Lessons on enzyme kinetics from quantitative proteomics

Lessons on enzyme kinetics from quantitative proteomics

Dan Davidi [a] and Ron Milo [a,1]

[a] Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 76100, Israel

[1] To whom correspondence should be addressed. Email: ron.milo@weizmann.ac.il.

Abstract

Enzyme kinetics are fundamental to an understanding of cellular metabolism and for crafting synthetic biology applications. For decades, enzyme characterization has been based on *in vitro* enzyme assays. However, kinetic parameters are only available for < 10% of reactions, and this data scarcity limits the predictive power of metabolic models. Here we review recent studies that leverage quantitative proteomics to gain insight into *in vivo* enzyme kinetics. We discuss findings on the relationship between *in vivo* and *in vitro* enzyme catalysis and show how proteomics can be used to characterize the efficiency of enzyme utilization across conditions. Lastly, the efficient use of enzymes is shown to rationalize preference for low energy-yield metabolic strategies, such as aerobic fermentation at high growth rate.

Highlights

1. Enzyme kinetic data is sparse, limiting the predictive power of metabolic models.
2. Quantitative proteomics can give insight into *in vivo* enzyme kinetics.
3. *In vitro* k_{cat} values and maximal *in vivo* catalytic rates generally concur.
4. Efficient use of enzymes rationalizes usage of low energy-yield metabolic pathways.

Introduction: kinetic data scarcity limits the scope of metabolic models

Enzyme kinetics are commonly characterized by the apparent parameters k_{cat} and K_M, i.e., the maximal turnover rate of the enzyme and the affinity (Michaelis constant) for reactants. Many models of cellular metabolism use k_{cat} and K_M as inputs to predict metabolic behaviors [1]. However, more than 100 years after the introduction of k_{cat} and K_M by Michaelis and Menten in 1913 [2], coverage remains poor. In Figure 1 we show the fraction of measured k_{cat} values out of all known metabolic reactions in *Escherichia coli, Saccharomyces cerevisiae, Arabidopsis thaliana* and *Homo sapiens*. We evaluated coverage on a per-organism basis because some k_{cat}

values vary by orders of magnitude between organisms [3–5]. For *E. coli*, the most extensively biochemically characterized organism, k_{cat} values are available for only ≈9% of its ≈2000 enzyme-reaction pairs. Note that the actual coverage is even lower, if one considers the scarcity of information on k_{cat} for the backwards direction in reactions that are physiologically reversible (about half of the reactions in *E. coli* [6]). For other organisms, the picture is even more grim.

Measured k_{cat} values span over six orders of magnitude [5]. This combination of wide dynamic range and extremely limited coverage highlights a major impediment to our ability to construct models and predict metabolic behaviors. For example, the ≈130 values reported in *E. coli* span a range well over a thousand fold with a standard deviation of 1.4 on a $\log_{10}$ scale. Thus, assignment of k_{cat} values to an unmeasured enzyme in *E. coli* by random sampling from the distribution (as is often done [7,8]), is likely to err by more than tenfold on average. This can lead to major prediction errors for example in models of protein burden. Unfortunately, the collection of kinetic constants remains painstaking and proceeds slowly.

As reviewed below, another concern in the current use of kinetic constants is that reported values are usually measured in non-physiological conditions, often without endogenous effectors present [9]. These *in vitro* measurements lack *in vivo* elements that are known to have central roles in modulating enzyme kinetics e.g., allosteric regulation and posttranslational modifications [10,11]. In this review we focus on recent progress on estimating enzyme kinetic parameters from high-throughput *in vivo* measurements, thereby alleviating data scarcity.

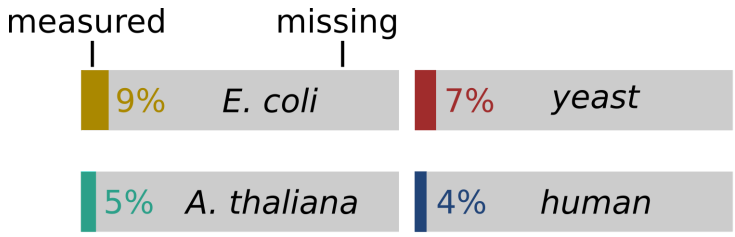

Figure 1: Measured turnover numbers (k_{cat}) cover only a small fraction of the metabolic reactions in each organism. The fraction of reactions with measured k_{cat} values in *E. coli*, *S. cerevisiae* (budding yeast), *A. thaliana* (Arabidopsis) and *H. sapiens* (human); measured k_{cat} values taken from [5]. The total numbers of currently known reactions are obtained from genome scale metabolic reconstructions and correspond to 2251 for *E. coli* [6], 1149 for budding yeast [12], 1363 for Arabidopsis [13] and 3673 for human [14]. Since

k_{cat} is defined per reaction direction and some reactions are physiologically reversible, the actual coverage is even lower.

The ubiquity of factors not accounted for in traditional *in vitro* assays

Most kinetic measurements in the literature are done *in vitro*, often in conditions that represent the cellular environment poorly [15]. A recent review by van Eunen and Bakker [16] discusses the differences between *in vivo* and *in vitro* conditions in enzyme assays, and shows that when using standardized *in vivo*-like media, enzyme kinetic values considerably deviate from reported *in vitro* measurements, casting doubt on the physiological relevance of some *in vitro* measurements [17,18].

Enzyme catalysis inside cells is affected by numerous factors. Several studies have quantified the reduction in net catalytic rates due to undersaturation of enzymes and backward flux caused by thermodynamic constraints [19–21]. Allosteric regulation of enzymes was also shown to have an important role in shaping enzyme catalysis *in vivo* [10,11,21]. However, allosteric regulators are often excluded from *in vitro* enzyme assays, since many are simply not known or the physiologically relevant concentration to use is unclear. Another key feature of enzyme catalysis *in vivo*, which usually cannot be mimicked in traditional assays, is the presence of covalent posttranslational modifications. Such regulatory mechanisms are long known in eukaryotic organisms and recent years have shown the ubiquity of covalent posttranslational modifications even in microbes [22].

Proteomics techniques now offer a unique opportunity to quantify covalent protein modifications on a large scale [23,24]. Schmidt *et al.* [25] and Licona-Cassani *et al.* [26] recently used phosphoproteomics [27] to show that phosphorylation of central metabolic enzymes is widespread across bacterial metabolism. Similar phosphorylation patterns have been observed in yeast, where about 20% of central metabolic enzymes were shown to change their degree of phosphorylation during growth under different conditions [28]. Methods like limited proteolysis coupled to MS and cross-linking MS are currently being used to look at conformational changes of enzymes and at assembly of enzymatic complexes *in vivo [29,30]*. As allosteric regulation is often manifested by conformational changes of enzymes, these studies will shed light on enzyme allostery. For a detailed review on the relation between MS proteomics and structural biology see [23]. Given the above, it is natural to ask: are *in vitro* parameters representative of what happens within cells? As we now proceed to describe, an ability to infer enzyme kinetics *in*

vivo with genome-wide coverage was recently demonstrated using metabolic flux predictions and mass spectrometry (MS) proteomics.

Inferring enzyme catalytic rates from fluxomics and proteomics

Proteomics techniques have revolutionized our ability to measure protein abundances inside cells on a genome-wide scale [31,32]. The rate of an enzymatic reaction can be obtained by dividing the flux through the enzyme by its abundance (flux per enzyme copy). Today, ^{13}C-labeling analysis and constraint based modeling methods offer quantification of metabolic fluxes [33,34]. Therefore, in order to calculate the *in vivo* catalytic rate of an enzyme, one could simply divide the flux through the enzyme (from flux analysis) by the abundance of that enzyme from quantitative proteomics (Figure 2). A rapidly expanding body of proteomic data currently reports on about 60 organisms under various growth conditions [35]. Flux estimates cover a large fraction of the metabolic network, at present mostly from constraint based modeling methods such as flux balance analysis but also increasingly using ^{13}C-labeling analysis [36]. Thus, the above approach holds great potential to infer *in vivo* enzymatic rates in many organisms.

Several caveats should be kept in mind when inferring enzymatic rates using omics data. First, inside cells, the mapping between fluxes and enzymes is not one-to-one. Some enzymes catalyze more than a single reaction and some reactions are catalyzed by more than a single enzyme [37–39]. Recent association studies are beginning to shed light on the broad range of non-specific enzymatic reactions and promises to further elucidate the mapping between enzymes and their associated *in vivo* functionalities [40,41]. Still, as reflected in metabolic scale reconstructions [42], the majority of enzymes are assigned to a single physiologically relevant reaction, and high specificity is thought to be crucial for the ordered metabolism of living organisms [43]. Second, flux measurements represent the net flux through the reaction (the difference between the forward and backward flux), and quantitative proteomics measurements represent the total amount of proteins and usually do not distinguish between active and inactive enzymes. Thus, by dividing the flux by the enzyme abundance, one does not get the actual rate of catalysis but the average net rate of the enzyme in the physiologically operative direction. Therefore, these rates are apparent rates of enzymes which are accordingly denoted as k_{app} values [44]. Whether and how it is possible to disentangle the kinetic constants k_{cat} and K_M from k_{app} is discussed below.

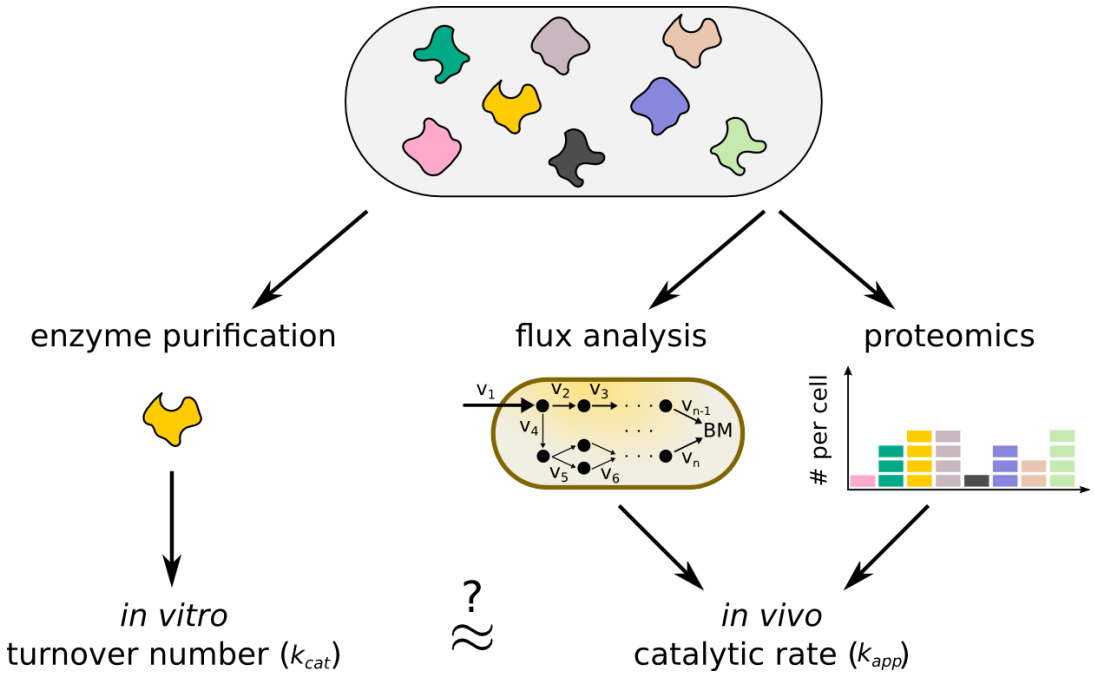

Figure 2: Using flux analysis and quantitative proteomics to calculate the rate of enzyme catalysis inside cells (k_{app}). Traditionally, *in vitro* enzyme assays are used to kinetically characterize enzymes. Such assays are labor intensive, involving purification of the enzyme of interest and a screen for favorable ambient conditions (pH, temperature, crowding agents, co-factors, etc.). To calculate the *in vivo* catalytic rate, denoted as k_{app}, one can divide the flux carried by the enzyme with the abundance of that enzyme to get flux per enzyme, i.e., the rate of the enzyme. Integration of quantitative proteomics and flux data enables such calculations for many enzymes simultaneously.

In vivo maximal rates recapitulate *in vitro* k_{cat} measurements

A recent study introduced a new parameter, k_{max}^{vivo}, which describes the maximum enzymatic rate *in vivo* [45]. To approximate k_{max}^{vivo}, the maximum k_{app} value across a large set of growth conditions was calculated. Comparison between *in vivo* based k_{max}^{vivo} values and *in vitro* k_{cat} values showed that *in vivo* and *in vitro* maximal rates generally concur - exhibiting a correlation of $r^2 = 0.62$ in log scale with a root mean square difference of 0.54. The prediction uncertainty is equivalent to about 3-fold compared to the >10-fold standard deviation of the k_{cat} distribution. The agreement between k_{max}^{vivo} and k_{cat} values highlights two important conclusions: (i) in most cases *in vitro* enzyme assays quite accurately reflect the maximal rates of enzyme *in vivo,* and

(ii) k_{max}^{vivo} values can be used as a method for the collection of maximal enzyme rates, thus providing a high-throughput avenue for ameliorating the acute data scarcity described above.

It appears that in the era of omics data, an alternative for laborious enzyme assays is to use fluxes and enzyme levels to infer enzymatic rates and to predict metabolic behaviors [46]. This approach, described in details in [45], most probably also reflects a more physiologically representative method for extracting rate constants of enzymes than the traditional *in vitro* assays. In some cases, such as enzymes with low expression levels, targeted proteomics at low throughput might be needed. A standing challenge is estimating K_M values *in vivo*. One possible approach would be to plot enzyme rates as a function of metabolite concentrations and infer K_M values. Yet this will require extra information regarding absolute metabolite concentrations, which is another current frontier of technological advancement. Such information will complement and expand our ability to accurately predict metabolic behavior within cells. We note that for most applications knowing k_{cat} is more crucial than knowing K_M values, because while the effect on enzyme rate from the saturation term ([47], based on K_M and substrate concentration) would very rarely give more than a 10-fold effect, the capacity term ([47], based on k_{cat}) often varies by more than 10-fold.

Defining the capacity utilization of metabolic reactions

Network-wide coverage of enzyme kinetic parameters can also provide lessons regarding the catalytic efficiency of cellular metabolism. The kinetic properties of enzymes, together with their mass fraction out of the proteome, represent the metabolic capacity of a cell. We can measure the catalytic efficiency of metabolism by the "capacity utilization" metric described below, with the hope of connecting a number of concepts described in recent years [45,48–51].

We start by defining the catalytic capacity of a reaction as $E \cdot k_{max}^{vivo}$, where E is the amount of the catalyzing enzyme. The capacity utilization of a reaction is defined as the ratio between the actual flux through the reaction, v, and the catalytic capacity (Figure 3A). We note that this definition is the inverse of a parameter termed the safety factor [52], which was previously defined as the ratio of the maximal reaction rate (at high substrate concentrations) to the reaction rate under actual physiological conditions. If the capacity utilization is zero, the associated enzyme is unutilized [49] - i.e., the enzyme is expressed but does not carry flux. Similarly, if the capacity utilization is 1 (i.e. 100%), the enzyme operates at its maximum rate, k_{max}^{vivo}, and therefore the enzyme is fully utilized (the safety factor is zero).

One can also think of this from the perspective of protein cost investment (e.g., see [51] and Goel *et al.* [50]). For such cost-associated nomenclature, it is beneficial to define E_{min} as the minimal amount of enzyme required to support the observed flux (v). E_{min} is realized when enzymes operate at their k_{max}^{vivo} and hence the ratio between E_{min} and the actual enzyme amount, E, equals the capacity utilization (Figure 3A). As reviewed below, the capacity utilization can be used to evaluate the extent to which cells exploit their enzymatic capabilities. A particular proteome resource allocation strategy can therefore be evaluated by examining the capacity utilization of all expressed enzymes. Interestingly, Noor *et al.* [51] and Goel *et al.* [50] raise conflicting observations regarding the tendency of bacterial cells to operate at high capacity utilization. We review these issues in the following sections.

To what extent do cells utilize their catalytic capacity across different conditions?

The proteome composition changes between conditions. O'Brien *et al.* [49] recently characterized the capacity utilization in *E. coli* across 16 different conditions, effectively harnessing the power of quantitative proteomics in the metabolic context. They compared protein levels measured by quantitative proteomics [53] to model-based predictions of the minimal protein demand for growth, i.e., E_{min} values. To compute E_{min}, O'Brien *et al.* used a genome-scale model (ME-model [54]), together with measured growth rates and assigned kinetic rates, which were randomly sampled from a global k_{cat} distribution [5]. The authors found that in most environments 50% of the proteome was unused, where this value includes the mass beyond the minimal requirement in underutilized enzymes as well as the mass invested in enzymes not supporting any flux under the given condition. Beyond the overall values reported by O'Brien *et al.* [49], it is of interest to portray the distribution of the capacity utilization of different enzymes in *E. coli*, as we present in Figure 3B. In order to calculate the capacity utilization, one requires a knowledge of the maximal potential rate. We chose to use k_{max}^{vivo} as the surrogate for maximal reaction rates *in vivo* (in contrast to using random sampling from the highly heterogenous distribution of k_{cat}). As can be observed in Figure 3B, for batch growth on either glucose or galactose, ≈30% of the enzymes (by mass rather than by number which we find less informative) are apparently idle - exhibiting a capacity utilization value below 10%. We note that this is based on flux analysis and some of these enzymes probably have a moonlighting activity [55]. For others, it could be that expression of unutilized enzymes is a result of cellular objectives other than minimizing enzyme costs, e.g., bet hedging [56].

Alternatively, such overexpression could simply be a result of wasteful metabolic regulation [49]. Finally, constraints imposed by the metabolic network may not allow all enzymes to operate at full capacity at the same time. For example, not all enzymes can be close to saturation concurrently [57]. These possibilities raise the question of how much of the observed underutilization is a result of each explanation - a question that is addressed for selected enzymes in [45].

O'Brien *et al.* [49] also showed that with the decrease in specific growth rate, the unused fraction of the proteome increased. They suggest that the growth rate is explained by the capacity utilization of the proteome. To support such a claim, one should be able to predict the fraction of unused proteome without using growth rate as input for calculating the flux values in a metabolic model.

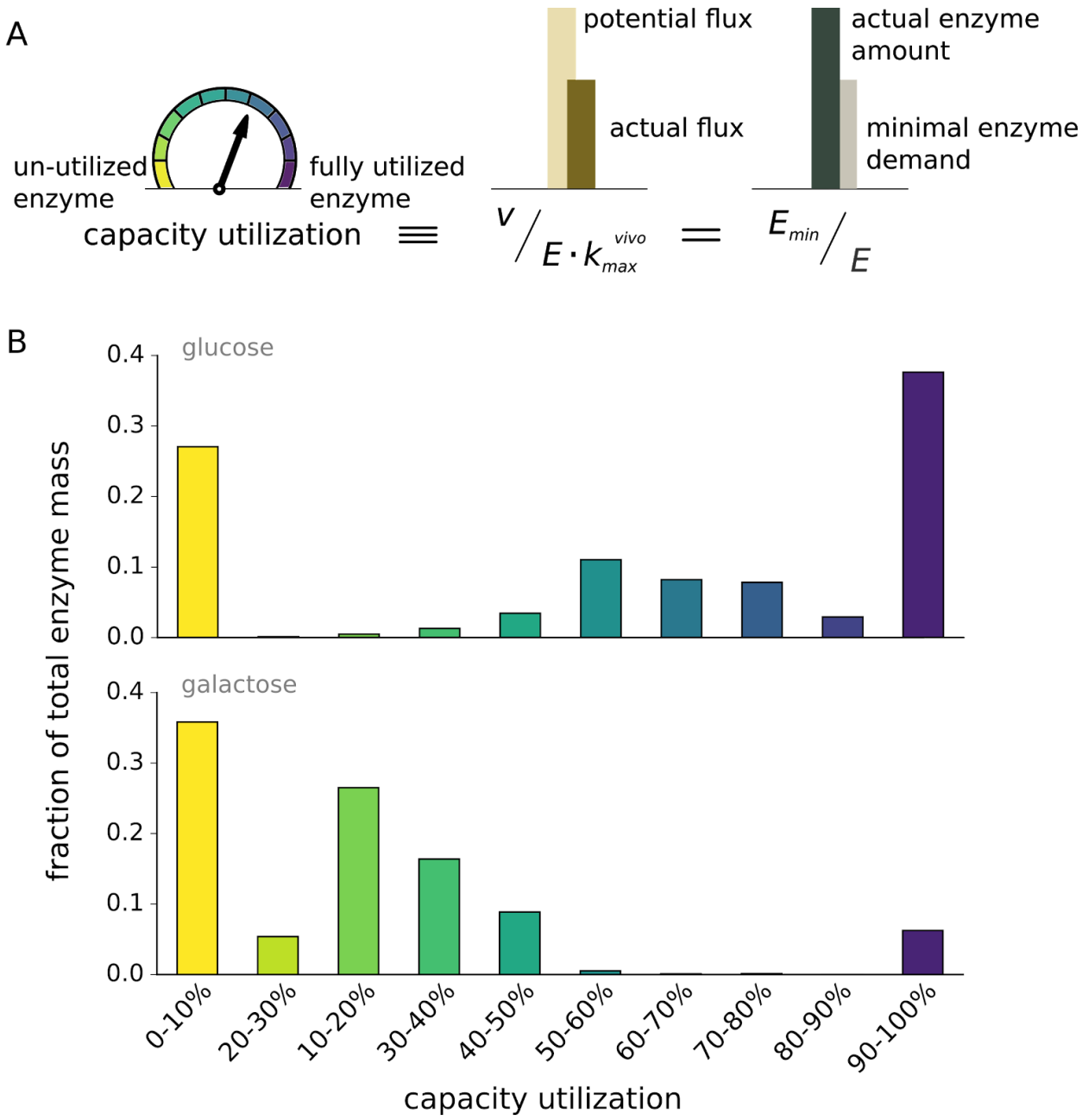

Figure 3: The capacity utilization of enzymes. A) The capacity utilization of a metabolic reaction is defined as the ratio between the actual flux supported by the reaction (v) and the maximal flux the reaction can potentially carry, given by $E \cdot k_{max}^{vivo}$. This ratio is equal to the ratio between the minimal enzyme requirement needed to support v (E_{min}) and the actual expressed enzyme level (E). The above equality holds since by definition v / k_{max}^{vivo} = E_{min}. B) The distribution of capacity utilization across all enzymes in *E. coli* during exponential growth on glucose (upper panel) and galactose (lower panel) as sole carbon sources. For calculating capacity utilization we used flux values from [19] and enzyme

abundance from [53]. Since flux measurements covered only central metabolic reactions, we used parsimonious Flux Balance Analysis (pFBA) [58] to expand flux estimates to the entire metabolic network.

Quantitative proteomics rationalizes usage of low yield metabolic strategies

In order to show the high relevance of metabolism to the resources economy of the cell, we show in Figure 4 the proteome composition of *E. coli* according to functional role across four growth conditions. As consistently observed in the figure, about half of the proteome mass is composed of enzymes (brown and yellow sectors). Thus, metabolism is a dominant factor when cells face the challenge of allocating their protein resources between different cellular functions - a challenge which can limit growth capacity, depending on how resources are distributed [59–64].

Noor *et al.* [51] recently suggested that *E. coli* operates to minimize overexpression of enzymes in its central metabolic network, given thermodynamic and kinetic limitations. Another recent work by Park *et al.* [48] provided evidence that in the glycolytic pathway of *E. coli* (and also in that of budding yeast and humans) most metabolite concentrations exceed the K_M values of enzymes, hence facilitating efficient enzyme activities. Efficient use of enzymes has also been shown to rationalize preference of low energy yield metabolic strategies. Following previous suggestions [44,54,65], Basan *et al.* [66] provided quantitative evidence that by performing aerobic respiration, *E. coli* was able to balance the proteome demand between energy biogenesis and biomass synthesis. The assumption that the proteome cost of energy of biogenesis by respiration exceeds that by fermentation was shown to agree with quantitative proteomics measurements [66]. A further refinement of the model presented by Basan *et al.* showed that molecular crowding constraints were important in explaining the switch from oxidative phosphorylation and overflow metabolism [67]. In fact, the presence of alternative glycolytic pathways in prokaryotes was also suggested to reflect a tradeoff between ATP yield and protein cost. Specifically, the protein cost of the canonical Embden-Meyerhof-Parnas (EMP) glycolytic pathway was shown to be higher than that of the Entner-Doudoroff (ED) pathway, whose ATP yield is half that of the EMP pathway [68].

Conversely, Goel *et al*. [50] recently observed that in *L. lactis,* across a 4-fold growth rates range, most glycolytic enzymes of *L. lactis* practically did not change in levels. That is, during slow growth there seems to be an excess of glycolytic enzymes which seems inefficient in terms of resource allocation. The authors suggest that the apparent lack of regulation on the levels of

glycolytic enzymes shows that enzyme levels are not under selection to minimize protein expression costs in *L. lactis*. It seems that the role of protein costs in shaping flux distribution is highly context dependent, for example changing between batch and chemostat growth or between transition states and adapted balanced growth states [61].

Figure 4: Proteomaps is a web-based tool for visualizing proteomics data, which provides a "snapshot" of how cells allocate their resources between different conditions [69]. Each polygon represents the mass fraction of a single protein and all proteins with similar functions are colored in shades of the same color. Here we visualized the proteome of *E. coli* under four different carbon source conditions: (1) glucose limited chemostat and batch growth on (2) galactose, (3) acetate or (4) glucose as sole carbon sources [53]. The specific growth rate (μ) in each condition is indicated above each map. The fraction allocated to metabolism represents about 50% of the proteome by mass across all conditions (shades of brown and

yellow). Labels indicate subsystems of metabolism, such as transport, glycolysis or the TCA cycle. Further explanation about proteomaps and how to use a web based tool to generate maps of your own data is available at http://www.proteomaps.net/.

Conclusion

MS proteomics increasingly supply a comprehensive picture of enzyme levels and states inside cells. As we aimed to highlight in this review, proteomics also provides a powerful tool for gaining kinetic parameters of enzymes *in vivo*. Proteomics-derived *in vivo* maximal rates both recapitulate *in vitro* k_{cat} measurements and offer a physiologically relevant approach for the collection of enzyme parameters. While further progress should be made to infer in vivo K_M values, in vivo k_{max}^{vivo} values can already be integrated into metabolic models and, given proteomics data, serve as upper bounds on reaction rates. This can be useful for synthetic biology applications - to predict potential cell growth and biosynthetic products production capacity and to improve the predictive power of resource allocation models, for example. Furthermore, as recent advances in MS-proteomics are now used to characterize protein modifications and structures *in vivo*, it may soon be possible to relate specific posttranslational regulation events to enzyme activities. In conclusion, the omics era promises to finally satisfy the long held desire of a more realistic understanding of *in vivo* enzyme kinetics.

Acknowledgements

The authors thank Elad Noor, Arren Bar-Even, Uri Barenholz, Avi Flamholz, Yinon Bar-On, Wolfram Liebermeister, Nathan Lewis, Karl Kochanowski, Noam Prywes, David Wernick, Alexander Geller and Lianet Noda for productive discussions and critical reading of the manuscript. This research was funded by the European Research Council (Project NOVCARBFIX 646827); the Israel Science Foundation (grant No. 740/16); Beck-Canadian Center for Alternative Energy Research; Dana and Yossie Hollander; Helmsley Charitable Foundation; The Larson Charitable Foundation; Wolfson Family Charitable Trust; Charles Rothschild; Selmo Nussenbaum. R.M. is the Charles and Louise Gartner professional chair.

References

* special interest

** outstanding interests

*1. Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H, Herwig R: ***Systems Biology: A Textbook***. John Wiley & Sons; 2009.

The authors provide a remarkable textbook for systems biology, covering computational and experimental approaches in systems biology, including introduction to metabolic modeling and hands-on experience. It is tailored for advanced students of biology, engineering, and computer science.

2. Michaelis L, Menten ML: **Die Kinetik der invertinwirkung**. *Biochem. Z.* 1913, **49**:352.

3. Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N, Dunn WB, Murabito E, Swainston N, Dada JO, Khan F, et al.: **A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes**. *FEBS Lett.* 2013, **587**:2832–2841.

4. Gautam A, Rishi P, Tewari R: **UDP-N-acetylglucosamine enolpyruvyl transferase as a potential target for antibacterial chemotherapy: recent developments**. *Appl. Microbiol. Biotechnol.* 2011, **92**:211–225.

*5. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R: **The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters.** *Biochemistry* 2011, **50**:4402–4410.

The authors performed global analysis of the kinetic parameters of enzymes, with focus on physiologically relevant reactions. The study quantitatively characterizes the wide dynamic rage of k_{cat} and K_M. The authors analyzed systematic driving forces shaping rate constants.

6. Orth JD, Conrad TM, Na J, Lerman J a., Nam H, Feist AM, Palsson BØ: **A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011.** *Mol. Syst. Biol.* 2011, **7**:535.

7. Liebermeister W, Klipp E: **Biochemical networks with uncertain parameters**. *IEE Proceedings-Systems Biology* 2005, **152**:97–107.

8. Lubitz T, Schulz M, Klipp E, Liebermeister W: **Parameter Balancing in Kinetic Models of**

Cell Metabolism. *J. Phys. Chem. B* 2010, **114**:16298–16303.

9. Van Eunen K, Bouwman J, Daran-Lapujade P, Postmus J, Canelas AB, Mensonides FIC, Orij R, Tuzun I, Van Den Brink J, Smits GJ, et al.: **Measuring enzyme activities under standardized in vivo-like conditions for systems biology**. *FEBS J.* 2010, **277**:749–760.

10. Goodey NM, Benkovic SJ: **Allosteric regulation and catalysis emerge via a common route.** *Nat. Chem. Biol.* 2008, **4**:474–482.

11. Kochanowski K, Sauer U, Noor E: **Posttranslational regulation of microbial metabolism**. *Curr. Opin. Microbiol.* 2015, **27**:10–17.

12. Duarte NC, Herrgård MJ, Palsson BØ: **Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model**. *Genome Res.* 2004, **14**:1298–1309.

13. Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T: **Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity**. *Proc. Natl. Acad. Sci. U. S. A.* 2012, **109**:339–344.

14. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, Srivas R, Palsson BØ: **Global reconstruction of the human metabolic network based on genomic and bibliomic data**. *Proc. Natl. Acad. Sci. U. S. A.* 2007, **104**:1777–1782.

15. Ringe D, Petsko GA: **How enzymes work**. *SCIENCE-NEW YORK THEN WASHINGTON-* 2008, **320**:1428.

16. van Eunen K, Bakker BM: **The importance and challenges of in vivo-like enzyme kinetics**. *Perspectives in Science* 2014/5, **1**:126–130.

17. Wright BE, Butler MH, Albe KR: **Systems analysis of the tricarboxylic acid cycle in Dictyostelium discoideum. I. The basis for model construction**. *J. Biol. Chem.* 1992, **267**:3101–3105.

18. Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, et al.: **Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry**. *Eur. J. Biochem.* 2000, **267**:5313–5329.

19. Gerosa L, van Rijsewijk BRBH, Christodoulou D, Kochanowski K, Schmidt TSB, Noor E, Sauer U: **Pseudo-transition Analysis Identifies the Key Regulators of Dynamic Metabolic Adaptations from Steady-State Data**. *cels* 2015, **1**:270–282.

20. Noor E, Bar-Even A, Flamholz A, Reznik E, Liebermeister W, Milo R: **Pathway Thermodynamics Highlights Kinetic Obstacles in Central Metabolism**. *PLoS Comput. Biol.* 2014, **10**.

21. Chubukov V, Gerosa L, Kochanowski K, Sauer U: **Coordination of microbial metabolism.** *Nat. Rev. Microbiol.* 2014, **12**:327–340.

22. Pisithkul T, Patel NM, Amador-Noguez D: **Post-translational modifications as key regulators of bacterial metabolic fluxes**. *Curr. Opin. Microbiol.* 2015, **24**:29–37.

23. Faini M, Stengel F, Aebersold R: **The Evolving Contribution of Mass Spectrometry to Integrative Structural Biology**. *J. Am. Soc. Mass Spectrom.* 2016, **27**:966–974.

24. Boja ES, Rodriguez H: **Mass spectrometry-based targeted quantitative proteomics: achieving sensitive and reproducible detection of proteins**. *Proteomics* 2012, **12**:1093–1110.

25. Schmidt A, Trentini DB, Spiess S, Fuhrmann J, Ammerer G, Mechtler K, Clausen T: **Quantitative phosphoproteomics reveals the role of protein arginine phosphorylation in the bacterial stress response**. *Mol. Cell. Proteomics* 2014, **13**:537–550.

26. Licona-Cassani C, Lim S, Marcellin E, Nielsen LK: **Temporal dynamics of the Saccharopolyspora erythraea phosphoproteome**. *Mol. Cell. Proteomics* 2014, **13**:1219–1230.

27. Picotti P: **Phosphoproteomics takes it easy**. *Nat. Biotechnol.* 2015, **33**:929–930.

28. Oliveira AP, Ludwig C, Picotti P, Kogadeeva M, Aebersold R, Sauer U: **Regulation of yeast central metabolism by enzyme phosphorylation**. *Mol. Syst. Biol.* 2012, **8**:623.

29. Fontana A, de Laureto PP, Spolaore B, Frare E, Picotti P, Zambonin M: **Probing protein structure by limited proteolysis**. *Acta Biochim. Pol.* 2004, **51**:299–321.

30. Sinz A: **The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks**. *Expert Rev. Proteomics* 2014, **11**:733–743.

31. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B: **Quantitative mass spectrometry in proteomics: a critical review**. *Anal. Bioanal. Chem.* 2007, **389**:1017–1031.

32. Nikolov M, Schmidt C, Urlaub H: **Quantitative mass spectrometry-based proteomics: an overview**. *Methods Mol. Biol.* 2012, **893**:85–100.

33. Sauer U: **Metabolic networks in motion: 13C-based flux analysis**. *Mol. Syst. Biol.* 2006, **2**:62.

34. Orth JD, Thiele I, Palsson BØ: **What is flux balance analysis?** *Nat. Biotechnol.* 2010, **28**:245–248.

35. Wang M, Weiss M, Simonovic M, Haertinger G, Schrimpf SP, Hengartner MO, von Mering C: **PaxDb, a database of protein abundance averages across all three domains of life**. *Mol. Cell. Proteomics* 2012, **11**:492–500.

36. Niedenführ S, Wiechert W, Nöh K: **How to measure metabolic fluxes: a taxonomic guide for 13C fluxomics**. *Curr. Opin. Biotechnol.* 2015/8, **34**:82–90.

37. Weng J-K, Philippe RN, Noel JP: **The rise of chemodiversity in plants**. *Science* 2012,

336:1667–1670.

38. Weng J-K, Noel JP: **The remarkable pliability and promiscuity of specialized metabolism**. *Cold Spring Harb. Symp. Quant. Biol.* 2012, **77**:309–320.

39. Notebaart RA, Szappanos B, Kintses B, Pál F, Györkei Á, Bogos B, Lázár V, Spohn R, Csörgő B, Wagner A, et al.: **Network-level architecture and the evolutionary potential of underground metabolism**. *Proc. Natl. Acad. Sci. U. S. A.* 2014, **111**:11762–11767.

40. Fuhrer T, Zampieri M, Sévin DC, Sauer U, Zamboni N: **Genomewide landscape of gene–metabolome associations in Escherichia coli**. *Mol. Syst. Biol.* 2017, **13**:907.

41. Sévin DC, Fuhrer T, Zamboni N, Sauer U: **Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in Escherichia coli**. *Nat. Methods* 2016, doi:10.1038/nmeth.4103.

42. Ebrahim A, Palsson BO, Lewis NE: **BiGG Models: A platform for integrating, standardizing and sharing genome-scale models**. *Nucleic acids* 2016, [no volume].

43. Dixon M: **EC Webb, Enzymes**. *Academic Press Inc. , New York, NY, 2nd ed. . I9f* 1964, **7**:647.

44. Valgepea K, Adamberg K, Seiman A, Vilu R: **Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins. *Mol. Biosyst.* 2013, **9**:2344–2358.

Analysis of genome-wide intracellular mRNA, protein concentrations and metabolic flux analysis in *E. coli* across a fivefold range of growth rates. Authors used absolute proteomics and flux measurements to determine the catalytic rate of enzymes *in vivo*, k_{app}, and showed that at fast growth rates many enzymes operate significantly faster than at slow growth rates.

45. Davidi D, Noor E, Liebermeister W, Bar-Even A, Flamholz A, Tummler K, Barenholz U, Goldenfeld M, Shlomi T, Milo R: **Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro kcat measurements. *Proc. Natl. Acad. Sci. U. S. A.* 2016, doi:10.1073/pnas.1514240113.

This work presents a methodology for deriving the maximal catalytic rates of enzymes inside cells from quantitative proteomics and flux data. The authors compared maximal *in vivo* rates to measured k_{cat} values and found good agreement. The presented approach can be used to accelerate the collection of rate constants of enzymes, which are more physiologically relevant.

46. Kotte O, Heinemann M: **A divide-and-conquer approach to analyze underdetermined biochemical models**. *Bioinformatics* 2009, **25**:519–525.

47. Noor E, Flamholz A, Liebermeister W, Bar-Even A, Milo R: **A note on the kinetics of**

enzyme action: A decomposition that highlights thermodynamic effects. *FEBS Lett.* 2013, **587**:2772–2777.

48. Park JO, Rubin SA, Xu Y-F, Amador-Noguez D, Fan J, Shlomi T, Rabinowitz JD: **Metabolite concentrations, fluxes and free energies imply efficient enzyme usage**. *Nat. Chem. Biol.* 2016, doi:10.1038/nchembio.2077.

49. O'Brien EJ, Utrilla J, Palsson BO: **Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments. *PLoS Comput. Biol.* 2016, **12**:e1004998.

The authors compared protein expression data from quantitative proteomics with model based predictions of minimal enzyme demands. They report that nearly half of the proteome mass is unused in certain environments. Accounting for the cost of this unused protein expression is suggested to explain >95% of the variance in growth rates of *Escherichia coli* across 16 distinct environments.

*50. Goel A, Eckhardt TH, Puri P, de Jong A, Branco Dos Santos F, Giera M, Fusetti F, de Vos WM, Kok J, Poolman B, et al.: **Protein costs do not explain evolution of metabolic strategies and regulation of ribosomal content: does protein investment explain an anaerobic bacterial Crabtree effect?** *Mol. Microbiol.* 2015, **97**:77–92.

In this work the authors studied growth rate-dependent relative transcription and protein ratios, enzyme activities, and fluxes of *L. lactis* in glucose-limited chemostats. They show that enzyme levels in central metabolism of *L. lactis* hardly change across a 4-fold range of growth rates.

*51. Noor E, Flamholz A, Bar-Even A, Davidi D, Milo R, Liebermeister W: **The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization**. *PLoS Comput. Biol.* 2016, **12**:e1005167.

The authors develop a scalable and robust framework which utilizes a full kinetic model and optimizes enzyme costs. They show that minimization of enzyme cost is a meaningful optimality principle for exponentially growing *E. coli* by showing that the framework is able to predict the measured enzyme abundances in central metabolism.

52. Weiss SL, Lee EA, Diamond J: **Evolutionary matches of enzyme and transporter capacities to dietary substrate loads in the intestinal brush border**. *Proc. Natl. Acad. Sci. U. S. A.* 1998, **95**:2117–2121.

53. Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heinemann M: **The quantitative and condition-dependent

Escherichia coli proteome. *Nat. Biotechnol.* 2015, doi:10.1038/nbt.3418.

The largest available quantitative proteomics study to date. This work provides genome scale protein quantification in *E. coli* across more than twenty different growth conditions, including chemostat growth, batch growth on different carbon sources and stress conditions.

54. O'Brien EJ, Lerman J a., Chang RL, Hyduke DR, Palsson BØ: **Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction**. *Mol. Syst. Biol.* 2013, **9**:693.

55. Mani M, Chen C, Amblee V, Liu H, Mathur T, Zwicke G, Zabad S, Patel B, Thakkar J, Jeffery CJ: **MoonProt: a database for proteins that are known to moonlight.** *Nucleic Acids Res.* 2015, **43**:D277–82.

56. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U: **Multidimensional optimality of microbial metabolism.** *Science* 2012, **336**:601–604.

57. Fell D, Cornish-Bowden A: *Understanding the control of metabolism*. Portland press London; 1997.

58. Lewis NE, Hixson KK, Conrad TM, Lerman J a., Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO, Lopez-Ferrer D, et al.: **Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models.** *Mol. Syst. Biol.* 2010, **6**:390.

59. Kafri M, Metzl-Raz E, Jona G, Barkai N: **The Cost of Protein Production**. *Cell Rep.* 2016, **14**:22–31.

60. Dekel E, Alon U: **Optimality and evolutionary tuning of the expression level of a protein**. *Nature* 2005, **436**:588–592.

61. Shachrai I, Zaslaver A, Alon U, Dekel E: **Cost of unneeded proteins in E. coli is reduced after several generations in exponential growth**. *Mol. Cell* 2010, **38**:758–767.

62. Dong H, Nilsson L, Kurland CG: **Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction**. *J. Bacteriol.* 1995, **177**:1497–1504.

63. Dean AM, Dykhuizen DE, Hartl DL: **Fitness as a function of beta-galactosidase activity in Escherichia coli**. *Genet. Res.* 1986, **48**:1–8.

64. Barenholz U, Keren L, Segal E, Milo R: **A Minimalistic Resource Allocation Model to Explain Ubiquitous Increase in Protein Expression with Growth Rate**. *PLoS One* 2016, **11**:e0153344.

65. Molenaar D, van Berlo R, de Ridder D, Teusink B: **Shifts in growth strategies reflect tradeoffs in cellular economics**. *Mol. Syst. Biol.* 2009, **5**:323.

*66. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T: **Overflow metabolism in Escherichia coli results from efficient proteome allocation**. *Nature* 2015, **528**:99–104.

The authors constructed a theoretical model which suggested that oxidative fermentation is more efficient in terms of proteome resource allocation, and showed that quantitative proteomics data supports the model predictions.

67. Vazquez A, Oltvai ZN: **Macromolecular crowding explains overflow metabolism in cells**. *Sci. Rep.* 2016, **6**:31007.

68. Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo R: **Glycolytic strategy as a tradeoff between energy yield and protein cost.** *Proc. Natl. Acad. Sci. U. S. A.* 2013, **110**:10039–10044.

69. Liebermeister W, Noor E, Flamholz A, Davidi D, Bernhardt J, Milo R: **Visual account of protein investment in cellular functions**. *Proc. Natl. Acad. Sci. U. S. A.* 2014, **111**:8488–8493.