Enzyme Classification

Class Assignment of Sequence-Unrelated Members of Enzyme Superfamilies by Activity-Based Protein Profiling**

*Nadim Jessani, Jason A. Young, Sandra L. Diaz, Matthew P. Patricelli, Ajit Varki, and Benjamin F. Cravatt**

Enzyme superfamilies are composed of members that share a common catalytic mechanism, but not necessarily sequence or structural homology.[1] The characterization of sequence-unrelated members of enzyme classes is an experimentally challenging endeavor which often requires extensive biochemical studies on purified preparations of enzymes.[2,3] Given these issues, in conjunction with the abundance of unannotated proteins provided by recent genome-sequencing projects, it is likely that cryptic members of many enzyme classes still exist in eukaryotic and prokaryotic proteomes. Herein, we demonstrate that a chemical proteomics method referred to as activity-based protein profiling (ABPP)[4,5] can be used to identify sequence-unrelated members of enzyme superfamilies on the basis of their reactivity with "mechanism-based" probes.

In previous studies, we analyzed a panel of human cancer cell lines with a set of fluorophosphonate (FP) based ABPP probes, [6] which target the serine hydrolase (SH) superfamily of enzymes.[7] FPs are well-characterized affinity labels for SHs; they phosphonylate irreversibly the active-site serine nucleophile in these enzymes.[8] Numerous FP-labeled enzymes were identified in these experiments, including proteases, lipases, and esterases. Most of these enzymes were readily assigned to the SH superfamily on the basis of database (BLAST) searches, which identified conserved sequence elements shared by members of this enzyme class. Notably, however, one FP target, sialic acid 9-*O*-acetylesterase (SAE), which was selectively expressed in melanoma cell lines (Figure 1 a), eluded such classification, as this enzyme shared no sequence homology with SHs or, for that matter, any other enzyme class. SAE was originally characterized by

[*] N. Jessani, J. A. Young, Prof. B. F. Cravatt
The Skaggs Institute for Chemical Biology and
Departments of Cell Biology and Chemistry
The Scripps Research Institute
La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-8023
E-mail: cravatt@scripps.edu
S. L. Diaz, Prof. A. Varki
Departments of Medicine and Cellular & Molecular Medicine
Glycobiology Research and Training Center
University of California, San Diego
La Jolla, CA 92093-0687 (USA)
M. P. Patricelli
Activx Biosciences
11025 N. Torrey Pines Rd, Suite 120, La Jolla, CA 92038 (USA)

[**] This research was supported by the NIH (CA087660 and GM32373) and the Skaggs Institute for Chemical Biology.

  DOI: 10.1002/ange.200463098

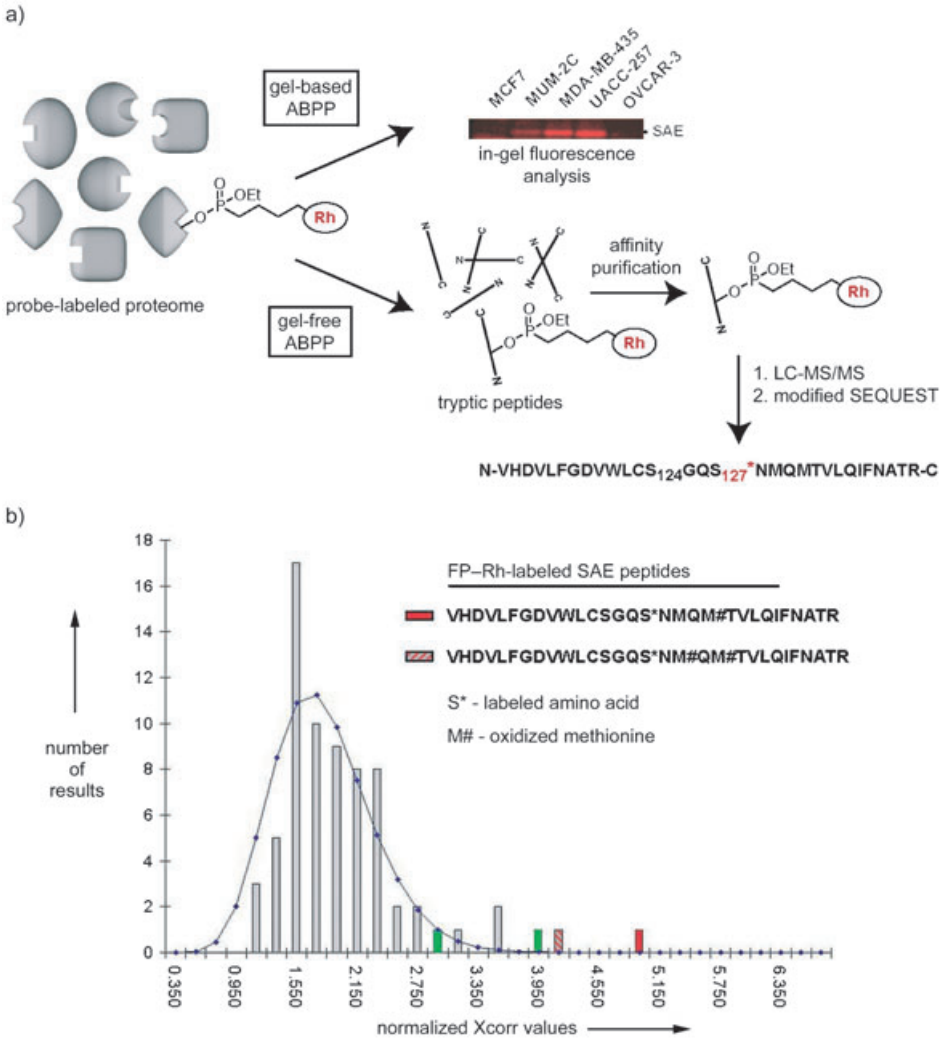

Figure 1. Identification of the site of FP–Rh labeling in SAE by gel-free ABPP. a) General strategy for gel-free ABPP as compared to gel-based ABPP. With gel-based ABPP, probe-labeled proteomes are analyzed by in-gel fluorescence scanning (shown is a representative analysis of FP–Rh-labeled SAE signals in the secreted proteomes of breast (MCF7), ovarian (OVCAR3), and melanoma (MUM-2C, MDA-MB-435, UACC-257) cancer cell lines). With gel-free ABPP, probe-labeled proteomes are denatured and, after reduction and alkylation of their thiol groups, digested with trypsin. Probe-modified peptides are then affinity isolated from the resulting mixture of tryptic peptides by bead-conjugated anti-rhodamine monoclonal antibodies (Activx Biosciences)[12,13] and analyzed by LC-MS/MS and the SEQUEST search algorithm to identify concurrently the protein targets of ABPP probes and the specific residues labeled by these probes. General experimental conditions for gel-free ABPP of SAE: Conditioned-media proteome (2 mL, 1 mg protein mL^{-1}) from COS-7 cells that recombinantly express human SAE was treated with FP–Rh (4 μM) for 1 h. The sample was then processed and analyzed as described above to produce MS/MS spectra of FP–Rh-labeled peptides. The spectra were searched against a human-protein database by using SEQUEST, and the resulting data were filtered and analyzed as described.[12,13] The normalized Xcorr values from the SEQUEST results (shown as bars in the graph in (b)) were plotted against the predicted distribution of false positive results in the same data set (stippled blue line). Red filled and red hatched bars indicate results that represent an SAE peptide, residues 112–141, with S127 as the predicted probe-modified residue. This peptide was identified in both singly methionine-oxidized and doubly methionine-oxidized forms (red filled and red hatched bars, respectively). The combined probability of the random occurrence of the SAE peptides with the indicated scores was calculated at less than 0.0001 %. Green bars indicate results that represent other SHs in the proteome with probe modifications on their active-site serine nucleophile.

Varki and colleagues as an enzyme that removes *O*-acetyl esters from the 9-position of naturally occurring sialic acids[9,10] and has been shown to be inhibited by diisopropyl fluorophosphate and the arginine-modifying reagents 2,3-butandione and phenylglyoxal.[11] Nonetheless, to date the specific residues involved in SAE catalysis have remained obscure, and, as a consequence, so too has the mechanistic classification of this enzyme. To determine whether SAE might represent a novel member of the SH superfamily, we set out to identify and characterize the site of FP labeling in this enzyme.

To further characterize the SAE–FP interaction, we employed a gel-free version of ABPP that enables sites of probe labeling to be determined directly for enzymes in complex proteomes (Figure 1a).[12,13] In this method, probe-treated proteomes are digested with trypsin, and the probe-labeled peptides are captured by affinity chromatography and identified by liquid chromatography–tandem mass spectrom-

etry (LC-MS/MS). COS-7 cells were transfected transiently with a cDNA that encodes human SAE, and the secreted proteome from these cells was treated with a rhodamine-tagged FP probe (FP–Rh[14]). Following trypsin digestion, FP–Rh-labeled peptides were captured by using anti-Rh antibody beads[12,13] and analyzed by LC-MS/MS. SEQUEST searches of the resulting LC-MS/MS data set identified a single FP-labeled SAE peptide (amino acids 112–141) with a predicted site of probe modification at serine 127 (S127; Figure 1b). Confidence in this assignment was bolstered by the detection of multiple forms of this probe-labeled peptide that differed in the degree of methionine oxidation (Figure 1b). In each case, the site of probe modification was designated as S127.

The mutagenesis of S127 to alanine produced an SAE variant (indicated as S127A) that expressed at wild-type levels in COS-7 cells, but failed to undergo labeling with FP–Rh (Figure 2a). In contrast, the mutagenesis of the neighboring serine residue S124 to alanine generated an enzyme S124A that reacted strongly with FP–Rh (Figure 2a). We next compared the catalytic activity of wild-type SAE (WT SAE) with that of the S124A and S127A mutants by using a radiolabeled-substrate assay. The WT and S124A SAE enzymes hydrolyzed the substrate 9-*O*-[*acetyl*-^{3}H]acetyl-*N*-acetylneuraminic acid ([^{3}H]Neu5,9Ac$_2$)[9] to similar extents (Figure 2b). In contrast, the S127A SAE mutant showed negligible catalytic activity (Figure 2b). Collectively, these data indicate that S127 fulfills two of the major requirements expected of a SH catalytic nucleophile: the residue is 1) specifically labeled by FP inhibitors and 2) essential for catalysis.

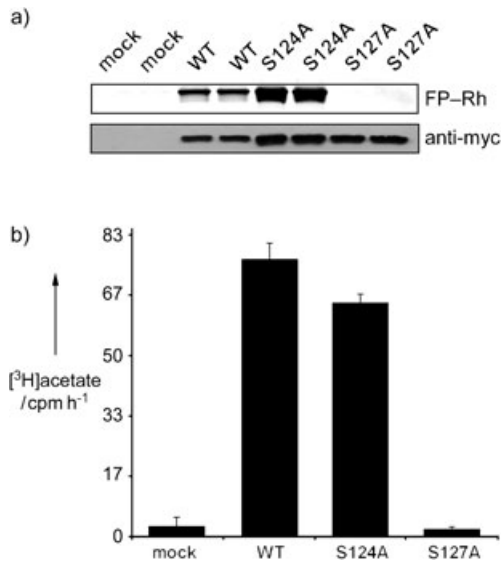

Figure 2. Comparative characterization of WT SAE and the S124A and S127A SAE mutants. SAE proteins were recombinantly expressed in COS-7 cells by transient transfection as myc-epitope-tagged fusion proteins (myc epitope: EQKLISEEDL) and affinity purified by using anti-myc agarose beads (sigma; 50-μL beads with 0.75 mL of proteome in 50 mM tris(hydroxymethyl)aminomethane (Tris), pH 7.5 (1 mg protein mL^{-1})). Following thorough washing of the beads with Tris (50 mM, pH 7.5, 2x0.5 mL), an aliquot of each SAE variant was treated with FP–Rh (4 μM) for 1 h. The reactions were then quenched with standard SDS-PAGE (sodium dodecylsulfate–polyacrylamide gel electrophoresis) loading buffer and analyzed by in-gel fluorescence scanning (a, upper panel) and western blotting with anti-myc antibodies (a, lower panel). The FP–Rh probe labeled both WT and S124A SAE, but not S127A SAE. Anti-myc blotting confirmed that each SAE protein was expressed at high levels in COS-7 cells. b) Comparison of the catalytic activity of WT SAE and the S124A and S127A SAE mutants. SAE variants were assayed on beads with a radiolabeled-substrate assay that follows the cleavage of [^{3}H]acetate from [^{3}H]Neu5,9Ac$_2$ (10000 cpm/reaction; cpm = counts per minute). This assay[11] and the biosynthetic preparation of [^{3}H]Neu5,9Ac$_2$[26] have been described previously. Results represent the averages of two independent trials and have been normalized to account for the relative quantity of enzyme present in each reaction (as estimated by western blotting). Mock cells were transfected with empty vector (pcDNA3mycHis) and served as a background control for all experiments.

If S127 serves as the catalytic nucleophile in SAE, this residue should be conserved among sequence-related homologues of this enzyme. BLAST searches identified over 30 enzymes in public databases that show sequence homology with SAE. Alignment of these sequences revealed that S127 is a conserved residue, even in the most distantly related homologues (which share less than 25% sequence identity; Figure 3). No other conserved serine residues were identified

SAE family members	S127
human	TLRVHD**V**LFG**D**V**W**LC**S**G**Q**S**N**MQMTVLQI
mouse	TLRVHD**V**LFG**D**V**W**LC**S**G**Q**S**N**MQMTVSQI
bact. theta	LLTLQN**V**LIG**E**V**W**FC**S**G**Q**S**N**MEMPMGGF
rhod. baltica	TKTFTD**V**LVG**E**V**W**IC**S**G**Q**S**N**MAWAVQSA
bact. frag.	----NN**V**LAG**E**V**W**LC**S**G**Q**S**N**MEFYLSWS
xanth. axon.	ELQVRD**V**LVG**D**V**W**LAG**G**Q**S**N**N**MEWPLAQA
pedio. pent.	--LLKK**V**RFG**R**V**I**LMAG**G**Q**S**N**N**VGFRMVQD

Figure 3. Sequence comparisons of SAE and representative bacterial homologues. Shown is the region surrounding S127 (red), the predicted serine nucleophile of SAE. S127 is completely conserved among all SAE sequence homologues present in public databases. Additional conserved residues are shown in bold. S124, which is mostly, but not completely, conserved, is shown in blue. SAE enzymes (Genbank accession number): human (NP_733746), mouse (CAA67214), *Bact. theta.* (*Bacteroides thetaiotaomicron*, AA078230), *Rhod. baltica* (*Rhodopirellula baltica*, CAD76603), *Bact. frag.* (*Bacteroides fragilis*, BAC48479), *Xanth. axon.* (*Xanthomonas axonopodis*, AAM36635), *Pedio. pent.* (*Pediococcus pentosaceus*, ZP_00323591).

in this group of proteins. These findings suggest that all SAE-related enzymes, which can be found in both higher eukaryotes and bacteria, are members of the SH superfamily.

SHs also possess a catalytic base to activate the serine nucleophile. Most commonly, this catalytic base is a histidine residue which bridges the serine nucleophile and an aspartate residue to form the classic Ser-His-Asp catalytic triad.[8] However, other arrangements of catalytic residues are also possible, including Ser-Asp[15] and Ser-Lys dyads,[16] as well as a Ser-Ser-Lys triad.[17,18] Sequence comparisons suggest that the SAE subfamily does not employ a Ser-His-Asp catalytic triad, as this group of enzymes lacks a conserved aspartate residue. In contrast, single conserved histidine and lysine residues (H377 and K381), as well as two conserved arginine residues (R301 and R387), are present in the SAE family. The sensitivity of SAE to 2,3-butandione and phenylglyoxal[11] suggests that one or both of the arginine residues may be important for substrate binding and/or catalysis. Future

studies should clarify the roles that these arginine residues, as well as H377 and K381, play in the SAE catalytic mechanism.

In summary, the results reported herein indicate that SAE and its sequence homologues constitute a novel branch of the SH superfamily. This discovery may enable the design of specific inhibitors of SAE enzymes that incorporate functional groups, such as α-keto heterocycles and trifluoromethyl ketones,[19] that show strong affinity for SH active sites. Such inhibitors would be of value for testing the functional significance of the 9-O acetylation of sialic acids, a modification that is expressed selectively on gangliosides in human melanoma cells.[20] More generally, it is interesting to consider whether the assignment of unannotated enzymes to other mechanistic classes may also be facilitated by ABPP probes. On the one hand, the extrapolation of mechanistic information with active-site-directed probes that target multiple classes of enzymes may prove challenging.[21,22] On the other hand, several ABPP probes do show preferential reactivity with a single class of enzymes (e.g., hydroxamate probes for metalloproteases,[23,24] E-64-based probes for cysteine proteases[25]). In these cases, like with the FP probes, ABPP may provide a useful method to uncover cryptic members of enzyme superfamilies that have resisted classification based on sequence comparisons.

Received: December 29, 2004
Published online: March 14, 2005

Keywords: activity-based protein profiling · enzymes · proteomics · serine hydrolases

[1] J. A. Gerlt, P. C. Babbitt, *Annu. Rev. Biochem.* **2001**, *70*, 209–246.
[2] M. P. Patricelli, M. A. Lovato, B. F. Cravatt, *Biochemistry* **1999**, *38*, 9804–9812.
[3] W. R. Tschantz, M. Sung, V. M. Delgado-Partin, R. E. Dalbey, *J. Biol. Chem.* **1993**, *268*, 27349–27354.
[4] N. Jessani, B. F. Cravatt, *Curr. Opin. Chem. Biol.* **2004**, *8*, 54–59.
[5] A. E. Speers, B. F. Cravatt, *ChemBioChem* **2004**, *5*, 41–47.
[6] N. Jessani, Y. Liu, M. Humphrey, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10335–10340.
[7] Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14694–14699.
[8] C. T. Walsh, *Enzymatic Reaction Mechanisms*, W. H. Freeman and Company, New York, **1979**.
[9] H. H. Higa, A. Manzi, S. Diaz, A. Varki, *Methods Enzymol.* **1989**, *179*, 409–415.
[10] M. J. Guimaraes, J. F. Bazan, J. Castagnola, S. Diaz, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, A. Varki, A. Zlotnik, *J. Biol. Chem.* **1996**, *271*, 13697–13705.
[11] B. K. Hayes, A. Varki, *J. Biol. Chem.* **1989**, *264*, 19443–19448.
[12] G. C. Adam, J. J. Burbaum, J. W. Kozarich, M. P. Patricelli, B. F. Cravatt, *J. Am. Chem. Soc.* **2004**, *126*, 1363–1368.
[13] 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**, in press.
[14] M. P. Patricelli, D. K. Giang, L. M. Stamp, J. J. Burbaum, *Proteomics* **2001**, *1*, 1067–1071.
[15] A. Dessen, J. Tang, H. Schmidt, M. Stahl, J. D. Clark, J. Seehra, W. S. Somers, *Cell* **1999**, *97*, 349–360.
[16] M. Paetzel, R. E. Dalbey, N. C. Strynadka, *Nature* **1998**, *396*, 186–190.
[17] M. H. Bracey, M. A. Hanson, K. R. Masuda, R. C. Stevens, B. F. Cravatt, *Science* **2002**, *298*, 1793–1796.
[18] S. Shin, T. H. Lee, N. C. Ha, H. M. Koo, S. Y. Kim, H. S. Lee, Y. S. Kim, B. H. Oh, *EMBO J.* **2002**, *21*, 2509–2516.
[19] D. Leung, C. Hardouin, D. L. Boger, B. F. Cravatt, *Nat. Biotechnol.* **2003**, *21*, 687–691.
[20] D. A. Cheresh, R. A. Reisfeld, A. Varki, *Science* **1984**, *225*, 844–846.
[21] G. C. Adam, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol.* **2002**, *20*, 805–809.
[22] K. T. Barglow, B. F. Cravatt, *Chem. Biol.* **2004**, *11*, 1523–1531.
[23] A. Saghatelian, N. Jessani, A. Joseph, M. Humphrey, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10000–10005.
[24] E. W. Chan, S. Chattopadhaya, R. C. Panicker, X. Huang, S. Q. Yao, *J. Am. Chem. Soc.* **2004**, *126*, 14435–14446.
[25] D. Greenbaum, K. F. Medzihradszky, A. Burlingame, M. Bogyo, *Chem. Biol.* **2000**, *7*, 569–581.
[26] H. H. Higa, A. Manzi, A. Varki, *J. Biol. Chem.* **1989**, *264*, 19435–19442.

