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

Nadim Jessani[†], Jason Young[†], Sandra L. Diaz[#], Matthew P. Patricelli[††], Ajit Varki[#], Benjamin F. Cravatt[†]*

[†]The Skaggs Institute for Chemical Biology and Departments of Cell Biology and Chemistry, The Scripps Research Institute, La Jolla, CA 92037; [††], Activx Biosciences, 1025 N. Torrey Pines Rd Suite 120, La Jolla, CA 92038; [#], Division of Biological Sciences, Glycobiology Research and Training Center, University of California at San Diego, La Jolla, CA 92093-0687, USA.

*To whom correspondence should be addressed: cravatt@scripps.edu

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

Enzyme superfamilies are comprised 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, often requiring extensive biochemical studies on purified preparations of enzymes.[2,3] Given these issues, in conjunction with the flurry 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. Here, 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 based on 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 superfamily of enzymes.[7] FPs are well-characterized affinity labels for SHs, irreversibly phosphonylating the active site serine nuclelophile in these enzymes.[8] Numerous FP - labeled enzymes were identified in these experiments, including proteases, lipases, and esterases. Most of these enzymes could be readily assigned to the SH superfamily based on database (BLAST) searches, which identified conserved sequence elements shared by members of this enzyme class. Interestingly, however, one FP target, sialic acid 9-*O*-acetylesterase (SAE), which was selectively expressed in melanoma cell lines (Figure 1), shared no sequence homology with SHs or, for that matter, any other enzyme class. SAE was originally characterized by 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. Curious 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 for enzymes directly in complex proteomes.[12,13] In this method, probe-treated proteomes are digested with trypsin, and the probe-labeled peptides captured by affinity chromatography and identified by liquid-chromatography-tandem mass spectrometry (LC-MS/MS). The secreted proteome from COS-7 cells transiently transfected with the human SAE cDNA was treated with FP-Rh and analyzed by gel-free ABPP. 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 on serine 127 (S127; Figure 2). Confidence in this assignment was bolstered by the detection of multiple forms of this probe-labeled peptide that differed in methionine-oxidation state. In each case, the site of probe modification was designated as S127.

Mutagenesis of S127 to alanine produced an SAE variant that expressed at wild type levels in COS-7 cells, but failed to label with a rhodamine-tagged FP probe (FP-Rh[14]) (Figure 3A). Conversely, mutagenesis of the neighboring serine residue, S124, to alanine

generated an enzyme that reacted avidly with FP-Rh (Figure 3A). We next compared the catalytic activity of WT-SAE and the S124A and S127A mutants using a radiolabeled substrate assay. Both 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 3B). In contrast, the S127A SAE mutant showed negligible catalytic activity (Figure 3B). Collectively, these data indicate that S127 fulfills two of the major requirements expected of a SH catalytic nucleophile, as the residue was: 1) specifically labeled by FP inhibitors, and 2) essential for catalysis.

If S127 serves as the catalytic nucleophile in SAE, this residue should be conserved among sequence-related homologs 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 homologs (which share less than 25% sequence identity) (Figure 4). No other conserved serine residues were identified 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, which bridges the serine nucleophile and an aspartate to from 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 are present in the SAE family (H377 and K381). Future studies should clarify whether these residues also play critical roles in the SAE catalytic mechanism.

In summary, the findings reported herein indicate that SAE and its sequence homologs constitute a novel branch of the SH superfamily. This classification was fortified by the reactivity of SAE with FP probes, a set of affinity agents that display exquisite selectivity for SHs. It is interesting to speculate on whether the assignment of enzymes to other mechanistic classes will also be facilitated by ABPP probes. On the one hand, extrapolating mechanistic information with active site-directed probes that target multiple classes of enzymes may prove challenging.[19,20] On the other hand, several ABPP probes do show preferential reactivity with a single class of enzymes (e.g., hydroxamate probes for MPs,[21,22] E-64-based probes for cysteine proteases[23]). In these cases, like with the FP probes, ABPP may provide a useful method to uncover cryptic members of enzyme superfamilies that have eluded classification based on sequence comparisons.

References

[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. U.S.A.* **2002**, *99*, 10335-10340.
[7] Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. U.S.A.* **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. Amer. 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. U.S.A.* **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] G. C. Adam, E. J. Sorensen, B. F. Cravatt, *Nat Biotechnol* **2002**, *20*, 805-809.
[20] K. T. Barglow, B. F. Cravatt, *Chem Biol* **2004**, *11*, 1523-1531.
[21] A. Saghatelian, N. Jessani, A. Joseph, M. Humphrey, B. F. Cravatt, *Proc Natl Acad Sci U S A* **2004**, *101*, 10000-10005.
[22] E. W. Chan, S. Chattopadhaya, R. C. Panicker, X. Huang, S. Q. Yao, *J Am Chem Soc* **2004**, *126*, 14435-14446.
[23] D. Greenbaum, K. F. Medzihradszky, A. Burlingame, M. Bogyo, *Chem. Biol.* **2000**, *7*, 569-581.
[24] H. H. Higa, A. Manzi, A. Varki, *J Biol Chem* **1989**, *264*, 19435-19442.

Figure Legends

Figure 1. SAE activity is enriched in human melanoma cell lines. A comparison of FP-Rh-labeled SAE signals in the secreted proteomes of representative breast (MCF7), ovarian (OVCAR3), and melanoma (MUM-2C, MDA-MB-435, UACC-257) cancer cell lines. Fluorescent signals shown in gray scale.

Figure 2. Identification of the site of FP-Rh labeling in SAE by gel-free ABPP. Conditioned media proteome (2 mL, 1 mg protein/mL) from COS-7 cells recombinantly expressing human SAE was labeled with FP-Rh (4 μM) for 1 hr. The probe-treated proteome was then digested with trypsin and FP-Rh labeled peptides affinity purified using anti-Rh monoclonal antibody beads as described.[12] The affinity purified Rh-labeled peptides were analyzed by LC-MS/MS and the resulting MS/MS spectra searched against a human protein database using the SEQUEST algorithm. The data were filtered and analyzed as described.[12,13] The normalized Xcorr values from the SEQUEST results (shown as bars in the plot) were plotted against the predicted distribution of false positive results in the same dataset (stipled blue line). Bars in red indicate results representing 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 SAE peptides randomly occurring with the indicated scores was calculated at less than 0.0001%. Green bars indicate results representing other SHs in the proteome with probe modifications on their active site serine nucleophile.

Figure 3. 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-tagged fusion proteins and affinity purified using anti-myc agarose beads [Invitrogen; 50 μL beads with 0.75 mL of proteome in 50 mM Tris, pH 7.5 (1 mg protein/mL)]. Following thorough washing of the beads (2X with 0.5 mL 50 mM Tris, pH 7.5), an aliquot of each SAE variant was treated with FP-Rh (4 μM) for 1 h. The reactions were then quenched with 1X standard SDS-PAGE 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 using a radiolabeled substrate assay that follows the cleavage of [^{3}H]-acetate from [^{3}H]-Neu5,9Ac$_2$ (10,000 cpm/reaction). This assay[11] and the biosynthetic preparation of [^{3}H]-Neu5,9Ac$_2$[24] have been previously described. 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.

Figure 4 . Sequence comparisons of SAE and representative bacterial homologs. Shown is the region surrounding S127 (red), the predicted serine nucleophile of SAE. S127 is

completely conserved among all SAE sequence homologs present in public databases. Additional conserved residues are shown in bold. S124, which is mostly, but not completely conserved, is shown in blue.

Figure 1

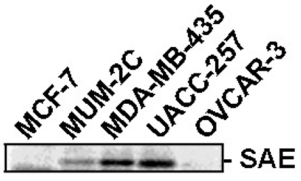


Figure 2

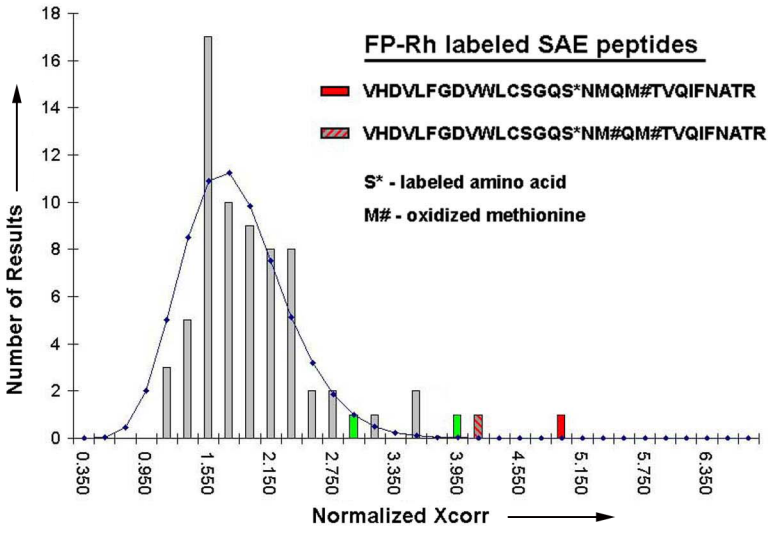

FP-Rh labeled SAE peptides
VHDVLFGDVWLCSGQS*NMQM#TVQIFNATR
VHDVLFGDVWLCSGQS*NM#QM#TVQIFNATR
S* - labeled amino acid
M# - oxidized methionine
Number of Results
Normalized Xcorr
0
2
4
6
8
10
12
14
16
18
0.350
0.950
1.550
2.150
2.750
3.350
3.950
4.550
5.150
5.750
6.350

Figure 3

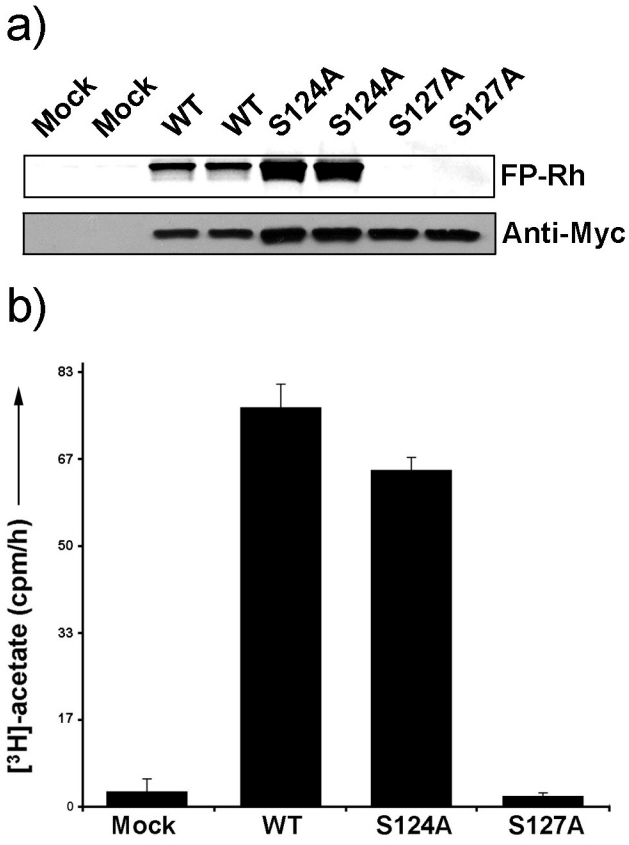
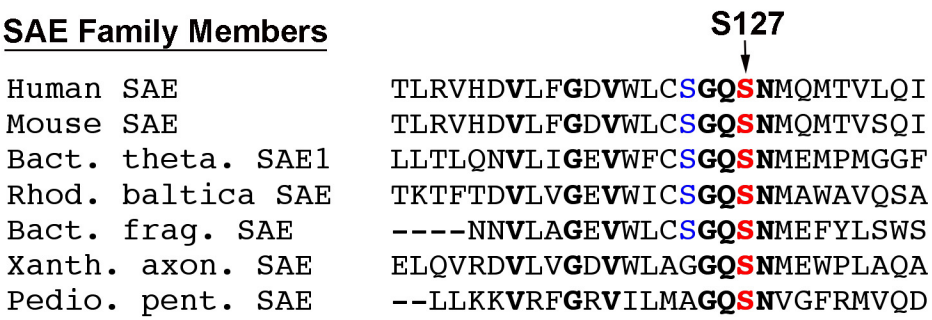

a)
Mock
Mock
WT
WT
S124A
S124A
S127A
S127A
FP-Rh
Anti-Myc
b)
[3H]-acetate (cpm/h)
83
67
50
33
17
0
Mock
WT
S124A
S127A

Figure 4


SAE Family Members
S127
Human SAE TLRVHDVLFGDVWLCSGQSNMQMTVLQI
Mouse SAE TLRVHDVLFGDVWLCSGQSNMQMTVSQI
Bact. theta. SAE1 LLTLQNVLIGEVWFCSGQSNMEMPMGGF
Rhod. baltica SAE TKTFTDVLVGEVWICSGQSNMAWAVQSA
Bact. frag. SAE ----NNVLAGEVWLCSGQSNMEFYLSWS
Xanth. axon. SAE ELQVRDVLVGDVWLAGGQSNMEWPLAQA
Pedio. pent. SAE --LLKKVRFGRVILMAGQSNVGFRMVQD