Identification of secondary binding sites on protein surfaces for rational elaboration of synthetic protein mimics

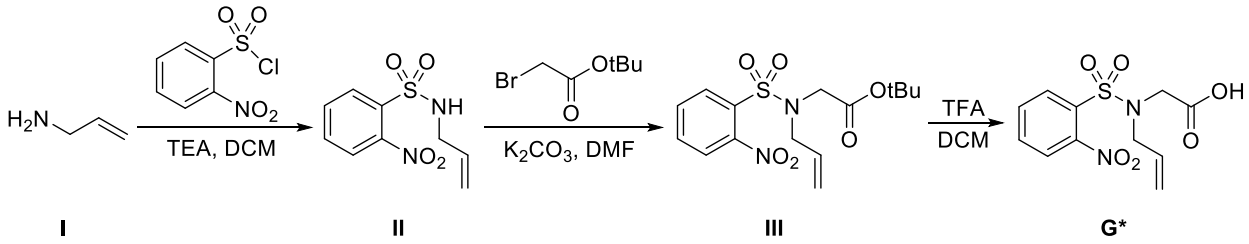
Justin M. Torner, Yuwei Yang, David Rooklin, Yingkai Zhang* and Paramjit S. Arora*

Department of Chemistry, New York University, 100 Washington Square East, New York, NY 10003

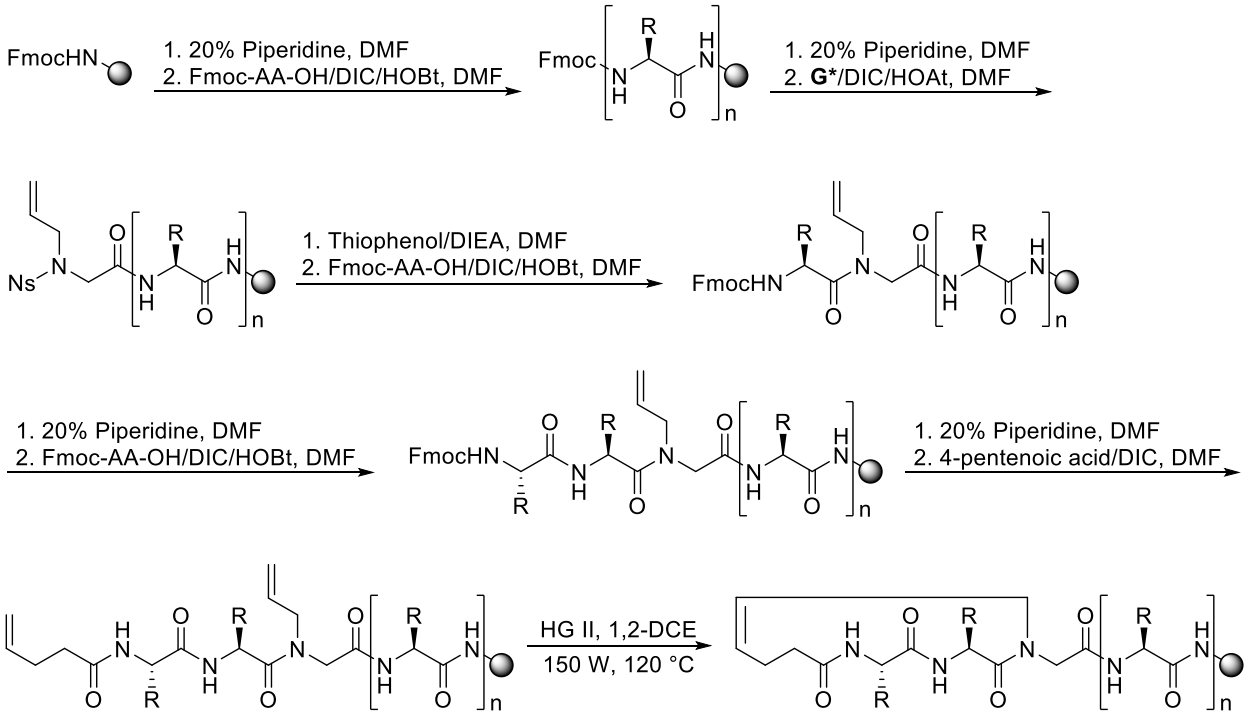
Email: arora@nyu.edu or yingkai.zhang@nyu.edu

Supporting Information

Supporting Figures



Scheme S1 Synthesis of **G***



Scheme S2 Solid phase synthesis of HBS peptides

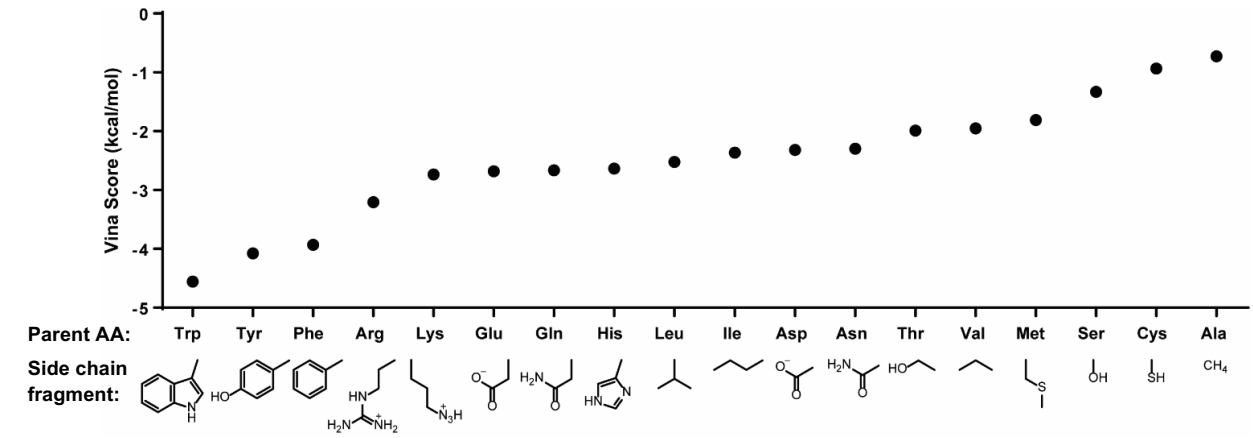


Figure S1 Autodock Vina scores of amino acid side chain fragments docked into the MdmX secondary binding site.

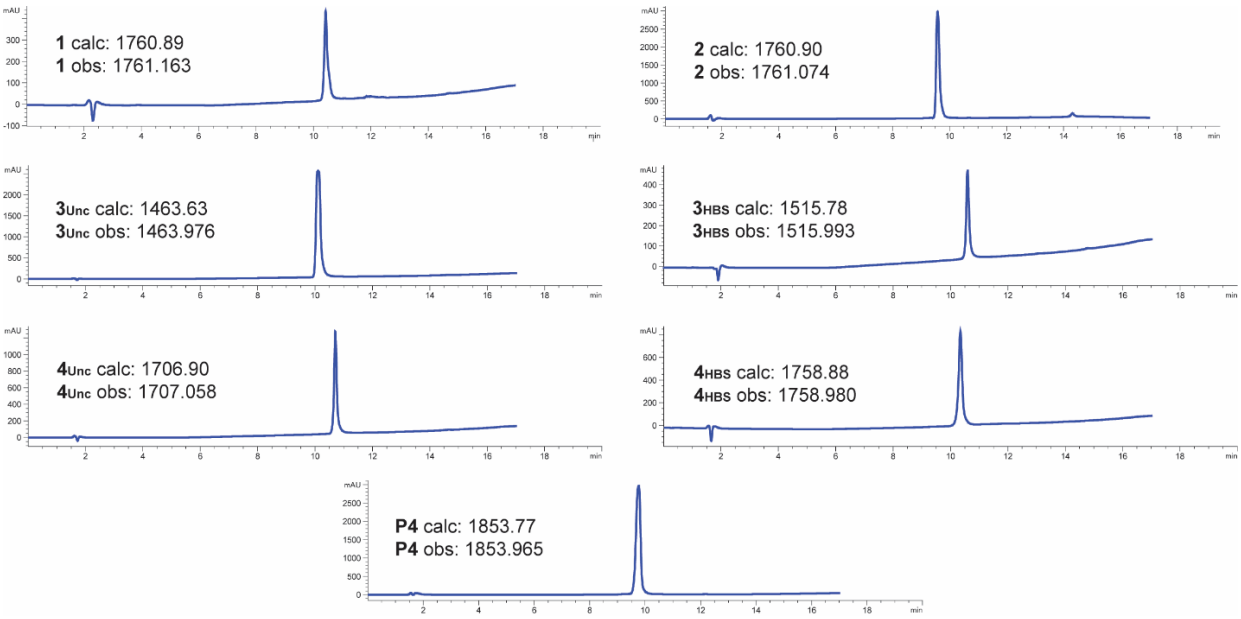


Figure S2 Peptide characterization. HPLC traces and MALDI-TOF m/z data are shown.

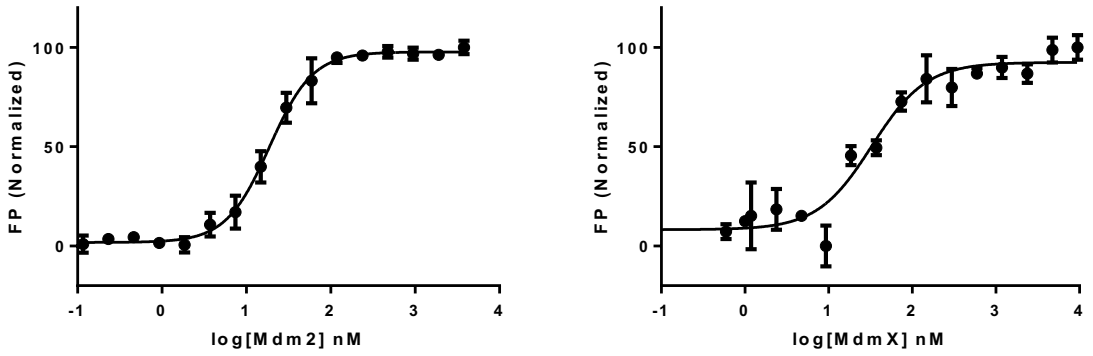
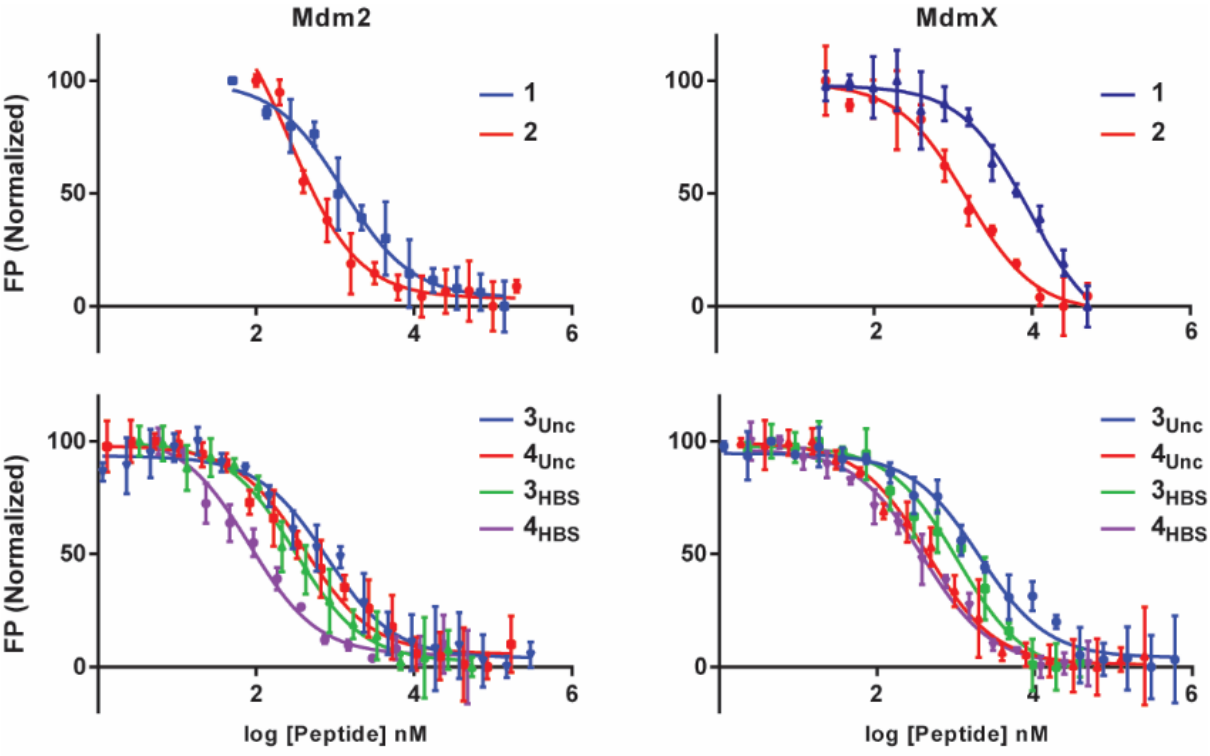


Figure S3 Direct binding of probe **P4** to Mdm2 (left) and MdmX (right)



Peptide		16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	K_i, Mdm2 (nM)	K_i, MdmX (nM)
1	Ac	Q	E	T	F	S	D	L	W	K	L	L	P	E	N				1049 ± 476	1420 ± 425
2	Ac	Q	E	T	F	S	D	L	W	K	L	L	P	G	W				355 ± 116	106 ± 51
3Unc	Ac	Q	E	G	F	S	D	L	W	K	L	L	S						271 ± 80	205 ± 61
4Unc	Ac	Q	E	G	F	S	D	L	W	K	L	L	S	G	W				157 ± 47	87 ± 21
3HBS	X	Q	E	G*	F	S	D	L	W	K	L	L	S						112 ± 46	123 ± 34
4HBS	X	Q	E	G*	F	S	D	L	W	K	L	L	S	G	W				4 ± 8	50 ± 16

Figure S4 Competitive fluorescence polarization binding curves with data summarized in table. X = 4-pentenoic acid. G* = *o*-nosyl-N-allylglycine. Line between X and G* represents HBS macrocyclization. All peptides are C-terminally amidated.

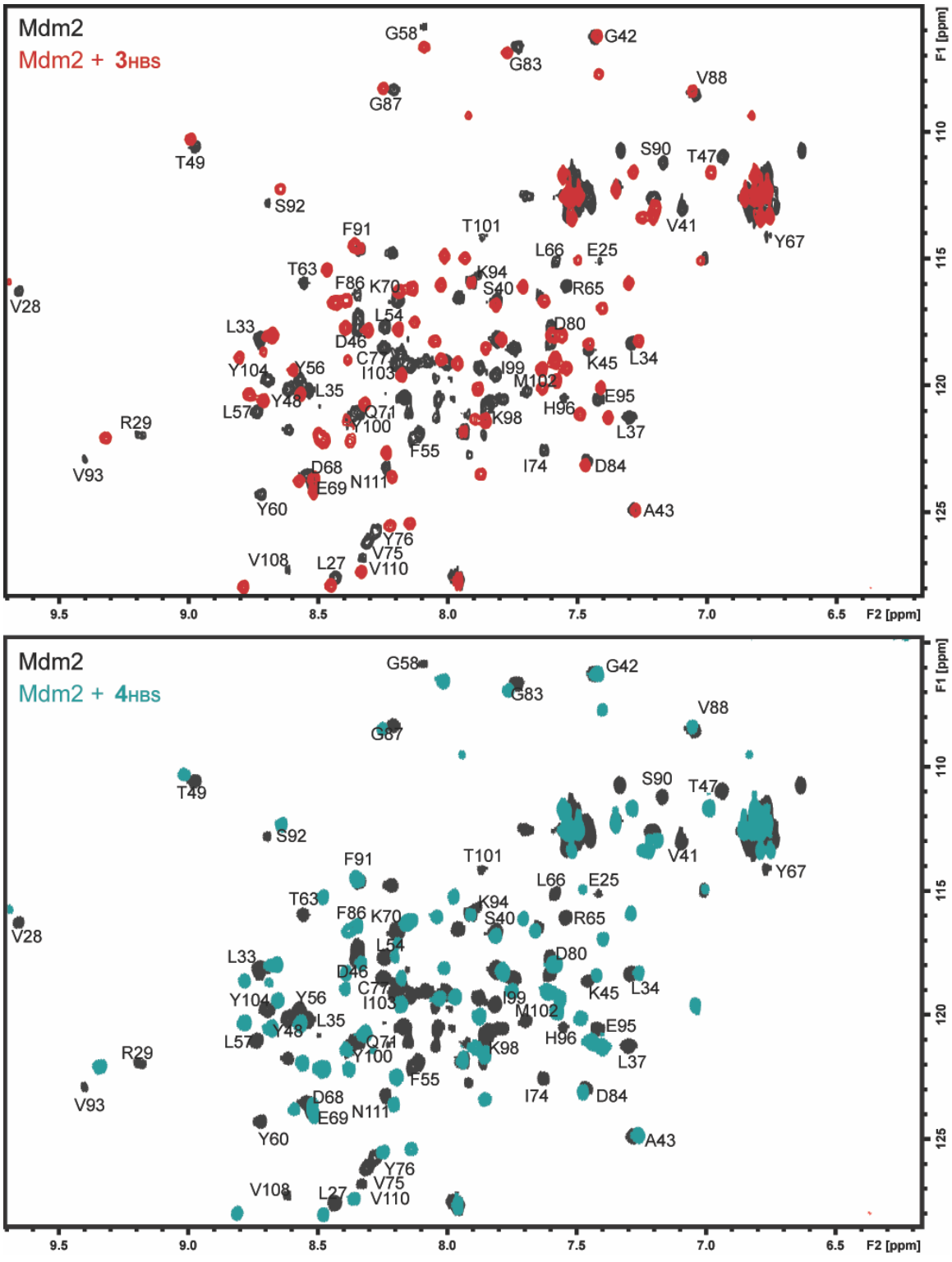


Figure S5 ^{15}N-^{1}H HSQC NMR spectra of ^{15}N-labelled Mdm2 apo (black), in complex with $\mathbf{3}_{HBS}$ (red), and in complex with $\mathbf{4}_{HBS}$ (teal).

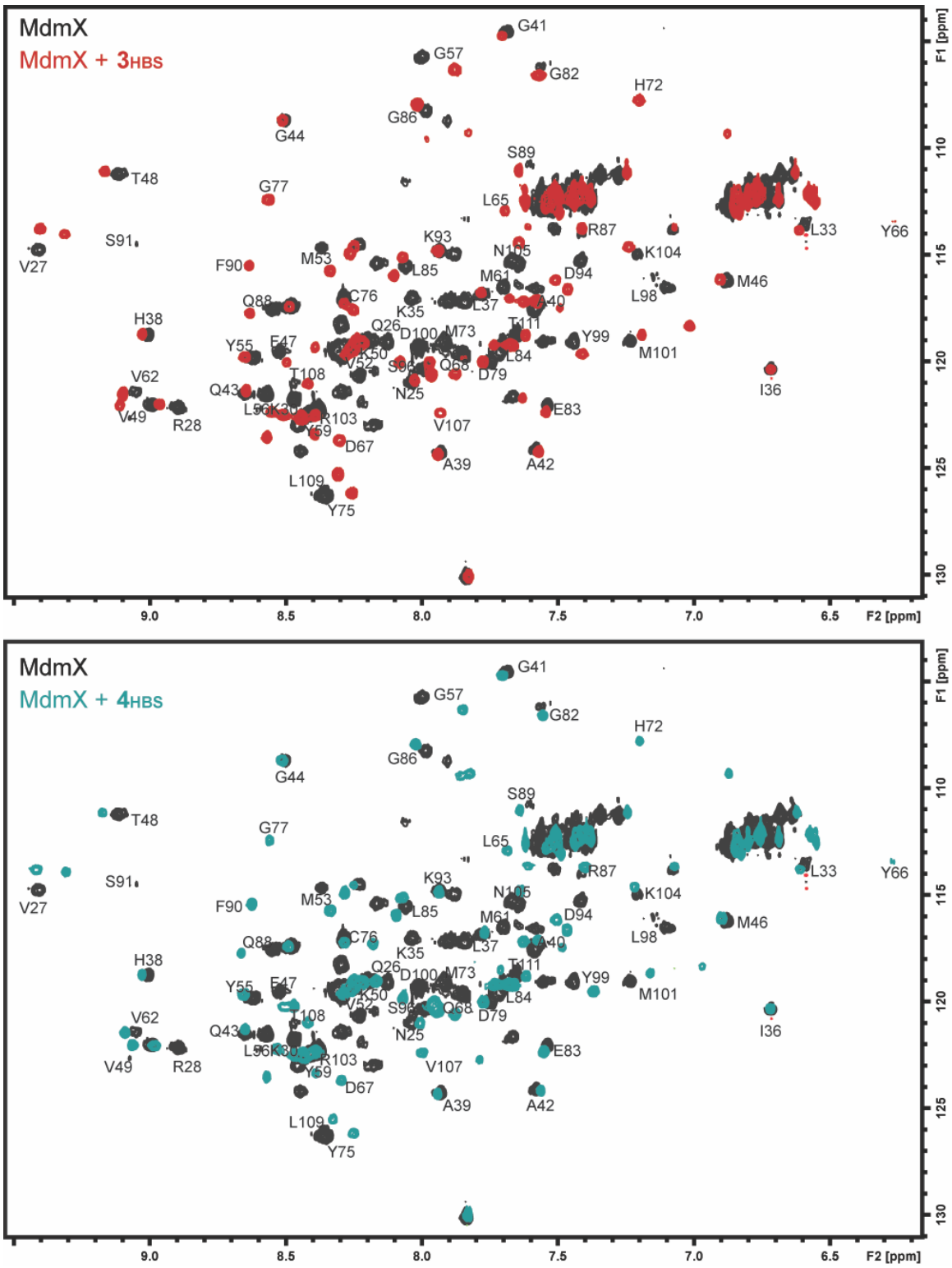


Figure S6 ^{15}N-^{1}H HSQC NMR spectra of ^{15}N-labelled MdmX apo (black), in complex with $\mathbf{3}_{\mathbf{HBS}}$ (red), and in complex with $\mathbf{4}_{\mathbf{HBS}}$ (teal).

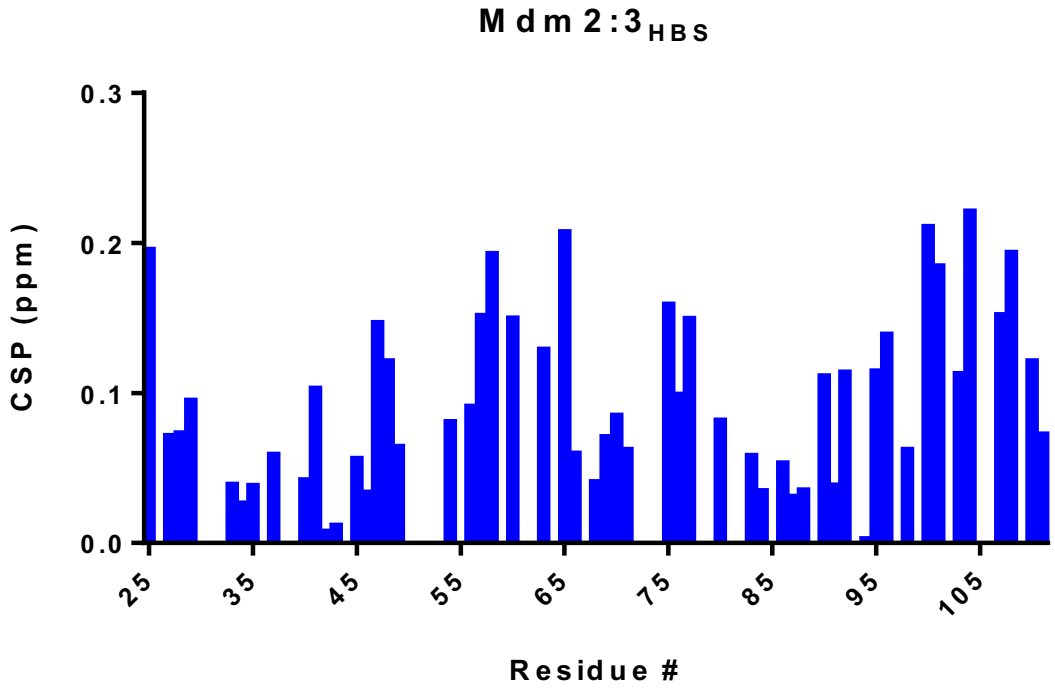


Figure S7 ^{15}N-^{1}H HSQC NMR chemical shift perturbations of Mdm2 upon addition of $\mathbf{3}_{\mathbf{HBS}}$

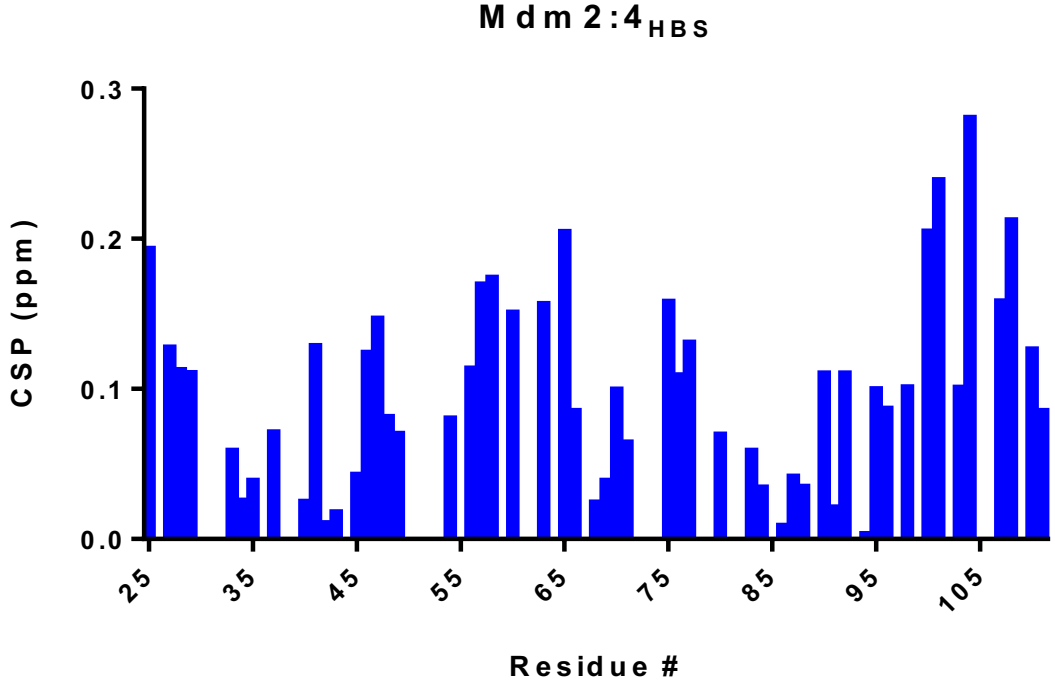


Figure S8 ^{15}N-^{1}H HSQC NMR chemical shift perturbations of Mdm2 upon addition of $\mathbf{4}_{\mathbf{HBS}}$

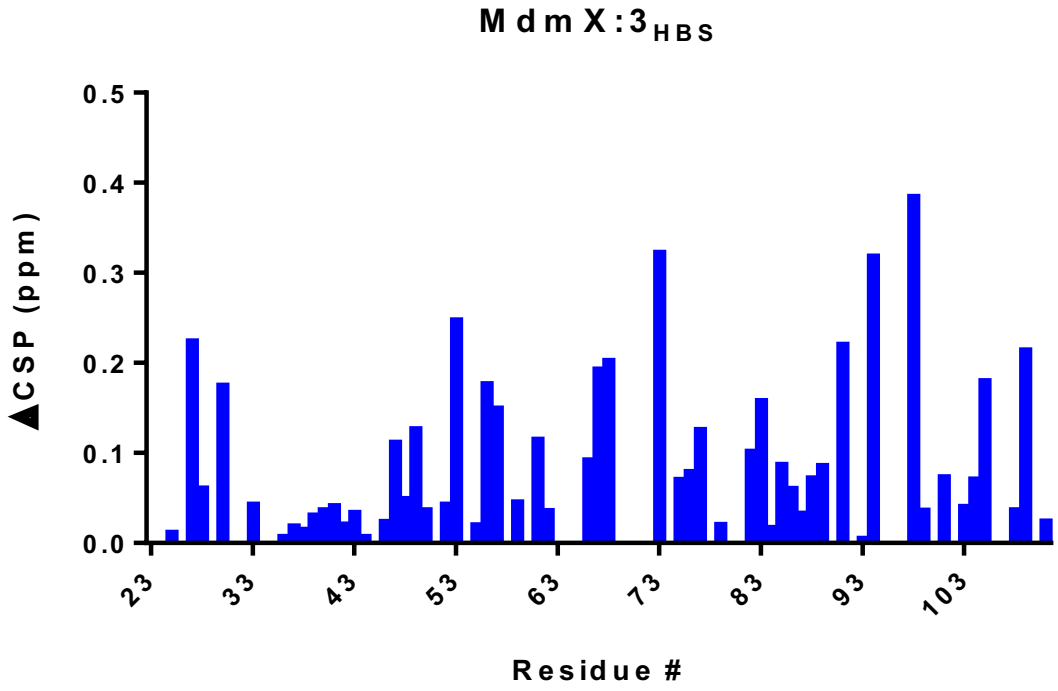


Figure S9 ^{15}N-^{1}H HSQC NMR chemical shift perturbations of MdmX upon addition of $\mathbf{3}_{\mathbf{HBS}}$

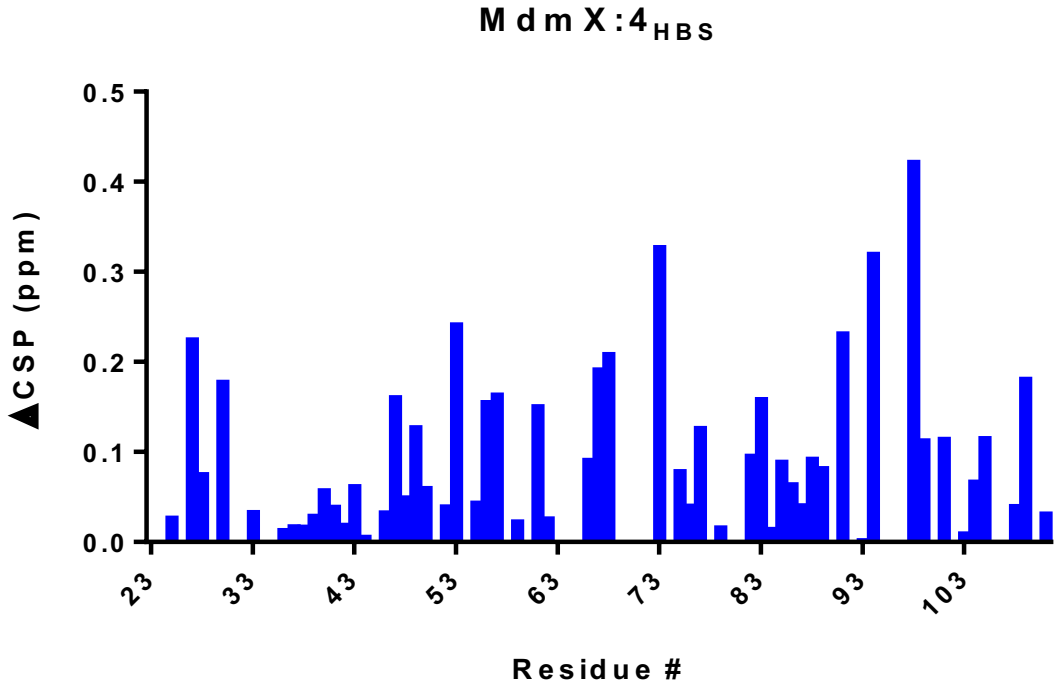


Figure S10 ^{15}N-^{1}H HSQC NMR chemical shift perturbations of MdmX upon addition of $\mathbf{4}_{\mathbf{HBS}}$

Supporting Methods

G* (*o*-Nosyl-*N*-allylglycine) Synthesis (CAS Registry Number 1521634-83-2). Under N_2, allylamine (**I**) (5 mmol, 374 μL) was dissolved in 7 mL DCM with triethylamine (5 mmol, 697 μL) on ice. *o*-nitrobenzenesulfonyl chloride (4.5 mmol, 1008 mg) was dissolved in 3 mL DCM and slowly added to the former solution. The mixture was brought to room temp and monitored by TLC (25% EtOAc/hexane). After 1.5 hr when *o*-NBS-Cl was consumed, the reaction was quenched with 10 mL 1M HCl (aq). The DCM layer was collected and the aqueous layer with was rinsed with 3x DCM. Organic layers were combined and concentrated *in vacuo*. The resulting yellow oil was dissolved in 10 mL EtOAc and rinsed 3x with brine. The organic layer was dried with Na_2SO_4, filtered, and concentrated *in vacuo* to afford Ns-allylamine (**II**) as a white solid in 93% yield.

Ns-allylamine (**II**) (15.9 mmol) and *tert*-butyl bromoacetate (20.7 mmol, 3052 μL) were dissolved in 50 mL DMF and stirred at room temp. To this solution was added K_2CO_3 (33.4 mmol, 4615 mg) portion-wise and the reaction was allowed to proceed overnight. Upon completion as determined by TLC (25% EtOAc/hexane), the reaction was quenched with ethanolamine (25 mmol, 2460 μL) and stirred for 1 hr. The mixture was diluted with 100 mL Et_2O and washed with 3x 1M HCl, 3x sat. $NaHCO_3$, and 3x brine. The organic later was dried over Na_2SO_4, filtered, and concentrated *in vacuo* to afford **III** in quantitative yield.

Compound **III** (2.9 mmol, 1.05 g) was dissolved in 20 mL of TFA/DCM (50:50) and stirred at room temperature for two hours until loss of starting material by TLC (25% EtOAc/hexane, drop of NEt_3). Product was concentrated *in vacuo* and recrystallized from EtOAc/hexane to afford **G*** as a white solid in quantitative yield.

***o*-Nosyl-*N*-allylglycine** [CAS Registry Number 1521634-83-2]. ESI-MS $[M+Na]^+$ Calc: 323.03, Obs: 323.0; ^1H NMR (500 MHz, CD_3CN) δ 7.99 (dd, 1H), 7.67 (m, 3H), 5.64 (m, 1H), 5.12 (m, 2H), 4.00 (s, 2H), 3.92 (d, J = 6 Hz, 2H); ^{13}C NMR (400 MHz, CD_3CN) δ 170.1, 148.7, 135.1, 133.5, 133.1, 133.0, 131.4, 125.0, 120.1, 51.6, 47.8

Peptide (1, 2, 3_{Unc}, 4_{Unc}) Synthesis. Peptides were synthesized by Fmoc SPPS on Rink Amide resin. Resin was thoroughly rinsed with DMF and DCM between steps. Fmoc deprotection was done with 20% piperidine in DMF with 0.1 M HOBt (2x 10 min). Amino acid coupling was performed by pre-activation of Fmoc-AA-OH with DIC and HOBt (5 eq. each with respect to resin, 10 min activation, 45 min coupling). N-termini were acetylated with 5 eq. each of Ac_2O and DIEA for 10 min. Cleavage from resin and global deprotection was done with 95% TFA, 2.5% triisopropylsilane, 2.5% water for two hours. Resin beads were removed by filtration and solvent was removed *in vacuo*. Solid peptide was obtained by trituration with cold diethyl ether followed by centrifugation and ether decantation. Peptides were purified by RP-HPLC on a C18 column running a gradient of 15-60% MeCN in water with 0.1% TFA followed by lyophilization.

HBS Peptide (3_{HBS}, 4_{HBS}) Synthesis. Peptides were synthesized by Fmoc SPPS on Rink Amide resin according to procedure reported by Miller *et al.*[1] Resin was thoroughly rinsed with DMF and DCM between steps. Fmoc deprotection was done with 20% piperidine in DMF with 0.1 M HOBt (2x 10 min). Amino acid coupling was performed by pre-activation of Fmoc-AA-OH with DIC and HOBt (5 eq. each with respect to resin, 10 min activation, 45 min coupling). **G*** and the following E were coupled with HOAt instead of HOBt. Deprotection of nosyl was achieved with 10 eq. each of thiophenol and DIEA in DMF (4 x 45 min). N-termini were acylated with 5 eq. each of 4-pentenoic acid and DIC in DMF. Peptides on resin were then thoroughly dried by washing with Et_2O, transferred to microwave tube, and placed under vacuum. To the resin was added 20 mol% Hoveyda-Grubbs II (Aldrich) and flushed with N_2. 3 mL dry 1,2-dichloroethane was added and the mixture was mixed under N_2 for 30 min. The tube was placed in a CEM Discover microwave device and subjected to 150 W and 120 °C with a 2 min ramp up and 15 min hold time. Cleavage from resin and global deprotection was done with 95% TFA, 2.5% triisopropylsilane, 2.5% water for two hours. Resin beads were removed by filtration and solvent was removed *in vacuo*. Solid peptide was obtained by trituration with cold diethyl ether followed by centrifugation and ether decantation. Peptides were purified by RP-HPLC on a C18 column running a gradient of 15-60% MeCN in water with 0.1% TFA followed by lyophilization.

FAM-P4 Fluorescent Probe Synthesis. Peptide was synthesized by Fmoc SPPS on Wang resin to afford a C-terminal acid as described above and reported by Czarna *et al.*[2] The N-terminus was capped with 5'-caroxyfluorescein (HOBt/DIC). Cleavage and purification were carried out as described above.

Protein Expression and Purification. A pET-14b vector coding for Mdm2 (25-117) with a 6xHis tag was obtained as a generous gift from the lab of Neal Zondlo. A pET-28b vector coding for MdmX (23-111) with a 6xHis tag was used for MdmX expression. Mdm2 expression was performed with 100 μg/mL ampicillin and MdmX expression was performed with 50 μg/mL kanamycin (both are referred to as "antibiotic" in the following). Plasmids were transformed into BL21(DE3) competent cells. One colony was inoculated in 5 mL LB with antibiotic overnight which was transferred to 1 L LB with antibiotic and shaken at 37 °C until an OD_{600} of 0.6. Expression was induced with 0.5 mM IPTG and the culture was shaken at room temperature for 6 hr. In the case of ^{15}N-labelled protein, LB was replaced with M9 minimal media supplemented with $^{15}NH_4Cl$. Cells were pelleted by centrifugation and stored at -80 °C. Buffers used for purification contained 20 mM tris-HCl pH 7.9, 1 M NaCl, and 5 mM β-mercaptoethanol. Cells were resuspended in 30 mL buffer with 5 mM imidazole, 10 mg lysozyme, 1 mM PMSF, 1 mM benzamidine, 1 μg/mL pepstatin A, and 1 μg/mL leupeptin. After the cells were sonicated on ice, lysate was collected by centrifugation and incubated with 1 mL pre-equilibrated Ni-NTA resin (Thermo) at 4 °C with gentle rocking for 30 min. Flow-through was collected and the resin was washed with buffer containing 20 mM imidazole. The resin was then treated with buffer fractions containing 60 to 300 mM imidazole for elution. Pure fractions identified by SDS-PAGE were pooled and dialyzed with a 3.5k MWCO slide-a-lyzer cassette (Thermo) in 1x PBS pH 7.4, 5 mM EDTA, 0.5 mM DTT, and 10% glycerol and stored flash-frozen in 10-20 μM aliquots at -80 °C.

Fluorescence Polarization (FP) Binding Assay. Binding affinities for peptide-protein interactions were determined with fluorescence polarization based on the method reported by Lao *et al.*[3] FP assays were performed in 1x PBS pH 7.4, 5 mM EDTA, 0.5 mM DTT, and 0.1% pluronic acid (Aldrich) on a DTX 880 Multimode Detector (Beckman) at 25° C, with 485 nm excitation and 525 nm emission. Protein was spin-concentrated with 3.5k MWCO centrifugal filter (Millipore Sigma) before experiments and direct binding with **P4** was always performed the day off competition experiments. Probe **P4** was held constant at 15 nM with protein titrated over 0 to 3.8 μM (Mdm2) and 9.5 μM (MdmX) Binding affinities reported are averages of three experiments and fitted to nonlinear regression sigmoidal dose-response model of GraphPad Prism 5.0.

$$K_{D1} = [R_T \times (1-F_{SB}) + L_{ST} \times F_{SB}^2]/F_{SB}-L_{ST}$$

R_T = [protein]

L_{ST} = [fluorescent probe]

F_{SB} = fraction fluorescent probe bound

The K_{D1} of **P4** was determined to be 11.22±2 nM and 26.5±10 nM to Mdm2 and MdmX, respectively. For competition experiments, wells contained a probe:protein concentration ratios corresponding to 0.7 fraction bound in the direct binding experiment. Peptide of interest was then titrated over 0 to 100 μM and the 96-well plate was agitated for 30 min. Peptide binding affinities, K_{D2}, were determined according to loss of fluorescence polarization from dissociation of probe-protein.

$$K_{D2} = K_{D1} \times F_{SB} \times [(L_T/(L_{ST} \times F_{SB} - (K_{D1} + L_{ST} + R_T) \times F_{SB} + R_T)) - 1/(1 - F_{SB})]$$

K_{D1} = K_D of probe-protein

R_T = [protein]

L_{ST} = [fluorescent probe]

F_{SB} = fraction peptide bound at EC_{50}

L_T = [peptide]

^{1}H-^{15}N HSQC NMR Spectroscopy. ^{15}N-labelled Mdm2 and MdmX were concentrated to 50 and 90 μM, respectively, in NMR buffer (25 mM tris-HCl pH 7.4, 150 mM NaCl, 0.1 mM EDTA, 0.02% NaN_3, 2 mM DTT, 10% D_2O) using a 3.5k MWCO centrifugal filter (Millipore Sigma). HSQC spectra were collected on a 600 MHz Bruker four-channel NMR at 298 K. Protein residue resonances were assigned as previously reported for Mdm2[4, 5] and MdmX.[6, 7] In separate experiments, HBS peptides $\mathbf{3}_{\mathbf{HBS}}$ and $\mathbf{4}_{\mathbf{HBS}}$ were titrated into solution with Mdm2 (0.5 eq peptide) and MdmX (1 eq peptide) and the differences were analyzed. Chemical shift perturbations were determined as previously described.[8]

$$\Delta\delta = \sqrt{0.5 \times (\delta_H^2 + 0.1\delta_N^2)}$$

CD Spectroscopy. Peptides $\mathbf{4}_{\mathbf{Unc}}$ and $\mathbf{4}_{\mathbf{HBS}}$ were dissolved at 50 μM in 50 mM sodium phosphate pH 7.4. Spectra were collected on a Jasco J-1500 CD Spectrometer as an average of 5 accumulations with scanning speed of 50 nm/min and digital integration time of 4 sec. Blank values were subtracted and data was normalized to mean residue ellipticity.

AlphaSpace Analysis and Fragment Screening. MDM protein structures are downloaded RCSB Protein Data Bank (PDB)[9] including 28 PDB files for Mdm2 and 20 PDB files for MdmX. Pocket mapping and comparison is performed using AlphaSpace 2.0 following the proposed protocol.[10] The structures are aligned, and in each structure β atoms are detected by AlphaSpace 2.0. All detected β atoms are clustered into ensemble pockets using average linkage[11] and β atoms in each structure are relabeled by the ensemble pocket ID. Pocket spaces and BScores are calculated using AlphaSpace 2.0. In order to constrain fragments in the pocket or pocket community of interest, detected β atoms are used to define the searching space of the docking algorithm, and a 2 Å buffer space was added on each searching dimension. Smina,[12] a fork of Autodock Vina,[13] is used to conduct fragment screening and the 20 amino acid side chains are used as the fragment library. Glycine and proline are not included in our analysis since the former does not have a side chain and the latter has its side chain fused with the backbone, which makes it hard to isolate the binding energy contribution of the side chain atoms and have a fair comparison with other amino acid side chains.

Supporting References

[1] Miller, S. E., Thomson, P. F., and Arora, P. S. (2014) Synthesis of Hydrogen-Bond Surrogate α-Helices as Inhibitors of Protein-Protein Interactions, *Curr. Protoc. Chem. Biol. 6*, 101-116.

[2] Czarna, A., Popowicz, G. M., Pecak, A., Wolf, S., Dubin, G., and Holak, T. A. (2009) High affinity interaction of the p53 peptide-analogue with human Mdm2 and Mdmx, *Cell Cycle 8*, 1176-1184.

[3] Lao, B. B., Drew, K., Guarracino, D. A., Brewer, T. F., Heindel, D. W., Bonneau, R., and Arora, P. S. (2014) Rational Design of Topographical Helix Mimics as Potent Inhibitors of Protein–Protein Interactions, *J. Am. Chem. Soc. 136*, 7877-7888.

[4] Stoll, R., Renner, C., Mühlhahn, P., Hansen, S., Schumacher, R., Hesse, F., Kaluza, B., Engh, R. A., Voelter, W., and Holak, T. A. (2000) Letter to the Editor: Sequence-specific 1H, 15N, and 13C assignment of the N-terminal domain of the human oncoprotein MDM2 that binds to p53, *J. Biomol. NMR 17*, 91-92.

[5] Uhrinova, S., Uhrin, D., Powers, H., Watt, K., Zheleva, D., Fischer, P., McInnes, C., and Barlow, P. N. (2005) Structure of Free MDM2 N-terminal Domain Reveals Conformational Adjustments that Accompany p53-binding, *J. Mol. Biol. 350*, 587-598.

[6] Chen, R., Zhou, J., Qin, L., Chen, Y., Huang, Y., Liu, H., and Su, Z. (2017) A Fusion Protein of the p53 Transaction Domain and the p53-Binding Domain of the Oncoprotein MdmX as an Efficient System for High-Throughput Screening of MdmX Inhibitors, *Biochemistry 56*, 3273-3282.

[7] Sanchez, M. C., Renshaw, J. G., Davies, G., Barlow, P. N., and Vogtherr, M. (2010) MDM4 binds ligands via a mechanism in which disordered regions become structured, *FEBS Lett. 584*, 3035-3041.

[8] Williamson, M. P. (2013) Using chemical shift perturbation to characterise ligand binding, *Prog. Nucl. Magn. Reson. Spect. 73*, 1-16.

[9] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235-242.

[10] Katigbak, J., Li, H., Rooklin, D., and Zhang, Y. (2020) AlphaSpace 2.0: Representing Concave Biomolecular Surfaces Using β-Clusters, *J. Chem. Inf. Model 60*, 1494-1508.

[11] Sokal, R. R., Michener, C. D., and Kansas, U. o. (1958) *A Statistical Method for Evaluating Systematic Relationships*, University of Kansas.

[12] Koes, D. R., Baumgartner, M. P., and Camacho, C. J. (2013) Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise, *J. Chem. Inf. Model 53*, 1893-1904.

[13] Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comp. Chem. 31*, 455-461.