Structural and biochemical studies on the chromo-barrel domain of male specific lethal (MSL3) reveal a binding preference for mono or dimethyl lysine 20 on histone H4

Stanley A. Moore, Yurdagul Ferhatoglu, Yunhua Jia, Rami A. Al-Jiab, Maxwell J. Scott

Supplemental Methods

Calculation of the protein UV absorbance ratio A_{280}/A_{260}

Extinction coefficients for tryptophan and tyrosine at 280 nm were the recommended average values for proteins published by Pace *et al.* 1995. They are 5500 L mol^{-1} cm^{-1} for Trp and 1490 L mol^{-1} cm^{-1} for Tyr. These are best values based on the UV absorption of N-acetyl-Trp-amide and N-acetyl-Tyr ethyl ester determined by Edelhoch (Edelhoch, 1967). For the extinction values of Trp and Tyr at 260 nm, we therefore measured the extinction values at 260 nm directly from the graphs published for the above-mentioned compounds (Edelhoch, 1967). They are 3300 L mol^{-1} cm^{-1} for N-acetyl-Trp-amide, and 600 L mol^{-1} cm^{-1} for N-acetyl-Tyr ethyl ester (Edelhoch, 1967). Phe does not absorb appreciably at 280 nm, but does absorb significantly at 260 nm (Wetlaufer, 1962). Therefore we used an extinction of Phe at 260 nm (145 L mol^{-1} cm^{-1}) read directly from the published UV absorption spectrum (Wetlaufer, 1962). There are no disulfides in any of the chromo-barrel domains so a correction for cystine was not applied. The dMSL3 chromo-barrel domain contains 4 Phe, 5 Tyr, and 1 Trp residues. The hMSL3 chromo-barrel domain contains 4 Phe, 2 Tyr, and 3 Trp residues. The dMRG15 chromo-barrel domain contains 2 Phe, 6 Tyr, and 3 Trp residues.

Table S1. List of DNA primers used in this study.

Primer sequence (5'to 3') [a]	**Description**
GACCTG**GGATCC**ACGGAGCTAAGGGACGAGACA	dMSL3 CBD forward
GACGA**GGAATTC**CTACTTGGCAGCTTCGGCCAG	dMSL3 CBD reverse
GACCTG**GGATCC**AGCGCGAGCGAGGGCATGAAATTT	hMSL3 CBD forward
GACGAG**GAATTC**CTAGCGAGCTACAGCTTTTCTTGCCAA	hMSL3 CBD reverse
GACCTG**GGATCC**GGAGAAGTAAAACCCGCGAAAG	dMRG15 CBD forward
GACGAC**GAATTC**CTACTGGCGAGCCAGTTCCTG	dMRG15 CBD reverse
GGAGAAAGTGCTGTGCTTC*GCAGCT*GACCCCACCAAGGCGCG	hMSL3 CBD EPAA forward
CGCGCCTTGGTGGGGTC*AGCTGC*GAAGCACAGCACTTTCTCC	hMSL3 CBD EPAA reverse
CCCACCAAGGCGCGAGTGCTG*GCC*GATGCCAAGATTGTTGTAG	hMSL3 CBD Y31A forward
CTACAACAATCTTGGCATC*GGC*CAGCACTCGCGCCTTGGTGGG	hMSL3 CBD Y31A reverse
GCGAGATAGTCCTTTGCTAC*GCCGCT*GATAAATCGAAGGCCCGTG	dMSL3 CBD EPAA forward
CACGGGCCTTCGATTTATC*AGCGGC*GTAGCAAAGGACTATCTCGC	dMSL3 CBD EPAA reverse
TCGAAGGCCCGTGTTCTT*GCC*ACCAGCAAGGTGCTAAACG	dMSL3 CBD Y31A forward
CGTTTAGCACCTTGCTGGT*GGC*AAGAACACGGGCCTTCGA	dMSL3 CBD Y31A reverse

[a] Restriction sites are highlighted in bold, sites of introduced mutations in italic.

SUPPLEMENTAL FIGURES

Figure S1. Ramachandran plot of the refined hMSL3 atomic structure. Drawn with Procheck (Laskowski et al., 1993; CCP4, 1994).

Figure S2. Structure of the hMSL3 chromo-barrel domain. (A) Backbone trace of the five hMSL3 monomers in the asymmetric unit of the crystal. Monomers are colored gold (subunit A), green (subunit B), magenta (subunit C), blue (subunit D) and dark orange (subunit E). The MSL3-specific loop between strands β1 and β2 is colored blue-green. Sulfate anions and CHES molecules are shown as stick models (oxygens red). Subunit D has a CHES molecule bound on the surface of the C-terminal helix. (B) Ribbon diagram of a hMSL3 dimer showing side chains on the β1-β2 loop. Residues making up the methyllysine pocket or making contacts across the dimer interface are drawn as stick models. Bound CHES molecules are drawn in green. (C) 2Fobs -DFcalc electron density map (drawn at 1.2 sigma) of the methyllysine binding pocket (final refined model) and bound CHES for subunit A in the hMSL3 CBD structure. (D) Structural superposition of the five independent copies of the hMSL3 chromo-barrel domain. Each chain is shown as a differently colored Cα trace, colored as in part (A). Bound CHES molecules and sulfates bound at the common sites in each monomer are shown as stick models and labeled according to the description in the text. Parts (A), (B) and (D) drawn with Molscript/Raster3D (Kraulis, 1991; Merrit and Bacon, 1997). Part (C) drawn with Coot (Emsley and Cowtan, 2004).

Figure S3. Comparison of the structures of the hMSL3 and hMRG15 CBDs. (A) The hMSL3 CBD is rendered as opaque qualitative electrostatic molecular surface (red negatively

charged, blue, positively charged). The green arrow points to the region containing conserved residues Lys10 and Asn79 and Arg80 in hMSL3 and binds a sulfate anion in the structure. The view is similar to that depicted in Figure 5A of the main article, with the C-terminal helix in the upper left quadrant, and the β2-β3 hairpin at the lower right. The exit from the methyllysine binding pocket is on the back side of the molecule. (B) The same orientation as (A), but showing the hMRG15 electrostatic surface. (C) Superposition of the structures of the hMSL3 CBD (gold trace, beige side chains shown, glu21 white), the hMRG15 CBD (magenta trace, pink side chains) (Zhang et al., 2006) and the 53BP1 Tudor domain 1 (blue trace and light blue side chains) (Botuyan et al. 2006). The side chains of the residues making up the methyllsyine binding pockets are shown as stick models. (A) and (B) Drawn with Pymol (www.pymol.org). (C) Drawn with Molscript/Raster3d (Kraulis, 1991; Merrit and Bacon, 1997).

Figure S4. Single SPR injections of methyllysine containing histone tail peptides over the immobilized dMSL3, hMSL3 and dMRG15 chromo-barrel domains. (A), (B) Surface plasmon resonance single injection steady state response plots for the immobilized dMSL3 CBD and a series of histone tail peptides at 100 μM concentration, in 150 mM NaCl, 100 mM HEPES, pH 7.5, 3 mM EDTA. See Experimental procedures for details of the SPR experiments and the peptide sequences used. (C), (D) Same as A,B, but for hMSL3 CBD. (E), (F), Same as A, B, but for dMRG15 CBD.

REFERENCES

Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., and Mer, G. (2006). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. Cell *127*, 1361-1373.

Collaborative Computational Project, No. 4 (1994). Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Edelhoch, H. (1967). Spectroscopic determination of Tryptophan and Tyrosine in proteins. Biochemistry *6*, 1948-1954.

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

Kraulis, P. (1991). Molscript: a program to produce both detailed and schematic plots of protein structures J. App. Cryst. *24*, 946-950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK - a

program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283-291.

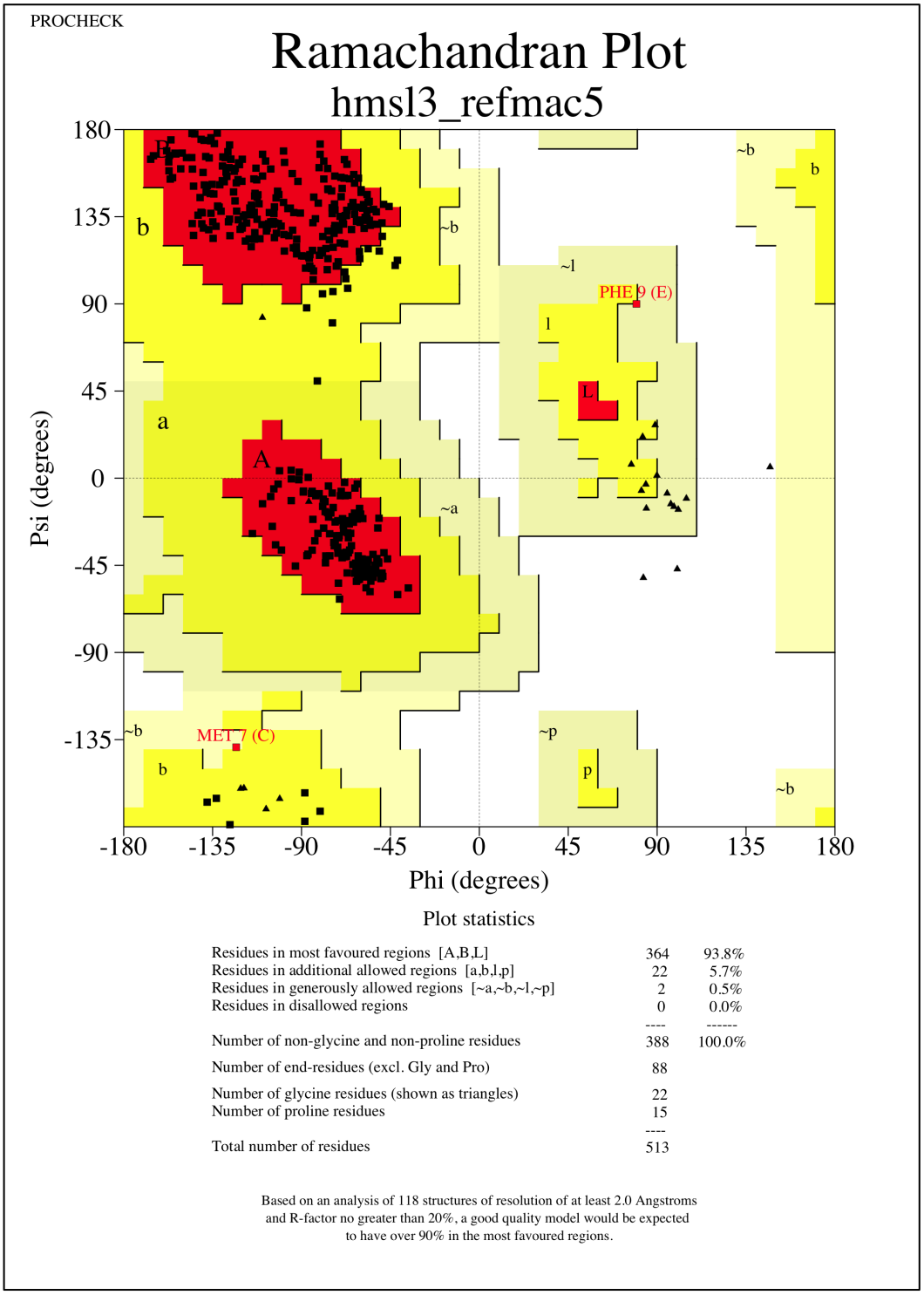
Merritt, E. and Bacon, D. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505-524.

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Prot. Sci. *4*, 2411-2423.

Wetlaufer, D. B. (1962). Ultra violet spectra of proteins and amino acids. Adv. Prot. Chem. *17*, 303-391.

Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q., and Ding, J. (2006). Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. Nucleic Acids Res. *34*, 6621-6628.

Figure S1



Plot statistics

Residues in most favoured regions [A,B,L]	364	93.8%
Residues in additional allowed regions [a,b,l,p]	22	5.7%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.5%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	388	100.0%
Number of end-residues (excl. Gly and Pro)	88	
Number of glycine residues (shown as triangles)	22	
Number of proline residues	15	
Total number of residues	513	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

hmsl3_refmac5_01.ps

Figure S2

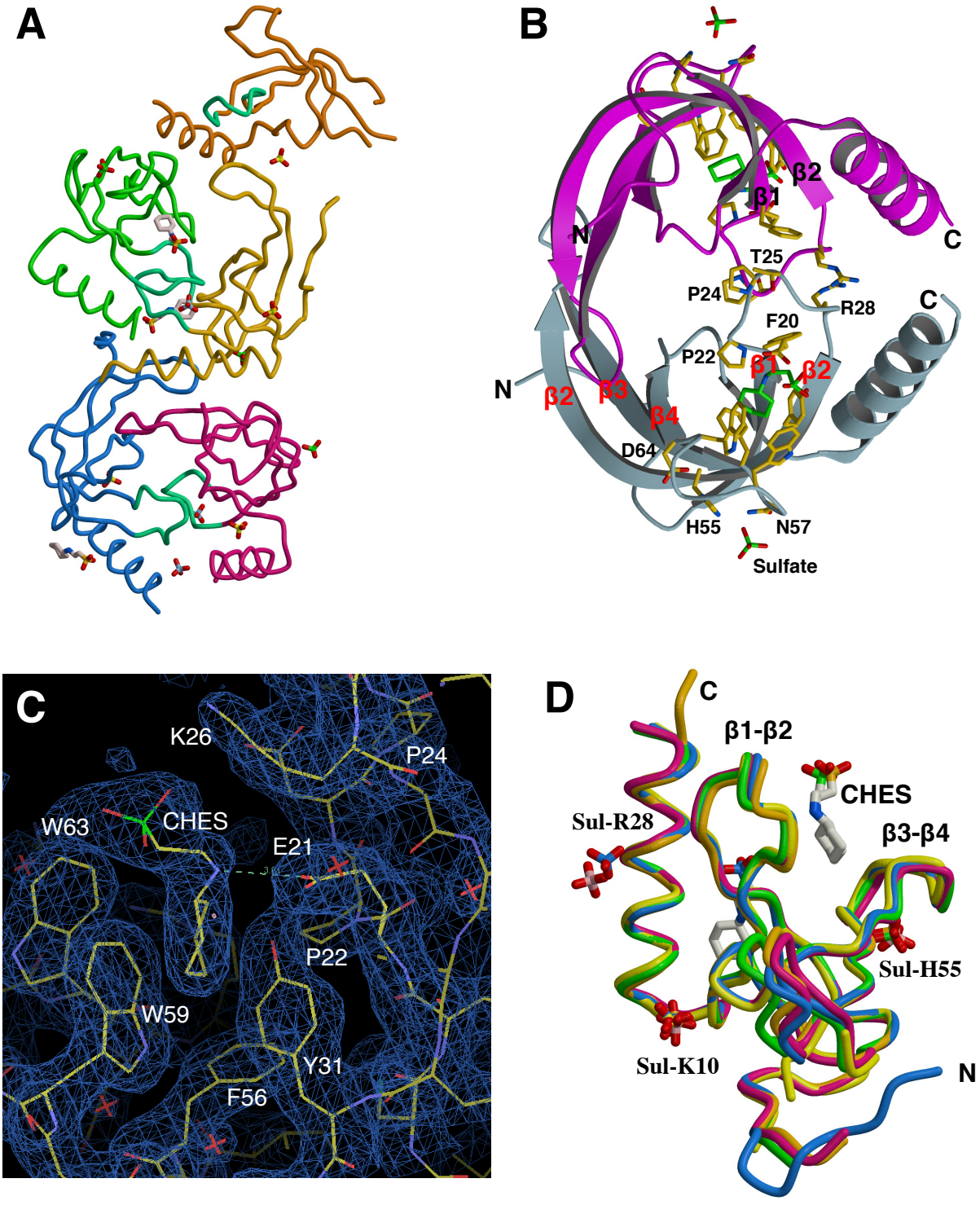
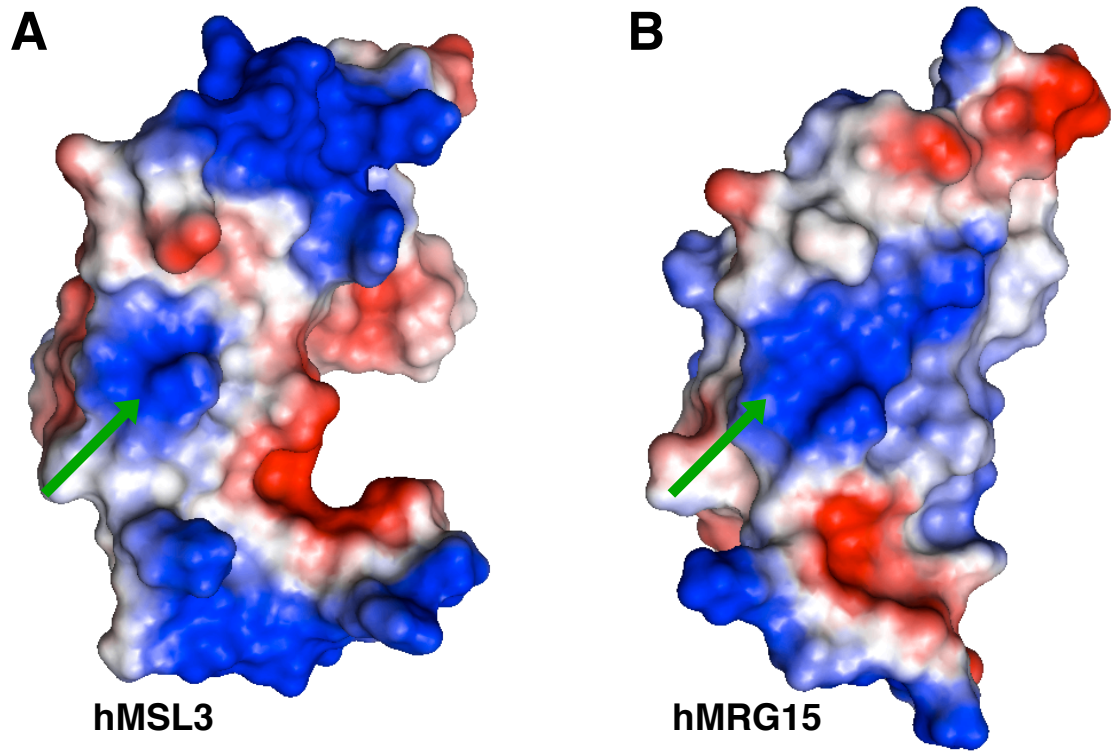

A
B
β2
β1
N
C
T25
P24
R28
C
F20
P22
β1
β2
N
β2
β3
β4
D64
H55
N57
Sulfate
C
K26
P24
W63
CHES
E21
P22
W59
Y31
F56
D
C
β1-β2
CHES
Sul-R28
β3-β4
Sul-H55
Sul-K10
N

Figure S3



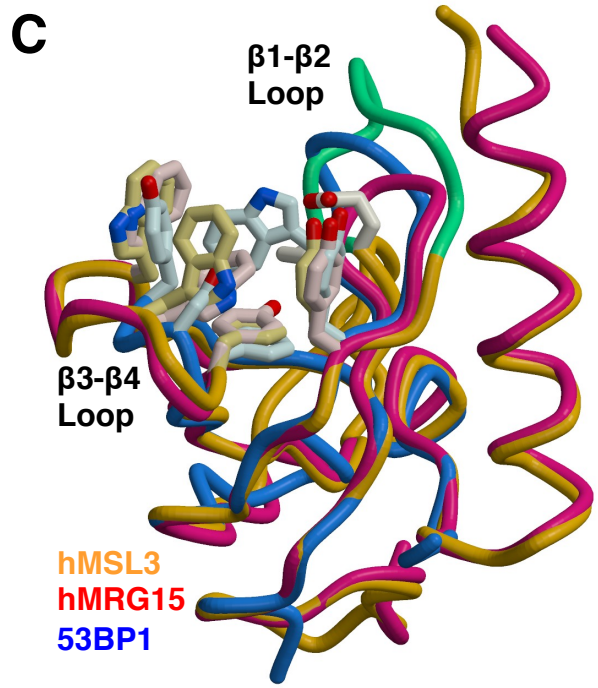
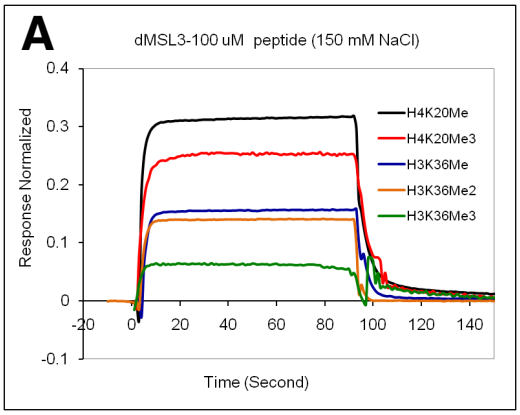
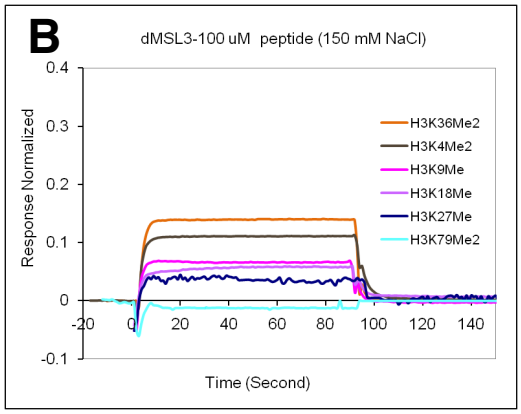
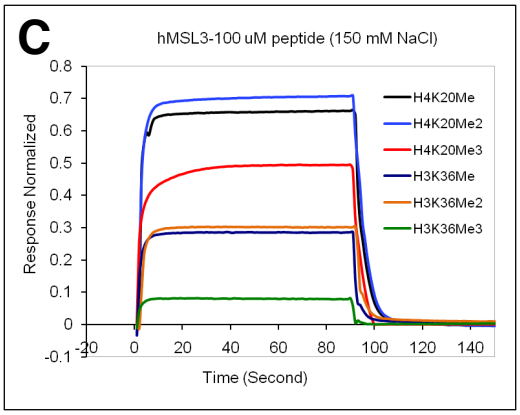
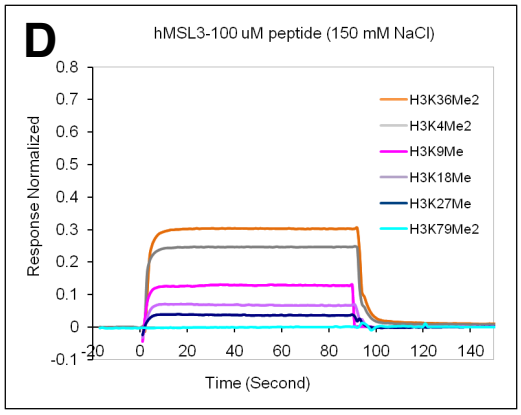


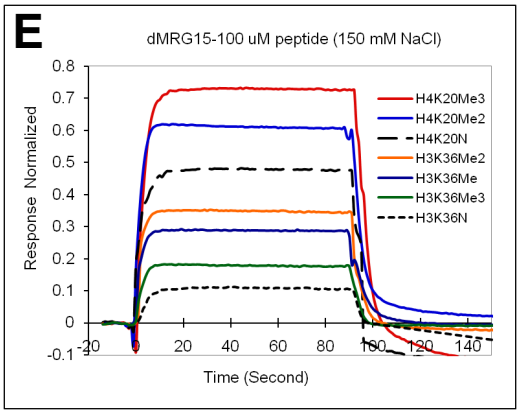
Figure S4

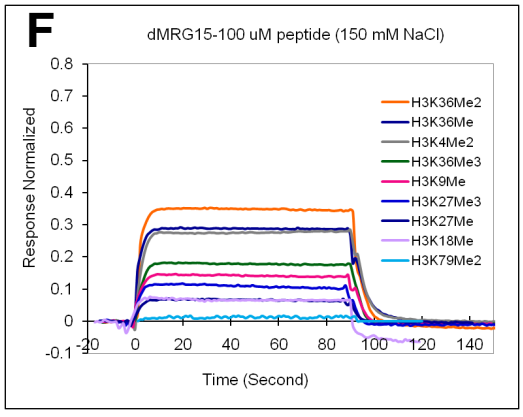

A
dMSL3-100 uM peptide (150 mM NaCl)
Response Normalized
Time (Second)
H4K20Me
H4K20Me3
H3K36Me
H3K36Me2
H3K36Me3


B
dMSL3-100 uM peptide (150 mM NaCl)
Response Normalized
Time (Second)
H3K36Me2
H3K4Me2
H3K9Me
H3K18Me
H3K27Me
H3K79Me2


C
hMSL3-100 uM peptide (150 mM NaCl)
Response Normalized
Time (Second)
H4K20Me
H4K20Me2
H4K20Me3
H3K36Me
H3K36Me2
H3K36Me3


D
hMSL3-100 uM peptide (150 mM NaCl)
Response Normalized
Time (Second)
H3K36Me2
H3K4Me2
H3K9Me
H3K18Me
H3K27Me
H3K79Me2


E
dMRG15-100 uM peptide (150 mM NaCl)
Response Normalized
Time (Second)
H4K20Me3
H4K20Me2
H4K20N
H3K36Me2
H3K36Me
H3K36Me3
H3K36N


F
dMRG15-100 uM peptide (150 mM NaCl)
Response Normalized
Time (Second)
H3K36Me2
H3K36Me
H3K4Me2
H3K36Me3
H3K9Me
H3K27Me3
H3K27Me
H3K18Me
H3K79Me2