



Contents lists available at ScienceDirect

Chinese Chemical Letters

journal homepage: [www.elsevier.com/locate/ccllet](http://www.elsevier.com/locate/ccllet)

## A new photolabeling probe for efficient enrichment and deep profiling of cell surface membrane proteome by mass spectrometry



Yuanyuan Li<sup>a,b</sup>, Chaoshuang Xia<sup>b</sup>, Hongxian Zhao<sup>b</sup>, Yuping Xie<sup>b</sup>, Yangjun Zhang<sup>b</sup>,  
Wanjun Zhang<sup>b,\*</sup>, Yongliang Yu<sup>a,\*</sup>, Jianhua Wang<sup>a</sup>, Weijie Qin<sup>b,c,\*</sup>

<sup>a</sup> Research Center for Analytical Sciences, Department of Chemistry, College of Sciences, Northeastern University, Shenyang 110819, China

<sup>b</sup> National Center for Protein Sciences Beijing, State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Lifeomics, Beijing 102206, China

<sup>c</sup> College of Basic Medicine, Anhui Medical University, Hefei 230032, China

### ARTICLE INFO

#### Article history:

Received 20 December 2021

Revised 13 February 2022

Accepted 25 March 2022

Available online 28 March 2022

#### Keywords:

Cell surface membrane proteins

Photolabeling probe

Enrichment

Mass-spectrometry

Proteomics

Derivatization

### ABSTRACT

The cell surface membrane proteome is a class of proteins encoded by ~25% of all protein-coding genes in living organisms and plays a key role in mediating communication between the cells and their surrounding environment. However, most cell surface membrane proteins (CSMPs) are naturally expressed at very low levels compared with intracellular proteins. The difficulties in their purification with high specificity further hinder the understanding of their structure and function. In this study, we developed a new photolabeling probe to achieve efficient tagging and facile enrichment of the CSMPs. The probe is composed of a lipid tail for cell surface localization, a polyethylene glycol (PEG) spacer for increased water solubility, two 4-(*N*-maleimido)benzophenone (MBP) groups for UV-active tagging of the CSMPs, and a biotin tag for subsequent isolation. Application of this photolabeling probe resulted in the successful enrichment and identification of 3098 annotated CSMPs in HT22 cells with close to 70% selectivity. The proposed photolabeling probe and enrichment strategy were demonstrated to be a powerful method for deep cell surface proteome profiling, representing one of the largest groups of current drug targets.

© 2022 Published by Elsevier B.V. on behalf of Chinese Chemical Society and Institute of Materia Medica, Chinese Academy of Medical Sciences.

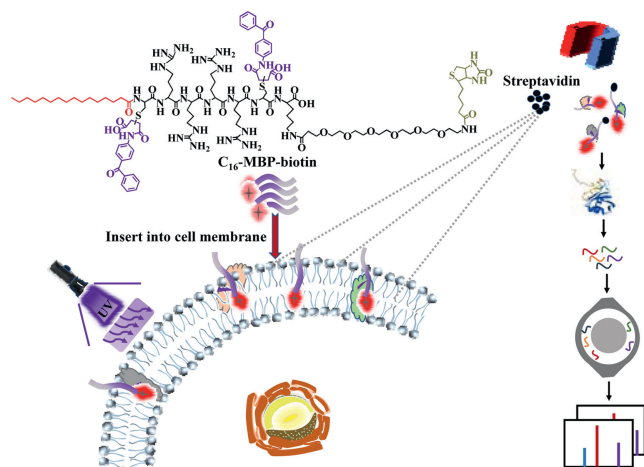
The plasma membrane at the cell surface is covered with proteins that play key roles in many cellular functions and activities, such as cell–cell interactions, pathogen recognition, ion transport and signal transduction [1–5]. It is estimated that approximately 25% of protein-coding genes in the human genome are integral membrane proteins [6]. These proteins include many cell surface receptors, ion channels, and transporter proteins, constituting approximately 70% of all Food and Drug Administration (FDA)-approved drug targets, and reflecting the pharmacological relevance of cell surface membrane proteins (CSMPs) [7,8]. Therefore, comprehensive analysis of CSMPs may facilitate a better understanding of their functions in various cellular activities and disease development to aid in the discovery of new biomarkers and drug targets [9].

Great efforts have been made in the fields of localization, identification, and quantification of CSMPs in recent years. For example, aptamers have been employed to obtain novel tumor markers/drug

targets from CSMPs for therapeutic purposes [10–12]. Real-time imaging methods using nanotechnology to investigate CSMPs have also been reported for disease diagnosis and therapeutic discovery [13–16]. Compared with the above targeted strategies that study one protein at a time, the throughput of proteomic analysis conducted by liquid chromatography–mass spectrometry (LC–MS) is obviously advantageous. Thousands of proteins can now be identified in a single LC–MS run on a routine-basis. However, CSMPs are commonly underrepresented in most large-scale proteomic studies, due to their relatively low abundance [17]. An essential step in cell surface proteome identification is the isolation and purification of CSMPs, which is complicated by their inherent hydrophobicity and heterogeneity. To achieve the efficient isolation of CSMPs, various approaches have been reported, which can be subcategorized into two major types. One type relies on the physicochemical characteristics of CSMPs, and the other depends on the chemical labeling of CSMPs. Compared with physicochemical isolation using high-speed centrifugation [18] and assisted dissolution by detergents [19–22], which are more suitable for the extraction of all membrane proteins including proteins from the plasma, nucleus and mitochondrial membrane, labeling methods using hydrazide

\* Corresponding authors.

E-mail addresses: [zwj2004zwj@126.com](mailto:zwj2004zwj@126.com) (W. Zhang), [yuy@mail.neu.edu.cn](mailto:yuy@mail.neu.edu.cn) (Y. Yu), [aunp\\_dna@126.com](mailto:aunp_dna@126.com) (W. Qin).



**Scheme 1.** C<sub>16</sub>-MBP-biotin probe-based photoactive labeling strategy for CSMP enrichment and large-scale identification by MS.

chemistry, “click chemistry” and enzymatic tagging on cell surface glycans/proteins are more specifically targeted at plasma membrane/cell surface proteins [23–26]. Although hundreds of CSMPs have been successfully identified, the proteome coverage and selectivity are still unsatisfactory, presumably due to the limited cell surface targeting ability of the reported methods.

Inspired by the lipophilic and hydrophobic properties of palmitic acid and its capability for membrane insertion [27,28], we developed a novel palmitic acid-based photoactive probe, C<sub>16</sub>-C(MBP)RRRRC(MBP)K-PEG<sub>6</sub>-biotin (C<sub>16</sub>-MBP-biotin, where C is a cysteine and R is an arginine) for CSMP labeling in this work. This probe is composed of C<sub>16</sub> for cell surface membrane insertion, a short peptide for improving membrane targeting, two 4-(*N*-maleimido)benzophenone (MBP) groups for photoactive-based protein labeling, PEG<sub>6</sub> to enhance the water solubility of the probe and biotin for specific enrichment.

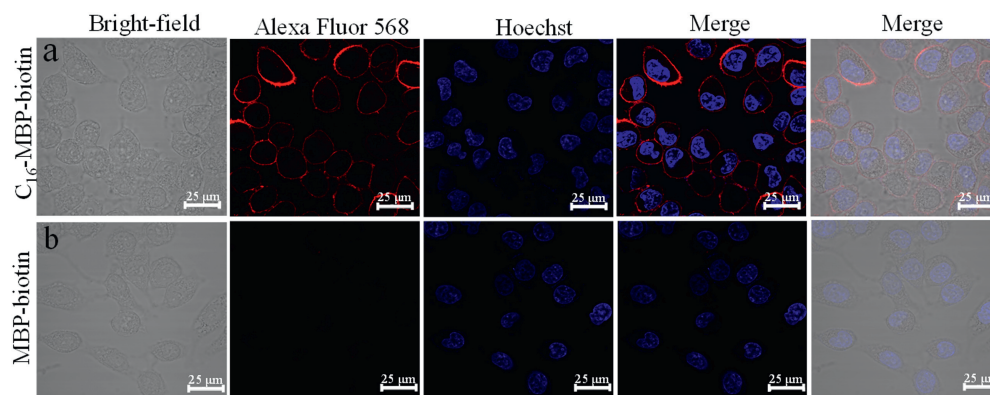
As shown in Scheme 1, the C<sub>16</sub>-MBP-biotin probe inserts into the negatively charged lipid bilayer of the cell membrane due to the lipid affinity of C<sub>16</sub> and the electrostatic interaction of the positively charged peptide sequence (CRRRRCK). After irradiation with 365 nm UV, benzophenone, which is part of the lipid bilayer-inserted C<sub>16</sub>-MBP-biotin probe, is activated to form an extremely reactive triplet biradical, and then abstracts the surrounding molecular hydrogen atom to transform it into a free radical [29]. Next, the radicals on the carbon atoms form covalent linkages with the adjacent amino acid residues and specifically label cell surface membrane proteins (CSMPs) via a C, H-insertion cross-linking [30–34]. Finally, the probe-tagged CSMPs can be specifically enriched via the strong binding between biotin and streptavidin beads after lysing the cells. After enriching the tagged proteins via the biotin handle, LC-MS analysis is carried out for CSMP identification. It is worth noting that the photoactive labeling step can be well controlled by turning on or off the UV light source to avoid false-positive tagging of the intracellular proteins during sample processing. Therefore, the proteins on the cell surface can be photochemically labeled and enriched in a controlled and highly specific manner. The C<sub>16</sub>-MBP-biotin probe was successfully applied in the large-scale enrichment of CSMPs from HT22 cells. A total of 3486 ± 120 proteins were enriched and identified from four replicate tests with good quantification reproducibility (Pearson coefficients all above 0.9), indicating the high enrichment efficiency and reproducibility of this method. In total, 3098 identified proteins were annotated as CSMPs by comparison with the PANTHER classification system (<http://pantherdb.org/>) [35] and the UniProt database (<http://www.uniprot.org>). After evaluation of the identifi-

cation results of C<sub>16</sub>-MBP-biotin by quantitative proteomic comparison with the nonmembrane targeting control probe without C<sub>16</sub> (C(MBP)RRRRC(MBP)-PEG<sub>6</sub>-biotin, referred to as MBP-biotin), we obtained 3883 identified proteins that passed stringent filtering (fold change ≥ 4, *P* ≤ 0.01), indicating the high enrichment selectivity of this method. Furthermore, 1275 proteins were predicted to have transmembrane domains according to TMHMM-2.0 [36] and Phobius [37].

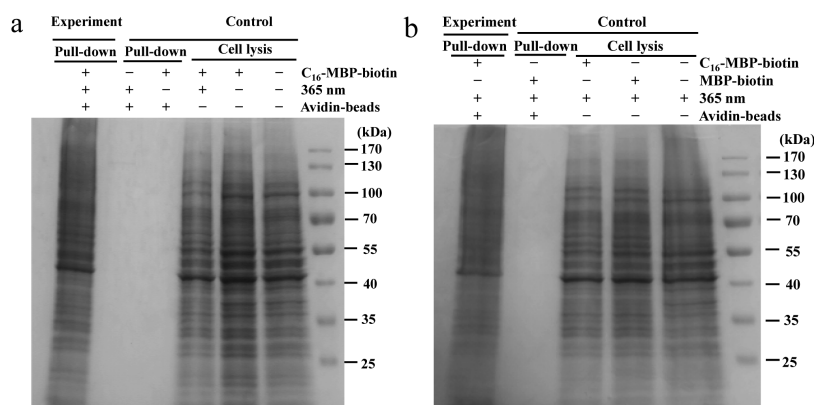
The key synthesis process of the C<sub>16</sub>-MBP-biotin probe is summarized in Scheme S1 in Supporting information and is described in detail in the Methods section. Briefly, C<sub>16</sub>-PEP-biotin (C<sub>16</sub>-CRRRRC-PEG<sub>6</sub>-biotin) was provided by Synpeptide Co., Ltd. (Shanghai, China) via the standard solid-phase peptide synthesis method. Subsequently, the sulfhydryl group of cysteine in the CRRRRC sequence was coupled with maleimide to enable easy introduction of the MBP group to the C<sub>16</sub>-PEP-biotin probe via a simple and efficient Michael reaction. The masses of the C<sub>16</sub>-MBP-biotin probe (Fig. S1 in Supporting information) and the control probe (MBP-biotin) (Fig. S2 in Supporting information) were measured by matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS) or electrospray ionization (ESI) to demonstrate the successful synthesis of the two probes. Purities of over 90% were obtained for the two probes by liquid chromatography separation (Figs. S3 and S4 in Supporting information). The synthesis of C<sub>16</sub>-MBP-biotin and MBP-biotin was confirmed by <sup>1</sup>H NMR and <sup>13</sup>C NMR (Figs. S5–S8 in Supporting information).

The feasibility of the C<sub>16</sub>-MBP-biotin probe for cell surface targeting was demonstrated using fluorescence imaging. We used Hoechst for nuclear labeling (blue) and Alexa Fluor 568-labeled streptavidin (red) to bind with the biotin of the C<sub>16</sub>-MBP-biotin probe to obtain the subcellular localization of the probe. The probe treated cells were thoroughly washed to ensure removal of the excess probes that were not bound on the cell surface. As shown in Fig. 1a, a strong red fluorescent signal of Alexa Fluor 568-labeled streptavidin was discovered for the cells treated with the C<sub>16</sub>-MBP-biotin probe. Red fluorescence was only observed on the surface of the cell with almost no internalization into the cytoplasmic region after 20 min of incubation, indicating successful cell surface insertion of the probe. In contrast, the control probe (MBP-biotin) could not be located at the cell surface and could be easily removed by repeated washing to avoid false-positive labeling (Fig. 1b).

The selectivity of C<sub>16</sub>-MBP-biotin probe-based CSMP labeling and enrichment was evaluated using HT22 cells and SDS-PAGE. In the experimental group shown in Fig. 2a, the C<sub>16</sub>-MBP-biotin probe, UV irradiation and streptavidin-bead enrichment were used to locate the probe on the cell surface, tag it with proteins and extract the proteins, respectively. The enriched proteins were eluted for SDS-PAGE analysis, and clear gel bands were observed. However, no gel bands could be found in the control groups (pull down) without either the C<sub>16</sub>-MBP-biotin probe or UV irradiation, indicating that these are two key factors for achieving successful protein tagging and enrichment. The effects of the C<sub>16</sub>-MBP-biotin probe and UV irradiation on the stability of proteins were also investigated by applying the corresponding conditions to the cell lysates and no protein degradation was observed (control, cell lysis, 1% v/v loading). To further evaluate the key function of C<sub>16</sub> in protein enrichment, the MBP-biotin control probe without C<sub>16</sub> was used, as shown in Fig. 2b. Omitting C<sub>16</sub> leads to complete abolishment of the gel bands, demonstrating that without C<sub>16</sub> insertion into the lipid bilayer, the probes could be completely removed from the cell by repeated washing and would not interfere with the subsequent protein enrichment process. After demonstrating the feasibility of C<sub>16</sub>-MBP-biotin-based protein labeling and enrichment, the labeling conditions were optimized by varying the amount of the C<sub>16</sub>-MBP-biotin probe used and the UV irradiation time. As revealed in Fig S9a (Supporting information), HT22



**Fig. 1.** Fluorescent confocal microscopy images of the cells treated with  $C_{16}$ -MBP-biotin (a) and MBP-biotin control probes (b). The cells were co-stained with Hoechst 33342 (1  $\mu\text{g}/\text{mL}$ , blue) and Alexa Fluor 568-labeled streptavidin (5  $\mu\text{g}/\text{mL}$ , red) as nucleus and biotin markers, respectively.



**Fig. 2.** (a) SDS-PAGE characterization of  $C_{16}$ -MBP-biotin probe-treated HT22 cells under different conditions; (b) SDS-PAGE characterization of the key role of  $C_{16}$  in protein enrichment.

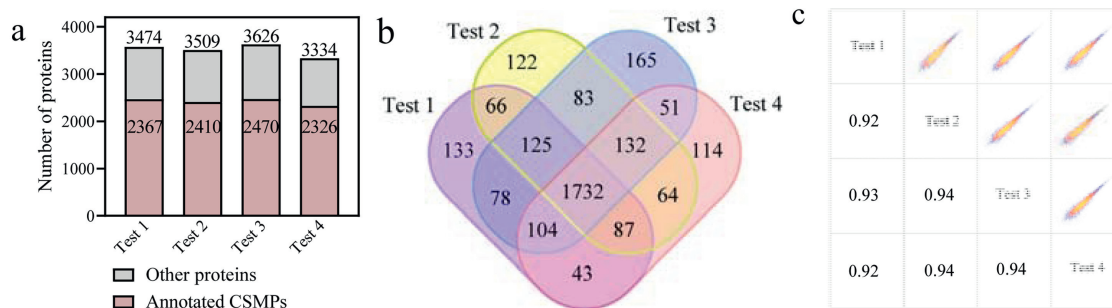
cells were tagged with the  $C_{16}$ -MBP-biotin probe by UV irradiation for different times, and cell lysates were collected for SDS-PAGE. No obvious protein degradation was observed for UV irradiation from 0.5 min to 10 min. Further characterization of the  $C_{16}$ -MBP-biotin probe-tagged proteins by western blotting using streptavidin staining showed that optimal tagging was achieved at a 2 min irradiation time (Fig. S9b in Supporting information). Almost no streptavidin staining was found in the control groups without the  $C_{16}$ -MBP-biotin probe and 365 nm UV, indicating the high labeling specificity of this method. Preincubation of the  $C_{16}$ -MBP-biotin probe with HT22 cells is critical for efficient protein labeling. Therefore, we evaluated the time and amount of  $C_{16}$ -MBP-biotin probe used for incubation with HT22 cells by characterizing the totally tagged protein quantity enriched by the streptavidin beads.  $C_{16}$ -MBP-biotin (15  $\mu\text{mol}/\text{L}$ ) and a 20 min incubation time were used in the following experiments (Figs. S10a and b in Supporting information).

After establishing the optimized workflow of the  $C_{16}$ -MBP-biotin probe for CSMP tagging and enrichment, we applied it in the large-scale cell surface proteome profiling of HT22 cells for further evaluation. We obtained  $3486 \pm 120$  proteins in four biological replicates, among which  $2393 \pm 62$  (3098 in total) were annotated as CSMPs (Fig. 3a and Table S1 in Supporting information). Close to 70% of the identified proteins by  $C_{16}$ -MBP-biotin probe enrichment were annotated as CSMPs. Furthermore, 75.4% of all the annotated CSMPs (2560) were identified in at least two tests, and the Pearson coefficient for each two tests was above 0.9, indicating good identification and quantification reproducibility (Figs. 3b and c). For the identified proteins, TMHMM-2.0 and Phobius were employed to predict whether they have transmembrane domains

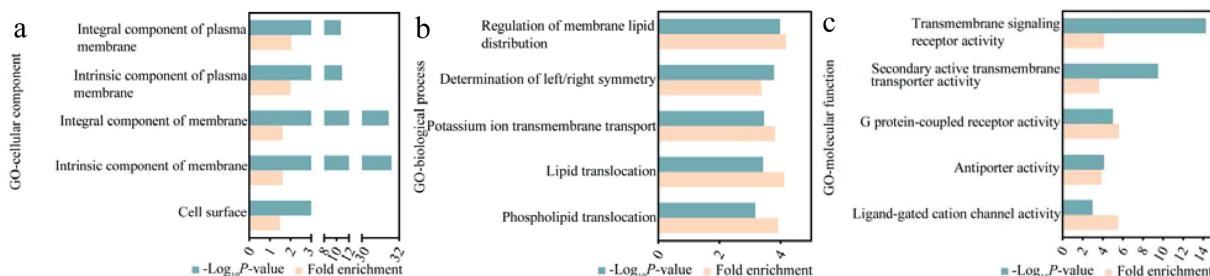
(TMDs). As a result, 1275 proteins in our results were predicted to have one to multiple TMDs. The distribution of the number of TMDs is displayed in Fig S11 (Supporting information).

Gene ontology (GO) annotation analysis of the  $C_{16}$ -MBP-biotin probe-enriched proteins was performed using the proteomic profiling data of the total cell lysates without enrichment as the reference. As shown in Fig. 4a, "Integral component of plasma membrane", "Intrinsic component of plasma membrane", "Integral component of membrane", "Intrinsic component of membrane" and "Cell surface" were significantly enriched in the GO terms of cellular component. For biological process, GO terms such as "Regulation of membrane lipid distribution", "Potassium ion transmembrane transport", "Lipid translocation", and "Phospholipid translocation" were significantly enriched (Fig. 4b). Similarly, cell surface-related GO terms such as "Transmembrane signaling receptor activity", "Secondary active transmembrane transporter activity", "G protein-coupled receptor activity", "Antiporter activity", and "Ligand-gated cation channel activity" were enriched in the molecular function category. The above results further demonstrated the high cell surface selectivity of our  $C_{16}$ -MBP-biotin probe (Fig. 4c).

Considering that nonspecific adsorption is inevitable in enrichment experiments, quantitative comparative proteomic analysis of the MBP-biotin control probe was conducted to further evaluate the enrichment efficiency and remove the nonspecifically adsorbed non-CSMPs. Without cell surface targeting by  $C_{16}$ , the MBP-biotin control probe could be easily removed by repeated washing, resulting in significantly reduced protein binding on the streptavidin beads and decreased protein identification. Only a very limited number of nonspecifically adsorbed proteins were found in the control samples. Volcano plots and differential protein screen-



**Fig. 3.** (a) Number of (cell surface membrane) proteins identified by  $C_{16}$ -MBP-biotin probe enrichment in the four biological replicates. (b, c) Identification and quantification reproducibility of the annotated CSMPs.



**Fig. 4.** GO enrichment analysis ((a) cellular component, (b) biological process and (c) molecular function) of the  $C_{16}$ -MBP-biotin probe-enriched proteins.

ing were conducted using Perseus [38]. By stringent filtering (fold change  $\geq 4$ ,  $P \leq 0.01$ ), 3883 differential proteins were obtained, as shown in the volcano plot in Fig S12a (Supporting information), resulting in an improved percentage of the annotated CSMPs after using the PANTHER tool and the UniPort database. Only 27% of the differential proteins were left without information to support their localization on the cell surface. The detailed information is provided in Table S2 (Supporting information). Consistently, GO analysis of the 3883 differential proteins revealed highly enriched GO terms that were almost all related to cell surface proteins (Fig. S12b in Supporting information).

We also compared our method with two previously used strategies for membrane protein extraction, the centrifugation method and the membrane protein extraction kit to further demonstrate the novelty and significance of our method. The centrifugation method for CSMPs isolation was performed according to the reported protocol [39]. The membrane protein extraction kit was used following the manufacturer's instructions. 3098, 2128 and 2331 CSMPs were enriched and identified by our method, the centrifugation method and the extraction kit. Our method exhibited 45.6% and 32.9% enhancement in the identification scale of CSMPs compared with the two existing methods. Further comparison revealed that the CSMPs enriched by our method efficiently covered those obtained by the two existing methods ( $\sim 80\%$ ) (Fig. S13 in Supporting information). We attributed the increased CSMPs enrichment by our method to the novel design of the  $C_{16}$ -MBP-biotin probes. (1) Due to its lipophilic and hydrophobic properties, palmitic acid can insert the  $C_{16}$ -MBP-biotin probe into the phospho-lipid bilayers of the cell membrane, leading to a close proximity between the probe and the target CSMPs to facilitate direct and specific interaction between them. (2) The two highly reactive benzophenone groups in the  $C_{16}$ -MBP-biotin probe provide facile and sensitive photoactive-based protein labeling, which is crucial for capturing the low abundant CSMPs. Furthermore, the enrichment reaction can be easily initiated and stopped by simply "on/off" the UV irradiation. (3) The covalent linkages formed between the  $C_{16}$ -MBP-biotin probes and the target CSMPs support a stringent cleaning process using harsh wash conditions to facil-

itate removal of non-CSMPs and preserve the CSMPs. As a result, improved CSMPs enrichment identification can be achieved.

In summary, we reported a photoactive labeling probe,  $C_{16}$ -MBP-biotin, that covalently tags CSMPs for large-scale enrichment and identification of the cell surface proteome by LC-MS. The palmitic acid  $C_{16}$  and MBP groups of the probe lead to successful insertion of the probe on the cell surface and efficient UV light-induced protein tagging. More than 3000 known CSMPs with approximately 70% enrichment selectivity were achieved. We expect that further application of the  $C_{16}$ -MBP-biotin probe may lead to deeper cell surface proteome coverage and facilitate basic mechanistic studies of cell membrane proteins and drug target screening.

#### Declaration of competing interest

The authors report no declarations of interest.

#### Acknowledgments

This study was supported by National Natural Science Foundation of China (Nos. 32088101, 22074158, 21904008), National Key R&D Program of China (Nos. 2021YFA1302604, 2021YFA1301601, 2017YFA0505002), and National Key Laboratory of Proteomics (Nos. SKLP-K201706, 2021-NCPSB-003).

#### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ccllet.2022.03.100.

#### References

- [1] R. Phillips, T. Ursell, P. Wiggins, P. Sens, Nature 459 (2009) 379–385.
- [2] L. Kuhlmann, E. Cummins, I. Samudio, T. Kislinger, Expert Rev. Proteom. 15 (2018) 259–275.
- [3] J. Li, S. Han, H. Li, et al., Cell 180 (2020) 373–386.
- [4] X. Zhen, P.H. Cheng, K. Pu, Small 15 (2019) 1804105.
- [5] J.W. DePierre, M.L. Karnovsky, J. Cell Biol. 56 (1973) 275–303.
- [6] UniProt consortium, Nucleic Acids Res. 46 (2018) 2699.
- [7] J.J. Liu, K. Sharma, L. Zangrandi, et al., Science 360 (2018) 1314.

- [8] B.T. Lobingier, R. Hüttenhain, K. Eichel, et al., *Cell* 169 (2017) 350–360.
- [9] Y. Zhang, X. Wu, W.A. Tao, *TrAC-Trends Anal. Chem.* 107 (2018) 21–30.
- [10] Z. Guo, Y. Liu, N. He, Y. Deng, L. Jin, *Chin. Chem. Lett.* 32 (2021) 40–47.
- [11] P. Kalra, A. Dhiman, W.C. Cho, J.G. Bruno, T.K. Sharma, *Front. Mol. Biosci.* 5 (2018) 41.
- [12] H.E. Zumrut, M.N. Ara, M. Fraile, G. Maio, P. Mallikaratchy, *Nucleic Acid Ther.* 26 (2016) 190–198.
- [13] Y. Chen, C. Xue, J. Wang, et al., *Chin. Chem. Lett.* 33 (2022) 1637–1642.
- [14] R.L. Grime, J. Goulding, R. Uddin, et al., *Nanoscale* 12 (2020) 11518–11525.
- [15] H. Liu, Z. Miao, Z. Zha, *Chin. Chem. Lett.* 33 (2022) 1673–1680.
- [16] C. Sevencan, R.S.A. McCoy, P. Ravisankar, et al., *Adv. Ther.* 3 (2020) 1900201.
- [17] S. Li, H. Luo, R. Lou, et al., *Sci. Adv.* 7 (2021) eabf0634.
- [18] X. Lai, *Electrophoresis* 34 (2013) 809–817.
- [19] Q. Zhao, F. Fang, Y. Shan, et al., *Anal. Chem.* 89 (2017) 5179–5185.
- [20] L.H. Urner, I. Liko, H.Y. Yen, et al., *Nat. Commun.* 11 (2020) 1–10.
- [21] Y. Liu, G. Yan, M. Gao, C. Deng, X. Zhang, *Anal. Bioanal. Chem.* 408 (2016) 3495–3502.
- [22] Y. Liu, G. Yan, M. Gao, C. Deng, X. Zhang, *Proteomics* 15 (2015) 3892–3900.
- [23] M.V. Oostrum, B. Campbell, C. Seng, et al., *Nat. Commun.* 11 (2020) 1–16.
- [24] F. Sun, S. Suttapitugsakul, R. Wu, *Anal. Chem.* 91 (2019) 4195–4203.
- [25] Y. Xie, S. Chen, Q. Li, et al., *Chem. Sci.* 12 (2021) 8767–8777.
- [26] Y. Li, Y. Wang, Y. Yao, et al., *Anal. Chem.* 93 (2021) 4542–4551.
- [27] L.H. Liu, W.X. Qiu, Y.H. Zhang, et al., *Adv. Funct. Mater.* 27 (2017) 1700220.
- [28] W. Ma, S.N. Sha, P.L. Chen, et al., *Adv. Healthc. Mater.* 9 (2020) 1901100.
- [29] S.A. Fleming, *Tetrahedron* 51 (1995) 12479–12520.
- [30] G. Dorman, G.D. Prestwich, *Biochemistry* 33 (1994) 5661–5673.
- [31] Y. Hatanaka, Y. Sadakane, *Curr. Top. Med. Chem.* 2 (2002) 271–288.
- [32] G.R. Mettam, L.B. Adams, How to prepare an electronic version of your article, in: B.S. Jones, R.Z. Smith (Eds.), *Introduction to the Electronic Age, E-Publishing Inc.*, New York, 2009, pp. 281–304.
- [33] P.J.A. Weber, A.G. Beck-Sickinger, *J. Pept. Res.* 49 (1997) 375–383.
- [34] R.E. Galaray, L.C. Craig, J.D. Jamieson, et al., *J. Biol. Chem.* 249 (1974) 3510–3518.
- [35] G.D. Prestwich, G. Dormán, J.T. Elliott, D.M. Marecak, A. Chaudhary, *Photochem. Photobiol.* 65 (1997) 222–234.
- [36] H. Mi, A. Muruganujan, J.T. Casagrande, P.D. Thomas, *Nat. Protoc.* 8 (2013) 1551–1566.
- [37] A. Krogh, B. Larsson, G. Von Heijne, E.L. Sonnhammer, *J. Mol. Biol.* 305 (2001) 567–580.
- [38] L. Käll, A. Krogh, E.L. Sonnhammer, *J. Mol. Biol.* 338 (2004) 1027–1036.
- [39] S. Tyanova, T. Temu, P. Sinitcyn, et al., *Nat. Methods* 13 (2016) 731–740.