A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition GDCh Chemie

www.angewandte.org

Accepted Article

Title: Exploration of the reactivity of multivalent electrophiles for affinity labeling: sulfonyl fluoride as a highly efficient and selective label

Authors: Kaori Sakurai, Nanako Suto, Shoichi Hosoya, and Shione Kamoshita

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202104347

Link to VoR: https://doi.org/10.1002/anie.202104347

WILEY-VCH

10.1002/anie.202104347

WILEY-VCH

RESEARCH ARTICLE

Exploration of the reactivity of multivalent electrophiles for affinity labeling: sulfonyl fluoride as a highly efficient and selective label

Nanako Suto,[a] Shione Kamoshita,[a] Shoichi Hosoya[b] and Kaori Sakurai*[a]

[a] N. Suto, S. Kamoshita, Prof. K. Sakurai
Department of Bioengineering and Life Science
Tokyo University of Agriculture and Technology
4-24-16,Naka-cho, Koganei-shi, Tokyo, 184-8588, Japan
E-mail: sakuraik@cc.tuat.ac.jp
[b] S. Hosoya
Institute of Research
Tokyo Medical and Dental University,
1-5-45, Yushima, Bunkyo-ku, Tokyo, 113-8510, Japan

Supporting information for this article is given via a link at the end of the document.

Abstract: Designing suitable affinity labels for elucidating small molecule target proteins is an important problem in chemical probe-based approaches to drug target discovery and functional analyses. Here we explored the reactivity of a set of multivalent electrophiles cofunctionalized with a carbohydrate ligand on gold-nanoparticles to achieve efficient affinity labeling for target protein analysis. Evaluation of the reactivity and selectivity of the electrophiles against three different cognate binding proteins identified aryl sulfonyl fluoride as the most efficient protein-reactive group in this study. We demonstrated that multivalent aryl sulfonyl fluoride probe **4** at 50 nM concentration achieved selective affinity labeling and enrichment of a model protein PNA in cell lysate, which was more effective than photoaffinity probe **1** with aryl azide group. Labeling site analysis by LC-MS/MS revealed that the nanoparticle-immobilized aryl sulfonyl fluoride group can target multiple amino acid residues around the ligand binding site of the target proteins. Our study highlighted the utility of aryl sulfonyl fluoride as a highly effective multivalent affinity label suitable for covalently capturing unknown target proteins.

Introduction

Determining biological targets of small molecules is critical for a comprehensive understanding of their cellular interactome so that we can better predict their utility.[1,2] Protein-reactive chemical probes have become essential tools to facilitate target identification and engagement analysis of small molecules.[3-7] For target identification, affinity labeling probes, in which a protein-reactive group is derivatized on a small molecule of interest, are employed to covalently crosslink otherwise reversible protein-ligand complexes for downstream functional analysis. Since there is no prior knowledge on the structural features of target proteins, it is a significant challenge to design a suitable protein-reactive group that can react with target proteins with high efficiency and exhibits a broad substrate scope, while having no or low reactivity toward non-targets.[1,8,9] However, no generally useful protein-reactive group yet fulfills these requirements of affinity labeling probes.

Two major types of functionalities are currently available as protein-reactive groups. Photoreactive groups such as benzophenone, aryl azide, and diazirine are uniquely suited for target identification based on their highly reactive photo-activated intermediates that can nonselectively react with any proximal amino acid residues.[2,10-12] Although photoaffinity labeling has received a renewed attention as a promising approach, the low labeling yields and selectivity hamper their routine use in detecting low-abundance or low-affinity proteins.[13] Another type of functionality involves electrophilic groups, a wide variety of which are used in affinity labeling,[14-17] activity-based protein profiling (ABPP),[3,18-22] reactivity-based protein profiling,[7,22,23] ligand-directed chemistries,[5,24,25] and covalent drugs[26-30] where a good balance of the reactivity and chemoselectivity of electrophiles enables targeting of locally reactive nucleophilic side chains in proteins such as enzyme active site residues. In contrast to photoreactive groups, the availability of a diverse array of electrophiles with a broad range of reactivity and chemoselectivity represents an attractive resource for researchers in designing protein-reactive probes. In most of earlier examples, these electrophilic groups served as a warhead in the ligand-binding site.[14-17,31,32] More recently, the Hamachi group extensively demonstrated that it is possible to selectively label surface-exposed nucleophilic residues at the periphery of the ligand binding site by various electrophilic probes in a ligand-dependent fashion[5,24,25] However, utilization of electrophiles for crosslinking unknown proteins has been limited to cases where the presence of a nucleophilic amino acid residue at the ligand-binding site of a target protein is known or predictable.[33-36] This is largely due to the inherent reactivity of electrophiles with free nucleophiles and their selectivity toward certain nucleophilic amino acid residues. Other factors determining the reactivity and selectivity of affinity labeling by electrophilic groups include the availability of accessible nucleophilic amino acid residues and their reactivity toward the electrophiles, which is often dependent on the local microenvironment.[5,37,38] Therefore, the general applicability of electrophiles for covalent capture of unknown target proteins with no structural information or prediction remains to be tested.

We recently reported gold nanoparticle-based photoaffinity probes that display small-molecule ligands and photoreactive groups in a multivalent fashion.[39-41] The gold-nanoparticle scaffold promotes high local concentrations of both the ligands

10.1002/anie.202104347

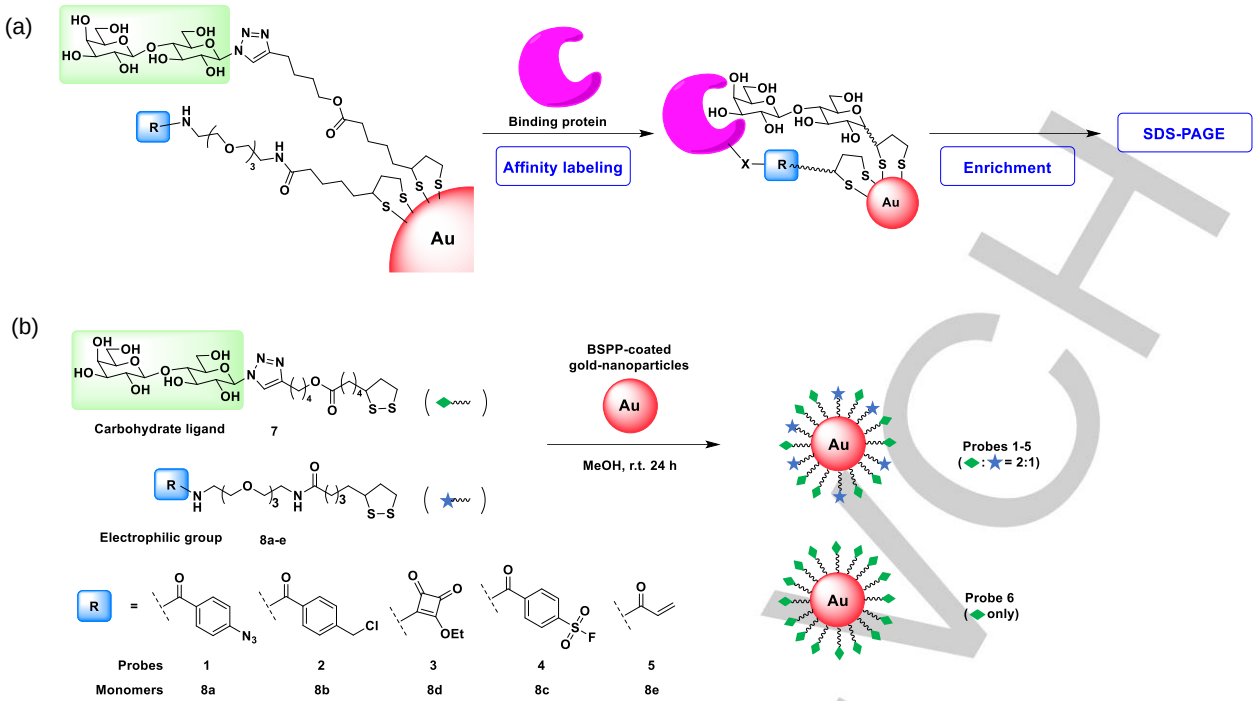


Figure 1. (a) Scheme showing gold nanoparticle-based affinity labeling for identifying specific carbohydrate-binding proteins. (b) Structures of gold-nanoparticle affinity labeling probes with a photoreactive group (**1**), different electrophilic groups (**2**-**5**) and a lactose-immobilized gold-nanoparticle probe (**6**).

and photoreactive groups, and reactions over a large surface area. As a consequence, these probes exert enhanced protein binding affinity and ligand-dependent reactivity. With the goal of expanding the utility of our gold-nanoparticle probes for target identification, we hypothesized that the low labeling efficiency of photoreactive groups could be addressed by employing electrophilic groups as an affinity label instead. A particular advantage of the gold-nanoparticle system is the modular nature of surface functionalization, which allows facile optimization of the probe design. While polar, nucleophilic amino acid residues such as lysine are preponderant on the protein surfaces, whether they happen to be within the reach of a probe greatly affects the outcome of labeling. We anticipated that the multivalent presentation of a suitable electrophilic group should overcome this problem by the avidity effect, allowing broad and simultaneous screening for labeling sites in the proximal area of the ligand binding site.

Here we designed and synthesized a set of gold nanoparticle-immobilized electrophiles as novel multivalent affinity labels and systematically characterized their reactivity, selectivity, and targetable amino acid residues. Among the four different types of electrophilic groups evaluated, aryl sulfonyl fluoride was identified as the most efficient and selective protein reactive group toward three different model carbohydrate-binding proteins. We showed that these gold nanoparticle-immobilized electrophiles can remain unreactive at high dilution (~nM concentration) but become highly reactive toward a protein that is bound to a ligand. The combination of these effects offers an ideal reactive group that can efficiently crosslink target proteins in a pseudo-intramolecular setting while being stable toward bimolecular nucleophilic substitution reactions with nonspecific proteins.[16] Moreover, the multivalent sulfonyl fluoride probe allowed labeling of target proteins at several of their surface residues around the ligand binding site, which demonstrated the utility of multivalency in maximizing protein capture efficiency,[42-46] which is critical for target identification.

Results and Discussion

To explore the potential of gold nanoparticle-based affinity labeling, we first employed lactose as a model low-affinity ligand. Lactose is known to bind several lectins such as peanut agglutinin (PNA), *Erythrina cristagalli* agglutinin (ECA), and *Ricinus communis* agglutinin (RCA) with a binding affinity in the high micromolar range.[47-49] For affinity labels, we employed a panel of four different electrophilic groups, comprising a benzyl chloride (BnCl), squaramide ester (SQ), aryl sulfonyl fluoride (SF), and acrylamide (AA) (Fig. 1). We reasoned that they represent a variety of protein-reactive groups with different reaction types and different sets of target amino acid residues. BnCl potentially reacts with a cysteine residue via SN2 reaction,[50] whereas SQ[51,52] and AA[28,53] serve as Michael acceptors in ABPP probes or protein crosslinking reagents to react with lysine and cysteine residues. SF has been found to be an excellent reactant as a part of covalent drugs or ABPP probes to promote sulfur (IV)-fluoride exchange (SuFEX) reactions[54,55] with several nucleophilic amino acid residues including cysteine, lysine, tyrosine, histidine, serine, and threonine.[56-61] We also employed the aryl azide (ArAz) group[10,62,63] to incorporate a photoaffinity label to provide a point of comparison with our previously developed gold nanoparticle-based photoaffinity probes.[39,40] Several factors were furthermore considered for the design of the nanoparticle probes that could affect their function. The gold-nanoparticles tend to agglomerate in aqueous solution, hence incorporating hydrophilic moiety is often critical. Since our model ligand, lactose is a hydrophilic molecule, it was expected to confer colloidal stability to the nanoparticle probes, as long as the relative ratio of the

10.1002/anie.202104347

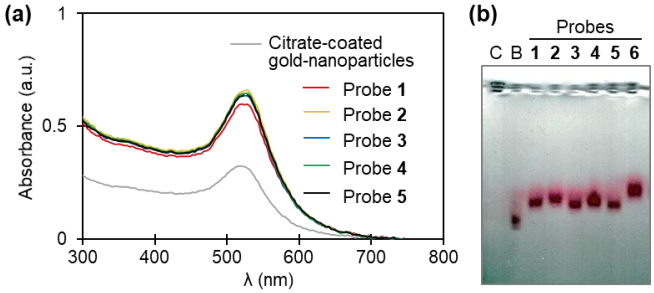


Figure 2. Characterization of functionalized gold-nanoparticle probes **1**-**5**. (a) UV-VIS spectra of citrated-coated gold-nanoparticles and probes **1**-**5**. (b) Agarose gel analysis of probes **1**-**6**. C: Citrate-coated gold-nanoparticles, B: BSPP-stabilized gold-nanoparticles.

electrophilic group, which is usually hydrophobic, was kept relatively low. To achieve optimal binding affinity and labeling efficiency, we anticipated that the ratio of ligand to a reactive group around 2:1 would be a good starting point since the high local concentration of both a ligand moiety and an electrophilic group should promote protein binding as well as efficient labeling reaction. It was also expected that the slightly higher fraction of a hydrophilic ligand to an electrophilic group should confer colloidal stability.

Therefore, we designed and synthesized five types of gold nanoparticle-based probes (**1**-**5**), with a lactose ligand and one of five different groups (four electrophilic groups and one photoreactive group) at a 2:1 ratio to evaluate which group enables efficient crosslinking of binding proteins. The lactose ligand and electrophiles were derivatized with lipoic acid for immobilizing onto the surface of gold-nanoparticles with approximately 13-nm diameter through bivalent Au-S bonds by a two-step ligand exchange method.[39-41] Functionalized gold-nanoparticles were characterized by transmission electron microscopy (TEM), UV-visible spectroscopy (UV-VIS), agarose gel electrophoresis, dynamic light scattering (DLS), and matrix-assisted laser desorption/ionization spectrometry (MALDI-MS) analysis (Fig. 2, Fig. S1-S2, Table S1). To verify that the presence of multivalent ligands on a gold-nanoparticle scaffold increases affinity,[39,64] the K_d values of lactose-functionalized gold-nanoparticle **6** were measured with PNA, ECA, and RCA by the previously developed pull-down assay.[39] Lactose-gold nanoparticle **6** exhibited high affinity toward all three lectins with K_d values of 26 nM (PNA), 8.7 nM (ECA), and 18 nM (RCA) (Fig. S3). Thus the multivalent probe achieved a large affinity enhancement of approximately 15000-fold compared to monomeric lactose (K_d = 400 μM for PNA, 323 μM for ECA, and 37 μM for RCA).[47-49,65] In addition, agglutination of nanoprobes by lectins may contribute to the overall affinity enhancement by forming large complexes composed of three-dimensional networks of lectins and probes. Such complexes are likely to be kinetically stable.[44-46]

The ligand-dependent reactivity of gold-nanoparticle probes **1**-**5** was investigated upon binding with PNA, ECA, or RCA. Probes **2**-**5** at a concentration of 50 nM were incubated with a lectin (0.4 μM as a monomer) at 4 °C for 2 h and the labeled and enriched proteins were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and fluorescence imaging. The control photoaffinity probe **1** was irradiated with UV at 365 nm at 4 °C for 30 min according to our previous study.[39] The labeling efficiency was determined based on the relative fluorescence intensity of a protein band corresponding to the labeled lectin to that of an input lectin in an SDS-PAGE gel. Among the five probes, SF probe **4** gave the highest labeling yields across all three lectins (25-42%) (Fig. 3a, Table S2).[66] ArAz probe **1** showed ~16% labeling of the three lectins. BnCl probe **2** showed little ligand-dependent reactivity toward all lectins. SQ probe **3** and AA probe **5** were rather unreactive toward PNA and ECA but were able to label RCA at 13-15% yield. RCA has 13 cysteine residues, of which Cys20 and Cys39 reside in proximity to lactose binding sites, while PNA and ECA have no cysteine residues.[67,68] Therefore, cysteine residues may be responsible for the higher labeling of RCA than PNA and ECA by SQ probe **3** and AA probe **5**. Importantly, all probes showed low nonselective labeling of BSA, which is not a lactose binding protein and is known to display multiple surface-exposed lysine and cysteine residues.[69] Time-course analysis of the labeling reaction of PNA with probes **2**-**5** showed that a longer reaction time (16 h) allowed SF probe **4** to achieve quantitative labeling (Fig. 3b). No further increase in the labeling products was observed for the other electrophilic probes (**2**, **3,** and **5**). These results underscored the unique reactivity of the SF group, which has sufficient reactivity toward proximal protein residues while being resistant toward hydrolysis.[54,55] Taken together, the comparative reactivity analysis found SF probe **4** to display high ligand-dependent reactivity in general, which was superior to the reactivity profile of ArAz photoaffinity probe **1**.

To evaluate the labeling selectivity, we next assessed the efficiency of affinity labeling by gold-nanoparticle probes **1**-**5** with a mixed solution of PNA and 5-fold excess BSA. Washing with a high concentration of lactose solution[70] confirmed that SF probe **4** indeed covalently labeled PNA with high efficiency (Fig. S4, lane 8). The time-course analysis demonstrated that SF probe **4** maintained labeling selectivity toward PNA over a duration of 16 h (Fig. 3c, Fig. S5). Its labeling yields were slightly lower than when reacted alone with PNA as shown in Fig. 3b, likely due to sequestration by BSA. We next reacted gold-nanoparticle probes (**1**-**5**) with HL-60 cell lysate containing PNA (0.8 w/w%). To ensure enrichment of covalently labeled proteins from unreacted lysate proteins, stringent washing conditions involving 2% SDS solution were applied. Despite the presence of excess amounts of various nonspecific proteins, all probes selectively enriched PNA (Fig. 3d, Fig. S6a). Moreover, SF probe **4** exhibited the highest labeling efficiency (37%) and selectivity, which were superior to those of photoaffinity probe **1** (Fig. 3d, lane1 and 4). The ligand-dependence of PNA labeling by SF probe **4** was verified by a negative control experiment where excess amounts of free lactose ligand inhibited the labeling reaction (Fig. 3d, lane 6). The labeling of PNA was also demonstrated by reacting probes **1**-**5** with fluorescently-tagged, FITC-PNA instead of PNA in cell lysate, which showed probe **4** as most efficient (Fig. S6b-c). These results thus suggested that SF is a promising protein reactive group suited for covalent labeling of target proteins in a complex proteome.

To characterize the amino acid residues targetable by the multivalent affinity labeling probes, we chose the SF probe **4**, which displayed the highest labeling efficiency for all three lectins and low nonspecific labeling, to determine its labeling sites. Since our gold-nanoparticle probes were designed to display both ligand **7** and protein reactive group **8c** in a multivalent fashion, it was anticipated that multiple sites in the proximity of the ligand binding site could be labeled. Using the optimal labeling conditions found

10.1002/anie.202104347

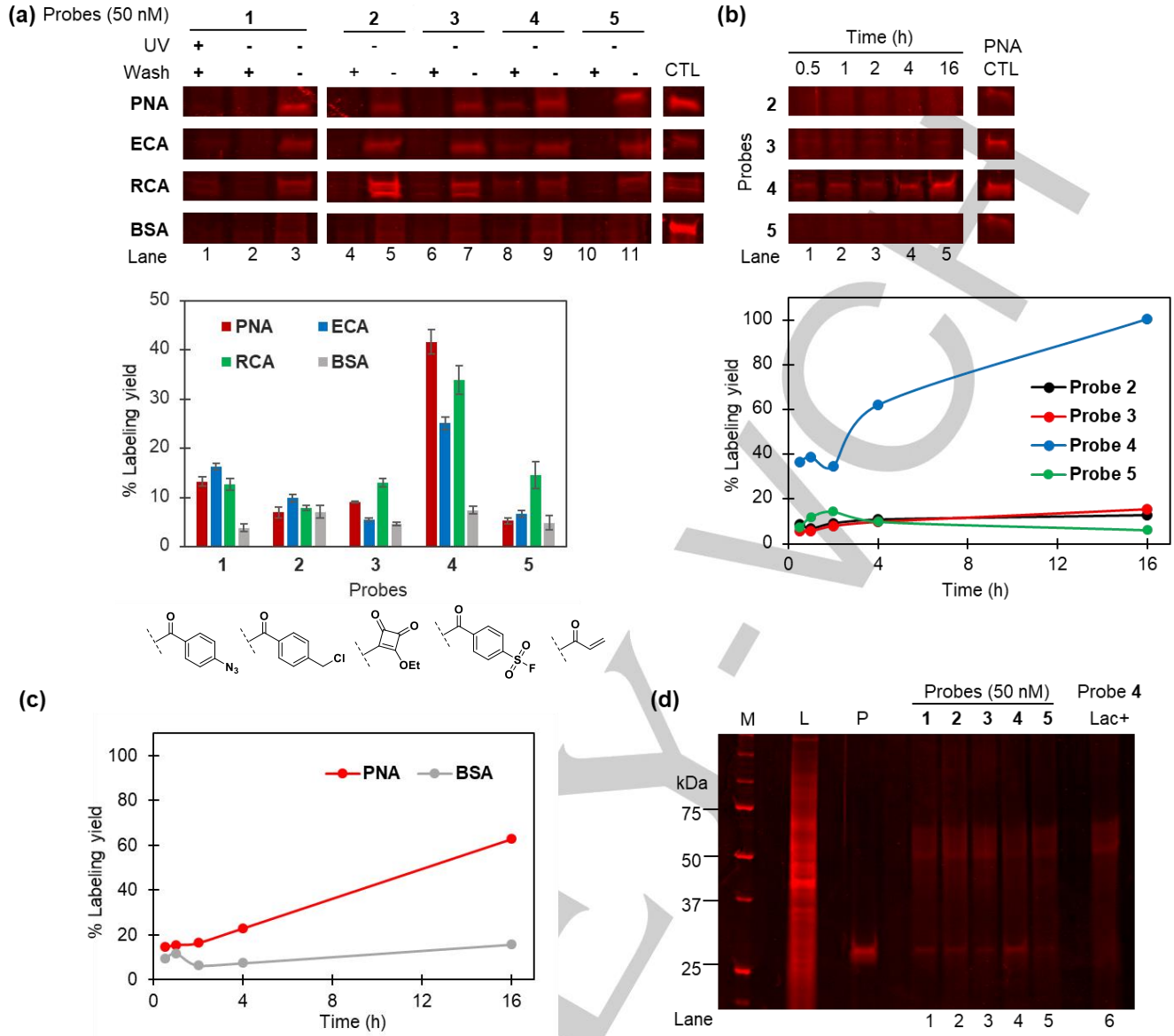


Figure 3. (a) Top: SDS-PAGE analysis of the affinity labeling of PNA, ECA, RCA, and BSA (monomer: 0.4 μM) by probes **1**-**5** (50 nM). CTL: protein used for a given reaction; Wash: washing three times with 0.5 M lactose-1% n-octyl-β-D-glucoside-6.5% PEG8000 in PBS. Bottom: labeling efficiency of each protein by probes **1**-**5**. The % yields were calculated based on the amount of protein used for a given reaction. Error bars show standard errors (n=3). (b) Top: SDS-PAGE time-course analysis of affinity labeling of PNA (0.1 μM) by probes **2**-**5** (50 nM). CTL: PNA used for a given reaction. Bottom: labeling efficiency of PNA by probes **2**-**5**. The % yields were calculated based on the amounts of PNA used for a given reaction. (c) Time course analysis of affinity labeling of PNA (100 nM) by probe **4** (50 nM) in the presence of excess BSA (1.5 μM). (d) SDS-PAGE analysis of affinity labeling of PNA (0.4 μg, 50 nM) spiked in cell lysate (5 μg) using probes **1**-**5** (50 nM). M: molecular weight marker; L: lysate 5 μg (1/10); P: 0.4 μg PNA; Lac+: 20 mM lactose.

in Fig. 3c, 10 μg of PNA or RCA was reacted with probe **4** (46 pmol) at 4 °C for 16 h, and the labeled proteins were purified by a few cycles of centrifugation and stringent washing with excess lactose followed by SDS-PAGE. The gel bands were excised and then subjected to in-gel trypsin digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.[71] Representative MS/MS spectra for a digested peptide fragment carrying a protein-reacted moiety are shown in Fig. 4a. Several peptides were detected with mass values increased by the expected mass values of the protein-reacted moiety **17** or its iodoacetamide derivatives **17a** and **17b**, its fragments **17c** and **17d**, or its oxidized derivative **17e** (Fig. S8; **17**: 547.16, **17a**: 663.22, **17b**: 665.19, **17c**: 359.13, **17d**: 184.99, **17e**: 579.15). The labeled peptides were identified from triplicate experiments, which revealed a remarkable scope of the labeling reaction by multivalent SF probe **4**. Multiple labeled sites close to the ligand binding site in PNA and RCA were detected at several types of amino acid residues (Fig. 4b, c, Fig. S9-10, Table S3). Among the labeled amino acid residues, tyrosine notably exhibited the highest occurrence (three residues for each lectin), which is consistent with previous studies using SF-containing covalent drugs or ABPP probes.[56-61] The SF group has been found to preferentially react with tyrosine and lysine and to a lesser degree with histidine, cysteine, serine, and threonine. Other labeled residues included lysine, aspartic acid, and asparagine, suggesting that SF probe **4** can potentially target diverse types of amino acid residues. These results supported our hypothesis that when the target proteins are specifically bound to ligands on the nanoparticle, multivalent SF groups are positioned in such a way that some can reach out and readily react with matching

10.1002/anie.202104347

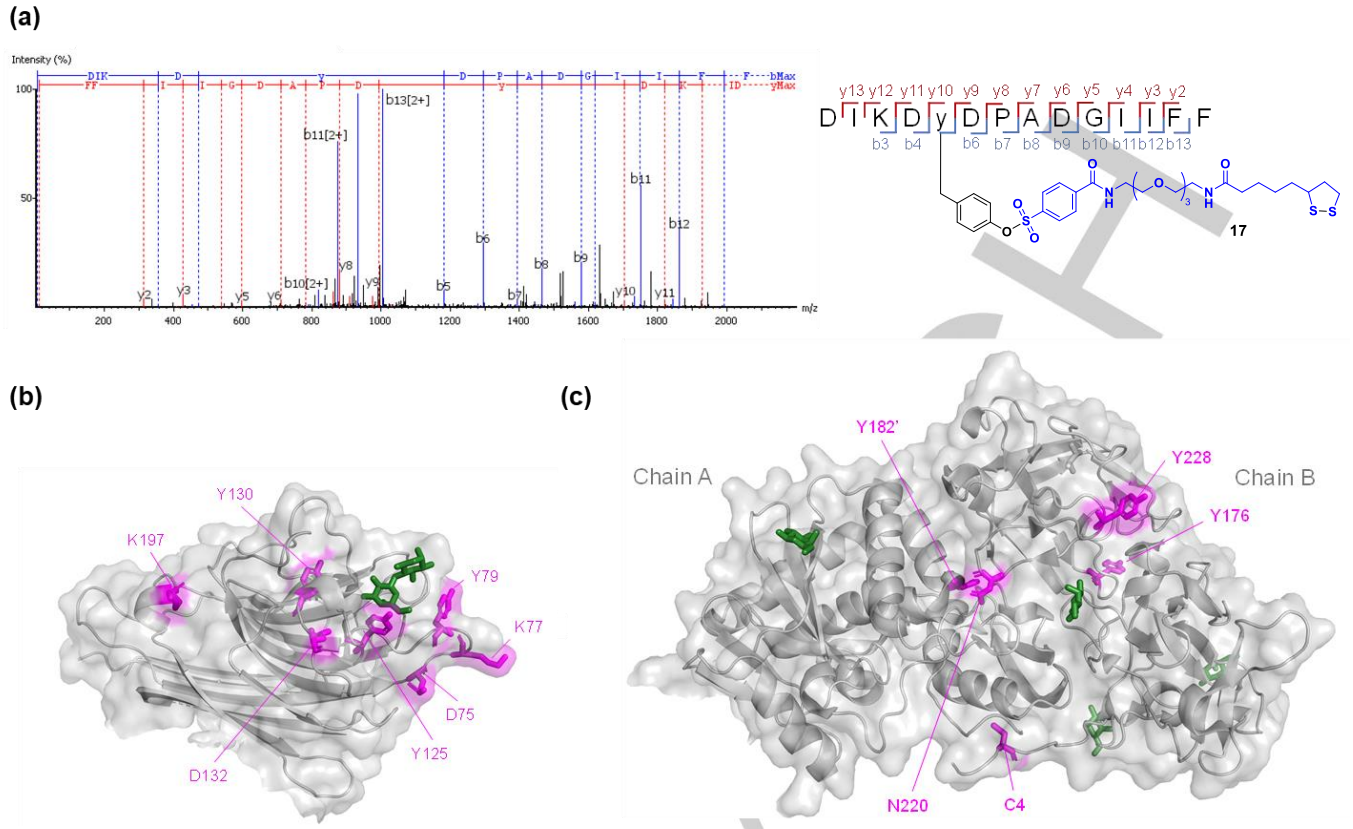


Figure 4. (a) LC-MS/MS spectrum of the peptide fragment labeled with **17**, supporting the labeling of PNA Y79 by SF probe **4**. The inferred labeled amino acid residues mapped onto the X-ray structure of (b) PNA (PDB: 2PEL) and (c) RCA (PDB: 1RZO) are visualized in magenta. Bound lactose (PNA) and galactose (RCA) in the crystal structures are shown as green sticks. The chain A residues for RCA are indicated by ' in the figure.

nucleophilic residues on the protein surface.

When mapped to the X-ray structures of ligand-bound PNA[72] and RCA,[73] the detected labeled sites are all located within approximately 15 Å and 18 Å of their ligand binding sites. The most prominent labeled sites in PNA reside within a loop region harboring Asp75-Lys112, which is in close proximity (~14 Å) to the lactose binding pocket. In RCA, which is a dimer of an α/β heterodimer, major labeled sites were found in the loop region composed of Asp220-Arg236 close to one of the three galactose binding sites in chain B.[67,68] These flexible loop regions may provide readily accessible nucleophilic residues for reacting with the SF groups on the gold-nanoparticle probes. Lysine is known to be one of the most abundant amino acid residues on protein surfaces.[74] Lysine, tyrosine, and aspartic acid, together comprise approximately 24% of the amino acid residues at the surfaces of proteins.[75] If we estimate the contact surface area between probe **4** and its target protein to be an area of a circle with a diameter of ~30 Å, then it seems possible for multivalent SF groups in a nanoparticle probe to find and react with one or more lysine and tyrosine residues in the proximity of the ligand binding site.

It has been shown that the reactivity of electrophiles toward monomeric amino acids is not necessarily reflected by their reactivity toward proteins.[76-78] We wondered whether the labeled amino acid residues in PNA and RCA could be generated by reaction with the SF group due to their inherent reactivity or due to a specific molecular setting in which the protein and the SF group are juxtaposed on a gold-nanoparticle surface. To test this notion, we reacted 4-(fluorosulfonyl) benzoic acid **18** at 1 mM with *N*-acetylated monomeric amino acids at 100 mM in hydroxyethylpiperazine ethanesulfonic acid (HEPES) buffer at pH 7.3 and analyzed the time-dcourse of the formation of the crosslinked products by LC/MS. As shown in Fig. 5a, the formation of an adduct (**23a**) was observed only with *N*-acetyl-lysine **19** after prolonged reaction time. In contrast, *N*-acetyl-tyrosine **20**, *N*-acetyl-aspartic acid **21**, or *N*-acetyl-asparagine **22** did not show reactivity toward a small-molecule sulfonyl fluoride reagent **18** even after 48 hours (Fig. S11). When the reactivity of these *N*-acetylated monomeric amino acids at 100 mM toward probe **4** at 50 nM in HEPES buffer (pH 7.3) was evaluated, none of the tested amino acids yielded an adduct after 48 hours (Fig. 5b, Fig. S12). The SF groups on probe **4** hydrolyzed only slowly over 48 hours. Our finding is consistent with the literature that aryl sulfonyl fluoride group is relatively stable in neutral aqueous solution and yet can be readily activated toward nucleophilic amino acid residues of the bound proteins.[38, 55] According to the mechanism of affinity labeling,[16] the different reactivity toward probe **4** observed for monomeric amino acid residues and the amino acid residues on the protein may be explained by the increased effective molarity of the latter relative to the SF groups. Additionally, the avidity effect of multivalent SF groups may contribute to a selective increase in the rate of affinity labeling by keeping the probe concentration at nanomolar. It thus suggested that the unique reactivity feature of aryl sulfonyl fluoride, with combined effects of the increased effective molarity and avidity effect promoted highly efficient affinity labeling by probe **4**. It is also possible that some of the labeled amino acids identified in PNA and RCA represent hyperactivated residues[79,80] by

10.1002/anie.202104347

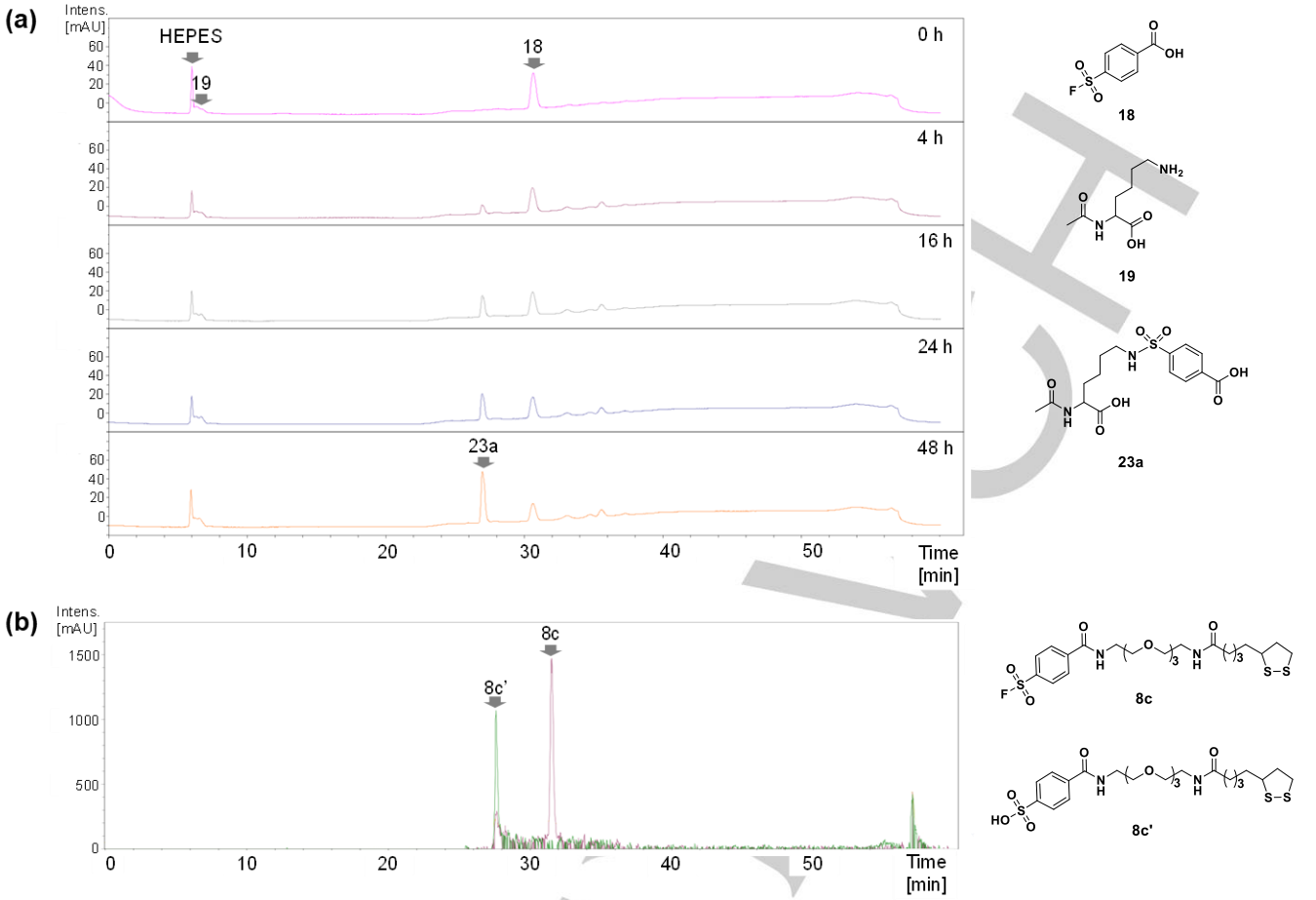


Figure 5. (a) HPLC traces monitoring the solution reactivity of **18** toward *N*-acetyl-lysine **19** at λ = 254 nm. Reaction mixtures after 0, 4, 16, 24, and 48 h were analyzed. **23a** was confirmed by ESI-MS. **23a**: LRMS (ESI-TOF) calculated for $C_{15}H_{21}N_2O_7S$ $[M+H]^+$: 373.11; found 373.11. (b) Extracted ion chromatograms for m/z=567.1669±0.0500 (**8c**: $C_{23}H_{36}FN_2O_7S_3$ $[M+H]^+$) and 565.1712±0.0500 (**8c'**: $C_{23}H_{37}N_2O_8S_3$ $[M+H]^+$) present in the reaction mixture generated by incubating SF probe **4** and **19** for 48 h.

proximal amino acid residues, which display inherently high reactivity toward electrophilic reagents.[81]

Conclusion

In summary, we developed a novel gold nanoparticle-based multivalent approach to covalently capture target proteins by electrophilic groups. Evaluating four different electrophilic groups against three different lectins identified aryl sulfonyl fluoride group as the most efficient and selective affinity label. Significantly, we showed that multivalent aryl sulfonyl fluoride probe **4** at 50 nM concentration achieved selective affinity labeling of PNA in cell lysate, which was more effective than photoaffinity probe **1** with aryl azide group. It thus demonstrated a promising utility of multivalent aryl sulfonyl fluoride affinity probe as a complementary target identification tool. LC-MS/MS analysis revealed that the gold nanoparticle-immobilized aryl sulfonyl fluoride groups can label several amino acid residues around the ligand binding site. It supported the view that multivalent display of electrophilic groups should lead to efficient affinity labeling due to the avidity effect by providing a large reactive surface area toward a target protein. Further analysis with monomeric amino acid derivatives validated that the aryl sulfonyl fluoride groups on gold-nanoparticles exhibit high reactivity toward various amino acids only when the proteins are bound to the nanoparticle surface. Notably, that SF group did not show any cross-reactivity with the lactose, a polyol moiety despite the proximity.[82] The cross-reactivity between the ligand moiety and electrophilic groups co-functionalized on the nanoparticles is undesirable as it would result in deactivated affinity labeling probes before use. When ligands contain nucleophilic functionalities such as phenols or amines, they are expected to display cross-reactivity with the electrophilic moiety cofunctionalized on the nanoparticles. In such cases, it may be more beneficial to pursue photoaffinity labeling for target analysis with photoreactive groups instead of electrophilic groups as labeling groups. Alternatively, the cross-reactivity between a ligand and an electrophilic group may be circumvented by optimizing linker length so that they would not be positioned to react with each other. Taken together, these results highlighted the unique reactivity of aryl sulfonyl fluoride group suited as a multivalent affinity label. Our findings have

implications for the new design of chemical probes for target protein analysis.

Acknowledgments

The authors thank Prof. Keiichi Noguchi for technical assistance with LC/MS analysis and TEM analysis, Prof. Naoyuki Sugiyama, and Dr. Seketsu Fukuzawa for helpful discussion on the LC-MS/MS analysis. This study was supported by KAKENHI 18K05331 and 17H06110 by JSPS. The authors are grateful for supports from JSPS A3 Foresight Program, ACBI and GIR programs.

Keywords: Affinity labeling • protein modifications • carbohydrates • nanoparticles • multivalency

[1] M. Schenone, V. Dančík, B. K. Wagner, P. A. Clemons, *Nat. Chem. Biol.* **2013**, *9*, 232-240.
[2] S. Ziegler, V. Pries, C. Hedberg, H. Waldmann, *Angew. Chem. Int. Ed.* **2013**, *52*, 2744-2792, *Angew. Chem.* **2013**, *125*, 2808-2859.
[3] D. Greenbaum, A. Baruch, L. Hayrapetian, Z. Darula, A. Burlingame, K. F. Medzihradszky, M. Bogyo, *Mol. Cell. Proteomics* **2002**, *1*, 60-68.
[4] B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383-414.
[5] K. Shiraiwa, R. Cheng, H. Nonaka, T. Tamura, I. Hamachi, *Cell Chem. Biol.* **2020**, *27*, 970-985.
[6] M. H. Wright, S. A. Sieber, *Nat. Prod. Rep.* **2016**, *33*, 681-708.
[7] A. M. Roberts, C. C. Ward, D. K. Nomura, *Curr. Opin. Biotechnol.* **2017**, *43*, 25-33.
[8] A. Blencowe, W. Hayes, *Soft Matter*, **2005**, *1*, 178-205.
[9] W. Zheng, G. Li, X. Li, *Arch. Pharm. Res.* **2015**, *38*, 1661-1685.
[10] F. Kotzyba-Hibert, I. Kapfer, M. Goeldner, *Angew. Chem. Int. Ed.* **1995**, *34*, 1296-1312, *Angew. Chem.* **1995**, *107*, 1391-1408.
[11] S. S. Ge, B. Chen, Y. Y. Wu, Q. S. Long, Y. L. Zhao, P. Y. Wang, S. Yang, *RSC Adv.* **2018**, *8*, 29428-29454.
[12] J. P. Holland, M. Gut, S. Klingler, R. Fay, A. Guillou, *Chem. Eur. J.* **2020**, *26*, 33-48.
[13] M. Hashimoto, Y. Hatanaka, *Eur. J. Org. Chem.* **2008**, *15*, 2513-2523.
[14] L. Wofsy, H. Metzger, S. J. Singer, *Biochemistry* **1962**, *1*, 1031-1039.
[15] F. Wold, *Methods Enzymol.* **1977**, *46*, 3-14.
[16] D. Givol, M. Wilchek, in *Encyclopedia of Immunology 2nd Ed.,* Vol. 2 (Eds.: P. Delve, I. Roitt), Academic Press, United States, **1998**, pp 50-52.
[17] T. Hayashi, I. Hamachi, *Acc. Chem. Res.* **2012**, *45*, 1460-1469.
[18] M. Uttamchandani, J. Li, H. Sun, S. Q. Yao, S. Q. *ChemBioChem* **2008**, *9*, 667-675.
[19] M. J. Niphakis, B. F. Cravatt, *Annu. Rev. Biochem.* **2014**, *83*, 341-377.
[20] P. Yang, K. Liu, *ChemBioChem* **2015**, *16*, 712-724.
[21] B. P. Kok, S. Ghimire, W. Kim, S. Chatterjee, T. Johns, S. Kitamura, J. Eberhardt, D. Ogasawara, J. Xu, A. Sukiasyan, S. M. Kim, C. Godio, J. M. Bittencourt, M. Cameron, A. Galmozzi, S. Forli, D. W. Wolan, B. F. Cravatt, D. L. Boger, E. Saez, *Nat. Chem. Biol.* **2020**, *16*, 997-1005.
[22] H. J. Benns, C. J. Wincott, E. W. Tate, M. A. Child, *Curr. Opin. Chem. Biol.* **2021**, *60*, 20-29.
[23] C. C. Ward, J. I. Kleinman, S. M. Brittain, P. S. Lee, C. Y. S. Chung, K. Kim, Y. Petri, J. R. Thomas, J. A. Tallarico, J. M. McKenna, M. Schirle, D. K. Nomura, *ACS Chem. Biol.* **2019**, *14*, 2430-2440.
[24] S. Tsukiji, I. Hamachi, *Curr. Opin. Chem. Biol.* **2014**, *21*, 136-143.
[25] T. Tamura, T. Ueda, T. Goto, T. Tsukidate, Y. Shapira, Y. Nishikawa, A. Fujisawa, I. Hamachi, *Nat. Commun.* **2018**, *9*, 1870.
[26] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug Discov.* **2011**, *10*, 307-317.
[27] J. M. Bradshaw, J. M. McFarland, V. O. Paavilainen, A. Bisconte, D. Tam, V. T. Phan, S. Romanov, D. Finkle, J. Shu, V. Patel, T. Ton, X. Li, D. G. Loughhead, P. A. Nunn, D. E. Karr, M. E. Gerritsen, J. O. Funk, T. D. Owens, E. Verner, K. A. Brameld, R. J. Hill, D. M. Goldstein, J. Taunton, *Nat. Chem. Biol.* **2015**, *11*, 525-531.
[28] T. A. Baillie, *Angew. Chem. Int. Ed.* **2016**, *55*, 13408-13421, *Angew. Chem.* **2016**, *128*, 13606-13619.
[29] E. A. Grossman, C. C. Ward, J. N. Spradlin, L. A. Bateman, T. R. Huffman, D. K. Miyamoto, J. I. Kleinman, D. K. Nomura, *Cell Chem. Biol.* **2017**, *24*, 1368-1376.e4.
[30] N. Shindo, H. Fuchida, M. Sato, K. Watari, T. Shibata, K. Kuwata, C. Miura, K. Okamoto, Y. Hatsuyama, K. Tokunaga, S. Sakamoto, S. Morimoto, Y. Abe, M. Shiroishi, J. M. M. Caaveiro, T. Ueda, T. Tamura, N. Matsunaga, T. Nakao, S. Koyanagi, S. Ohdo, Y. Yamaguchi, I. Hamachi, M. Ono, A. Ojida, *Nat. Chem. Biol.* **2019**, *15*, 250-258.
[31] J. L. Counihan, B. Ford, D. K. Nomura, *Curr. Opin. Chem. Biol.* **2016**, *30*, 68-76.
[32] M. J. C. Long, Y. Aye, *Cell Chem. Biol.* **2017**, *24*, 787-800.
[33] B. G. Hoffstrom, A. Kaplan, R. Letso, R. S. Schmid, G. J. Turmel, D. C. Lo, B. R. Stockwell, *Nat. Chem. Biol.* **2010**, *6*, 900-906.
[34] P. Y. Yang, K. Liu, M. H. Ngai, M. J. Lear, R. Wenk, S. Q. Yao, *J. Am. Chem. Soc.* **2010**, *132*, 656-666.
[35] T. Wirth, K. Schmuck, L. F. Tietze, S. A. Sieber, *Angew. Chem. Int. Ed.* **2012**, *51*, 2874-2877, *Angew. Chem.* **2012**, *124*, 2928-2931.
[36] J. A. Clulow, E. M. Storck, T. Lanyon-Hogg, K. A. Kalesh, L. H. Jones, E. W. Tate, *Chem. Commun.* **2017**, *53*, 5182-5185.
[37] U. Haedke, E. V. Küttler, O. Vosyka, Y. Yang, S. H. L. Verhelst, *Curr. Opin. Chem. Biol.* **2013**, *17*, 102-109.
[38] O. O. Fadeyi, L. R. Hoth, C. Choi, X. Feng, A. Gopalsamy, E. C. Hett, R. E., Jr. Kyne, R. P. Robinson, L. H. Jones, *ACS Chem. Biol.* **2017**, *12*, 2015-2020.
[39] K. Sakurai, Y. Hatai, A. Okada, *Chem. Sci.* **2016**, *7*, 702-706.
[40] K. Sakurai, A. Kato, K. Adachi, *Bioorg. Med. Chem.* **2018**, *28*, 3227-3230.
[41] K. Mori, K. Sakurai, *Org. Biomol. Chem.* **2021**, *19*, 1268-1273.
[42] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Angew. Chem. Int. Ed.* **2006**, *45*, 2348-2368, *Angew. Chem.* **2006**,*118*, 2408-2429.
[43] M. Mammen, S. K. Choi, G. M. Whitesides, *Angew. Chem. Int. Ed.* **1998**, *37*, 2754- 2794, *Angew. Chem.* **1998**, *110*, 2908-2953.
[44] P. I. Kitov, D. R. Bundle, *J. Am. Chem. Soc.* **2003**, *125*, *52*, 16271-16284.
[45] J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555-578.
[46] R. J. Pieters, *Org. Biomol. Chem.* **2009**, *7*, 2013-2025.
[47] A. J. Cagnoni, O. Varela, M. L. Uhrig, J. Kovensky, *Eur. J. Org. Chem.* **2013**, *5*, 972-983.
[48] C. Svensson, S. Teneberg, C. L. Nilsson, A. Kjellberg, F. P. Schwarz, N. Sharon, U. Krengel, *J. Mol. Biol.* **2002**, *321*, 69-83.
[49] S. Sharma, S. Bharadwaj, A. Surolia, S. K. Podder, *Biochem. J.* **1998**, *333*, 539-542.
[50] G. Chen, A. Heim, D. Riether, D. Yee, Y. Milgrom, M. A. Gawinowicz, D. Sames, *J. Am. Chem. Soc.* **2003**, *125*, *27*, 8130-8133.
[51] C. Dingels, F. Wurm, M. Wagner, H. A. Klok, H. Frey, *Chem. Eur. J.* **2012**, *18*, 16828-16835.
[52] M. Fonvielle, N. Sakkas, L. Iannazzo, C. L. Fournis, D. Patin, D. Mengin-Lecreulx, A. El-Sagheer, E. Braud, S. Cardon, T. Brown, M. Arthur, M. Etheve-Quelquejeu, *Angew. Chem. Int. Ed.* **2016**, *55*, 13553-13557, *Angew. Chem.* **2016**,*128*, 13751-13755.
[53] K. M. Backus, B. E. Correia, K. M. Lum, S. Forli, B. D. Horning, G. E. González-Páez, S. Chatterjee, B. R. Lanning, J. R. Teijaro, A. J. Olson, D. W. Wolan, B. F. Cravatt, *Nature*, **2016**, *534*, 570-574.
[54] J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2014**, *53*, 9430-9448, *Angew. Chem.* **2014**,*126*, 9545-9545
[55] A. S. Barrow, C. J. Smedley, Q. Zheng, S. Li, J. Dong, J. E. Moses, *Chem. Soc. Rev*. **2019**, *48*, 4731-4758.
[56] R. F. Colman, *FASEB J.* **1997**, *11*, 217-226.
[57] E. C. Hett, H. Xu, K. F. Geoghegan, A. Gopalsamy, R. E. Kyne, C. A. Menard, A. Narayanan, M. D. Parikh, S. Liu, L. Roberts, R. P. Robinson, M. A. Tones, L. H. Jones, *ACS Chem. Biol.* **2015**, *10*, 1094-1098.
[58] A. Narayanan L. H. Jones, *Chem. Sci.* **2015**, *6*, 2650-2659.
[59] Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame, J. Taunton, *J. Am. Chem. Soc.* **2017**, *139*, *2*, 680-685.
[60] B. Yang, H. Wu, P. D. Schnier, Y. Liu, J. Liu, N. Wang, W. F. DeGrado, L. Wang, *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 11162-11167.
[61] L. H. Jones, J. W. Kelly, *RSC Med. Chem.* **2020**, *11*, 10-17.

[62] E. Leyva, M. S. Platz, G. Persy, J. Wirz, *J. Am. Chem. Soc.* **1986**, *108*, *13*, 3783-3790.

[63] ArAz group serves as a photoaffinity label by generating nitrene upon UV irradiation. It is also known to further isomerize from nitrene into highly electrophilic dehydroazepine via a ring-expansion, which can react with lysine residues.

[64] D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo, D. Zhou, *J. Am. Chem. Soc.* **2020**, *142*, *42*, 18022-18034.

[65] The K_d values calculated based on the concentration of multivalent probe **6** are compared with those of monomeric lactose because the apparent binding affinity per probe is important for the apparent labeling efficiency of the multivalent probes.

[66] The band tailing observed is likely due to the presence of the lipoate moiety, which co-eluted with the labeled proteins upon cleavage of the functionalities from the gold-nanoparticle surface.

[67] R. Banerjee, K. Das, R. Ravishankar, K. Suguna, A. Surolia, M. Vijayan, *J. Mol. Biol.* **1996**, *259*, 281-296. (PDB: 2PEL)

[68] A. G. Gabdoulkhakov, Y. Savochkina, N. Konareva, R. Krauspenhaar, S. Stoeva, S. V. Nikonov, W. Voelter, C. Betzel, A. M. Mikhailov, DOI:10.2210/pdb1RZO/pdb. The Protein Data Bank, 2004. (PDB:1RZO)

[69] K. A. Majorek, P. J. Porebski, A. Dayal, M. D. Zimmerman, K. Jablonska, A. J. Stewart, M. Chruszcz, W. Minor, *Mol. Immunol.* **2012**, *52*, 174-182.

[70] H. Otsuka, Y. Akiyama, Y. Nagasaki, K. Kataoka, *J. Am. Chem. Soc.* **2001**, *123*, *34*, 8226-8230.

[71] V. L. Mendoza, R. W. Vachet, *Mass Spectrom. Rev.* **2009**, *28*, 785-815.

[72] N. V. Konareva, A. G. Gabdulkhakov, S. Eschenburg, S. Stoeva, A. N. Popov, R. Krauspenhaar, M. E. Andrianova, Y. Savochkina, I. I. Agapov, A. G. Tonevitskii, A. N. Kornev, V. V. Kornilov, V. N. Zaitsev, W. Voelter, C. Betzel, S. V. Nikonov, B. K. Vainshtein, A. M. Mikhailov, *Crystallogr. Rep.* **2001**, *46*, 792-800.

[73] RCA, which is a dimer of an α/β heterodimer, has one galactose binding site in chain A and three in chain B, of which that of chain B involving Glu35, Lys40, and Asp46 is considered a primary binding site.

[74] T. E. Creighton, *Proteins: structures and molecular properties*, WH Freeman, New York, **1993**, p2-4.

[75] A. D. White, A. K. Nowinski, W. Huang, A. J. Keefe, F. Sun, S. Jiang, *Chem. Sci.* **2012**, *3*, 3488-3494.

[76] E. Weerapana, G. M. Simon, B. F. Cravatt, *Nat. Chem. Biol.* **2008**, *4*, 405-407.

[77] L. M. McGregor, M. L. Jenkins, C. Kerwin, J. E. Burke, K. M. Shokat, *Biochemistry* **2017**, *56*, 3178-3183.

[78] L. Petri, P. Ábrányi-Balogh, I. Tímea, G. Pálfy, A. Perczel, D. Knez, M. Hrast, M. Gobec, I. Sosič, K. Nyíri, B. G. Vértessy, N. Jänsch, C. Desczyk, F. J. Meyer-Almes, I. Ogris, S. G. Grdadolnik, L. G. Iacovino, C. Binda, S. Gobec, G. M. Keserű, *ChemBioChem* **2020**, *22*, 743-753.

[79] E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. D. Dillon, D. A. Bachovchin, K. Mowen, D. Baker, B. F. Cravatt, *Nature* **2010**, *468*, 790-795.

[80] S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia, B. F. Cravatt, *Nat. Chem.* **2017**, *9*, 1181-1190.

[81] An electrostatic representation of PNA suggested that the protein surface around the labeled sites were negatively charged.

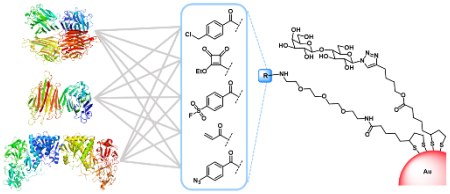
[82] A mass data search of the LC/MS data in Fig. 5(b) did not provide a mass peak corresponding to a crosslinked product that could be formed between lactose-PEG-lipoate **7** and sulfonyl fluoride-PEG lipoate **8c**, indicating that the lactose ligand moiety of **7** did not cross-react with aryl sulfonyl fluoride of **8c** on the same nanoparticle.

10.1002/anie.202104347

WILEY-VCH

RESEARCH ARTICLE

Entry for the Table of Contents



Chemical probes were prepared by multivalently presenting electrophilic groups and ligands on gold-nanoparticles and their reactivities as novel affinity labeling probes were explored. Compared to a photoaffinity probe bearing aryl azide group, aryl sulfonyl fluoride probe at a nanomolar concentration achieved more efficient and selective labeling of model proteins in cell lysate. Labeling site analysis revealed that the nanoparticle-immobilized aryl sulfonyl fluoride groups can target multiple amino acid residues around the ligand binding site of the target proteins.

Institute and/or researcher Twitter usernames: @TUAT_all