Supplementary Information

Development of nucleus-targeted histone-tail-based photoaffinity probes to profile the epigenetic interactome in native cells

Yu Wang[1†], Jian Fan[2†], Xianbin Meng[3†], Qingyao Shu[2†], Yincui Wu[1], Guo-Chao Chu[4*], Rong Ji[1], Yinshan Ye[1], Xiangwei Wu[4], Jing Shi[2], Haiteng Deng[3], Lei Liu[4*], Yi-Ming Li[1,5*]

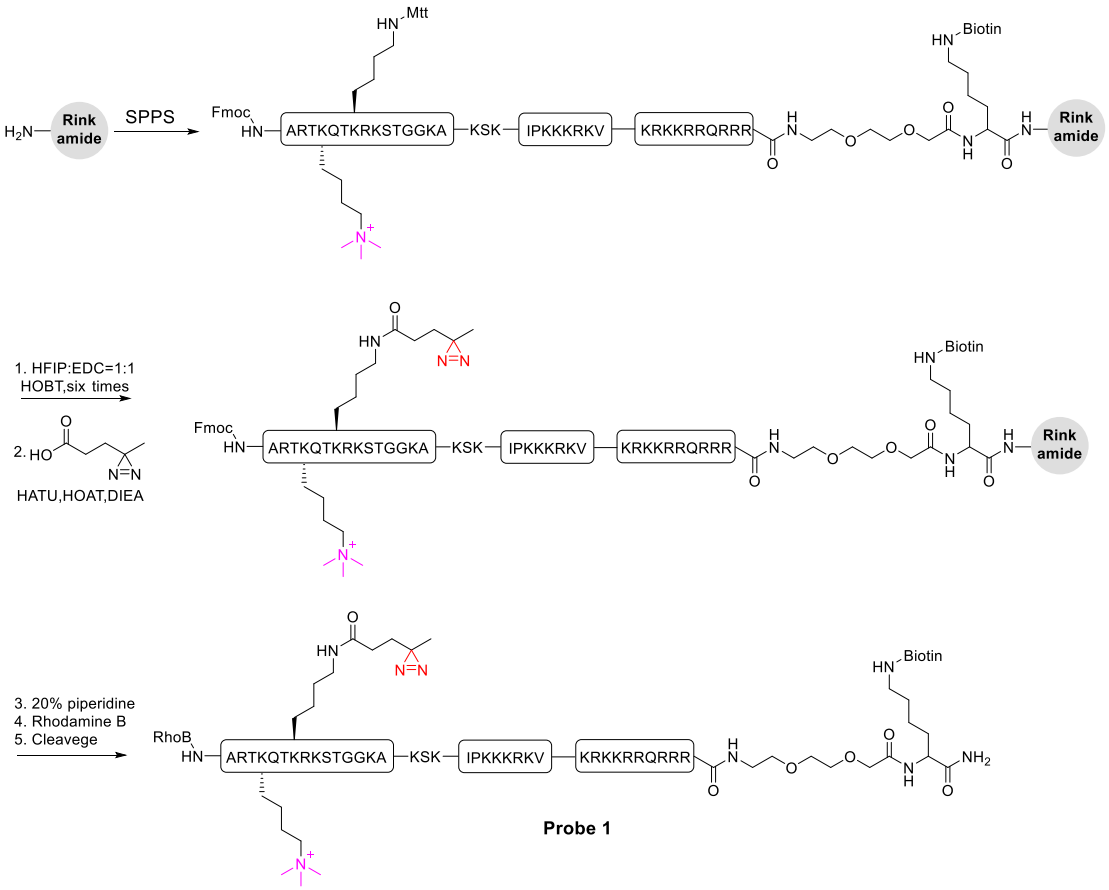
[1] School of Food and Biological Engineering, Engineering Research Center of Bio-process, Ministry of Education, Key Laboratory of Animal Source of Anhui Province, Hefei University of Technology, Hefei, 230009, China.

[2] Department of Chemistry, Hefei National Laboratory of Physical Science at Microscale, University of Science and Technology of China, Hefei 230026, China.

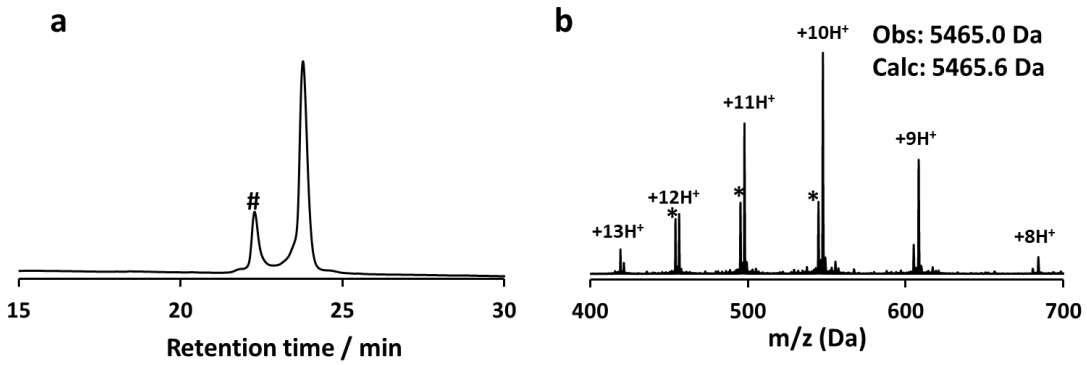
[3] MOE Key Laboratory of Bioinformatics, School of Life Sciences, Tsinghua University, Beijing 100084, China

[4] New Cornerstone Science Laboratory, Tsinghua-Peking Center for Life Sciences, Ministry of Education Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology, Department of Chemistry, Tsinghua University, Beijing 100084, China.

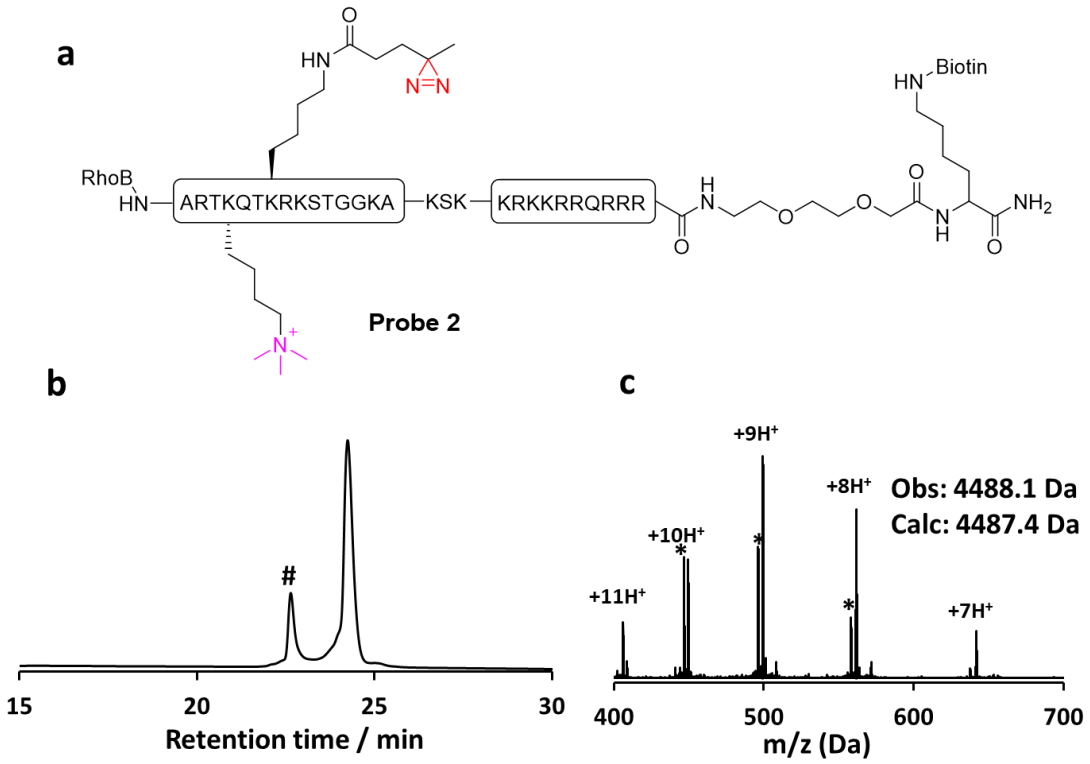
[5] Beijing Institute of Life Science and Technology, Beijing, 102206, China



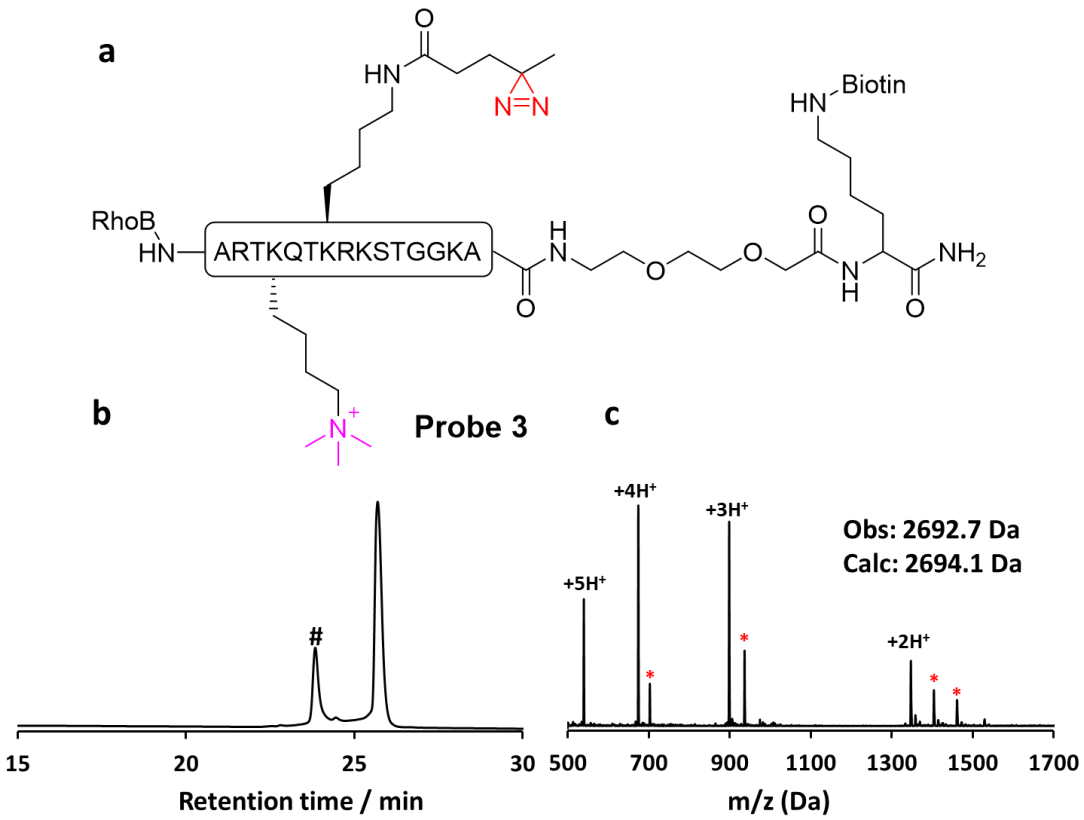
Supplementary Fig. 1. Synthetic route for probe **1**.



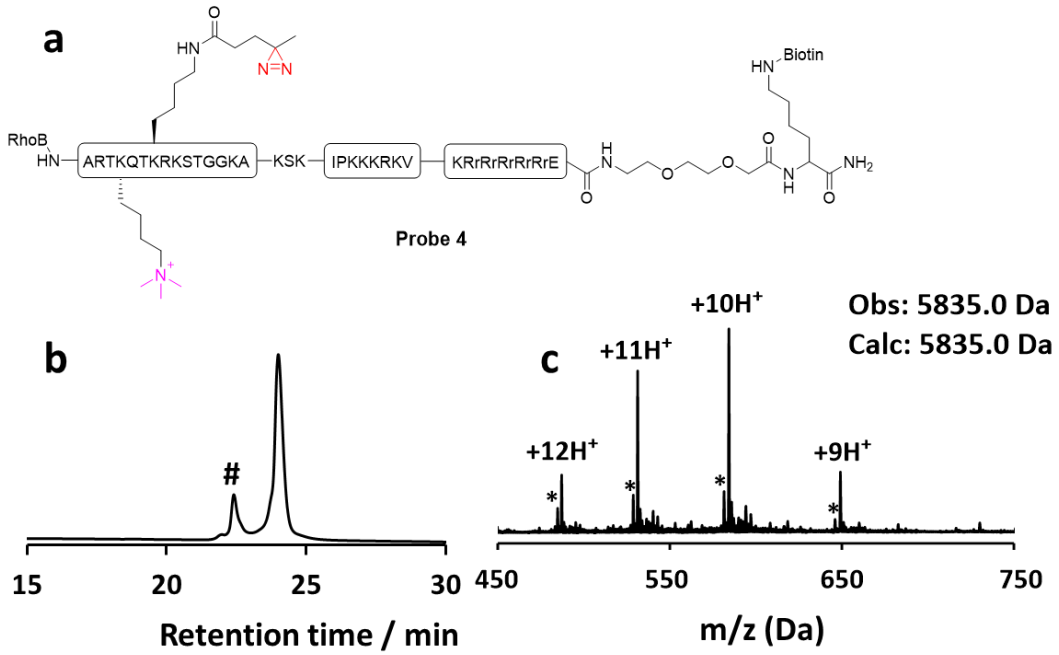
Supplementary Fig. 2. a) Analytical HPLC of probe **1**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. b) ESI-MS analysis of probe **1**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



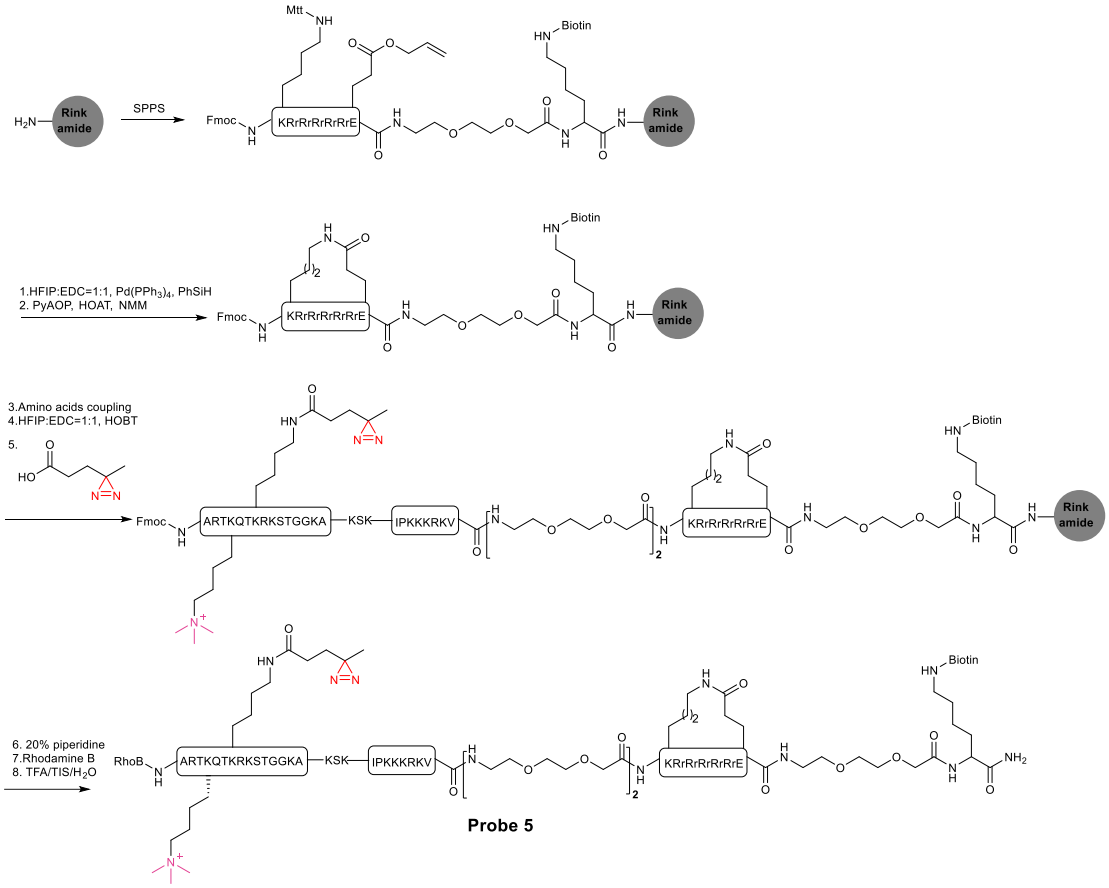
Supplementary Fig. 3. a) Chemical structure of probe **2**. b) Analytical HPLC of probe **2**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. c) ESI-MS analysis of probe **2**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



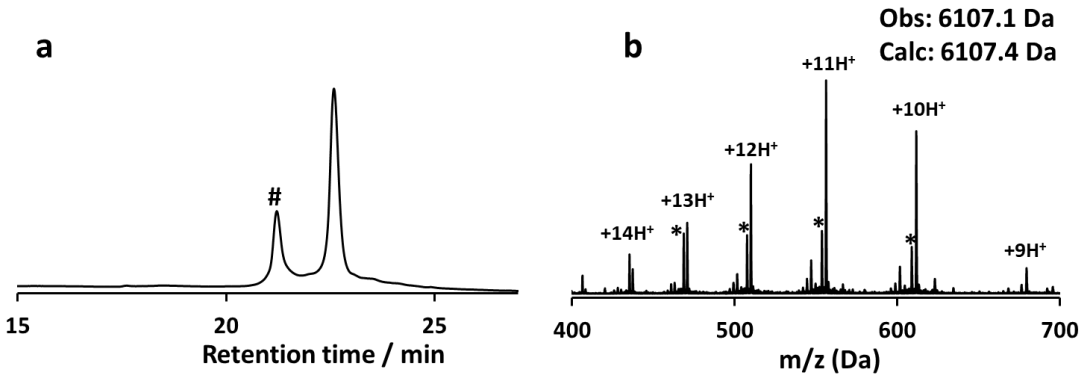
Supplementary Fig. 4. a) Chemical structure of probe **3**. b) Analytical HPLC of probe **3**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. c) ESI-MS analysis of probe **3**. Peaks in ESI-MS marked with "*" corresponded to TFA adducts.



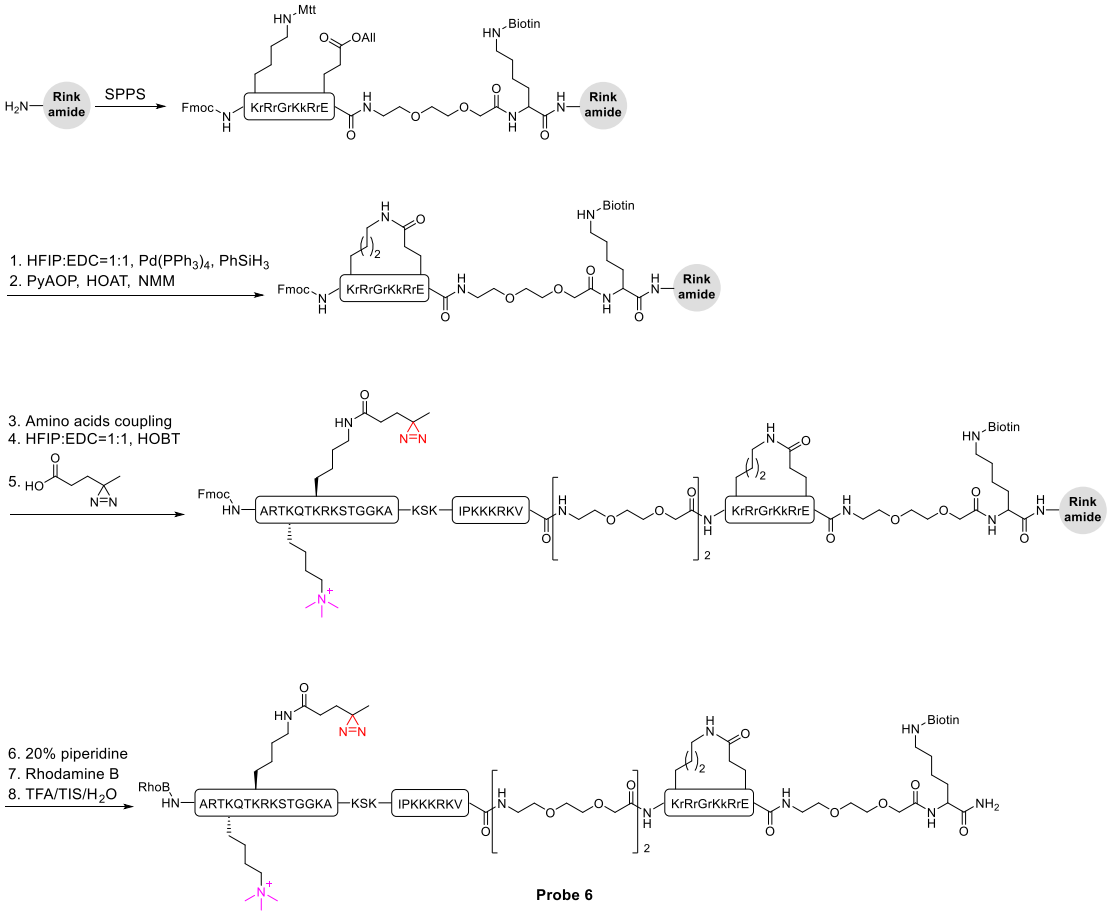
Supplementary Fig. 5. a) Chemical structure of probe **4**. b) Analytical HPLC of probe **4**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. c) ESI-MS analysis of probe **4**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



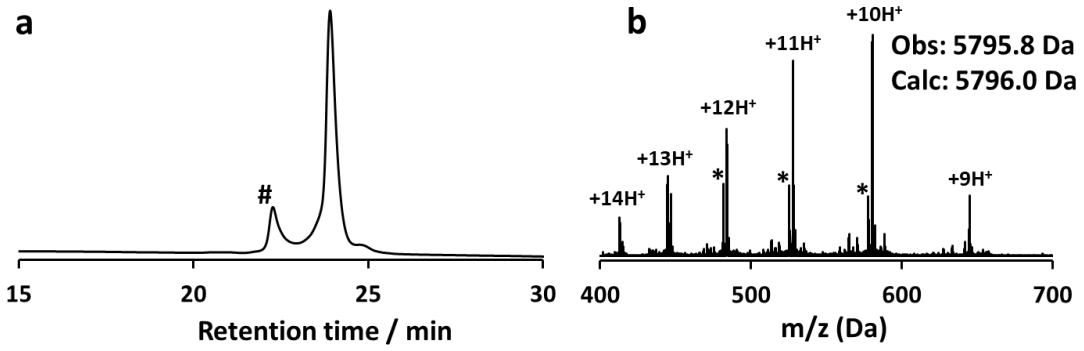
Supplementary Fig. 6. Synthetic route for probe **5**.



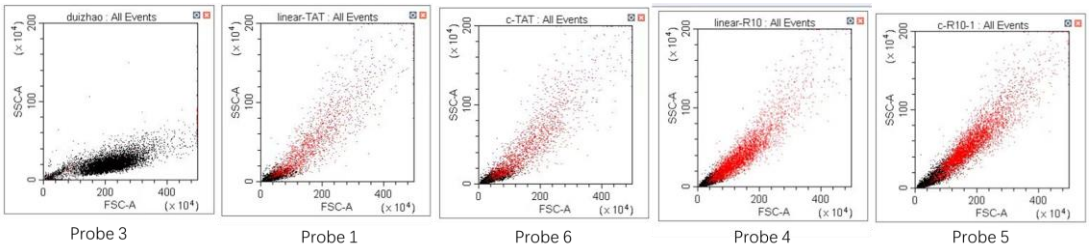
Supplementary Fig. 7. a) Analytical HPLC of probe **5**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. b) ESI-MS analysis of probe **5**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



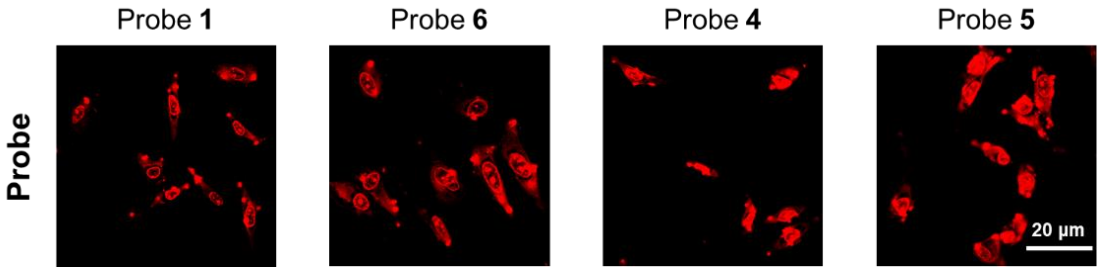
Supplementary Fig. 8. Synthetic route for probe **6**.



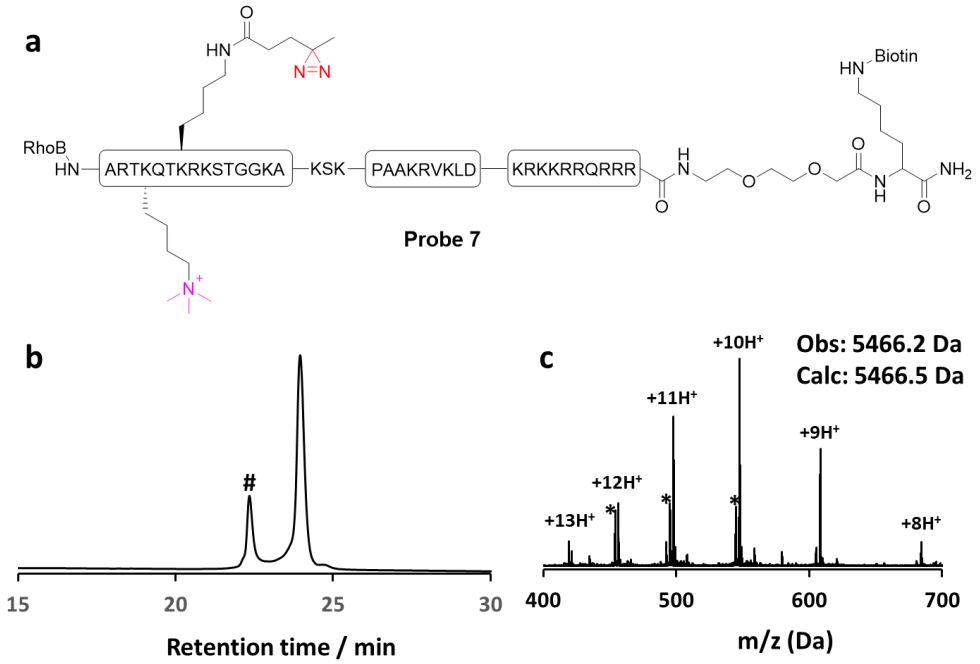
Supplementary Fig. 9. a) Analytical HPLC of probe **6**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. b) ESI-MS analysis of probe **6**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



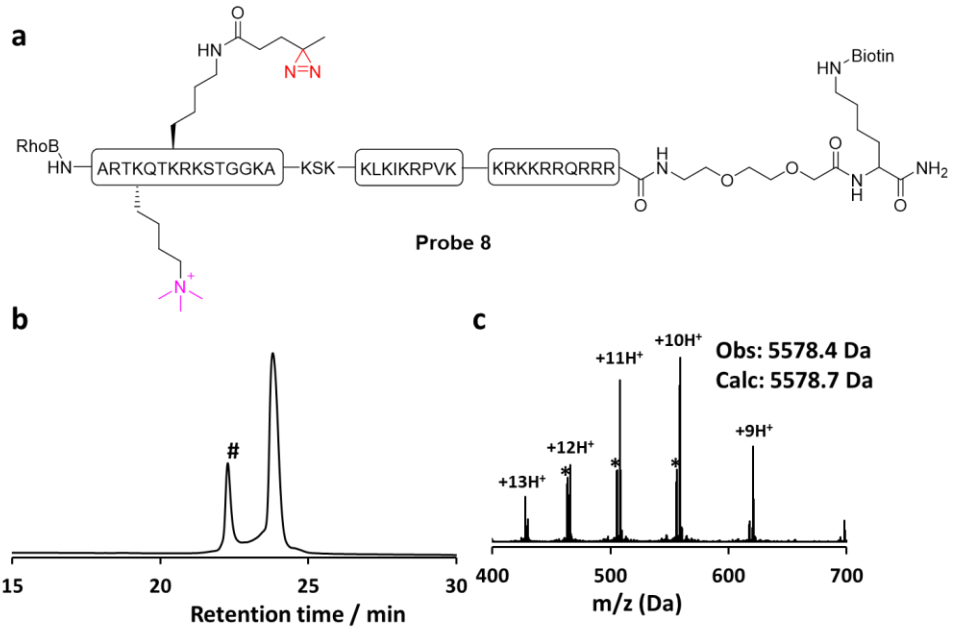
Supplementary Fig. 10. Gating strategy for flow cytometry data.



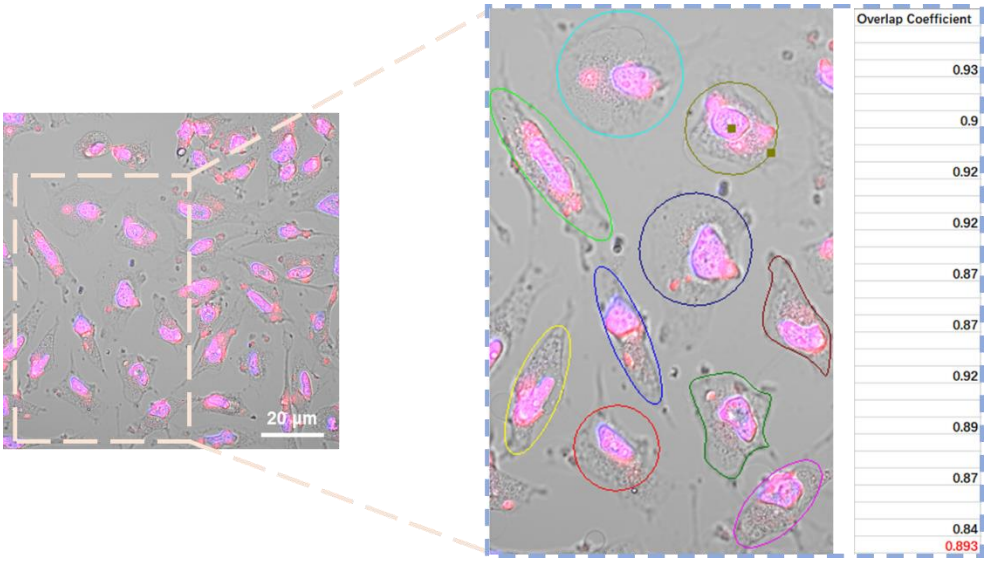
Supplementary Fig. 11. Confocal microscopy images of Hela cells treated with probe **1**, **4**-**6** for 1 h at 37℃, respectively. Probe **1**, **4**-**6** were visualized using Rho B fluorescence (red channel). Confocal microscopy images shown in (Supplementary Fig. 11.) are representative of independent biological replicates (n = 3).



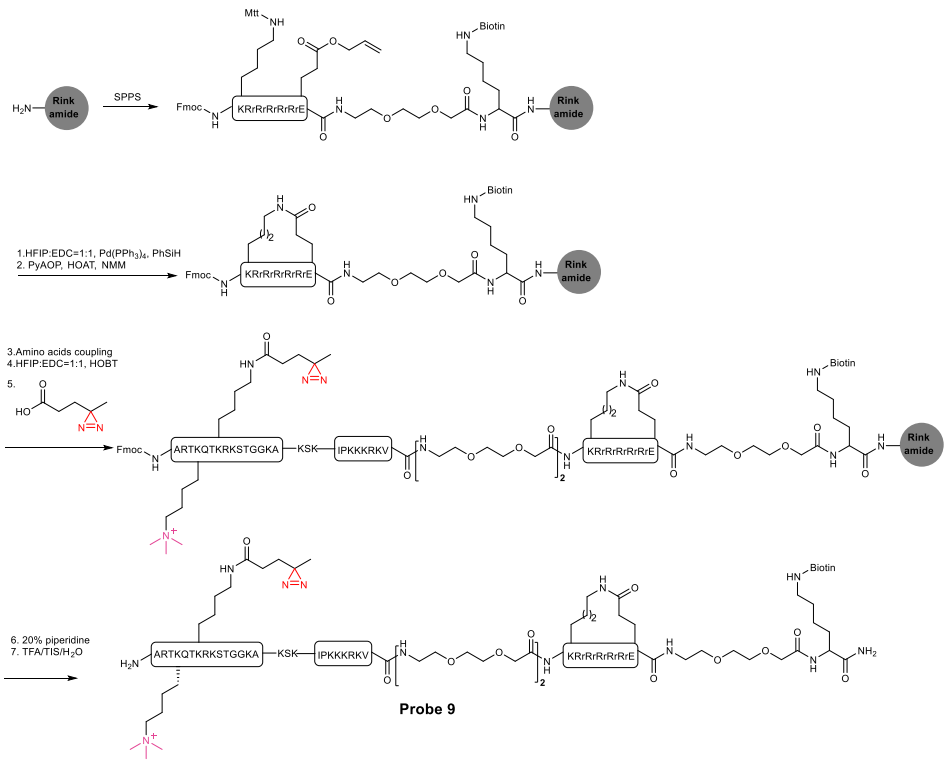
Supplementary Fig. 12. a) Chemical structure of probe **7**. b) Analytical HPLC of probe **7**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. c) ESI-MS analysis of probe **7**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



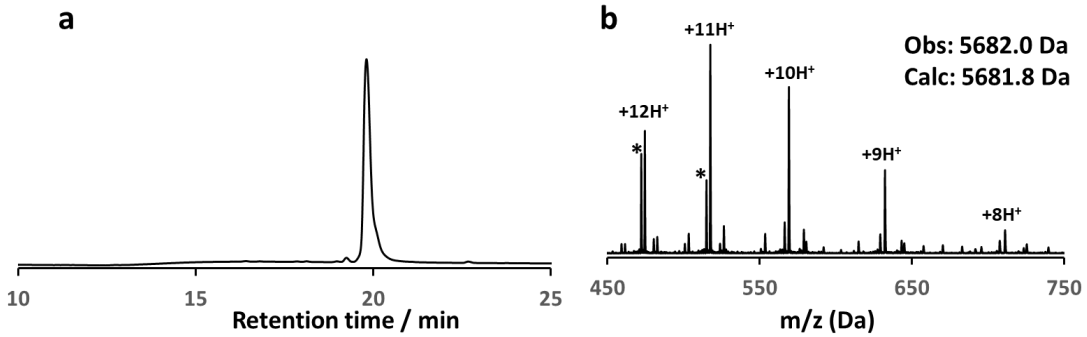
Supplementary Fig. 13. a) Chemical structure of probe **8**. b) Analytical HPLC of probe **8**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. c) ESI-MS analysis of probe **8**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



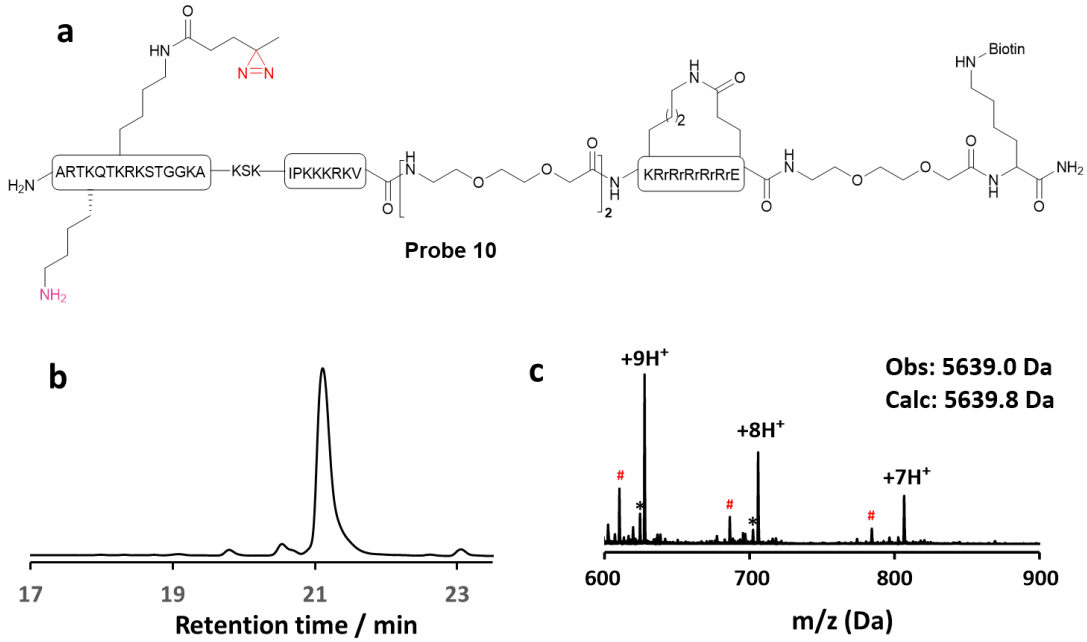
Supplementary Fig. 14. Confocal microscopy images of Hela cells treated with probe **5**. Probe **5** were visualized using Rho B fluorescence (red channel), and Hoechst 33258 was utilized for nuclear staining (blue channel). Overlap coefficient was measured by ZEN imaging software. Confocal microscopy images shown in (Supplementary Fig. 14.) are representative of independent biological replicates (n = 2).



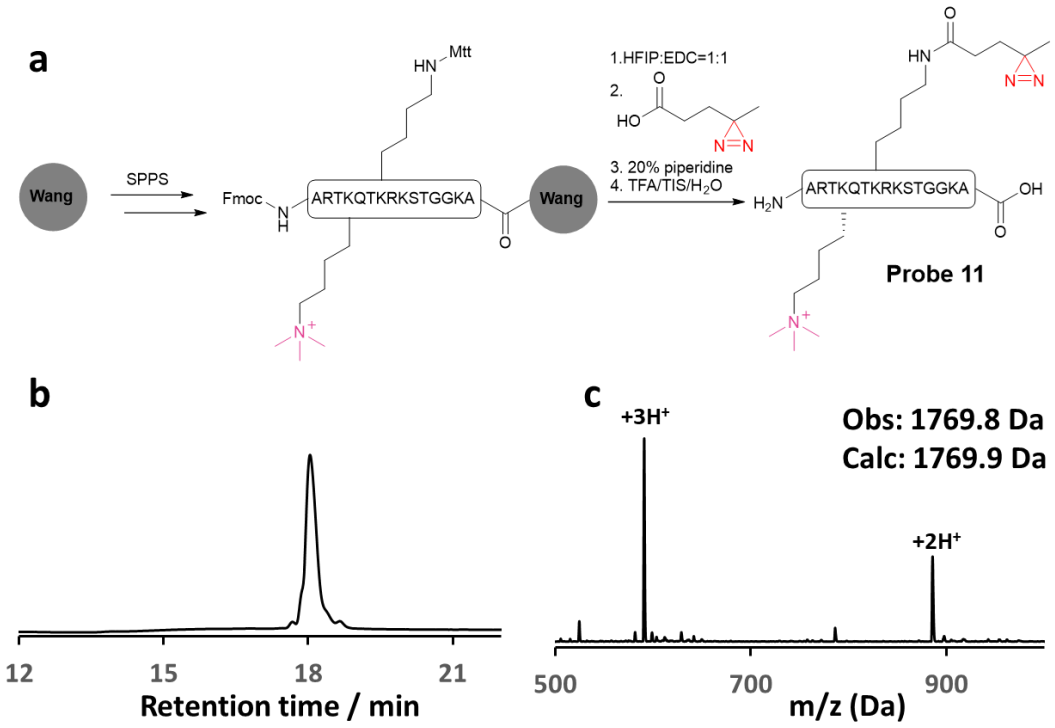
Supplementary Fig. 15. Synthetic route for probe **9**.



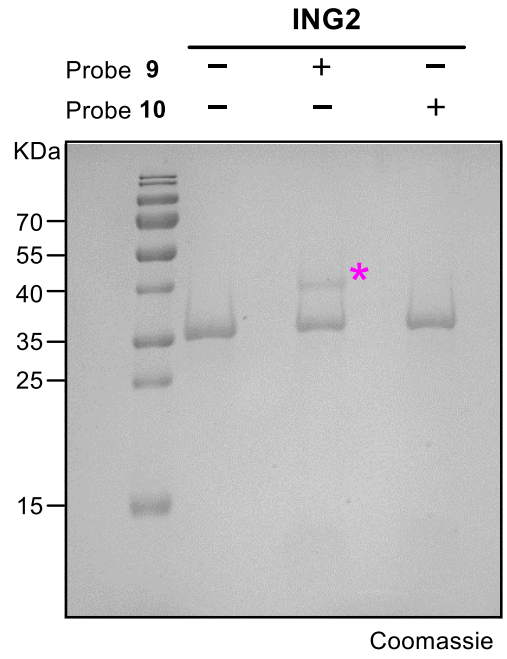
Supplementary Fig. 16. a) Analytical HPLC of probe **9**. b) ESI-MS analysis of probe **9**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



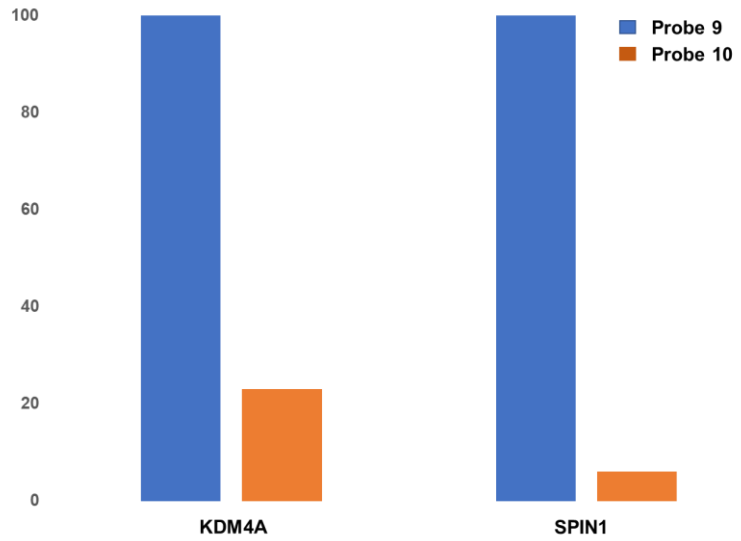
Supplementary Fig. 17. a) Chemical structure of probe **10**. b) Analytical HPLC of probe **10**. c) ESI-MS analysis of probe **10**. "*" indicates the molecular weight of diazirine after loss of N_2. # indicates the molecular weight of an arginine lost.



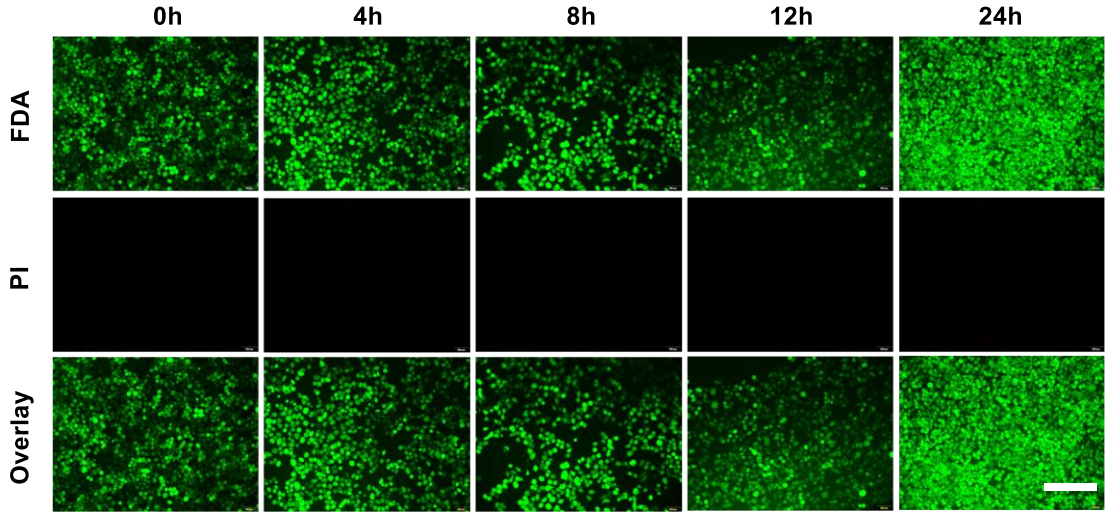
Supplementary Fig. 18. a) Synthetic route for probe **11**. b) Analytical HPLC of probe **11**. c) ESI-MS analysis of probe **11**.



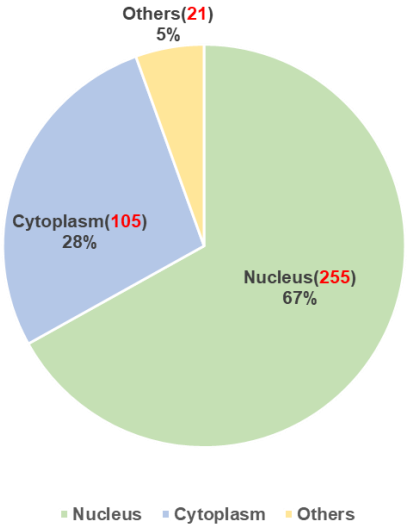
Supplementary Fig. 19. SDS-PAGE analysis of labeling of ING2 by probe **9**. Probe **10** was used as comparison. Gel image shown in (Supplementary Fig. 19) is representative of independent biological replicates (n = 2).



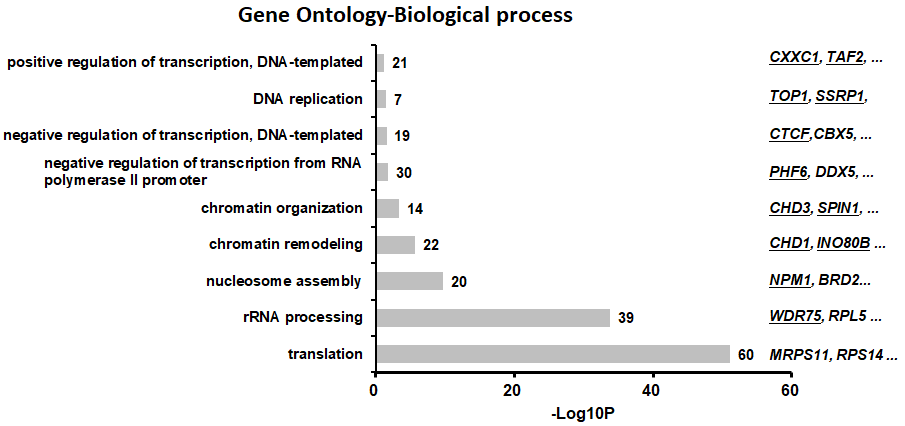
Supplementary Fig. 20. Quantification of SPIN1 and KDM4A captured by probe 9 and probe 10 (n=1).



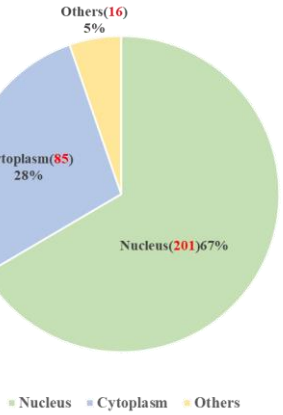
Supplementary Fig. 21. Live/dead cell assay for Hela cells treated with probe **9**. (FDA/green: live cells, PI/red: dead cells). Scale bar: 200 μm.



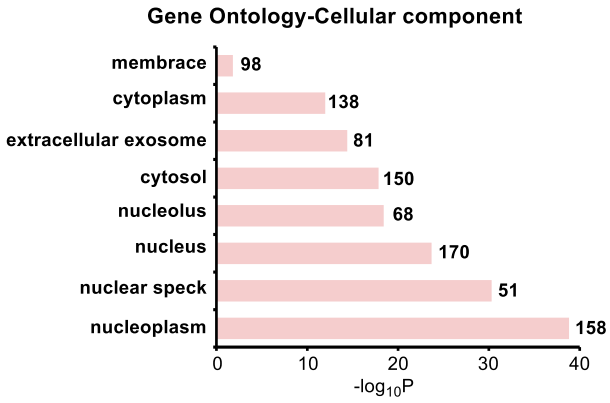
Supplementary Fig. 22. The subcellular location of proteins enriched by our H3K4me3 probe in Hela cells (from UniprotKB keywords).



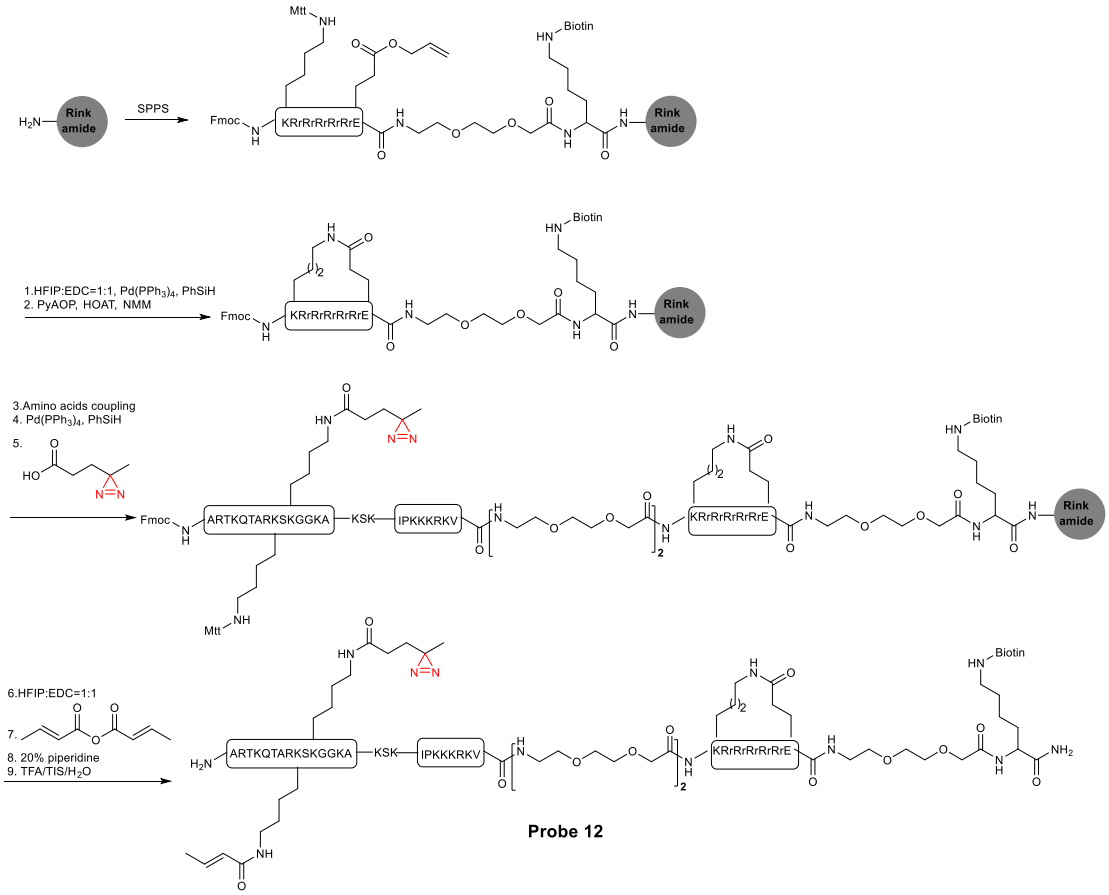
Supplementary Fig. 23. GO analysis (biological process) for enriched proteins from Hela cells using probe **9**. The number of proteins in each GO term is shown.



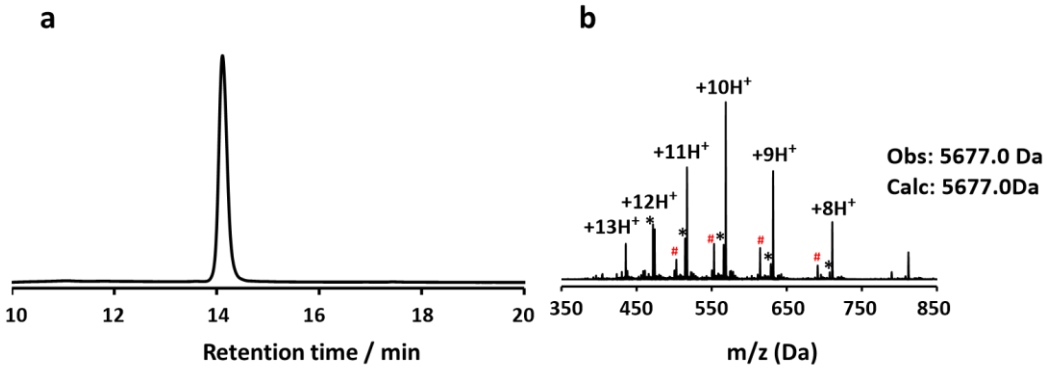
Supplementary Fig. 24. The subcellular location of proteins enriched by our H3K4me3 probe in RAW264.7 cells (from UniprotKB keywords).



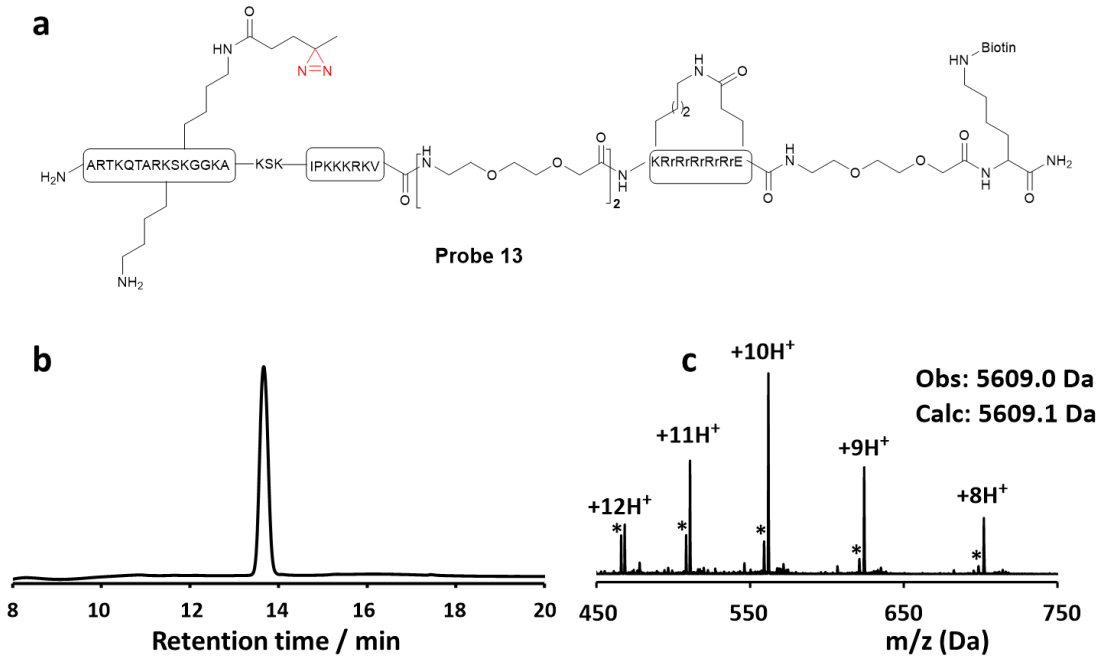
Supplementary Fig. 25. GO analysis (cellular components) of enriched proteins from RAW264.7 cells using probe **9**. The number of proteins in each GO term is shown.



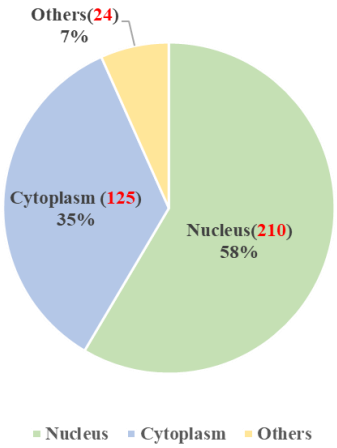
Supplementary Fig. 26. Synthetic route for probe **12**.



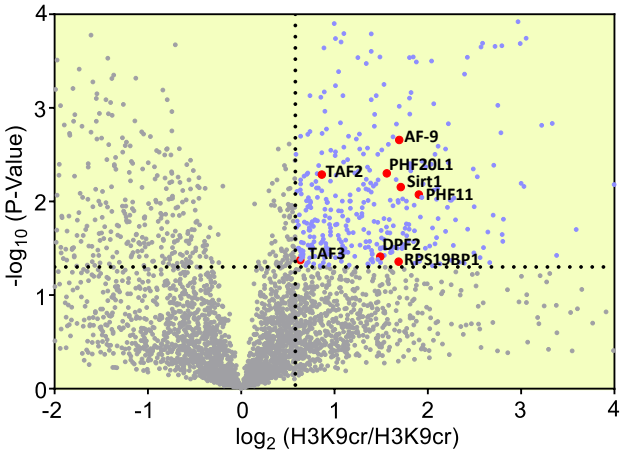
Supplementary Fig. 27. a) Analytical HPLC of probe **12**. b) ESI-MS analysis of probe **12**. "*" indicates the molecular weight of diazirine after loss of N_2. # indicates the molecular weight of an arginine lost.



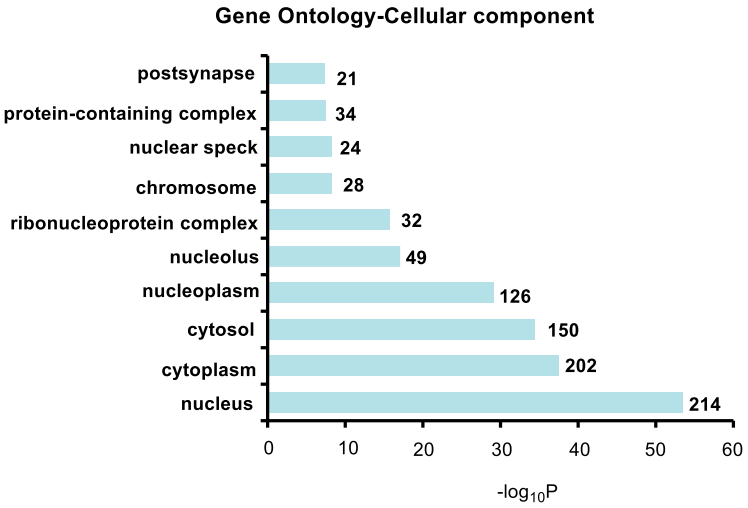
Supplementary Fig. 28. a) Chemical structure of probe **13**. b) Analytical HPLC of probe **13**. c) ESI-MS analysis of probe **13**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



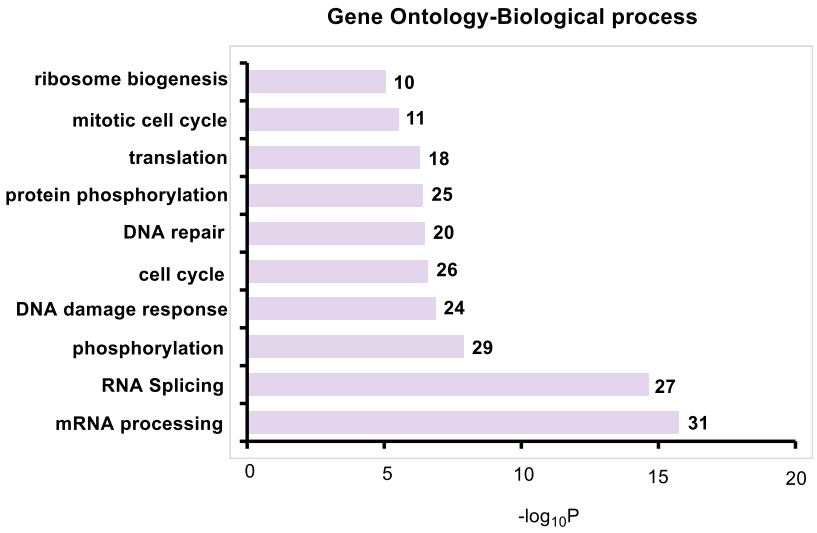
Supplementary Fig. 29. The subcellular location of proteins enriched by our H3K9cr probe in RAW264.7 cells (from UniprotKB keywords).



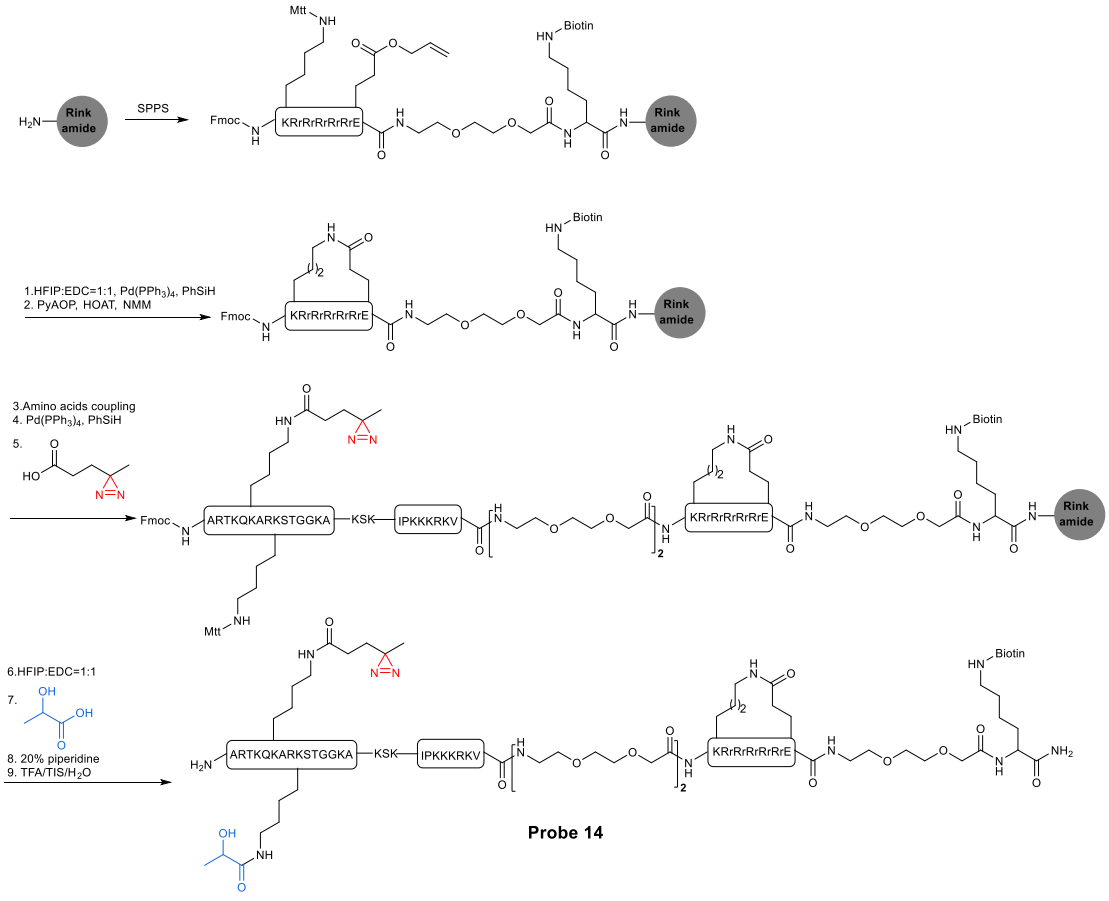
Supplementary Fig. 30. Volcano plots of mass spectrometry results. Hits significantly enriched by probe **12** ($p < 0.05$, >1.5-fold change) are colored blue. Some established H3K9cr reader proteins are highlighted and labeled in red. A two-sided test was used.



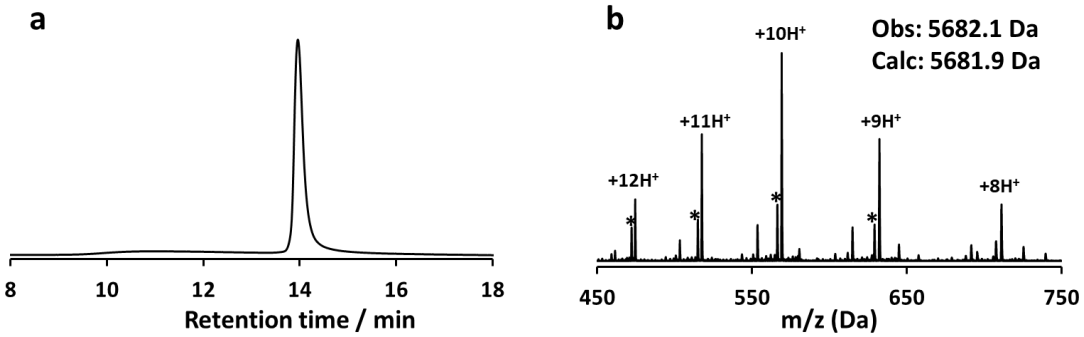
Supplementary Fig. 31. GO analysis (cellular components) of enriched proteins from RAW264.7 cells using probe **12**. The number of proteins in each GO term is shown.



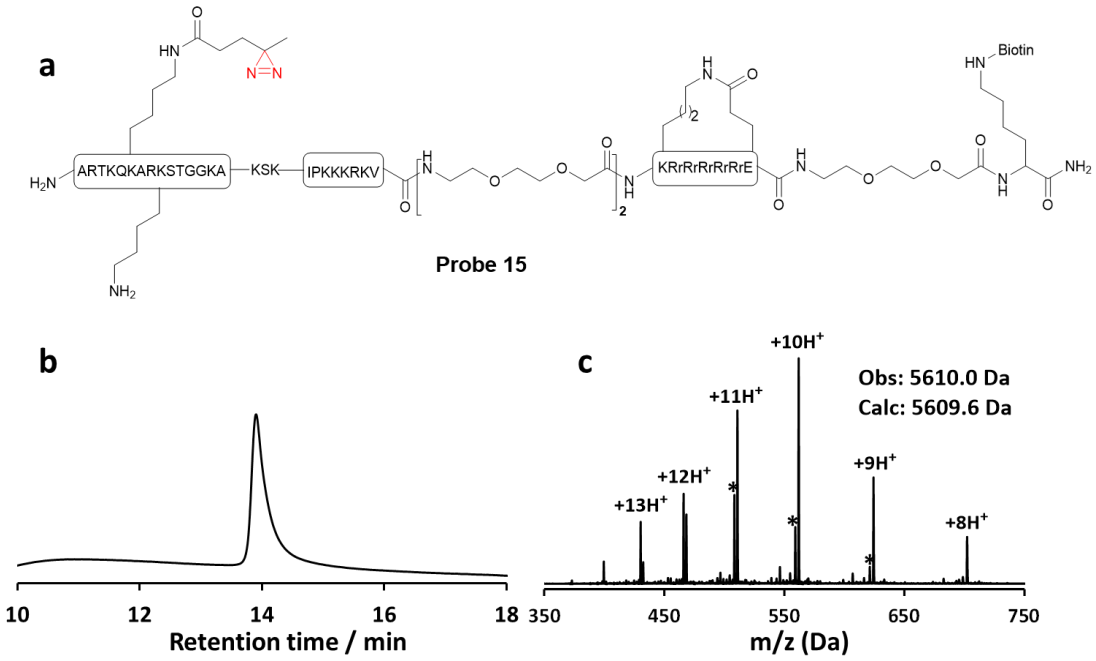
Supplementary Fig. 32. GO analysis (biological process) for enriched proteins from RAW264.7 cells using probe **12**. The number of proteins in each GO term is shown.



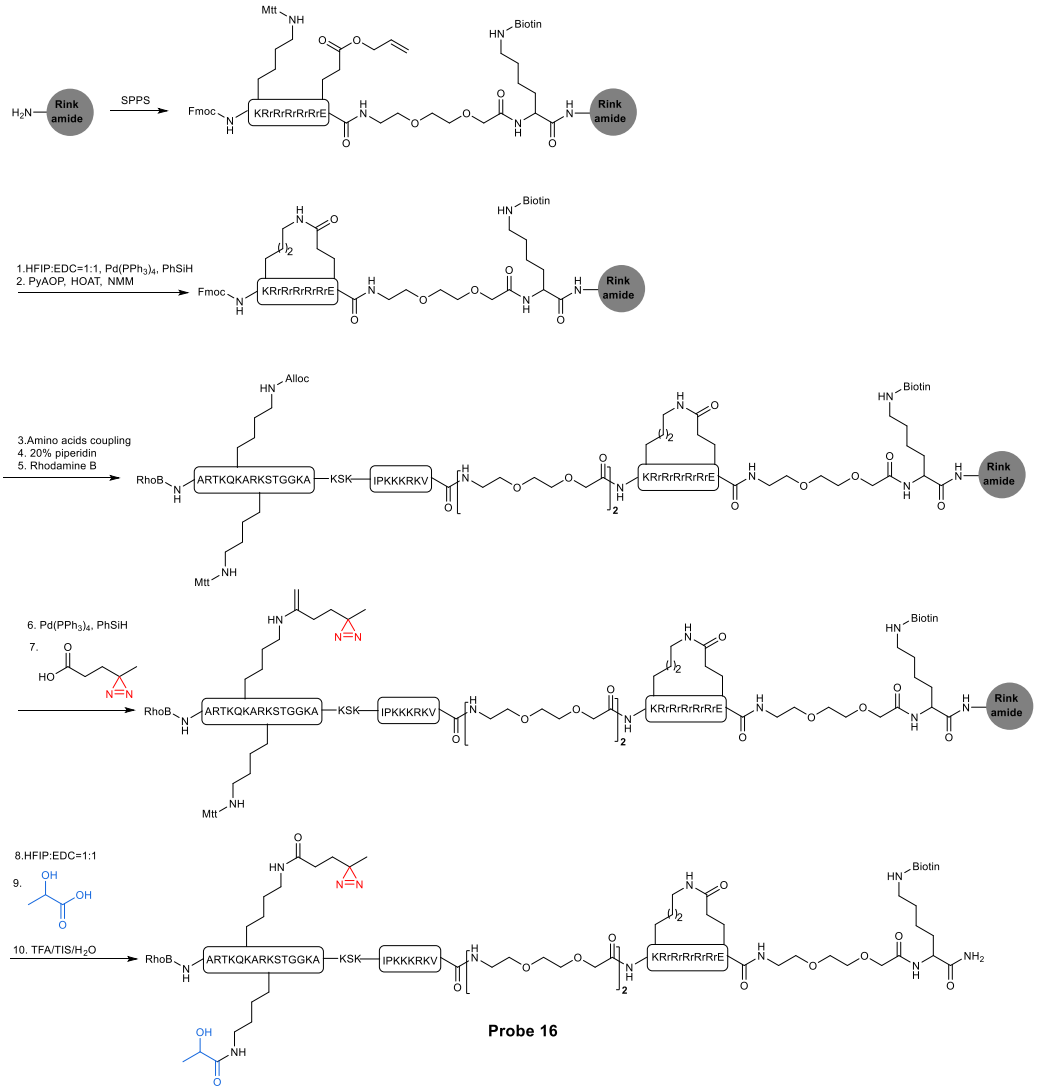
Supplementary Fig. 33. Synthetic route for probe **14**.



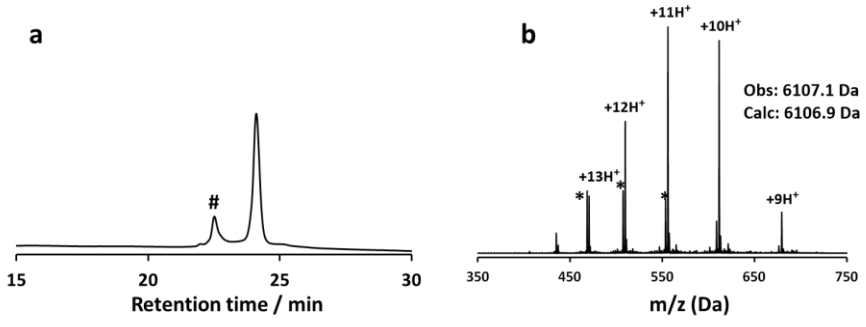
Supplementary Fig. 34. a) Analytical HPLC of probe **14**. b) ESI-MS analysis of probe **14**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



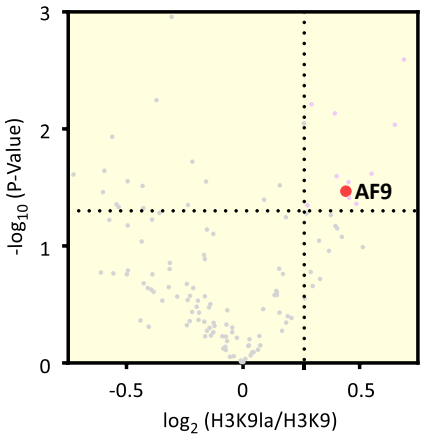
Supplementary Fig. 35. a) Chemical structure of probe **15.** b) Analytical HPLC of probe **15**. c) ESI-MS analysis of probe **15**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



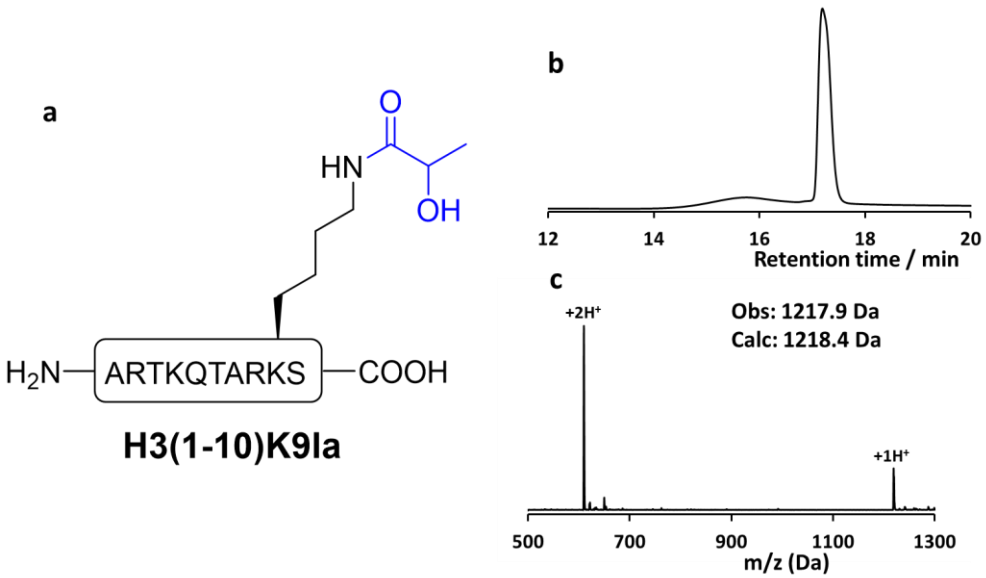
Supplementary Fig. 36. Synthetic route for probe **16**.



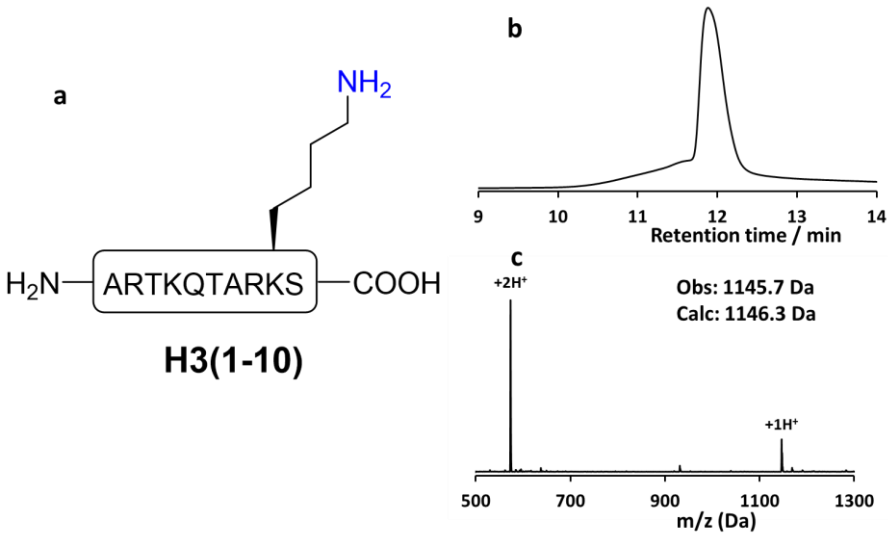
Supplementary Fig. 37. a) Analytical HPLC of probe **16**. "#" indicates that RhoB has a reversible equilibrium between non-fluorescent spironolactone and fluorescent zwitterions, thus producing two peaks in the chromatogram[1]. b) ESI-MS analysis of probe **16**. "*" indicates the molecular weight of diazirine after loss of N_2[2].



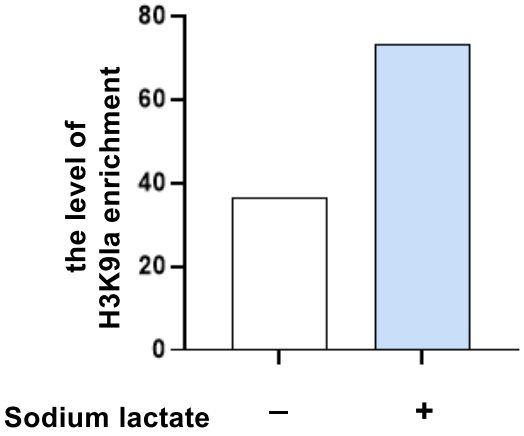
Supplementary Fig. 38. Volcano plots of the mass spectrometry results. Hits significantly enriched by probe **12** ($p < 0.05$, >1.2-fold change) are colored red. A two-sided test was used.



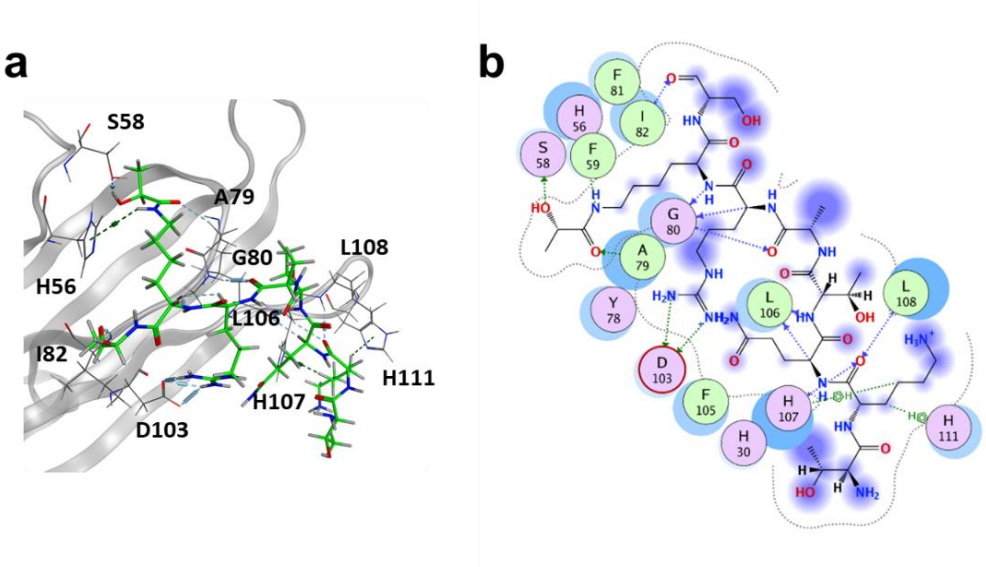
Supplementary Fig. 39. a) Chemical structure of peptide H3(1-10)K9la. b) Analytical HPLC of peptide H3(1-10)K9la. c) ESI-MS analysis of peptide H3(1-10)K9la.



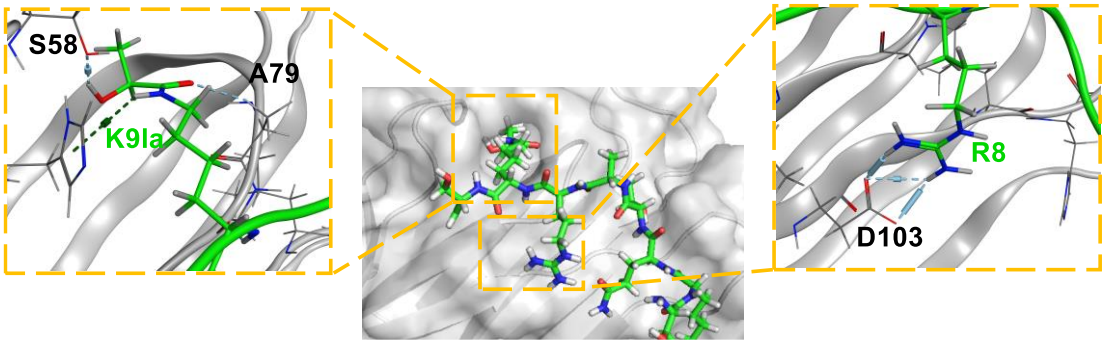
Supplementary Fig. 40. a) Chemical structure of peptide H3(1-10). b) Analytical HPLC of peptide H3(1-10). c) ESI-MS analysis of peptide H3(1-10).



Supplementary Fig. 41. Statistical quantification map of band grey values in result IB AF9 (n=1).



Supplementary Fig. 42. a) Stereo view of hydrogen bonding network involving H3 side chains and residues in $AF9_{YEATS}$. The secondary structure of the $AF9_{YEATS}$ was displayed in gray cartoon, the carbon atoms of the amino acid structure are displayed in gray line, the carbon atoms of the ligand are displayed in green sticks, the oxygen atoms are red, and the nitrogen atoms are blue. The light blue dashed line represents hydrogen bonds, while the light green dashed line represents arene-cation interactions. b) LIGPLOT diagram listing critical contacts between the H3(1-10)K9la peptide and the AF9 YEATS domain. Arrows represent hydrogen bonds, and dashed lines represent arene-cation interaction.



Supplementary Fig. 43. Detailed interaction map of the K9la readout by $AF9_{YEATS}$.

Supplementary Table 1. Protein hits observed in the H3K4me3 sample.

	Protein	Known?	Reference
H3K4me3	SPIN1	Yes	3
	CHD1	Yes	4
	CHD3	Yes	5
	TAF3	Yes	6
	TAF2	By similarity to TAF3	
	CXXC1	Yes	7
	PHF20L1	By similarity to PHF8	
	PHF6	By similarity to PHF8	
	WDR3	By similarity to WDR5	
	WDR70	By similarity to WDR5	
	WDR75	By similarity to WDR5	
	INO80	Yes	8
	INO80B	Yes	8
	TOP1	Yes	9
	NPM1	Yes	10
	RBM39	Yes	11
	CTCF	Yes	12
	SSRP1	Yes	9
	ZCCHC10	NO	
	ZCCHC17	NO	
	ZNF136	NO	

Supplementary Table 2. Well-characterized reader proteins of H3K4me3 enriched by our probe and two recent studies.

domain	This study	Burton, A. J et al.[13]	Li, W. et al[14]
PHD	TAF3、TAF2、PHF20L1、PHF6、CXXC1	ING5、CXXC1、PHF2、KDM5A、KDM5B、DIDO1、KDM7A、PHF8、BPTF、PHF23、ING1、ING2、ING3、ING4	ING2、ING1、PHF8、PYGO2、CXXC1、DIDO1、PHF23、KDM5C、BPTF
Chromeo	CHD1、CHD3	CHD1	
WD40	WDR3、WDR70、WDR75		WDR5
Tudor	SPIN1	SPIN1	CCDC101、SPIN3、SPIN1、KDM4A、KDM4C
Others	INO80、TOP1、NPM1、RBM39、CTCF、SSRP1、INO80B	MORC3、ORC3、BRWD2	KCMF1

Supplementary Table 3. Well-characterized reader proteins of H3K4me3 enriched by our probe in Hela cells and RAW264.7 cells.

domain	Hela	RAW 264.7
PHD	TAF3、TAF2、PHF20L1、PHF6、CXXC1	TAF3、BRPF1
Chromeo	CHD1、CHD3	CHD1、CHD9、CHD6
WD40	WDR3、WDR70、WDR75	
Tudor	SPIN1	SPIN1
Others	INO80、TOP1、NPM1、RBM39、CTCF、SSRP1、INO80B	SSRP1、SUPT16H、RBM39、NPM1、INO80B

Supplementary Table 4. Isolated mass and percent yields of probe 1-16, H3(1-10)K9la and H3(1-10)

Probe	Isolated mass (mg)	Percent yields (%)
1	**12.1**	**37.2**
2	**16.0**	**38.0**
3	**21.3**	**40.3**
4	**15.2**	**32.2**
5	**16.6**	**35.3**
6	**15.4**	**30.0**
7	**18.2**	**37.2**
8	**18.8**	**39.1**
9	**40.3**	**38.3**
10	**45.1**	**35.2**
11	**20.3**	**40.1**
12	**35.8**	**21.4**
13	**39.8**	**33.2**
14	**44.1**	**35.6**
15	**37.8**	**36.4**
16	**40.1**	**36.6**
H3(1-10)K9la	**5.0**	**60.8**
H3(1-10)	**5.2**	**70.5**

Supplementary Reference

1. D. Si, Q. Li, Y. Bao, J. Zhang, L. Wang, Fluorogenic and Cell-Permeable Rhodamine Dyes for High-Contrast Live-Cell Protein Labeling in Bioimaging and Biosensing. *Angew. Chem. Int. Ed.* **62**, e202307641 (2023).

2. T. Yang, Z. Liu, X. D. Li, Developing diazirine-based chemical probes to identify histone modification 'readers' and 'erasers'. *Chem. Sci.* **6**, 1011-1017 (2015).

3. W. Wang, Z. Chen, Z. Mao, H. Zhang, X. Ding, S. Chen, X. Zhang, R. Xu, B. Zhu, Nucleolar protein Spindlin1 recognizes H3K4 methylation and stimulates the expression of rRNA genes. *EMBO Rep.* **12**, 1160-1166 (2011).

4. J. F. Flanagan, L. Z. Mi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad, S. Khorasanizadeh, Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**, 1181-1185 (2005).

5. Y. Hu, D. Liu, X. Zhong, C. Zhang, Q. Zhang, D. X. Zhou, CHD3 protein recognizes and regulates methylated histone H3 lysines 4 and 27 over a subset of targets in the rice genome. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5773-5778 (2012).

6. S. M. Lauberth, T. Nakayama, X. Wu, A. L. Ferris, Z. Tang, S. H. Hughes, R. G. Roeder, H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021-36 (2013).

7. J. H. Lee, C. M. Tate, J. S. You, D. G. Skalnik, Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J. Biol. Chem.* **282**, 13419-13428 (2007).

8. Z. Shao, Y. Bai, E. Huq, H. Qiao, LHP1 and INO80 cooperate with ethylene signaling for warm ambient temperature response by activating specific bivalent genes. *bioRxiv.* doi: 10.1101/2024.03.01.583049 (2024).

9. A. Husain, N. A. Begum, T. Taniguchi, H. Taniguchi, M. Kobayashi, T. Honjo, Chromatin remodeller SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic instability. *Nat. Commun.* **7**, 10549 (2016).

10. X. Cui, D. Huo, Q. Wang, Y. Wang, X. Liu, K. Zhao, Y. You, J. Zhang, C. Kang, RUNX1/NPM1/H3K4me3 complex contributes to extracellular matrix remodeling via enhancing FOSL2 transcriptional activation in glioblastoma. *Cell Death Dis.* **15**, 98 (2024).

11. P. K. Puvvula, Y. Yu, K. R. Sullivan, H. Eyob, J. Rosenberg, A. Welm, C. Huff, A. M. Moon, Inhibiting an RBM39/MLL1 epigenomic regulatory complex with dominant-negative peptides disrupts cancer cell transcription and proliferation. *Cell Rep.* **35**, 109156 (2021).

12. Y. S. Bedi, H. Wang, K.N. Thomas, A. Basel, J. Prunier, C. Robert, M. C. Golding, Alcohol induced increases in sperm Histone H3 lysine 4 trimethylation correlate with increased placental CTCF occupancy and altered developmental programming. *Sci. Rep.* **12**, 8839 (2022).

13. A. J. Burton, M. Haugbro, L. A. Gates, J. D. Bagert, C. D. Allis, T. W. Muir, In situ chromatin interactomics using a chemical bait and trap approach. *Nat. Chem.* **12**, 520-527 (2020).

14. J. Lin, X. Bao, X. D. Li, A tri-functional amino acid enables mapping of binding sites for posttranslational-modification-mediated protein-protein interactions. *Mol. Cell* **81**, 2669-2681.e9 (2021).

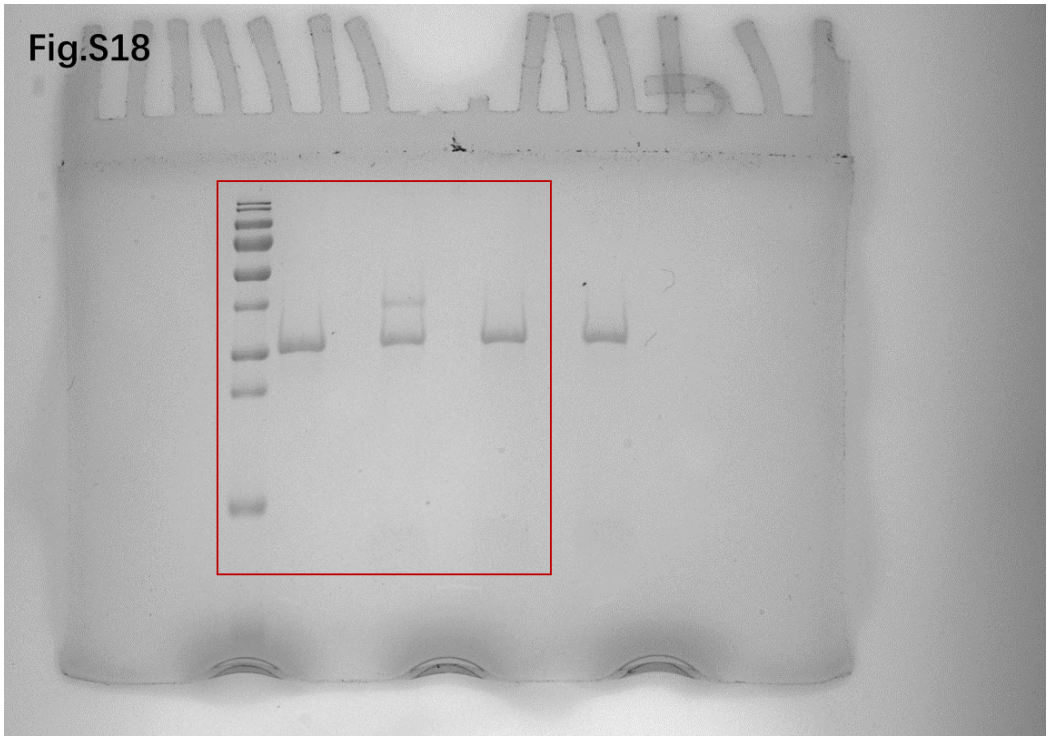


Fig.S18