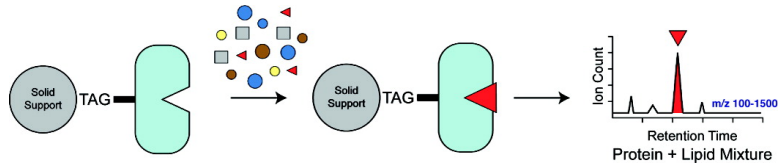
Communication

A Global Metabolite Profiling Approach to Identify Protein#Metabolite Interactions

Ranitendranath Tagore, Horatio R. Thomas, Edwin A. Homan, Ali Munawar, and Alan Saghatelian

J. Am. Chem. Soc., **2008**, 130 (43), 14111-14113 • DOI: 10.1021/ja806463c • Publication Date (Web): 03 October 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 10/03/2008

A Global Metabolite Profiling Approach to Identify Protein−Metabolite Interactions

Ranitendranath Tagore, Horatio R. Thomas, Edwin A. Homan, Ali Munawar, and Alan Saghatelian*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received August 14, 2008; E-mail: saghatelian@chemistry.harvard.edu

Biological systems utilize a series of elementary interactions to carry out an incredible array of cellular and physiological processes. During protein synthesis, for example, protein−DNA, protein−protein, and protein−RNA complexes are all necessary to produce a polypeptide.[1] In addition to these biopolymer systems, small molecules also play a vital role in biology through protein−metabolite interactions (PMIs).[2] The development of methods that identify connections betweens proteins and metabolites will complement current functional proteomic approaches. Here, we outline a process for assigning PMIs that relies on global metabolite profiling[3] to characterize protein-bound metabolites.

In the initial step of the workflow, a protein is immobilized onto a solid support through an affinity handle, such as a glutathione S-transferase (GST) fusion protein[4] (Figure 1). The bound protein is subsequently incubated with a mixture of cellular metabolites that are isolated from cells and tissues where the protein is naturally expressed. During the incubation step, metabolites can bind to the protein resulting in a protein–metabolite complex on the solid support. After isolation of this complex away from other metabolites by filtration, the protein is eluted, and the eluate is analyzed by using a liquid chromatography−mass spectrometry (LC−MS)-based global metabolite profiling platform. In contrast to "targeted" metabolite analysis methods, global profiling quantifies metabolites based on their absolute mass ion intensity (MSII), which obviates the need for internal standards and allows measurements of both known and structurally novel metabolites.[3] Finally, comparison of the LC−MS chromatograms from the protein sample to the appropriate controls, using the XCMS[5] software package, can identify specific protein bound metabolites in an unbiased manner.

Examples with radiolabeled lipids,[6] as well as targeted[7] and untargeted[8] MS detection methods have demonstrated the feasibility of immobilized protein-mediated metabolite enrichment. Moreover, immobilized proteins have also been used to screen for drug candidates from mixtures of synthetic compounds through similar enrichment protocols.[9] Using a series of lipid binding proteins (LBPs), we attempt to combine protein-mediated metabolite enrichment with global metabolite profiling. If successful, the integration of these methods will enable the discovery of PMIs without requiring any prior knowledge of the structure of the metabolite.

LBPs carry out a variety of different functions in the cell, ranging from lipid transport to lipid sensing and signaling.[10−12] We chose cytosolic retinoic acid binding protein 2 (CRABP2),[13] a tight binder of retinoic acid (RA), to develop this approach. *In vivo*, CRABP2 regulates aspects of RA metabolism[14] and signaling[15] through the binding and transport of RA. Recombinant expression of CRABP2 as a GST fusion protein (CRABP-GST) provided active protein, as determined by fluorescent binding assays with RA and 1-anilinonaphthalene-8-sulfonic acid (ANS) (Supporting Information).

Incubation of CRABP-GST at room temperature for thirty minutes with glutathione resin saturated the resin with protein

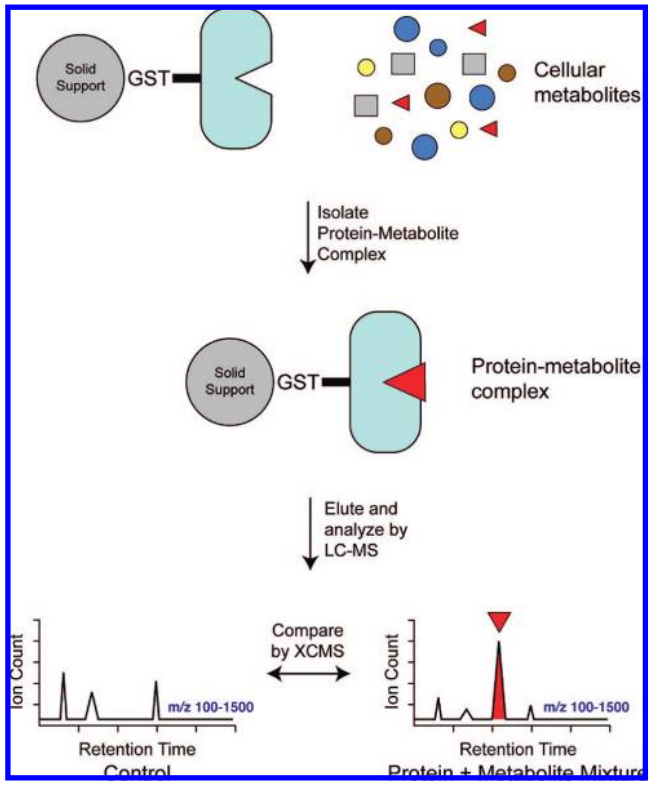


Figure 1. Assignment of protein−metabolite interactions using global metabolite profiling. An immobilized metabolite binding protein is incubated with a mixture of small molecule metabolites, resulting in the formation of a protein−metabolite complex between the protein and its natural binding partner. The protein−metabolite complex is separated from the unbound metabolites and subsequently eluted from the solid support. Global metabolite profiling of the eluate and comparison to the appropriate controls (no protein and protein without lipid), using the XCMS software package, identifies those metabolites specifically enriched by the protein.

(Supporting Information). The final amount of active CRABP-GST on the solid support was estimated to be 5 μg/μL of resin, based on RA binding experiments and the amount of protein released upon elution with glutathione. We developed the enrichment protocols using a binary mixture of RA and ^{13}C-oleic acid (^{13}C-OA). Starting with two lipids, instead of a complex lipid mixture, simplified the optimization of the lipid enrichment protocols. The selection of the unnatural ^{13}C-OA sidestepped any challenges that could arise through background ion interference from naturally occurring ^{12}C-OA while optimizing the enrichment step.

Resin bound CRABP-GST (5 nmol total protein) was incubated with RA and ^{13}C-OA (4 nmol, 20 μM) for 1 h at room temperature. After incubation, the unbound lipid was rapidly filtered away from the resin, which was then briefly washed. Following the elution of

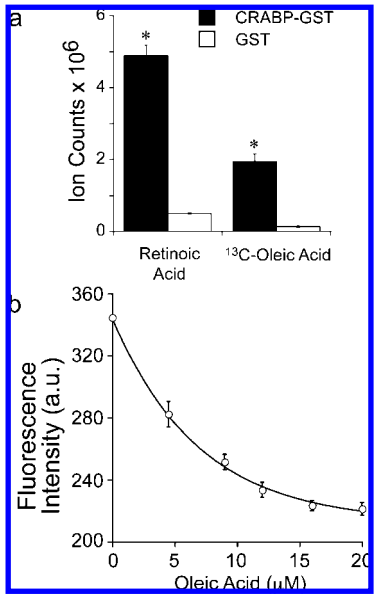


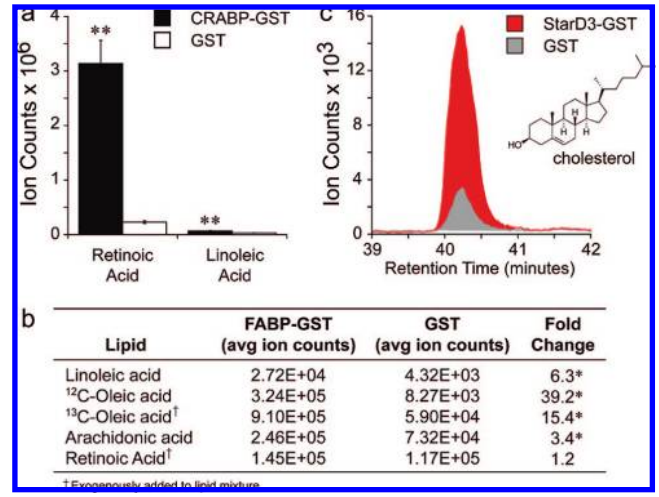
Figure 2. Enrichment experiments and binding of fatty acids to CRABP-GST. (a) Binary mixture of RA and ^{13}C-OA was added to resin-bound CRABP-GST. After incubation, the resin-bound protein was isolated, washed, and eluted prior to LC−MS analysis. Comparison of the ion counts from the CRABP-GST sample (black) to the GST control (white) showed that CRABP-GST enriched RA and ^{13}C-OA. (b) Fatty acid binding site on CRABP-GST was demonstrated using displacement of ANS by OA in a solution-based fluorescence assay ([CRABP] = 10 μM, [ANS] = 500 μM, ex. 400 nm, em. 500 nm). (*, p-value < 0.01, $N = 4$, Student's t test.)

the CRABP-GST-metabolite complex with a glutathione solution (10 mM), the eluate was analyzed directly using negative mode ionization LC−MS.

We analyzed the data by integrating the peak areas for RA and ^{13}C-OA in the LC−MS chromatograms from CRABP-GST, GST, and no lipid samples. The highest level of RA was found in the CRABP-GST sample (Figure 2a), which demonstrated the ability to detect CRABP-GST mediated binding and enrichment of RA. Unexpectedly, the data also revealed that ^{13}C-OA was enriched in the CRABP-GST sample. This enrichment was surprising because CRABP2 has not been reported to bind fatty acids, but it is possible that the fusion of CRABP2 to GST created a binding site for fatty acids. We confirmed fatty acid binding to CRABP-GST by displacing ANS from CRABP-GST with OA[16] (Figure 2b). These initial experiments demonstrate the identification and discovery of binding interactions between proteins and small molecule metabolites from simple lipid mixtures.

In practice, however, this feat needs to be accomplished from a much more complex lipid mixture, such as the lipid extract from a cell or a tissue, to be considered a generally useful approach for identifying natural binding partners. Thus, we set out to test whether we could identify CRABP2 binding partners by enriching RA from a complex mixture of cellular lipids. In these experiments, the complex lipid mixture was composed of fresh lipid extract from mouse tissue (brain)[3] and exogenously added RA and ^{13}C-OA. Analysis of this complex mixture by LC−MS identified fatty acids, RA (added), phospholipids, acyl glycerols, cholesterol esters, and cholesterol.[3]

A portion of this mixture (corresponding to 1 nmole of RA) was incubated with CRABP-GST bound beads. After incubation, the beads were washed, the CRABP-GST was eluted, and the eluate was analyzed by LC−MS. These experiments were performed in positive and negative ionization modes to provide coverage over the entire metabolome. In addition, three control experiments were



b

Lipid	FABP-GST (avg ion counts)	GST (avg ion counts)	Fold Change
Linoleic acid	2.72E+04	4.32E+03	6.3*
^{12}C-Oleic acid	3.24E+05	8.27E+03	39.2*
^{13}C-Oleic acid†	9.10E+05	5.90E+04	15.4*
Arachidonic acid	2.46E+05	7.32E+04	3.4*
Retinoic Acid†	1.45E+05	1.17E+05	1.2

†Exogenously added to lipid mixture.

Figure 3. Identification of specific binding partners from complex lipid mixtures. (a) RA and C18:2 free fatty acid are specifically enriched in CRABP-GST samples (black) from a complex mixture of brain lipids, which were spiked with RA and ^{13}C-OA. (b) Same experiment performed with FABP-GST results in the enrichment of free fatty acids (linoleic acid, ^{12}C-OA, ^{13}C-OA, and arachidonic acid) but not RA. (c) StarD3 mediated enrichment of cholesterol from a brain lipid extract. (*, p-value <0.05; **, p-value <0.01, $N = 3-4$, Student's t test.)

also analyzed (CRABP-GST without lipid, GST with lipid, and GST without lipid). The samples that lack any added lipid reveal those background ions that are associated with protein prior to the enrichment step.

In this case, we tested whether XCMS could be used to identify enriched ions from these samples, which circumvents any need for prior structural knowledge of the protein bound metabolites. XCMS aligns, quantifies, and statistically ranks differences between two sets of LC−MS chromatograms (e.g., CRABP-GST vs GST samples) based on the ion intensity of a metabolite peak. The data analysis was divided into two steps: (1) automated analysis using XCMS followed by (2) manual processing of the XCMS output files (Supporting Information). Because XCMS is only designed to perform pairwise comparisons, manual processing was necessary to identify the overlap in CRABP-GST enriched ions from the XCMS analysis of CRABP-GST with lipid vs CRABP-GST without lipid and CRABP-GST with lipid vs GST with lipid samples. Despite this manual intervention, overall this process was still completely unbiased since the initial lists of ions were all generated automatically.

After analysis of the CRABP-GST samples with XCMS, and manual cross-referencing of the XCMS output files, we were left with two statistically significant changes (m/z 299 and 279). Based on known elution times, and coelution with standards, as well as the accurate masses we confirmed that these ions correspond to RA and linoleic acid (C18:2 free fatty acid) (Figure 3a). Moreover, the fold change for RA (~14-fold) is much larger than linoleic acid (2-fold), as expected for CRABP. These results validate our data analysis method and, more generally, highlight the utility of global metabolite profiling in identifying natural PMIs from complex lipid mixtures.

The lack of ^{13}C-OA enrichment is an obvious, but not surprising, difference when comparing the analysis of complex mixtures to those using only binary mixtures. We speculate that during our initial experiments ^{13}C-OA was binding to a nonspecific binding site(s) on CRABP-GST. When the same experiment is repeated in the presence of a complex lipid mixture the nonspecific binding site(s) were occupied by a variety of different metabolites, which

means that no single species, including OA, was significantly enriched. Thus, the use of a complex lipid mixture benefits the analysis because it provided an internal mechanism to reduce the signal associated with nonspecific binding events.

The next goal was to determine whether this method is general and can be used for other classes of LBPs. Specifically, we looked at the binding interactions of fatty acid binding protein 2 (FABP2)[17] and StAR-related lipid transfer (START) domain containing 3 (StarD3).[18] FABP2 is an intracellular intestinal fatty acid binding protein that has been linked to metabolic disorders and cardiovascular disease.[19] We expressed FABP2 as a GST fusion (FABP-GST) and confirmed the activity upon binding of OA using an ANS displacement assay[20] (Supporting Information). Resin bound FABP-GST was incubated with brain lipid extracts, with exogenously added RA (negative control) and ^{13}C-OA (positive control), washed, and eluted as described previously. Analysis of the eluate by global metabolite profiling and XCMS revealed the specific enrichment of fatty acids, but not RA, by FABP from the lipid extracts (Figure 3b). Indeed, oleic (^{12}C and ^{13}C), linoleic, and arachidonic acid were all enriched 3−40-fold in the FABP-GST samples, consistent with observed lipid specificity of FABP2.[21]

Lastly, to ensure that the affinity enrichment is not biased toward anionic lipids, we examined StarD3. StarD3 is a member of the StarD family of LBPs that are found in mammalian systems and is responsible for binding and shuttling cholesterol from the outer plasma membrane to organelles within the cell for use in metabolism.[22] The binding activity of the StarD3-GST fusion was confirmed using cholesterol-nitrobenzoxadiazole (ch-NBD),[23] a fluorescent cholesterol derivative that increases its fluorescence upon binding to StarD3 (Supporting Information). Resin bound StarD3-GST was incubated with the lipid mixture, washed, eluted, and analyzed by global metabolite profiling. After data analysis with XCMS, the natural ligand of StarD3, cholesterol, was identified as the only significant difference between the samples and controls (Figure 3c). Together the examples of CRABP, FABP, and StarD3 underscore the potential and generality of global metabolite profiling in the analysis of protein−lipid interactions.

The global metabolite profiling platform described here provides broad coverage across the lipid metabolome and will enable the classification of other LBPs, even if the structures of their lipid binding partners are currently unknown.[3] Moreover, the use of different analytical conditions will expand the metabolome coverage[24,25] and in the process enable the identification of binding interactions between proteins and other classes of metabolites, such as hydrophilic small molecules.[8] In the future, we plan to apply this approach to discover the interactions between metabolites and orphan lipid binding proteins, such as nuclear receptors.[12] Additionally, this method might also be extended for the identification of interactions between proteins and synthetic small molecules,[26] providing a parallel strategy for small molecule screening.

Acknowledgment. We thank Dr. James Cardia for helpful discussions. We thank the Burroughs Wellcome Fund for a Career Award in the Biomedical Sciences (A.S.), Harvard University for start-up funds, and the NSF for a predoctoral fellowship (E.A.H.).

Supporting Information Available: Detailed experimental procedures and characterization of all proteins is described. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 6th ed.; W. H. Freeman: New York, 2006.
(2) Devane, W. A.; Hanus, L.; Breuer, A.; Pertwee, R. G.; Stevenson, L. A.; Griffin, G.; Gibson, D.; Mandelbaum, A.; Etinger, A.; Mechoulam, R. *Science* **1992**, *258*, 1946–1949.
(3) Saghatelian, A.; Trauger, S. A.; Want, E. J.; Hawkins, E. G.; Siuzdak, G.; Cravatt, B. F. *Biochemistry* **2004**, *43*, 14332–14339.
(4) Smith, D. B.; Johnson, K. S. *Gene* **1988**, *67*, 31–40.
(5) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 779–787.
(6) Rodriguez-Agudo, D.; Ren, S.; Hylemon, P. B.; Redford, K.; Natarajan, R.; Del Castillo, A.; Gil, G.; Pandak, W. M. *J. Lipid Res.* **2005**, *46*, 1615–1623.
(7) Ortlund, E. A.; Lee, Y.; Solomon, I. H.; Hager, J. M.; Safi, R.; Choi, Y.; Guan, Z.; Tripathy, A.; Raetz, C. R.; McDonnell, D. P.; Moore, D. D.; Redinbo, M. R. *Nat. Struct. Mol. Biol.* **2005**, *12*, 357–363.
(8) Morozov, V. N.; Morozova, T. Y.; Johnson, K. L.; Naylor, S. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2430–2438.
(9) Annis, D. A.; Nickbarg, E.; Yang, X.; Ziebell, M. R.; Whitehurst, C. E. *Curr. Opin. Chem. Biol.* **2007**, *11*, 518–526.
(10) Bernlohr, D. A.; Simpson, M. A.; Hertzel, A. V.; Banaszak, L. J. *Annu. Rev. Nutr.* **1997**, *17*, 277–303.
(11) Vivanco, I.; Sawyers, C. L. *Nat. Rev. Cancer* **2002**, *2*, 489–501.
(12) Mangelsdorf, D. J.; Thummel, C.; Beato, M.; Herrlich, P.; Schutz, G.; Umesono, K.; Blumberg, B.; Kastner, P.; Mark, M.; Chambon, P.; Evans, R. M. *Cell* **1995**, *83*, 835–839.
(13) Fiorella, P. D.; Napoli, J. L. *J. Biol. Chem.* **1991**, *266*, 16572–16579.
(14) Bucco, R. A.; Zheng, W. L.; Davis, J. T.; Sierra-Rivera, E.; Osteen, K. G.; Chaudhary, A. K.; Ong, D. E. *Biochemistry* **1997**, *36*, 4009–4014.
(15) Schug, T. T.; Berry, D. C.; Shaw, N. S.; Travis, S. N.; Noy, N. *Cell* **2007**, *129*, 723–733.
(16) Kane, C. D.; Bernlohr, D. A. *Anal. Biochem.* **1996**, *233*, 197–204.
(17) Plotz, P. H.; Berk, P. D.; Scharschmidt, B. F.; Gordon, J. K.; Vergalla, J. *J. Clin. Invest.* **1974**, *53*, 778–785.
(18) Lowe, J. B.; Sacchettini, J. C.; Laposata, M.; McQuillan, J. J.; Gordon, J. I. *J. Biol. Chem.* **1987**, *262*, 5931–5937.
(19) Tsujishita, Y.; Hurley, J. H. *Nat. Struct. Biol.* **2000**, *7*, 408–414.
(20) Baier, L. J.; Sacchettini, J. C.; Knowler, W. C.; Eads, J.; Paolisso, G.; Tataranni, P. A.; Mochizuki, H.; Bennett, P. H.; Bogardus, C.; Prochazka, M. *J. Clin. Invest.* **1995**, *95*, 1281–1287.
(21) Chuang, S.; Velkov, T.; Horne, J.; Porter, C. J. H.; Scanlon, M. J. *J. Med. Chem.* **2008**, *51*, 3755–3764.
(22) Zimmerman, A. W.; van Moerkerk, H. T.; Veerkamp, J. H. *Int. J. Biochem. Cell Biol.* **2001**, *33*, 865–8676.
(23) Alpy, F.; Tomasetto, C. *J. Cell Sci.* **2005**, *118*, 2791–2801.
(24) Reitz, J.; Gehrig-Burger, K.; Strauss, J. F., 3rd; Gimpl, G. *FEBS J.* **2008**, *275*, 1790–1802.
(25) Yuan, J.; Fowler, W. U.; Kimball, E.; Lu, W.; Rabinowitz, J. D. *Nat. Chem. Biol.* **2006**, *2*, 529–530.
(26) Carlson, E. E.; Cravatt, B. F. *Nat. Methods* **2007**, *4*, 429–435.
(27) Duffner, J. L.; Clemons, P. A.; Koehler, A. N. *Curr. Opin. Chem. Biol.* **2007**, *11*, 74–82.

JA806463C