymo Research, D5020) according to the manufacturer's instructions. PCR products were cloned into the pMD-18-T vector, followed by sequencing analysis. PCR primers were designed according to Widschwendter et al., 2000.

Supplemental References

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*, 1948-1954.

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

Shi, YJ., Sawada, J-I., Sui, GC., El Bachir, A., Whetstine, JR., Lan, F., Ogawa, H., Luke, M. P-S., Nakatani, Y., and Shi, Y. (2003). Coordicated histone modifications mediated by a CtBP co-repressor complex. Nature. *442*, 735-738.

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

Zwart, P.H., Afonine, P.V., Grosse-Kunstleve, R.W., Hung, L.W., Loerger, T.R., McCoy, A.J., McKee, E., Moriarty, N.W., Read, R.J., Sacchettini, J.C., et al. (2008). Automated structure solution with the PHENIX suite. Methods Mol. Biol. *426,* 419-435.