Review

CellPress

Special Issue: Systems Biology

Why measure enzyme activities in the era of systems biology?

Mark Stitt[1] **and Yves Gibon**[2]

[1] Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam-Golm, Germany
[2] INRA, University of Bordeaux, UMR 1332 Fruit Biology and Pathology, F-33883 Villenave d'Ornon, France

Information about the abundance and biological activities of proteins is essential to reveal how genes affect phenotypes. Over the past decade, mass spectrometry (MS)-based proteomics has revolutionized the identification and quantification of proteins, and the detection of post-translational modifications. Interpretation of proteomics data depends on information about the biological activities of proteins, which has created a bottleneck in research. This review focuses on enzymes in central metabolism. We examine the methods used for measuring enzyme activities, and discuss how these methods provide information about the kinetic and regulatory properties of enzymes, their turnover, and how this information can be integrated into metabolic models. We also discuss how robotized assays could enable the genetic networks that control enzyme abundance to be analyzed.

Protein world

Powerful analytic and computational methods have led to a revolution in identifying and measuring the abundance of RNA species [1]. However, most RNA species do not have a biological activity *per se*; rather they act as a template for the synthesis of proteins. Proteins interact in networks to generate phenotypes [2]. Protein abundance depends on not only transcript levels, but also the rate of translation and of protein degradation [3]. The properties and biological activities of proteins depend on their amino acid sequence and the resulting 3D structure. Importantly, the biological activity of proteins is modified and tuned by processes such as differential splicing, proteolysis, subcellular targeting, assembly into complexes, post-translational modification, and the levels of substrates and effectors. This increases flexibility and adaptability and shortens the time scale of information flow, enabling biological networks to rapidly take on new tasks and adjust to new conditions (Box 1). Therefore, information about the abundance and the biological activities of proteins is essential if we want to understand how gene function and genetic variation affect phenotypes. In this review, we focus on the special case of enzymes in central metabolism to illustrate the importance of obtaining information about the abundance and functional properties of proteins, and integrating this with information about network structure.

Corresponding author: Stitt, M. (mstitt@mpimp-golm.mpg.de).

1360-1385/$ – see front matter
 http://dx.doi.org/10.1016/j.tplants.2013.11.003

Proteomics and enzymology are complementary

The study of protein structure and function has been at the core of biochemistry for decades. However, methods that enable the large-scale identification of proteins and measurement of their abundance have only become available over the past decade. These methods are based on tandem liquid chromatography and MS [4–6]. This rapidly expanding research field is being driven by improvements in instrumentation, increasingly powerful algorithms for analyzing large data sets, and an explosion of data about protein sequences inferred from genome and transcriptome sequencing. It is now possible to perform a quantitative analysis of essentially the entire yeast proteome [6] and a large part of the proteome of a mammalian cell culture [4]. In plants, the greatest number of proteins quantified per sample to date is approximately 2000 [7,8]. The lower coverage may reflect the complexity of cell types in plant organs and is likely to increase in the near future. MS is also providing powerful methods to detect isoforms due to differential splicing or proteolysis and to monitor protein modifications, such as phosphorylation [9–12] and redox modification [13,14].

One might ask: why, in the 21st century, is it still necessary to assay enzymes? The key point is that enzyme assays provide functional information (Box 1), such as what reaction(s) an enzyme catalyzes. When performed under optimized conditions, they enable measurement of maximum activity, or capacity, which is a fast and cheap proxy for protein abundance. Ideally, abundance should be cross-checked in baseline studies using antibodies or MS. Activity assays are also the only way to learn about regulatory characteristics: specific activity (k_{cat}), substrate kinetics, cooperativity, whether the enzyme is regulated by effector molecules, as well as quantitative information about the affinity for substrates and effectors (e.g., K_m, K_a, and K_i) and how these are modified by sequence changes or protein modification. Information about capacities and properties is essential to understand how enzymes operate in metabolic networks (see below). Of all types of protein, enzymes are arguably the class for which the most

Box 1. Gene expression, protein abundance, and enzyme activity

Linking gene expression to enzyme activity

Transcriptomics provides a comprehensive readout of signaling activity. However, the impact of changes in transcript abundance on protein abundance will depend on multiple parameters affecting transcript stability, translation, protein assembly, and protein degradation. Thus, dramatic changes in gene expression can have little impact on protein abundance, and protein abundance can change independently of change in transcription rates. Protein activity depends not only on protein abundance, but also on substrate and product availability, and often on a plethora of small-molecular-weight effectors and on various post-translational modifications (thiol reduction illustrated in Figure I). Proteomics is able to quantify protein abundance and detect post-translational events. When k_{cat} is known, enzyme capacity can be estimated from protein abundance. Enzymology collects functional information about proteins (e.g., k_{cat}, K_m, and K_i). Each enzyme feature under study requires a dedicated protocol and, to date, it has not been possible to perform enzymology on a genome scale.

Relation between protein abundance, catalytic capacity, and the *in vivo* rate of catalysis

Protein abundance can be measured either by quantitative proteomics, or by proxy by assaying enzyme activities under standardized conditions that enable maximal catalytic activity and include appropriate controls to safeguard against loss of activity during extraction and assay, and by using published data on k_{cat} to interconvert values for maximum activity and protein abundance (Figure II).

Activity *in vivo* is lower, often much lower, than the maximum activity (Figure II). This can be due to many factors, including operation of the reverse reaction, limiting substrate, inhibitors, limiting levels of activators and incomplete post-translational activation. *In vivo* activity can, at best, be estimated. This requires detailed knowledge of the kinetic and regulatory properties of the enzyme, or modeling of *in vivo* flux over the enzyme from experimental analyses of pathway flux.

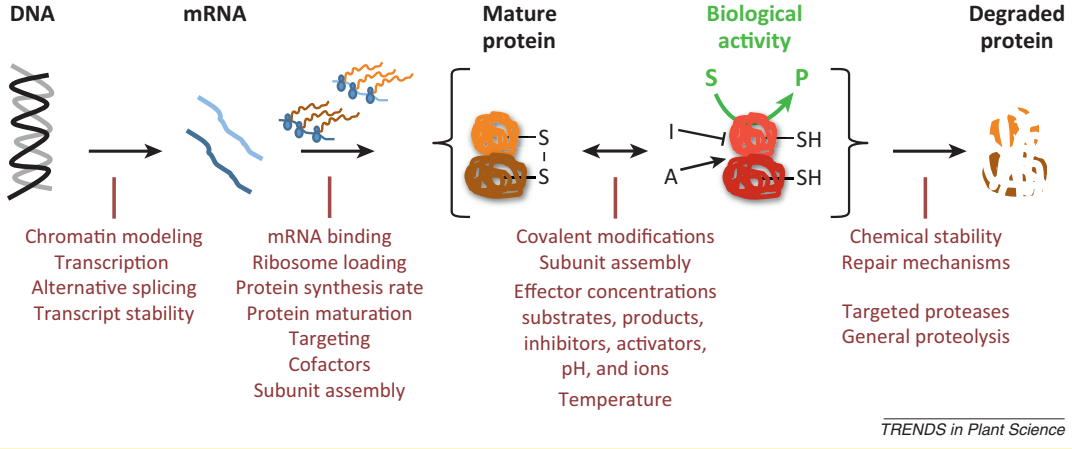

Figure I. Linking gene expression to enzyme activity, shown here for thiol reduction.

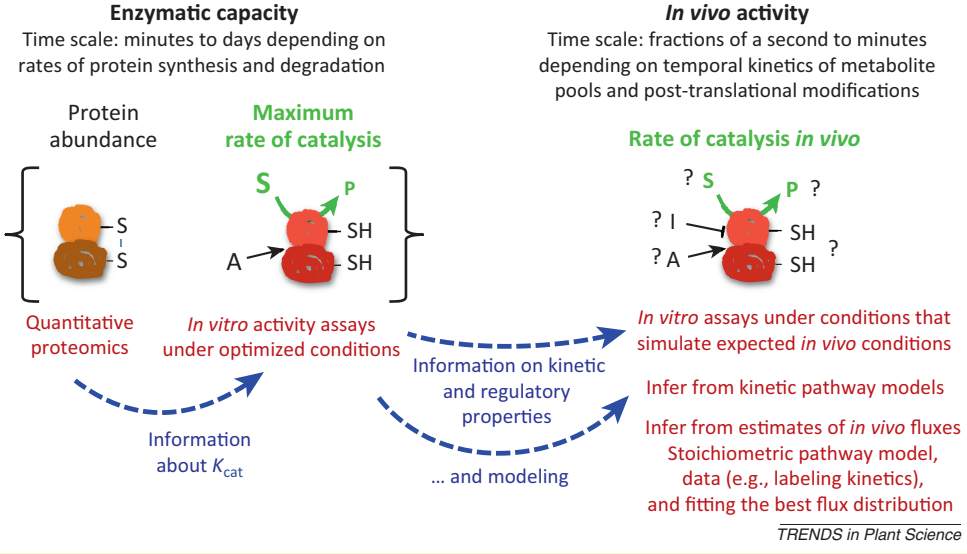

Figure II. Relation between protein abundance, catalytic capacity, and the *in vivo* rate of catalysis.

functional information can be gained, at least with current biochemical and cellular methods.

Proteomics and enzymology provide complementary information (Figure 1). Proteomics is an increasingly powerful method for inventorying protein types, abundance, sequence, modifications, and putative regulatory motifs [e.g., AtProteome [7] (http://fgcz-atproteome.unizh.ch/); Promex [15] (http://promex.pph.univie.ac.at/promex/); The Plant Proteome Database [16] (http://ppdb.tc.cornell.edu/); and PhosPhAt [17] (http://phosphat.mpimp-golm.mpg.de/)].