Metabolism as a signal generator in bacteria

Review Article

Author(s):
Ledezma Tejeida, Daniela Elizabeth; Schastnaya, Evgeniya; Sauer, Uwe

Publication date:
2021-12

Permanent link:
https://doi.org/10.3929/ethz-b-000517650

Originally published in:
Current Opinion in Systems Biology 28, https://doi.org/10.1016/j.coisb.2021.100404

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Systems Biology

Metabolism as a signal generator in bacteria

Daniela Ledezma-Tejeida[1], Evgeniya Schastnaya[1,2] and Uwe Sauer[1]

Abstract

Bacteria constantly monitor their environment to adapt their inner makeup. Beyond providing chemical sustenance, metabolism provides most of the feedback on the cellular environment via metabolite binding to regulatory proteins or mRNA. Although first metabolite-protein interactions were discovered more than 60 years ago, identification of new interactions is still technically challenging and time-consuming. Here, we compiled and quantified the current knowledge on metabolite-protein interactions and review recent advances in the identification of interactions and in understanding how metabolites act as signals to transcription factors, two-component systems, protein kinases, and riboswitches. New systematic methods of metabolite-protein identification and omics integration will accelerate the pace of discovery, a remaining challenge is understanding of functionality and the coordination of local and global metabolic signals across different regulatory layers.

Addresses

[1] Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

[2] Life Science Zurich PhD Program on Systems Biology, Zurich, Switzerland

Corresponding author: Sauer, Uwe (sauer@imsb.biol.ethz.ch)

Current Opinion in Systems Biology 2021, **28:**100404

This review comes from a themed issue on **Metabolic Networks (2022)**

Edited by **Sarah-Maria Fendt** and **Markus Ralser**

For complete overview of the section, please refer the article collection - Metabolic Networks (2022)

Available online 28 October 2021

https://doi.org/10.1016/j.coisb.2021.100404

Keywords

Regulation, Metabolism, Transcription factors, Kinases, Signaling.

Introduction

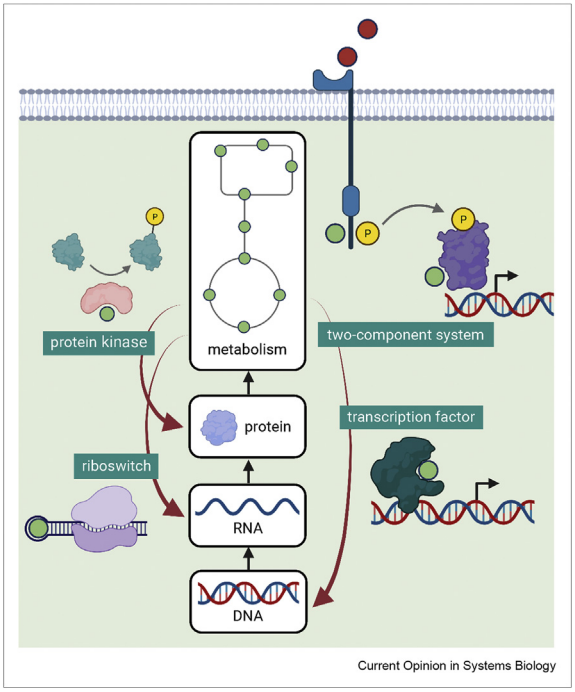
From a microbial perspective, the world is a fiercely competitive place with continuously changing conditions. Reacting in time is therefore essential for microbes to maintain their highly dynamic states and rapidly adapt to environmental cues. Various levels of intertwined regulation act on different time scales, from the slower transcriptional regulation to faster translational or post-translational regulation driven by dedicated regulatory proteins such as transcription factors or protein kinases. These regulators respond to various input signals that include binding to other proteins or small molecules [1]. Because there are many more conditions that need to be sensed than there are regulatory proteins, many input signals are intermediates of cellular metabolism that integrate a wide range of environmental perturbations into a smaller set of defined intracellular signals [2–6]. Such internal metabolic signals can trigger transcriptional, translational, or post-translational regulation, but the underlying metabolite-protein interactions are fleeting in nature and therefore difficult to measure experimentally [5,7,8]. Consequently, we do not know the input signals for many regulators, and those that we do know are typically not well understood in quantitative terms.

Here, we review recent progress in identifying and understanding metabolic input signals into microbial regulation systems. Because direct allosteric regulation of metabolic enzymes through small molecules was recently reviewed [9], we focus on small-molecule binding to transcription factors (TFs), two-component systems (TCSs), protein kinases and other regulatory proteins, and riboswitches — highlighting the role of intracellular metabolism as a signal generator (Figure 1).

Transcription factors

With the arguably best-studied transcriptional regulatory network, *Escherichia coli* has more than 300 predicted TFs in its genome [10], 75% of which encode a small-molecule-binding domain [11] whose ligands may be metabolic intermediates. By mining the specialized databases RegulonDB [12], EcoCyc [13], and the literature, we identified 124 effectors with experimental evidence of binding 93 TFs, a number way below the expected 75% of TFs potentially binding small molecules. Most evidence for the known effector-TF interactions is based on *in vitro* studies, and the low number of identified effectors can be attributed to the time-consuming nature of these techniques. Recent examples are the identification of 2-oxoglutarate binding to NifA in *Herbaspirillum seropedicae* [14], arachidonic acid binding to FadR in enterohemorrhagic *E. coli* [15], and L-ascorbate and α-D-galacturonate binding to PlaR (previously YiaJ) in *E. coli* K12 [16].

Figure 1



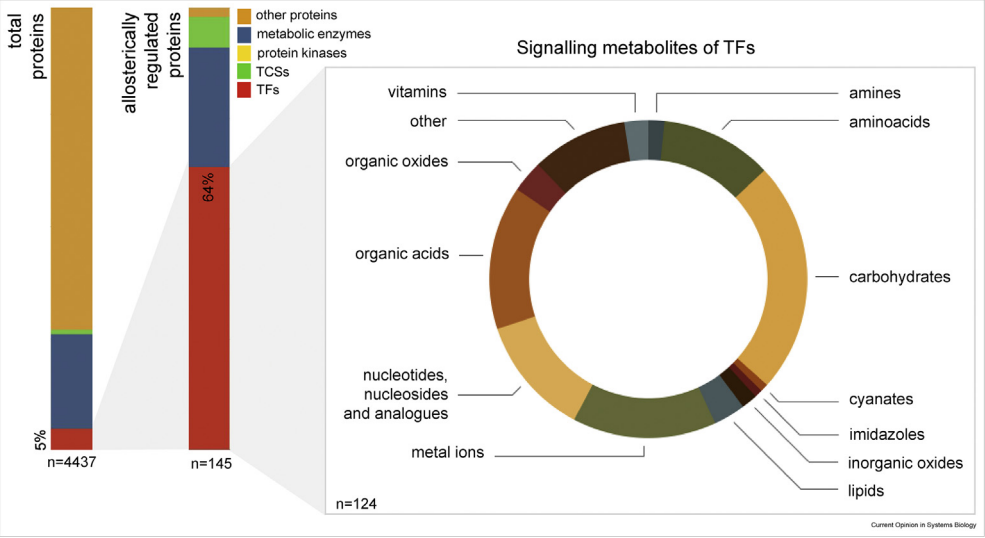
Metabolites as signals. Metabolism provides signaling inputs into various cellular regulatory systems that in turn have an effect on the metabolic composition of the cell, thus maintaining homeostasis.

The first systematic approach to identify metabolite-TF interactions combined transcriptomics and metabolomics data from *E. coli* shifted between growth and starvation [17]. Functional metabolite-TF interactions were identified by inferring the activity of 209 TFs through network component analysis [18] and comparing it to the abundances of 123 metabolites to pinpoint the metabolite-TF pairs that approximated Hill kinetics. Twelve out of 26 known metabolite-TF interactions were recovered, and 39 were newly predicted for 30 TFs, three of which were further validated *in vitro*. Although this discovery method likely has a high false-positive rate, and predictions still require validation, combining *in vivo* and *in vitro* evidence helped not only to identify interactions but also demonstrated their functional relevance in the cell. Key to learning causal interactions was the use of dynamic data as was also recently demonstrated for TF-gene relationships [19], because molecules rapidly responding to a perturbation are more likely to be directly related than slower responders that often result from indirect consequences. With the vast majority of gene expression studies focusing on the network below the TFs, the progress of effector-TF identification has so far been slow albeit steady. Approaches like the abovementioned hold promise for making systematic identification of metabolite signals more common and significantly increasing the pace of discovery, thereby putting us on a road to eventually fathom the entire process from signal recognition to the regulated processes, which in turn will often affect the input signals again.

Despite these challenges, the compendium of effector interactions is much larger for TFs than for other types of proteins; for example, TFs represent only 5% of *E. coli*'s encoded proteins but constitute ~60% of the proteins with experimental evidence for allosteric regulation by small molecules (Figure 2, Supplementary file 1). Carbohydrates, metal ions, and organic acids account for half of the known effectors, and nucleosides, nucleotides, and amino acids make up an additional 23%. More than 60% of the TFs have only one known effector, and the vast majority of effectors bind only one TF, with zinc as the exception of binding four regulators: ZntR; Zur; NrdR; and NikR. The lack of systematic approaches to identify effector-TF interactions makes it impossible to assess whether the high percentage of TFs binding only one effector represents a general principle or rather reflects our limited knowledge.

Once an interaction has been identified, the question that arises is how does effector binding impacts gene expression? For the most part, effector binding alters the TF conformation, and consequently the effector-TF complex associates with or dissociates from DNA [15,20–22]. An alternative mechanism is a TF whose conformation is altered in the presence of its effector but does not dissociate from DNA. For example, structural studies demonstrated that OxyR from *Corynebacterium glutamicum* binds to DNA through two-binding sites, and interaction with H_2O_2 promotes the release of only one of them, which is enough to allow polymerase binding and activation of gene expression [23]. The relationship between effector concentration and gene expression has been commonly described by a Hill equation [24], which is in a fragile equilibrium. A recent study of 10^5 random mutations on the *E. coli* TF LacI demonstrated that substitutions within 7 Å of the ligand-binding pocket have a 67% probability of turning the LacI repressor into an activator in the presence of its inducer [25]. Moreover, only two substitutions are enough to alter the wild-type sigmoidal shape of the Hill equation. Instead of the inducer concentration correlating linearly with gene expression for a continuous range of concentrations, linear correlations appeared at low and high inducer concentrations but not in middle ranges, suggesting that bound monomeric and fully-bound dimeric LacI had regulatory activity but not dimeric LacI with only one inducer molecule bound. These results highlight the flexibility of the effector-TF-gene interaction that enables rapid evolutionary innovation with most transcriptional research focusing on TFs and their downstream targets, it is often overlooked that most gene expression changes depend on the cell's rate of growth, so-called global regulation, in *E. coli* affecting about 85% of the promoters [26] and explaining about 70% of the measured expression

Figure 2



E. coli proteins with experimental evidence of allosteric regulation compared with all encoded proteins in *E. coli*. Although transcription factors (TFs) represent only 5% of total proteins, they account for 64% of known metabolite-protein interactions. Subpanel shows the chemical classes of 124 TF-binding effectors. Total proteins were extracted from EcoCyc [13], metabolic proteins were extracted from the study reported by Orth et al. [80], TFs and two-component systems (TCSs) (from RegulonDB, protein kinase classification was extracted manually from the literature. Chemical classification of metabolites was obtained from the study reported by Wishart et al. [81]. All information is available on supplementary file 1.

changes in central metabolism [27]. For any type of quantitative understanding of gene expression, the interplay between global regulation and TFs must be considered. One such example was demonstrated for the specific case of arginine biosynthesis in *E. coli*, where binding of the effector arginine to the repressor ArgR is responsible for the on/off switch of gene expression, but the magnitude of the expression during the on phase is dependent on global regulation [28]. ArgR belongs to the class of best-studied TFs where the presence of an amino acid or nutrient induces its transport and represses its biosynthesis or induces its catabolism [29]. Although the functional relevance of these TFs is easy to explain, quantitative understanding of their regulatory dynamics has been so far achieved in only few cases. A recent example is a theoretical analysis of XapR from *E. coli* [30], demonstrating that its on/off switch behavior depends on the presence of two-binding sites upstream of the genes coding for the transporter and catabolizer of its effector xanthosine. Achieving such understanding is even harder for more complex regulons such as those controlled by CRP or Cra that jointly regulate 555 genes in *E. coli*. Here, it is already difficult to say how many signals and regulated genes are necessary to orchestrate a particular adaptation and how this is coordinated. For *E. coli* central metabolism, a combination of reporter genes, metabolomics, and a mathematical model of metabolite-TF interactions demonstrated that 70% of the gene expression changes can be explained by global regulation and only two signal inputs, that is, binding of cAMP and fructose 1,6-bisphosphate to CRP and Cra, respectively [27]. The unexpected negative effect of catabolism promoting cAMP-CRP activity on anabolic promoters that were not CRP targets [27] was later explained by an indirect mechanism that may be the competition for limited capacity of the expression machinery [32]. The cAMP-CRP signal appears to mediate a total glycolytic flux feedback that is also one key driver of sequential consumption of certain carbon substrates [33]. Collectively, these findings support a model where few metabolites mediate global, heuristic responses that compensate for many possible perturbations at the same time, as opposed to evolving a precise regulatory circuit in response to each environmental challenge that *E. coli* may encounter. Thus, first principles are being unraveled for smaller subsystems, but we struggle with quantitative understanding of signals and expression coordination at the whole genome level.

Even less studied is signal integration when multiple effectors act on the same TF. Existing evidence comes almost exclusively from structural studies of TFs bound to their effectors where it was observed that different effectors promote different DNA affinities. For example, *Helicobacter pylori*'s NikR structure bound to nickel and zinc showed that the dynamic range of conformational states of the TF decreases in the presence of nickel, favoring a DNA-bound state [34]. The presence of zinc promotes an unstable interaction with DNA, consistent with the previously observed inability

of the NikR–zinc complex to induce gene expression [35]. In the absence of systematic studies on multi-effector TFs, we do not even know whether effectors act competitively, additively, or synergistically. Addressing these questions is challenging because they require *in vivo* evidence of the effector impact on the TF regulatory activity. The most direct evidence that can be obtained nowadays is based on correlations between effector abundance and promoter activity of the TF-regulated genes. Although the current paradigm of TF activity assumes that all target genes are regulated at the same time, as is exemplified by methods to infer TF targets or activity through transcriptomics data [18,36], it has been observed that only 25% of TFs regulate genes involved in the same biological process [37]. Hence, effectors might be adding a layer of coordination by selectively activating a subset of regulon genes under very specific conditions.

Two-component systems

Bacteria can sense environmental changes through so-called TCSs, consisting of a membrane-bound sensor histidine kinase that phosphorylates a downstream response regulator, typically acting as a TF [38]. Input signals are extracellular or intracellular, propagating either through direct interaction with the kinase or indirectly via accessory proteins that respond to signals [1]. In model bacteria, the induction signals are frequently known or linked to general environmental conditions, such as pH, osmotic pressure, or nutritional deprivation, but for most bacteria, the specific signaling molecules and the sensing mechanisms remain largely uncharacterized [39]. Here, we summarize recent work on deciphering metabolic effectors of TCS activity.

Many studies focused on intestinal pathogens and their interaction with the host. The first butyrate-sensing bacterial TCS was identified in *Campylobacter jejuni*, enabling it to spatially orient in the gut by monitoring metabolic signals and preferentially infect the lower intestine via BumSR that senses butyrate indirectly and promotes the expression of colonization factors [40]. For *Salmonella enterica* serovar Typhimurium, a mechanism of *lldPRD* operon regulation induced by host-derived L-lactate and oxygen was proposed [41]. In *Staphylococcus aureus*, ArlRS was hypothesized to be activated by increased levels of 3-phosphoglycerate or 1,3-bisphosphoglycerate [42]. Several studies focused on the ubiquitous bacterial second messenger cyclic-di-guanosine-5′-monophosphate (cyclic-di-GMP). The transition between virulence and swimming motility of phytopathogenic bacterium *Xanthomonas campestris* has been demonstrated to be allosterically regulated by cyclic-di-GMP binding to a sensor kinase, RavS, which enhances phosphorylation of response regulator RavR [43]. In *Mycobacterium smegmatis*, cyclic-di-GMP mediates adaptation to low-nutrient environments by binding to the sensor kinase PdtaS [44] and promotes oxidative stress tolerance by directly binding to the response regulator DevR, thereby enhancing its phosphorylation through the sensor kinase DevS [45]. For *E. coli*, it has been shown that depending on its concentration pyruvate activates either the low-affinity PyrSR or high-affinity BtsSR, controlling different sets of target genes [46,47]. Overall, these new insights highlight the importance of metabolite-mediated TCS signaling in pathogenesis. Given the diversity of bacterial TCSs and their activating stimuli, the current knowledge of specific metabolites sensed by TCSs is probably just the tip of the iceberg, and we can expect many more metabolic ligands to be discovered in the future.

Although TCS response regulators are generally considered to be regulated by their upstream kinase, they can also be activated by alternative kinases or directly by acetyl phosphate or inhibited by interaction with proteins or metabolites [48]. As suggested previously for other response regulators, recent findings in *Mycobacterium tuberculosis* show that acetyl phosphate can nonenzymatically phosphorylate the response regulator DevR and acetylate MtrA, both influencing the pathogenicity-related regulons of dormancy [49,50]. Surprisingly, acetylation overrides the phosphorylation of MtrA through its sensor kinase MtrB, suggesting that the metabolic state of the cell, as reflected by intracellular acetyl-phosphate levels, dominates over the extracellular stimulus transmission. Cumulative evidence for activation of numerous response regulators by acetyl phosphate across various bacteria further strengthens the role of metabolism in modulating TCS signal transduction, independent of the sensor kinases.

Protein kinases

Beyond the primarily transcription-regulating TCSs, bacteria also contain classical kinases that can post-translationally modify many different proteins. Recent phosphoproteomic experiments revealed phosphorylation of about 25% of all proteins in *E. coli*, suggesting that phosphoregulation might be more important in bacteria than previously considered [9,51]. Different from TFs, the input signals to most kinases are unknown even in the better characterized eukaryotic kinase networks [52], let alone in bacteria. Because the sensor kinases of bacterial TCSs and several eukaryotic kinases such as PKA, AMPK, and mTORC1 sense metabolites [53,54], it is tempting to speculate that at least some of the bacterial protein kinases also respond to metabolic signals. Indeed, several recent studies unraveled metabolic inputs into the modulation of kinase-mediated processes in bacteria. Proteome-wide metabolite binding identified fructose-1,6-bisphosphate as an effector of the bifunctional kinase/phosphorylase PpsR in *E. coli* [55], and an *in vitro* study demonstrated that the YeaG kinase phosphorylates isocitrate lyase only in the

presence of malate [56]. Similarly, the *Legionella pneumophila* Lpg2603 protein kinase remains inactive until it is allosterically activated by the host-derived inositol hexakisphosphate on infection [57]. Accumulating evidence thus suggests that many metabolite signal inputs to protein kinases are yet to be discovered.

Beyond dedicated protein kinases, also the many metabolic kinases could potentially exhibit moonlighting function as protein kinases, as was demonstrated for pyruvate kinase, phosphoglycerate kinase 1, and ketohexokinase in human cells [58]. In addition to their canonical metabolic activity, they phosphorylate various targets *in vivo* and regulate multiple cellular processes, such as the Warburg effect, gene expression, autophagy, and apoptosis. Although there is currently no evidence that metabolic kinases phosphorylate also protein substrates in *E. coli*, it is not unlikely that such moonlighting activity can at least partially explain the gap between around 3000 detected phosphosites and less than 20 annotated kinases with serine/threonine/tyrosine phosphorylation activity [9,59]. Given that metabolic kinases are typically subject to allosteric regulation by multiple metabolites, their putative moonlighting activity might also respond to metabolite input signals.

Metabolite interaction with other proteins

The vast majority of studies focus on identifying the input signals to the various regulatory proteins, but metabolites interact also with other proteins. A combination of limited proteolysis and mass spectrometry identified 1447 novel, physical protein interactions between proteins present in whole-cell lysates and 20 central metabolites in *E. coli*, of which 28% involved noncatalytic proteins and 21% proteins with unknown function [55]. Subsequently, ligand-detected nuclear magnetic resonance (NMR) screening of 29 central *E. coli* enzymes exposed to 55 polar metabolites identified 76 new interactions with some proteins binding up to 11 metabolites [60]. These results agree with a previous observation that protein regulation through metabolite binding is ubiquitous [61]. An aggregation of independently validated enzyme−metabolite interactions in humans and *E. coli* showed that binding of metabolites structurally similar to an enzyme's substrates is a common occurrence, but evolution has selected strategies to avoid these interactions when they are detrimental to the cell [61]. Together, these results advocate for a scenario where plenty of metabolite-protein interactions are waiting to be identified, even among proteins with no known function, and new avenues of research will include the dynamics of competing metabolites and the functional relevance of the interactions.

Riboswitches

Beyond proteins, metabolites can also bind to secondary RNA structures of so-called riboswitches that are typically located in the 5' untranslated region (UTR) of coding mRNAs. Metabolite binding induces conformational changes that affect the transcription or translation rate of the mRNA, for example, by obstructing access to the Shine-Dalgarno sequence [62]. Most known riboswitches regulate genes in pathways that involve the binding metabolite, hence creating feedback inhibition. Riboswitches are classified depending on the structure they adopt to bind their ligand. At least 38 classes have been experimentally validated to bind 27 metabolites, including coenzymes, nucleotide derivatives, second messengers, amino acids, and others [63]. Computational prediction of RNA regions that can potentially act as riboswitches suggested hundreds of 'orphan' riboswitches yet without known ligands [64−66]. Identifying such putative ligands is hampered by the requirement of *a priori* hypotheses of interacting metabolites for biochemical validation, which is challenging when the regulated gene has no known function or obvious candidates do not bind. Since 2004, ligands have been identified for only 16 previously orphan riboswitches [67], the most recent are new classes of tetrahydrofolate [68], guanidine [69,70] and glycine [71] riboswitches. Diversity analysis suggested that riboswitch classes follow a power-law distribution, where few classes have many members that occur in many different species, whereas most classes have few members and occur in only some species [72]. Current research supports this model and suggests that there are hundreds of riboswitch classes still waiting to be identified [66], presaging that we have only scratched the surface of metabolite binding to riboswitches.

Conclusions

Accumulating knowledge of input signals into microbial regulation systems supports the notion of metabolism as an intracellular signal generator for TFs, TCSs, protein kinases, and riboswitches. About 90% of the known metabolite-protein interactions are with TFs and metabolic enzymes that represent less than a quarter of all encoded proteins (Figure 2), suggesting that our knowledge originating primarily from biochemical work on individual proteins may be biased. Although lagging behind other interaction studies, the expanding portfolio of systematic discovery methods in bacteria [5,7,17,55,60] and eukaryotes [73,74] holds promise to expand the regulatory interactome space. Although higher-throughput approaches suffer from high rates of false-positive hits, prioritization of candidates for validation, or choice of relevant conditions for screening, first applications based on physical interactions [55] or correlating dynamic metabolite levels with transcript levels [17] already expanded the regulatory interaction

space substantially. The presented evidence suggests that we are far from a comprehensive metabolite-protein interactome in bacteria and even further from understanding functionality. In particular, the latter is hampered by the highly dynamic nature of the weak interactions between metabolites and macromolecules and the many overlapping regulation processes, delineation of which requires advanced computational methods to integrate multiple data types [75,76], as was demonstrated for identifying expression controlling metabolites in *E. coli* [17,32,77]. To truly accelerate the pace of understanding regulatory functionality, it will be necessary to combine both *in vivo* and *in vitro* high-throughput approaches for the same regulators, such that functional and interaction evidence can be combined and false positives kept to a minimum. Tedious validation efforts can then focus on high confidence cases. Inevitably, the ultimate demonstration of understanding requires some type of mathematical formulation that quantitatively relates input signals to regulatory output and is able to accurately capture the dynamic transition to a new steady state, ideally with predictive power.

Although the regulatory metabolite-protein interactome is daunting, microbes appear to rely on relatively few general metabolic signals that report on the global cell state and control large modules [2,27,33,79]. Thus, there is simplicity in the apparent complexity. Our increasing knowledge on regulatory metabolic feedback is already being exploited for devising dynamic control systems in biotechnology [78] and is also crucial for pharmacological interventions. Beyond identifying these general signals, a remaining challenge is to understand the coordination of local and global metabolic signals across different regulatory layers and time scales, which again will ultimately require computational models to quantitatively demonstrate the extent of our understanding.

Conflict of interest statement

The authors declare no conflict of interest.

Acknowledgements

DLT acknowledges funding through an EDH Postdoctoral Fellowship, and part of this work was funded through a grant of the Botnar Child Care Foundation to US and DLT. Figure 1 was created with BioRender.com.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.coisb.2021.100404.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Krell T, Lacal J, Busch A, Silva-Jiménez H, Guazzaroni ME, Ramos JL: **Bacterial sensor kinases: diversity in the recognition of environmental signals**. *Annu Rev Microbiol* 2010, **64**:539–559.

2. Chubukov V, Gerosa L, Kochanowski K, Sauer U: **Coordination of microbial metabolism**. *Nat Rev Microbiol* 2014, **12**:327–340. 2014 125.

3. Wegner A, Meiser J, Weindl D, Hiller K: **How metabolites modulate metabolic flux**. *Curr Opin Biotechnol* 2015, **34**:16–22.

4. Ryan DG, Murphy MP, Frezza C, Prag HA, Chouchani ET, O'Neill LA, Mills EL: **Coupling Krebs cycle metabolites to signalling in immunity and cancer**. *Nat Metab* 2018, **1**:16–33. 2019 11.

5. ** Rinschen MM, Ivanisevic J, Giera M, Siuzdak G: **Identification of bioactive metabolites using activity metabolomics**. *Nat Rev Mol Cell Biol* 2019, **20**:353–367. 206 2019.
Review of recent examples and methods used to understand how the metabolome impacts other layers of cellular organization mostly in eukaryotes. Covers direct metabolite-protein interactions but also indirect effects of a metabolite on metabolic fluxes or cancer biogenesis.

6. Kosmacz M, Sokołowska EM, Bouzaa S, Skirycz A: **Towards a functional understanding of the plant metabolome**. *Curr Opin Plant Biol* 2020, **55**:47–51.

7. Diether M, Sauer U: **Towards detecting regulatory protein–metabolite interactions**. *Curr Opin Microbiol* 2017, **39**: 16–23.

8. Luzarowski M, Skirycz A: **Emerging strategies for the identification of protein-metabolite interactions**. *J Exp Bot* 2019, **70**: 4605–4617.

9. Gruber CH, Diether M, Sauer U: **Conservation of metabolic regulation by phosphorylation and non-covalent small-molecule interactions**. *Cell Syst* 2021, **12**:538–546.

10. Pérez-Rueda E, Tenorio-Salgado S, Huerta-Saquero A, Balderas-Martínez YI, Moreno-Hagelsieb G: **The functional landscape bound to the transcription factors of Escherichia coli K-12**. *Comput Biol Chem* 2015, **58**:93–103.

11. Babu MM, Teichmann SA: **Evolution of transcription factors and the gene regulatory network in Escherichia coli**. *Nucleic Acids Res* 2003, **31**:1234–1244.

12. Santos-Zavaleta A, Salgado H, Gama-Castro S, Anchez P, Erez MS, Omez-Romero LG, Ledezma-Tejeida D, Santiago García-Sotelo J, Alquicira-Hernándezhernhernández K, Jos´ LJ, Muñiz J, *et al.*: **RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12**. *Nucleic Acids Res* 2019, **47**, https://doi.org/10.1093/nar/gky1077.

13. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M, *et al.*: **The EcoCyc database: reflecting new knowledge about Escherichia coli K-12**. *Nucleic Acids Res* 2017, **45**:D543–D550.

14. Stefanello AA, Oliveira MAS de, Souza EM, Pedrosa FO, Chubatsu LS, Huergo LF, Dixon R, Monteiro RA: **Regulation of Herbaspirillum seropedicae NifA by the GlnK PII signal transduction protein is mediated by effectors binding to allosteric sites**. *Biochim Biophys Acta Protein Proteonomics* 2020, **1868**:140348.

15. Ellermann M, Jimenez AG, Pifer R, Ruiz N, Sperandio V: **The canonical long-chain fatty acid sensing machinery processes arachidonic acid to inhibit virulence in enterohemorrhagic escherichia coli**. *mBio* 2021, **12**:1–17.

16. Shimada T, Yokoyama Y, Anzai T, Yamamoto K, Ishihama A: **Regulatory role of PlaR (YiaJ) for plant utilization in Escherichia coli K-12**. *Sci Rep* 2019, **9**:1–12.

17. ** Lempp M, Farke N, Kuntz M, Freibert SA, Lill R, Link H: **Systematic identification of metabolites controlling gene expression in E. coli**. *Nat Commun* 2019, **10**.
Dynamic metabolomics and transcriptomics of *E. coli* to identify metabolites whose abundance correlates with the activity of 209 TFs. Integration of the data predicted 39 novel TF-metabolite-TF interactions and validate three of them.

18. Liao JC, Boscolo R, Yang Y, Tran LM, Sabatti C, Roychowdhury VP: **Network component analysis:**

reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A* 2003, **100**:15522–15527.

19. Hackett SR, Baltz EA, Coram M, Wranik BJ, Kim G, Baker A, Fan M, Hendrickson DG, Berndl M, McIsaac RS: **Learning causal networks using inducible transcription factors and transcriptome-wide time series**. *Mol Syst Biol* 2020, **16**:1–15.

20. Lara J, Diacovich L, Trajtenberg F, Larrieux N, Malchiodi EL, Fernández MM, Gago G, Gramajo H, Buschiazzo A: **Mycobacterium tuberculosis FasR senses long fatty acyl-CoA through a tunnel and a hydrophobic transmission spine**. *Nat Commun* 2020, **11**:1–13.

21. He X, Ni D, Zhang H, Li X, Zhang J, Fu Q, Liu Y, Lu S: **Zinc-mediated conformational preselection mechanism in the allosteric control of DNA binding to the zinc transcriptional regulator (ZitR)**. *Sci Rep* 2020, **10**:1–12.

22. Vigouroux A, Meyer T, Naretto A, Legrand P, Aumont-Nicaise M, Di Cicco A, Renoud S, Doré J, Lévy D, Vial L, *et al.*: **Characterization of the first tetrameric transcription factor of the GntR superfamily with allosteric regulation from the bacterial pathogen Agrobacterium fabrum**. *Nucleic Acids Res* 2021, **49**: 529–546.

23. Pedre B, Young D, Charlier D, Á Mourenza, Rosado LA, Marcos-Pascual L, Wahni K, Martens E, de la Rubia AG, Belousov VV, *et al.*: **Structural snapshots of OxyR reveal the peroxidatic mechanism of H2O2 sensing**. *Proc Natl Acad Sci U S A* 2018, **115**:E11623–E11632.

24. Alon U: *An introduction to systems biology*. Chapman and Hall/CRC; 2019.

25. * Tack DS, Tonner PD, Pressman A, Olson ND, Levy SF, Romantseva EF, Alperovich N, Vasilyeva O, Ross D: **The genotype-phenotype landscape of an allosteric protein**. *Mol Syst Biol* 2021, **17**:1–26.

Screening responses of 10^5 *E. coli* LacI mutants to different inducer concentrations quantified the fragile equilibrium of the relationship between effector concentration and expression level of regulated genes. It was demonstrated that few mutations are sufficient to generate new dynamics or even turning a repressor into an activator.

26. Keren L, Zackay O, Lotan-Pompan M, Barenholz U, Dekel E, Sasson V, Aidelberg G, Bren A, Zeevi D, Weinberger A, *et al.*: **Promoters maintain their relative activity levels under different growth conditions**. *Mol Syst Biol* 2013, **9**:1–17.

27. Kochanowski K, Gerosa L, Brunner SF, Christodoulou D, Nikolaev YV, Sauer U: **Few regulatory metabolites coordinate expression of central metabolic genes in *Escherichia coli***. *Mol Syst Biol* 2017, **13**:903.

28. Gerosa L, Kochanowski K, Heinemann M, Sauer U: **Dissecting specific and global transcriptional regulation of bacterial gene expression**. *Mol Syst Biol* 2013, **9**:1–11.

29. Martinez-Antonio A, Janga SC, Salgado H, Collado-Vides J: **Internal-sensing machinery directs the activity of the regulatory network in Escherichia coli**. *Trends Microbiol* 2006, **14**: 22–27.

30. Laxhuber KS, Morrison MJ, Chure G, Belliveau NM, Strandkvist C, Naughton KL, Phillips R: **Theoretical investigation of a genetic switch for metabolic adaptation**. *PLoS One* 2020, **15**:1–23.

32. ** Kochanowski K, Okano H, Patsalo V, Williamson J, Sauer U, Hwa T: **Global coordination of metabolic pathways in Escherichia coli by active and passive regulation**. *Mol Syst Biol* 2021, **17**:1–14.

Identified CRP and its signal metabolite cAMP as the main regulator of proteome allocation when *E. coli* coordinates catabolism and anabolism. Additionally, CRP-cAMP indirectly promotes a decrease in the necessary resources to produce anabolic enzymes. A good example for broad effects of a signaling metabolite through a global transcription factor.

33. Okano H, Hermsen R, Kochanowski K, Hwa T: **Regulation underlying hierarchical and simultaneous utilization of carbon substrates by flux sensors in Escherichia coli**. *Nat Microbiol* 2020, **5**:206–215.

34. * Baksh KA, Pichugin D, Prosser RS, Zamble DB: **Allosteric regulation of the nickel-responsive NikR transcription factor from Helicobacter pylori**. *J Biol Chem* 2021, **296**:100069.

From structural information of NikR bound to its effectors nickel and zinc, a mechanism was proposed that explains why NikR regulates gene expression in presence of one effector but not the other.

35. van Vliet AH, Kuipers EJ, Waidner B, Davies BJ, de Vries N, Penn CW, Vandenbroucke-Grauls CM, Kist M, Bereswill S, Kusters JG: **Nickel-responsive induction of urease expression in Helicobacter pylori is mediated at the transcriptional level**. *Infect Immun* 2001, **69**:4891–4897.

36. Martínez-Flores I, Pérez-Morales D, Sánchez-Pérez M, Paredes CC, Collado-Vides J, Salgado H, Bustamante VH: **In silico clustering of Salmonella global gene expression data reveals novel genes co-regulated with the SPI-1 virulence genes through HilD**. *Sci Rep* 2016, **6**:1–12.

37. Ledezma-Tejeida D, Altamirano-Pacheco L, Fajardo V, Collado-Vides J: **Limits to a classic paradigm: most transcription factors in E. coli regulate genes involved in multiple biological processes**. *Nucleic Acids Res* 2019, https://doi.org/10.1093/nar/gkz525.

38. Gao R, Bouillet S, Stock AM: **Structural basis of response regulator function**. *Annu Rev Microbiol* 2019, **73**:175–197.

39. Li X, Lv X, Lin Y, Zhen J, Ruan C, Duan W, Li Y, Xie J: **Role of two-component regulatory systems in intracellular survival of Mycobacterium tuberculosis**. *J Cell Biochem* 2019, **120**: 12197–12207.

40. * Goodman KN, Powers MJ, Crofts AA, Trent MS, Hendrixson DR: **Campylobacter jejuni BumSR directs a response to butyrate via sensor phosphatase activity to impact transcription and colonization**. *Proc Natl Acad Sci U S A* 2020, **117**:11715–11726.

Identified the first bacterial butyrate-sensing two-component system BumSR, which enables *Campylobacter jejuni* to spatially orient in the gut and preferentially infect the lower intestinal niches. The non-canonical sensor BumS acts as a phosphatase, rather than a kinase, sensing butyrate indirectly and affecting the transcription of colonization factors through the cognate BumR response regulator.

41. Gillis CC, Winter MG, Chanin RB, Zhu W, Spiga L, Winter SE: **Host-derived metabolites modulate transcription of salmonella genes involved in L-Lactate utilization during gut colonization**. *Infect Immun* 2019, **87**.

42. Párraga Solórzano PK, Yao J, Rock CO, Kehl-Fiea TE: **Disruption of glycolysis by nutritional immunity activates a two-component system that coordinates a metabolic and antihost response by staphylococcus aureus**. *mBio* 2019, **10**:1321–1340.

43. Cheng ST, Wang FF, Qian W: **Cyclic-di-GMP binds to histidine kinase RavS to control RavS-RavR phosphotransfer and regulates the bacterial lifestyle transition between virulence and swimming**. *PLoS Pathog* 2019, **15**, e1007952.

44. Hariharan VN, Yadav R, Thakur C, Singh A, Gopinathan R, Singh DP, Sankhe G, Malhotra V, Chandra N, Bhatt A, *et al.*: **Cyclic di-GMP sensing histidine kinase PdtaS controls mycobacterial adaptation to carbon sources**. *Faseb J* 2021, **35**, e21475.

45. Hu Q, Zhang J, Chen Y, Hu L, Li W, He ZG: **Cyclic di-GMP co-activates the two-component transcriptional regulator DevR in Mycobacterium smegmatis in response to oxidative stress**. *J Biol Chem* 2019, **294**:12729–12742.

46. Miyake Y, Inaba T, Watanabe H, Teramoto J, Yamamoto K, Ishihama A: **Regulatory roles of pyruvate-sensing two-component system PyrSR (YpdAB) in *Escherichia coli* K-12**. *FEMS Microbiol Lett* 2019, **366**:9.

47. Ogasawara H, Ishizuka T, Yamaji K, Kato Y, Shimada T, Ishihama A: **Regulatory role of pyruvate-sensing BtsSR in biofilm formation by Escherichia coli K-12**. *FEMS Microbiol Lett* 2020, **366**:251.

48. Desai SK, Kenney LJ: **To ~P or Not to ~P? Non-canonical activation by two-component response regulators**. *Mol Microbiol* 2017, **103**:203–213.

49. * Sharma S, Kumari P, Vashist A, Kumar C, Nandi M, Tyagi JS: **Cognate sensor kinase-independent activation of Mycobacterium tuberculosis response regulator DevR (DosR) by acetyl phosphate: implications in anti-mycobacterial drug design**. *Mol Microbiol* 2019, **111**:1182–1194.

Demonstrated that acetyl phosphate can non-enzymatically phosphorylate and activate DevR response regulator of the DevRS/DosT two-component system independent of its cognate sensor kinases. These results suggest that anti-mycobacterial therapy aimed at a kinase might be not as effective as targeting DevR directly.

50. * Singh KK, Athira PJ, Bhardwaj N, Singh DP, Watson U, Saini DK: **Acetylation of response regulator protein MtrA in M. tuberculosis regulates its repressor activity**. *Front Microbiol* 2021, **11**:516315.

Showed that MtrA response regulator can be non-enzymatically acetylated with acetyl phosphate, and this modification overrides the effect of phosphorylation. It was unexpected that passive acetylation can dominate over enzymatic phosphorylation and opens new questions about the functional role of a widespread acetylation in bacteria.

51. Macek B, Forchhammer K, Hardouin J, Weber-Ban E, Grangeasse C, Mijakovic I: **Protein post-translational modifications in bacteria**. *Nat Rev Microbiol* 2019, **17**:651–664.

52. Ochoa D, Jonikas M, Lawrence RT, El Debs B, Selkrig J, Typas A, Villén J, Santos SD, Beltrao P: **An atlas of human kinase regulation**. *Mol Syst Biol* 2016, **12**:888.

53. Wang YP, Lei QY: **Metabolite sensing and signaling in cell metabolism**. *Signal Transduct Target Ther* 2018, **3**:1–9.

54. Milanesi R, Coccetti P, Tripodi F: **The regulatory role of key metabolites in the control of cell signaling**. *Biomolecules* 2020, **10**:862.

55. Piazza I, Kochanowski K, Cappelletti V, Fuhrer T, Noor E, Sauer U, Picotti P: **A map of protein-metabolite interactions reveals principles of chemical communication**. *Cell* 2018, **172**:358–372. e23.

56. Sultan A, Jers C, Ganief TA, Shi L, Senissar M, Køhler JB, Macek B, Mijakovic I: **Phosphoproteome study of Escherichia coli devoid of ser/thr kinase YeaG during the metabolic shift from glucose to malate**. *Front Microbiol* 2021, **12**:657562.

57. * Sreelatha A, Nolan C, Park BC, Pawłowski K, Tomchick DR, Tagliabracci VS: **A Legionella effector kinase is activated by host inositol hexakisphosphate**. *J Biol Chem* 2020, **295**:6214–6224.

Identified Lpg2603 as an atypical protein kinase in *Legionella pneumophila*, that phosphorylates the substrate only in the presence of inositol hexakisphosphate, which allosterically activates it. This is the first bacterial kinase that requires the eukaryote-specific host-derived inositol hexakisphosphate for activity.

58. Lu Z, Hunter T: **Metabolic kinases moonlighting as protein kinases**. *Trends Biochem Sci* 2018, **43**:301–310.

59. Semanjski M, Germain E, Bratl K, Kiessling A, Gerdes K, Macek B: **The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli to promote multidrug tolerance**. *Sci Signal* 2018, **11**:eaat5750.

60. Diether M, Nikolaev Y, Allain FH, Sauer U: **Systematic mapping of protein-metabolite interactions in central metabolism of Escherichia coli**. *Mol Syst Biol* 2019, **15**, e9008.

61. Alam MT, Olin-Sandoval V, Stincone A, Keller MA, Zelezniak A, Luisi BF, Ralser M: **The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization**. *Nat Commun* 2017, **8**:1–13.

62. Bédard ASV, Hien EDM, Lafontaine DA: **Riboswitch regulation mechanisms: RNA, metabolites and regulatory proteins**. *Biochim Biophys Acta - Gene Regul Mech* 2020, **1863**:194501.

63. Mccown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR: **Riboswitch diversity and distribution**. *RNA* 2017, **23**:995–1011.

64. Weinberg Z, Lünse CE, Corbino KA, Ames TD, Nelson JW, Roth A, Perkins KR, Sherlock ME, Breaker RR: **Detection of 224 candidate structured RNAs by Comparative analysis of specific subsets of intergenic regions**. *Nucleic Acids Res* 2017, **45**:10811–10823.

65. Stav S, Atilho RM, Mirihana Arachchilage G, Nguyen G, Higgs G, Breaker RR: **Genome-wide discovery of structured noncoding RNAs in bacteria**. *BMC Microbiol* 2019, **19**:1–18.

66. Brewer KI, Greenlee EB, Higgs G, Yu D, Mirihana Arachchilage G, Chen X, King N, White N, Breaker RR: **Comprehensive discovery of novel structured noncoding RNAs in 26 bacterial genomes**. *RNA Biol* 2021:1–16. 00.

67. Sherlock ME, Breaker RR: **Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes**. *RNA* 2020, **26**:675–693.

68. Chen X, Arachchilage GM, Breaker RR: **Biochemical validation of a second class of tetrahydrofolate riboswitches in bacteria**. *RNA* 2019, **25**:1091–1097.

69. Salvail H, Balaji A, Yu D, Roth A, Breaker RR: **Biochemical validation of a fourth guanidine riboswitch class in bacteria**. *Biochemistry* 2020, **59**:4654–4662.

70. Lenkeit F, Eckert I, Hartig JS, Weinberg Z: **Discovery and characterization of a fourth class of guanidine riboswitches**. *Nucleic Acids Res* 2020, **48**:12889–12899.

71. Munyati-Othman N, Appasamy SD, Damiri N, Emrizal R, Alipiah NM, Ramlan EI, Firdaus-Raih M: **Regulation of Glycine cleavage and detoxification by a highly conserved Glycine riboswitch in burkholderia spp**. *Curr Microbiol* 2021, **78**:2943–2955.

72. Ames TD, Breaker RR: **Bacterial riboswitch discovery and analysis**. *Chem Biol Nucleic Acids* 2010, https://doi.org/10.1002/9780470664001.ch20.

73. Wang Z, Pisano S, Ghini V, Kadeřávek P, Zachrdla M, Pelupessy P, Kazmierczak M, Marquardsen T, Tyburn J-M, Bouvignies G, *et al.*: **Detection of metabolite-protein interactions in complex biological samples by high-resolution relaxometry: toward interactomics by NMR**. *J Am Chem Soc* 2021, https://doi.org/10.1021/jacs.1c01388.

74. Venegas-Molina J, Molina-Hidalgo FJ, Clicque E, Goossens A: **Why and how to dig into plant metabolite-protein interactions**. *Trends Plant Sci* 2021, **26**:472–483.

75. Noor E, Cherkaoui S, Sauer U: **Biological insights through omics data integration**. *Curr Opin Struct Biol* 2019, **15**:39–47.

76. Wörheide MA, Krumsiek J, Kastenmüller G, Arnold M: **Multi-omics integration in biomedical research - a metabolomics-centric review**. *Anal Chim Acta* 2021, **1141**:144–162.

77. Ledezma-Tejeida D, Ishida C, Collado-Vides J: **Genome-wide mapping of transcriptional regulation and metabolism describes information-processing units in Escherichia coli**. *Front Microbiol* 2017, **8**.

78. Ni C, Dinh CV: **Prather KLJ: dynamic control of metabolism**. *Annu Rev Chem Biomol Eng* 2021, **12**:519–541.

79. Erickson DW, Schink SJ, Patsalo V, Williamson JR, Gerland U, Hwa T: **A global resource allocation strategy governs growth transition kinetics of Escherichia coli**. *Nature* 2017, **551**:119–123.

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

81. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, *et al.*: **HMDB 4.0: the human metabolome database for 2018**. *Nucleic Acids Res* 2018, **46**:D608–D617.