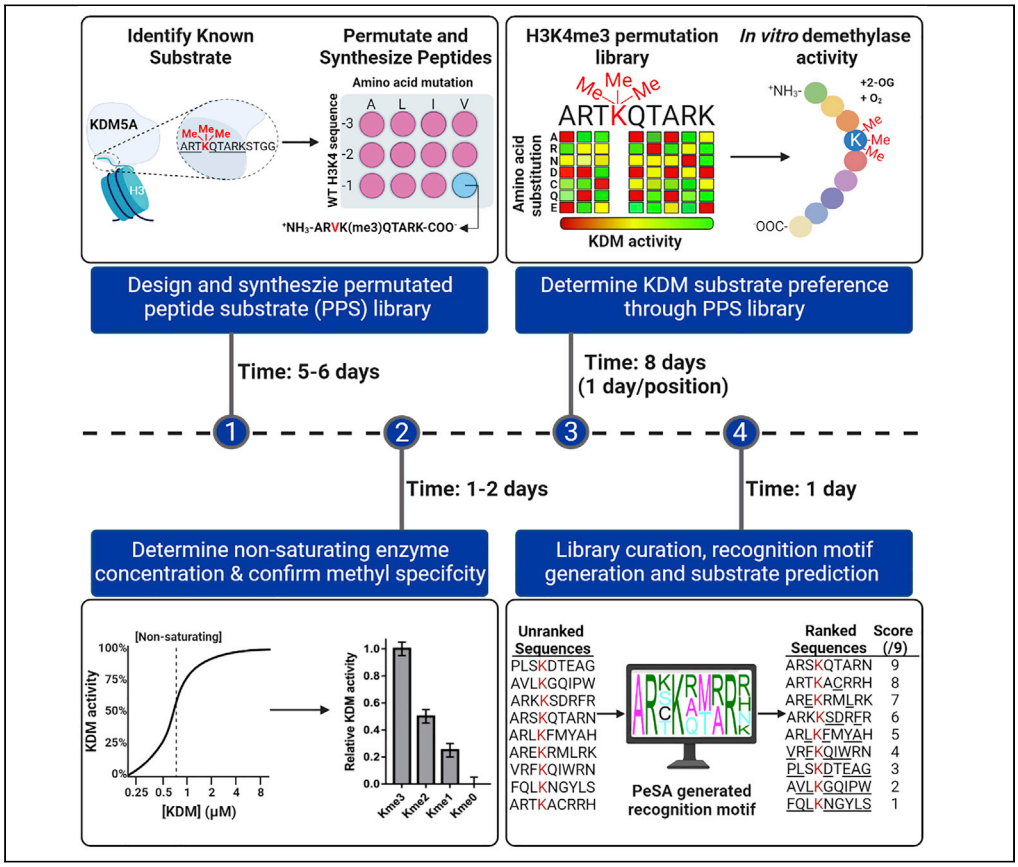
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Evaluation of Jumonji C lysine demethylase substrate preference to guide identification of *in vitro* substrates



Within the realm of lysine methylation, the discovery of lysine methyltransferase (KMTs) substrates has been burgeoning because of established systematic substrate screening protocols. Here, we describe a protocol enabling the systematic identification of JmjC KDM substrate preference and *in vitro* substrates. Systematically designed peptide libraries containing methylated lysine residues are used to characterize enzyme-substrate preference and identify new candidate substrates *in vitro*.

Matthew Hoekstra, Anand Chopra, William G. Willmore, Kyle K. Biggar

kyle_biggar@carleton.ca

Highlights

Use of a permutated substrate library to define JmjC KDM recognition motifs

JmjC KDM activity is measured via luminescent detection of succinate

Recognition motifs enable prediction of novel *in vitro* substrates of JmjC KDMs

Hoekstra et al., STAR Protocols 3, 101271
June 17, 2022
https://doi.org/10.1016/j.xpro.2022.101271

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Evaluation of Jumonji C lysine demethylase substrate preference to guide identification of *in vitro* substrates

Matthew Hoekstra,[1,2] Anand Chopra,[1,2] William G. Willmore,[1] and Kyle K. Biggar[1,3,4,*]

[1]Institute of Biochemistry and Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada

[2]These authors contributed equally

[3]Technical contact

[4]Lead contact

*Correspondence: kyle_biggar@carleton.ca
https://doi.org/10.1016/j.xpro.2022.101271

SUMMARY

Within the realm of lysine methylation, the discovery of lysine methyltransferase (KMTs) substrates has been burgeoning because of established systematic substrate screening protocols. Here, we describe a protocol enabling the systematic identification of JmjC KDM substrate preference and *in vitro* substrates. Systematically designed peptide libraries containing methylated lysine residues are used to characterize enzyme-substrate preference and identify new candidate substrates *in vitro*.

For complete details on the use and execution of this protocol, please refer to Hoekstra and Biggar (2021).

BEFORE YOU BEGIN

The protocol below describes the specific steps for determining a substrate recognition motif for the KDM5A demethylase, as well as how to leverage this information for *in vitro* substrate discovery. Results for KDM3A are shown alongside KMD5A to highlight (1) how KDM substrate preference deviates between enzymes and (2) that the method is not solely applicable to a single enzyme.

Theoretical considerations

The procedure for determining JmjC KDM substrate recognition motifs involves monitoring enzyme activity towards a library of peptides representing systematic mutations from the canonical substrate sequence (i.e., peptide permutation library). Analogously, this approach has been established for KMTs (Dhayalan et al., 2011; Kudithipudi et al., 2014; Lanouette et al., 2015; Rathert et al., 2008a, 2008b). Thus, this methodology is limited to those JmjC KDMs with a known peptide substrate. Many methyl-modifying enzymes have established substrates within the histone code (Hyun et al., 2017). For example, the canonical histone substrate for the KDM5 family of enzymes is histone H3 trimethylated at lysine-4 (i.e., H3-K4me3).

Practical considerations

To note, the methodology begins assuming a pure source of recombinant enzyme is available. Recombinant KDM5A and KDM3A enzymes were expressed and purified as described (Krishnan and Trievel, 2016; Rose et al., 2012). Furthermore, although our example monitors JmjC KDM activity via detection of succinate (i.e., turnover of the 2-oxoglutarate cofactor), there are other methods for detecting *in vitro* JmjC KDM activity. This includes other techniques monitoring cofactor turnover and formaldehyde formation that would also be viable alternatives for detecting enzyme activity ((Hirsilä et al., 2003; Kivirikko and Myllylä, 1982; Krishnan et al., 2012; Luo et al., 2006).

Table 1. Example of permutation of the H3-K4me3 peptide sequence at the -3 position (underlined)

Mutation	Peptide sequence
A(WT)	ART(Kme3)QTARKSTGGKA
R	RRT(Kme3)QTARKSTGGKA
N	NRT(Kme3)QTARKSTGGKA
D	DRT(Kme3)QTARKSTGGKA
C	CRT(Kme3)QTARKSTGGKA
Q	QRT(Kme3)QTARKSTGGKA
E	ERT(Kme3)QTARKSTGGKA
G	GRT(Kme3)QTARKSTGGKA
H	HRT(Kme3)QTARKSTGGKA
I	IRT(Kme3)QTARKSTGGKA
L	LRT(Kme3)QTARKSTGGKA
K	KRT(Kme3)QTARKSTGGKA
M	MRT(Kme3)QTARKSTGGKA
F	FRT(Kme3)QTARKSTGGKA
P	PRT(Kme3)QTARKSTGGKA
S	SRT(Kme3)QTARKSTGGKA
T	TRT(Kme3)QTARKSTGGKA
W	WRT(Kme3)QTARKSTGGKA
Y	YRT(Kme3)QTARKSTGGKA
V	VRT(Kme3)QTARKSTGGKA

Design permutated peptide substrate (PPS) library sequences

Timing: 1–2 h

Permutation libraries consist of peptide sequences whereby a given residue position is mutated to all other naturally occurring amino acids while leaving the remainder of the sequence unaltered. Systematically performing this for multiple residue positions enables the assessment of amino acid specificity of peptide-protein interactions over a defined window. The specificity of numerous KMTs has been mapped via this approach and the epitope is generally defined by the residues occurring directly proximal to the modification site. Structural analysis of JMJD14, a plant KDM5 enzyme, found that residues directly proximal to the H3-K4me3 target site are important for substrate recognition (Yang et al., 2018). Thus, the permutated window consisted of residues -3 to +5 amino acids relative to the H3-K4me3 methylation site (ARTKQTARKSTGGKA; K4 position bold, permutated window underlined).

1. Retrieve peptide sequence known to permit demethylase activity (e.g., the H3-K4 sequence [ART(Kme3)QTARKSTGGKA] was used for KDM5A).
2. At each residue position within the desired window (e.g., −3 to +5 relative to the methylation site), generate 19 other sequences whereby the wild-type residue is exchanged to another naturally occurring amino acid (Table 1 shows an example for the -3 position only).

   Optional: Add a tryptophan residue to the C-terminal end of the peptide sequence, separated from the peptide sequence by a flexible linker (e.g., 6-aminohexanoic acid (ahx)), to enable quantification of peptide concentration through tryptophan fluorescence.

Preparation of peptides, cofactors, and buffer stocks

Timing: 1 day

Due to the focus of the protocol described herein, the nonessential need to obtain peptides via the same methodology used, and the broad availability of peptide synthesis protocols, a detailed synthesis

protocol is not described. To note, the peptides were synthesized following standard Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry, on an automated ResPep SL peptide synthesizer (Intavis), at a scale of 2 μmol following procedures previously described (McKenna et al., 2021; Wei et al., 2018).

3. Dissolve dry peptides in 1× phosphate-buffered saline (PBS; pH 7), or another activity-compatible buffer), to a final concentration of 10 mM.

   Note: Given that peptides were synthesized at a scale of 2 μmol, dissolve the peptides in 200 μL of 1× PBS to obtain a concentration of 10 mM.

   Note: Peptides synthesized in-house may be highly acidic due to residue trifluoroacetic acid and may require pH to be adjusted to 7 by gradual addition of sodium hydroxide (a 5 M NaOH solution was used to add low volumes of 5–10 μL to peptides).

   ⚠ CRITICAL: all buffers and reagents should be prepared in ATP-free water. This includes cofactors and buffers described below.

   Optional: If peptides are synthesized in-house, it is highly recommended to quantify the peptides to ensure proper concentration. If an ahx-Trp is added to the C-terminal ends of peptide sequences, concentration may be determined by absorbance at 280 nm and the extinction coefficient of the peptide (available using the Expasy ProtParam tool; https://web.expasy.org/protparam/).

4. Prepare 4× peptide stocks (40 μM) by diluting peptides in 1× PBS. Next, prepare a 1 mM peptide stock solution for the wild-type substrate peptide in 1× PBS.

   ⏸ **Pause point:** Peptide stocks should be aliquoted and stored at −20°C until use.

5. Make 10 mM solutions of ascorbic acid, α-ketoglutarate, and Fe(II)SO_4 in a volume of 1 mL using ATP-free water, in separate 1.5 mL Eppendorf tubes.
   a. Dilute 10 mM cofactor stocks to 1 mM with ATP-free water.
6. Make a 40 mL solution of 0.5 M HEPES in a 50 mL conical tube. Dissolve HEPES in 30 mL of ATP-free water and adjust pH to 7.5 using 5 M NaOH. Top up volume to 40 mL to yield a final concentration of 0.5 M HEPES.

   Note: Stores HEPES solution at 4°C protected from light.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
$KDM5A_{1-801}$-His	SGC Oxford	n/a
$KDM3A_{515-1317}$-His-Flag	(Rose et al., 2012)	n/a
Fmoc-Ala-OH	P3 Biosystems	Cat# 41004
Fmoc-Arg(Pbf)-OH	P3 Biosystems	Cat# 41002
Fmoc-Asn(Trt)-OH	P3 Biosystems	Cat# 41007
Fmoc-Asp(OtBu)-OH	P3 Biosystems	Cat# 41019
Fmoc-Cys(Trt)-OH	P3 Biosystems	Cat# 41008
Fmoc-Glu(OtBu)-OH	P3 Biosystems	Cat# 41005
Fmoc-Gln(Trt)-OH	P3 Biosystems	Cat# 41011
Fmoc-Gly-OH	P3 Biosystems	Cat# 41010
Fmoc-His(Trt)-OH	P3 Biosystems	Cat# 41017
Fmoc-Ile-OH	P3 Biosystems	Cat# 41018
Fmoc-Leu-OH	P3 Biosystems	Cat# 41003

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fmoc-Lys(Boc)-OH	P3 Biosystems	Cat# 41001
Fmoc-Met-OH	P3 Biosystems	Cat# 41020
Fmoc-Phe-OH	P3 Biosystems	Cat# 41013
Fmoc-Pro-OH	P3 Biosystems	Cat# 41009
Fmoc-Ser(tBu)-OH	P3 Biosystems	Cat# 41006
Fmoc-Thr(tBu)-OH	P3 Biosystems	Cat# 41016
Fmoc-Trp(Boc)-OH	P3 Biosystems	Cat# 41012
Fmoc-Tyr(tBu)-OH	P3 Biosystems	Cat# 41014
Fmoc-Val-OH	P3 Biosystems	Cat# 41015
Fmoc-Lys(Boc,Me)-OH	P3 Biosystems	Cat# 47238
Fmoc-Lys(Me)2-OH.HCl	P3 Biosystems	Cat# 47237
Fmoc-Lys(Me)3-OH chloride	Sigma-Aldrich	Cat# F5062
HBTU	P3 Biosystems	Cat# 31001
NMM	Alfa Aesar	Cat# A12158
Rink Amide MBHA Resin	P3 Biosystems	Cat# 52002
Acetic anhydride	Fisher Scientific	Cat# 108-24-7
Piperidine	Sigma-Aldrich	Cat# 104094
DMF	Fisher Scientific	Cat# D119-20
Ethanol	Commercial Alcohols	Cat# P016EAAN
DCM	Acros Organics	Cat# 354800025
Trifluoroacetic acid	Fisher Scientific	Cat# L06374
TIPS	Acros Organics	Cat# 214922500
Ethyl Ether	Fisher Scientific	Cat# E138-4
HEPES	BioShop	Cat# HEP005.1
$Fe(II)SO_4$	BDH Chemicals	Cat# B28400
α-ketoglutarate	Sigma-Aldrich	Cat# K-3752
Ascorbic acid	J.T. Baker Chemical Co.	Cat# B581.5
Tris(2-carboxyethyl)phosphine (TCEP) HCl salt	Sigma-Aldrich	Cat# 646547
Bovine serum albumin	BioShop	Cat# ALB001.100
Dimethyl sulfoxide	Bio Basic Canada	Cat# D0231
NaCl	BioShop	Cat# SOD002.1
KCl	BioShop	Cat# POC308.500
Sodium phosphate monobasic monohydrate	Sigma-Aldrich	Cat# S9638
Sodium phosphate dibasic anhydrous	BioShop	Cat# SPD307.500
Potassium phosphate monobasic	BioShop	Cat# PPM666.1
Acetic acid	Anachemia	Cat# 00598-463
Critical commercial assays		
Succinate-Glo JmjC Demethylase/Hydroxylase Assay	Promega Corporation (Alves et al., 2018)	Cat# V7990
Software and algorithms		
Peptide Specificity Analyst (PeSA)	(Topcu and Biggar, 2019)	https://doi.org/10.5281/zenodo.6323540
Other		
BioTek Cytation 5 microplate reader	BioTek	Cat# BTCYT5M
Greiner 384-well plate, white	Greiner Bio-One	Cat# 781075

MATERIALS AND EQUIPMENT

2× Reaction Buffer/Peptide Mix

Reagent	Final concentration	Amount
0.5 M HEPES (pH 7.5)	50 mM	50 μL
1 mM ascorbic acid	200 μM	100 μL
1 mM $Fe(II)SO_4$	20 μM	10 μL

(*Continued on next page*)

Continued		
Reagent	Final concentration	Amount
1 mM 2-oxoglutarate	20 μM	10 μL
DMSO	2%	10 μL
ATP-free water	n/a	310 μL
1 mM peptide	20 μM	10 μL
Total	**n/a**	**500 μL**

2× Reaction Buffer		
Reagent	Final concentration	Amount
0.5 M HEPES (pH 7.5)	50 mM	50 μL
1 mM ascorbic acid	200 μM	100 μL
1 mM $Fe(II)SO_4$	20 μM	10 μL
1 mM 2-oxoglutarate	20 μM	10 μL
DMSO	2%	10 μL
ATP-free water	n/a	320 μL
Total	**n/a**	**500 μL**

Succinate Detection Reagent I		
Reagent	Final concentration	Amount
Succinate-Glo™ Solution	n/a	5 μL
Acetoacetyl-CoA 100×	1×	5 μL
Succinate-Glo™ Buffer	n/a	500 μL
Total	**n/a**	**510 μL**

Alternatives: Here a peptide list comprising the known methylproteome was scored using JmjC KDM recognition motifs. This peptide list is provided (Table S1, accessed from PhosphoSitePlus on 12-03-2020), however any peptide list can be used (Hornbeck et al., 2012).

STEP-BY-STEP METHOD DETAILS

Determine non-saturating enzyme concentration

Timing: 4 h

Given a fixed set of reaction parameters (e.g., cofactor concentration, pH, temperature, etc.), a user can determine the optimal concentration range of JmjC KDM yielding a non-saturated signal. Establishing a non-saturating enzyme concentration is key for downstream steps when comparing the effect of different peptides on JmjC KDM activity. This is achieved by dilution series of the given KDM from the highest concentration possible to a low or sub nanomolar concentration (e.g., <1–10 nM). In our prototypic example of KDM5A and KDM3A, by dilution series we determined enzyme concentrations yielding non-saturating signal (Figure 1).

Note: To maximize the signal-to-background ratio, we recommend users of this methodology establish optimal conditions (e.g., pH, temperature, etc.) prior to beginning this workflow. This may be done experimentally or by literature search.

Note: Commercially available *Promega Succinate-Glo™* assay reagents are stored at −80°C. Begin thawing reagent components on ice ~1 h before use. For Acetoacetyl-CoA 100×, specifically, thaw at 22°C–24°C for 5 min before addition to Succinate-Glo™ Buffer.

1. Place a white 384-well microplate on a Peltier device and set temperature to 4°C.

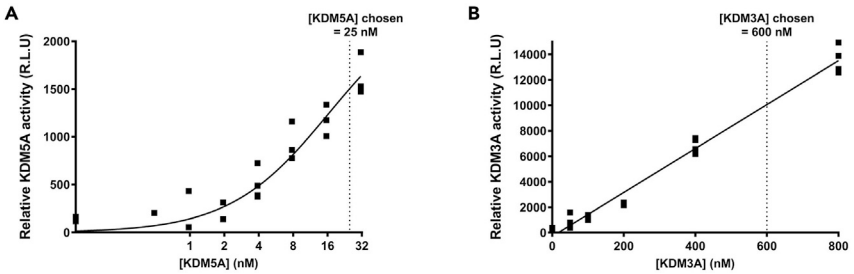


Figure 1. Determining non-saturating JmjC KDM concentration

(A and B) Two-fold dilution series of recombinant (A) KDM5A and (B) KDM3A enzymes, observing activity towards H3-K4me3 and H3-K9me2 peptides, respectively. The data represents the average of three luminescent readings and error bars represent the SEM (n=3). Data for KDM5A was previously published (Hoekstra and Biggar, 2021).

2. Prepare 2× Reaction Buffer/Peptide Mix.
3. Add 2.5 µL of 2× Reaction Buffer/Peptide Mix to all experimental wells in the 384-well microplate.
4. Thaw JmjC KDM stock protein on ice.
5. Perform a 20 µL dilution series (e.g., 4× or 2×) of the JmjC KDM protein in 1.5 mL Eppendorf tubes on ice in the same buffer used initially for storage (see Table 2 for an example dilution range).

   Note: Save at least a 20 µL aliquot of storage buffer for the "no enzyme" control.

   Alternatives: Dilution series may be performed in higher or lower volumes depending on the number of replicate reactions performed.

6. Set Peltier device to 23°C.

   Alternatives: Other temperatures for the JmjC KDM reaction may be used (e.g., optimal temperature may be determined prior to this protocol or published in literature).

7. To initiate the reaction, add 2.5 µL of each JmjC KDM dilution to a specific microplate well containing the 2× Reaction Buffer/Peptide Mix (Table 2 shows an example plate layout).
   a. Set the 384-well microplate on a plate shaker for 2 min at 22°C–25°C.
   b. Centrifuge the plate at 240 × *g* for 1 min at 23°C.

   Note: For the no enzyme control (NEC), add 2.5 µL of storage buffer containing an absence of JmjC KDM protein.

8. Allow the demethylation reaction to occur for 60 min at 23°C.
9. Add 5 µL of Succinate Detection Reagent I to each assay well.
   a. Shake the 384-well microplate for 2 min at 22°C–25°C.
   b. Centrifuge the plate at 240 × *g* for 1 min at 23°C and incubate for 60 min at 23°C.

   Note: The time of JmjC reaction may be adjusted, however, the reaction time used in this step must be kept consistent when performing all downstream JmjC KDM reactions. For example, reactions with KDM3A occurred for 3 h.

10. Add 10 µL of Succinate Detection Reagent II.
    a. Shake the 384-well microplate for 2 min at 22°C–25°C.
    b. Centrifuge the plate at 240 xg for 1 min at 23°C and incubate for 10 min at 23°C.
11. Read the luminescence from each well using a microplate reader.

Table 2. Example of plate layout for JmjC 2× dilution series

Column	1	2	3	4	5	6	7	8	9	10	11	12	13
Row A-C (n=3)	NEC	10	5.0	2.5	1.3	0.63	0.31	0.16	0.078	0.039	0.020	0.01	0.005

Initial concentration of purified KDM5A was 20 μM, yielding a final reaction concentration of 10 μM. Concentrations listed are in μM.

12. Calculate the average luminescence signal across technical replicates.
13. For each condition, subtract the average signal from that of the average of the NEC to account for background luminescence.
14. Using data analysis/graphing software, such as GraphPad Prism, visualize the data to determine non-saturating and linear dose responsive JmjC KDM concentrations (Figure 1). Choose a concentration of protein within the mid-linear range of relative activity.
    a. See Troubleshooting 1 if JmjC KDM activity is difficult to observe.

Validate KDM methyl-state preference

Timing: 4 h

This step aims to validate the methyl-state preference (i.e., mono-, di-, tri-methylation) of the recombinant JmjC KDM for an established substrate. This is important as many JmjC KDMs have been shown to be capable of discriminating between different methyl-states of the same substrate. As the peptide libraries used for this method are synthesized as one methyl-state, this validation should be performed prior to commercially ordering or synthesizing the full peptide permutation library. Given the optimal concentration of the JmjC KDM determined in the previous step, the JmjC KDM activity can be monitored by performed reactions with peptide substrates of different methylation states. For assay validation, it is important to include an unmethylated substrate and a 'no peptide control' (NPC) alongside the standard NEC.

⚠ CRITICAL: Some JmjC KDMs can convert 2-oxoglutarate to succinate in an appreciable extent in the absence of methylated peptide substrate. As a result, a 'no peptide control (NPC) is critical to use going forward to assess the true level of relative demethylation activity (Table 3).

15. Thaw 40 μM wild-type substrate (including null-, mono-, di-, and tri-methyl state peptides) on ice.
16. Repeat steps 1–3, except using 2× Reaction Buffer (similar to 2× Reaction Buffer/Peptide Mix minus peptide substrate).
    a. Keep 2× Reaction Buffer on ice for same day use.
17. Add 1.25 μL of 40 μM wild-type peptide substrates, individually, to each experimental well containing 2× Reaction Buffer.
    a. For the NPC, use 1.25 μL of 1× PBS.

Note: If peptides were dissolved in a buffer other than 1× PBS, use the peptide buffer for the NPC.

18. Thaw JmjC KDM stock protein on ice and dilute to 4× the optimal concentration determined in step 14.

Table 3. Example of plate layout for testing JmjC KDM activity towards differentially methylated peptides (monomethylation [Kme1], demethylation [Kme2], and trimethylation [Kme3])

Column	1	2	3	4	5	6
Row A-C (n=3)	NEC	NPC	Null	Kme1	Kme2	Kme3

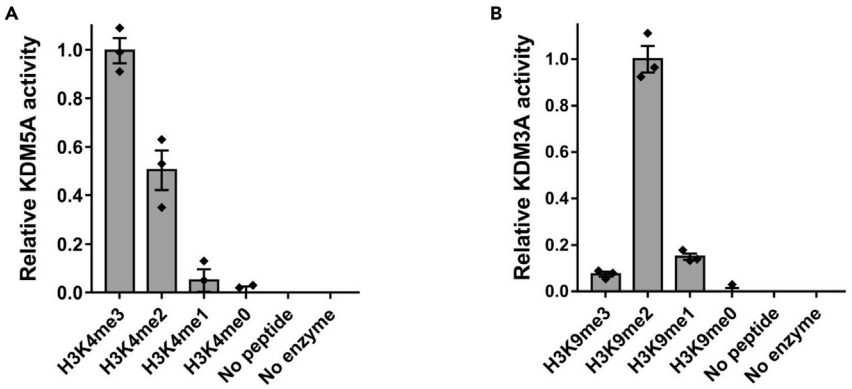


Figure 2. Validation of JmjC KDM methyl-state preference

(A and B) Methyl-state specificity of (A) KDM5A and (B) KDM3A towards differentially methylated H3-K4 and H3-K9 peptides, respectively. KDM activity towards non-methylated, mono-, di-, and tri-methylated (i.e., me0/1/2/3) peptides at the indicated residues was assessed. Data represents the mean luminescence and standard deviation (n=3) normalized to the max signal. Data for KDM5A was previously published (Hoekstra and Biggar, 2021).

a. Considering the number of experimental conditions requiring enzyme (5 in total), the volume of JmjC KDM needed per reaction (1.25 µL), and the number of replicates (e.g., n=3), make at least 20.63 µL of 2× JmjC KDM (this is 10% excess of the minimum volume needed (18.75 µL)).
b. Perform the protein dilution with the JmjC storage buffer.

19. Set Peltier device to 23°C (or appropriate constant temperature).
20. Add 1.25 µL of 4× JmjC KDM dilution to all experimental wells, except for the NEC. Add 1.25 µL JmjC KDM storage buffer to the NEC.
21. Repeat steps 8–13.
22. For each condition, subtract the signal from that of the NPC to account for demethylation-uncoupled 2-oxoglutarate turnover.

Optional: Represent methyl-state preference as relative JmjC KDM activity; normalize subtracted luminescent signals to that of the peptide displaying the highest signal. Again, using data analysis/ graphing software visualize the data to represent methyl-state preference (Figure 2).

Determine JmjC KDM substrate preference by permutated peptide substrate (PPS) library

Timing: 4 h (per mutation position)

These steps aim to map JmjC KDM substrate preference by monitoring activity towards all peptides in the PPS library. At this point, all peptides within the library should be methylated according to the preferred methyl-state (e.g., H3-K4me3 and H3-K9me2 for KDM5A and KDM3A, respectively). The experimental set-up is performed in the same manner as the previous section (methyl-state validation), except using permutated peptides alongside the wild-type peptide and reaction controls (Table 4). In analyzing the data relative to the positive control (wild-type peptide) and negative controls (no peptide and no enzyme controls), one can determine the relative effect of individual amino acid mutations on JmjC KDM activity.

Note: To limit variability in the start of reaction time between experimental conditions; it is recommended to assess JmjC KDM activity towards peptides representing one residue position of the PPS library at a time.

23. Thaw 40 µM peptide stocks for the PPS library position being tested on ice.

Table 4. Example of plate layout for PPS library experiments

	Experimental permutated peptides																				Controls	
Column	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Row A-D (n=4)	R	H	D	E	N	Q	A	G	P	I	L	V	M	F	Y	W	S	T	C	K	NPC	NEC

The following is set up for mutations in one residue position. Each single letter amino acid designation is representative of the amino acid substitution occurring in the mutated peptide substrate.

24. Make fresh 2× Reaction Buffer (e.g., for each residue position, 242 μL of 2× Reaction Buffer is required (22 conditions × 4 replicates × 2.5 μL = 220 μL + 10% extra).
25. Repeat steps 1–3, using the freshly prepared 2× Reaction Buffer.
26. Add 1.25 μL of 4× (40 μM) peptide, individually, to experimental wells containing 2× Reaction Buffer. For the NPC, use 1.25 μL of 1× PBS.

Note: Again, if peptides were dissolved in a buffer other than 1× PBS, use that buffer for the NPC.

⚠ CRITICAL: Ensure each run of the experiment contains a reaction condition with a wild-type peptide. All mutation positions have a wild-type peptide. However, if mutation positions are not run in full, a wild-type peptide must be included.

27. Thaw JmjC KDM stock on ice and dilute to 4× the optimal concentration determined in the previous step in JmjC storage buffer.

Note: Consider the number of experimental conditions requiring enzyme (e.g., the volume of JmjC KDM needed per reaction well (1.25 μL), and the number of replicates (e.g., n=4), make 115.5 μL of 2× JmjC KDM (this is 10% excess of what is needed; 21 conditions × 4 replicates × 1.25 μL = 105 μL)).

28. Set Peltier device to 23°C (or appropriate constant temperature).
29. Add 1.25 μL of 4× JmjC KDM solution to all experimental peptide conditions (wild-type peptide and mutant peptides) and the NPC well. Add 1.25 μL JmjC KDM storage buffer to the NEC well.
30. Repeat steps 8–13.
31. For each condition, subtract the average signal of the NPC to account for demethylation-uncoupled 2-oxoglutarate turnover.
32. To assess the relative effect of mutations on JmjC KDM activity; normalize substrate luminescent signals to that of the wild-type peptide.

Note: Normalization is achieved from dividing the luminescent signals of reactions with experimental peptides by that of the WT peptide. This is critical to gauge the relative effect of individual mutations on enzyme activity.

33. Repeat steps 23–32 for each mutation position.

Optional: Represent relative activity data as a heat-map to visualize JmjC KDM substrate preference (Figure 3) using basic data analysis software (e.g., Microsoft Excel).

JmjC KDM recognition motif generation, substrate prediction, and *in vitro* validation

Timing: 1–2 h

The major aim of this step is to use JmjC KDM substrate preference to prioritize peptides to be tested for *in vitro* KDM activity. Peptide Specificity Analyst (PeSA) software can be used to easily produce a

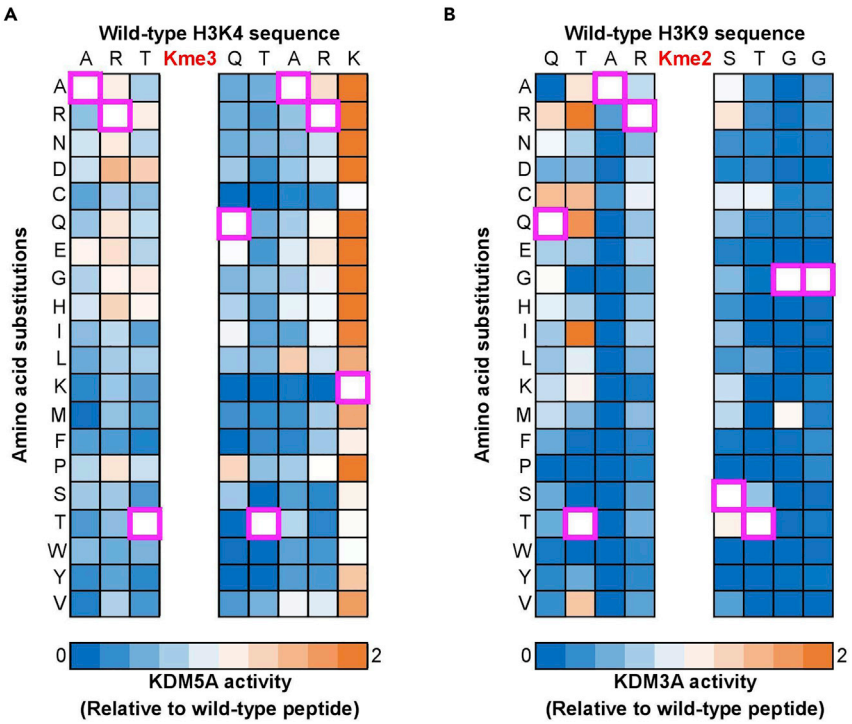


Figure 3. Representation of KDM substrate preference as a heat-map

(A and B) Relative activity of recombinant (A) KDM5A and (B) KDM3A were assessed towards permutation libraries of their corresponding canonical substrates of preferred methyl state (i.e., H3-K4me3 and H3-K9me2, respectively). X-axis and y-axis represent the wild-type peptide sequence and amino acid substitutions, respectively. Location of wild-type peptide (i.e., relative activity 1.0) spots are defined by the pink borders.

candidate KDM recognition motif (Topcu and Biggar, 2019). In turn, these candidate recognition motifs can then be used to score and prioritize the methyllysine proteome for candidate substrates.

34. Download and open PeSA software (Topcu and Biggar, 2019; also available at https://github.com/EmineTopcu/PeSA).
35. Select "Permutation Array" tab at the menu.
36. Click "Load from File" and upload your JmjC KDM quantification matrix.

Note: The quantification matrix can be uploaded with either the relative activity values or the background subtracted luminescent values. If the latter, normalization to the wild-type luminescent values in each position can be performed within the software to obtain relative activity values. Under "Normalized", select "Per Row/Column" (see Table S2 for an example of the KDM5A quantification matrix).

Note: Experimentally, only the residue positions flanking the methylation site are assessed. However, to maintain the central target lysine within the visualization of the recognition motif; in the quantification matrix one can assign the central methylated K the relative value of "1" and all other mutations in this central position as "0".

37. Set the "Threshold" to the desired the value.

Note: In this context the "Threshold" defines the minimum level of relative JmjC KDM activity for an amino acid substitution to be included within the candidate recognition motif. Keep in mind that stringency increases with higher threshold values.

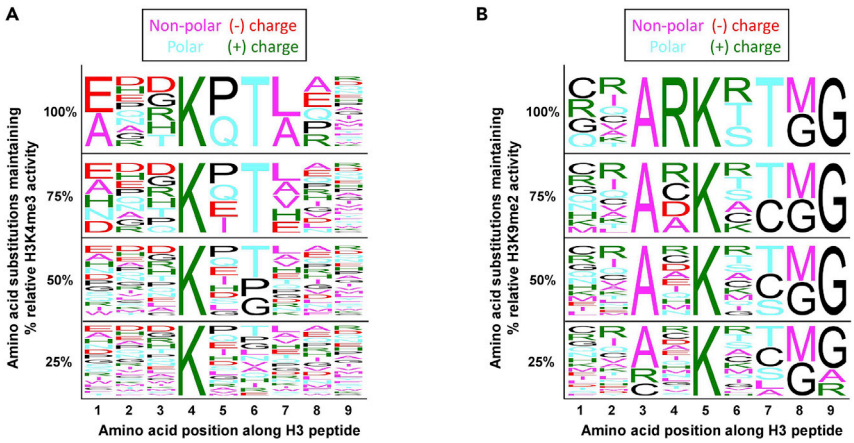


Figure 4. Representation of the screening data as PeSA generated sequence motif

(A and B) The recognition motifs for (A) KDM5A and (B) KDM3A depict amino acid substitutions maintaining a minimum level relative activity defined by the threshold values on the y-axis.

Alternatives: Another threshold value may be used and should be considered when Troubleshooting problem 2. Figure 4 represents KDM5A and KDM3A recognition motifs set at various threshold values (e.g., 0.25, 0.5, 0.75, and 1.0).

38. Obtain the recognition motif under the "motif" tab on the left-hand side of the window.
39. Score peptides that possess windows which reflect known methylation sites in the methylproteome, using the PeSA generated candidate recognition motif.
    a. Each position in the queried sequence which shares an amino acid with the recognition motif contributes a value of 1 to the total score. An example of scoring is provided in Table 5.

Note: Given the recognition motif produced and Kme position, ensure that the queried windows are of the same length and Kme residue in question is in the correct position. The PeSA score reflects the number of residues in the queried peptide matching those in the recognition motif at their corresponding positions. As a result, the predictions were made based on a 9-residue recognition motif, and therefore, the highest possible score obtainable is a total PeSA score of 9 (i.e., for windows exactly matching the recognition motif) (Table 5).

40. Determine mean (μ) and standard deviation (σ) of PeSA scores across all queried peptide sequences.
41. To define high-ranking peptide sequences, determine the PeSA score corresponding to 2 standard deviations above the mean.
    a. For KDM5A, using the 0.5 relative activity recognition motif, a PeSA score of 9 reflects high-ranking peptide sequences (μ = 6.6, σ = 1.2) (Figure 5A).
42. Repeat steps 23–32 in the previous section, except using high-ranking peptides instead of PPS library peptides to validate the *in vitro* activity of substrate predictions/rankings (Figure 5B).

Table 5. Example of motif-based scoring

		Position and tolerable substitutions within motif									
Protein-site	Queried sequence	-4 [CRGQ]	-3 [RIQCVAKT]	-2 [A]	-1 [R]	0 [K]	+1 [RTS]	+2 [T]	+3 [MG]	+4 [G]	PeSA score
H3-K9 (known)	QTARKSTGG	Q	T	A	R	K	S	T	G	G	9
RPA2-K693 (unknown)	CQMGKQTMG	C	Q	M	G	K	Q	T	M	G	6

Example of motif-based scoring. Example of scoring 9-mer sequences with KDM3A motif (defined by amino acids maintaining at least 100% relative activity). Tolerable amino acids at each position are shown in brackets. Matching and mismatching residues, compared to the motif, are shown in green and red, respectively. PeSA score is defined by the number of matching residues.

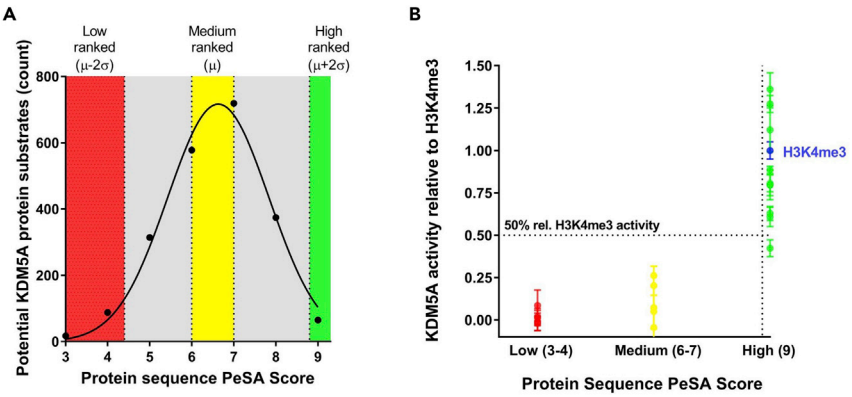


Figure 5. Distribution of peptide scores and testing *in vitro* KDM5A activity

(A) Gaussian distribution of PeSA scores of methylproteome peptides scored with the 0.5 relative activity KDM5A recognition motif. High- and low-ranking substrates are defined by PeSA scores occurring 2 standard deviations (σ = 1.2) above (or equal to) or below (or equal to) the population mean (μ = 6.6), respectively.

(B) Relative KDM5A activity towards differentially classified peptides sequences. Substrates were randomly chosen within their respective classification and KDM5A activity is shown as relative to the H3-K4me3 substrate.

a. See Troubleshooting 2 if JmjC KDM shows no activity to none, or few, of the predicted JmjC KDM substrates

Note: If the focus is purely on JmjC KDM substrate discovery, we recommend testing as many of the high-ranking peptides as possible, if not all. However, if the user is also interested in assessing whether the method accurately predicts or enriches in highly active *in vitro* substrates for the select JmjC KDM, one can assess a select number of random substrates with high-, medium-, and low-ranking PeSA scores (Figure 6 demonstrates this analysis for KDM5A).

EXPECTED OUTCOMES

In our experience, PeSA scores of large lists of peptide sequences (e.g., methylproteome; n=2,155 sequences) follow a Gaussian distribution. For example, Figure 5A shows the distribution of PeSA scores of methylproteome sequences, scored with the 0.5 relative activity recognition motif of KDM5A. Thus, applying appropriate thresholds, based on standard deviation, to classify substrates as 'high-ranking' will result in a small number of candidate substrates to test *in vitro* JmjC KDM activity.

If testing relative JmjC KDM activity towards predicted substrates, we expect PeSA scores to correlate with *in vitro* JmjC KDM activity. For example, we sought to test KDM5A activity towards a handful of high, medium- and low-ranking substrates (Figure 5B). We were able to observe that none of the medium and low-ranking substrates displayed any significant KDM5A activity (activity defined as greater than 50% of H3-K4me3 activity). Furthermore, we observed KDM5A activity towards 90% of high-ranked peptides. This verifies that the 0.5 relative activity KDM5A recognition motif was accurate in predicting peptide sequences that would be amenable to KDM5A activity *in vitro*. Ultimately, this added step of validation does prioritize several potential substrates to refine and focus further validation efforts on.

LIMITATIONS

Limitation 1

A limitation to consider is that the relative importance of the individual amino acids at each position in the recognition motif is determined in the context of a fixed sequence and thus sequence bias may influence the results. This is a consequence of using permutation-based exploration of enzyme

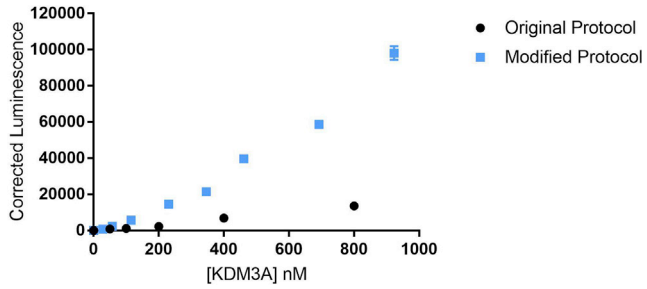


Figure 6. Luminescent signal from KDM3A reaction, detected with the original (black) and modified (blue) assay protocols

Results for the modified assay protocol were plotted over the results of the original assay protocol displayed in Figure 1B.

specificity. It is certainly possible that the relative importance of a given amino acid is only observable when occurring in the presence of specific amino acids at other positions. Supporting this, and in the context of methyl-binding domains, JMJD2A-double Tudor domain binding to permutations of H3-K23me3, H3-K4me3, and H4-K20me3 show distinct specificities depending on wild-type sequence used (Liu et al., 2013).

Limitation 2

The protocol defined here assumes recognition of substrates by a given enzyme is specified by the residues directly proximal to the modification site (i.e., −4 to +5 positions). Although this has been shown to be the case for many KMT enzymes, and for Suv39H2 the +5 position also determines specificity (Schuhmacher et al., 2015), recognition of substrates may also be determined by more distant interactions. Additionally, on the H3 tail, more distal sequence elements were recently shown to be important for KDM5A-dependent demethylation of the H3-K4 site (Petronikolou et al., 2020).

TROUBLESHOOTING

Problem 1

Low level of observable enzyme activity (i.e., low signal-to-background ratio). This problem is referring to corresponding protocol step 14.

Potential solution

A low observable enzyme activity could be due to several factors such as reaction buffer composition, pH, temperature, time, as well as substrate and cofactor concentration. Each factor mentioned may be optimized individually. Additionally, additives may affect enzyme activity. For example, KDM3A activity may be improved nearly two-fold by addition of TCEP at certain concentrations, whereas higher concentrations of sodium chloride hinder KDM3A activity (Yu et al., 2014). Furthermore, peptide length may influence the level of enzyme activity. For example, the affinity of KDM3A for the H3-K9 substrate has been observed to be nearly 200-fold greater when using 21-mer peptides compared to 15-mer peptides (Goda et al., 2013; Yu et al., 2014).

Furthermore, if the level of uncoupled 2-oxoglutarate conversion (i.e., succinate production in absence of peptide) is relatively high compared to succinate formation in the presence of peptide; true demethylation activity may be difficult to observe using this assay or other assays detecting 2-oxoglutarate turnover. If this is the case, we suggest using assays directly assessing demethylation (e.g., formaldehyde turnover). However, if a significant difference is still able to be observed between the two conditions (i.e., JmjC KDM activity in the presence of peptide versus no peptide), succinate detection may still be a feasible technique. In this case, the uncoupled 2-oxoglutarate conversion must be considered as a baseline.

Finally, as JmjC KDMs belong to the Fe(II)/2-oxoglutarate-dependent family of dioxygenases they are susceptible to product inhibition conferred by the formation of succinate. The potency of inhibition varies depending on the specific enzyme of interest. To attenuate this effect, the assay can be performed in a format such that any succinate formed is immediately converted during the demethylation reaction, rather than converting succinate at the end of the JmjC KDM reaction (e.g., incorporating coupling enzymes in the KDM reaction). To achieve this, Succinate-Glo Solution and Acetoacetyl-CoA, provided in the assay kit, may be incorporated in the demethylation reaction by ensuring each of these components are diluted 100-fold in the final assay volume. Although the materials in the kit utilized is proprietary, succinate may be directly converted to succinyl-CoA by 3-oxoacid-CoA-transferase in the presence of acetoacetyl-CoA. For KDM3A, specifically, we found that the modified protocol increases the luminescent signal (Figure 6).

Problem 2

The JmjC KDM assessed shows no activity to none, or few, of the predicted JmjC KDM substrates. This problem is referring to corresponding protocol step 42.

Potential solution

No activity towards any of the substrate predictions could be due to several factors, such as limitations inherent to permutation-based exploration of specificity. The latter cannot be addressed by adjusting variables within this protocol and would need to be resolved by assessing specificity in another manner or by using multiple substrate permutations. However, the adjustable variables include reducing the stringency of substrate predictions to be more permissive for substrate discovery (i.e., use recognition motif defined by a lower activity threshold). It may also be beneficial to test activity of the given JmjC KDM at multiple substrate concentrations. Finally, and related to Limitation 2, the critical positions defining specificity may exist outside of the permutated positions and thus it may be beneficial to expand this window and re-predicting substrates based on this expanded recognition motif.

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact Dr. Kyle K. Biggar (kyle_biggar@carleton.ca).

Materials availability

This protocol is not associated with any newly generated materials.

Data and code availability

The protocol includes the methylproteome peptide list (Table S1) accessed from PhosphoSitePlus on 12-03-2020 (Hornbeck et al., 2012).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xpro.2022.101271.

ACKNOWLEDGMENTS

This work was supported by a National Science and Engineering Research Council (NSERC) Canada Discovery grant to K.K.B. (grant no. RGPIN-2016-06151) and W.G.W. (grant no. RGPIN-2017-06414). A.C. held a Canada Graduate Scholarship – Doctoral (CGS-D) from the NSERC of Canada. The KDM5A Sf9 construct was a generous gift from the Structural Genomics Consortium Oxford group. The KDM3A construct was a generous gift from Nicola Burgess-Brown.

AUTHOR CONTRIBUTIONS

Methodology, M.W. and A.C.; data curation, M.H. and A.C.; validation, M.H.; writing – original draft, M.H. and A.C.; writing – review & editing, A.C., K.K.B., and W.G.W.; conceptualization, K.K.B.; resources, K.K.B.; funding acquisition, K.K.B.; supervision, K.K.B. and W.G.W.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Alves, J., Vidugiris, G., Goueli, S.A., and Zegzouti, H. (2018). Bioluminescent high-throughput succinate detection method for monitoring the activity of JMJC histone demethylases and Fe(II)/2-oxoglutarate-dependent dioxygenases. SLAS Discov. *23*, 242–254. https://doi.org/10.1177/2472555217745657.

Dhayalan, A., Kudithipudi, S., Rathert, P., and Jeltsch, A. (2011). Specificity analysis-based identification of new methylation targets of the SET7/9 protein lysine methyltransferase. Chem. Biol. *18*, 111–120. https://doi.org/10.1016/J.CHEMBIOL.2010.11.014.

Goda, S., Isagawa, T., Chikaoka, Y., Kawamura, T., and Aburatani, H. (2013). Control of histone H3 lysine 9 (H3K9) methylation state via cooperative two-step demethylation by Jumonji domain containing 1A (JMJD1A) homodimer *. J. Biol. Chem. *288*, 36948–36956. https://doi.org/10.1074/JBC.M113.492595.

Hirsilä, M., Koivunen, P., Günzler, V., Kivirikko, K.I., and Myllyharju, J. (2003). Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. J. Biol. Chem. *278*, 30772–30780. https://doi.org/10.1074/JBC.M304982200.

Hoekstra, M., and Biggar, K.K. (2021). Identification of in vitro JMJD lysine demethylase candidate substrates via systematic determination of substrate preference. Anal. Biochem. *633*, 114429. https://doi.org/10.1016/J.AB.2021.114429.

Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. Exp. Mol. Med. *49*, e324. https://doi.org/10.1038/emm.2017.11.

Kivirikko, K.I., and Myllylä, R. (1982). Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. Methods Enzymol. *82*, 245–304. https://doi.org/10.1016/0076-6879(82)82067-3.

Krishnan, S., Collazo, E., Ortiz-Tello, P.A., and Trievel, R.C. (2012). Purification and assay protocols for obtaining highly active Jumonji C demethylases. Anal. Biochem. *420*, 48–53. https://doi.org/10.1016/J.AB.2011.08.034.

Krishnan, S., and Trievel, R.C. (2016). Purification, biochemical analysis, and structure determination of JmjC lysine demethylases. In Methods in Enzymology (Academic Press Inc.), pp. 279–301. https://doi.org/10.1016/bs.mie.2016.01.023.

Kudithipudi, S., Lungu, C., Rathert, P., Happel, N., and Jeltsch, A. (2014). Substrate specificity analysis and novel substrates of the protein lysine methyltransferase NSD1. Chem. Biol. *21*, 226–237. https://doi.org/10.1016/j.chembiol.2013.10.016.

Lanouette, S., Davey, J.A., Elisma, F., Ning, Z., Figeys, D., Chica, R.A., and Couture, J.F. (2015). Discovery of substrates for a SET domain lysine methyltransferase predicted by multistate computational protein design. Structure *23*, 206–215. https://doi.org/10.1016/J.STR.2014.11.004.

Liu, H., Galka, M., Mori, E., Liu, X., Lin, Y.f., Wei, R., Pittock, P., Voss, C., Dhami, G., Li, X., et al. (2013). A method for systematic mapping of protein lysine methylation identifies new functions for HP1β in DNA damage repair. Mol. Cell *50*, 723. https://doi.org/10.1016/J.MOLCEL.2013.04.025.

Luo, L., Pappalardi, M.B., Tummino, P.J., Copeland, R.A., Fraser, M.E., Grzyska, P.K., and Hausinger, R.P. (2006). An assay for Fe(II)/2-oxoglutarate-dependent dioxygenases by enzyme-coupled detection of succinate formation. Anal. Biochem. *353*, 69–74. https://doi.org/10.1016/J.AB.2006.03.033.

McKenna, M., Balasuriya, N., Zhong, S., Li, S.S.C., and O'Donoghue, P. (2021). Phospho-form specific substrates of protein kinase B (AKT1). Front. Bioeng. Biotechnol. *8*, 1580. https://doi.org/10.3389/FBIOE.2020.619252/BIBTEX.

Petronikolou, N., Longbotham, J.E., and Fujimori, D.G. (2020). Extended recognition of the histone H3 tail by histone demethylase KDM5A. Biochemistry *59*, 647–651. https://doi.org/10.1021/acs.biochem.9b01036.

Hornbeck, P.V., Kornhauser, J.M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res. *40*, D261–D270. https://doi.org/10.1093/NAR/GKR1122.

Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008a). Protein lysine methyltransferase G9a acts on non-histone targets. Nat. Chem. Biol. *4*, 344–346. https://doi.org/10.1038/nchembio.88.

Rathert, P., Zhang, X., Freund, C., Cheng, X., and Jeltsch, A. (2008b). Analysis of the substrate specificity of the Dim-5 histone lysine methyltransferase using peptide arrays. Chem. Biol. *15*, 5–11. https://doi.org/10.1016/J.CHEMBIOL.2007.11.013.

Rose, N.R., Woon, E.C.Y., Tumber, A., Walport, L.J., Chowdhury, R., Li, X.S., King, O.N.F., Lejeune, C., Ng, S.S., Krojer, T., et al. (2012). The plant growth regulator daminozide is a selective inhibitor of the human KDM2/7 histone demethylases. J. Med. Chem. *55*, 6639. https://doi.org/10.1021/JM300677J.

Schuhmacher, M.K., Kudithipudi, S., Kusevic, D., Weirich, S., and Jeltsch, A. (2015). Activity and specificity of the human SUV39H2 protein lysine methyltransferase. Biochim. Biophys. Acta *1849*, 55–63. https://doi.org/10.1016/J.BBAGRM.2014.11.005.

Topcu, E., and Biggar, K.K. (2019). PeSA: a software tool for peptide specificity analysis. Comput. Biol. Chem. *83*, 107145. https://doi.org/10.1016/j.compbiolchem.2019.107145.

Wei, R., Kaneko, T., Liu, X., Liu, H., Li, L., Voss, C., Liu, E., He, N., and Li, S.S.C. (2018). Interactome mapping uncovers a general role for numb in protein kinase regulation. Mol. Cell. Proteomics *17*, 2216–2228. https://doi.org/10.1074/MCP.RA117.000114.

Yang, Z., Qiu, Q., Chen, W., Jia, B., Chen, X., Hu, H., He, K., Deng, X., Li, S., Tao, W.A., et al. (2018). Structure of the arabidopsis JMJ14-H3K4me3 complex provides insight into the substrate specificity of KDM5 subfamily histone demethylases. Plant Cell *30*, 167–177. https://doi.org/10.1105/tpc.17.00666.

Yu, W., Eram, M.S., Hajian, T., Szykowska, A., Burgess-Brown, N., Vedadi, M., and Brown, P.J. (2014). A scintillation proximity assay for histone demethylases. Anal. Biochem. *463*, 54–60. https://doi.org/10.1016/J.AB.2014.06.023.