DOI: 10.1002/cbic.200900010

Unraveling the Secrets of Protein–Metabolite Interactions

Stephan A. Sieber*[a]

Genome-sequencing projects have provided a wealth of information on gene identities in many prokaryotic and eukaryotic organisms. Technological advancements in biological, chemical and analytical sciences have provided methods in the fields of genomics and proteomics that allow molecular, cellular and (patho)physiological functions to be monitored and investigated for the full complement of genes and proteins.[1] Although mRNA gene-expression data and proteomic analyses led in many cases to the identification of disease-associated enzyme targets, one important component of cellular physiology has not been addressed—the physiological substrates of these enzymes, also referred to as metabolites.[2]

The metabolome comprises the full complement of small molecules (metabolites) within a cell, tissue or whole organism,[3] encompassing a broad diversity of molecules including peptides, carbohydrates, lipids, nucleosides, and the catabolic products of exogenous molecules.[4] Many of these metabolites contribute to numerous physiologically and pathologically relevant metabolic and signaling pathways. The elucidation of these pathways contains several challenges and requires new strategies for rapidly and systematically identifying the metabolic substrates of relevant enzymes. Contrary to genomic and proteomic methods, metabolites share no direct link to the genetic code and are instead products of enzymatic networks in cells and tissues.[2] Moreover, metabolites are composed of diverse chemical structures and vary in their physical properties, and their composition is not a result of sequence-dependent processing, as observed in RNA or proteins. These distinguishing features make the metabolome a challenging part of biomolecular space that requires new customized tools for its analysis and to unravel its secrets.

Rising to this challenge, Saghatelian and colleagues recently introduced a global metabolite-profiling approach that allowed protein–metabolite interactions (PMI) to be directly identified.[5] One huge advantage of this method is that, in principle, even the protein interactions of unknown metabolites can be detected. The workflow of this approach starts with immobilization of the protein of interest by a glutathione S-transferase (GST) tag onto solid support (Figure 1). The activity of the immobilized protein is tested by appropriate assays (for example, fluorescent substrates) to ensure that the catalytic power is not perturbed by the immobilization process. Subsequent incubation with mixtures of cellular metabolites originating from cells or tissues in which the protein of interest is naturally expressed leads to capture of the dedicated metabolite and results in a protein–metabolite complex on the solid support. Excess metabolites are subsequently washed away, and the desired protein–metabolite complex is selectively eluted from the solid support by the release of GST-tagged protein with an excess of glutathione. The eluate containing the protein–metabolite complex is subsequently analyzed by liquid

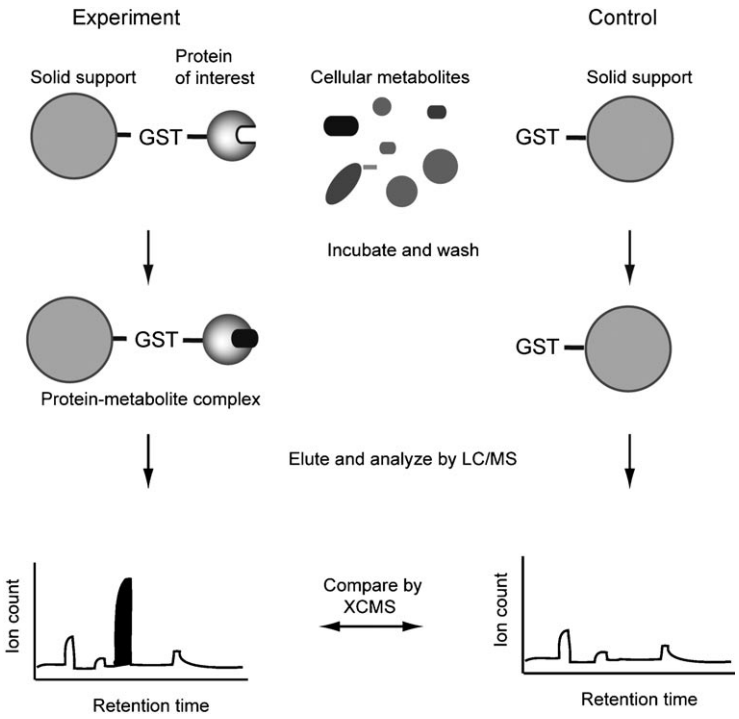


Figure 1. Assignment of protein–metabolite interactions. An immobilized protein is incubated with a mixture of small-molecule metabolites to form a protein–metabolite complex. Elution from the solid support and subsequent analysis by LC/MS together with comparisons of appropriate controls (for example, no protein on solid support) using the XCMS software, identifies those metabolites that are specifically enriched by the protein.

[a] *Dr. S. A. Sieber*
Center for integrated Protein Science Munich CiPSM
Department of Chemistry and Biochemistry
Ludwig-Maximilians-Universität München
Butenandtstrasse 5–13
81377 Munich (Germany)
Fax: (+49) 89-2180-77756
E-mail: stephan.sieber@cup.uni-muenchen.de

chromatography–mass spectrometry (LC/MS) analysis, a method that is adapted from a global metabolite-profiling strategy introduced earlier.[6]

In global metabolite profiling, the LC/MS analysis of metabolites is carried out in an untargeted mode, quantifying metabolites based on their absolute mass ion intensity (MII). Since different metabolites exhibit different ionization efficiencies, only relative changes of the same metabolite in different samples can be quantified. Therefore, the relative levels are accurate within the linear range of the mass spectrometer, which is 3–4 orders of magnitude.[7] No internal standards are required, and measurements of both known and novel metabolites can be carried out.

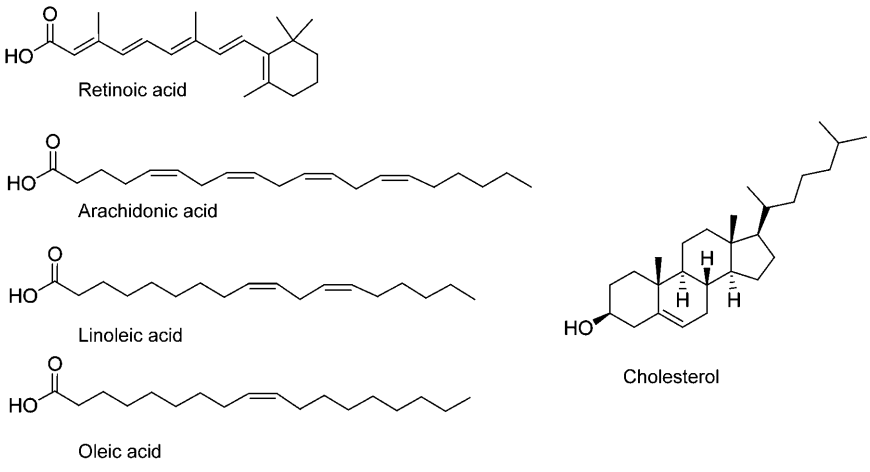
To identify the mass of the protein-bound metabolite, LC/MS chromatograms from the eluate and appropriate controls (metabolites without immobilized protein and immobilized protein without metabolites) are compared by using a specialized software (XCMS) that identifies mass peaks of those metabolites that are specifically enriched by the protein.[8] XCMS is designed to detect, align, quantify (based on area), and statistically rank peaks in LC/MS chromatograms through the pairwise comparison of two different data sets. The value of this experimental strategy results from the powerful combination of protein metabolite enrichment with untargeted global metabolite profiling that allows new PMIs to be detected without requiring any prior knowledge about the metabolite structure.

Saghatelian and colleagues first demonstrated the general utility of their methodology by studying a characterized protein–metabolite pair as a proof of concept.[5] For this initial test, GST affinity-tagged cytosolic retinoic acid binding protein 2 (CRABP2) was recombinantly expressed and immobilized on solid support. CRABP2 is a tight binder of retinoic acid (RA) that regulates its metabolism and signaling.[9,10] Incubation of the immobilized CRABP2 with a binary mixture of the dedicated lipid RA as well as ^{13}C-oleic acid (not a known substrate of CRABP2) for 1 h was followed by washing and elution of the protein-bound metabolite. LC/MS analysis of the eluate and the corresponding control without protein revealed the highest levels of RA in the CRABP2-GST sample, thus demonstrating that it is possible to detect protein-mediated binding and enrichment. Subsequent incubation of CRABP2-GST with a more complex lipid extract from mouse tissue delivered two enriched masses that corresponded to RA as well as linoleic acid (LA), as validated by coelution with corresponding standards.[5] As expected for CRABP2, the enrichment for RA was much higher than for LA; this emphasizes the value of this approach for identifying natural PMIs from complex lipid mixtures.

Following up on these results, the authors looked at the binding interactions of fatty acid binding protein 2 (FABP2), an important intracellular intestinal lipid-binding protein that has been linked to metabolic disorders and cardiovascular disease.[11] Contrary to CRABP2, FABP2 exhibits no affinity for RA and should, if the method works, reveal completely different metabolites. According to the general experimental layout, the enzyme was expressed with a GST-tag, immobilized on solid support, and incubated with brain lipid extracts. Elution of the proteins and LC/MS analysis revealed the specific binding and enrichment of oleic, linoleic, and arachidonic acids in the FABP2-GST sample but not in the corresponding controls.[5] These results are in good agreement with the current knowledge about FABP2, which is reported to exhibit lipid specificity. Controls that were exogenously spiked with RA did not lead to an enrichment of RA, thus confirming the unbiased and specific outcome of these experiments.

The authors had so far conducted their experiments solely with anionic lipids. In order to exclude a certain bias, an additional PMI between the lipid binding protein StarD3 and the uncharged compound cholesterol was investigated. StarD3 is a member of the StarD family of lipid binding proteins that are found in mammalian systems. It is responsible for binding and shuttling cholesterol from the outer plasma membrane to organelles within the cell for use in metabolism.[12] StarD3-GST resin was incubated with the lipid mixture, washed, and eluted, and the enriched metabolites were investigated by LC/MS. Data analysis confirmed that cholesterol was the only chemical compound that was significantly enriched in the protein sample compared to the controls.[5] This last experiment completed a set of assays that were necessary to confirm the potential value of the methodology towards the unbiased, reliable and specific identification of natural uncharacterized PMIs. Although useful, it should be noticed that it is beyond the scope of this method to use MS signal intensities to directly measure the corresponding affinities of protein–metabolite interactions.

The global metabolite-profiling platform for PMIs is an intriguing novel approach for a better understanding, characterization, and classification of lipid binding proteins, even if the structures of their dedicated partners are unknown.



Continuous efforts to establish searchable databases with all the chemical and physical properties of known metabolites together with advancements in the design of high-resolution analytical tools will further increase the speed and reliability of metabolite identification. Challenging future tasks for this methodology include the investigation of less-characterized PMIs such as nuclear hormone receptor–metabolite interactions or interactions between drug libraries and proteins. For these studies, the method needs to be validated also for hydrophilic compounds, and, additionally, the expression and immobilization of difficult-to-handle membrane proteins have to be established. With this envisioned broad applicability, the method has a huge potential to unravel the many secrets of PMIs.

Acknowledgements

I thank Prof. Dr. Thomas Carell and his group for their generous support and the excellent working environment. I gratefully acknowledge funding by the Emmy Noether Program of the Deutsche Forschungsgemeinschaft (DFG), a DFG grant (SFB 749), a stipend by the Römer-Stiftung, support by the Fonds der chemischen Industrie and by the Center for integrated Protein Science Munich CiPS^M.

Keywords: affinity capture · lipids · mass spectrometry · metabolites · proteins

[1] A. Saghatelian, B. F. Cravatt, *Nat. Chem. Biol.* **2005**, *1*, 130.
[2] A. Saghatelian, B. F. Cravatt, *Life Sci.* **2005**, *77*, 1759.
[3] O. Fiehn, *Plant Mol. Biol.* **2002**, *48*, 155.
[4] A. Saghatelian, B. F. Cravatt, *Curr. Opin. Chem. Biol.* **2005**, *9*, 62.
[5] R. Tagore, H. R. Thomas, E. A. Homan, A. Munawar, A. Saghatelian, *J. Am. Chem. Soc.* **2008**, *130*, 14111.
[6] A. Saghatelian, S. A. Trauger, E. J. Want, E. G. Hawkins, G. Siuzdak, B. F. Cravatt, *Biochemistry* **2004**, *43*, 14332.
[7] W. Lu, B. D. Bennet, J. D. Rabinowitz, *J. Chromatogr. B* **2008**, *871*, 236.
[8] C. A. Smith, E. J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, *Anal. Chem.* **2006**, *78*, 779.
[9] R. A. Bucco, W. L. Zheng, J. T. Davis, E. Sierra-Rivera, K. G. Osteen, A. K. Chaudhary, D. E. Ong, *Biochemistry* **1997**, *36*, 4009.
[10] T. T. Schug, D. C. Berry, N. S. Shaw, S. N. Travis, N. Noy, *Cell* **2007**, *129*, 723.
[11] Y. Tsujishita, J. H. Hurley, *Nat. Struct. Biol.* **2000**, *7*, 408.
[12] A. W. Zimmerman, H. T. van Moerkerk, J. H. Veerkamp, *Int. J. Biochem. Cell Biol.* **2001**, *33*, 865.

Received: January 7, 2009
Published online on March 2, 2009

