Electronic Supplementary Material (ESI) for Chemical Communications

Structural insights into how 5-hydroxymethylation influences transcription factor binding

Lukas Lercher,†[a] Michael A. McDonough,†[a] Afaf H. El-Sagheer,[bc] Armin Thalhammer,[a] Skirmantas Kriaucionis,[d] Tom Brown,*[b] and Christopher J. Schofield*[a]

[a] Department of Chemistry and the Oxford Centre for Integrative Systems Biology, Chemistry Research Laboratory, Oxford, UK.

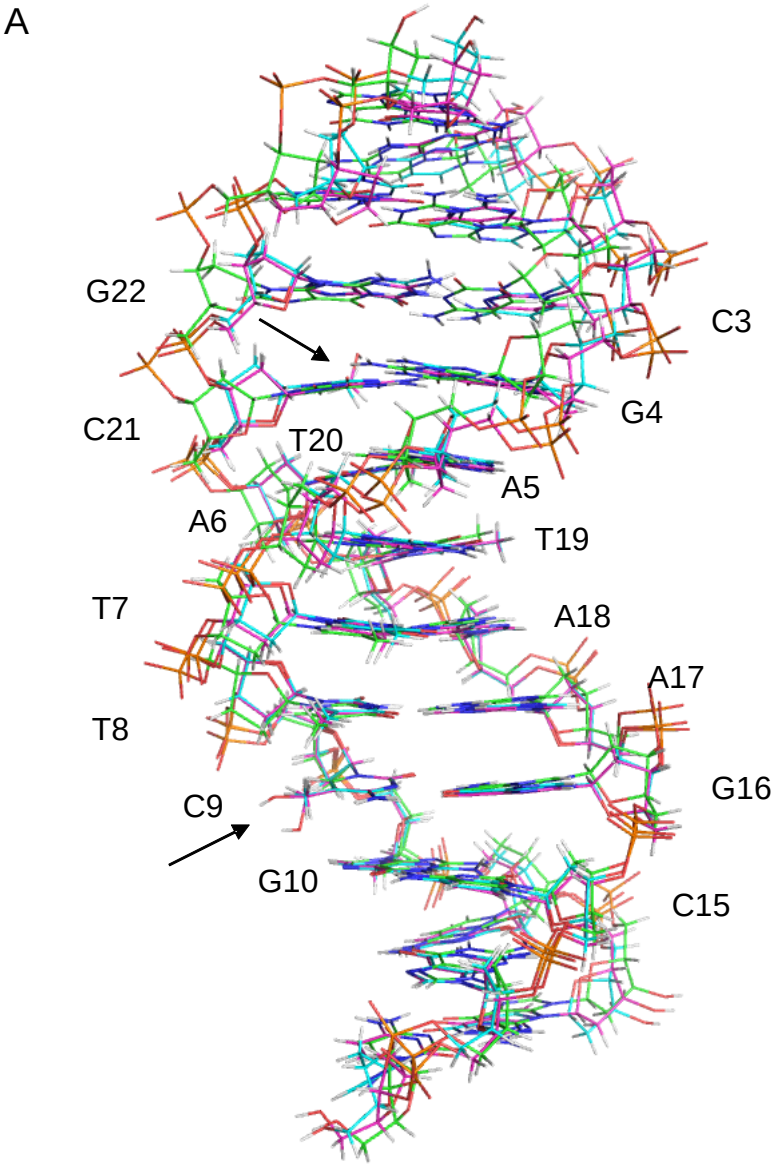
[b] Department of Chemistry, Chemistry Research Laboratory, Oxford, OX1 3TA, UK

[c] Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez, Egypt

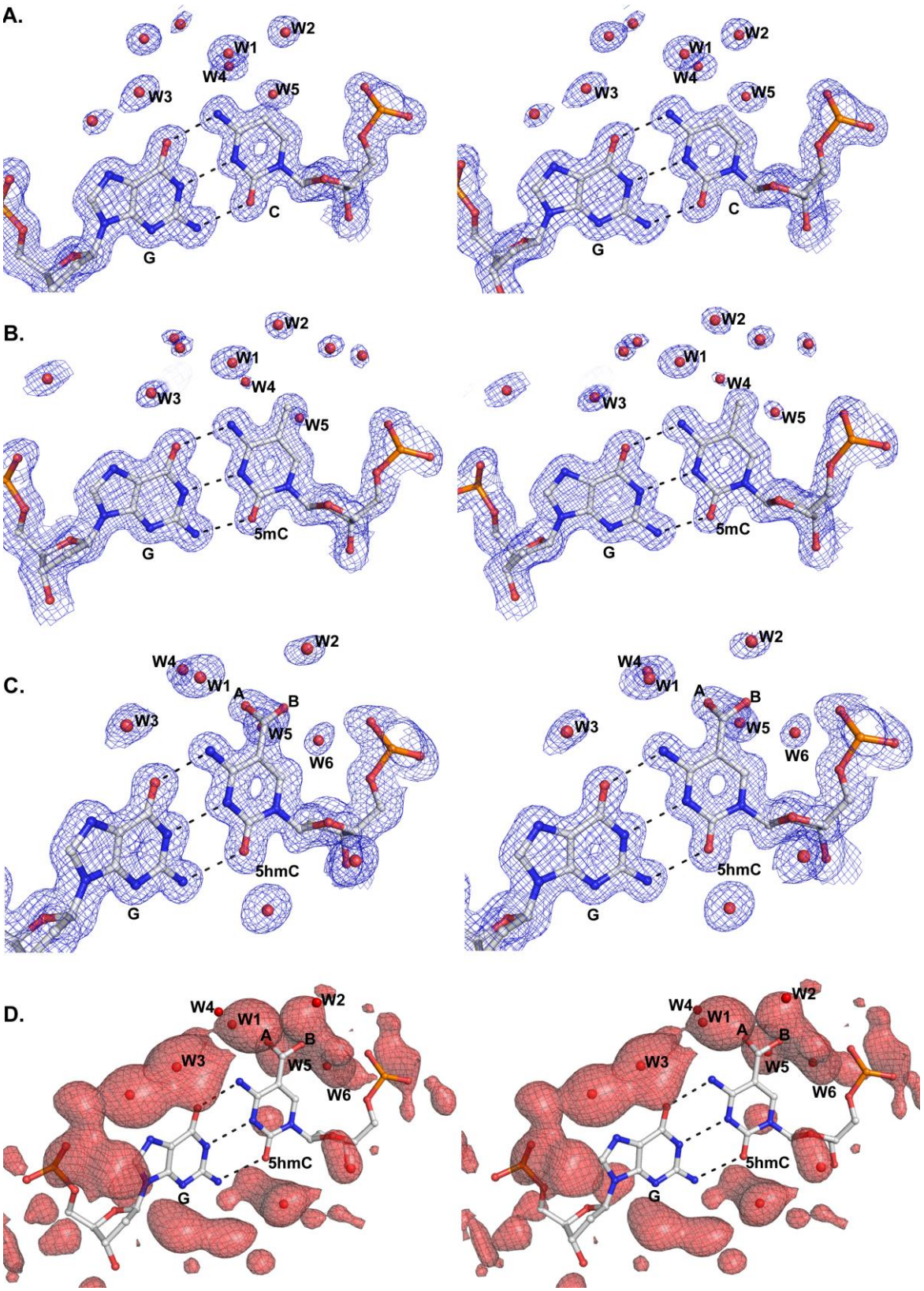
[d] Laboratory of Epigenetic Mechanisms, Ludwig Institute for Cancer Research, The University of Oxford, ORCRB, Oxford OX37DQ, UK

† These authors contributed equally to this work.

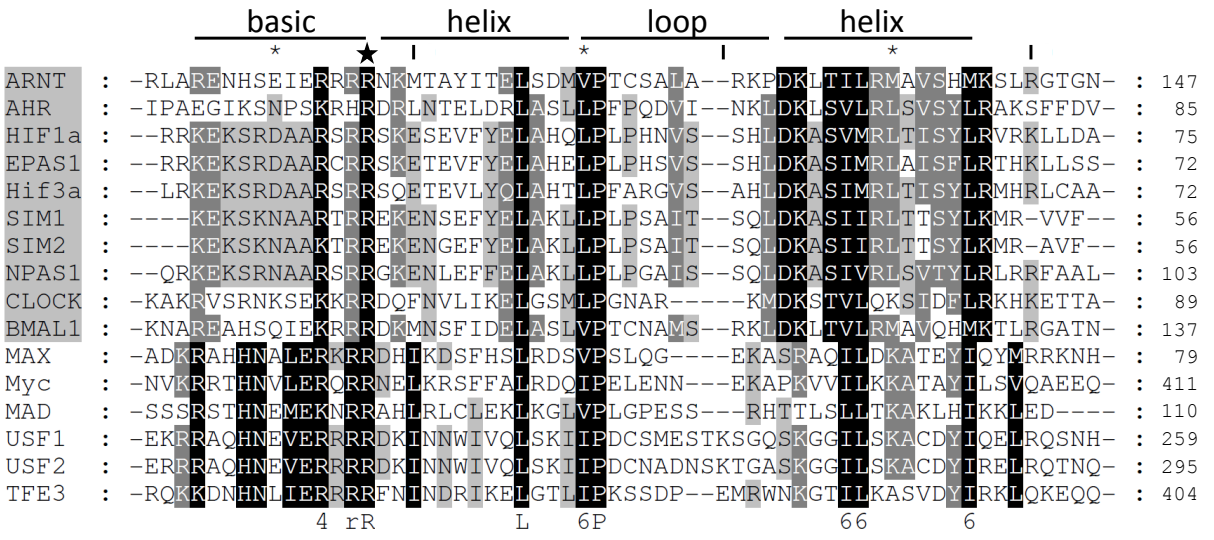
* E-mail: Christopher.schofield@chem.ox.ac.uk (C.J.S) or tb2@soton.ac.uk (T.B.)



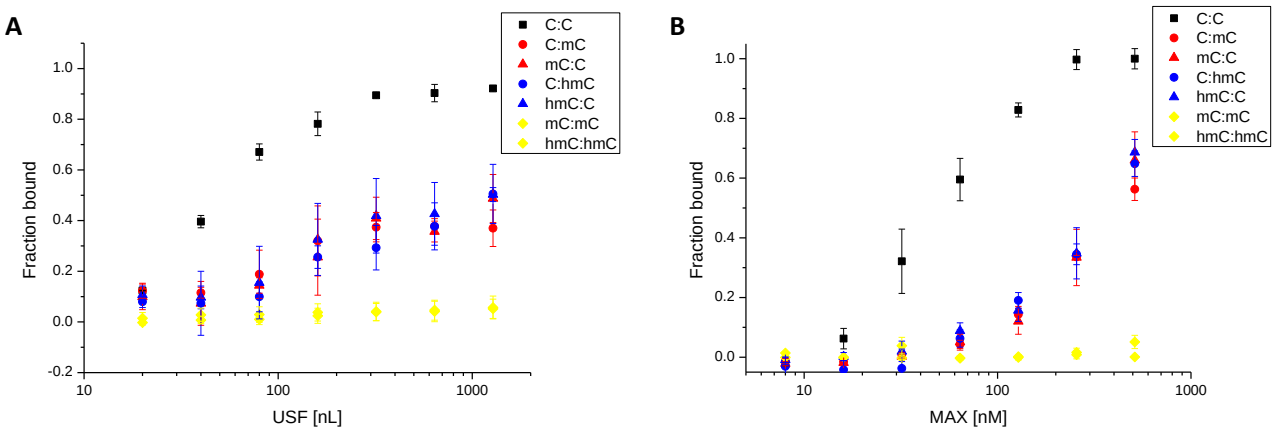
Supplementary Figure S1 (A) Overlay of the cytosine (green; 4C64), 5-methylcytosine (cyan; 4C63) and 5-hydroxymethylcytosine (pink; 4C5X) containing dsDNA. The sites of modification are marked with arrows. (B) Sequence of the used duplex with numbering. Modified sites are underlined.



Supplementary Figure S2 Stereoviews of the water structure around the C9-G basepairs of (A) unmodified C9 (4C64), (B) mC9(4C63), (C) hmC9(4C5X), and (D) SWS[1] plot of unmodified cytosine.

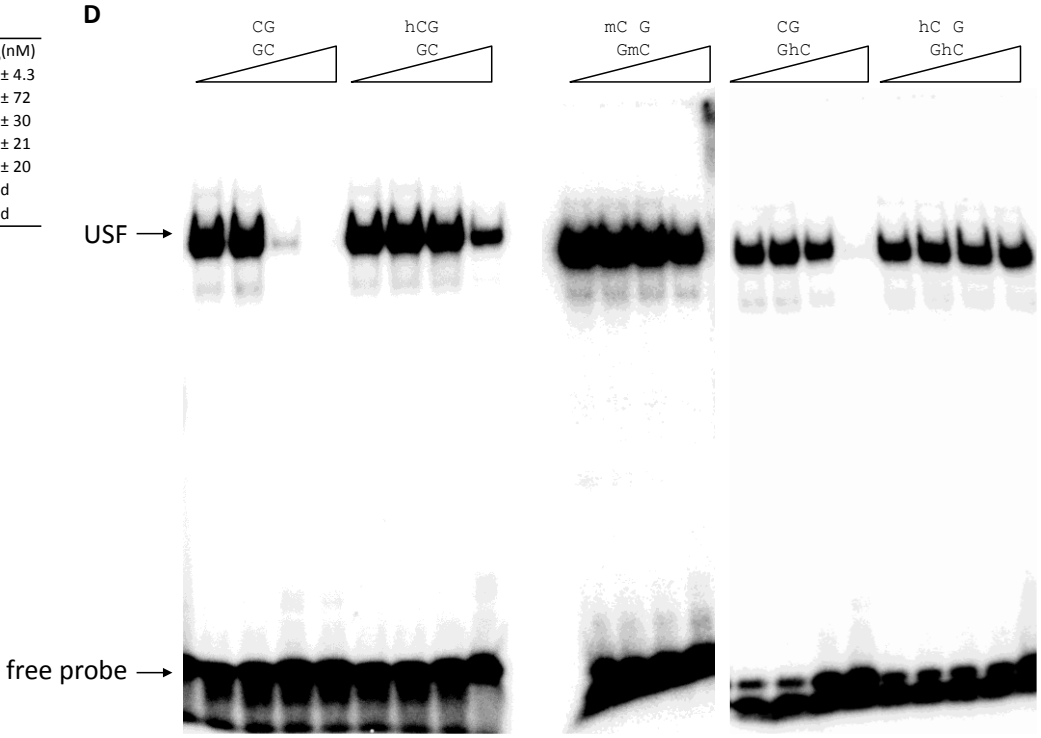


Supplementary Figure S3 Sequence alignment of the bHLH region of different bHLH-PAS (highlighted in grey) and bHLH-ZIP proteins. The conserved Arg that can form a hydrogen bond with the central G (Arg35 in MAX[2], Arg211 in USF[3], Arg47 in CLOCK[4], Arg85 in BMAL1[5]) at the CpG is highlighted by the star. Sequences were retrieved from UniProt (www.uniprot.org). ARNT, Aryl hydrocarbon receptor nuclear translocator, P27540; AHR, Aryl hydrocarbon receptor; P35869; HIF1a, Hypoxia-inducible factor 1-alpha , Q16665; EPAS1, Endothelial PAS domain-containing protein 1, Q99814 ; Hif3a, Hypoxia-inducible factor 3-alpha, Q9Y2N7; SIM1, Single-minded homolog 1, P81133; SIM2, Single-minded homolog 2, Q14190; NPAS1, Neuronal PAS domain-containing protein 1, Q99742; CLOCK, Circadian locomoter output cycles protein kaput, O08785 ; BMAL1, Aryl hydrocarbon receptor nuclear translocator-like protein 1, Q9WTL8; MAX, Myc-associated factor X, P61244; Myc, Myc proto-oncogene protein, P01106 ; MAD, Max dimerization protein 1, Q05195; USF1, Upstream stimulatory factor 1, P22415; USF2, Upstream stimulatory factor 2, Q15853; TFE3, Transcription factor E3, P19532.

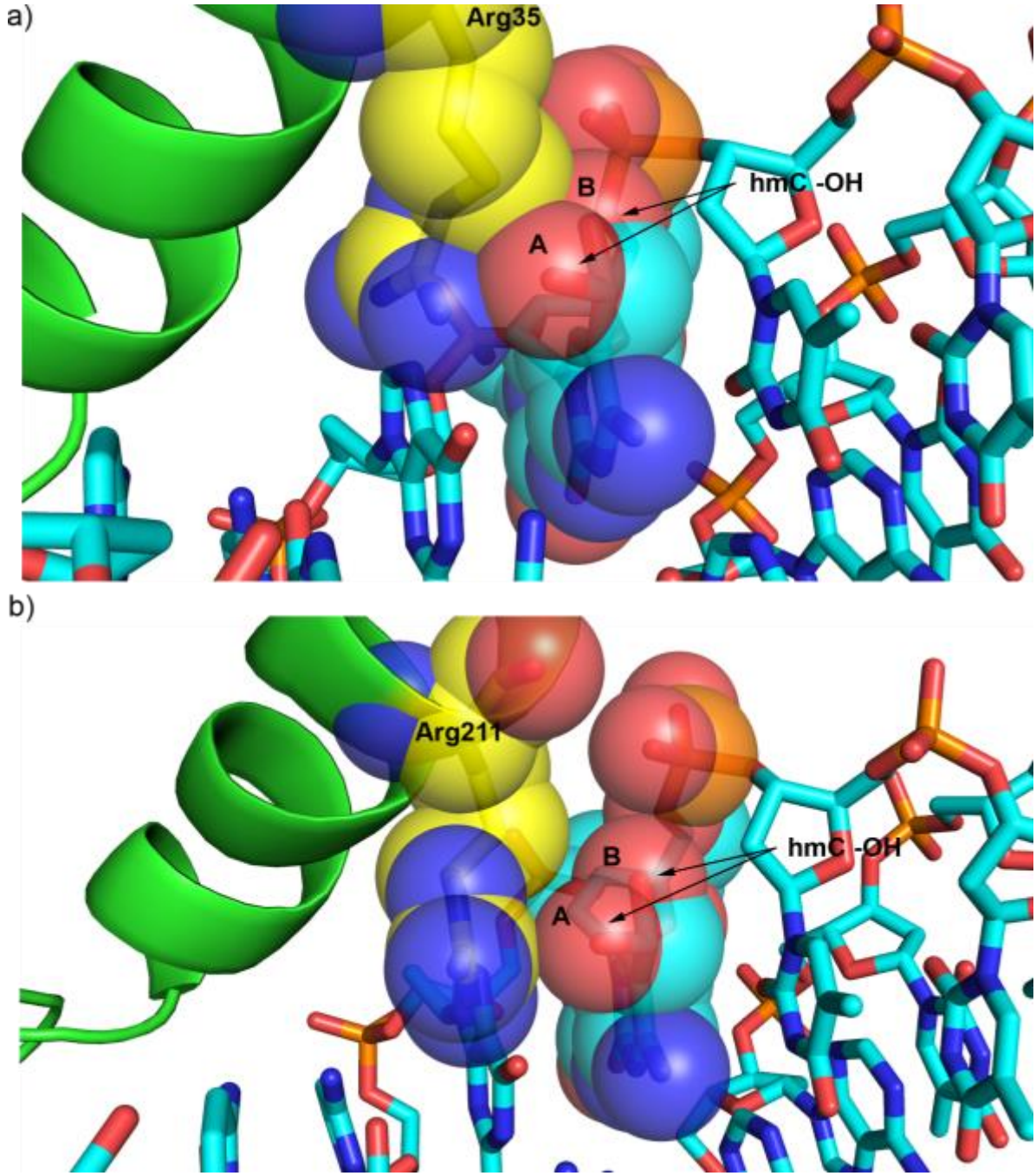


C

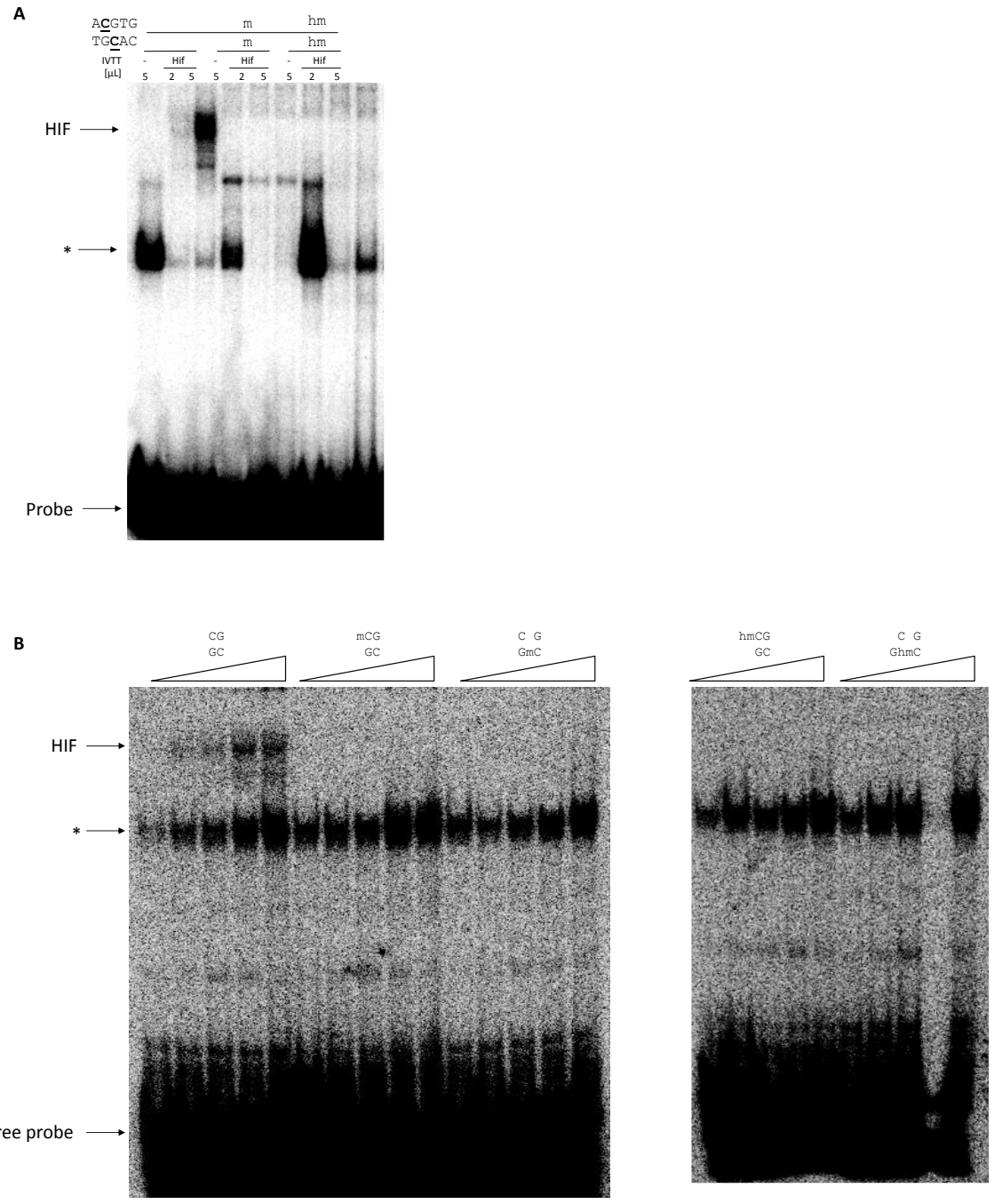
	Kd_{app}(nM)
C:C	51.3 ± 4.3
C:mC	342 ± 72
mC:C	366 ± 30
C:hmC	338 ± 21
hmC:C	325 ± 20
mC:mC	nd
hmC:hmC	nd



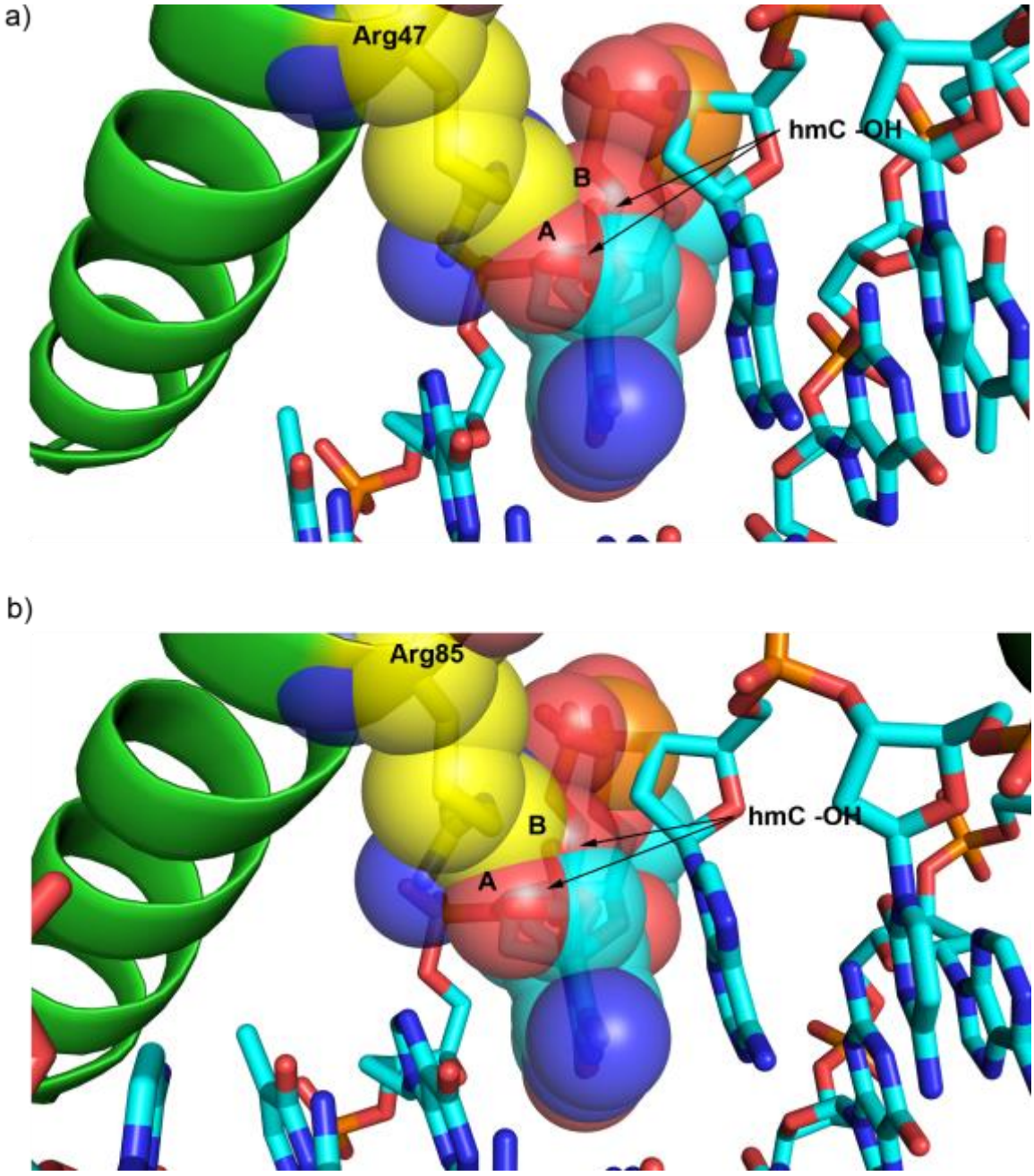
Supplementary Figure S4 Titration curves of MAX (A) and USF (B) measured by electrophoretic mobility shift assays (EMSAs) with radiolabeled dsDNA probes bearing different cytosine C-5 modifications at the central CpG of the E-Box sequence. The EMSA modification state of the cytosine in the central CpG of each series is given in the legend. Experiments were performed in triplicate. For detailed experimental procedure and used sequences see Supplemental Methods. (C) The affinity of MAX for the E-Box sequences containing different cytosine modifications as determined by EMSAs[6]. In the case of mC:mC and hmC:hmC only weak / non-specific binding was observed as evidenced by a slight reduction of free probe but no formation of a distinct complex. (nd: not determined) (D) Competition of a radiolabeled unmodified E-Box-USF complex with unlabeled probes containing different modifications. The central CpG modification state is given above the lanes.



Supplementary Figure S5 Structural overlay of the hmC9 containing duplex (4C5X) with the MAX-DNA complex (A; adapted from pdb entry 1AN2[7]) and the USF-DNA complex (B; adapted from pdb entry 1AN4[3]) showing the potential clash of the C5 modification with the arginine (Arg35 in MAX, Arg211 in USF). Both the mayor (A; 70% occupancy) and the minor (B; 30% occupancy) of the hmC9 alcohol are depicted.



Supplementary Figure S6 Modification of the central CpG in the hypoxic response element (HRE; ACGTG) abolishes binding to hypoxia-inducible factor (HIF). Full-length HIF1α and HIF1β (ARNT) were first produced by *in vitro* transcription translation (IVTT) and then incubated with radiolabeled DNA probes containing a HRE (see experimental section for sequence). The obtained complexes were separated on a 5% PAGE gel. (A) EMSA with IVTT HIF and different unmodified, fully methylated or fully hydroxymethylated probes derived from the erythropoietin **(**EPO) promoter[6, 8]. Different modifications were introduced by substitution of dCTP with dmCTP or dhmCTP during the PCR amplification[9]. (B) Titrations of increasing amounts of IVTT translated HIF1-α/β with differently modified probes**.**



Supplemental Figure S7. Structural overlay of the hmC9 containing duplex (4C5X) with the CLOCK-DNA complex (A) and the BMAL1-DNA complex (B; both adapted from pdb entry 4H10[5]) showing the potential clash of the C5 modification with the arginine (Arg47 in CLOCK, Arg85 in BMAL1). Both the mayor (A; 70% occupancy) and the minor (B; 30% occupancy) of the hmC9 alcohol are depicted.

Structure	Initial Crystallization drop conditions (12 μL or 10 μL total initial volume)	Crystallization well conditions
Unmodified (4C64), 5mC (4C63) and 5hmC (4C5X)	0.27 mM DNA sample 8.3 % methylpentanediol (MPD) 5.0 mM Spermine-HCL 8.3 mM Na Cacodylate pH 7.5 33.3 mM $MgCl_2$	40% MPD

Supplementary Table S1 Crystallization conditions

	Unmodified	5mC	5hmC
PDB ID	4C64	4C63	4C5X
Data Collection and Processing			
X-ray source	Diamond I04-1	Diamond I04-1	Diamond I04-1
Wavelength (Å)	0.917300	0.917300	0.917300
Resolution (Å) (outer shell)	25.32-1.30 (1.37-1.30)	25.34-1.30 (1.37-1.30)	34.2-1.30 (1.37-1.30)
Space Group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit Cell Dimensions (Å)	25.15 39.84 65.60	24.45 40.06 65.41	25.30 40.27 64.88
Total # of reflections	103,920 (11,809)	202,083 (27,411)	89,233 (13,443)
# unique reflections	16,471 (2,312)	16,355 (2,215)	16,569 (2,325)
Redundancy	6.3 (5.1)	12.4 (12.4)	5.4 (5.8)
Completeness (%)	98.0 (95.9)	98.4 (93.9)	98.0 (96.5)
I/σI	32.1 (6.1)	18.0 (2.3)	26.2 (4.6)
Rp.i.m. (%)	1.0 (10.9)	1.8 (35.7)	1.3 (16.5)
Rmerge (%)	2.8 (22.3)	5.8 (123.0)*	2.6 (36.9)
Refinement			
resolution	1.320	1.320	1.30
Rcryst(%)	0.1385	0.1497	0.1421
Wilson B ($Å^2$)	16.8	19.2	15.6
Rfree(%)	0.1823	0.1848	0.1734
Deviations from ideal			
Bonds (Å)	0.017	0.021	0.022
Angles (˚)	2.1	2.4	2.5
Average B factors ($Å^2$)	23.8	24.9	22.1
# waters	131	94	136

Supplementary Table S2. Crystallographic data collection, processing and structure refinement statistics.

*Note: The high redundancy data for the 5mC structure contributes to the exaggerated high Rmeas in the high resolution bin – the I/σI value in this case is a more reasonable measure of the data quality.

C		Minor Groove		Major Groove	
		P-P	Refined	P-P	Refined
1	cg/cg	---	---	---	---
2	gc/gc	---	---	---	---
3	**cg/cg**	**14.3**	**---**	**17.9**	**---**
4	ga/tc	11.9	11.8	17.5	17.4
5	aa/tt	10	9.9	17.7	17.4
6	at/at	9.5	9.5	16.7	16
7	tt/aa	8.9	8.9	17.8	17.8
8	tc/ga	9.4	9.4	18.4	18.3
9	**cg/cg**	**10.5**	**---**	**17.7**	**---**
10	gc/gc	---	---	---	---
11	cg/cg	---	---	---	---

mC		Minor Groove		Major Groove	
		P-P	Refined	P-P	Refined
1	cg/cg	---	---	---	---
2	gc/gc	---	---	---	---
3	**cg/cg**	**13.9**	**---**	**18**	**---**
4	ga/tc	11.8	11.7	17.9	17.8
5	aa/tt	9.9	9.9	17.2	17
6	at/at	9.3	9.3	16.4	15.8
7	tt/aa	8.8	8.8	17.5	17.5
8	tc/ga	9.3	9.3	18	18
9	**cg/cg**	**10.5**	**---**	**16.8**	**---**
10	gc/gc	---	---	---	---
11	cg/cg	---	---	---	---

hmC		Minor Groove		Major Groove	
		P-P	Refined	P-P	Refined
1	cg/cg	---	---	---	---
2	gc/gc	---	---	---	---
3	**cg/cg**	**13.9**	**---**	**18.7**	**---**
4	ga/tc	11.7	11.7	18.7	18.6
5	aa/tt	9.6	9.5	18.4	18.2
6	at/at	9.2	9.2	15.6	15.2
7	tt/aa	9.1	9.1	16.9	16.8
8	tc/ga	9.7	9.6	18.5	18.5
9	**cg/cg**	**10.9**	**---**	**18.8**	**---**
10	gc/gc	---	---	---	---
11	cg/cg	---	---	---	---

Supplementary Table S3 Groove widths calculated with 3DNA[10] for different duplexes. hmC containing base pairs are highlighted.

Strand I	C				mC				hmC			
base	tm	P	Puck		tm	P	Puck		tm	P	Puck	
C-1	-5.2	38.6	169.2	C2'-endo	2.3	39	158.4	C2'-endo	0.8	38.8	160.2	C2'-endo
G-2	1.5	34	158.3	C2'-endo	-1.2	34.2	162.7	C2'-endo	6.8	34	149.8	C2'-endo
C-3	33.6	35.5	53.2	C4'-exo	36.6	38.1	52.1	C4'-exo	30.9	33.1	94.2	O4'-endo
G-4	-2.1	35.7	164.7	C2'-endo	-6.4	34	171.9	C2'-endo	9.7	46.4	149.1	C2'-endo
A-5	6.4	32.5	150.3	C2'-endo	1.1	34.9	159.3	C2'-endo	9.8	36.8	146.5	C2'-endo
A-6	18	35.3	131.9	C1'-exo	17.3	34	132.5	C1'-exo	8.6	31.1	145.9	C2'-endo
T-7	26.4	38.7	118.8	C1'-exo	25.8	38.4	120.4	C1'-exo	29.7	40.6	116.6	C1'-exo
T-8	25.9	35.6	115	C1'-exo	19.1	33.1	126.3	C1'-exo	26.9	39.9	120.4	C1'-exo
C-9	**9.2**	**33.9**	**146.4**	**C2'-endo**	**12.3**	**36.5**	**142.9**	**C1'-exo**	**17.4**	**33.5**	**131.1**	**C1'-exo**
G-10	15.7	46.9	142.2	C1'-exo	16.6	52.8	143	C1'-exo	15.5	43.3	140.7	C1'-exo
C-11	-2.8	36.1	166.3	C2'-endo	-1.6	37.4	164.2	C2'-endo	4.5	35.2	154.6	C2'-endo
G-12	34.9	37.8	96.5	O4'-endo	42.8	44.7	89.7	O4'-endo	28.6	35.9	110.7	C1'-exo

Strand II	C				mC				hmC			
base	tm	P	Puck		tm	P	Puck		tm	P	Puck	
C-13	20.8	40.8	11.1	C3'-endo	27.4	42	21.7	C3'-endo	19	38.5	9.7	C3'-endo
G-14	23.1	41.4	14.5	C3'-endo	12.3	33	138.1	C1'-exo	24.2	41.7	16.7	C3'-endo
C-15	12.1	44.8	146.2	C2'-endo	10.9	48	149.3	C2'-endo	13.7	43.2	143.5	C1'-exo
G-16	30	40.2	113.9	C1'-exo	29.5	39.3	111.9	C1'-exo	37.8	41	92.6	O4'-endo
A-17	23.3	42.7	128.4	C1'-exo	15.4	38.4	138.4	C1'-exo	14.6	38.5	139.9	C1'-exo
A-18	26.5	37.8	116.8	C1'-exo	27	39.2	118.2	C1'-exo	27.7	39.2	117.4	C1'-exo
T-19	18.8	32.6	126.6	C1'-exo	19.6	35.1	128.3	C1'-exo	24	37.1	121.6	C1'-exo
T-20	-6.3	33.7	172.3	C2'-endo	-4.6	31.3	170.1	C2'-endo	-8	32.5	175.8	C2'-endo
C-21	**-8.9**	**32.5**	**176.6**	**C2'-endo**	**-5**	**29.7**	**170.4**	**C2'-endo**	**-11.8**	**36.3**	**179.4**	**C2'-endo**
G-22	32.2	38.3	37.2	C4'-exo	31.2	35.8	41	C4'-exo	31.4	41.5	28.6	C3'-endo
C-23	17.6	31.1	126.8	C1'-exo	14.4	34.9	136.9	C1'-exo	26.2	38.8	118	C1'-exo
G-24	12.5	40.6	145	C2'-endo	6.9	42.4	152.3	C2'-endo	7.1	28.5	147.8	C2'-endo

Supplementary Table S4 Sugar pucker values and conformations for different duplexes calculated with 3DNA[10]. The differently modified base is highlighted.

Additional background information

Crystallographic investigations of B-DNA reveal high similarity between C and 5mC duplexes, with methylation causing slight minor groove compaction[11]. Structures of 5mC in Z-[12, 13], A-[14, 15] and E-[16] DNA forms (in CpG contexts) and of 5mC-binding proteins (MBD[17], SRA[18], zinc finger[19]) complexed with 5mC DNA are also reported. A MBD4 structure in complex with 5hmC containing DNA is reported[20], but the relatively low resolution (2.4 Å) and presence of protein, does not enable analysis of the effect of 5hmC on isolated dsDNA structure.

Supplementary Methods

Oligonucleotide synthesis, purification and analysis

Standard DNA phosphoramidites, solid supports and additional reagents were purchased from Link Technologies and Applied Biosystems Ltd. Oligonucleotides were synthesized using an Applied Biosystems 394 automated DNA/ RNA synthesizer using a standard 1.0 μmole phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by an automated trityl cation conductivity monitoring facility and in all cases were >98.0%. β-Cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 35 s, and the coupling time for the 5-methyl-2´-deoxycytidine monomer was 60 s and for 5-hydroxmethyl-2´-deoxycytidine monomer (5hmC) it was extended to 360 s. Cleavage of the oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution (60 min. room temp.) followed by heating in a sealed tube (72 h, 65 °C). The fully deprotected oligonucleotides were purified by reversed-phase HPLC on a Gilson system using a Luna 10 μL C8 100Å pore Phenomenex 10x250 mm column with a gradient of acetonitrile in ammonium acetate (0% to 50% buffer B over 20 min, flow rate 4 mL/min), (buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate, pH 7.0, with 50% acetonitrile). Elution was monitored by UV absorption at 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-10 columns (GE Healthcare). The purified oligonucleotides were characterized by mass spectra on a Bruker micrOTOF^{TM} II focus ESI˙TOF MS instrument in ES^- mode.

Crystallography

Crystals were grown at 20˚C by the hanging drop vapour diffusion method in 24 well Linbro plates using 22mm round cover slips and sealed using vacuum grease using the conditions in Supplementary Table S1. Crystals appeared within a few days and were harvested after 1 week. The MPD in the equilibrated drops was sufficient to serve as cryo-protectant and crystals were harvested directly from the crystal growth drop using a nylon loop and plunged into liquid nitrogen to cyro-cool. Crystals were stored under liquid nitrogen until they were mounted on a goniometer at the synchrotron beamline under a nitrogen gas stream at 100K. Complete data sets were collected for single crystals of each modification type and independently indexed, integrated and scaled using Xia2[21] or XDS/SCALA[22] as indicated. The structures were solved by molecular replacement using PHASER[23] and PDB entry 1BNA[24] and refined using PHENIX[25].

Protein methods

MAX and USF were cloned from isolated cDNA into pGEM-T Easy vectors (Promega) using the following primers MAX_FW: *ATGAGCGATAACGATGACATCGAGGTGG*, MAX_RW: *TTAGCTGGCCTCCATCCGGAGCTTC*, USF_FW: *ATGAAGGGGCAGCAGAAAACAGCTG*, USF_RW: *GGCCCAAAGCCCCTGAATCCCCA*. MAX was subsequently subcloned into pET15b vector. MAX was expressed in BL21(DE3) Rosetta 2 and purified by IMAC and gel filtration. Plasmids encoding for hypoxic inducible factor (HIF)1-α and HIF1-β were a kind gift from Prof. Christoph W. Pugh. USF and Hif1-α/β were produced by IVTT using the TnT® Coupled Reticulocyte Lysate System (T7 for HIF, SP6 for USF; Promega; L4601 and L4611) using the manufacturer recommended conditions. For the control lanes, the manufacturer supplied luciferase plasmid was added to the IVTT mixture instead of the HIF1-α and -β plasmids.[26]

EMSA probe preparation

For PCR generated HIF probes using Pfu Ultra (Agilent) substitution mCTP (NEB, N0356S) or hmCTP (bioline, BIO-39046) for CTP using the manufacture recommended conditions. Following template and primers were used EPO_template: AGGGGTGGAGGGGGCTGGGCCCTACGTGCTGTCTCACACAGCCTGTCTGACCTCTCGAC, EPO_FW: GGTGGAGGGGGCTGGGCCCTA, EPO_RW: CGAGAGGTCAGACAGGCTGTGTGAGACAGC.

The EPO template corresponds to the sequence responsible for the hypoxic upregulation of the *erythropoietin / epo* gene that contains a HRE (+97-156bp from the 3' end; the EPO_template corresponds to 38,138,574 - 38,138,632 of human chromosome 7 genomic scaffold (ref|NW_004078032.1).

The following DNA sequences were used:
EBOX_FW_C: CTCAGGCACCA**C**GTGGTGGGGGAT,
EBOX_RW_C: ATCCCCCACCA**C**GTGGTGCCTGAG,
EBOX_FW_mC: CTCAGGCACCA**mC**GTGGTGGGGGAT,
EBOX_RW_mC: ATCCCCCACCA**mC**GTGGTGCCTGAG,
EBOX_FW_hmC: CTCAGGCACCA**hmC**GTGGTGGGGGAT,
EBOX_RW_hmC: ATCCCCCACCA**hmC**GTGGTGCCTGAG,
HIF_FW_C: GCCCTA**C**GTGCTGTCTCACACAGCCT,
HIF_RW_C: AGGCTGTGTGAGACAGCA**C**GTAGGGC,
HIF_FW_mC: GCCCTA**mC**GTGCTGTCTCACACAGCCT,
HIF_RW_mC: AGGCTGTGTGAGACAGCA**mC**GTAGGGC,
HIF_FW_hmC: GCCCTA**hmC**GTGCTGTCTCACACAGCCT,
HIF_RW_hmC: AGGCTGTGTGAGACAGCA**hmC**GTAGGGC.

Oligonucleotides obtained from Sigma (unmodified), Eurofins (mC) or ATDBio (hmC) and resuspended in 10 mM TRIS pH 7.5. Concentrations were determined by measuring the A260 using the nanodrop. Theoretical absorption values and molecular weight were calculated using the web-based IDT oligo analyzer (http://eu.idtdna.com/analyzer/applications/oligoanalyzer/). The probes were annealed at an equimolar ratio (5 µM in 50 µL) in a PCR machine by cooling from 95°C to 4°C at 0.5°C/min. Annealing was confirmed by PAGE gel (15 %, 0.5x TBE; data not shown). After annealing, the concentration was again measured (Nanodrop ND-1000). Probes (50 ng) were phosphorylated using polynucleotide kinase (10 u, NEB, M0201) with ATP, [γ-^{32}P] (1 µL, Perkin Elmer, NEG002A250UC) for 30 min at 37°C. Radiolabelled probes were separated from free ATP using the QIAquick Nucleotide Removal Kit (Qiagen, 28304) as described in the manufacture manual.

Electrophoretic mobility shift assays (EMSAs)[6]

For assays with hypoxic inducible factor (HIF), the binding assays contained 2-5 µL of the IVTT mixture, 2 µL 10x binding buffer (final: 10 mM Tris pH 7.4, 50 mM NaCl, 50 mM KCl, 1 mM Ethylenediaminetetraacetic acid (EDTA), 5 mM DTT, 5% glycerol), 75ng dI/dC (1 µg/µL stock) diluted to 19 µL with water and incubated on ice for 10 min. Then 1 µL of radio-labeled DNA probe (0.08 ng/µL diluted from stock; see above for sequence) was added and the mixture was incubated on ice for an additional 30 min. The complete 20 µL EMSA mixture was then loaded onto a 5% PAGE gel (0,5x TBE (45 mM Tris-borate, 1 mM EDTA, pH 8.3), 0.7 mm, cast 1 day in advance and prerun at 240 V for 1h) and run at 240 V for 3.45-4 hours. The gel was put on Whatman paper, covered with saran wrap and dried at 80°C for 1 h. A phosphoscreen (Biorad) was exposed to the gel for 24-72 h and imaged using a Personal Molecular Imager (PMI, Biorad). Images were processed using the Quantity one analysis software (Biorad).

The USF binding assays consisted of 1 µL binding buffer (as above), 0.1 µL dI/dC (stock 1 µg/µL; 100ng final), IVTT mixture (1 µL for fig 3 and S6; 0.02-1.28 µL for titration curves), 1 µL probe (0.05ng/µL; see above for sequence) and water to 10 µL. The reactions were incubated on ice for 20 minutes and electrophoresis was performed as described above, but at room temperature.

MAX binding assays consisted of 1 µL binding buffer (as above), 0.1 µL dI/dC (stock 1 µg/µL; 100ng final), 1 µL MAX (2.5 µM or 2x dilutions for titration curves), 1 µL probe (0.05ng/µL; see above for sequence) and water to 10 µL. Electrophoresis was performed as for the USF assays.

1. P. Auffinger and Y. Hashem, *Bioinformatics*, 2007, **23**, 1035-1037.
2. D. L. C. Solomon, B. Amati and H. Land, *Nucleic Acids Research*, 1993, **21**, 5372-5376.
3. A. R. Ferre-D'Amare, P. Pognonec, R. G. Roeder and S. K. Burley, *Embo J*, 1994, **13**, 180-189.
4. Z. Wang, Y. Wu, L. Li and X.-D. Su, *Cell Res*, 2013, **23**, 213-224.
5. N. Huang, Y. Chelliah, Y. Shan, C. A. Taylor, S.-H. Yoo, C. Partch, C. B. Green, H. Zhang and J. S. Takahashi, *Science*, 2012, **337**, 189-194.
6. I. Kvietikova, R. H. Wenger, H. H. Marti and M. Gassmann, *Nucleic Acids Research*, 1995, **23**, 4542-4550.
7. A. R. Ferre-D'Amare, G. C. Prendergast, E. B. Ziff and S. K. Burley, *Nature*, 1993, **363**, 38-45.
8. R. H. Wenger, I. Kvietikova, A. Rolfs, G. Camenisch and M. Gassmann, *European Journal of Biochemistry*, 1998, **253**, 771-777.
9. M. Mellén, P. Ayata, S. Dewell, S. Kriaucionis and N. Heintz, *Cell*, 2012, **151**, 1417-1430.
10. X. J. Lu and W. K. Olson, *Nucleic Acids Research*, 2003, **31**, 5108-5121.
11. U. Heinemann and M. Hahn, *Journal of Biological Chemistry*, 1992, **267**, 7332-7341.
12. S. Fujii, A. H. Wang, G. van der Marel, J. H. van Boom and A. Rich, *Nucleic Acids Res*, 1982, **10**, 7879-7892.
13. G. P. Schroth, T. F. Kagawa and P. S. Ho, *Biochemistry*, 1993, **32**, 13381-13392.
14. C. Mayer-Jung, D. Moras and Y. Timsit, *Journal of Molecular Biology*, 1997, **270**, 328-335.
15. C. Mayer-Jung, D. Moras and Y. Timsit, *Embo J*, 1998, **17**, 2709-2718.
16. J. M. Vargason, B. F. Eichman and P. S. Ho, *Nat Struct Mol Biol*, 2000, **7**, 758-761.
17. I. Ohki, N. Shimotake, N. Fujita, J.-G. Jee, T. Ikegami, M. Nakao and M. Shirakawa, *Cell*, 2001, **105**, 487-497.
18. H. Hashimoto, J. R. Horton, X. Zhang, M. Bostick, S. E. Jacobsen and X. Cheng, *Nature*, 2008, **455**, 826-829.
19. B. A. Buck-Koehntop, R. L. Stanfield, D. C. Ekiert, M. A. Martinez-Yamout, H. J. Dyson, I. A. Wilson and P. E. Wright, *Proceedings of the National Academy of Sciences*, 2012, **109**, 15229-15234.
20. J. Otani, K. Arita, T. Kato, M. Kinoshita, H. Kimura, I. Suetake, S. Tajima, M. Ariyoshi and M. Shirakawa, *Journal of Biological Chemistry*, 2013, **288**, 6351-6362.

21. G. Winter, *Journal of Applied Crystallography*, 2010, **43**, 186-190.
22. P. Evans, *Acta Crystallographica Section D*, 2006, **62**, 72-82.
23. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *Journal of Applied Crystallography*, 2007, **40**, 658-674.
24. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proceedings of the National Academy of Sciences*, 1981, **78**, 2179-2183.
25. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallographica Section D*, 2010, **66**, 213-221.
26. M. E. Cockman, N. Masson, D. R. Mole, P. Jaakkola, G.-W. Chang, S. C. Clifford, E. R. Maher, C. W. Pugh, P. J. Ratcliffe and P. H. Maxwell, *Journal of Biological Chemistry*, 2000, **275**, 25733-25741.