

# Affinity Purification of Methyllysine Proteome by Site-Specific Covalent Conjugation

Rui Wang,<sup>†,‡</sup> Mei Huang,<sup>†</sup> Linting Li,<sup>§</sup> Tomonori Kaneko,<sup>†</sup> Courtney Voss,<sup>†</sup> Liang Zhang,<sup>\*,‡,||</sup> Jiang Xia,<sup>\*,§,||</sup> and Shawn S. C. Li<sup>\*,†</sup>

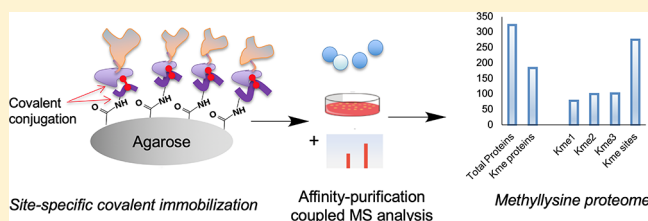
<sup>†</sup>Department of Biochemistry and Siebens-Drake Medical Research Institute, Schulich School of Medicine and Dentistry, Western University, London, Ontario N6A 5C1, Canada

<sup>‡</sup>Department of Biomedical Sciences, City University of Hong Kong, Kowloon, Hong Kong Special Administrative Region, China

<sup>§</sup>Department of Chemistry, The Chinese University of Hong Kong, Shatin, Hong Kong Special Administrative Region, China

## Supporting Information

**ABSTRACT:** A basic but critical step in targeted proteomics by mass spectrometry is the separation of the targeted proteins from the complex mixture of the whole proteome by affinity purification. The bait protein is usually immobilized on the surface of a solid support to enable affinity-based purification of the targeted proteome. Here, we developed a site-specific covalent immobilization of the bait protein through affinity-guided covalent coupling (AGCC) of a single cysteine residue of an SH2 domain (utilized as an affinity tag for the protein target) with an engineered ligand peptide. Site-specific covalent immobilization of a methyllysine-binding protein HP1 $\beta$  chromodomain on the agarose resin was used to purify the methyllysine proteome from the whole-protein mixture. This new bait immobilization led to a notably low background in the affinity purification step, markedly outperforming the conventional (His)<sub>6</sub> tag–nickel nitrilotriacetic acid (Ni-NTA) immobilization method. Subsequent analysis of the purified proteome identified 275 lysine methylated sites and 184 methylated proteins from 332 HP1 $\beta$  CD-binding proteins, including 30 novel methylated proteins. This work demonstrates that a robust site-specific covalent protein immobilization method is well-suited for proteomic analysis of low-abundance proteins. This method also enables the identification of new methylated proteins and methylation sites in the methyllysine proteome.



Affinity purification coupled with mass spectrometry (AP/MS) has gained increasing momentum in proteomics research, given its robust applicability in dissecting protein complexes and profiling post-translational modifications (PTMs). In general, the affinity purification step requires immobilizing a bait protein on the surface of various types of solid supports.<sup>1,2</sup> Conventionally, immobilization of proteins can be achieved through noncovalent interactions, such as hydrophobic, polar, or ionic interactions.<sup>3–5</sup> However, noncovalent methods suffer from rather weak interactions and risk possible protein loss during washing. Alternatively, antibodies or affinity reagents anchored to surfaces can be used to purify endogenous proteins or proteins expressed with epitope tags, such as His, Flag, or hemagglutinin (HA) peptide tags, glutathione S-transferase (GST) or green fluorescent protein (GFP) protein tags, etc.<sup>6,7</sup> However, the proteins are still not covalently anchored, and the varying binding efficiencies and specificities of different antibodies or affinity reagents are still major issues.<sup>2,8</sup> Direct covalent conjugation is certainly preferred for stable protein immobilization. However, conventional reactions in the bioconjugation toolbox rely on the chemical reactivity of natural amino acids and thus lack site selectivity.<sup>9–11</sup> For example, surface-displayed *N*-hydroxysuccinimide (NHS) ester will react with all the available lysine residues, giving a mixed population of differently oriented

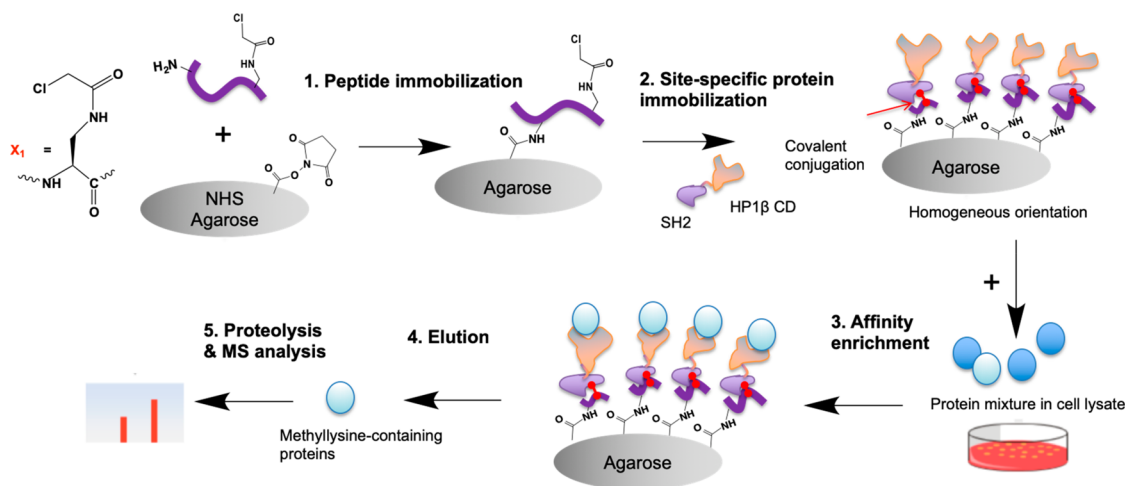
proteins on the surface, with some having their active sites shielded or blocked (Figure 1). Therefore, in some cases, random covalent immobilization may shield some of the protein binding, cause signal loss, and give suboptimal data quality.<sup>10,11</sup>

We have developed a strategy called affinity-guided covalent conjugation (AGCC) that enables efficient site-specific covalent conjugation of engineered and natural proteins through protein-binding domains such as PDZ, SH2 and SH3.<sup>12,13</sup> In AGCC, the cysteine– $\alpha$ -chloroacetyl conjugation reaction proceeds rapidly when both groups are positioned in close proximity through protein–peptide binding interaction, but very slowly when they are out of reach of each other.<sup>14,15</sup> This feature therefore confers different chemical reactivity to cysteine residues at different locations and yields site-selective conjugation. Here we envision that the AGCC reaction will facilitate stable tethering of bait proteins on the surface of the solid support for highly efficient AP/MS proteomics.<sup>16</sup> Specifically, we showcase the immobilization of the heterochromatin protein 1 $\beta$  chromodomain (HP1 $\beta$  CD),<sup>17</sup> which is

Received: June 21, 2018

Accepted: November 5, 2018

Published: November 5, 2018



**Figure 1.** Schematic illustration of affinity-guided covalent coupling (AGCC)/mass spectrometry strategy and its application to identification of methyllysine-containing proteins from the proteome. A step-by-step protocol can be found in [Supporting Information](#).

known to bind methyllysine and used for profiling of the lysine methylation proteome.<sup>18–20</sup> Specifically, we employed the site-specific covalent conjugation reaction between a phospholipase C  $\gamma$ 1 C-terminal SH2 domain (PLC $\gamma$ 1-c-SH2)<sup>21</sup> and its engineered ligand peptide PGFpYXEAN, derived from the X–Y linker of PLC $\gamma$ 1 (where X is an unnatural amino acid with an  $\alpha$ -chloroacetyl group at its side chain). PGFpYXEAN was immobilized on NHS-activated agarose resin through the N-terminal amino group of the linker. Then HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein was covalently immobilized by reaction with PGFpYXEAN through the SH2 domain, allowing the entire HP1 $\beta$  CD domain to be displayed on the surface (Figure 1). The HP1 $\beta$  CD-coated agarose was then used to pull down proteins with methyllysine modification from cell lysate prior to identification by mass spectrometry.

## RESULTS AND DISCUSSION

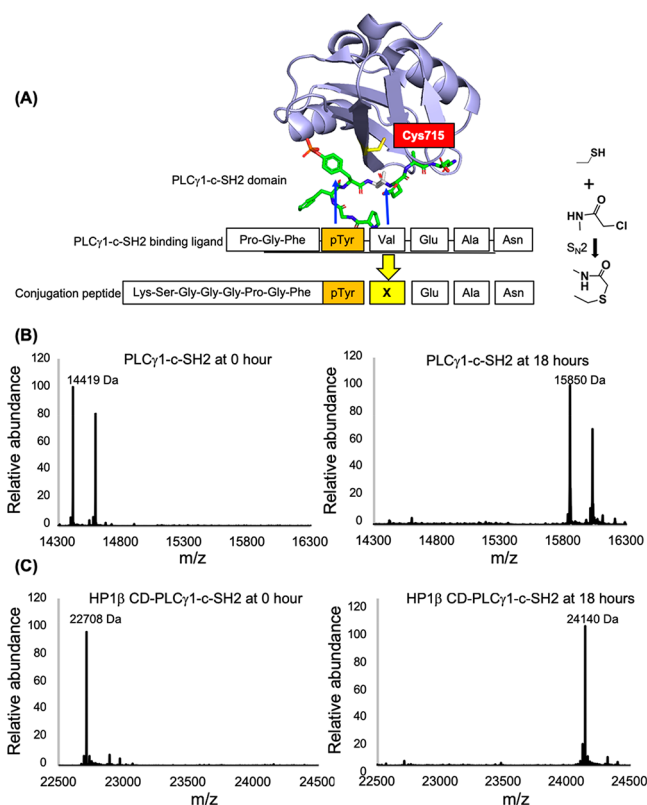
**Design of Site-Specific Covalent Conjugation Reaction.** PLC $\gamma$ 1-c-SH2 domain binds the phosphotyrosine peptide PGFpYVEAN from the X–Y linker of PLC $\gamma$ 1 (PDB ID 4K45) with high affinity (Figure 2A).<sup>22</sup> In this interaction, PLC $\gamma$ 1-c-SH2 contains a cysteine residue that is close to the peptide-binding site and points to the pY + 1 position. We designed a derivative of PGFpYVEAN by replacing pY + 1 valine with an unnatural amino acid X, (2*S*)-2-amino-3-[( $\alpha$ -chloroacetyl)amino]propionic acid. An extended linker (KSGGG) is added to the N-terminus of PGFpYVEAN, extending the space between the solid support and the bait protein. The final ligand peptide KSGGGPGFpYXEAN (Figure S1), named BC-k, contains amino groups at the N terminus for *N*-hydroxysuccinimide (NHS)-mediated conjugation with the solid support, an NHS resin.<sup>23</sup> Mixing the recombinantly expressed and purified PLC $\gamma$ 1-c-SH2 and the ligand peptide in phosphate-buffered saline (PBS) yields a stable covalent cross-link between the peptide and the protein, as shown by the mass of the product (Figure 2B).

The HP1 $\beta$  chromodomain (HP1 $\beta$  CD) is known to have an affinity for methylated lysine 9 of histone H3 (H3K9), as well as methylated lysine residues on various other proteins.<sup>24,25</sup> Therefore, HP1 $\beta$  CD has served as a robust tool to enrich methylated proteins and peptides for proteomic analysis. Therefore, we aim to employ AGCC to efficiently display HP1 $\beta$  CD on a solid surface for affinity purification. We first

designed, expressed, and purified recombinant HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein and examined its reaction with the BC-k ligand peptide (Figure 2C, Figure S2). Incubation of PLC $\gamma$ -c-SH2 (MW 14 549) or HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 (MW 22 708) with conjugation ligand BC-k (MW 1466.56) in excess overnight (18 h) gave their covalently coupled complexes (MW 15 850 and 24 140, respectively) as confirmed by liquid chromatography/mass spectrometry (LC/MS). This indicates that the PLC $\gamma$ 1-c-SH2 domain in the HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein maintained its correct folding and pTyr-peptide recognition property. This is consistent with the modularity of both PLC $\gamma$ 1-c-SH2 and HP1 $\beta$  CD domains.

The kinetics of the conjugation reaction was then scrutinized in greater detail. Bis-Tris polyacrylamide gel electrophoresis (PAGE) clearly indicated an irreversible covalent linkage between PLC $\gamma$ 1-c-SH2 or HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 and the BC-k peptide (Figure S3A). As expected, HP1 $\beta$  CD protein alone cannot react with the BC-k peptide, confirming the conjugation reaction is specific between the SH2 protein and the BC-k peptide. Increasing the peptide-to-protein ratio and reaction time resulted in increased conjugation yields, to more than 60% conversion. Serendipitously, we found that the presence of imidazole in the reaction solution promoted the conjugation reaction. At an imidazole concentration of 300 mM, HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 reacted with peptide BC-k more efficiently (80% completion) as well as PLC $\gamma$ 1-c-SH2 (90% completion) at 1:1 protein-to-peptide ratio (Figure S3B). Other molecules, such as reducing agents, did not affect the reaction. The promoting effect of imidazole has been well-documented in some protein ligation reactions.<sup>26,27</sup> It has been proposed that the ester reacts first with imidazole to form a highly reactive acyl-imidazole intermediate, which subsequently reacts rapidly with cysteine thiol.<sup>26</sup> A similar mechanism may explain the imidazole-promoted cysteine–chloroacetyl conjugation reaction: substitution of chlorine with imidazole may form a reactive acetyl-imidazole intermediate to promote the cysteine nucleophilic attack.

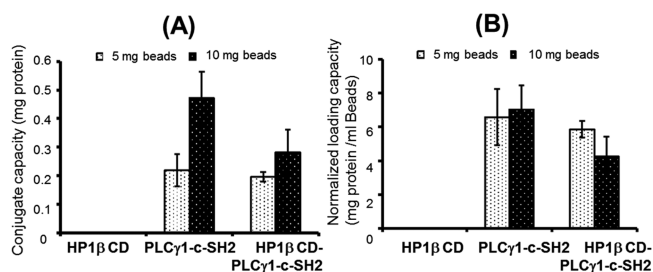
**Covalent Immobilization of HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 through the BC-k Peptide.** Next, we sought to covalently immobilize HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 through the BC-k peptide onto NHS-functionalized agarose resin. The BC-k peptide was first covalently coupled on NHS-functionalized agarose resin through reaction of N-terminal amino groups



**Figure 2.** Covalent reaction between PLC $\gamma$ 1-c-SH2 domain and phosphotyrosine peptide. (A) Model depicting binding of the conjugation peptide BC-k (KSGGGPGFpYXEAN) derived from the X–Y linker to the PLC $\gamma$ 1-c-SH2 domain based on the structure in PDB ID 4K45. There is a desired reactive Cys715 residue exposed on the surface of PLC $\gamma$ 1-c-SH2 domain and around its pTyr binding pocket. The  $\alpha$ -chloroacetyl moiety is incorporated to the pTyr + 1 position in BC-k peptide, allowing it to be specifically and covalently conjugated to PLC $\gamma$ 1-c-SH2 domain via affinity-guided conjugation reaction (depicted on the right). (B, C) MS spectra of products from affinity-guided conjugation reaction between conjugation peptide and (B) PLC $\gamma$ 1-c-SH2 domain or (C) HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2, before and after 18 h of incubation. The masses of PLC $\gamma$ 1-c-SH2, HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2, and their respective conjugation complexes with BC-k are identified on the spectra. The mass spectra were collected on a Bruker Reflex IV matrix-assisted laser desorption ionization time-of-flight (MALDI TOF) mass spectrometer.

with NHS moieties (the peptide was used in excess to ensure maximal coverage of the NHS groups, and the unreacted NHS groups on the resin were quenched afterward). Then HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein was added to BC-k-coated resin for covalent conjugation reaction. We quantified that 5 or 10 mg of BC-k-functionalized resin could immobilize 0.20 or 0.30 mg of HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 protein and 0.22 or 0.49 mg of PLC $\gamma$ 1-c-SH2 protein, respectively (Figure 3A). After normalizing the protein loading to 10 mg starting resin in dry weight, we achieved immobilization of 4 mg/mL HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 or 7 mg/mL PLC $\gamma$ 1-c-SH2 (Figure 3B). Such loading capacity is on par with commercially available GST or His beads. These results indicate that the SH2 domain in the fusion protein HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 efficiently reacts with the BC-k peptide displayed on the surface of agarose.

Next, we investigated whether the surface-displayed HP1 $\beta$  CD protein is capable of binding to the methylated lysine-containing proteins. As HP1 $\beta$  CD is known to recognize the

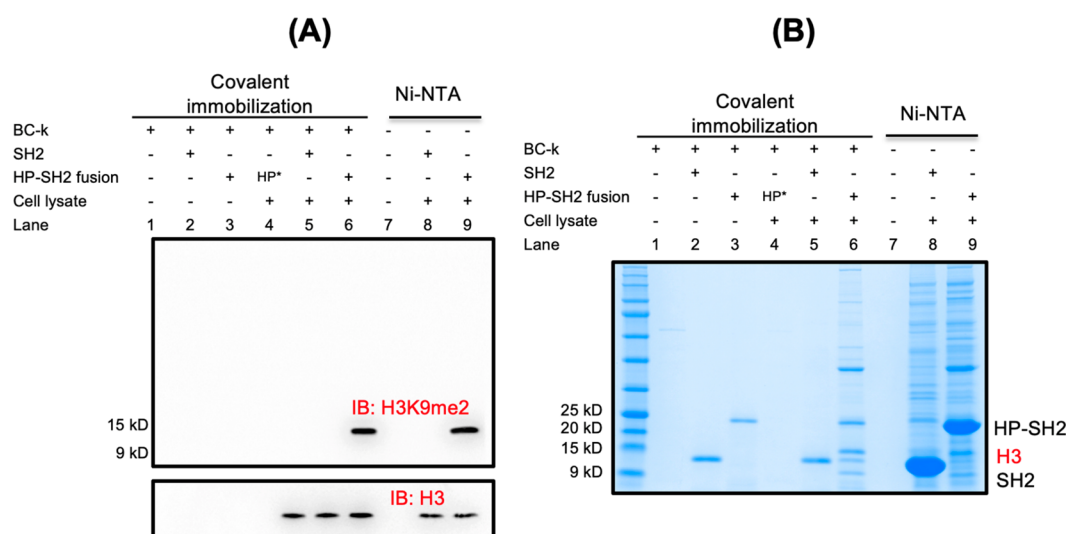


**Figure 3.** Covalent immobilization of HP1 $\beta$  CD on NHS-agarose through BC-k–PLC $\gamma$ 1-c-SH2 conjugation reaction and preliminary pull-down experiment. (A) Quantifying the loading capacity of the covalent immobilization reaction. The amount of the conjugated protein was quantified as milligrams of total protein input (before conjugation) minus milligrams of free protein in supernatant (after conjugation). (B) Normalized loading capacity of the two-step conjugation procedure. The fusion protein largely retained the reactive efficiency with BC-k-coated resin as compared with PLC $\gamma$ 1-c-SH2 alone, except for a slight decrease of 20% in the 10 mg resin case.

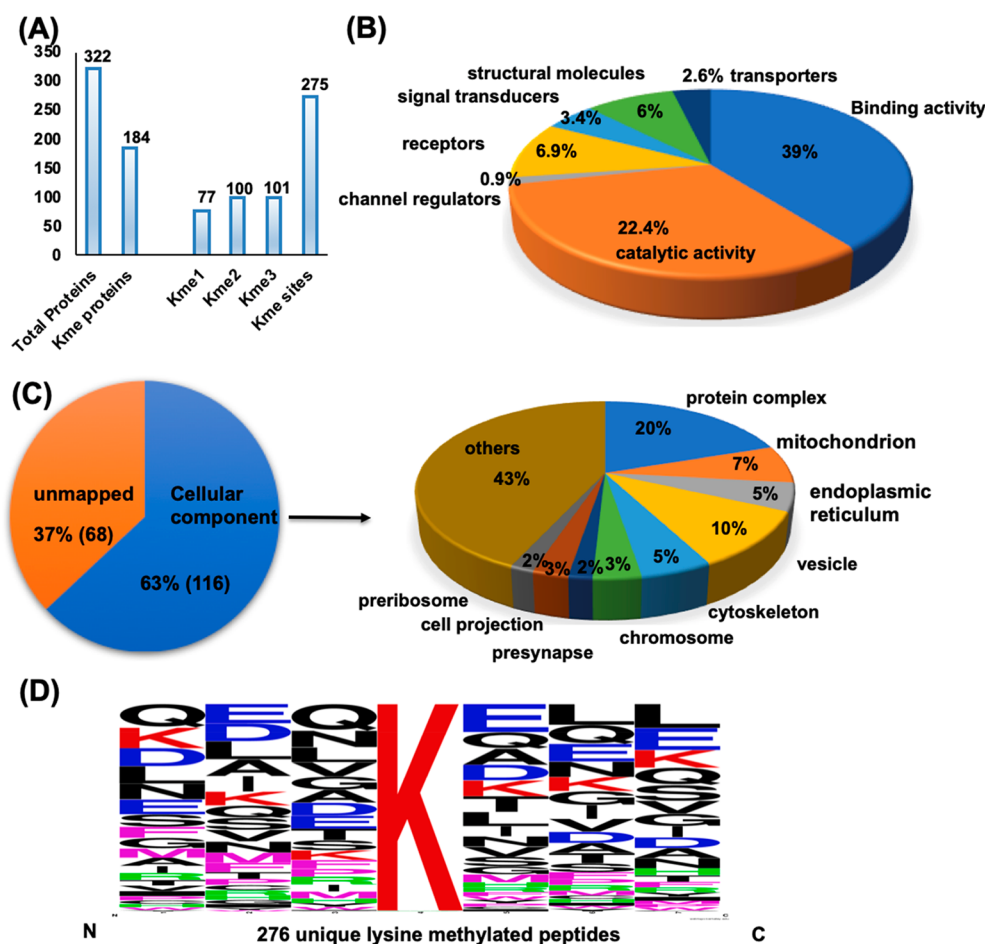
histone proteins H3K9 and H3K23 in a methylation-dependent manner, we tested whether surface-displayed HP1 $\beta$  CD (through HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion) could pull down methylated histone protein H3K9me2. MCF-7 cells were homogenized in RIPA buffer and the lysates were incubated overnight with conjugated agarose resin at 4 °C. After washing, the bound proteins were eluted and resolved by Bis-Tris 10% PAGE. The gel was stained by Coomassie blue or silver staining and probed with antibodies specific for H3K9me2 or H3, respectively. We observed that H3K9me2 protein could be successfully pulled down if the immobilization procedure included both BC-k peptide and HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein (Figure 4A,B, lane 6). In contrast, H3K9me2 could not be purified when the agarose resins only displayed the BC-k peptide or the PLC $\gamma$ 1-c-SH2 fusion protein (Figure 4A,B, lanes 4 and 5). This indicates that the pulldown is specifically through interaction between surface-immobilized HP1 $\beta$  CD and H3K9me2. In addition, we utilized nickel nitrilotriacetic acid (Ni-NTA) agarose to anchor His-tagged HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein and performed a similar pulldown experiment. Western blot assay demonstrated a similar level of purified H3K2me2 (Figure 4A,B, lane 9) compared to the AGCC method (Figure 4A,B, lane 6). However, Coomassie dye staining showed a vast difference whereby the AGCC covalent immobilization gave a much lower background, indicating that the resin coated with HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 minimizes nonspecific interactions with lysate proteins. This is probably because AGCC leads to much more homogeneous orientation of HP1 $\beta$  CD. Furthermore, under the denaturing conditions of the gel electrophoresis, the SH2 protein and the HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2 fusion protein are also eluted from the Ni-NTA/His-mediated purification. On the contrary, all the covalent immobilization conditions (lanes 1–6 in Figure 4B) showed much lower amount of the immobilized proteins, indicating that covalent anchoring effectively prevents the leakage of the proteins displayed on the surface. This offers a great advantage to reduce the level of noise peptides in the following mass spectrometric analysis.

**Profiling the Methylome in MCF-7 Cells by the Covalent Immobilization Strategy.** Next we investigated whether the AGCC-displayed HP1 $\beta$  CD can facilitate profiling of proteins containing methylated lysine through interaction





**Figure 4.** Pulldown of methylated histone proteins by covalently immobilized HP1 $\beta$  CD protein resolved by Bis-Tris PAGE and probed by (A) Western blotting and (B) Coomassie staining analysis. SH2 refers to the PLC $\gamma$ 1-c-SH2 protein; HP-SH2 fusion refers to HP1 $\beta$  CD-PLC $\gamma$ 1-c-SH2 fusion protein; HP\* refers to HP1 $\beta$  CD protein. The covalent immobilization group (lanes 1–6) is compared with the Ni-NTA group (lanes 7–9). The covalent immobilization group gives much lower background than the Ni-NTA group but the same level of pull-down of H3K9me2 protein.



**Figure 5.** Bioconjugation AP/MS-based enrichment analysis. (A) Numbers of identified proteins, lysine-methylated (Kme) proteins, and lysine-methylated sites based on bioconjugation AP/MS. (B, C) Classification of identified lysine-methylated proteins based on their cellular functions (B) and intracellular locations (C) from Gene Ontology analysis. The 184 identified lysine-methylated proteins were categorized according to their known or predicted functions or cellular localizations. Protein with no annotations in Gene Ontology (left of panel C) were defined as unmapped proteins (37%). (D) Motif representation of identified lysine methylated peptides along with a consensus logo (frequency plot of seven amino acids flanking lysine), generated by Weblogo. N and C denote the N- and C-termini of the peptides.

with HP1 $\beta$  CD proteins from the lysate of MCF-7 cells. The affinity-enriched proteins from MCF-7 cell lysate were resolved on sodium dodecyl sulfate (SDS)–PAGE. The gel was then divided into six fractions before in-gel digestion and mass spectrometric analysis. This led to the identification of 322 proteins associated with HP1 $\beta$  CD (Table S1), of which 184 were lysine-methylated (Kme, Figure 5A and Table S2). From these methylated proteins, we identified 275 different lysine methylation sites in total, including 101 trimethylation sites, 100 dimethylation sites, and 77 monomethylation sites. Upon comparison to the 2977 known methylated proteins (PhosphoSitePlus),<sup>28</sup> 154 (84%) of the methyllysine-containing proteins that we identified have been reported to be methylated, while 30 (16%) were novel methylated proteins. In a previous effort, we utilized recombinant HP1 $\beta$  CD that was tethered to the resin through nonspecific covalent conjugation.<sup>18</sup> In that study, 109 HP1 $\beta$  CD-binding proteins were identified, among which 29 were methylated on 40 lysine residues.<sup>18</sup> Apparently, the AGCC strategy allowed us to discover more lysine-methylated proteins and sites, providing an effective tool for affinity enrichment and profiling of the methyllysine proteome by MS.

We then analyzed the identified lysine-methylated proteins for their subcellular localizations with Gene Ontology (GO) term analysis. Out of the 184 lysine-methylated proteins that we identified, cytoplasmic proteins constitute the majority (Figure 5B), in addition to ones that are distributed in a wide range of cellular organelles (Figure 5C). Previous studies have reported that methylated lysine sites are associated with E/L/MK motifs.<sup>29</sup> We performed motif analyses of the peptides containing methyllysine that were identified by AGCC-anchored HP1 $\beta$  CD (Figure 5D). Although we did observe LK and KE motifs in the identified methylated lysine peptides, they do not constitute a dominant population. As a result, there were no obvious motifs for methylation sites. The newly identified methyllysine sites through this new technology will drive further exploration of the physiological role of lysine methylation in normal or pathological states.

## CONCLUSION

Expanding bioconjugation chemistry has provided novel opportunities for specific enrichment of PTM proteins that are less abundant. In the present work, we developed a new protein immobilization strategy based on affinity-guided covalent conjugation (AGCC) that takes advantage of the Cys-chloroacetyl nucleophilic S<sub>N</sub>2 reaction between an engineered BC-k peptide and PLC $\gamma$ 1-c-SH2 domain. The positioning of the cysteine allowed a nucleophile close to the PLC $\gamma$ 1-c-SH2 binding pocket without harming the native folding of proteins. Using this strategy, we efficiently immobilized HP1 $\beta$  CD as a bait for proteomic profiling of the proteins with methylated lysine. Our approach is highlighted by its ability to specifically and covalently conjugate the tag protein (PLC $\gamma$ 1-c-SH2) with minimal interference on the bait protein, HP1 $\beta$  CD. We envision that the AGCC strategy leads to homogeneous orientation of the bait protein while maintaining its structure and activity, an important feature that facilitates the application of bioconjugation for AP/MS study. This work showcases how the development of site-specific bioconjugation reaction drives the advancement of protein immobilization methods in AP/MS and accelerates our understanding of more epigenetic functions from PTM.

## MATERIALS AND METHODS

**Functionalizing Peptides on NHS-Activated Resin, Followed by Protein Conjugation.** Purified reactive peptide BC-k was added to NHS-activated agarose with a peptide-to-NHS group ratio of at least 2:1. The conjugation reaction lasted for 1 h at room temperature, before a centrifugation step (1000g for 1 min) to remove excess peptide, and the reaction was quenched with 1 M Tris at pH 7.5. The agarose resin was collected and washed three times with PBS. BC-k peptide-functionalized agarose resin was incubated with 1 mg of fusion protein (His–HP1 $\beta$  CD–PLC $\gamma$ 1-c-SH2) to allow the site-specific covalent conjugate reaction between the surface-immobilized reactive peptides and the SH2 fusion protein in PBS (pH 7.4) at 4 °C overnight. This step gives AGCC resins; that is, resins with HP-SH2 fusion protein covalently conjugated on the surface.

**Affinity Purification of Methyllysine Proteins.** MCF-7 cell lysate (3 mg) was mixed with 5 mg of AGCC resins (covalently functionalized with HP-SH2 fusion) and then incubated with end-to-end rotation at 4 °C overnight. After washing, the slurry was resuspended in SDS–PAGE loading buffer, and 20  $\mu$ L was loaded for PAGE analysis followed by Coomassie blue staining or Western blotting with antibodies specific to H3K9me2 and H3, respectively, to give the gel picture of the covalent immobilization group. For comparison, His-tag-labeled HP-SH2 fusion protein or PLC $\gamma$ 1-c-SH2 in the same amount were immobilized on Ni-NTA agarose and used to pull down proteins from the same amount of lysate. This gives the Ni-NTA group. The immunocomplexes were captured by prepared bioconjugated NHS resin or Ni–NTA agarose. Notably, under the condition of gel electrophoresis, all the immobilized HP-SH2 fusion protein or PLC $\gamma$ 1-c-SH2 protein were dissociated from the Ni-NTA resins, but only a small portion of proteins were dissociated from the covalent immobilization protocol, corresponding to the small population of nonspecific noncovalent immobilization.

**Tryptic Digestion and LC-MS/MS Analysis.** Immunoprecipitated proteins on beads were washed five times with cold 20 mM ammonium bicarbonate. Iodoacetamide in 20 mM ammonium bicarbonate was added to beads, and the mixture was incubated in the dark for 1 h. Digestion was performed with Promega sequencing-grade trypsin (1:50) incubated overnight at 37 °C and stopped by acidifying to a final concentration of 0.1% trifluoroacetic acid (TFA). The harvested supernatants containing the peptide mixture were desalted and dried in a SpeedVac. After desalting on C18-Zip tips and drying in the SpeedVac, the obtained samples were loaded onto a Thermo Easy-Spray analytical column (75  $\mu$ m i.d.  $\times$  500 mm) C18 column with an Easy-nLC 1000 chromatography pump. For each analysis, we reconstituted peptides in 20  $\mu$ L of 0.1% TFA and loaded 4  $\mu$ L onto the column for four times running. Peptides in each running were separated on a 125 min (5–40% acetonitrile) gradient. Mass spectra were collected on a Q-Exactive hybrid quadrupole-Orbitrap mass spectrometer coupled to an Easy-nLC 1000 system (ThermoFisher). The spectrometer was set in full MS/data-dependent-MS2 TopN mode: mass analyzer over  $m/z$  range of 400–1600 with a mass resolution of 70 000 (at  $m/z$  = 200), 35 NEC (normalized collision energy), 2.0  $m/z$  isolation window, and 15 s dynamic exclusion. Detailed experimental procedures are described in Supporting Information.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.8b02796.

Additional text with general workflow of the AGCC method and detailed experimental procedures; three figures showing MS analysis of peptide BC-k, molecular design of fusion protein, and conjugation reaction kinetics; two tables with complete lists of the 332 identified proteins and the identified lysine methylated proteins (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail liangzhang.28@cityu.edu.hk (L.Z.).

\*E-mail jiangxia@cuhk.edu.hk (J.X.).

\*E-mail sli@uwo.ca (S.S.C.L.).

### ORCID

Jiang Xia: 0000-0001-8112-7625

### Present Address

<sup>||</sup>L.Z.: City University of Hong Kong Shenzhen Research Institute, Shenzhen, Guangdong, 518057, China.

### Author Contributions

R.W. and M.H. contributed equally. S.S.C.L. and J.X. designed the experiments with input from L.Z. M.H., R.W., L.L., T.K., and C.V. performed the experiments. R.W., M.H., J.X., S.S.C.L. and L.Z. wrote the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was partially funded by grants (to S.S.C.L.) from the Canadian Institute of Health Research and the Natural Science and Engineering Council of Canada, by University Grants Committee of Hong Kong (GRF Grants 14304915, 14321116, 14306317, and AoE/M-09/12) and Food and Health Bureau (HMRF 15140052), National Natural Science Foundation of China (21628201), and CUHK (Direct Grants 4053110, 3132686, 3132687, 3132235, and 3132238) to J.X., as well as by Shenzhen Science and Technology Program (20170407162100014) and the Research Grant Council of Hong Kong (21101917) to Z.L. S.S.C.L. holds a Canada Research Chair (Tier I) in Molecular and Epigenetic Basis of Cancer.

## ■ REFERENCES

- (1) Phizicky, E.; Bastiaens, P. I.; Zhu, H.; Snyder, M.; Fields, S. *Nature* **2003**, *422*, 208–215.
- (2) Chen, Y. X.; Triola, G.; Waldmann, H. *Acc. Chem. Res.* **2011**, *44*, 762–773.
- (3) Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S. *Bioconjugate Chem.* **2008**, *19*, 786–791.
- (4) van Swieten, P. F.; Leeuwenburgh, M. A.; Kessler, B. M.; Overkleeft, H. S. *Org. Biomol. Chem.* **2005**, *3*, 20–27.
- (5) Tomizaki, K. Y.; Usui, K.; Mihara, H. *ChemBioChem* **2005**, *6*, 782–799.
- (6) Zhou, Y.; Guo, T.; Tang, G.; Wu, H.; Wong, N. K.; Pan, Z. *Bioconjugate Chem.* **2014**, *25*, 1911–1915.
- (7) Yasun, E.; Gulbakan, B.; Ocoy, I.; Yuan, Q.; Shukoor, M. I.; Li, C.; Tan, W. *Anal. Chem.* **2012**, *84*, 6008–6015.

- (8) Miller, L. W.; Cai, Y.; Sheetz, M. P.; Cornish, V. W. *Nat. Methods* **2005**, *2*, 255–257.
- (9) Darmanis, S.; Nong, R. Y.; Hammond, M.; Gu, J.; Alderborn, A.; Vanelid, J.; Siegbahn, A.; Gustafsdottir, S.; Ericsson, O.; Landegren, U.; Kamali-Moghaddam, M. *Mol. Cell. Proteomics* **2010**, *9*, 327–335.
- (10) Cha, T.; Guo, A.; Zhu, X. Y. *Proteomics* **2005**, *5*, 416–419.
- (11) Lin, P. C.; Weinrich, D.; Waldmann, H. *Macromol. Chem. Phys.* **2010**, *211*, 136–144.
- (12) Lu, Y.; Huang, F.; Wang, J.; Xia, J. *Bioconjugate Chem.* **2014**, *25*, 989–999.
- (13) Yu, Y.; Liu, M.; Ng, T. T.; Huang, F.; Nie, Y.; Wang, R.; Yao, Z. P.; Li, Z.; Xia, J. *ACS Chem. Biol.* **2016**, *11*, 149–158.
- (14) Wang, J.; Yu, Y.; Xia, J. *Bioconjugate Chem.* **2014**, *25*, 178–187.
- (15) Wang, P.; Zhang, S.; Meng, Q.; Liu, Y.; Shang, L.; Yin, Z. *Org. Lett.* **2015**, *17*, 1361–1364.
- (16) Blomster, H. A.; Imanishi, S. Y.; Siimes, J.; Kastu, J.; Morrice, N. A.; Eriksson, J. E.; Sistonen, L. *J. Biol. Chem.* **2010**, *285*, 19324–19329.
- (17) Luciani, J. J.; Depetris, D.; Missirian, C.; Mignon-Ravix, C.; Metzler-Guillemain, C.; Megarbane, A.; Moncla, A.; Mattei, M. G. *Eur. J. Hum. Genet.* **2005**, *13*, 41–51.
- (18) Liu, H.; Galka, M.; Mori, E.; Liu, X.; Lin, Y. F.; Wei, R.; Pittcock, P.; Voss, C.; Dhimi, G.; Li, X.; Miyaji, M.; Lajoie, G.; Chen, B.; Li, S. *Mol. Cell* **2013**, *50*, 723–735.
- (19) Peters, A. H.; O'Carroll, D.; Scherthan, H.; Mechtler, K.; Sauer, S.; Schofer, C.; Weipoltshammer, K.; Pagani, M.; Lachner, M.; Kohlmaier, A.; Opravil, S.; Doyle, M.; Sibilia, M.; Jenuwein, T. *Cell* **2001**, *107*, 323–337.
- (20) Nakayama, J.-i.; Rice, J. C.; Strahl, B. D.; Allis, C. D.; Grewal, S. I. *Science* **2001**, *292*, 110–113.
- (21) Groesch, T. D.; Zhou, F.; Mattila, S.; Geahlen, R. L.; Post, C. B. *J. Mol. Biol.* **2006**, *356*, 1222–1236.
- (22) Irvin, B. J.; Williams, B. L.; Nilson, A. E.; Maynor, H. O.; Abraham, R. T. *Mol. Cell. Biol.* **2000**, *20*, 9149–9161.
- (23) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.
- (24) Liu, H.; Galka, M.; Iberg, A.; Wang, Z.; Li, L.; Voss, C.; Jiang, X.; Lajoie, G.; Huang, Z.; Bedford, M. T.; Li, S. S. *J. Proteome Res.* **2010**, *9*, 5827–5836.
- (25) Sridharan, R.; Gonzales-Cope, M.; Chronis, C.; Bonora, G.; McKee, R.; Huang, C.; Patel, S.; Lopez, D.; Mishra, N.; Pellegrini, M.; Carey, M.; Garcia, B. A.; Plath, K. *Nat. Cell Biol.* **2013**, *15*, 872–882.
- (26) Fang, G.; Cui, H.; Zheng, J.; Liu, L. *ChemBioChem* **2010**, *11*, 1061–1065.
- (27) Payne, R. J.; Ficht, S.; Greenberg, W. A.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2008**, *47*, 4411.
- (28) Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J. M.; Latham, V.; Skrzypek, E. *Nucleic Acids Res.* **2015**, *43*, D512–D520.
- (29) Kaur, I.; Zeeshan, M.; Saini, E.; Kaushik, A.; Mohammed, A.; Gupta, D.; Malhotra, P. *Sci. Rep.* **2016**, *6*, 35432.