Proteomics

DOI: 10.1002/anie.201408580

An Isotope-Coded Fluorogenic Cross-Linker for High-Performance Target Identification Based on Photoaffinity Labeling**

Takenori Tomohiro, Shota Morimoto, Toshiya Shima, Junya Chiba, and Yasumaru Hatanaka*

Abstract: *A photoaffinity labeling (PAL)-based method for the rapid identification of target proteins is presented in which a high-performance chemical tag, an isotope-coded fluorescent tag (IsoFT), can be attached to the interacting site by irradiation. Labeled peptides can be easily distinguished among numerous proteolytic digests by sequential detection with highly sensitive fluorescence spectroscopy and mass spectrometry. Subsequent MS/MS analysis provides amino acid sequence information with a higher depth of coverage. The combination of PAL and heterogeneous target-selecting techniques significantly reduces the amount of time and protein required for identification. An additional photocleavable moiety successfully accelerated proteomic analysis using cell lysate. This method is a widely applicable approach for the rapid and accurate identification of interacting proteins.*

The identification of the target protein of a ligand and a structure-based analysis of its interaction site are indispensable for a fundamental understanding of the cellular functions of both proteins and small molecules in a living system.[1] Photoaffinity labeling (PAL) is a powerful chemical tagging technique that is used as a straightforward method for identifying specific ligand–receptor interactions and the ligand-binding site by using a photoactivatable ligand to covalently cross-link the interacting molecules.[2] It can be applied to low-affinity proteins, membrane-bound proteins, and protein complexes that are difficult to access by conventional affinity-based purification methods. Typical PAL-based target protein identification is achieved by peptide mass fingerprinting (PMF) analysis of proteolytic digestion products of the purified labeled protein. Accurate identification has generally required sequencing the labeled peptide that is part of the interacting surface. Recent efforts have focused on efficient purification of the labeled products with several separation tags, such as a biotin,[3] a perfluoroalkyl group,[4] a clickable tag (e.g., azide/ethynyl groups) as a post-labeling technique,[5] or an oligonucleotide tag.[6] However, the purification process is often complicated by inevitable contamination that is due to non-specific adsorption and significant loss during handling because of the different physical properties of individual peptide fragments. This compromises the ability to identify the target peak in LC-MS profiles, although high-performance MS equipment and analysis software can be used for exhaustive analysis of all fragments.

Recently, we reported a novel technique for fluorescence labeling at the cross-linked position of an interacting protein by sequential photoreactions.[7] This technique enabled the selective detection of the labeled fragments with high sensitivity as fluorescence peaks in the HPLC profile of the proteolytic digests without any further chemical post-treatments and accelerated the analytical process by LC-MS.[8] However, the resulting MS and MS/MS profiles still remain complex in proteomics, owing to contaminants in the fluorescence peaks. Alternatively, the use of stable-isotope labeling in combination with MS has been an efficient method for quantitative proteomics and allows low-abundance labeled peptides to be selectively detected by their distinctive isotopic patterns after enzymatic digestion.[9] Examples of stable-isotope labeling techniques include the use of mixed isotope tags [isotope-coded affinity tag (ICAT),[10] stable-isotope labeling with amino acids in cell culture (SILAC)],[11] isobaric tags [isobaric tags for relative and absolute quantitation (iTRAQ),[12] deuterium isobaric amine reactive tags (DiART)],[13] and fluorous PAL as an isotope-coded photoaffinity tag with a perfluoroalkyl chain.[14] Finally, a stable-isotope-based method has been integrated with our fluorophore-tagging technique, which uses an isotope-coded fluorescent tag (IsoFT) as a fully functionalized cross-linker to enable the rapid identification of the labeled peptides without requiring highly purified targets (Figure 1).

The cross-linker unit consists of a diazirinyl group and an *ortho*-hydroxycinnamate unit with an ethyl group at the α-position. The *ortho*-hydroxycinnamate unit undergoes *E* to *Z* isomerization under UV irradiation to form a coumarin fluorophore by an intramolecular substitution reaction with concurrent release of a ligand molecule.[7b,8] The fluorescence tagging occurs only by irradiation and does not require further chemical treatments. The combined use of an ethyl group introduced at the C3 position of coumarin with a deuterated ethyl group yielded a doublet peak with a mass difference of 5.03 u ($z = 1$) for labeled products in the electrospray ionization (ESI) mass spectrum (Figure 1). While the reaction unit has many functions that are advantageous for LC-MS-based target identification, the structure maintains the size of a typical cross-linker unit, which thus has minimal effects on the affinity of the ligand for the target protein. Herein, we evaluated the performance of IsoFT for

[*] Dr. T. Tomohiro, S. Morimoto, T. Shima, Dr. J. Chiba, Dr. Y. Hatanaka
Graduate School of Medicine and Pharmaceutical Sciences
University of Toyama
2630 Sugitani, Toyama 930-0194 (Japan)
E-mail: ttomo@pha.u-toyama.ac.jp

[**] This work was financially supported by Grants-in-Aid for Scientific Research (20390032) from MEXT, the Tamura Science and Technology Foundation, and the Fugaku Trust for Medicinal Research. We thank Dr. Tomoko Okada, AIST, for her kind support with preparation of the HeLa cell lysate.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201408580.

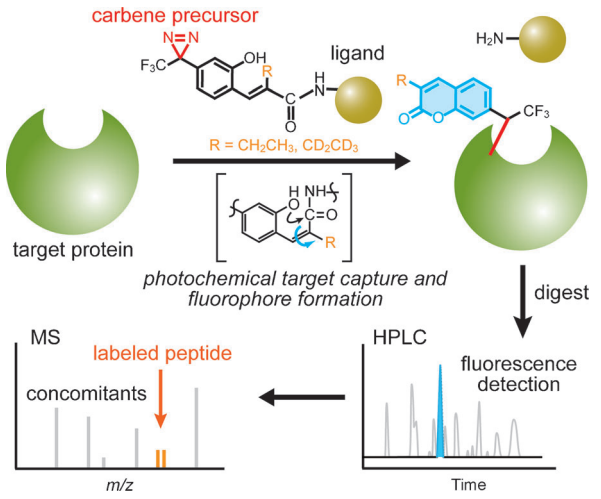

Figure 1. A new strategy for target identification using PAL with IsoFT simplifies the identification of the target peak in both HPLC and MS analyses.

target identification through specific PAL capture of avidin with biotin probe **1** (Figure 2) in a four-protein mixture and then in a mixture of the biotin-dependent enzymes in the proteome.

The probe was prepared by coupling the cross-linker unit with biotin through an ethylene diamine moiety (see the Supporting Information). Probes **1a** and **1b** were dissolved in 5% DMSO/water at a 1:1 molar ratio to prepare the stock solution. The probe (final concentration = 0.1 μM) was incubated with a protein mixture of avidin, bovine serum albumin, transferrin, and carbonic anhydrase (50 ng each) for one hour at room temperature in a 100 mM PBS solution (5 μL). Then, the sample solution was irradiated with 365 nm light by using a 250 W high-pressure mercury lamp with a bandpass filter (fwhm = 10 nm) for 15 seconds at 0°C. After reductive alkylation through treatment with iodoacetamide and removal of small compounds by ultrafiltration (Amicon Ultra-0.5 10 K), the sample solution was irradiated with 365 nm light for 15 minutes at 37°C. The photoproducts were digested with trypsin/Lys-C overnight at 37°C. The tryptic digestion products were then analyzed by LC-MS/MS.

Detection of the fluorescence emission at 410 nm (λ_{ex} = 340 nm) clearly showed two major peaks at 17.3 and 19.5 minutes, which correspond to coumarin-labeled products (Figure 2A, lower HPLC profile), among numerous unlabeled fragments (upper HPLC profile). These fluorescence peaks showed a decrease in the amount of photoproducts formed in the presence of biotin as a competitor (Supporting Information, Figure S1), which indicates the specificity of the probe. Among the MS peaks in the second fluorescence peak, we chose a doublet peak at m/z 1122.5 and 1124.9 (z = 2) as the target, which was expected to be a peptide fragment from residues 27–45 of avidin (2245.0 and 2250.0 for $[M+H]^+$ + coumarin, Figure 2B). The MS/MS spectrum acquired for the singly charged $[M+H]^+$ ions displayed series of b_n+255 and 260 u ions and y_n+255 and 260 u ions with a coumarin modification when it was recorded with an isolation width of 9 m/z units around the mean value of both MS peaks (Figure 2C). The mass difference, Δm = 5 u, may be suitable for MS/MS-based sequencing, as the number of MS peaks can be limited by reducing the mass range around the target peak for isolation. All of the coumarin-labeled fragments appeared as doublet peaks, whereas unmodified fragments appeared as singlet peaks. The mass numbers and differences of the doublet peaks (b_{15}–b_{18}, y_5–y_{18}) should first be confined so that complete sequencing with a high depth of coverage can be achieved without difficulty. The labeled site was identified as Ser41 (indicated by an asterisk) at position 15 on the N-terminus of the 19-amino-acid peptide (GEFTGTYTTAVTATS*NEIK). The small coumarin tag allows for analysis of all MS/MS fragments in a relatively small mass range without

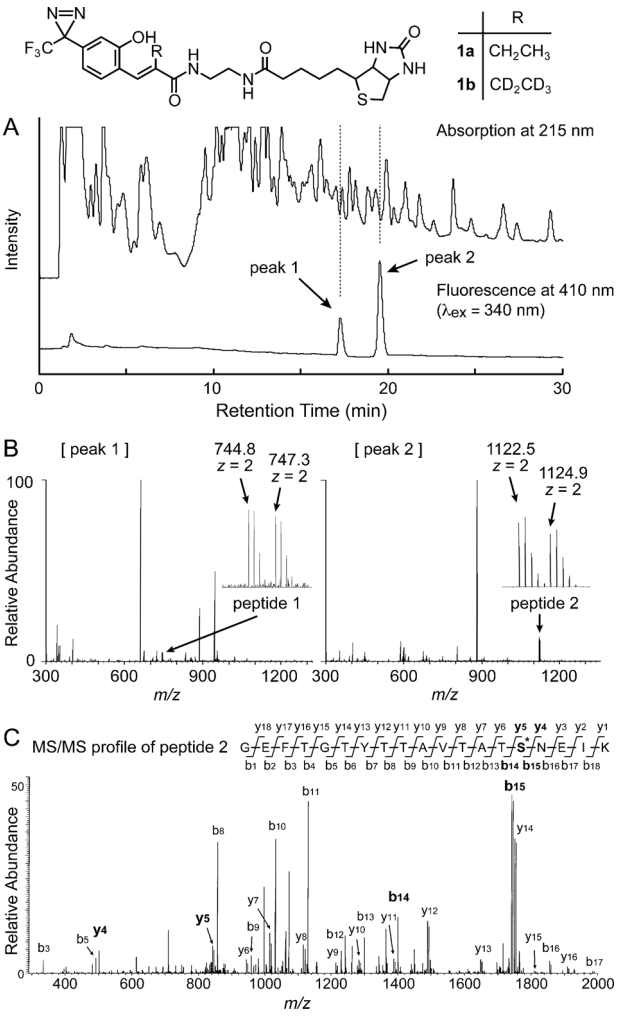

Figure 2. Structure of photoprobe **1** and analysis of photoproducts in the protein mixture. A) Reverse-phase HPLC profiles of digested products detected by absorption at 215 nm (upper trace) and fluorescence at 410 nm (λ_{ex} = 340 nm, lower trace). In this experiment, the sample was prepared in a PBS solution (10 μL) with probe **1** (1 μM) and several proteins, including avidin, bovine serum albumin, transferrin, and carbonic anhydrase (1 μg each), to detect all peptide fragments by UV absorption. B) ESI-MS profile of peaks 1 and 2. C) Tandem MS/MS profile of peptide 2. The spectrum displays the b_n+255 and 260 u as well as the y_n+255 and 260 u series of coumarin-tagged fragments as doublet peaks. The labeled serine residue is indicated by an asterisk.

decreasing the MS resolution. Furthermore, this tag was stable under denaturation conditions, digestion procedures, and collision-induced dissociation at typical energy levels, and it could succeed in avoiding complex fragmentation by releasing the ligand molecule (Figure 2C). The labeled peptide 1 in the first fluorescence peak was identified as SSVN*DIGDDWK (residues 101–111, Asn104 indicated by an asterisk was the labeled site). These labeled positions are at reasonable sites on the surface of the biotin-binding pocket, as supported by X-ray crystallography data of the avidin–biotin complex (PDB ID: 1AVD; Figure S3).[15] These results indicate that compared to conventional PAL methods, this approach leads to greater accuracy in the identification of the target protein by determination of multiply labeled peptides. The simplicity of the procedure and the clarity of the target signal can significantly reduce the amount of protein necessary for analysis (nanogram amounts), whereas conventional procedures require a large amount of the target protein owing to the low labeling yield, significant losses during purification, and further repeated control experiments.

Next, we performed PAL-based target proteomics using HeLa cell lysate with probe **1** for the identification of proteins that require a biotin for their enzymatic reactions, such as biotin-dependent carboxylases. The aim was to evaluate the performance of the target enrichment utilizing the cleavage property for functional proteomics, as the labeled proteins can be selectively and mildly recovered from the support by the second photoreaction.[8] After cross-linking the lysate (cytoplasm, 2 mg in total) with probe **1** (100 μM) in a PBS solution (500 μL) at 0 °C by the first photoreaction, labeled proteins were enriched and isolated by trapping on avidin-immobilized agarose gel as these were biotinylated. After thoroughly washing with SDS buffer solution (0.2 %) at 0 °C, the target proteins could be efficiently recovered from the gel accompanied by a coumarin moiety. Furthermore, this process can avoid co-elution of intrinsically biotinylated proteins. Reverse-phase HPLC of the digests of eluted proteins clearly showed some fluorescence peaks (Figure 3A) whose intensity decreased in the presence of biotin as a competitor (Figure S4). The ESI-MS profile of peak 1 eluted at 18.6 minutes is presented in Figure 3B and clearly shows a doublet peak at m/z 473.70 and 476.19 ($z = 2$). The following MS/MS analysis revealed the sequence to be DT*QAMK (with coumarin) and to be a part of pyruvate carboxylase (PC), an enzyme that catalyzes the biotin-dependent production of oxaloacetate. The labeled site, Thr1092, which is indicated by an asterisk, is located in the central area of the tetrameric structure and oriented to the catalytic pocket, which is supported by X-ray crystallography data of the human PC (PBD ID: 3BG3; Figure S5).[16] This is a reasonable position as the C-terminal biotin carboxyl carrier protein (BCCP) domain has been shown to move allosterically depending on substrate binding.[17] Furthermore, MS analysis of peak 2 (22.2 min) indicated that this peak might arise from acetyl-CoA carboxylase (ACC)-1 or 2, based on the labeled peptides assigned as IT*IGNK. Although human ACCs have not been crystallized in their entirety, this region is expected to be located adjacent to the BCCP domain in those proteins, which is similar to what has been observed

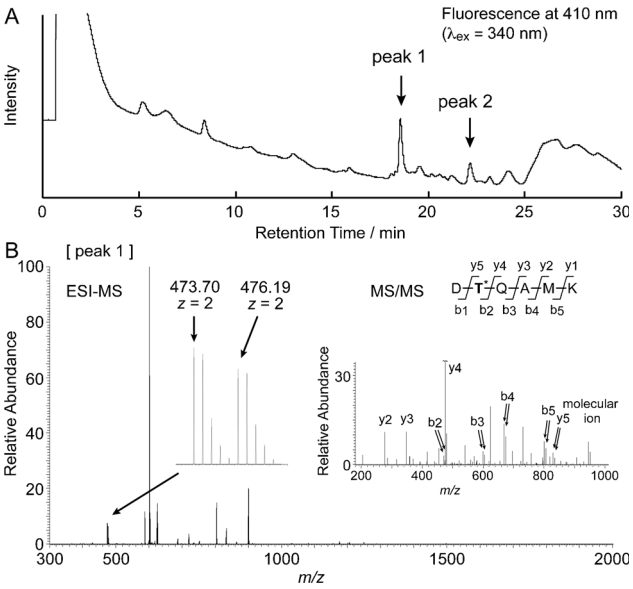

Figure 3. A) Reverse-phase HPLC profiles of digested products detected by fluorescence at 410 nm ($\lambda_{ex} = 340$ nm). The sample was prepared in a PBS solution (500 μL) with probe **1** (100 μM) and HeLa cell extracts (2 mg) and irradiated with 365 nm light at 0 °C. After treatment with avidin-immobilized gel, the labeled proteins were eluted by the second photoreaction and digested with trypsin/Lys-C. B) ESI-MS profile of peak 1 eluted at 18.6 minutes and MS/MS profile of the signals at m/z 473.7 and 476.2 ($z = 2$).

with PC. As PMF analysis of the eluted proteins still indicated the presence of many non-specifically adsorbed proteins (Table S1), sequencing of the labeled peptide should be important in providing more accurate information on the target protein and the ligand-binding site.

In conclusion, the PAL-based IsoFT strategy provides a method for the rapid identification of target proteins, which was supported by sequencing of the labeled peptide that constitutes the ligand-binding pocket. Identification can be achieved quickly through simple procedures, including brief irradiation and proteolytic digestion, with intelligible judgment of the target signal. Furthermore, the photocleavable moiety of the cross-linker enabled the efficient enrichment of the labeled protein and proteomic target screening. This simple procedure for target capture and identification has the potential to enhance the proteomic platform in the field of molecular biology.

Received: August 26, 2014
Revised: October 2, 2014
Published online: November 7, 2014

Keywords: isotopic labeling · mass spectrometry · photoaffinity labeling · proteins · recognition

[1] a) B. R. Stockwell, *Nature* **2004**, *432*, 846–854; b) B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383–414.
[2] a) B. Domon, R. Aebersold, *Science* **2006**, *312*, 212–217; b) A. Sinz, *Mass Spectrom. Rev.* **2006**, *25*, 663–682; c) K. J. Pacholarz,

R. A. Garlish, R. J. Taylor, P. E. Barran, *Chem. Soc. Rev.* **2012**, *41*, 4335 – 4355.
[3] a) Y. Hatanaka, M. Hashimoto, Y. Kanaoka, *Bioorg. Med. Chem.* **1994**, *2*, 1367 – 1373; b) T. Tomohiro, M. Hashimoto, Y. Hatanaka, *Chem. Rec.* **2005**, *5*, 385 – 395; c) Y. Luo, C. Blex, O. Baessler, M. Glinski, M. Dreger, M. Sefkow, H. Köster, *Mol. Cell. Proteomics* **2009**, *8*, 2843 – 2856.
[4] a) Z. Song, Q. Zhang, *Org. Lett.* **2009**, *11*, 4882 – 4885; b) A. B. Kumar, J. M. Anderson, R. Manetsch, *Org. Biomol. Chem.* **2011**, *9*, 6284 – 6292.
[5] a) D. Lapinsky, *J. Bioorg. Med. Chem.* **2012**, *20*, 6237 – 6247; b) C. X. Song, C. He, *Acc. Chem. Res.* **2011**, *44*, 709 – 717; c) J. S. Cisar, B. F. Cravatt, *J. Am. Chem. Soc.* **2012**, *134*, 10385 – 10388; d) C. H. Sohn, H. D. Agnew, J. E. Lee, M. J. Sweredoski, R. L. J. Graham, G. T. Smith, S. Hess, G. Czerwieniec, J. A. Loo, J. R. Heath, R. J. Deshaies, J. L. Beauchamp, *Anal. Chem.* **2012**, *84*, 2662 – 2669; e) A. L. MacKinnon, J. Taunton, *Curr. Protocol. Chem. Biol.* **2009**, *1*, 55 – 73; f) M. Hashimoto, Y. Hatanaka, *Chem. Pharm. Bull.* **2005**, *53*, 1510 – 1512.
[6] G. Li, Y. Liu, Y. Liu, L. Chen, S. Wu, Y. Liu, X. Li, *Angew. Chem. Int. Ed.* **2013**, *52*, 9544 – 9549; *Angew. Chem.* **2013**, *125*, 9723 – 9728.
[7] a) T. Tomohiro, K. Kato, S. Masuda, H. Kishi, Y. Hatanaka, *Bioconjugate Chem.* **2011**, *22*, 315 – 318; b) T. Tomohiro, H. Inoguchi, S. Masuda, Y. Hatanaka, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 5605 – 5608.
[8] S. Morimoto, T. Tomohiro, N. Maruyama, Y. Hatanaka, *Chem. Commun.* **2013**, *49*, 1811 – 1813.
[9] D. R. Müller, P. Schindler, H. Towbin, U. Wirth, H. Voshol, S. Hoving, M. O. Steinmetz, *Anal. Chem.* **2001**, *73*, 1927 – 1934.
[10] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994 – 999.
[11] S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **2002**, *1*, 376 – 386.
[12] P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D. J. Pappin, *Mol. Cell. Proteomics* **2004**, *3*, 1154 – 1169.
[13] D. Zeng, S. Li, *Chem. Commun.* **2009**, 3369 – 3371.
[14] Z. Song, W. Huang, Q. Zhang, *Chem. Commun.* **2012**, *48*, 3339 – 3341.
[15] O. Livnah, E. A. Bayer, M. Wilchek, J. L. Sussman, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5076 – 5080.
[16] S. Xiang, L. Tong, *Nat. Struct. Mol. Biol.* **2008**, *15*, 295 – 302.
[17] A. D. Lietzan, M. St. Maurice, *J. Biol. Chem.* **2013**, *288*, 19915 – 19925.

