THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 52, pp. 40879–40890, December 24, 2010

Structural and Biochemical Studies on the Chromo-barrel Domain of Male Specific Lethal 3 (MSL3) Reveal a Binding Preference for Mono- or Dimethyllysine 20 on Histone H4*[S]

Received for publication, April 14, 2010, and in revised form, September 19, 2010 Published, JBC Papers in Press, October 12, 2010, DOI 10.1074/jbc.M110.134312

Stanley A. Moore[‡1], **Yurdagul Ferhatoglu**[‡], **Yunhua Jia**[‡], **Rami A. Al-Jiab**[§], **and Maxwell J. Scott**[§2]

From the [‡]*Department of Biochemistry, University of Saskatchewan, Saskatoon, SK S7N 5E5, Canada and the* [§]*Institute of Molecular Biosciences, Massey University, Private Bag 11-222, Palmerston North, New Zealand*

We have determined the human male specific lethal 3 (hMSL3) chromo-barrel domain structure by x-ray crystallography to a resolution of 2.5 Å ($r = 0.226$, $R_{free} = 0.270$). hMSL3 contains a canonical methyllysine binding pocket made up of residues Tyr-31, Phe-56, Trp-59, and Trp-63. A six-residue insertion between strands β_1 and β_2 of the hMSL3 chromo-barrel domain directs the side chain of Glu-21 into the methyllysine binding pocket where it hydrogen bonds to the NH group of a bound cyclohexylamino ethanesulfonate buffer molecule, likely mimicking interactions with a histone tail dimethyllysine residue. *In vitro* binding studies revealed that both the human and *Drosophila* MSL3 chromo-barrel domains bind preferentially to peptides representing the mono or dimethyl isoform of lysine 20 on the histone H4 N-terminal tail (H4K20Me$_1$ or H4K20Me$_2$). Mutation of Tyr-31 to Ala in the hMSL3 methyllysine-binding cage resulted in weaker *in vitro* binding to H4K20Me$_1$. The same mutation in the *msl3* gene compromised male survival in *Drosophila*. Combined mutation of Glu-21 and Pro-22 to Ala in hMSL3 resulted in slightly weaker *in vitro* binding to H4K20Me$_1$, but the corresponding *msl3* mutation had no effect on male survival in *Drosophila*. We propose MSL3 plays an important role in targeting the male specific lethal complex to chromatin in both humans and flies by binding to H4K20Me$_1$. Binding studies on the related dMRG15 chromo-barrel domain revealed that MRG15 prefers binding to H4K20Me$_3$.

Nuclear histone acetyltransferase (HAT)[3] enzymes are found in multiprotein complexes that acetylate specific lysine residues on the N-terminal tails of histone proteins, thereby regulating nucleosome structure, chromatin packaging, and gene expression (1–15). MOF, a conserved member of the MYST (Moz, Ysb2, Sas2, Tip60) family of HAT enzymes, functions as the catalytic subunit in a number of distinct HAT complexes that target gene promoters (8), large contiguous domains of chromatin (3, 4, 14, 15), or non-histone proteins such as p53 (5–7). The precise targeting and substrate specificity of MOF relies on the presence of components distinct from the catalytic subunit (3, 4, 7, 8). Specifically, in the MOF-containing *Drosophila* male specific lethal (MSL) complex, the MSL3 protein is required for chromatin targeting, nucleosome binding, histone tail substrate recognition, and maximal MOF HAT activity (16–20).

The most well studied MOF-containing complex is the *Drosophila melanogaster* male specific lethal or MSL complex that binds selectively to large regions of the X-chromosome in male flies (14, 15, 21–26) where it is enriched at the 3′ ends of actively transcribed genes (27–30) and acetylates lysine 16 on histone H4 (H4K16Ac) (22, 31), thereby balancing male X-chromosomal gene expression. The *Drosophila* MSL complex contains the dMSL1, dMSL2, and dMSL3 proteins, the RNA/DNA helicase MLE, the HAT enzyme MOF, and one of two apparently functionally redundant non-coding RNAs (roX1 and roX2) (reviewed in Refs. 14, 15, and 21). The absence of any of the MSL components results in male lethality (14, 15, 23–26). The precise specificity of MOF for H4K16 and the targeting to specific domains of the male X-chromosome is determined by other components of the complex, in particular MSL1 and MSL3 (16–20, 23, 26, 28–30, 32).

A similar MOF-containing human MSL complex has also been identified, and although it contains homologous hMSL1, hMSL2, and hMSL3 subunits, it does not contain the MLE or RNA components, and has nothing to do with dosage compensation (3, 4). However, the human MSL complex does function as a global histone H4K16 acetyltransferase (3–5). Human MOF (hMOF) and the hMSL proteins also play a role in cell-cycle progression and DNA repair (3–5).

The *Drosophila* and human MSL3 proteins contain a highly conserved N-terminal chromo-barrel domain (CBD) (residues 2–91; 50% identical) plus a conserved C-terminal MRG domain (residues 196–512 in dMSL3; 25% identical to hMSL3) (3, 4, 33–35). The CBD of *Drosophila* MSL3 is known to contribute to nucleic acid and nucleosome binding in the *Drosophila* MSL complex (18–20), transcriptional up-regulation on the *Drosophila* male X-chromosome (18), and proper targeting and spreading of the dosage compensation complex *in*

* This work was supported by Canadian Institutes for Health Research Grant MOP 79377 (to S. M.), the Saskatchewan Health Research Foundation regional partnership program, and bridge funding from the University of Saskatchewan College of Medicine.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S4 and Table S1.

[1] To whom correspondence should be addressed: A3 Health Sciences Bldg., 107 Wiggins Rd., Saskatoon, SK S7N 5E5, Canada. Tel.: 306-966-4381; Fax: 306-966-4390; E-mail: stan.moore@usask.ca.

[2] Present address: Dept. of Genetics, Campus Box 7614, North Carolina State University, Raleigh, NC.

[3] The abbreviations used are: HAT, histone acetyltransferase; MSL, male specific lethal; CBD, chromo-barrel domain; bis-tris, 2-[bis(2-hydroxyethyl)-amino]-2-(hydroxymethyl)propane-1,3-diol; CAPS, 3-(cyclohexylamino)-1-propanesulfonic acid; CHES, 2-(cyclohexylamino)ethanesulfonic acid; r.m.s., root mean square.

vivo (20). The CBD of dMSL3 likely contributes to targeting of the MSL complex to chromatin regions containing specific histone tail lysine methylation modifications (19, 20). The C-terminal MRG domain of dMSL3 is responsible for interactions with dMSL1, but does not interact directly with dMOF (16, 17). Through interactions bridged by dMSL1, dMSL3 stimulates the HAT activity of dMOF and controls its substrate specificity (16, 17).

The human MSL3 protein has been less well studied, but two versions of hMSL3 have been identified associated with hMOF, one contains the full-length protein, and the other lacks the N-terminal chromo-barrel domain (3, 4), the particular isoform of hMSL3 found associated with the hMSL complex is tissue-dependent (3). In contrast to the *Drosophila* MSL complex where dMSL3 interacts directly with dMSL1 (16, 17), hMSL3 has been shown to interact directly with hMOF via its conserved C-terminal MRG domain (4).

To better understand the role played by MSL3 in the MSL HAT complex and histone tail recognition, we have undertaken structural and biochemical studies of the highly conserved MSL3 CBD and compared its *in vitro* histone tail binding to that of the related MRG15 CBD (9, 35). After submission of this manuscript, an independent structural and biochemical study on the *Drosophila* and human MSL3 CBDs was published (36). The results and conclusions of that study differ in part from those presented in this report and will be discussed in comparison with our findings.

EXPERIMENTAL PROCEDURES

Cloning and Protein Expression—The *D. melanogaster* MSL3 CBD (amino acid residues 2–91) was cloned using PCR from a *msl3* cDNA. The *Homo sapiens* MSL3 CBD was cloned using a hMSL3 cDNA kindly provided by Dr. Edwin Smith (Stowers Research Institute). The dMRG15 CBD was cloned from a drmg15 cDNA acquired from the *Drosophila* Genomics Resource Center, Bloomington, IN. The *D. melanogaster msl3* gene, a *msl3-TAP* tag fusion (20, 28) was kindly provided by Dr. Mitzi Kuroda, Harvard Medical School. The PCR primers used for cloning the dMSL3 (residues 2–91), hMSL3 (residues 2–93), and dMRG15 (residues 2–90) CBDs from their respective cDNAs are provided in the supplemental data (supplemental Table S1). The forward primers all contain a BamHI site and the reverse primers an EcoRI site and stop codon, and were designed for use with the P-GEX-6P3 GST fusion vector (GE Healthcare). PCR amplification was carried out using *Pfu* high fidelity DNA polymerase using standard reaction conditions. PCR product purification, and ligation with a gel-purified BamHI/EcoRI-digested vector were carried out using standard protocols. The cloning of dMSL3-CBD-GEX6P3, dMRG15-CBD-GEX6P3, and hMSL3-CBD-GEX6P3 was verified by restriction endonuclease digestion of the purified plasmid and DNA sequencing. Competent *Escherichia coli* BL21(DE3) cells were transformed with the purified plasmid, and an overnight culture was diluted 1/200 into fresh LB containing ampicillin (concentrations = 100 μg/ml). Cells were grown at 37 °C to an optical density (OD) of 0.6 and induced with 0.1 mM isopropyl β-D-thiogalactopyranoside. The temperature was then shifted to 25 °C and cells were grown for approximately 16 h. Cells were pelleted by centrifugation at 8000 × *g*, suspended in lysis buffer, frozen overnight, and subjected to two passes through a French Press.

Site-specific Mutagenesis—The Stratagene QuikChange mutagenesis kit was used for generating all point mutants. Mutations were confirmed by DNA sequencing. Primers for mutation of hMSL3 and dMSL3 are provided under supplemental Table S1.

Generation of Transgenic Drosophila—Transgenic *Drosophila* were generated as described (19). Briefly a mixture of 300 ng/ml of plasmid DNA containing attB and *msl3-TAP* and 0.5 mg/ml of ϕC31 capped integrase mRNA were injected into embryos from the y w; attP1; msl3^1/TM3,Sb line. The injection stock carries an attP1 integration site on chromosome 2R. G_0 individuals were crossed with y w; +/CyO; msl3^1/TM3, Sb, and the transgenic offspring identified by green fluorescent eyes. For complementation, we crossed y w; attP1 y+{p[gfp+ *msl3-TAP**-pGreeni]}; msl3^1/TM3,Sb males with y w; msl3^1 homozygous females. The rescue rates were calculated by dividing the number of rescued males by the number of their stubble (TM3) brothers. Statistical tests of significance were performed used χ-squared analysis. MSL3TAP* was MSL3 WT, MSL3 Glu^{21}-Pro^{22}-Ala, or MSL3 Tyr^{31}-Ala.

Protein Purification—E. coli cell lysates were pretreated with DNase I and then clarified by centrifugation at 15,000 × *g* for 1 h and then loaded onto a 10-ml glutathione-Sepharose column pre-equilibrated in PBS buffer and washed extensively. Each fusion protein was eluted with 0.1 mM reduced glutathione, and then dialyzed into Precission protease cleavage buffer. Precission protease was then incubated with the GST fusion overnight at 4 °C, followed by dialysis to remove any remaining glutathione, and then again subjected to glutathione-Sepharose affinity chromatography to remove the GST tag and remaining protease. Partially purified proteins were then subjected to ion-exchange chromatography (dMSL3: Source S, pH 5.0, 50 mM sodium malonate; hMSL3: Source S, pH 6.5, 50 mM bis-tris propane; dMRG15: source Q, pH 8.0, 20 mM Tris) and a 0.1 to 1.0 M NaCl gradient followed by analytical gel-filtration chromatography (Superdex 70), which resulted in highly pure recombinant proteins.

Estimation of Nucleic Acid Contaminants—We used an ethidium bromide fluorescence assay (37) to assess the presence of contaminating double-stranded nucleic acid. 10 μl of a 100 μg/ml of solution of calf thymus DNA was diluted into 2 ml of a 0.5 μg/ml of solution of ethidium bromide in 5 mM Tris-HCl, 0.5 mM EDTA, pH 8.1, yielding a fluorescence of 5.6 units at 600 nm (excitation 430 nm). The assay buffer blank gave a fluorescence of 0.858. 10 μl of each purified protein sample at 2.8 mg/ml concentration was diluted in 2 ml of assay buffer to give fluorescence readings of 0.855 for hMSL3, 0.818 for dMSL3, and 0.881 for dMRG15. Hence, there was less than 0.002 μg of double-stranded nucleic acid per μg of protein in our samples. The far UV absorption spectrum of each protein at 2.8 mg/ml exhibited only a small shoulder absorbance at 260 nm, the observed A_{280}:A_{260} ratios for the three protein samples (undiluted) was measured to be 1.67 for hMSL3 (1.67 calculated), 1.82 for dMSL3 (1.88 calculated),

and 1.86 for dMRG15 (1.84 calculated), indicating less than 2% contamination from nucleic acids. Details of protein absorbance ratio calculations are provided under supplemental data. Extinction values for Phe, Tyr, and Trp at 260 and 280 nm were from the literature (38, 39).

Protein Crystallization—Purified protein was concentrated to 12 mg/ml and crystallization conditions for the purified hMSL3 CBD were found using the Wizard II commercial screen (Emerald Biosciences) from a 2.0 M ammonium sulfate solution buffered at pH 10.5 with 0.1 M CAPS. The initial crystals were clusters of small plates. Single crystals of the hMSL3 CBD were grown by repeated seeding in solutions containing 1.6–2.0 M ammonium sulfate, pH 8.0–9.0, and 0.1 M CHES. Seeded crystals took from 1–2 weeks to grow to a maximal size of 0.02 × 0.1 × 0.1 mm as thin plates. Crystals for flash-freezing were prepared with a mother liquor containing 18% (w/v) glycerol. X-ray diffraction data to 2.5-Å resolution were collected on a crystal flash frozen in a stream of liquid nitrogen vapor at the Canadian Light Source Small-gap undulator beamline (R_{merge} = 0.078 to 2.5-Å resolution (see Table 1). Data processing indicated the Laue group was 2/m, space group C2, but with a β angle = 89.6°. The data would not merge in higher symmetry Laue classes or in a primitive monoclinic setting.

Structure Determination and Refinement—Diffraction data were processed with HKL2000 (40) (Table 1). The structure was solved using molecular replacement, the search model was the refined x-ray structure of the human MRG15 CBD (PDB 2F5K) (41) and PHASER (42) located five molecules in the asymmetric unit. Subsequent model building and refinement using Coot (43), CNS (44), and finally Refmac 5.5 in CCP4 (45–47) resulted in a refined model with excellent stereochemistry (Table 1) as illustrated by a Ramachandran plot (48) (supplemental Fig. S1). During refinement, 5% of reflections were randomly put aside as a test set for calculation of the free *R* factor. Figures depicting the contents of the asymmetric unit and electron density for one of the hMSL3 methyllysine binding pockets are provided under supplemental Fig. S2. For Fig. 1, sequence alignments were calculated with T-Coffee (49) and rendered using Espript (50). Molecular superpositions were carried out using LSQKAB in CCP4 (45). All molecular figures were drawn using either Molscript/Raster3D (51, 52) or PyMOL. Qualitative electrostatic surfaces were drawn using the adaptive Poisson-Boltzmann solver as implemented in PyMOL and are contoured at ±75 mV (units of kT/*e*). The refined atomic coordinates and structure factors for the hMSL3 CBD have been deposited with the RCSB protein data bank (PDB code 3OB9).

Synthetic Peptides—Synthetic peptides with the indicated sequences and post-translational modifications were purchased from Sigma-Genosys or New England Peptide at greater than 95% purity. The peptides used in this study all contained an unmodified N terminus and an amidated C terminus unless indicated otherwise: H4K20Me$_{1-3}$, GKGGAKRHR(KMe$_{1-3}$)VLRDY; H3K4Me$_2$, ART(KMe$_2$)QTARKSTGGKAY; H3K36Me$_{1-3}$, APSTGGV(KMe$_{1-3}$)KPHRYR; H3K9Me$_1$, KQTAR(KMe$_1$)-STGGKAY; H3K18Me$_1$, GGKAPR(KMe$_1$)QLATY; H3K27Me$_1$, TKAAR(KMe$_1$)SAPSTGY; H3K79Me$_2$, AQDY(KMe$_2$)TDLR.

Surface Plasmon Resonance—The dMSL3, hMSL3, and dMRG15 chromo-barrel domains were covalently linked to carboxyl methylcellulose-based sensor chips (CM5, GE-Healthcare) using *N*-ethyl-*N*-(3-dimethylaminopropyl)carbodiimide and N-hydroxysuccinimide (NHS), and quenched with ethanolamine according to the manufacturer's instructions. Typically 4000–5000 response units (2000 for dMRG15) of protein were immobilized per experiment. All measurements were carried out on a Biacore X instrument. CBD histone tail peptide dissociation constants were calculated using a Langmuir 1:1 steady state equilibrium binding model ($A + B \leftrightarrows AB$) with Scrubber. Steady state affinity binding of peptides to the immobilized CBDs was measured for peptide concentrations ranging from either 1 to 800 μM or 1 μM to 3.8 mM. Typically, complete binding data for at least two peptides were measured per CM5 chip, reproducibility from chip to chip was excellent. Peptide concentrations were determined by tyrosine UV absorbance at 276 nm using a Nanodrop device, as each synthetic histone tail peptide was designed to contain a C-terminal tyrosine residue. Peptides were dissolved in a running buffer of 50 mM HEPES pH 7.4, 3 mM EDTA, and 150 or 250 mM NaCl as indicated.

RESULTS AND DISCUSSION

Tertiary Structure of the hMSL3 Chromo-barrel Domain—The structure of the human MSL3 CBD has been fully refined to 2.5-Å resolution (*R*-factor = 0.226; R-free = 0.270) (Table 1 and supplemental Figs. S1 and S2). There are five independent copies of the hMSL3 CBD in the asymmetric unit of the crystal lattice and the tertiary structures of the five subunits are very similar, (backbone atom pairwise r.m.s. differences of 0.68 to 0.74 Å) (supplemental Fig. S2). The hMSL3 CBD structure is similar to that of the MRG15 CBD (41), folding as a 5-stranded antiparallel β-barrel domain with a C-terminal α-helix (backbone atom rms difference of 0.99 Å for 68 equivalent residues) (Fig. 1 and supplemental Fig. S3). Similar to the CBD of MRG15 and other tudor and chromo-barrel domains (41, 53–57), a canonical methyllysine binding pocket is evident at one end of the hMSL3 β-barrel domain, with important residues coming from surface loops between strands β_1-β_2 and β_3-β_4 (Figs. 1 and 2). In hMSL3, there is a highly conserved six-residue insertion relative to MRG15, at the loop between strands β_1 and β_2, the proximity of this loop to the presumed methyllysine binding pocket suggests this loop may be important for methyllysine binding specificity (Fig. 1). The five hMSL3 CBD molecules observed in the crystal exhibit minimal disorder or heterogeneity at the methyllysine binding pocket and four of the five monomers pack as two almost identically arranged dimers (a modest 670 Å^2 is buried at the dimer interface) (supplemental Fig. S2). The fact that the hMSL3 CBD preferentially packs as a dimer in the crystals is unusual as most chromo-barrel domains, tudor domains, and chromodomains, including those of MRG15, and 53BP1, tend to pack as monomers (41, 53). However, we see no evidence of either hMSL3 or dMSL3 CBD dimerization in solution using analytical gel filtration chromatography (not shown). Hence the biological significance of hMSL3 CBD dimerization in the crystal lattice is unclear.

TABLE 1

Structure refinement statistics for *H. sapiens* MSL3 chromo-barrel domain

Data collection and refinement statistics	40.0-2.50 Å (2.568-2.500)[a]
Unit cell parameters (C2) (*a*, *b*, *c*) (Å) β (°)	179.18, 36.697, 85.556, 90.39
Wavelength (Å)	0.97934
Resolution (Å)	40-2.50 (2.59-2.50)[a]
No. unique reflections[b]	19,514 (1789)[a]
Redundancy[b]	3.6 (2.9)[a]
Completeness (%)[b]	99.0 (91.8)[a]
Average I/σ^b	16.6 (2.1)[a]
R_{merge}[b,c]	0.078 (0.374)[a]
Refinement resolution limits	40.0-2.50 Å (2.57-2.50)[a]
No. of reflections in working set	18,515 (1219)[a]
No. of reflections in test set	997 (60)[a]
R_{work}[d]	0.2256 (0.271)[a]
R_{free}[e]	0.2702 (0.344)[a]
No. of amino acid residues	361 (3733 atoms)
No. of water molecules and sulfates	57 and 17
No. of ligands (CHES)	3
Average *B*-factor (Å^2)[d,f]	27.8
R.m.s. deviations B bonded MC atoms (Å^2)[f]	0.39
R.m.s. deviations B bonded SC atoms (Å^2)[f]	0.76
R.m.s. deviations B angle MC atoms (Å^2)[f]	1.20
R.m.s. deviations B angle SC atoms (Å^2)[f]	1.97
R.m.s. deviations bond lengths (Å)[f]	0.0090
R.m.s. deviations angles (°)[f]	1.09
Residues in preferred Ramachandran (%)[g]	93.8 (5.7)

[a] Values in parentheses correspond to the highest resolution shell.
[b] Data processing statistics calculated using Denzo/HKL2000 (37).
[c] $R_{merge} = \Sigma_{hkl} \Sigma_i | \langle I(hkl)_{obs} \rangle - I(hkl)_{obs,\ i}| / \Sigma_{hkl,i} I(hkl)_{obs,i}$, where $I(hkl)_{obs,i}$ is the individual measurement of an *hkl* intensity and $\langle I(hkl)_{obs} \rangle = \Sigma_i I(hkl)_{obs,i}/n$; where $i = 1$ to n individual reflections are measured.
[d] $R_{work} = \Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$, where $|F_{obs}(hkl)|$ and $|F_{calc}(hkl)|$ are the observed and calculated amplitudes, respectively, for the structure factor $F(hkl)$.
[e] R_{free} is the equivalent of R_{work} for 5% of the reflections (randomly selected), which were not used in structure refinement.
[f] B-factor (TLS component not included) and r.m.s. deviation values were calculated with Refmac as implemented in CCP4 (41–43).
[g] The Ramachandran plot was generated with Procheck in CCP4, residues in allowed regions are in parentheses (41, 44).

An Aromatic Cage Methyllysine Binding Pocket in hMSL3—Residues making up the presumed methyllysine binding pocket of hMSL3 are Glu^{21}, Pro^{22}, and Tyr^{31} from the loop between strands β_1 and β_2, and Phe^{56}, Trp^{59}, and Trp^{63} from the loop between strands β_3 and β_4 of the β-barrel (Figs. 1 and 2). In the crystal structure, the deep binding pocket is occupied by the piperazine moiety of a CHES buffer molecule for two of the five independent copies of the CBD (Fig. 1 and supplemental Fig. S2). The binding of the piperazine moiety of CHES in a methyllysine pocket is reminiscent of the morpholino ring of MES binding in the methyllysine binding pocket of the h-l(3)mbt repeat protein and is thought to mimic the presence of a mono- or dimethyllysine residue (58). Other conserved residues on the hMSL3 β_1-β_2 loop near the methyllysine binding pocket are Asp^{23}, Thr^{25}, and Lys^{26}. These residues are conserved in MSL3 sequences and are positioned at the opening of the methyllysine pocket such that their side chains could form hydrogen bonds with the backbone peptide groups flanking the methyllysine residue in a histone tail substrate (Figs. 1 and 2). Importantly, the analogous β_1-β_2 loop is much shorter in the structure of the hMRG15 CBD and in place of Glu^{21} and Pro^{22} in hMSL3, there is a conserved histidine residue that forms one side of the hMRG15 methyllysine binding pocket (41) (Fig. 2). Otherwise, the overall shape and local environment of the hMSL3 methyllysine binding pocket are very similar to that of hMRG15, suggesting that these proteins may recognize a similar ligand (Fig. 2) (41).

The methyllysine binding pocket in hMSL3 also closely resembles the dimethyllysine binding pocket in the tandem tudor domain of human 53BP1 and monomethyllysine binding pockets in mbt repeat proteins that recognize the respective methylated form of lysine 20 on histone H4 (53–55) (Fig. 2). The Glu^{21} carboxylate side chain that makes up the side of the methyllysine binding pocket on the β_1-β_2 loop of hMSL3 resembles the positioning of the side chain of Asp^{1521} in 53BP1, required for dimethyllysine binding, but on the opposite β_3-β_4 loop (53) (Fig. 2). Based on simple steric and hydrogen bonding arguments, the presence of Glu^{21} in the hMSL3 methyllysine pocket would be expected to modulate the binding specificity of hMSL3 in the direction of binding mono- or dimethyllsyine residues relative to trimethyllysine. A trimethyllysine residue binding in the pocket would most likely result in unfavorable van der Waals interactions between the carboxylate group of Glu^{21} and the trimethylammonium group of the trimethyllysine. The side chain of Pro^{22} on the β_1-β_2 loop would likely constrict the space inside the methyllysine binding pocket enough so that only a mono- or dimethyllysine residue can easily be accommodated. Neither human MRG15 nor yeast Eaf3 CBDs have a carboxylate residue or a proline residue facing into their respective methyllysine binding pockets (41, 56, 57) (Fig. 2).

The hMSL3 CBD structure exhibits a prominent surface groove between the methyllysine binding pocket and helix α_1 that could easily accommodate a short extended peptide (Fig. 2). Residues on a similar surface groove in the structure of the related Eaf3 CBD undergo chemical shift changes upon binding of the H3K36Me$_3$ peptide (57). In hMSL3, this surface depression is also partially negatively charged and contains several polar side chains (Tyr^{31}, Asp^{32}, Asn^{57}, Trp^{59}, and Gln^{83}) that could potentially hydrogen bond to the backbone peptide groups from the amino-terminal residues of a bound histone tail (Fig. 2*D*). In support of this idea, the side chain indole nitrogen of Trp^{59} of hMSL3 makes a H-bond to the backbone oxygen from the β_2-β_3 hairpin loop on a nearby hMSL3 molecule in the crystal lattice, hence possibly mimicking the contact of a histone tail peptide backbone carbonyl group to the CBD (not shown). The hMSL3 surface depression also contains a small hydrophobic patch consisting of residues Val^{29}, Ala^{87}, and Ala^{90} that may participate in peptide binding (Fig. 2*D*).

The MSL3 CBD Preferentially Binds H4K20Me$_1$ or H4K20Me$_2$ in Vitro—The binding of CHES to the hMSL3 CBD and structural similarities with the 53BP1 methyllysine binding pocket suggest that hMSL3 will have a binding preference for peptides containing mono- or dimethyllysine. Furthermore, the presence of a highly conserved hMSL3 β_1-β_2 insert at the hMSL3 methyllysine binding pocket relative to MRG15 suggests that there may be different methyllysine binding specificities for these two otherwise closely related CBDs. Hence, we used histone tail peptide sequences containing methylation modifications corresponding to common histone H3 and H4 N-terminal epigenetic modifications in *Drosophila* and humans (59), and screened peptide binding to the hMSL3, dMSL3, and dMRG15 CBDs using surface plasmon resonance (Figs. 3 and 4 and supplemental Fig. S4). The

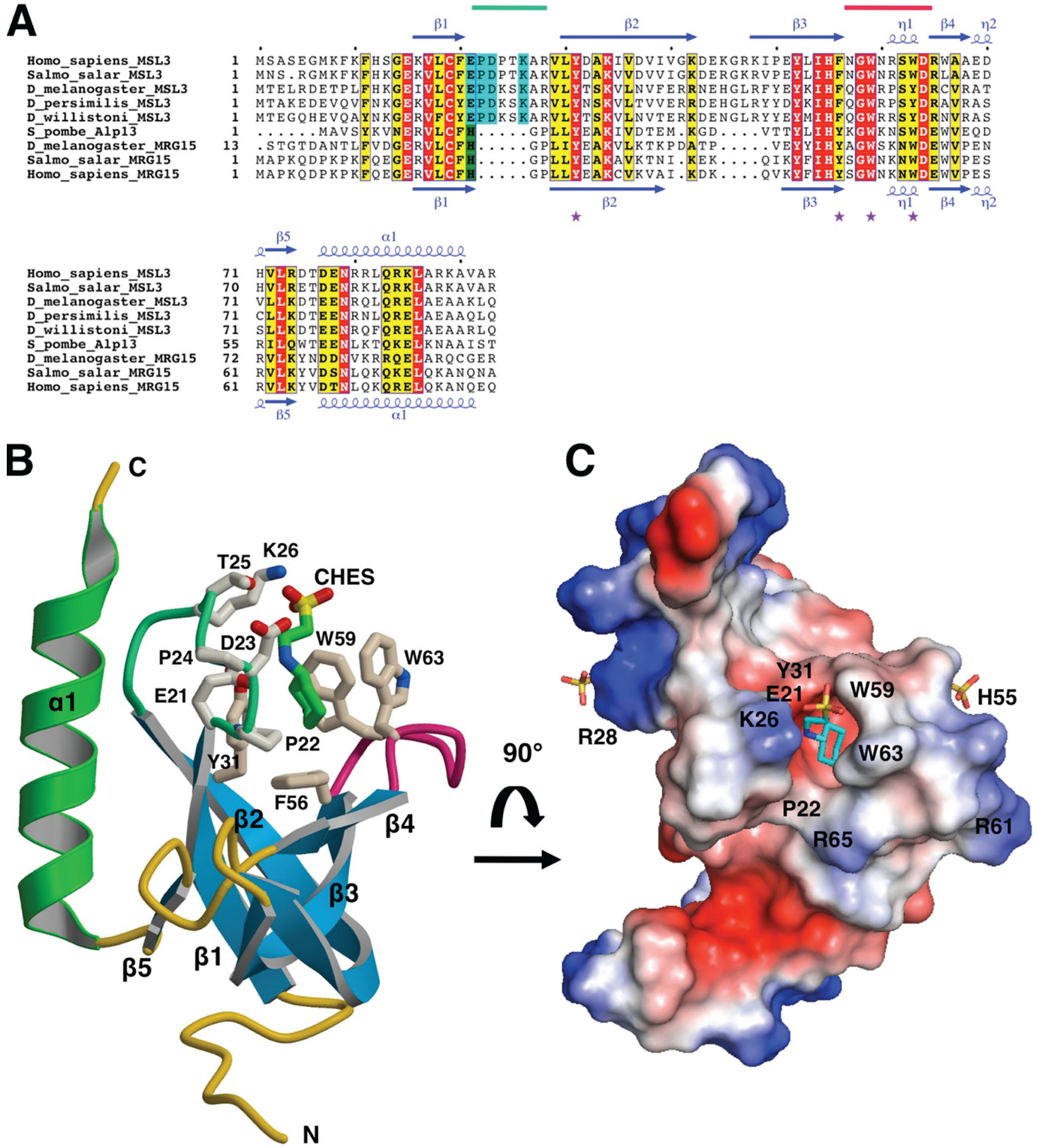

FIGURE 1. **Tertiary structure of human MSL3 chromo-barrel domain.** *A*, multiple sequence alignment of MSL3 and MRG15 CBD sequences from higher eukaryotes. Strictly conserved residues are shaded *red*, highly conserved residues *yellow*. A six-residue insertion specific to MSL3 is shaded *blue*, whereas a histidine residue found in the methyllysine binding cage of MRG15 is shaded *green*. Secondary structure assignments were derived from x-ray structures of hMSL3 (this work) and hMRG15 CBDs (PDB 2F5K) (41). The locations of the β1-β2 and β3-β4 loops making up the methyllysine binding pocket are marked with *light green* and *magenta lines*, respectively, *above* the sequences. Aromatic cage residues of the methyllysine binding pocket are marked with *purple stars*. The N-terminal 12 residues (MGEVKPAKVENY) of dMRG15 are not shown for clarity. *B*, ribbon diagram of hMSL3 residues 5–93 (subunit A in the crystal). The MSL3-specific loop between strands β1 and β2 is colored *light green*. Amino acid side chains associated with the presumed methyllysine binding pocket and a bound CHES buffer molecule are shown as stick models. *C*, qualitative electrostatic surface rendering of the hMSL3CBD (*red* negatively charged; *blue* positively charged, see "Experimental Procedures") and bound CHES and sulfate anions are shown as stick representations. The hMSL3 CBD is rotated ~90 degrees toward the viewer relative to *B*, looking directly into the methyllysine binding pocket.

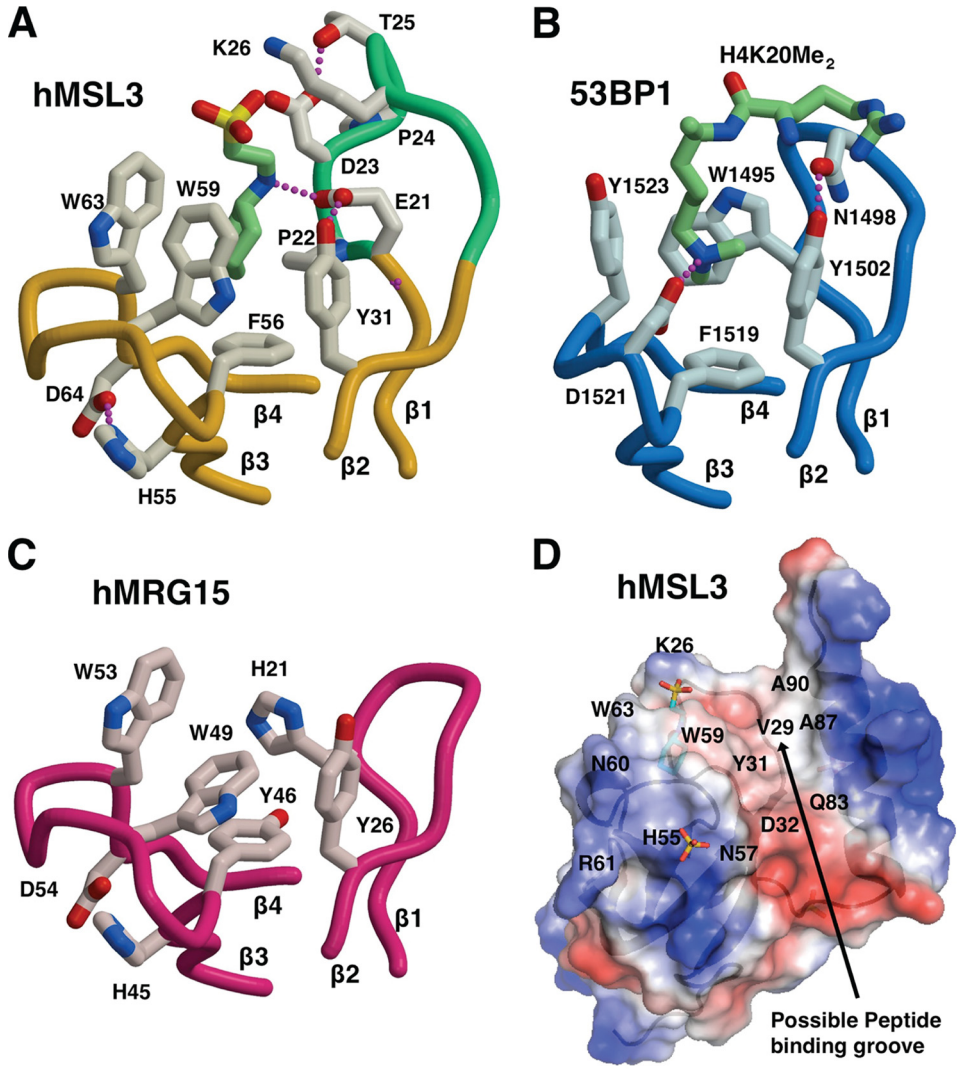

FIGURE 2. **Comparison of methyllysine binding pockets in hMSL3, hMRG15, and h53BP1.** *A*, methyllysine binding pocket in the hMSL3 CBD depicting the bound CHES buffer molecule (*green* carbons), colored as described in the legend to Fig. 1*B*. *B*, methyllysine binding pocket in the 53BP1 tandem tudor domain in *blue* (53) showing the bound H4K20Me$_2$ peptide with carbons colored *light green*. Only the first tudor domain is shown. *C*, methyllysine binding pocket in the hMRG15 CBD drawn in *magenta* with side chain carbons in *light pink* (41). Each pocket is drawn in an identical orientation, showing similar structural elements. Hydrogen bonds are drawn as *purple dotted lines*. Residues are labeled according to the text and the published structures. *D*, a potential peptide interaction surface on the hMSL3 CBD, the view is rotated ~180° about the vertical axis relative to Fig. 1*B*. The hMSL3 CBD is rendered as a semi-transparent electrostatic surface (*red* negatively charged, *blue* positively charged) overlaid onto a ribbon diagram of the molecule. The bound CHES molecule (*blue* carbons) and the sulfate anion near His^{55} are shown as stick representations. Residues mentioned in the text are labeled accordingly.

majority of binding experiments with the MSL3 CBDs were conducted under fairly stringent conditions (250 mM NaCl) to avoid possible nonspecific ionic interactions of the predominantly positively charged histone tail peptides with either the CBD or the CM5 sensor chip surface (see "Experimental Procedures"). For the H4K20Me$_1$ and H4K20Me$_3$ peptides, we also measured the binding affinity for dMSL3 and hMSL3 CBDs at a more physiological salt concentration of 150 mM (Table 2). We carried out a series of single injection experiments for a larger number of peptides for all three CBDs at 50 and 100 μM peptide concentrations in 150 mM NaCl running buffer (100 μM shown) to rule out significant binding by other sequences and ensure that the buffer salt concentration did not affect the relative affinities of peptide binding (supplemental Fig. S4). The results demonstrate that the relative orders of binding affinity of the tested peptides for both the hMSL3 and dMSL3 chromo-barrel domains are highly reproducible and independent of the running buffer salt concentrations tested. The preferential order of peptide binding we observe for both the hMSL3 and dMSL3 chromo-barrel domains is H4K20Me$_1$ ≈ H4K20Me$_2$ > H4K20Me$_3$ > H3K36Me$_1$ ≈ H3K36Me$_2$ > H3K4Me$_2$ > H3K36Me$_3$ (Table 2, Figs. 3 and 4, and supplemental Fig. S4). Hence, the MSL3 CBD in both humans and *Drosophila* preferentially binds monomethyl or dimethyllysine histone tail peptides over trimethyl-modified lysine peptides and H4 tail sequences over H3, suggesting that the binding specificity is conserved across species, and that the biological function is to recognize and bind H4K20Me$_1$ or H4K20Me$_2$ (Figs. 3 and 4). Based on our fitted K_d values for the relevant peptides in buffer containing 250 mM NaCl, the preference for H4K20Me$_1$ or H4K20Me$_2$ over H3K36Me$_3$ is approximately a factor of 50 (Table 2). The K_d for hMSL3 and

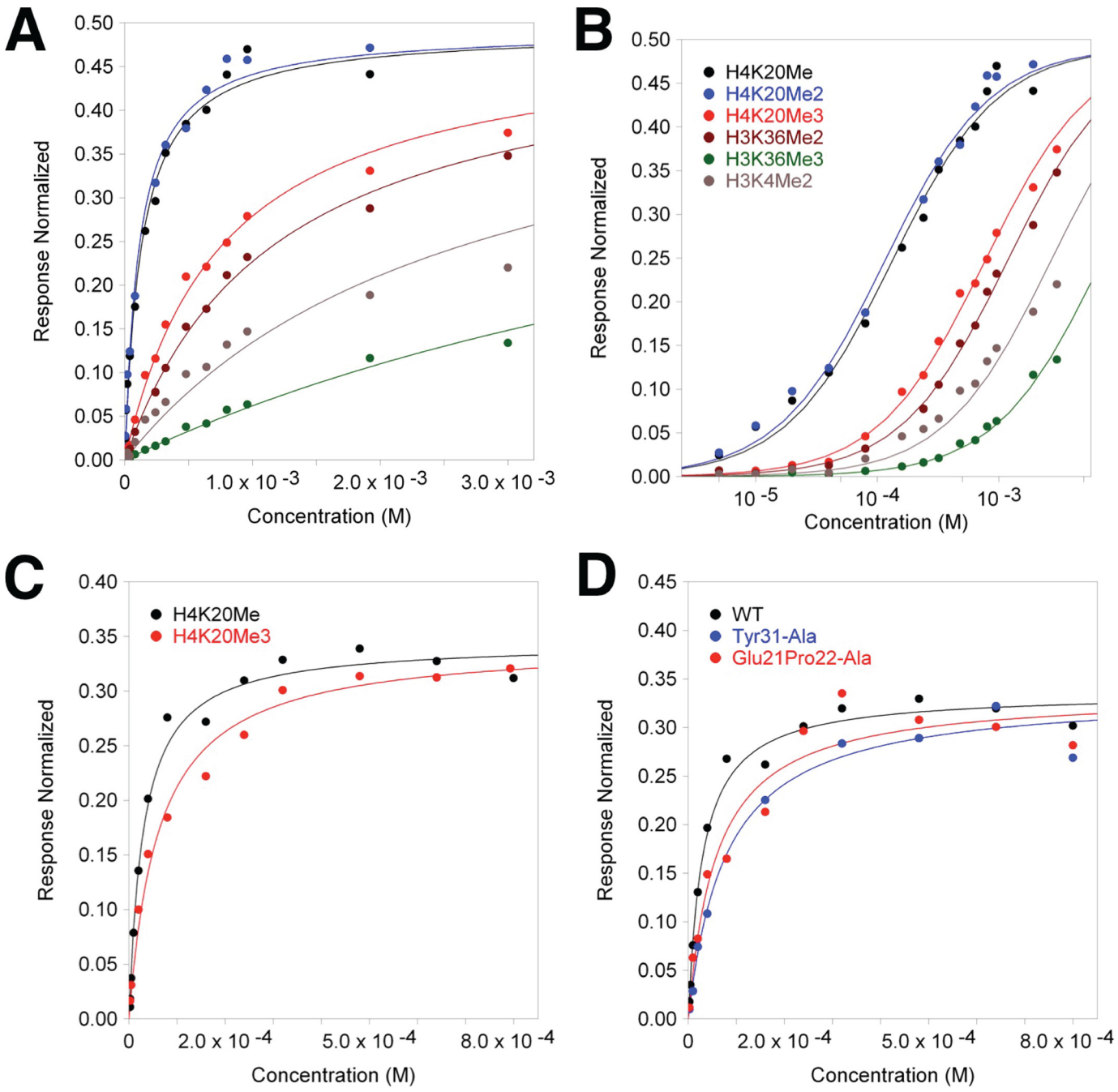

FIGURE 3. **Binding affinity of the human MSL3 chromo-barrel domain for the indicated methyllysine containing histone tail peptides.** *A*, surface plasmon resonance steady state equilibrium binding for methylated histone tail peptides over the hMSL3 chromo-barrel domain (in response units normalized to maximum theoretical occupancy of ligand) in 250 mM NaCl, 3 mM EDTA, 100 mM HEPES pH 7.5. Individual data points and fitted curves based on the calculated K_d values are shown for each peptide series. *B*, semilog plot of data presented in *A*. The legend applies to both *A* and *B*. *C*, binding data for hMSL3 CBD with $H4K20Me_1$ and $H4K20Me_3$ in 150 mM NaCl buffer. *D*, binding data for WT hMSL3, the Glu^{21}-Pro^{22}-Ala and Y31A hMSL3 CBD mutants with $H4K20Me_1$ in 150 mM NaCl running buffer.

$H4K20Me_1$ in 150 mM NaCl buffer is 31 μM, the value for the dMSL3 CBD under identical conditions is 224 μM, a factor of seven weaker (Fig. 4 and Table 2). We do not have an explanation for why the dMSL3 CBD binds to $H4K20Me_1$ more weakly than its human counterpart.

Comparison with a hMSL3 CBD:DNA H4K20 Structure—Recently, a structure of the hMSL3 CBD bound to duplex DNA and $H4K20Me_1$ was published (36). In contrast to our results, that study concluded that methylated histone tail binding to hMSL3 or dMSL3 CBDs occurred only in the presence of double-stranded DNA and that histone tail binding was significant only for $H4K20Me_1$ and not other histone tail sequences or higher degrees of methylation. In contrast, our results suggest that hMSL3 is capable of binding $H4K20Me_1$ with reasonable affinity in the absence of nucleic acids as our purified CBDs were deemed essentially nucleic acid free (see "Experimental Procedures"). We do not have an explanation for the reported binding discrepancy between the two studies.

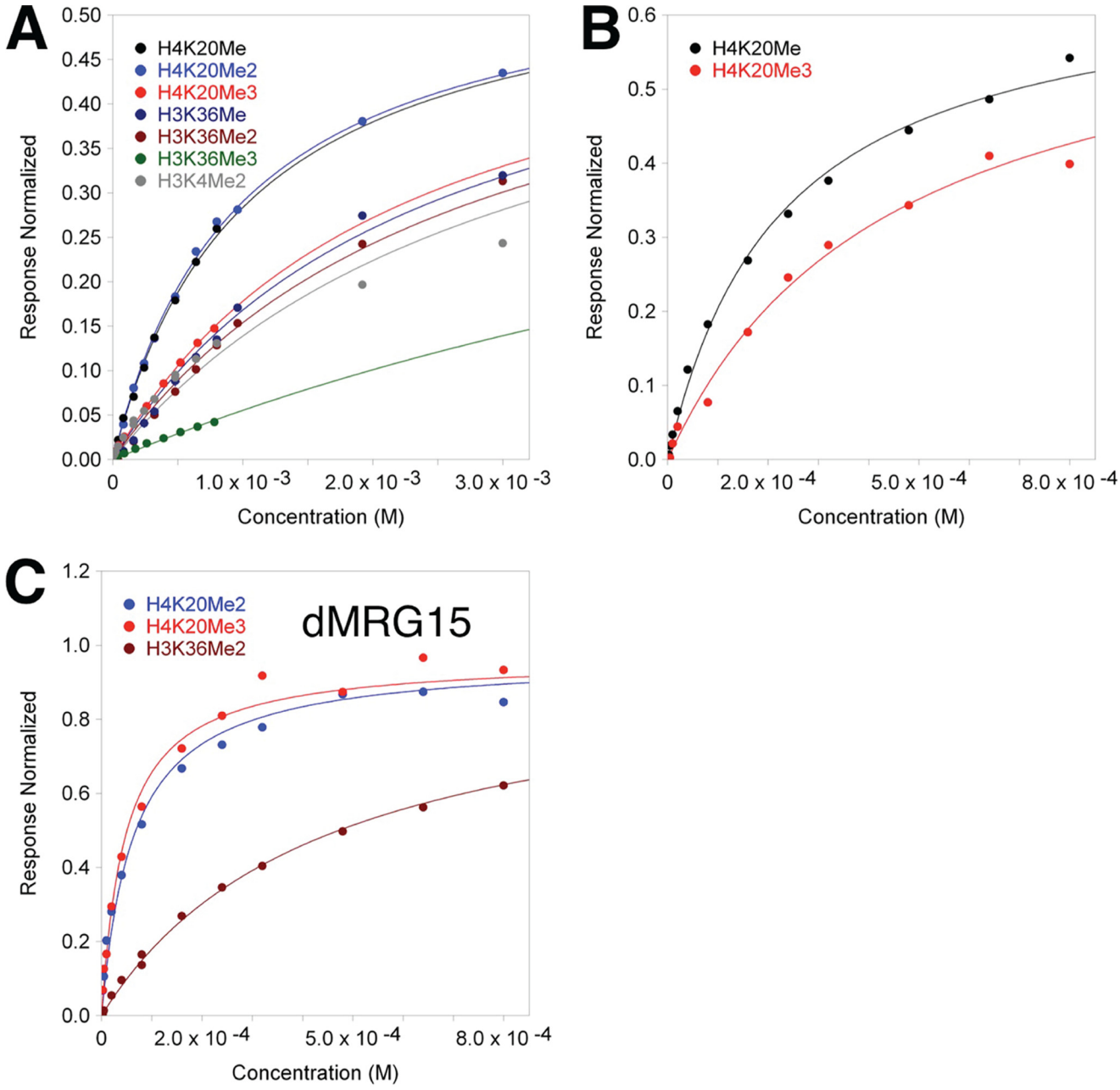

FIGURE 4. **Binding affinity of the *D. melanogaster* MSL3 and MRG15 chromo-barrel domains for methyllysine containing histone tail peptides.** *A*, surface plasmon resonance steady state equilibrium binding (in response units normalized to maximum theoretical occupancy of ligand) for the *D. melanogaster* CBD and histone tail peptides in 250 mM NaCl, 3 mM EDTA, 100 mM HEPES pH 7.5. Individual data points and fitted curves based on the calculated K_d values are shown for each peptide series. *B*, as in part A, but 150 mM NaCl and only H4K20Me$_1$ and H4K20Me$_3$. *C*, SPR binding of the dMRG15 CBD to H4K20Me$_3$, H4K20Me$_2$, and H3K36Me$_2$ in 150 mM NaCl running buffer.

Our study used surface plasmon resonance with a fixed amount of immobilized protein and differing concentrations of added peptides to determine dissociation constants. The other study used fluorescence polarization of a fixed amount of fluoresceinated peptide in combination with varying concentrations of the recombinant protein to quantify interactions (36). Although we do not doubt that double-stranded DNA binding to the MSL3 CBD could modulate subsequent *in vitro* or *in vivo* methylated histone tail binding, our results suggest that the MSL3 CBD alone is sufficient for *in vitro* binding to H4K20Me$_1$ and discrimination over other histone tail sequences. We also note that the other study used dMSL3 and hMSL3 constructs each containing an extra nine N-terminal residues (MKKHHHHHH) to facilitate protein purification and an extra eight residues at the C terminus for hMSL3 (^{94}LRSTGRKK^{101}) (36). The hMSL3 and dMSL3 constructs used in our study contain five extra N-terminal residues (GPLGS) from the cloning vector. Least squares superposition of our atomic coordinates with that of the hMSL3 CBD in the hMSL3·DNA·peptide complex (PDB code 3M9P) revealed only modest (0.6 to 1.10 Å) pairwise r.m.s. differences in the positions of the backbone atoms, and very few changes in amino acid side chain conformation, suggesting the tertiary structure of the hMSL3 CBD is essentially the same in the

TABLE 2
Calculated dissociation constants for peptide chromo-barrel domain surface plasmon resonance steady state affinity measurements

Peptide	[NaCl]	K_d hMSL3[a,b]	K_d dMSL3[a,b]
	mM	μM	
$H4K20Me_1$	250	116(4)	1003(3)
$H4K20Me_2$	250	98(2)	980(3)
$H4K20Me_3$	250	700(20)	2230(7)
$H3K36Me_1$	250	ND[c]	2420(7)
$H3K36Me_2$	250	1070(3)	2730(8)
$H3K36Me_3$	250	6500(3)	9400(7)
$H3K4Me_2$	250	2460(7)	3140(9)
$H4K20Me_1$	150	31(9)	224(6)
$H4K20Me_3$	150	100(5)	441(9)

[a] Based on a 1:1 Langmuir binding steady state equilibrium, calculated using Scrubber (see "Experimental Procedures").
[b] K_d residuals from fitting are provided in parentheses.
[c] ND, not determined.

absence or presence of nucleic acids. However, our hMSL3 atomic model does form dimers, and does contain an extra 3–4 residues visible at the N terminus for four of the five subunits, namely Gly^6 to Phe^9. Differences at the C terminus of the two independent hMSL3 structure determinations are negligible, as residues 92–101 of the hMSL3·DNA complex are not included in the refined model of that structure and hence are likely disordered in solution (36).

Histone Tail Binding by the dMRG15 CBD—The MRG15 protein contains an N-terminal CBD highly similar to MSL3 (Figs. 1 and 2), and a more distantly related C-terminal MRG domain (9, 35). MRG15 also associates with Tip60, a MYST HAT highly similar to MOF as part of a large multiprotein complex that functions in gene regulation and DNA repair through acetylation of histones phospho-H2Av and H4 (9–13). Hence, the conserved domain organization of MSL3 and MRG15, plus their association with highly similar HAT enzymes, suggest a related biological function for the two proteins (3, 4, 9). However, the CBD of hMRG15 was previously reported to bind $H3K36Me_3$, although binding of methylated H4K20 peptides to MRG15 was not tested in that study (41). Similar binding studies on the related yeast Eaf3 CBD also reported preferential binding to $H3K36Me_2$ or $H3K36Me_3$ (56, 57), but one of the studies also suggested that H4-based peptides bound nearly as well (57). Given the high degree of structural similarity between the hMSL3 and hMRG15 CBDs, we therefore measured the *in vitro* binding of methylated histone tail sequences to the dMRG15 CBD to see if there was any similarity with the MSL3 CBD binding profile. The results are presented in Fig. 4*C* and supplemental Fig. S4. We found that the CBD of dMRG15 binds preferentially to $H4K20Me_3$ (K_d = 48 μM) and with similar affinity for $H4K20Me_2$ (K_d = 61 μM) in 150 mM NaCl containing buffers, although the single injection profile suggests consistently stronger binding for $H4K20Me_3$ (supplemental Fig. S4). Binding to $H3K36Me_3$ or $H3K36Me_2$ (K_d = 430 μM) is significantly weaker (Fig. 4*C* and supplemental Fig. S4), indicating that the dMRG15 CBD prefers binding $H4K20Me_3$ or $H4K20Me_2$ over $H3K36Me_3$ or $H3K36Me_2$. Hence both MRG15 and MSL3 CBDs have similar histone H4 tail binding profiles, preferring H4K20 over H3 sequences, although MSL3 prefers a lower degree of methylation of Lys^{20}, and MRG15 prefers trimethylation. As the structures of the hMSL3 and hMRG15 CBDs are highly similar (Figs. 1 and 2), the different binding affinities for $H4K20Me_1$ *versus* $H4K20Me_3$ suggest that subtle differences in the methyllysine binding pockets of these proteins confer discrimination for the degree of methylation of H4K20.

TABLE 3
Rescue frequencies of transgenic *Drosophila* carrying *msl3-TAP* point mutations

Construct	Non-Sb males (rescued)	Sb males	Rescue frequency	*p* value
			%	
msl3-TAP WT	456	544	83.8	0.0542
Y31A	103	323	31.9	1.069e-14
Gly^{21}-Pro^{22}-Ala	349	313	111	0.350

Effect of MSL3 Methyllysine Binding Pocket Mutations on Histone Tail Binding in Vitro and on Male Survival in D. melanogaster—To test the possible contribution of Glu^{21} and Pro^{22} of MSL3 to binding mono- or dimethyllysine at H4K20, we simultaneously mutated both residues to alanine in hMSL3. As measured by SPR, the binding of the EP to AA mutant to $H4K20Me_1$ was somewhat weaker than for wild type hMSL3 (K_d = 59 μM) (Fig. 3*D*), suggesting that these residues make a modest contribution to the binding of $H4K20Me_1$. We followed up on this observation by making the same point mutations in a TAP-tagged version of the *Drosophila msl3* gene (20, 28) and observed the male *versus* female survival ratios in homozygous offspring carrying a stably integrated mutated *msl3-TAP* gene (Table 3). The EPAA double mutant had essentially wild type male survival rates, again suggesting that Glu^{21} and Pro^{22} in MSL3 play a minimal role in dosage compensation or targeting of MSL3 to a particular histone modification on the *Drosophila* male X-chromosome. However, the overall structure of the MSL3 β_1-β_2 loop or other nearby residues may also contribute to the observed $H4K20Me_1$ *in vitro* binding preference.

To test the contribution of residue Tyr^{31} to methyllysine binding in hMSL3, we made a Tyr^{31} to Ala mutation. The *in vitro* binding of the Y31A mutant to $H4K20Me_1$ was significantly weaker than the wild-type chromo-barrel domain, suggesting a larger contribution to methyllysine binding (Fig. 3; K_d = 78 μM). Male flies homozygous for the Y31A mutant did not survive as well as females (30.9% male survival; Table 3), suggesting that Tyr^{31} does play a role in dosage compensation, consistent with its contribution to $H4K20Me_1$ binding *in vitro*. In a previously published study of *msl3* chromo-barrel domain mutants, the Tyr^{31} residue was mutated to alanine along with two adjacent residues (Leu^{30} and Thr^{32}) (20). However, the triple amino acid mutant had a 0% rescue frequency and was unstable. The resultant mutant dMSL3-LYTA protein could not be detected by Western blots of heterozygous LYT30A cell extracts (20). The instability of MSL3 produced by this triple mutation deemed it uninformative in terms of specific dosage compensation function analysis. As the mutation of Leu^{30} is likely to disrupt the hydrophobic core of the dMSL3 CBD, we chose to test a less drastic mutation. In our Y31A transgenic line only 30.9% of *msl3* null homozygous males are rescued by the mutant construct. Furthermore, we could easily purify the corresponding hMSL3 Y31A mutant CBD and conduct SPR binding studies, suggest-

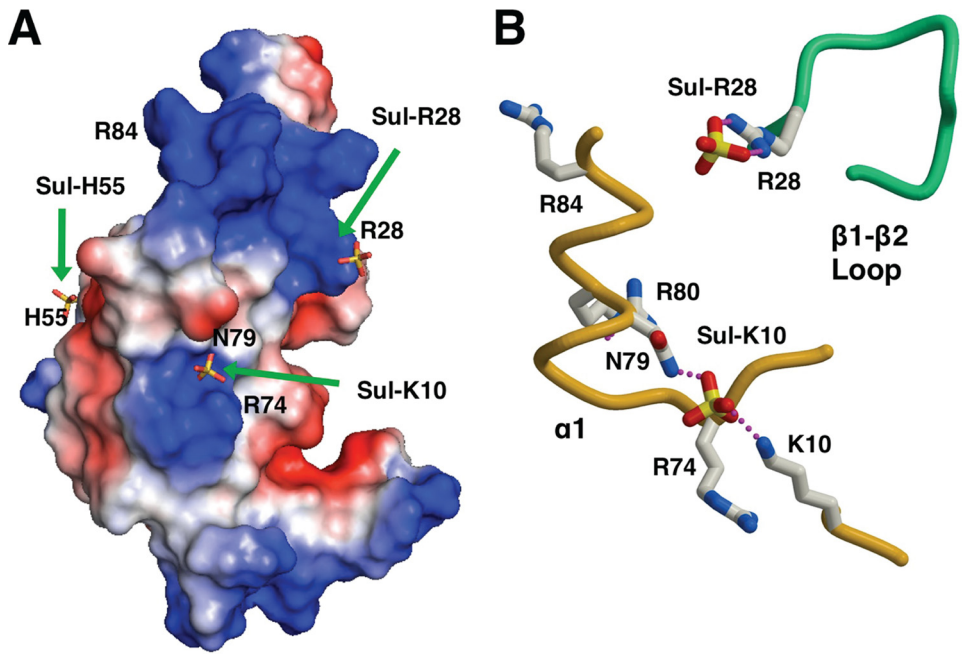

FIGURE 5. **An electropositive surface on the hMSL3 CBD binds sulfate anions.** *A*, the hMSL3 CBD is rendered as a opaque qualitative electrostatic surface (*red* negatively charged, *blue* positively charged, see "Experimental Procedures"). Sulfate anions bound at Lys^{10}-Asn^{79}, His^{55}, and Arg^{28} are shown as stick models (sulfur, *yellow*; oxygen, *red*). *Green arrows* point to regions that may interact with nucleic acids. *B*, ball and stick representation of two sulfate anions bound by conserved residues near Arg^{28} and Lys^{10} on the surface represented in *A*. Backbone trace is colored as described in the legend to Fig. 1*B*.

ing that the single mutant has a much less severe effect on the integrity and folding of the MSL3 protein.

Possible Nucleic Acid Binding Surfaces on the MSL3 CBD—In addition to binding methylated histone tails, we observed that both hMSL3 and dMSL3 chromo-barrel domains bound strongly to nucleic acids when purified from bacterial cell lysates, and remained bound after glutathione-Sepharose affinity purification, an observation consistent with the recent determination of a hMSL3 CBD·DNA complex (36). We could only effectively remove contaminating nucleic acid using ion-exchange chromatography (results not shown). Previously published studies have more specifically shown that the dMSL3 chromo-barrel domain is necessary for efficient *in vitro* binding of dMSL3 to nucleosomes (18, 19). Analysis of the hMSL3 CBD electrostatic surface revealed two positively charged surface patches (Fig. 5) and three sites where sulfate anions preferentially bind in the five hMSL3 monomers (Fig. 5 and supplemental Fig. S2). These three sulfate-binding sites involve residues Lys^{10}, Arg^{28}, and His^{55}, respectively (Fig. 5). Residues Lys^{10}, His^{55}, Arg^{74}, Asn^{79}, and Arg^{84} on the hMSL3 CBD that contact sulfate anions in the crystal structure are highly conserved in both MSL3 and MRG15 sequences (Fig. 1), suggesting that the MRG15 CBDs may also bind nucleic acids or nucleosome core particles in a similar manner to MSL3 (supplemental Fig. S3). We also note that the sulfate ion bound consistently at His^{55} in the five subunits of our structure overlaps almost perfectly with the position of a phosphate group of bound duplex DNA in the recently published hMSL3 CBD·DNA structure (36). Hence the observed sulfate binding sites in our hMSL3 CBD structure could represent binding sites for the phosphodiester backbone of nucleic acids (Fig. 5 and supplemental Fig. S2). Based on the conservation of amino acid residues in the vicinity of these sites in MSL3 and MRG15 sequences, nucleosome recognition could be important in both MSL3 and MRG15 function in their respective HAT complexes.

MSL3 CBD Function in MSL HAT Complex Targeting and Dosage Compensation—Previous studies in yeast have suggested an interaction between the Eaf3 CBD and $H3K36Me_3$ (60, 61). However, the preferential *in vitro* association of the related MSL3 and MRG15 CBDs with methylated H4K20 sequences suggests that the binding of $H3K36Me_3$ by the individual proteins is not significant in metazoans. The *in vitro* binding of $H4K20Me_1$ or $H4K20Me_2$ over $H3K36Me_2$ or other histone tail modifications associated with gene activation presented in this study strongly points to the N terminus of histone H4 being the likely *in vivo* binding target for the MSL3 and MRG15 CBDs. Supporting a role for $H4K20Me_1$ in the recruitment of the MSL complex, $H4K20Me_1$ is an abundant modification on histone tails in *Drosophila* and humans, and its presence correlates positively with the level of transcriptional activity along human genes, at least as well as for $H3K36Me_3$ (59, 62–67). Furthermore, other studies have shown that $H4K20Me_1$ is found to be preferentially deposited on nucleosomes residing within exons (68, 69), a pattern mirrored in the X-chromosome binding of the *Drosophila* MSL complex (30). However, there is no published polytene chromosome immunofluorescence evidence for $H4K20Me_1$ being enriched at the sites of *Drosophila* MSL complex binding (70, 71). Consistent with a role for the dMSL3 CBD in histone tail recognition *in vivo*, we see dosage compensation effects in a point mutation of the dMSL3 CBD (Y31A) that weakens $H4K20Me_1$ binding *in vitro*. In addition, the related MRG15 CBD prefers binding methylated H4K20, albeit the trimethyl modification, suggesting that recognition of H4K20 methylation

marks is a conserved feature of the MSL3/MRG15 gene family in higher eukaryotes.

As a primary function of the MSL complex is to acetylate lysine 16 on histone H4, it is not entirely surprising that the MSL3 CBD binds preferentially to the two most abundant H4 tail modifications found *in vivo* in humans and *Drosophila* (59, 62, 63). It also seems plausible that the binding of histone H4 tails by the MSL3 chromo-barrel domain may play a role in presenting H4 tails for acetylation by MOF within the MSL complex. Given we have shown that the dMSL3 and hMSL3 CBDs bind preferentially to $H4K20Me_1$ *in vitro*, MSL3 may help to target the MSL complex to specific chromatin regions at least in part by binding to $H4K20Me_1$.

Similarly, the demonstrated *in vitro* binding of $H4K20Me_3$ to the MRG15 component of the Tip60 complex may assist the targeting of Tip60 to regions of heterochromatin, as $H4K20Me_3$ is highly enriched in human and *Drosophila* heterochromatin (61, 70, 71) and Tip60 prefers binding to heterochromatic regions enriched in $H3K9Me_3$ (and presumably $H4K20Me_3$) during double-stranded DNA repair (13). Importantly, Tip60 and MRG15 are associated with gene repression functions in *Drosophila* and are classified as part of the Polycomb group (12). Work presented in this study suggests that the MSL complex may be recruited to active genes in part by MSL3 binding to $H4K20Me_1$, whereas the repressive Tip60 HAT complex may be targeted to heterochromatin by MRG15 being recruited to sites enriched in $H4K20Me_3$. Hence, chromo-barrel domain containing HAT subunits appear capable of exquisitely discriminating the methylation status of lysine 20 on histone H4 and the corresponding biological context.

Acknowledgments—We thank Bradley McLellan, Massey University, New Zealand, for cloning the dMSL3 chromo-barrel domain fragment; George Wong for purification conditions for the dMSL3 chromo-barrel domain; and Andrew Welham who carried out the initial characterization of the dMSL3 and hMSL3 chromo-barrel domains. Ewa Kerc contributed to the cloning of the hMSL3 chromo-barrel domain. Robert Shearer cloned the D. melanogaster MRG15 chromo-barrel domain and conducted the initial purification studies. We thank Edwin Smith for generously providing cDNAs for the hMSL proteins, and Prof. Mitzi Kuroda, Harvard Medical School, for providing the msl3-TAP construct for Drosophila transgenic studies. David Sanders (Dept of Chemistry, University of Saskatchewan) provided access to the Nanodrop device. Jeremy Lee graciously conducted the ethidium bromide fluorescence assays. SPR data were collected at the Saskatchewan Structural Sciences Centre, University of Saskatchewan. We thank Pawel Grochulski and James Gorin of the Canadian Light Source CMCF small-gap undulator beamline for timely access and expert assistance with data collection. The Canadian Light Source is funded by the Canadian Institutes for Health Research (CIHR), the National Science and Engineering Council of Canada (NSERC), the Canadian foundation for innovation (CFI) and the University of Saskatchewan.

REFERENCES

1. Carrozza, M. J., Utley, R. T., Workman, J. L., and Côté, J. (2003) *Trends Genet.* **19,** 321–329
2. Rea, S., Xouri, G., and Akhtar, A. (2007) *Oncogene* **26,** 5385–5394
3. Smith, E. R., Cayrou, C., Huang, R., Lane, W. S., Côté, J., and Lucchesi, J. C. (2005) *Mol. Cell. Biol.* **25,** 9175–9188
4. Taipale, M., Rea, S., Richter, K., Vilar, A., Lichter, P., Imhof, A., and Akhtar, A. (2005) *Mol. Cell. Biol.* **25,** 6798–6810
5. Gupta, A., Sharma, G. G., Young, C. S., Agarwal, M., Smith, E. R., Paull, T. T., Lucchesi, J. C., Khanna, K. K., Ludwig, T., and Pandita, T. K. (2005) *Mol. Cell. Biol.* **25,** 5292–5305
6. Sykes, S. M., Mellert, H. S., Holbert, M. A., Li, K., Marmorstein, R., Lane, W. S., and McMahon, S. B. (2006) *Mol. Cell* **24,** 841–851
7. Li, X., Wu, L., Corsa, C. A., Kunkel, S., and Dou, Y. (2009) *Mol. Cell* **36,** 290–301
8. Raja, S. J., Charapitsa, I., Conrad, T., Vaquerizas, J. M., Gebhardt, P., Holz, H., Kadlec, J., Fraterman, S., Luscombe, N. M., and Akhtar, A. (2010) *Mol. Cell* **38,** 827–841
9. Doyon, Y., Selleck, W., Lane, W. S., Tan, S., and Côté, J. (2004) *Mol. Cell. Biol.* **24,** 1884–1896
10. Squatrito, M., Gorrini, C., and Amati, B. (2006) *Trends Cell Biol.* **16,** 433–442
11. Kusch, T., Florens, L., Macdonald, W. H., Swanson, S. K., Glaser, R. L., Yates, J. R., 3rd, Abmayr, S. M., Washburn, M. P., and Workman, J. L. (2004) *Science* **306,** 2084–2087
12. Qi, D., Jin, H., Lilja, T., and Mannervik, M. (2006) *Genetics.* **174,** 241–251
13. Sun, Y., Jiang, X., Xu, Y., Ayrapetov, M. K., Moreau, L. A., Whetstine, J. R., and Price, B. D. (2009) *Nat. Cell Biol.* **11,** 1376–1382
14. Gelbart, M. E., and Kuroda, M. I. (2009) *Development* **136,** 1399–1410
15. Lucchesi, J. C. (2009) *Curr. Opin. Genet. Dev.* **19,** 550–556
16. Morales, V., Straub, T., Neumann, M. F., Mengus, G., Akhtar, A., and Becker, P. B. (2004) *EMBO J.* **23,** 2258–2268
17. Morales, V., Regnard, C., Izzo, A., Vetter, I., and Becker, P. B. (2005) *Mol. Cell. Biol.* **25,** 5947–5954
18. Buscaino, A., Legube, G., and Akhtar, A. (2006) *EMBO Rep.* **7,** 531–538
19. Larschan, E., Alekseyenko, A. A., Gortchakov, A. A., Peng, S., Li, B., Yang, P., Workman, J. L., Park, P. J., and Kuroda, M. I. (2007) *Mol. Cell* **28,** 121–133
20. Sural, T. H., Peng, S., Li, B., Workman, J. L., Park, P. J., and Kuroda, M. I. (2008) *Nat. Struct. Mol. Biol.* **15,** 1318–1325
21. Scott, M. J., and Li, F. (2008) *RNA Biol.* **5,** 13–16
22. Bone, J. R., Lavender, J., Richman, R., Palmer, M. J., Turner, B. M., and Kuroda, M. I. (1994) *Genes Dev.* **8,** 96–104
23. Gu, W., Wei, X., Pannuti, A., and Lucchesi, J. C. (2000) *EMBO J.* **19,** 5202–5211
24. Demakova, O. V., Kotlikova, I. V., Gordadze, P. R., Alekseyenko, A. A., Kuroda, M. I., and Zhimulev, I. F. (2003) *Chromosoma* **112,** 103–115
25. Copps, K., Richman, R., Lyman, L. M., Chang, K. A., Rampersad-Ammons, J., and Kuroda, M. I. (1998) *EMBO J.* **17,** 5409–5417
26. Gorman, M., Franke, A., and Baker, B. S. (1995) *Development* **121,** 463–475
27. Smith, E. R., Allis, C. D., and Lucchesi, J. C. (2001) *J. Biol. Chem.* **276,** 31483–31486
28. Alekseyenko, A. A., Larschan, E., Lai, W. R., Park, P. J., and Kuroda, M. I. (2006) *Genes Dev.* **20,** 848–857
29. Legube, G., McWeeney, S. K., Lercher, M. J., and Akhtar, A. (2006) *Genes Dev.* **20,** 871–883
30. Gilfillan, G. D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P. B. (2006) *Genes Dev.* **20,** 858–870
31. Smith, E. R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R. G., Allis, C. D., and Lucchesi, J. C. (2000) *Mol. Cell. Biol.* **20,** 312–318
32. Scott, M. J., Pan, L. L., Cleland, S. B., Knox, A. L., and Heinrich, J. (2000) *EMBO J.* **19,** 144–155
33. Koonin, E. V., Zhou, S., and Lucchesi, J. C. (1995) *Nucleic Acids Res.* **23,** 4229–4233
34. Prakash, S. K., Van den Veyver, I. B., Franco, B., Volta, M., Ballabio, A., and Zoghbi, H. Y. (1999) *Genomics* **59,** 77–84
35. Bertram, M. J., and Pereira-Smith, O. M. (2001) *Gene* **266,** 111–121
36. Kim, D., Blus, B. J., Chandra, V., Huang, P., Rastinejad, F., and Khorasanizadeh, S. (2010) *Nat. Struct. Mol. Biol.* **17,** 1027–1029
37. Morgan, A. R., Lee, J. S., Pulleyblank, D. E., Murray, N. L., and Evans,

D. H. (1979) *Nucleic Acids Res.* **7,** 547–569
38. Wetlaufer, D. B. (1962) *Adv. Protein Chem.* **17,** 303–391
39. Edelhoch, H. (1967) *Biochemistry* **6,** 1948–1954
40. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
41. Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q., and Ding, J. (2006) *Nucleic Acids Res.* **34,** 6621–6628
42. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr.* **40,** 658–674
43. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
44. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
45. Collaborative Computational Project, No. 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
46. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
47. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr. D Biol. Crystallogr.* **57,** 122–133
48. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
49. Notredame, C., Higgins, D. G., and Heringa, J. (2000) *J. Mol. Biol.* **302,** 205–217
50. Gouet, P., Courcelle, E., Stuart, D. I., and Métoz, F. (1999) *Bioinformatics* **15,** 305–308
51. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
52. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
53. Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., and Mer, G. (2006) *Cell* **127,** 1361–1373
54. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007) *Nat. Struct. Mol. Biol.* **14,** 1025–1040
55. Li, H., Fischle, W., Wang, W., Duncan, E. M., Liang, L., Murakami-Ishibe, S., Allis, C. D., and Patel, D. J. (2007) *Mol. Cell* **28,** 677–691
56. Sun, B., Hong, J., Zhang, P., Dong, X., Shen, X., Lin, D., and Ding, J. (2008) *J. Biol. Chem.* **283,** 36504–36512
57. Xu, C., Cui, G., Botuyan, M. V., and Mer, G. (2008) *Structure* **16,** 1740–1750
58. Wang, W. K., Tereshko, V., Boccuni, P., MacGrogan, D., Nimer, S. D., and Patel, D. J. (2003) *Structure* **11,** 775–789
59. Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) *Cell* **129,** 823–837
60. Joshi, A. A., and Struhl, K. (2005) *Mol. Cell* **20,** 971–978
61. Li, B., Gogol, M., Carey, M., Lee, D., Seidel, C., and Workman, J. L. (2007) *Science* **316,** 1050–1054
62. Yang, H., and Mizzen, C. A. (2009) *Biochem. Cell Biol.* **87,** 151–161
63. Pesavento, J. J., Yang, H., Kelleher, N. L., and Mizzen, C. A. (2008) *Mol. Cell. Biol.* **28,** 468–486
64. Talasz, H., Lindner, H. H., Sarg, B., and Helliger, W. (2005) *J. Biol. Chem.* **280,** 38814–38822
65. Vakoc, C. R., Sachdeva, M. M., Wang, H., and Blobel, G. A. (2006) *Mol. Cell. Biol.* **26,** 9185–9195
66. Yang, H., Pesavento, J. J., Starnes, T. W., Cryderman, D. E., Wallrath, L. L., Kelleher, N. L., and Mizzen, C. A. (2008) *J. Biol. Chem.* **283,** 12085–12092
67. Sakaguchi, A., Karachentsev, D., Seth-Pasricha, M., Druzhinina, M., and Steward, R. (2008) *Genetics* **179,** 317–322
68. Schwartz, S., Meshorer, E., and Ast, G. (2009) *Nat. Struct. Mol. Biol.* **16,** 990–995
69. Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., and Guigó, R. (2009) *Nat. Struct. Mol. Biol.* **16,** 996–1001
70. Nishioka, K., Rice, J. C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J. T., Allis, C. D., and Reinberg, D. (2002) *Mol. Cell* **9,** 1201–1213
71. Ebert, A., Lein, S., Schotta, G., and Reuter, G. (2006) *Chromosome Res.* **14,** 377–392