Activity-Based Proteomics

DOI: 10.1002/anie.200603187

Substrate Mimicry in an Activity-Based Probe That Targets the Nitrilase Family of Enzymes**

*Katherine T. Barglow and Benjamin F. Cravatt**

A principal goal of proteomics is to assign functions to each of the proteins encoded by prokaryotic and eukaryotic genomes.[1,2] A protein's function can be defined on multiple levels that reflect its coincident impact on biochemical, cellular, and physiological processes. Activity-based protein profiling (ABPP) is a functional proteomics method that utilizes active-site-directed chemical probes to characterize enzymes in native biological samples.[3–5] ABPP probes have proven useful for relating the activity state of enzymes to higher-order physiological and pathological processes, including cancer pathogenesis,[6–8] parasite invasiveness,[9] and obesity.[10] The extent to which these chemical tools might also offer insights into the structures of endogenous substrates of enzymes remains a pertinent, but largely unanswered question. Herein we show that profiling proteomes with a library of protein-reactive chemical probes can generate structure–activity data that are reflective of the natural substrates of enzymes. Thus, ABPP can facilitate the annotation of protein function at multiple levels.

Previous ABPP studies revealed markedly distinct proteome reactivity profiles for members of a library of rhodamine-tagged dipeptide α-chloroacetamide (α-CA) probes.[10] Some protein targets showed exclusive or preferential reactivity with a single α-CA, whereas others showed a more promiscuous profile, reacting with several members of the probe library. Among the more intriguing proteins that fell into the former class was a 44-kDa liver protein that exclusively reacted with Leu-Asp α-CA, where Leu and Asp represent R^2 and R^1, respectively (Figure 1a). We sought to further characterize the probe-labeling profile of this protein by testing its reactivity with an expanded library of α-CA probes in which either the R^1 or R^2 position was held constant as leucine and the complementary site varied with one of twenty amino acid side chains (Figure 1a). The 44-kDa protein was again found to react specifically with Leu-Asp α-CA, and showed negligible labeling with other probes including the structurally related Leu-Glu, Leu-Gln, and Asp-Leu α-CAs (Figure 1b and the Supporting Information).

The 44-kDa protein was enriched from mouse liver proteomes by labeling with a trifunctional Leu-Asp α-CA

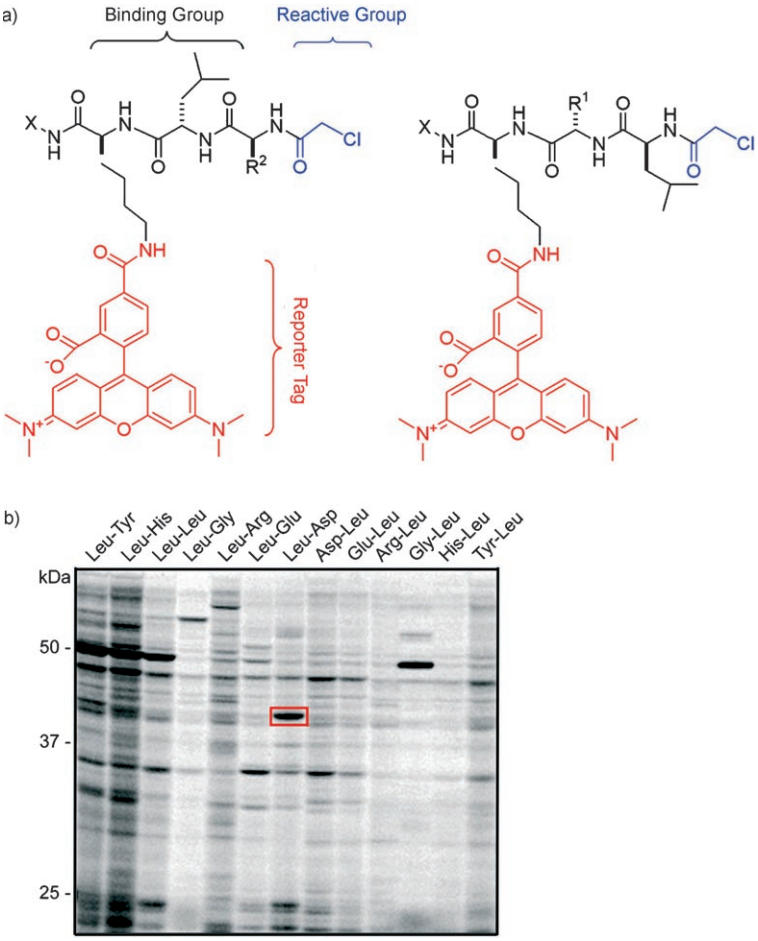

Figure 1. A library of dipeptide α-CA probes for ABPP. a) Structure of the dipeptide α-CA, where R^1 and R^2 are 15 different natural amino acids, phospho-serine, phospho-tyrosine, D-leucine, D-proline, or D-aspartate (see the Supporting Information for a complete summary of the dipeptide α-CA probes). For proteome screening, X = H; for protein identification, X = lysine-biotin. The α-CA reactive group and rhodamine tag are colored in blue and red, respectively. b) Proteome reactivity profiles for representative members of the α-CA probe library. Probes (20 μM) were allowed to react with mouse-liver proteome (2 mg protein mL^{-1}) for 1 h at room temperature and the reactions analyzed by SDS-PAGE and in-gel fluorescence scanning. A 44-kDa protein that reacted selectively with the Leu-Asp α-CA probe is highlighted by a red box. See the Supporting Information for comparison of the reactivity of this protein with Leu-Asp and Leu-Asn α-CA. Fluorescent gel images are shown in grayscale.

[*] K. T. Barglow, Prof. B. F. Cravatt
The Skaggs Institute for Chemical Biology and
Departments of Cell Biology and Chemistry
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-8023
E-mail: cravatt@scripps.edu

[**] This research was supported by the NIH (CA087660), the Howard Hughes Medical Institute (predoctoral fellowship to K.T.B.), and the Skaggs Institute for Chemical Biology. We thank M. Patricelli (Activx Biosciences) for assistance with the identification of probe sites of labeling for nitrilase enzymes and the members of the Cravatt lab for helpful discussions and critical reading of the manuscript.

Supporting information for this article, which also contains details of the materials and methods, is available on the WWW under http://www.angewandte.org or from the author.

probe coupled to both biotin and rhodamine, followed by avidin affinity chromatography and SDS-PAGE. Liquid chromatography (LC)–MS analysis of a trypsin digest of the gel-excised 44-kDa protein identified it as ureidopropionase-β (Upβ), a 43.9-kDa member of the nitrilase superfamily (see the Supporting Information). To confirm this identification, a cDNA (complementary DNA) encoding the full-length mouse Upβ (appended with a C-terminal myc epitope tag) was subcloned into the pcDNA3.1 expression vector and transiently transfected into COS-7 cells. A strong Leu-Asp α-CA-reactive 44-kDa protein was detected in Upβ-transfected cells, but not mock-transfected cells (Figure 2a, left). Like native liver Upβ, the recombinantly expressed form of this enzyme showed nearly exclusive labeling with Leu-Asp α-CA (Figure 2b).

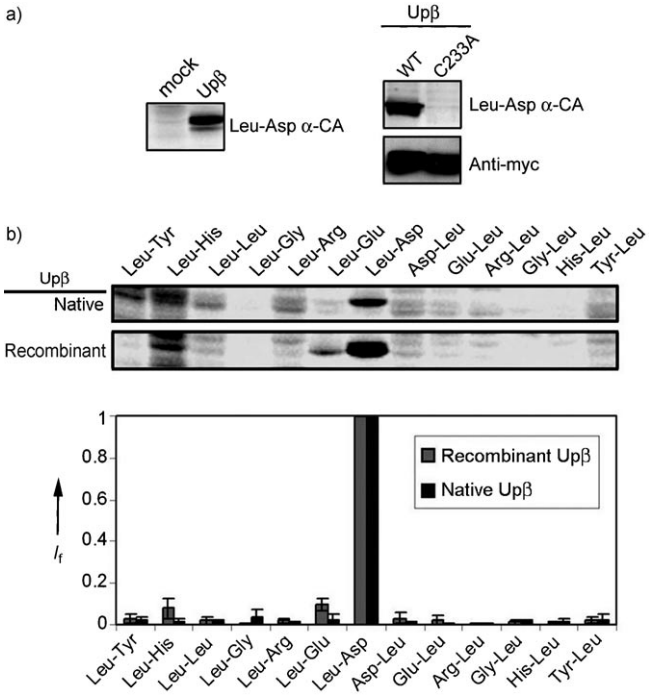

Figure 2. Identification and characterization of Upβ as a target of Leu-Asp α-CA. a) Left: Upβ-transfected, but not mock-transfected, COS-7 cells possess a 44-kDa protein that reacted with Leu-Asp α-CA; right: mutation of the cysteine nucleophile of Upβ to alanine (C233A) eliminated labeling by Leu-Asp α-CA. b) Native (liver) and recombinant (transfected COS-7) Upβ enzymes selectively reacted with Leu-Asp α-CA compared with other dipeptide α-CA probes. Note that low signals were also observed in the molecular-mass range of Upβ for the Leu-Tyr and Leu-His probes; however, these probes showed much higher overall background proteome reactivity (see Figure 1b), suggesting that these signals may be nonspecific. Top: fluorescent gel images shown in grayscale. Bottom: quantification of fluorescent signals. Data represent the average ± standard deviation for four independent experiments. I_f = relative fluorescence, WT = wild type.

The nitrilase family is a phylogenetically ancient class of enzymes with multiple mammalian members that utilize a conserved Cys-Lys-Glu catalytic triad to cleave C–N bonds, including nitriles, carbamates, ureas, and amides.[11,12] Upβ, in particular, catalyzes the hydrolysis of *N*-carbamoyl β-alanine (CarβAla) to β-alanine, which is a central step in the pyrimidine catabolic pathway.[13] Mutation of Upβ in humans leads to a buildup of CarβAla, which exerts neurotoxic effects resulting in severe neuropathology and movement disorders.[13] In considering possible explanations for the high selectivity displayed by Upβ for Leu-Asp α-CA, we noted a striking structural similarity between this probe and Upβ's endogenous substrate CarβAla (Figure 3). This structural relation suggested that Leu-Asp α-CA bound and labeled Upβ at the active site. This premise was confirmed by mapping the site of probe modification, which was identified by tandem MS analysis as cysteine 233, the catalytic nucleophile of Upβ (Table 1).[11] Mutation of this residue to alanine produced a Upβ variant (C233A) that was expressed at wild-type levels in transfected COS-7 cells but failed to react with Leu-Asp α-CA (Figure 2a, right).

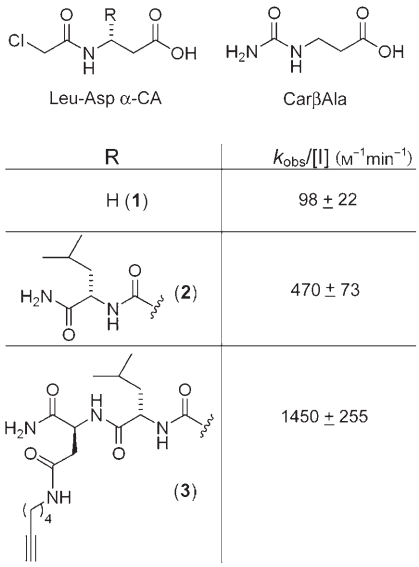

Figure 3. Leu-Asp α-CA is structurally similar to the natural Upβ substrate CarβAla. Top: structural comparison of Leu-Asp α-CA and CarβAla. Bottom: a comparison of the rates of inactivation of Upβ by analogues of Leu-Asp α-CA. The structure of the rhodamine-tagged Leu-Asp probe can be inferred from Figure 1.

Table 1: Sites of α-CA probe labeling for members of nitrilase family.

Enzyme	Labeled peptide[a]	Site of labeling[a]
Upβ	IAVNIC*YGR	C233
Nit1	VGLAIC*YDMRFPELSLK	C199

[a] Probe-labeled peptides and sites of labeling were determined by using an LC–MS platform for ABPP as described previously.[20,21] Proteomes were treated with an equimolar mixture of Leu-Asp, Leu-D-Asp, and D-Leu-D-Asp α-CA probes (20 μM of each probe).

Further evidence that Leu-Asp α-CA utilized substrate mimicry to achieve active-site modification of Upβ was sought by screening analogues of this probe for their ability to block enzyme labeling. Such competitive ABPP studies, when performed in a concentration- and time-dependent manner,

can provide an estimate of the relative potency with which inhibitors target the active site of an enzyme.[14] An α-CA was designed that most closely resembled the Upβ substrate CarβAla, and this agent (**1**) was found to block probe labeling of Upβ with a $k_{obs}/[I]$ value of $98 \pm 22\,M^{-1}\,min^{-1}$ (I = competitive inhibitor, Figure 3). The addition of a leucine group (agent **2**) accelerated the rate of inactivation of Upβ, providing an agent that displayed a $k_{obs}/[I]$ value of $470 \pm 73\,M^{-1}\,min^{-1}$ (Figure 3). The introduction of a third amino acid derivative possessing an extended alkynyl side chain (agent **3**) further increased the rate of Upβ inactivation ($k_{obs}/[I] = 1450 \pm 255\,M^{-1}\,min^{-1}$, Figure 3) and enabled confirmation of covalent labeling of the enzyme by a click-chemistry reaction with an azide-rhodamine reporter group[15,16] (see the Supporting Information). These results indicate that an α-CA probe containing the core features of the CarβAla substrate is capable of covalently labeling the active site of Upβ, but that additional features of Leu-Asp α-CA further enhance this reaction.

Given that labeling of Upβ by Leu-Asp α-CA occurred on the enzyme's cysteine nucleophile, an essential catalytic residue conserved among all nitrilases,[11,12] we next tested whether α-CA probes might label other members of this enzyme family. Analysis of mouse-liver proteomes by using the MS-based platform ABPP-MudPIT (Multidimensional Protein Identification Technology)[17] identified nitrilase 2 (Nit2), along with Upβ, as specific targets of Leu-Asp α-CA (see the Supporting Information). Mouse nitrilase 1 (Nit1) was recombinantly expressed as a myc-tagged fusion protein in COS7 cells and treated with an expanded panel of α-CA probes that included agents where the chirality and order of the Leu-Asp dipeptide scaffold were systematically varied. Strong labeling was observed for Nit1 with several of the dipeptide α-CA probes (Figure 4 a and b), which contrasted with the more selective labeling pattern of Upβ (Figures 2b and 4b). Tandem MS analysis revealed that probe labeling of Nit1 occurred on the enzyme's cysteine nucleophile (C199; Table 1). These data indicate that dipeptide α-CAs serve as effective ABPP probes for multiple members of the nitrilase enzyme class.

In summary, we report the first example of active-site-directed proteomic probes that target the nitrilase family of enzymes. Nitrilases perform important functions in a wide range of organisms, including humans. For example, Upβ is a key component of the pyrimidine catabolic pathway, catalyzing the conversion of CarβAla to β-alanine.[13] This reaction appears to constitute the dedicated endogenous function of Upβ, as is reflected in the enzyme's restricted substrate selectivity, probe reactivity, and tissue distribution profiles (specific expression in liver; http://symatlas.gnf.org/SymAtlas/), as well as the metabolic defects that arise from its genetic mutation.[13] On the other hand, Nit1 is more ubiquitously expressed in mammalian tissues,[18] and the disruption of this enzyme leads to multiple phenotypes in mice, including accelerated cell growth and increased incidence of chemically induced tumors.[18] These physiological findings, coupled with the more promiscuous probe reactivity profile of Nit1, suggest that it may display a broader metabolic function than Upβ. Confirmation of this hypothesis will require identification of endogenous substrates for Nit1.

The additional finding that optimal probes for Upβ mimicked the structure of the enzyme's natural substrate has broader implications for enzyme annotation by using ABPP methods. Although other ABPP probes have been designed based on the substrate preferences of their target enzyme classes (e.g., negatively charged acyloxymethyl ketones for caspases[19]), our data with Upβ suggest provocatively that the relationships between enzymes and substrates can be "discovered" de novo by screening libraries of structurally diverse probes. By extension, we speculate that the probe-reactivity profiles of uncharacterized enzymes may also contain examples of substrate mimicry and, thus, provide insights into the endogenous biochemical functions of these proteins.

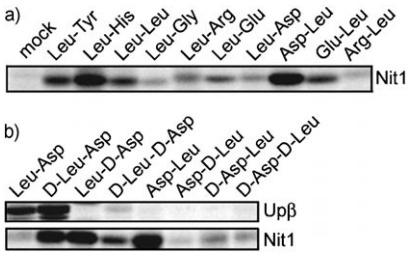

Figure 4. Identification and characterization of Nit1 as a target for dipeptide α-CA probes. a) Nit1-transfected, but not mock-transfected, COS7 cells possess a 35-kDa protein that reacted with several members of the α-CA probe library (mock-transfected proteome shown was allowed to react with Leu-Tyr α-CA). b) Comparison of the reactivity profiles of Upβ and Nit with a chiral library of Leu-Asp and Asp-Leu α-CA probes.

Received: August 5, 2006
Published online: October 11, 2006

Keywords: amino acids · biological activity · nitrilases · proteomics · screening

[1] P. C. Babbitt, *Curr. Opin. Chem. Biol.* **2003**, *7*, 230.
[2] A. Saghatelian, B. F. Cravatt, *Nat. Chem. Biol.* **2005**, *1*, 130.
[3] Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14694.
[4] A. E. Speers, B. F. Cravatt, *ChemBioChem* **2004**, *5*, 41.
[5] A. B. Berger, P. M. Vitorino, M. Bogyo, *Am. J. Pharmacogenomics* **2004**, *4*, 371.
[6] N. Jessani, Y. Liu, M. Humphrey, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10335.
[7] N. Jessani, M. Humphrey, W. H. McDonald, S. Niessen, K. Masuda, B. Gangadharan, J. R. Yates, 3rd, B. M. Mueller, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13756.
[8] J. A. Joyce, A. Baruch, K. Chehade, N. Meyer-Morse, E. Giraudo, F. Y. Tsai, D. C. Greenbaum, J. H. Hager, M. Bogyo, D. Hanahan, *Cancer Cell* **2004**, *5*, 443.
[9] D. C. Greenbaum, A. Baruch, M. Grainger, Z. Bozdech, K. F. Medzihradszky, J. Engel, J. DeRisi, A. A. Holder, M. Bogyo, *Science* **2002**, *298*, 2002.
[10] K. T. Barglow, B. F. Cravatt, *Chem. Biol.* **2004**, *11*, 1523.
[11] P. Bork, E. V. Koonin, *Protein Sci.* **1994**, *3*, 1344.
[12] C. Brenner, *Curr. Opin. Struct. Biol.* **2002**, *12*, 775.

[13] A. B. van Kuilenburg, R. Meinsma, E. Beke, B. Assmann, A. Ribes, I. Lorente, R. Busch, E. Mayatepek, N. G. Abeling, A. van Cruchten, A. E. Stroomer, H. van Lenthe, L. Zoetekouw, W. Kulik, G. F. Hoffmann, T. Voit, R. A. Wevers, F. Rutsch, A. H. van Gennip, *Hum. Mol. Genet.* **2004**, *13*, 2793.
[14] D. Leung, C. Hardouin, D. L. Boger, B. F. Cravatt, *Nat. Biotechnol.* **2003**, *21*, 687.
[15] A. E. Speers, G. C. Adam, B. F. Cravatt, *J. Am. Chem. Soc.* **2003**, *125*, 4686.
[16] A. E. Speers, B. F. Cravatt, *Chem. Biol.* **2004**, *11*, 535.
[17] N. Jessani, S. Niessen, B. Q. Wei, M. Nicolau, M. Humphrey, Y. Ji, W. Han, D. Y. Noh, J. R. Yates III, S. S. Jeffrey, B. F. Cravatt, *Nat. Methods* **2005**, *2*, 691.
[18] S. Semba, S. Y. Han, H. R. Qin, K. A. McCorkell, D. Iliopoulos, Y. Pekarsky, T. Druck, F. Trapasso, C. M. Croce, K. Huebner, *J. Biol. Chem.* **2006**, *281*, 28244.
[19] D. Kato, K. M. Boatright, A. B. Berger, T. Nazif, G. Blum, C. Ryan, K. Chehade, G. S. Salvensen, M. Bogyo, *Nat. Chem. Biol.* **2005**, *1*, 33.
[20] G. C. Adam, J. J. Burbaum, J. W. Kozarich, M. P. Patricelli, B. F. Cravatt, *J. Am. Chem. Soc.* **2004**, *126*, 1363.
[21] E. S. Okerberg, J. Wu, B. Zhang, B. Samii, K. Blackford, D. T. Winn, K. R. Shreder, J. J. Burbaum, M. P. Patricelli, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4996.

