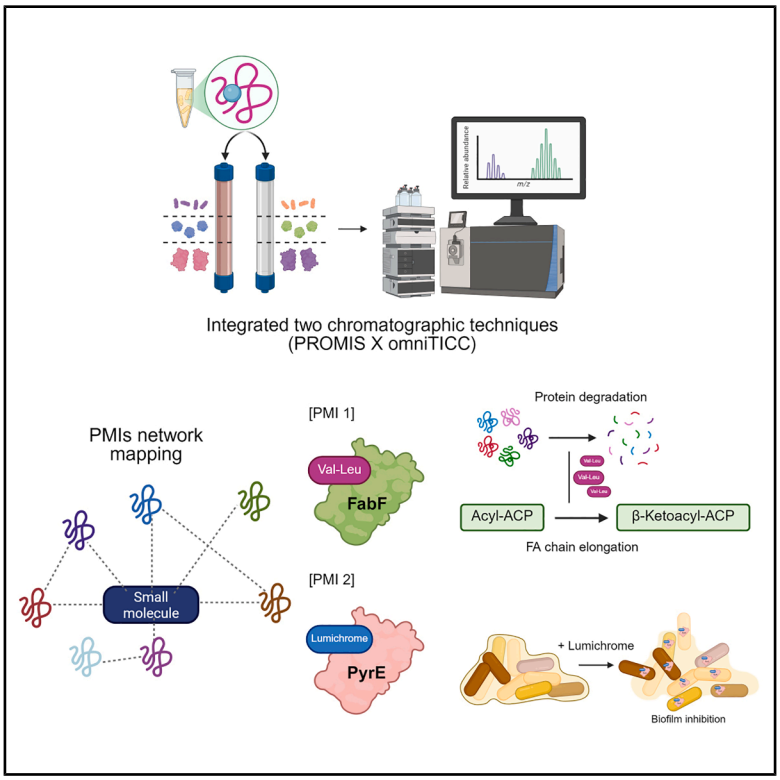
iScience

Article

Mapping protein-metabolite interactions in *E. coli* by integrating chromatographic techniques and co-fractionation mass spectrometry

Graphical abstract



Authors

Mateusz Wagner, Jieun Kang, Catherine Mercado, ..., Frank C. Schroeder, Dirk Walther, Aleksandra Skirycz

Correspondence

skirycza@msu.edu

In brief

Microbiology; Omics; Metabolomics

Highlights

- Integrating omniTICC and PROMIS improves accuracy of PMI networks
- Val-Leu binding to FabF suggests a link between protein degradation and lipid metabolism
- Lumichrome binding to PyrE links flavins to biofilm formation

Wagner et al., 2025, iScience *28*, 112611
June 20, 2025
https://doi.org/10.1016/j.isci.2025.112611

CellPress

iScience

CellPress
OPEN ACCESS

Article

Mapping protein-metabolite interactions in *E. coli* by integrating chromatographic techniques and co-fractionation mass spectrometry

Mateusz Wagner,[1,2,10] Jieun Kang,[1,3,10] Catherine Mercado,[1,3] Venkatesh P. Thirumlaikumar,[1,4] Michal Gorka,[5] Hanne Zillmer,[6] Jingzhe Guo,[7] Romina I. Minen,[1,8] Caroline F. Plecki,[1,9] Katayoon Dehesh,[7] Frank C. Schroeder,[1,2] Dirk Walther,[6] and Aleksandra Skirycz[1,2,3,11,*]

[1]Boyce Thompson Institute, Ithaca, NY 14853, USA
[2]Cornell University, Ithaca, NY 14853, USA
[3]Michigan State University, East Lansing, MI 48824, USA
[4]Purdue University, West Lafayette, IN 47907, USA
[5]Celon Pharma, 05092 Lomianki, Poland
[6]Max-Planck-Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany
[7]University of California, Riverside, Riverside, CA 92521, USA
[8]DKFZ German Cancer Research Center, 69120 Heidelberg, Germany
[9]Syracuse University, Syracuse, NY 13244, USA
[10]These authors contributed equally
[11]Lead contact
*Correspondence: skirycza@msu.edu
https://doi.org/10.1016/j.isci.2025.112611

SUMMARY

Toward characterization of protein-metabolite interactomes, we recently introduced PROMIS, a co-fractionation-based mass spectrometry approach. However, the challenge lies in distinguishing true interactors from coincidental co-elution when a metabolite co-fractionates with numerous proteins. To address this, we integrated two chromatographic techniques—size exclusion and ion exchange—to enhance the mapping of protein-metabolite interactions (PMIs) in *Escherichia coli*. This integration aims to refine the PMI network by considering size and charge characteristics, resulting in 994 interactions involving 51 metabolites and 465 proteins. The PMI network is enriched for known and predicted interactions, providing validation. Furthermore, analyzing protein targets for different metabolites revealed functional insights, such as a connection between proteinogenic dipeptides and fatty acid biosynthesis. Notably, we uncovered an inhibitory interaction between the riboflavin degradation product lumichrome and orotate phosphoribosyltransferase, a key enzyme in *de novo* pyrimidine synthesis affecting biofilm formation. In summary, our integrated chromatographic approach significantly advances PMI mapping.

INTRODUCTION

Protein-metabolite interaction (PMI) networks remain largely uncharacterized even in simple model systems, although metabolites play an essential role in regulating protein activities, protein-protein interactions (PPIs), protein subcellular localization, and protein condensed states.[1–4] Small molecule regulation is known for almost all protein classes. Notably, nearly 75% of the 300 transcription factors (TFs) in *Escherichia coli* have predicted small-molecule binding domains.[5] PMI networks are often complex and highly dynamic, responding rapidly to environmental and developmental cues. Recent studies highlight the potential for a single protein to engage with multiple small molecule ligands under different conditions, facilitating the discovery of previously unknown regulatory connections.[6–12]

Biochemical methods for the identification of PMIs range from the classical affinity purifications that exploit either a metabolite or a protein bait to identify partners to more recent approaches that measure ligand-induced changes in the protein properties such as thermal stability or accessibility to a protease around the binding site.[1,2,13–15] The methods mentioned above have their unique advantages and drawbacks but depend on availability of a metabolite or protein bait. We previously developed PROMIS (Protein-Metabolite Interactions using Size separation),[16] a co-fractionation mass spectrometry (CF-MS) approach for proteome and metabolome-wide mapping of protein-metabolite complexes to circumvent this limitation. CF-MS methods combine separation of native complexes with mass spectrometry-based analysis of the obtained fractions and use the similarity of elution profiles to delineate interactors.[17] CF-MS approaches have been instrumental in generating comprehensive PPI networks across model and non-model organisms, e.g.[18–25] In its original implementation, PROMIS relied on size separation chromatography.[26] Whenever a metabolite is bound to a protein

complex it would separate into earlier-eluting, protein-containing, high molecular weight fractions. In contrast, unbound metabolites would separate into late-eluting, protein-free, low molecular weight fractions. Fractions are analyzed using non targeted proteomic and metabolomic analyses, and the resulting small-molecule and protein profiles are searched for co-eluting metabolites and proteins. Recent PROMIS studies in Arabidopsis[16] and yeast[27,28] reported hundreds of metabolic features separating with protein complexes, of which only a subset could be annotated and many more awaiting chemical identification. Independently, Li et al.[29] identified 461 putative interactions, including a regulatory pairing between isopentenyl adenine and the ribosome, using a size exclusion chromatography (SEC)-based CF-MS workflow in the thermophilic fungus *Chaetomium thermophilum*.

However, in CF-MS experiments a single metabolite often co-fractionates with hundreds of different proteins, of which only a subset may represent ''true interactions'', while the presence of most other proteins is due to coincidental co-elution. Identifying the true binders constitutes the most critical challenge in using CF-MS–based methods for characterization of PMI networks. Based on previous work on PPIs,[17] we hypothesized that integrating PROMIS with two orthogonal chromatographic methods may greatly improve the ability to distinguish between true and coincidental co-elution. As a second chromatography to complement SEC we specifically focused on ion exchange (IEX), given that many similarly sized protein complexes will differ in charge and vice versa. Separation of protein drug complexes has been previously demonstrated by Chan et al. Their method, termed TICC (Target Identification by Chromatographic Co-elution), relies on the distinctive shift observed in a small molecule's chromatographic elution profile upon binding to its protein target.[30] Chan et al.'s experimental strategy encompassed two native IEX separations of the cell lysate incubated with the studied compounds and a control sample containing just the drug that enabled differentiating between the elution profiles of free drug versus drug-protein complex. More recently, TICC was used to separate antibiotic-protein complexes from bacteria.[31]

The central objective of our research presented in this study is to assess whether integrating IEX (TICC) and SEC (PROMIS) can enhance the accuracy of mapping endogenous PMI networks in cell lysates. For this purpose, we selected *E. coli* due to the relatively low complexity of its proteome and metabolome, coupled with the abundance of known or predicted interactions available from databases such as STITCH.[32] As anticipated, we found that integration of PROMIS and TICC facilitates distinguishing between true and coincidental co-elution events, greatly improving our ability to map PMI networks.

RESULTS

Ion exchange is suitable for the separation of protein-metabolite complexes from complex cell lysates

This study aimed to test whether using two orthogonal chromatographies, SEC and IEX, could significantly increase the scope of CF-MS for delineating PMI networks. Our overall experimental strategy is outlined in Figures 1A–1C. Analogous to previous studies,[16,28] PROMIS separation using SEC of *E. coli* lysate yielded 40 protein-containing fractions spanning proteins and protein complexes from approximately 10 MDa to 10 kDa. The collected SEC fractions were analyzed by untargeted liquid chromatography-mass spectrometry (LC-MS) metabolomics and LC-MS/MS proteomics (Figures 1D–1F). In a separate experiment, we subjected the same lysate sample to IEX fractionation (IEX_S fraction set). To distinguish between metabolite peaks corresponding to free vs. protein-bound metabolites, we additionally applied the same IEX fractionation protocol to heat-denatured lysate (IEX_C fraction set). The IEX_S separation yielded 24 fractions of which 17 contained proteins, which were subjected to untargeted metabolomics and proteomics (Figures 1G–1I). SEC and IEX separations were performed for two independent experiments yielding four CF-MS datasets. Metabolic features were annotated by using a library of authentic standards. Following the strategy outlined in the original TICC study[30] the comparison of metabolite elution profiles between IEX_C and IEX_S focused on eliminating peaks representing unbound metabolites, as determined by the control separation (Figures 2A and 2B). Note that although protein-bound metabolites were identified primarily based on shifts of their elution maxima in IEX_S vs. IEX_C, we also considered metabolite peaks whose intensity was strongly increased compared to the control separation as potentially protein-bound (Figure 2B, see an example of niacinamide); assuming that the elution profile will be a sum of a free and a protein-bound metabolite. The profile of differential metabolite elution comparing IEX_S and IEX_C we refer to as the omniTICC profile.[30] The ''omni'' points to the untargeted nature of the approach, enabling characterization of endogenous protein-metabolite complexes.

Proteomics and metabolomics of the SEC and IEX fractions revealed 1479 proteins and 58 known metabolites, annotated using a compound reference library, based on PROMIS and omniTICC elution profile across two independent separations (Figure 2C; Data S1, S2, S3, and S4). Functional analysis of the 1479 proteins revealed enrichment of categories broadly associated with metabolism. The 58 annotated metabolites included amino acids, nucleotides, cofactors, dipeptides, cyclic dipeptides, and diverse metabolic intermediates. Notably, among annotated metabolic features, a larger number of compounds separated with protein complexes in the IEX compared to SEC separations (Figure 2C). The compounds with an omniTICC profile that were absent in one or both PROMIS separations included e.g., dipeptides, metabolic intermediates, and cofactors (Data S2).

Next, we assessed the similarity of the PROMIS and omniTICC elution profiles of the 58 metabolites and 1479 proteins by calculating the Pearson correlation coefficient (PCC) for both replicates (Figure 2D). By scrambling the elution profiles, we additionally simulated random distributions. As expected, the distributions of experimental PCCs were shifted toward higher values compared to the random distributions. Approximately 70% and 85% of the PROMIS and omniTICC elution profiles, for both proteins and metabolites, passed the arbitrary PCC 0.7 cut-offs that we used in the past to define reproducibility.[27,33] The obtained results for PROMIS are on par with what we previously reported in yeast and Arabidopsis[16,28] but also point to

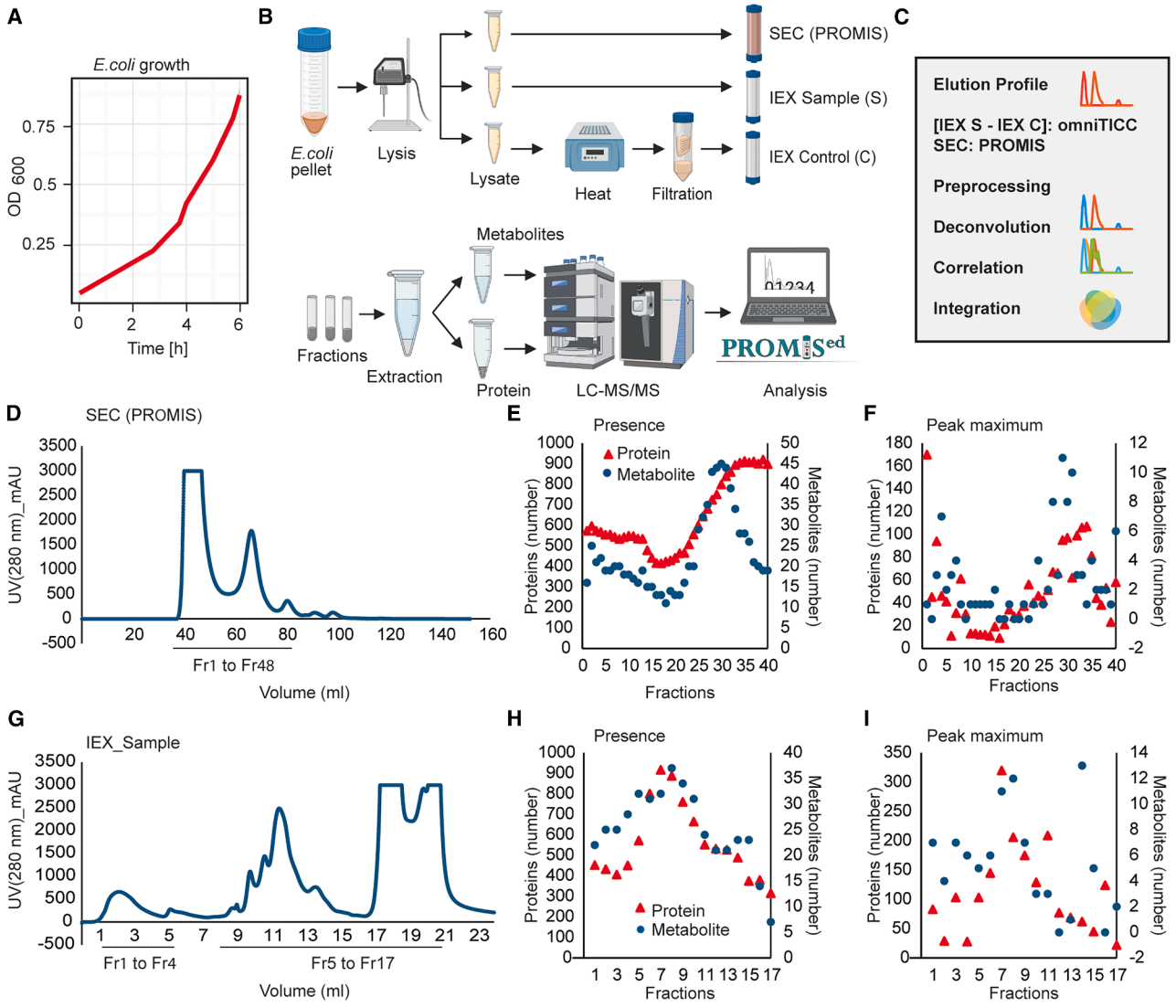


Figure 1. Overview of the experimental strategy

(A) *E. coli* growth curve. Bacterial cells were collected at the exponential phase ($OD_{600\ nm}$ = 0.8).

(B and C) Schematic representation of the experimental strategy and data analysis steps.

(D) Chromatogram of the absorption at 280 nm of SEC separation.

(E) Number of proteins and metabolites detected in the analyzed SEC fractions.

(F) Number of proteins and metabolites with an elution peak maximum in the analyzed SEC fractions.

(G) Chromatogram of the absorption at 280 nm of IEX_sample separation.

(H) Number of proteins and metabolites detected in the analyzed IEX_sample fractions.

(I) Number of proteins and metabolites with an elution peak maximum in the analyzed IEX_sample fractions.(E and H) Analysis was done on the preprocessed and normalized data of the 1479 proteins and 58 metabolites used to construct the PMI network.

(F and I) Analysis was done on the preprocessed, normalized and deconvoluted data of 1479 proteins and 58 metabolites used to construct PMI network. (E, F, H, and I) Preprocessing, normalization and deconvolution of the elution profiles was performed using standard settings in the PROMISed data analysis app.[33]

higher reproducibility of the IEX for separating protein-metabolite complexes. The heatmap representation also reflects the reproducible nature of PROMIS and omniTICC elution profiles (Figure 2E). Our results demonstrate that IEX can be effective at separating protein-metabolite complexes within a complex cellular lysate either as stand-alone separations or complementing SEC.

Functional insight from global analysis of the protein-metabolite interactome

To construct the protein-metabolite interactome of *E. coli*, we looked for protein-metabolite pairings that co-elute in both the PROMIS and omniTICC separations, across two independent repeats. Pre-processing, deconvolution, correlation, and integration steps were performed using the standard analysis

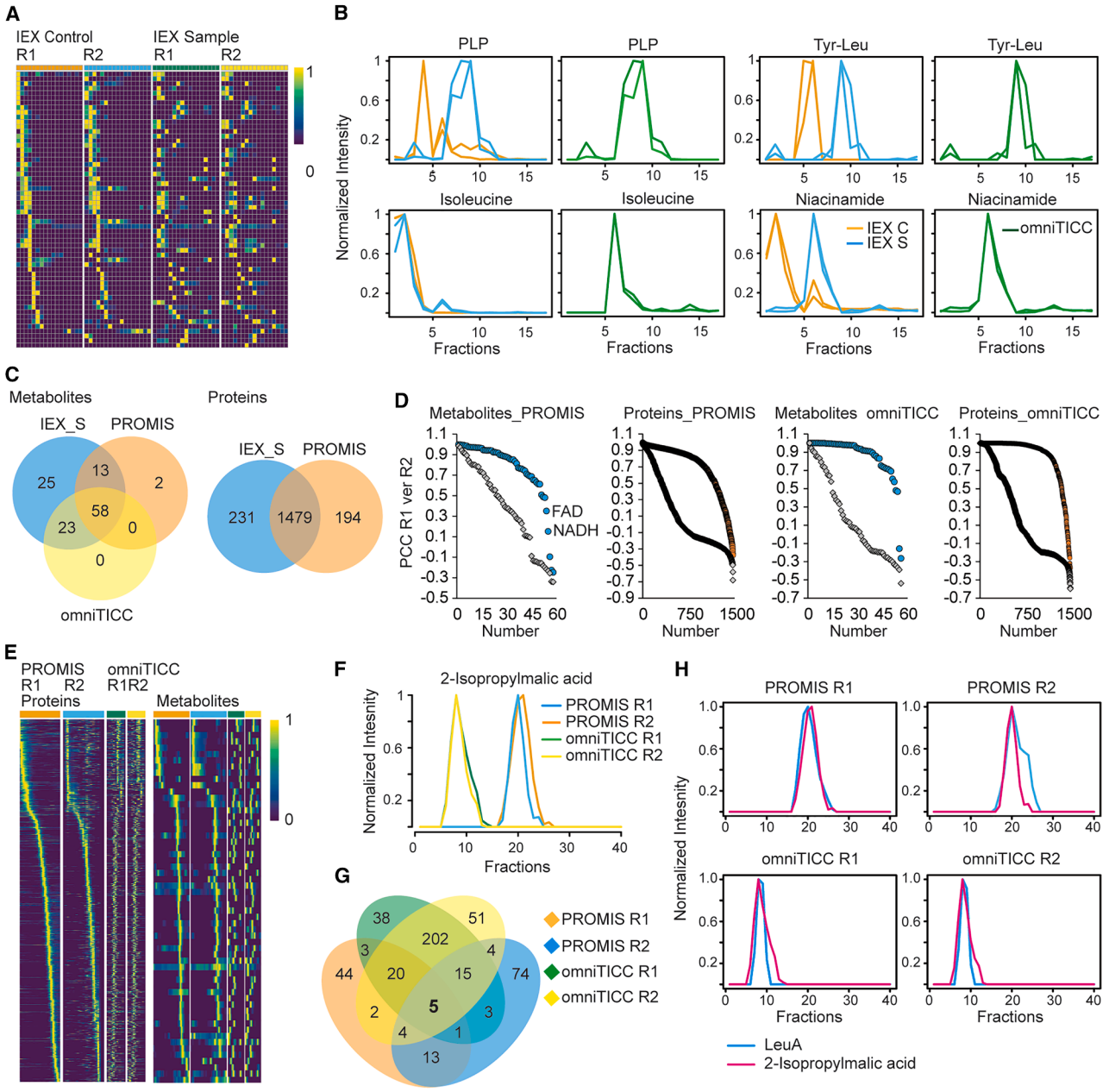


Figure 2. Overview of the PROMIS and omniTICC separations

(A) Heatmap representation of normalized metabolite elution profiles across examined fractions in the IEX control and IEX sample separations.

(B) Example elution profiles of four representative metabolites obtained from IEX control (orange line) and IEX sample (blue line) separations and the resulting omniTICC profile (green line).

(C) Venn diagram for overlap of metabolite and protein elution profiles in two PROMIS and IEX sample separations. Also shown are metabolites with differential omniTICC profiles.

(D) PCCs as calculated between elution profiles of metabolites (blue dots) and proteins (orange dots) from two PROMIS and omniTICC separations (Replicate 1 and Replicate 2). The graphs represent PCC distributions. Also shown are random PCC distributions calculated between randomized metabolite and protein elution profiles (gray diamonds).

(E) Heatmap representation of the protein and metabolite normalized elution profiles across PROMIS (orange and blue lines) and omniTICC (green and yellow lines) separations from two independent experiments.

(F) A representative example of an elution profile of a metabolite, 2-isopropylmalic acid, in the PROMIS and omniTICC separations from two independent experiments.

(G) Proteins co-eluting with 2-isopropylmalic acid in the four separations (Pearson correlation coefficient >0.8).

(H) Co-elution of 2-isopropylmalic acid and its well-known interactor, LeuA, (E–K) R1 and R2 stand for two independent biological replicates. (A–H) Preprocessing and normalization of the elution profiles was performed using standard settings in the PROMISed data analysis app.[33]

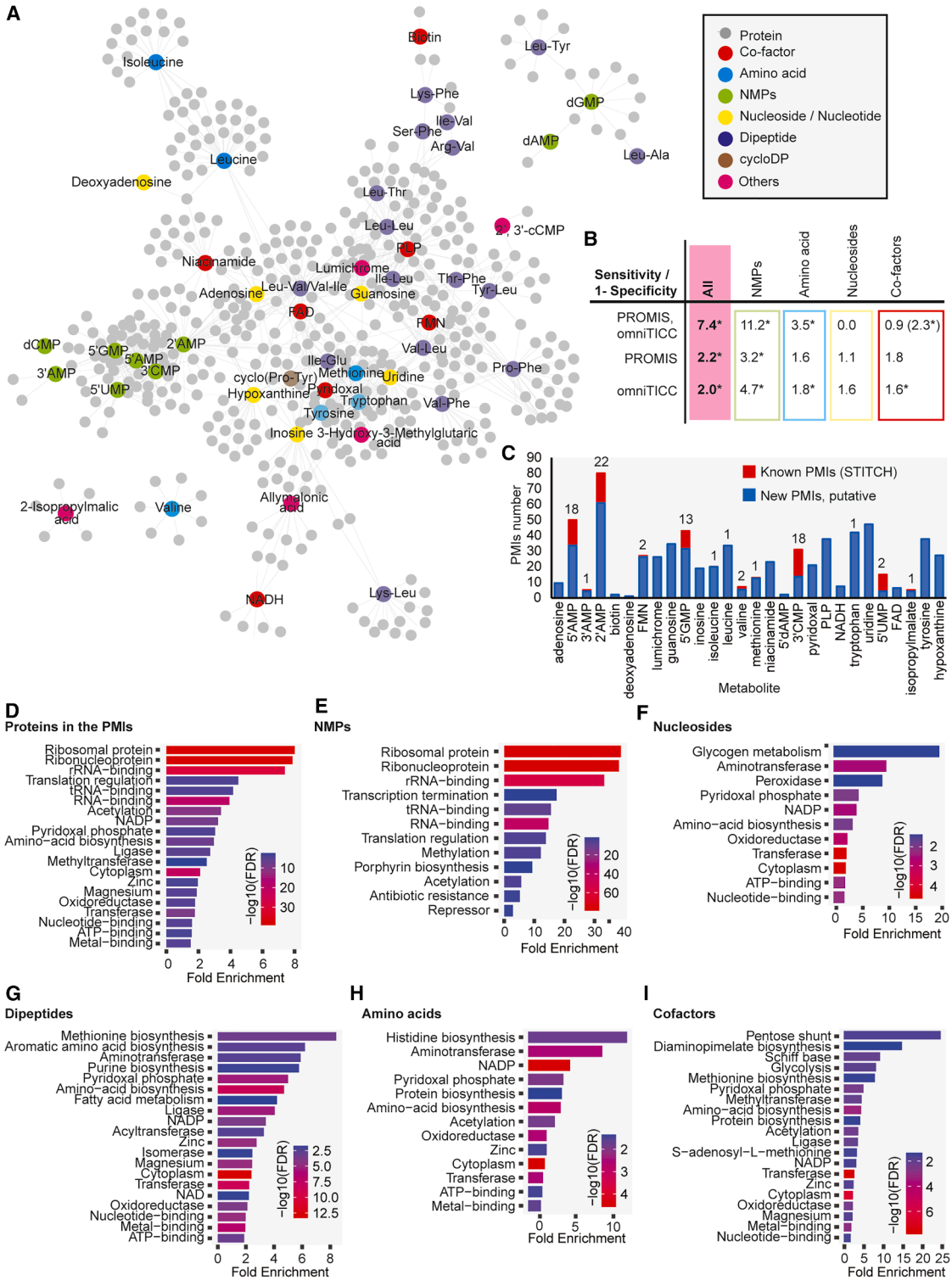


Figure 3. *E. coli* PMI network derived from integration of PROMIS and omniTICC separations

(A) The edges between metabolites and proteins derive from co-elution.

(B) Known PMIs deposited in the STITCH database (confidence score >0.4, based on the experimental evidence) were used to calculate sensitivity/1-specificity ratio for the entire network, and subnetworks of the different compound classes. Co-factors in brackets - edges derived from the two

(legend continued on next page)

settings embedded in the PROMISed app.[33] An example analysis is outlined in Figures 2F–2H for isopropylmalate, an intermediate in leucine biosynthesis. In a single PROMIS or omniTICC separation, isopropylmalate co-elute (Pearson correlation coefficient (PCC) > 0.8)) with tens of proteins, but only five proteins co-elute (PCC >0.8) across all four datasets (two repeats of PROMIS and omniTICC, Figure 2G). We kept the replicates separate rather than calculating an average profile following the recommendations from the recent guidelines for CF-MS experiments.[34,35] The assumption is that if the same protein partner across replicates drives the metabolite elution profile, these will co-fractionate even if the metabolite elution profiles are dissimilar between replicates. Hence, some shifts in the elution profile can even be advantageous in distinguishing between true and coincidental co-elution. Among these five putative protein targets, LeuA, an enzyme involved in leucine synthesis, is a known isopropylmalate binder (Figure 2H). The remaining four proteins co-eluting with 2-isopropylmalic acid remain putative interactors, as it can't be excluded that more than one protein contributes to the 2-isopropylmalic acid elution profile. Based on the docking analysis, the prime candidates are purB and thrS with the predicted nM-μM binding affinities (Table S1).

Analogous analysis was performed for all 58 metabolites to generate a global map of the *E. coli* PMI network. Correlation data for each separation can be found in the Data S5A–S5D. The resulting network encompasses 51 metabolites, 465 protein nodes, and 994 PMIs (Figure 3A; Data S6). The 0.8 PCC cut-off to delineate putative interactors was determined by optimizing the true positive (sensitivity, true targets) to false positive (1 minus specificity, coincidental coelution) ratio (Figure 3B). This optimization was based on the evaluation of 1012 known interactions retrieved from the STITCH database (based on the experimental evidence, confidence score >0.4) between 32 metabolites and 430 proteins in our dataset. 26 metabolites in our network, such as 18 dipeptides and cyclic dipeptides, have no known protein targets deposited in STITCH. The network recapitulated 92 of the STITCH interactions (Figures 3C and S1; Table S2 and S3;), which is 7.4 times more than expected by chance, based on the results of the positive likelihood ratio (sensitivity/(1-specificity)) (Figure 3B). In comparison, using PROMIS or omniTICC alone captures approximately two times more STITCH interactions than expected by chance (Figure 3B).

Of note, although based on the positive likelihood ratio, we settled on the 0.8 PCC, in the past analogous analysis for the yeast[27,28] and Arabidopsis[16] PROMIS experiments favored the less stringent 0.7 PCC cut-off. Therefore, when looking for interactions, we prioritize these passing the 0.8 PCC but also consider the 0.7 PCC. Coordinates for both networks, referred to as 0.7 and 0.8 PCC networks, can be found in the Data S6A and S6B.

We also queried our network against the list of putative PMIs from two recent *E. coli* studies[6,9]; obtained using two different targeted, metabolite-centric methods. Diether et al.[9] investigated the binding of 55 metabolites to the 29 proteins from the central metabolism in a semi-throughput *in vitro*-based ligand-detected NMR screen. Piazza et al.[6] obtained their data using LiP-small molecule mapping (LiP-SMap) method focusing on 20 metabolites. The LiP-SMap method examines the change in the accessibility of a protein to a limited protease treatment in the presence of a ligand. One significant advantage of LiP-SMap is that the results can be used to model the ligand binding site and are performed under native conditions. The LiP-SMap dataset comprises 1678 PMIs for 20 metabolites. Of these 20 metabolites, valine, methionine, and GMP are in our dataset, and for these three compounds, Piazza et al.[6] reports 29 interactions: five for valine, five for GMP, and 19 for methionine. Two of the five valine interactors, ilvI and ilvN, are also co-eluting in our dataset, albeit with a lower cut-off of 0.7. Of note, for all three metabolites—valine, isoleucine, and GMP—we retrieved other known interactions (Figure S1). Hence, we speculate that the discrepancy between Piazza et al.[6] and our study points to the two methods favoring the detection of a different subset of interactions. However, the two methods start with cell lysate; they vary in several ways. For instance, unlike in CF-MS, in the LiP-SMap experiment, metabolites of interest are added in excess to the lysate, whereas endogenous metabolites are depleted by filtration. Moreover, the source of the starting material used to prepare cell lysate varied between the two studies, and the intracellular *E. coli* metabolome is known to be highly dynamic, which may further affect the retrieved PMIs. Analogously, of the 98 interactions reported by Diether, only 16 could potentially be recapitulated in our network, specifically for isoleucine, tryptophan, AMP, and FAD, due to the vastly different sets of studied metabolites. However, the 16 interactions are absent from our network; we did find other known interactions for tryptophan, AMP, and isoleucine (Figure S1). Hence, we argue that the discrepancy again points to the two methods favoring the detection of a different subset of interactions. The technique used by Diether et al.[9] relies on the *in vitro* binding between recombinant proteins and a ligand, which is very different from the CF-MS that operates in the complex cellular lysate. Possible reasons why CF-MS only retrieves a subset rather than all cellular PMIs are discussed in the conclusion section.

Next, to learn which compounds drive the positive likelihood ratio, in the 0.8 PCC network, nucleotide monophosphates (NMPs), nucleosides, amino acids, and cofactors were examined separately (Figure 3B). The PROMIS & omniTICC-derived PMI network outperformed PROMIS or omniTICC alone in recapitulating STITCH interactions for NMPs and amino acids but not for cofactors (Figure 3B). Neither the single nor combined approaches captured more nucleoside STITCH-retrieved targets than expected by chance. Close inspection of the data revealed that, unlike amino acids, NMPs, and nucleosides, the elution profiles of cofactors, particularly nicotinamide adenine dinucleotide

omniTICC, but just a single PROMIS separation. Asterisk indicates significance calculated using Fisher's exact test p-value <0.01, https://www.socscistatistics.com.

(C) Number of retrieved PMIs for the subset of metabolites with known interactions reported in STITCH (referred to as known PMIs). Red bar and number refer to the known interactions.

(D–I) Functional over-representation of the protein targets of the different compound classes was calculated and visualized using ShinyGO 0.77 app.[36]

(NADH) and flavin adenine dinucleotide (FAD), differed between the two PROMIS separations (Figures 2D and S2A). When we considered proteins cofractionating with cofactors in the two omniTICC separations but only in one of the two PROMIS experiments, we found that the positive likelihood ratio increased from 0.89 to 2.3 (Figure 2B); an improved performance compared to PROMIS or omniTICC alone. Hence, when metabolites elution profiles largely differ between replicates, our results indicate that they should be considered separately as driven by different proteins.

Functional analysis of the 465 protein nodes captured in the network revealed enrichment of categories broadly related to translation and cellular metabolism, particularly amino acid biosynthesis and carbon metabolism (Figure 3D). Specifically, the enrichment of proteins associated with translation was driven by interactions between NMPs and ribosomes (Figure 3E). While the role of NMPs in ribosome function remains to be clarified, it is worth noting that 2′,3′-cyclic nucleotides (2′,3′-cNMPs) were recently shown to inhibit translation by directly binding to the bacterial ribosome.[37] Moreover, ribosomal subunits were enriched among the putative interactors of 3′,5′-cAMP in Arabidopsis.[38] In contrast to NMPs, which clustered together, the five nucleosides: guanosine, adenosine, uridine, inosine, deoxyadenosine, and nucleotide hypoxanthine had distinct sets of putative targets (Figure 3F). Among these proteins are several TFs, proteins involved in RNA modifications, and enzymes of glycogen metabolism. In addition to acting as metabolic intermediates, nucleosides are known for their regulatory and signaling functions, and hence it is expected that they may have multiple and specific protein targets. For instance, hypoxanthine binding to the TF PurR enhances repression of its target genes, thereby affecting the metabolism and transport of purine and pyrimidine nucleotides.[39] The PuR and hypoxanthine interaction was also captured in PROMIS and omniTICC separations (Figure S2B).

Corroborating our previous work in plants and yeast,[16,27,28] the PMI network contained multiple proteinogenic dipeptides. With some exceptions, dipeptides are typically products of protein degradation.[40–42] Notably, of the 17 dipeptides in the network, 13 have at least one branched chain amino acid, valine, leucine, or isoleucine (Figure 3A). The various dipeptides, including the retro dipeptides such as Tyr-Leu and Leu-Tyr, have a unique set of putative targets distinct from those delineated for individual amino acids, including valine, leucine, or isoleucine (Figure 3A). Distinct activities of dipeptides and amino acids have been reported in previous studies.[40,43–47] Dipeptides were found in complexes with enzymes of central carbon metabolism in plants and yeast.[28,48] Our *E. coli* dipeptide-enzyme interaction network is enriched in enzymes of amino acid metabolism, particularly aromatic amino acids, and lysine biosynthesis (Figure 3G). Given that dipeptides act as a readout of nitrogen status and a source for amino acids, their association with amino acid synthesis is unsurprising.[49] Furthermore, the concept of end-product inhibition in amino acid synthesis, well-described in bacteria for regulating amino acid levels,[50] leads us to speculate that dipeptides may serve analogous functions. Additionally, dipeptide targets include dipeptide transporter (DppA) and oligopeptide transporters (OppA and SapA). Dipeptide regulation of peptide transport has been reported before in yeast.[51,52] Finally, the lists of putative targets for amino acids include expected categories such as amino acid biosynthesis enzymes (Figure 3H), whereas the molecular terms overrepresented among putative targets of co-factors include ‘‘pyridoxal phosphate’’ and ‘‘NADP binding’’ and encompass multiple enzymes of carbon and amino acid metabolism (Figure 3I).

These results demonstrate that integrating two chromatographic techniques in a co-fractionation mass spectrometry approach can enhance PMI mapping capabilities and that resulting PMI networks may provide broad functional insights.

Identification of regulatory PMIs in lipid and nucleotide metabolism

The *E. coli* PROMIS and omniTICC-derived interaction network presented here contains nearly a thousand putative interactions. One way to prioritize PMIs for binding and functional characterization is to select associations that may have escaped traditional binding screens usually biased toward well-established small molecule effectors. Dipeptides are good examples of compounds reported to have diverse, only recently uncovered bioactivities; however, they remain understudied.[40] Protein interactors of dipeptides in our network comprised multiple enzymes, also including enzymes associated with fatty acid biosynthesis. This observation may suggest a regulatory interplay between protein degradation and lipid metabolism. To follow up on this intriguing hypothesis, we selected a representative interaction between a dipeptide, Val-Leu, and FabF (Figure 4A), which catalyzes the conversion of palmitoyl-ACP to 3-keto-cis-vaccenoyl-ACP, a crucial step in the type II fatty acid elongation cycle.[53] Moreover, based on PCC FabF was the top co-eluting protein for Val-Leu. To investigate potential binding between FabF and Val-Leu, we employed microscale thermophoresis (MST). MST exploits the differential movement of the free ligand versus ligand-protein complex in a microscopic temperature gradient.[54] The measured binding constant (K_d) of 12–14 μM between FabF and Val-Leu (Figures 4B and 4C) falls within the range of reported dipeptide concentrations.[55] Only one binding pocket was predicted with high confidence for FabF, using computational binding site prediction (see STAR Methods), coinciding with the binding site for the co-crystallized ligand platencin in the PDB structural record for FabF (PDB-ID 3HO9) (Figures 4D–4F). We used this binding pocket for molecular docking analysis and predicted Val-Leu binding with binding affinity within the range of nM to μM, in reasonable agreement with the experimentally determined binding constant. In addition to FabF, the list of putative Val-Leu interactors comprises 31 more proteins, many associated with amino acid metabolism. We specifically tested SerC involved in serine metabolism but found no evidence for binding. This highlights the importance of follow-up experiments to identify true binders from co-elution profiles. It underscores the need for applying diverse analytical methods to cross-validate interactions and improve the resolution of PMI profiles, paving the way for more accurate characterization in future studies.

One strategy that can be employed to prioritize interactions builds on the known conservation of regulatory PMI.[56] Hence, we wondered if we could tap into previously obtained PROMIS

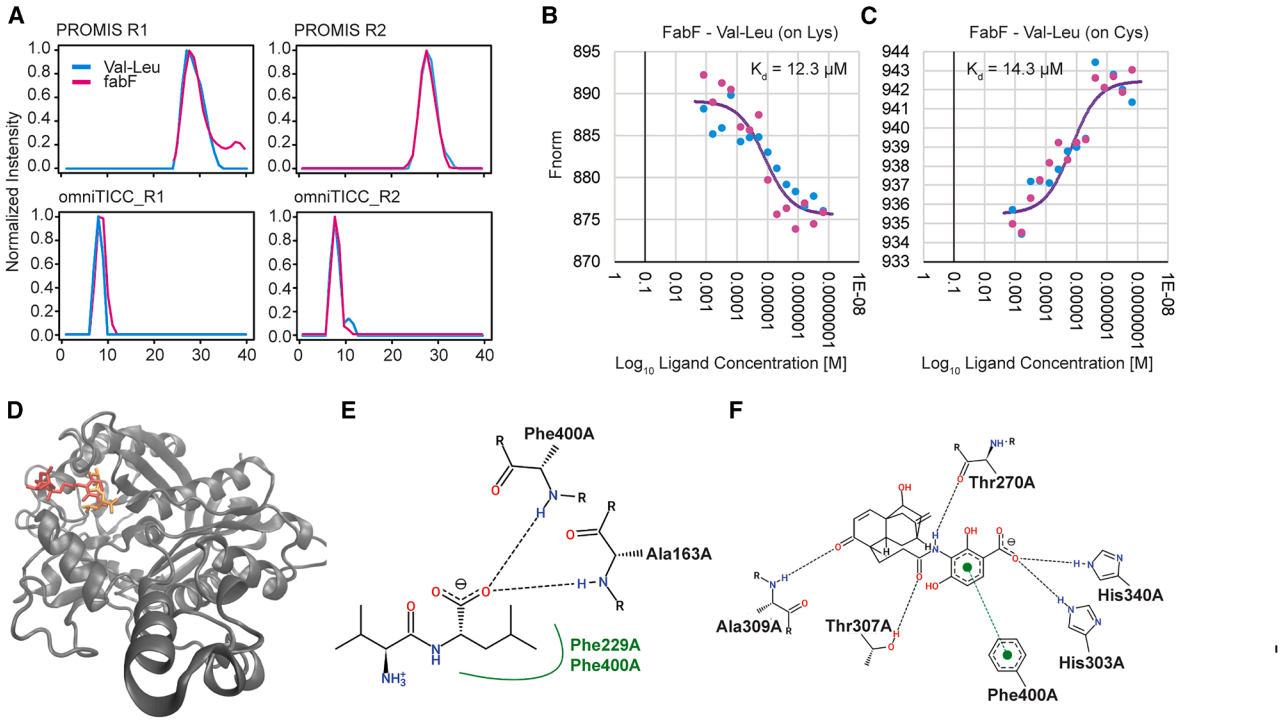


Figure 4. Binding events derived from the PMI network

(A) Co-elution profiles between dipeptide Val-Leu and enzyme FabF.

(B and C) MST binding curves; ligand corresponds to Val-Leu. The FabF protein was labeled using two different dyes, NHS dye that labels lysine and NHS-Maleimide dye that labels cysteine. Each binding curve is derived from two independent titrations and was fitted using Monolith analysis software.

(D–F) Potential interaction mode between FabF (PDB ID: 3HO9) and Val-Leu (PubChem ID: 6993118) obtained via molecular docking. (D) Ribbon diagram of the 3D-structure of FabF with the reference (co-crystallized in PDB file) ligand, platencin (PubChem ID: 16745128), in red, and the top docking pose of Val-Leu in orange. The predicted binding affinity for this top docking pose was estimated to lie in the nM to µM range, and within a nM range for platencin. (E and F) 2D ligand-protein interaction diagrams of FabF and Val-Leu (E) and platencin (F). The dashed lines represent hydrogen bonds, while the green spline segment represents hydrophobic contacts between Val-Leu and FabF.

datasets from plants and yeast to assist in prioritizing PMIs from *E. coli* based on presumed conservation. As a proof-of-concept example, we chose the photoconversion product of the cofactor riboflavin, lumichrome,[57] which we noted was present in the PROMIS separations from *E. coli*, yeast[28] and Arabidopsis.[16] Using a less stringent PCC cut-off of 0.7, as previously used with our yeast[28] and Arabidopsis studies,[16] we found that five of the 82 putative lumichrome targets in *E. coli* had a homolog in yeast and Arabidopsis that was also co-fractionating with lumichrome (Figure 5A). Of those five proteins, two had been previously associated with a related flavin molecule, the co-factor flavin mononucleotide (FMN) (Figure 5B), and we decided to focus on one of them, the enzyme orotate phosphoribosyltransferase (PyrE) (Figure 5C). Notably, the riboflavin and *de novo* pyrimidine synthesis pathways are biochemically connected, as they both rely on the intermediates of the pentose phosphate pathway. PyrE plays a pivotal role in *de novo* uridine monophosphate (UMP) biosynthesis by catalyzing a reaction between orotic acid and 5-phosphoribosyl-1-pyrophosphate (PRPP), ultimately yielding orotidine monophosphate (OMP).[59] In eukaryotes, the homolog of pyrE is called uridine monophosphate synthase (UMPS) and combines the activities of two enzymes, orotate phosphoribosyl transferase and orotidine-5′-decarboxylase. To test if lumichrome interacts with purified PyrE protein, we again employed MST. A control binding assay confirmed the interaction between PyrE and its well-known ligand PRPP, with a K_d value of 8.21 µM (Figure 5D), and we demonstrated binding between PyrE and lumichrome, with a binding affinity of 90.3 µM (Figure 5D). To further confirm PyrE interaction with lumichrome, we conducted a thermal shift assay (TSA). This method relies on the premise that ligand binding often influences the melting temperature of its protein target. We measured a small but reproducible shift in the melting temperature of PyrE according to an increase in the concentration of the ligands within the range of 74.17°C–74.76°C for PRPP and 74.17°C–74.91°C for lumichrome (Figure 5E). These observations support that lumichrome in fact binds to PyrE.

Using the PDB entry 6TAK (excluding the co-crystallized ligand orotate) as the target protein structure, *in silico* molecular docking to the pocket with the highest probability predicted lumichrome binding within an affinity range of µM to mM (Figures 5F–5H). Of note, while the predicted binding site encompasses the binding pocket of orotate, lumichrome was predicted to bind to the larger pocket adjacent to the minor pocket occupied by orotate. Interestingly, orotate and lumichrome share an interaction edge (O=C-NH-C=O, Figure 5F), suggesting a similar

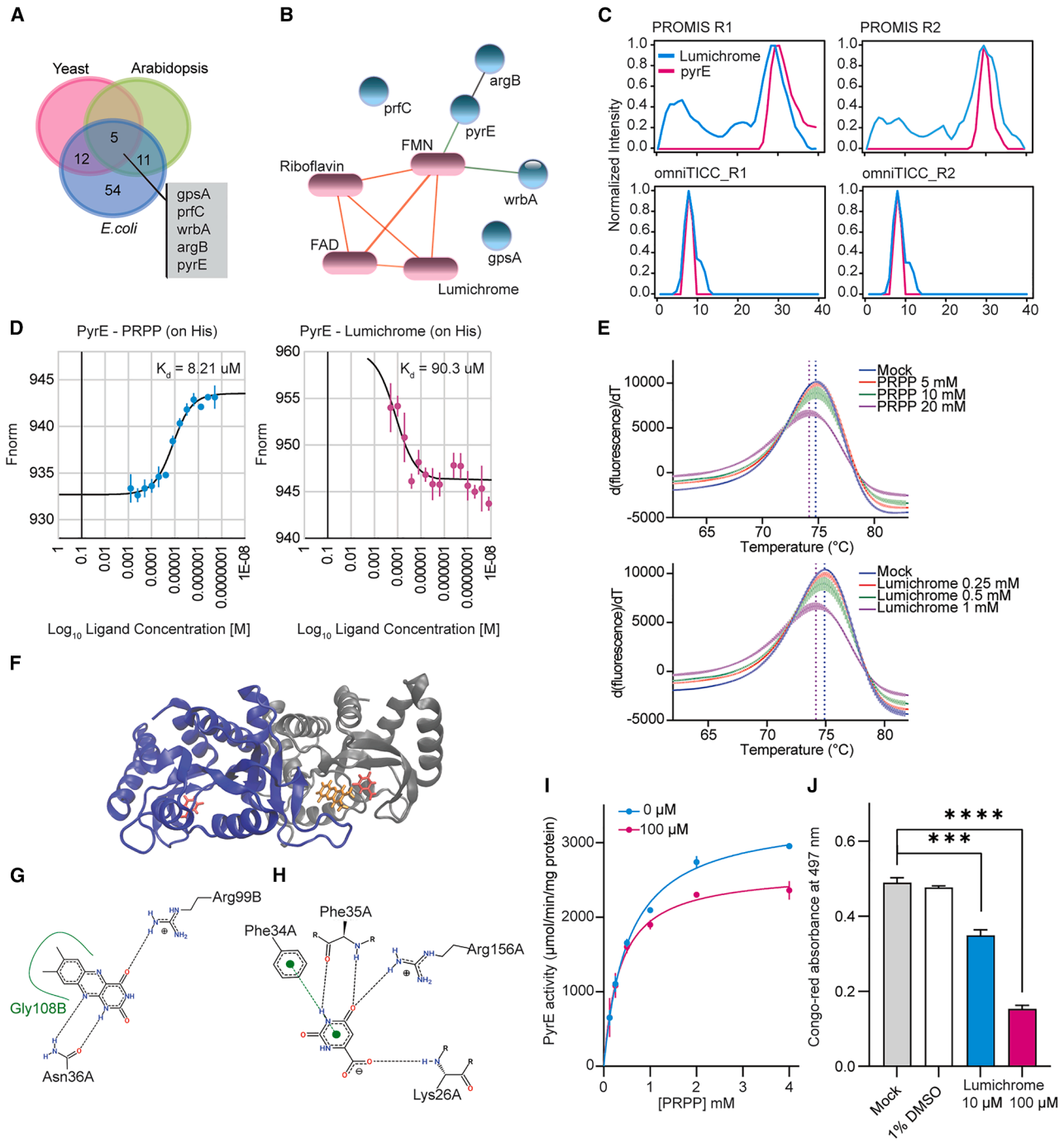


Figure 5. Binding events derived from the PMI network

(A) Venn representation focused on the overlap of *E. coli* putative lumichrome targets with proteins co-eluting with lumichrome in *A. thaliana*[16] and *S. cerevisiae*[28] in PROMIS experiments. Homology was derived using DIOPT ortholog prediction tool.[58]

(B) The STITCH derived interactions between the five proteins co-eluting with lumichrome in bacteria, yeast and plants with the different flavin compounds. Edges are based on the experimental, database and literature evidence with the confidence score >0.4.

(C) Co-elution profiles between lumichrome and enzyme PyrE.

(D) MST binding curves for PyrE-PRPP and PyE-lumichrome. PyrE was labeled at the position of His-tag. Each curve was generated from three independent titrations. Data are mean ± SE of $n = 3$.

(E) Melting curves of PyrE upon treatment with PRPP and lumichrome. Each curve was derived from three replicates. Data are mean ± SE of $n = 3$.

(F–H) Potential interaction mode between PyrE (PDB ID: 6TAK) and lumichrome (PubChem ID: 5326566) obtained via molecular docking. (F) Ribbon diagram of the 3D-structure of PyrE. Chain A of PyrE is shown in gray and chain B in blue. The reference (co-crystallized in PDB file) ligand, orotate (PubChem ID: 1492348),

(legend continued on next page)

binding mode. However, lumichrome is larger and does not fit into the orotate pocket. Visual inspection of the 3D structure reveals that a "lid" formed by a short beta-sheet segment may swing open, allowing lumichrome to bind in place of orotate. However, the available apo-structure (6TAI) does not show this lid as open, and enzymatic analysis (see below) argues against a competitive binding. An extended molecular docking simulation with a flexible protein backbone position will be necessary to investigate the binding details further.

After confirming the binding of lumichrome to PyrE, we investigated the impact of this binding on intrinsic PyrE activity. A continuous spectrophotometric method was used to monitor the activity of OPRTase in presence of PRPP and orotate. This method measures decreasing absorbance of orotate at 295 nm as OPRTase catalyzes conversion of orotate into OMP.[60] Enzyme activity reached saturation as the concentration of PRPP increased from 0.125 mM to 4 mM (Figure 5I). The presence of 100 μM of lumichrome reduced PyrE activity by approximately 20% at a concentration of 4 mM. The minimum tested lumichrome concentration that significantly decreased PyrE activity was 25 μM, whereas an increase in lumichrome concentration above 100 μM and up to 400 μM did not yield a stronger inhibitory effect (data not shown). The inhibitory effect of lumichrome on PyrE prompted us to investigate the effect of lumichrome on biofilm formation, since the *de novo* nucleotide biosynthesis pathway is closely linked to production of curli amyloid fibers, which constitute the extracellular matrix of biofilms in *E. coli*.[61] The operon encoding TF and genes associated with curli assembly and transport is strongly affected by inactivation of UMP biosynthetic genes, indicating a significant interplay between nucleotide metabolism and biofilm formation.[61] Notably, the inactivation of PyrE impairs curli production in *E. coli*, a phenomenon monitored through the binding between Congo-red dye and curli.[62] In our experiment, treatment of lumichrome resulted in the inhibitory effect on Congo-red binding in a concentration-dependent manner (Figure 5J). This observation is consistent with the reduced bacterial motility and biofilm formation in Δ*pyrE* mutant strain of *Pseudomonas aeruginosa* (PA01)[63] and *E. coli*.[64] Additionally, inhibition of biofilm formation in *E. coli* was observed with a decrease in PyrE activity due to acetylation at lysine 26 and 103 position.[64]

While the exact physiological significance requires further investigation, our findings indicate that the inhibitory effect of lumichrome on PyrE could contribute to the regulation of biofilm formation in bacteria. In this vein, previous studies have reported two noteworthy observations. Firstly, riboflavin and lumichrome have been identified as signaling molecules in bacterial quorum sensing (QS) and plant-microbe interactions.[65,66] Secondly, it has been reported that blocking the pyrimidine biosynthetic pathway affects riboflavin production in the filamentous hemiascomycete *Ashbya gossypii*.[67] Considering these findings, we hypothesize that lumichrome may function as a modulator for these interconnected pathways, including pyrimidine biosynthesis, riboflavin production, and QS, regulating the flux of metabolic intermediates. This regulation could be achieved through lumichrome's interaction with PyrE, potentially explaining the inhibition of biofilm formation in bacteria, as QS serves as an upstream regulatory component in the biofilm production process. Furthermore, based on the co-elution of lumichrome and UMPS in the PROMIS datasets from Arabidopsis and yeast, we speculate that the regulatory interplay between riboflavin pyrimidine nucleotide metabolism may be evolutionarily conserved. Notably, plants, fungi, and many bacteria encode the complete enzymatic machinery for *de novo* riboflavin and pyrimidine biosynthesis,[68,69] supporting our speculations.

In a broader sense, the lumichrome and pyrE/UMPS pairing is an excellent example of the main strength of using CF-MS for identifying PMIs, which is its untargeted nature. In contrast to, e.g., common amino acid and carbohydrate derivatives, more obscure vitamin catabolism products such as lumichrome would likely not be selected for a targeted PMI study.

DISCUSSION

Here, we present a PMI network for *E. coli* obtained from a CF-MS–based strategy integrating two orthogonal separations, SEC and IEX. The main advantage of CF-MS–based approaches to assessing PMIs is that they offer a direct and largely unbiased view of PMIs and retrieve interactions that likely would be missed by other methods, such as the lumichrome-PyrE pairing described here. Although lumichrome is a conserved metabolite in prokaryotic and eukaryotic organisms, its potential roles and mode of action are poorly understood. Given the importance of pyrimidine synthesis for organismal health and resilience, assessing the conservation of lumichrome regulation of uridine monophosphate synthase will be interesting. It is worth noting that, although the effect of lumichrome on the PyrE activity *in vitro* is relatively weak, it does translate into a significant phenotype *in vivo*. Our work further supports results from a recent LiP-MS study of metabolite interactions in the bacterial Calvin cycle.[70] Another example of an association uncovered by our PMI network is the interaction of leucine-containing dipeptides and the fatty acid metabolism enzyme, FabF. Potential roles of dipeptides regulating lipid metabolism are intriguing, and the FabF-Val-Leu binding data provide motivation for future analysis of the functional consequences of this interaction on FabF activity and fatty acid metabolism.

Recent guidelines for CF-MS experiments provide the community with valuable recommendations for experimental design and data analysis, especially for defining co-elution.[34,35]

and the top docking pose of lumichrome are shown in red and orange, respectively. The predicted binding affinity for this top docking pose was estimated to lie in the μM to mM range, and within a nM range for orotate. (G and H) 2D ligand-protein interaction diagrams of PyrE and lumichrome (G) and orotate (H). The dashed lines represent hydrogen bonds, while the green spline segment represents hydrophobic contacts between the lumichrome and PyrE.

(I) PyrE activity upon treatment with 100 μM of lumichrome. The activity was determined by calculating the amount of enzyme required to convert 1 μmol of orotic acid to orotidine monophosphate (OMP) per minute. Data are mean ± SE of $n = 3$.

(J) The inhibitory effect of lumichrome on Congo-red binding to biofilm components. Bacterial supernatant was measured at 497 nm after incubating with lumichrome for 24 h. Data are mean ± SE of $n = 3$. Significance was estimated using unpaired two-tailed Student's *t* test; ***: $p < 0.001$, ****: $p < 0.0001$.

Although these guidelines were originally developed for PPIs, they are similarly applicable to PMIs. The emerging consensus is that integrating different chromatographic techniques is a superior strategy for improving the recall of known complexes, as also demonstrated by our work. Here, we used the two most common chromatographic separations: size exclusion and IEX. However, other methods, such as hydrophobic chromatography, could also be implemented. Additionally, although it would require larger sample sizes and increase time requirements, an alternative to running two chromatographic separations in parallel would be to test two-dimensional (2D) separation, where a sample is first separated into multiple fractions using one chromatography (first dimension), then each of the resulting fractions is further separated using a different, orthogonal chromatography (second dimension). In addition to using orthogonal chromatographic separations, it was noted that analyzing replicates separately produces better results than calculating an average. The degree of dissimilarity between replicates may even be advantageous in distinguishing between true and coincidental co-elution. However, as demonstrated here for the cases of FAD and NADH, when elution profiles differ greatly, it is necessary to consider the possibility that they are driven by different PMIs. In the current study, each separation was performed in duplicate, which was principally limited by the large number of proteomics and metabolomics measurements required for each separation. An additional replicate would certainly be advantageous for resolving protein targets of compounds with dissimilar elution profiles, such as NADH or FAD, but the gains for many other compounds with more reproducible elution would likely be incremental. For data analysis, the Pearson correlation is regarded as a well-suited metric for discerning co-elution.[34,35] Herein, we used STITCH-derived interactions to calculate the PCC cut-off, which resulted in the best positive likelihood ratio. However, we also present a less stringent network obtained using the 0.7 PCC that we used previously in yeast[28] and Arabidopsis.[16] In comparison to PPIs, there are significantly fewer reported PMIs, which makes it more challenging to accurately infer the best cut-offs. We anticipate that this will change in the future as the field progresses.

As pointed out before, although we refer to the PROMIS and omniTICC-derived interactions as high-confidence compared to those derived from a single chromatography, all inferred PMIs are obviously putative and require independent validation. Because validation is time-consuming, prioritization is a critical step, as in any interactomics study. Prioritization may be based on prior knowledge or additional experiments. Here, we explored co-elution across multiple PROMIS datasets and the conserved nature of small-molecule binding events. A different, highly promising avenue is *in silico* modeling of the binding events.[71,72] We anticipate that binding predictions will become instrumental in refining CF-MS-derived PMI networks and selecting true binding events. The astounding success of AlphaFold[73] and that of similar such programs,[74,75] combined with the development of novel machine-learning methods, e.g., Graph Neural Networks,[76,77] Attention mechanisms,[78,79] have catapulted the field of structure-based computational studies of compound-protein binding events into a new era. Despite these breakthroughs, faithfully estimating binding energies and discerning detailed binding modes remains challenging.[80,81] However, as demonstrated here, presently available methods can already deliver corroborating evidence supporting experimental data and can be applied as an additional filter. Without a doubt, computational binding predictions will become even more instrumental in refining CF-MS-derived PMI networks and selecting true binding events.

Limitations of the study

As mentioned above, the primary advantage of using CF-MS to study PMIs is its untargeted and exploratory nature. Hundreds of metabolic features separate with protein complexes in a single experiment. Our datasets revealed co-fractionation of hundreds of metabolic features with specific proteins or protein complexes and, beyond the small subset of compounds that we succeeded in annotating, contain many more unknowns that await chemical characterization. However, CF-MS-based approaches also have important limitations. False positives may arise from coincidental co-elution, and CF-MS is biased toward stable complexes and will likely miss transient interactions. CF-MS will also favor abundant proteins and metabolites accessible to MS identification and quantification. Integrating different types of chromatography can help distinguish between true and coincidental co-elution, but it is important to note that this approach may miss interactions favored by one chromatographic condition. For example, we noted that IEX datasets contained metabolites absent in the PROMIS separations that, therefore, were excluded from the downstream analysis. Hence, it is unsurprising that we recapitulated only a subset of the STITCH-derived PMIs, analogously to the poor recall of previously reported PMIs.[6,9] Moreover, since the PMI landscape is highly dynamic and affected by multiple developmental and environmental cues, a single CF-MS experiment can only recapitulate the subset of PMIs relevant to one particular biological state.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Aleksandra Skirycz (skirycza@msu.edu).

Materials availability

Plasmids generated in this study are available from the lead contact upon request.

Data and code availability

- Proteomics data have been deposited in MassIVE repository,[82] with accession numbers MassIVE: MSV000093709; MassIVE: MSV000093733; MassIVE: MSV000094003; MassIVE: MSV000094057. Metabolomics data have been deposited in MetaboLights repository,[83] with accession number MetaboLights: MTBLS9098. All datasets are publicly available as of the date of publication. The accession numbers are listed in the key resources table.
- This article does not report original code. A web-based tool, PROMISed, used to analyze and visualize the PMI networks is available at https://github.com/DennisSchlossarek/PROMISed.
- Processed proteomics and metabolomics data are available in supplementary material. Any additional information required to reanalyze the data reported in this article is available from the lead contact upon request.

ACKNOWLEDGMENTS

We are grateful to Änne Michaelis for excellent technical assistance. The authors would like to acknowledge the support from Boyce Thompson Institute for Plant Research, Cornell University, Michigan State University, and Max Planck Society. We are thankful to Biogenity and DTU Proteomics Core for performing the proteomic sample measurements and contribution to data analysis. This work was supported by the U.S. National Science Foundation (grant 2427055 awarded to A.S.).

AUTHOR CONTRIBUTIONS

M.W. designed the experiments, conducted experiments. J.K. designed the experiments, conducted experiments, and wrote the manuscript. C.M. conducted experiments. V.P.T. performed proteomics measurements and analyzed proteomics data. M.G. analyzed proteomics data. R.I.M. established the FPLC protein purification method. C.F.P. prepared proteomics samples. J. Z. and K.D. provided FabF protein expression construct and revised the manuscript. F.C.S. assisted in data analysis and writing the manuscript. H.Z. and D. W. performed docking analysis. A.S. designed experiments, analyzed data, wrote the manuscript, and supervised the study.

DECLARATION OF INTERESTS

F.C.S. is a founder of, consultant, and a stockholder for Ascribe Bioscience and Holoclara Inc. Other than that, the authors declare that they have no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used Grammarly to edit/improve the English. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODELS AND STUDY PARTICIPATION DETAILS
- METHOD DETAILS
  - Size exclusion chromatography (SEC)
  - Ion-exchange chromatography (IEX)
  - Metabolite and protein extraction
  - Proteomics
  - Metabolomics
  - Data analysis
  - Protein purification
  - Microscale thermophoresis (MST)
  - Thermal shift assay (TSA)
  - Orotate phosphoribosyltransferase assay
  - Congo-red binding assay
  - Molecular docking
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.112611.

Received: July 11, 2024
Revised: January 21, 2025
Accepted: May 5, 2025
Published: May 8, 2025

REFERENCES

1. Venegas-Molina, J., Molina-Hidalgo, F.J., Clicque, E., and Goossens, A. (2021). Why and How to Dig into Plant Metabolite-Protein Interactions. Trends Plant Sci. *26*, 472–483.
2. Diether, M., and Sauer, U. (2017). Towards detecting regulatory protein–metabolite interactions. Curr. Opin. Microbiol. *39*, 16–23.
3. Kosmacz, M., Sokołowska, E.M., Bouzaa, S., and Skirycz, A. (2020). Towards a functional understanding of the plant metabolome. Curr. Opin. Plant Biol. *55*, 47–51.
4. Baker, S.A., and Rutter, J. (2023). Metabolites as signalling molecules. Nat. Rev. Mol. Cell Biol. *24*, 355–374.
5. Ledezma Tejeida, D.E., Schastnaya, E., and Sauer, U. (2021). Metabolism as a signal generator in bacteria. Curr. Opin. Syst. Biol. *28*, 100404. https://doi.org/10.3929/ETHZ-B-000517650.
6. Piazza, I., Kochanowski, K., Cappelletti, V., Fuhrer, T., Noor, E., Sauer, U., and Picotti, P. (2018). A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. Cell *172*, 358–372.e23.
7. Tian, M., von Dahl, C.C., Liu, P.-P., Friso, G., van Wijk, K.J., and Klessig, D. F. (2012). The combined use of photoaffinity labeling and surface plasmon resonance-based technology identifies multiple salicylic acid-binding proteins. Plant J. *72*, 1027–1038.
8. Lim, Y.T., Prabhu, N., Dai, L., Go, K.D., Chen, D., Sreekumar, L., Egeblad, L., Eriksson, S., Chen, L., Veerappan, S., et al. (2018). An efficient proteome-wide strategy for discovery and characterization of cellular nucleotide-protein interactions. PLoS One *13*, e0208273.
9. Diether, M., Nikolaev, Y., Allain, F.H., and Sauer, U. (2019). Systematic mapping of protein-metabolite interactions in central metabolism of Escherichia coli. Mol. Syst. Biol. *15*, e9008.
10. Link, H., Kochanowski, K., and Sauer, U. (2013). Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. Nat. Biotechnol. *31*, 357–361.
11. Hackett, S.R., Zanotelli, V.R.T., Xu, W., Goya, J., Park, J.O., Perlman, D.H., Gibney, P.A., Botstein, D., Storey, J.D., and Rabinowitz, J.D. (2016). Systems-level analysis of mechanisms regulating yeast metabolic flux. Science *354*, aaf2786. https://doi.org/10.1126/science.aaf2786.
12. Lempp, M., Farke, N., Kuntz, M., Freibert, S.A., Lill, R., and Link, H. (2019). Systematic identification of metabolites controlling gene expression in E. coli. Nat. Commun. *10*, 4463.
13. Wagner, M., Zhang, B., Tauffenberger, A., Schroeder, F.C., and Skirycz, A. (2021). Experimental methods for dissecting the terra-incognita of protein-metabolite interactomes. Curr. Opin. Syst. Biol. *28*, 100403.
14. Luzarowski, M., and Skirycz, A. (2019). Emerging strategies for the identification of protein–metabolite interactions. J. Exp. Bot. *70*, 4605–4618.
15. Orsak, T., Smith, T.L., Eckert, D., Lindsley, J.E., Borges, C.R., and Rutter, J. (2012). Revealing the Allosterome: Systematic Identification of Metabolite–Protein Interactions. Biochemistry *51*, 225–232.
16. Veyel, D., Sokolowska, E.M., Moreno, J.C., Kierszniowska, S., Cichon, J., Wojciechowska, I., Luzarowski, M., Kosmacz, M., Szlachetko, J., Gorka, M., et al. (2018). PROMIS, global analysis of PROtein–metabolite interactions using size separation in Arabidopsis thaliana. J. Biol. Chem. *293*, 12440–12453.
17. Schlossarek, D., Zhang, Y., Sokolowska, E.M., Fernie, A.R., Luzarowski, M., and Skirycz, A. (2023). Don't let go: co-fractionation mass spectrometry for untargeted mapping of protein-metabolite interactomes. Plant J. *113*, 904–914.
18. Skinnider, M.A., Scott, N.E., Prudova, A., Kerr, C.H., Stoynov, N., Stacey, R.G., Chan, Q.W.T., Rattray, D., Gsponer, J., and Foster, L.J. (2021). An atlas of protein-protein interactions across mouse tissues. Cell *184*, 4073–4089.e17.
19. Wan, C., Borgeson, B., Phanse, S., Tu, F., Drew, K., Clark, G., Xiong, X., Kagan, O., Kwan, J., Bezginov, A., et al. (2015). Panorama of ancient metazoan macromolecular complexes. Nature *525*, 339–344.

20. Mallam, A.L., Sae-Lee, W., Schaub, J.M., Tu, F., Battenhouse, A., Jang, Y. J., Kim, J., Wallingford, J.B., Finkelstein, I.J., Marcotte, E.M., and Drew, K. (2019). Systematic Discovery of Endogenous Human Ribonucleoprotein Complexes. Cell Rep. *29*, 1351–1368.e5.

21. Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A census of human soluble protein complexes. Cell *150*, 1068–1081.

22. Lee, Y., and Szymanski, D.B. (2021). Multimerization variants as potential drivers of neofunctionalization. Sci. Adv. *7*, eabf0984. https://doi.org/10.1126/sciadv.abf0984.

23. Lee, Y., Okita, T.W., and Szymanski, D.B. (2021). A co-fractionation mass spectrometry-based prediction of protein complex assemblies in the developing rice aleurone-subaleurone. Plant Cell *33*, 2965–2980.

24. Aryal, U.K., Xiong, Y., McBride, Z., Kihara, D., Xie, J., Hall, M.C., and Szymanski, D.B. (2014). A proteomic strategy for global analysis of plant protein complexes. Plant Cell *26*, 3867–3882.

25. McWhite, C.D., Papoulas, O., Drew, K., Cox, R.M., June, V., Dong, O.X., Kwon, T., Wan, C., Salmi, M.L., Roux, S.J., et al. (2020). A Pan-plant Protein Complex Map Reveals Deep Conservation and Novel Assemblies. Cell *181*, 460–474.e14.

26. Veyel, D., Kierszniowska, S., Kosmacz, M., Sokolowska, E.M., Michaelis, A., Luzarowski, M., Szlachetko, J., Willmitzer, L., and Skirycz, A. (2017). System-wide detection of protein-small molecule complexes suggests extensive metabolite regulation in plants. Sci. Rep. *7*, 42387.

27. Schlossarek, D., Luzarowski, M., Sokołowska, E.M., Thirumalaikumar, V. P., Dengler, L., Willmitzer, L., Ewald, J.C., and Skirycz, A. (2022). Rewiring of the protein–protein–metabolite interactome during the diauxic shift in yeast. Cell Mol. Life Sci. *79*, 550. https://doi.org/10.1007/s00018-022-04569-8.

28. Luzarowski, M., Vicente, R., Kiselev, A., Wagner, M., Schlossarek, D., Erban, A., de Souza, L.P., Childs, D., Wojciechowska, I., Luzarowska, U., et al. (2021). Global mapping of protein–metabolite interactions in Saccharomyces cerevisiae reveals that Ser-Leu dipeptide regulates phosphoglycerate kinase activity. Commun. Biol. *4*, 181.

29. Li, Y., Kuhn, M., Zukowska-Kasprzyk, J., Hennrich, M.L., Kastritis, P.L., O'Reilly, F.J., Phapale, P., Beck, M., Gavin, A.-C., and Bork, P. (2021). Coupling proteomics and metabolomics for the unsupervised identification of protein–metabolite interactions in Chaetomium thermophilum. PLoS One *16*, e0254429.

30. Chan, J.N.Y., Vuckovic, D., Sleno, L., Olsen, J.B., Pogoutse, O., Havugimana, P., Hewel, J.A., Bajaj, N., Wang, Y., Musteata, M.F., et al. (2012). Target identification by chromatographic co-elution: monitoring of drug-protein interactions without immobilization or chemical derivatization. Mol. Cell. Proteomics *11*, 016642.

31. Schäkermann, S., Wüllner, D., Yayci, A., Emili, A., and Bandow, J.E. (2021). Applicability of Chromatographic Co-Elution for Antibiotic Target Identification. Proteomics *21*, e2000038.

32. Kuhn, M., Szklarczyk, D., Franceschini, A., Campillos, M., von Mering, C., Jensen, L.J., Beyer, A., and Bork, P. (2010). STITCH 2: an interaction network database for small molecules and proteins. Nucleic Acids Res. *38*, D552–D556.

33. Schlossarek, D., Luzarowski, M., Sokołowska, E., Górka, M., Willmitzer, L., and Skirycz, A. (2021). PROMISed: A novel web-based tool to facilitate analysis and visualization of the molecular interaction networks from co-fractionation mass spectrometry (CF-MS) experiments. Comput. Struct. Biotechnol. J. *19*, 5117–5125.

34. Skinnider, M.A., and Foster, L.J. (2021). Meta-analysis defines principles for the design and analysis of co-fractionation mass spectrometry experiments. Nat. Methods *18*, 806–815.

35. Pang, C.N.I., Ballouz, S., Weissberger, D., Thibaut, L.M., Hamey, J.J., Gillis, J., Wilkins, M.R., and Hart-Smith, G. (2020). Analytical Guidelines for co-fractionation Mass Spectrometry Obtained through Global Profiling of Gold Standard Saccharomyces cerevisiae Protein Complexes. Mol. Cell. Proteomics *19*, 1876–1895.

36. Ge, S.X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics *36*, 2628–2629.

37. Chauhan, S.S., Marotta, N.J., Karls, A.C., and Weinert, E.E. (2022). Binding of 2',3'-Cyclic Nucleotide Monophosphates to Bacterial Ribosomes Inhibits Translation. ACS Cent. Sci. *8*, 1518–1526.

38. Chodasiewicz, M., Fuentealba, N.F., Franz, P., Luzarowski, M., Moreno, J. C., Childs, D., Masiuk, A., Saluja, M., Sampathkumar, A., and Skirycz, A. (2022). A novel role of 3,5-cAMP in the regulation of actin cytoskeleton in Arabidopsis. Preprint at bioRxiv. https://doi.org/10.1101/2022.01.31.478439.

39. Cho, B.-K., Federowicz, S.A., Embree, M., Park, Y.-S., Kim, D., and Palsson, B.Ø. (2011). The PurR regulon in Escherichia coli K-12 MG1655. Nucleic Acids Res. *39*, 6456–6464.

40. Minen, R.I., Thirumalaikumar, V.P., and Skirycz, A. (2023). Proteinogenic dipeptides, an emerging class of small-molecule regulators. Curr. Opin. Plant Biol. *75*, 102395.

41. Thirumalaikumar, V.P., Wagner, M., Balazadeh, S., and Skirycz, A. (2021). Autophagy is responsible for the accumulation of proteogenic dipeptides in response to heat stress in Arabidopsis thaliana. FEBS J. *288*, 281–292. https://doi.org/10.1111/febs.15336.

42. Tsukahara, T., Yamagishi, S., Neyama, H., and Ueda, H. (2018). Tyrosyl-tRNA synthetase: A potential kyotorphin synthetase in mammals. Peptides *101*, 60–68. https://doi.org/10.1016/j.peptides.2017.12.026.

43. Cheng, L., Tanaka, M., Yoshino, A., Nagasato, Y., Takata, F., Dohgu, S., and Matsui, T. (2023). A memory-improving dipeptide, Tyr-Pro, can reach the mouse brain after oral administration. Sci. Rep. *13*, 16908.

44. Ichinose, T., Moriyasu, K., Nakahata, A., Tanaka, M., Matsui, T., and Furuya, S. (2015). Orally administrated dipeptide Ser-Tyr efficiently stimulates noradrenergic turnover in the mouse brain. Biosci. Biotechnol. Biochem. *79*, 1542–1547.

45. Mizushige, T., Uchida, T., and Ohinata, K. (2020). Dipeptide tyrosyl-leucine exhibits antidepressant-like activity in mice. Sci. Rep. *10*, 2257.

46. Zhang, Z., Zhao, Y., Wang, X., Lin, R., Zhang, Y., Ma, H., Guo, Y., Xu, L., and Zhao, B. (2016). The novel dipeptide Tyr-Ala (TA) significantly enhances the lifespan and healthspan of Caenorhabditis elegans. Food Funct. *7*, 1975–1984.

47. Naka, K., Jomen, Y., Ishihara, K., Kim, J., Ishimoto, T., Bae, E.-J., Mohney, R.P., Stirdivant, S.M., Oshima, H., Oshima, M., et al. (2015). Dipeptide species regulate p38MAPK-Smad3 signalling to maintain chronic myelogenous leukaemia stem cells. Nat. Commun. *6*, 8039.

48. Moreno, J.C., Rojas, B.E., Vicente, R., Gorka, M., Matz, T., Chodasiewicz, M., Peralta-Ariza, J.S., Zhang, Y., Alseekh, S., Childs, D., et al. (2021). Tyr-Asp inhibition of glyceraldehyde 3-phosphate dehydrogenase affects plant redox metabolism. EMBO J. *40*, e106800.

49. Tegeder, M., and Rentsch, D. (2010). Uptake and partitioning of amino acids and peptides. Mol. Plant *3*, 997–1011.

50. Chubukov, V., Gerosa, L., Kochanowski, K., and Sauer, U. (2014). Coordination of microbial metabolism. Nat. Rev. Microbiol. *12*, 327–340.

51. Xia, Z., Turner, G.C., Hwang, C.-S., Byrd, C., and Varshavsky, A. (2008). Amino acids induce peptide uptake via accelerated degradation of CUP9, the transcriptional repressor of the PTR2 peptide transporter. J. Biol. Chem. *283*, 28958–28968.

52. Du, F., Navarro-Garcia, F., Xia, Z., Tasaki, T., and Varshavsky, A. (2002). Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain. Proc. Natl. Acad. Sci. USA *99*, 14110–14115.

53. Edwards, P., Nelsen, J.S., Metz, J.G., and Dehesh, K. (1997). Cloning of the fabF gene in an expression vector and in vitro characterization of recombinant fabF and fabB encoded enzymes from Escherichia coli. FEBS Lett. *402*, 62–66.

54. Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. Assay Drug Dev. Technol. *9*, 342–353.

55. Heidenreich, E., Pfeffer, T., Kracke, T., Mechtel, N., Nawroth, P., Hoffmann, G.F., Schmitt, C.P., Hell, R., Poschet, G., and Peters, V. (2021). A Novel UPLC-MS/MS Method Identifies Organ-Specific Dipeptide Profiles. Int. J. Mol. Sci. *22*, 9979. https://doi.org/10.3390/ijms22189979.

56. Gruber, C.H., Diether, M., and Sauer, U. (2021). Conservation of metabolic regulation by phosphorylation and non-covalent small-molecule interactions. Cell Syst. *12*, 538–546. https://doi.org/10.1016/j.cels.2021.04.009.

57. Treadwell, G.E., and Metzler, D.E. (1972). Photoconversion of riboflavin to lumichrome in plant tissues. Plant Physiol. *49*, 991–993.

58. Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., and Mohr, S.E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinf. *12*, 357.

59. Kim, M.-K., Song, H.-E., Kim, Y.S., Rho, S.-H., Im, Y.J., Lee, J.H., Kang, G. B., and Eom, S.H. (2003). Crystallization and preliminary X-ray crystallographic analysis of orotate phosphoribosyltransferase from Helicobacter pylori. Mol. Cells *15*, 361–363.

60. Scapin, G., Ozturk, D.H., Grubmeyer, C., and Sacchettini, J.C. (1995). The crystal structure of the orotate phosphoribosyltransferase complexed with orotate and alpha-D-5-phosphoribosyl-1-pyrophosphate. Biochemistry *34*, 10744–10754.

61. Garavaglia, M., Rossi, E., and Landini, P. (2012). The pyrimidine nucleotide biosynthetic pathway modulates production of biofilm determinants in Escherichia coli. PLoS One *7*, e31252.

62. Hammar, M., Arnqvist, A., Bian, Z., Olsén, A., and Normark, S. (1995). Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in Escherichia coli K-12. Mol. Microbiol. *18*, 661–670.

63. Niazy, A.A., Lambarte, R.N.A., and Alghamdi, H.S. (2022). de novo pyrimidine synthesis pathway inhibition reduces motility virulence of Pseudomonas aeruginosa despite complementation. J. King Saud Univ. Sci. *34*, 102040.

64. Lozano-Terol, G., Gallego-Jara, J., Sola-Martínez, R.A., Ortega, Á., Martínez Vivancos, A., Cánovas Díaz, M., and de Diego Puente, T. (2023). Regulation of the pyrimidine biosynthetic pathway by lysine acetylation of E. coli OPRTase. FEBS J. *290*, 442–464.

65. Rajamani, S., Bauer, W.D., Robinson, J.B., Farrow, J.M., 3rd, Pesci, E.C., Teplitski, M., Gao, M., Sayre, R.T., and Phillips, D.A. (2008). The vitamin riboflavin and its derivative lumichrome activate the LasR bacterial quorum-sensing receptor. Mol. Plant Microbe Interact. *21*, 1184–1192.

66. Dakora, F.D., Matiru, V.N., and Kanu, A.S. (2015). Rhizosphere ecology of lumichrome and riboflavin, two bacterial signal molecules eliciting developmental changes in plants. Front. Plant Sci. *6*, 700.

67. Silva, R., Aguiar, T.Q., Oliveira, C., and Domingues, L. (2019). Physiological characterization of a pyrimidine auxotroph exposes link between uracil phosphoribosyltransferase regulation and riboflavin production in Ashbya gossypii. N. Biotechnol. *50*, 1–8.

68. Averianova, L.A., Balabanova, L.A., Son, O.M., Podvolotskaya, A.B., and Tekutyeva, L.A. (2020). Production of Vitamin B2 (Riboflavin) by Microorganisms: An Overview. Front. Bioeng. Biotechnol. *8*, 570828.

69. Fischer, M., Römisch, W., Saller, S., Illarionov, B., Richter, G., Rohdich, F., Eisenreich, W., and Bacher, A. (2004). Evolution of vitamin B2 biosynthesis: structural and functional similarity between pyrimidine deaminases of eubacterial and plant origin. J. Biol. Chem. *279*, 36299–36308.

70. Sporre, E., Karlsen, J., Schriever, K., Asplund-Samuelsson, J., Janasch, M., Strandberg, L., Karlsson, A., Kotol, D., Zeckey, L., Piazza, I., et al. (2023). Metabolite interactions in the bacterial Calvin cycle and implications for flux regulation. Commun. Biol. *6*, 947.

71. Thieme, S., and Walther, D. (2022). Biclique extension as an effective approach to identify missing links in metabolic compound–protein interaction networks. Bioinform. Adv. *2*, vbac001. https://doi.org/10.1093/bioadv/vbac001.

72. Walther, D. (2023). Specifics of Metabolite-Protein Interactions and Their Computational Analysis and Prediction. Methods Mol. Biol. *2554*, 179–197.

73. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.

74. Verkuil, R., Kabeli, O., Du, Y., Wicky, B.I.M., Milles, L.F., Dauparas, J., Baker, D., Ovchinnikov, S., Sercu, T., and Rives, A. (2022). Language models generalize beyond natural proteins. Preprint at bioRxiv, 521521. https://doi.org/10.1101/2022.12.21.521521.

75. Wu, R., Ding, F., Wang, R., Shen, R., Zhang, X., Luo, S., Su, C., Wu, Z., Xie, Q., Berger, B., et al. (2022). High-resolution de novo structure prediction from primary sequence. Preprint at bioRxiv, 500999. https://doi.org/10.1101/2022.07.21.500999.

76. Nguyen, T., Le, H., Quinn, T.P., Nguyen, T., Le, T.D., and Venkatesh, S. (2021). GraphDTA: predicting drug-target binding affinity with graph neural networks. Bioinformatics *37*, 1140–1147.

77. Nguyen, N.-Q., Jang, G., Kim, H., and Kang, J. (2023). Perceiver CPI: a nested cross-attention network for compound-protein interaction prediction. Bioinformatics *39*, btac731. https://doi.org/10.1093/bioinformatics/btac731.

78. Zhao, Q., Duan, G., Yang, M., Cheng, Z., Li, Y., and Wang, J. (2023). AttentionDTA: Drug-Target Binding Affinity Prediction by Sequence-Based Deep Learning With Attention Mechanism. IEEE/ACM Trans. Comput. Biol. Bioinform. *20*, 852–863.

79. Campana, P.A., and Nikoloski, Z. (2023). Self- and cross-attention accurately predicts metabolite-protein interactions. NAR Genom. Bioinform. *5*, lqad008.

80. Li, Y., Fan, Z., Rao, J., Chen, Z., Chu, Q., Zheng, M., and Li, X. (2023). An overview of recent advances and challenges in predicting compound-protein interaction (CPI). Med. Rev. *3*, 465–486.

81. Basciu, A., Callea, L., Motta, S., Bonvin, A.M.J.J., Bonati, L., and Vargiu, A. V. (2022). Chapter Two - No dance, no partner! A tale of receptor flexibility in docking and virtual screening. In Annual Reports in Medicinal Chemistry, J. Caballero, ed. (Academic Press), pp. 43–97.

82. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. *50*, D543–D552.

83. Yurekten, O., Payne, T., Tejera, N., Amaladoss, F.X., Martin, C., Williams, M., and O'Donovan, C. (2024). MetaboLights: open data repository for metabolomics. Nucleic Acids Res. *52*, D640–D646.

84. Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005). Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. DNA Res. *12*, 291–299.

85. Sokolowska, E.M., Schlossarek, D., Luzarowski, M., and Skirycz, A. (2019). PROMIS: Global analysis of PROtein-metabolite interactions. Curr. Protoc. Plant Biol. *4*, e20101.

86. Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

87. Thirumalaikumar, V.P., Fernie, A.R., and Skirycz, A. (2023). Untargeted Proteomics and Metabolomics Analysis of Plant Organ Development. Methods Mol. Biol. *2698*, 75–85.

88. Guo, J., Van De Ven, W.T., Skirycz, A., Thirumalaikumar, V.P., Zeng, L., Zhang, Q., Balcke, G.U., Tissier, A., and Dehesh, K. (2024). An evolutionarily conserved metabolite inhibits biofilm formation in Escherichia coli K-12. Nat. Commun. *15*, 10079.

89. Shimosaka, M., Fukuda, Y., Murata, K., and Kimura, A. (1985). Purification and properties of orotate phosphoribosyltransferases from Escherichia coli K-12, and its derivative purine-sensitive mutant. J. Biochem. *98*, 1689–1697.

90. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

91. Jakubec, D., Skoda, P., Krivak, R., Novotny, M., and Hoksza, D. (2022). PrankWeb 3: accelerated ligand-binding site predictions for experimental and modelled protein structures. Nucleic Acids Res. *50*, W593–W597.

92. Fricker, P.C., Gastreich, M., and Rarey, M. (2004). Automated drawing of structural molecular formulas under constraints. J. Chem. Inf. Comput. Sci. *44*, 1065–1078.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli strain BW25113	Coli Genetics Stock Center	CGSC#: 7636
Escherichia coli BL21 (DE)	NEB	C2527I
Escherichia coli W3110S	ATCC	ATCC 27325
Escherichia coli strain K-12 expressing PyrE	ASKA collection	JW1081-AM
Chemicals, peptides, and recombinant proteins		
Protease inhibitor cocktail	Sigma Aldrich	Cat#P9599
Trypsin/Lys-C	Promega	Cat#V5073
MST RED-NHS dye	Nanotemper Technologies	Cat#MO-L011
MST RED-MALEIMIDE dye	Nanotemper Technologies	Cat#MO-L014
Protein thermal shift dye	Applied Biosystem	Cat# 4461146
Critical commercial assays		
Size exclusion chromatography column	Sepax	SRT-10 SEC-300
Ion-exchange chromatography column	Bio-Rad	Enrich Q, 5/50 mm
C18 Sep-Pak column plate	Waters	Cat#186003966
His-Bind purification kit	Merck	Cat#70239-M
HisTrap column	Cytiva	Cat#29051021
HSS T3 C18 reversed-phase column	Waters	Cat#186003539
C18 Acclaim Pepmap column	Thermo Fisher	Cat#164570
Deposited data		
Proteomics	MassIVE	MSV000093709; https://doi.org/10.25345/C5VT1H131
Proteomics	MassIVE	MSV000093733; https://doi.org/10.25345/C5RX93Q6M
Proteomics	MassIVE	MSV000094003; https://doi.org/10.25345/C5W95101G
Proteomics	MassIVE	MSV000094057; https://doi.org/10.25345/C51N7XZ36
Metabolomics	MetaboLights	MTBLS9098
Recombinant DNA		
pET-28a expression vector	Sigma Aldrich	69864
Software and algorithms		
Proteome Discoverer, version 2.5	Thermo Fisher	https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/proteome-discoverer-software.html?erpType=Global_E1
Spectronaut 15.5TM	Biognosys	https://biognosys.com/software/spectronaut/
Expressionist Refiner MS 12.0	Genedata	https://www.genedata.com/platform/expressionist
PROMIsed app	Schlossarek et al.[33]	https://github.com/DennisSchlossarek/PROMISed
SeeSAR software v13.0.1	BioSolveIT	https://www.biosolveit.de/products/seesar/
Other		
Amicon centrifugation filter	Millipore-Sigma	Cat#UFC9010
NGC Quest 10 Chromatography System	Bio-Rad	https://www.bio-rad.com/en-us/product/ngc-10-medium-pressure-chromatography-systems?ID=MFCVAQ4VY
Exploris 480 mass spectrometer	Thermo Fisher	https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-systems/orbitrap-lc-ms/orbitrap-exploris-mass-spectrometers/orbitrap-exploris-480-mass-spectrometers.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Q-Exactive mass spectrometer	Thermo Fisher	https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-systems/orbitrap-lc-ms/q-exactive-orbitrap-mass-spectrometers.html
Monolith	Nanotemper Technologies	https://nanotempertech.com/monolith/

EXPERIMENTAL MODELS AND STUDY PARTICIPATION DETAILS

E. coli strain BW25113 was grown on the M9 minimal medium. For each separation, 2 L of culture were harvested at OD = 0.8 and then the cells were washed thrice with 50 mM ammonium bicarbonate and 1.5 mM $MgCl_2$, pH 7.5 and flash frozen for future use. The *E. coli* expression strain received from ASKA collection[84] and *E. coli* BL21 (DE3) expression strains engineered to express a His-tagged PyrE and FabF, respectively, were grown in LB medium at 37°C, followed by induction with IPTG. *E. coli* W3110S used for biofilm characterization was cultured overnight at 37°C in LB medium.

METHOD DETAILS

Size exclusion chromatography (SEC)

The bacterial pellet was suspended in the cold lysis buffer (5 mL per 0.5 L of the culture). Composition of the lysis buffer: 50 mM Ammonium Bicarbonate, pH 7.5; 1.5 mM $MgCl_2$; 150 mM NaCl; 5 mM dithiothreitol (DTT); 1 mM phenylmethylsulfonyl fluoride (PMSF); 1× protease inhibitor cocktail (Sigma Aldrich, cat. no. P9599). Cells were lysed on ice with a sonication probe. The lysate was subjected to an initial centrifugation of 7000 × g, 4°C for 10 min, and the acquired supernatant was centrifuged again with the same parameters. The supernatant was further ultracentrifuged for 1 h at 165052 × g, 4°C. The supernatant was concentrated on a Amicon centrifugation filter (Millipore-Sigma) with the 10 kDa cutoff to about 2.5 mL. The supernatant was transferred to a new falcon tube and centrifuged again at 18000 × g, 4°C for 15 min to remove protein aggregates. Protein concentration was estimated with a Bradford assay. The lysate was diluted to 20 mg/mL and 2 mL were taken for the SEC separation. SEC was performed with an SRT-10 SEC-300 (Sepax) column connected to an NGC Quest 10 Chromatography System (Bio Rad) operating at 4 °C. Composition of the SEC buffer: 50 mM Ammonium Bicarbonate, pH 7.5; 1.5 mM $MgCl_2$; 150 mM NaCl). The flow rate was set to 7 mL/min. We collected a total of 48 fractions containing 40–80 mL of elution. The fractions were lyophilized and stored at −80 °C for proteomic and metabolomic analysis. Detailed protocol can be find in Sokolowska et al.[85]

Ion-exchange chromatography (IEX)

The bacterial pellet was suspended in the cold lysis buffer (5 mL per 0.5 L of the culture). Cells were lysed on ice with a sonication probe. Composition of the lysis buffer: 50 mM Ammonium Bicarbonate, pH 7.5; 1.5 mM $MgCl_2$; 5 mM dithiothreitol; 1 mM phenylmethylsulfonyl fluoride; 1× protease inhibitor cocktail (Sigma Aldrich, cat. no. P9599). The lysate was subjected to an initial centrifugation at 7000 × g, 4°C for 10 min, and the acquired supernatant was centrifuged again with the same parameters. The supernatant was further ultracentrifuged for 1 h at 165052 × g, 4°C. The supernatant was concentrated on the Amicon centrifugation filter (Milllipore-Sigma) with the 10 kDa cutoff to about 5 mL, after which 15 mL lysis buffer was added and concentrated again. The supernatant was transferred to a new falcon tube and centrifuged again at 18000 × g, 4°C for 15 min to remove protein aggregates. Protein concentration was estimated with a Bradford assay. The sample concentration was diluted to 10 mg/mL and 2 mL were taken for the IEX sample separation. For the control separation, the 2.5 mL of 10 mg/mL protein lysate was denatured in 95°C for 15 min and subsequently centrifuged in 18000 × g, 4°C for 15 min. The supernatant was filtered through the Amicon centrifugation filter (Milllipore-Sigma) with the 10 kDa cutoff to allow all of the sample through. The flow-through was collected for IEX control separation. 2.0 mL of soluble fraction corresponding to 20 mg of protein, or control metabolite only sample, were used for the separations. IEX was performed with a Enrich Q, 5/50 mm (Bio Rad) column connected to an NGC Quest 10 Chromatography System (Bio Rad) operating at 4°C. Composition of the IEX buffers; Buffer A: 50 mM Ammonium Bicarbonate, pH 7.5; 1.5 mM $MgCl_2$, Buffer B: 50 mM Ammonium Bicarbonate, pH 7.5, 1.5 mM $MgCl_2$, 1 M NaCl. The flow rate was set to 0.9 mL/min. Column was equilibrated with Buffer A. Elution gradient was set from 0% to 100% buffer B over the 20 mL. The collected 1 mL fractions were lyophilized and stored at −80°C for proteomic and metabolomic analysis.

Metabolite and protein extraction

The extraction protocol was adapted and modified from.[16] The lyophilized fractions were treated with a methyl *tert*-butyl ether (MTBE)/methanol/water solvent mixture to extract proteins and metabolites. Prior to further processing, equal amounts of the protein pellet and polar fraction were dried in a centrifugal evaporator and kept at −80°C.

Proteomics

IEX Rep 1 and 2, PROMIS Rep1

Sample preparation, measurements and data analysis were performed by DTU Proteomics Core, Lyngby, Denmark. 50 μL Lysis buffer (consisting of 6 M Guanidinium Hydrochloride, 10 mM TCEP, 40 mM CAA, 50 mM HEPES pH 8.5) was used to resuspend 40 lyophilized SEC fractions and 34 lyophilized IEX fractions. After being boiled at 95°C for 5 min, samples were sonicated for 5 min. Pierce Rapid Gold BCA (Thermo) was used to determine the protein concentration of the fractions, and 10 μg were taken forward for digestion. Samples were diluted 1:3 with 10% Acetonitrile, 50 mM HEPES pH 8.5. LysC (MS grade, Wako) was added in a 1:50 (enzyme to protein) ratio, and samples were incubated at 37°C for 4 h. After further dilution with 10% Acetonitrile, 50 mM HEPES pH 8.5. Trypsin (MS grade, Sigma) was added in a 1:100 (enzyme to protein) ratio and samples were incubated overnight at 37°C. Enzyme activity was quenched by adding 2% trifluoroacetic acid (TFA) to a final concentration of 1%. A SOLAμ SPE plate[86] (HRP, Thermo) was used for peptide desalting. For filter activation we used 200 μl of 100% Methanol (HPLC grade, Sigma), then 200ul of 80% Acetonitrile, 0.1% formic acid. For filter equilibration we used 2 × 200 μL of 1% TFA, 3% Acetonitrile. The samples were loaded on the filter using centrifugation at 1,500× rpm. After washing the filters twice with 200 μl of 0.1% formic acid, the peptides were eluted into clean 1.5mL Eppendorf tubes using 40% Acetonitrile, 0.1% formic acid. The eluted peptides were concentrated in an Eppendorf Speedvac, and re-constituted in 20 μl of 2% Acetonitrile, 1% TFA containing iRT peptides (Biognosys). For the ion exchange fractions, we loaded 500 ng onto EvoSep stagetips according to the manufacturer's protocol. For each IEX fraction, peptides were analyzed using the pre-set '15 samples per day' method on the EvoSep One instrument. Peptides were eluted over an 88-min gradient. For each SEC pool, peptides were loaded onto a 2cm C18 trap column (ThermoFisher 164705), connected in-line to a 15 cm C18 reverse-phase analytical column (Thermo EasySpray ES804A) using 100% Buffer A (0.1% Formic acid in water) at 750 bar, the Thermo EasyLC 1200 HPLC system, and the column oven operating at 30°C. Peptides were eluted over a 140 min gradient going from 10 to 23% of Buffer B (80% acetonitrile, 0.1% formic acid) in 95 min, then to 38% Buffer B in 30 min, then to 60% Buffer B in 5 min, ending with 10 min at 95% Buffer B. The flow was set to 250 nL/min. Spectra were acquired with an Orbitrap ExplorisTM 480 instrument (Thermo Fisher Scientific) FAIMS ProTM Interface (ThermoFisher Scientific) set to CV of −45 V. Full MS spectra were collected at a resolution of 120,000, with normalized AGC target of 300% or maximum injection time of 45 ms and a scan range of 345–1500 *m/z*. AGC target for the DIA experiment was set to 1000%. The precursor mass range was set to 350–1400 m/z with 48 isolation windows of 21.5 Da and 1 Da overlap. The resolution was set to 15,000 with maximum injection time set to auto and a normalized collision energy of 30%. MS performance was verified for consistency by running complex cell lysate quality control standard. Raw files from the DIA experiments were analyzed using Spectronaut 15.5TM (Biognosys) using the BGS Factory Settings. PROMIS Rep2: Protein pellets were processed as described in[87] with the modifications below. Protein pellets were dissolved in a denaturation buffer (2 M thiourea and 6 M Urea dissolved in Ammonium bicarbonate (Ambic) (50mM). Protein concentration was determined with Bradford assay. ∼50 μg of proteins were subjected to protein digestion. The protein extracts were reduced by the addition of DTT (100 μM) for 1 h at room temperature followed by (300 μM) iodoacetamide for alkylation at dark. DTT (100 μM) was used to quench the residual iodoacetamide for 10 min at room temperature. Trypsin/Lys-C mixture (Mass Spec Grade, Promega) was used 16 h to digest the[88] proteins. Digested peptide samples were desalted using C18 sep-pak column plates. The eluted peptides were transferred to Eppendorf low-bind tubes and then dried using speed vac and resuspended using a resuspension buffer (5% acetonitrile in 0.1% formic acid). ∼1 μg of the peptides were injected for analysis. The peptide mixtures were separated using a nano- LC (Dionex ultimate 3000) using an acclaim pepmap C18 column. A flow rate was kept constant at 300 nL min-1 for the peptide separation. Gradients were kept the same as.[88] A nano bore stainless-steel emitter (Fisher Scientific) was used to spray the peptide samples. Peptides were analyzed using an Orbitrap-Exploris-480 MS mass spectrometer in data-dependent acquisition (DDA) mode with a cycle time of 3 s. Standard mass spectrometry parameters were maintained as described in Thirumalaikumar et al.[87] The instrument was operated in positive ion mode with the settings described in Guo et al.[88] Full scan spectra were acquired at a resolution of 60,000 over an m/z range of 350–1650, with 40% set at RF lens. The AGC target was kept at the standard setting and IT (injection time) mode was set to Auto. Data-dependent MS2 scans, monoisotopic peak determination was kept for peptides, with a charge state selection from 2 to 8 and intensity threshold of 5000. Higher-energy collisional dissociation (HCD) with a normalized collision energy of 30% was used for Fragmentation. The isolation window was set to m/z 1.6, and the MS2 spectra were acquired at an Orbitrap resolution of 15,000 with RF lens at 50%, with standard AGC target and maximum IT (injection time) was set to auto. HeLa digests (Pierce, 88329) were analyzed before and after each experiment. Raw data were processed using Proteome Discoverer (version 2.5, Thermo Fisher Scientific) according to the manufacturer's instructions. Database searches were conducted using SEQUEST HT against an *E. coli* protein database downloaded from UniProt.

Metabolomics

Following metabolite extraction, ultra-performance liquid chromatography (UPLC) coupled with a Q-Exactive mass spectrometer (Thermo Fisher Scientific) was used to measure metabolites in positive and negative ionization modes, as described earlier.[16] Specifically, 100 μL of water was used to suspend the dried polar phase, helped by 5 min-long sonication. Following the 10min 20,800 g, RT centrifugation step to remove any left-over debris samples were transferred to UPLC glass vial. Separation was performed taking advantage of an HSS T3 C18 reversed-phase column; 3 μL of the sample was loaded for each ionization mode. We used the following mobile phase solutions: buffer A (0.1% formic acid in H_2O) and buffer B (0.1% formic acid in ACN). Separation of metabolites at 400 μL/min was performed using the following gradient: 1 min 1% LC-MS mobile phase buffer B, 11 min linear gradient from 1%

to 40% buffer B, 13 min linear gradient from 40% to 70% buffer B, then 15 min linear gradient from 70% to 99% buffer B, and hold a 99% buffer B concentration until 16 min. Starting from 17 min, a linear gradient from 99% to 1% buffer B was applied. Column was re-equilibrated for 3 min with 1% buffer B. Following settings were applied to acquire mass spectra: mass range from 100 to 1500 m/z, resolution set to 25,000, loading time restricted to 100 m, AGC target set to 1e6, capillary voltage to 3 kV with a sheath gas flow and auxiliary gas value of 60 and 20, respectively. The capillary temperature was set to 250 °C and skimmer voltage to 25 V. Expressionist Refiner MS 12.0 (Genedata AG, Basel, Switzerland) was used for processing the LC-MS data. Metabolites were annotated using an in-house library of authentic reference compounds allowing 10 ppm mass deviation and dynamic retention time deviation (maximum 0.1 min).

Data analysis

Elution profiles were analyzed using standard settings in the PROMIsed app,[33] including pre-processing, deconvolution, and integration steps. To calculate the Pearson Correlation Coefficient, we used pre-processed and deconvoluted elution profiles, and the N/A values were substituted by zero.[33] The omniTICC profiles are derived from the manual curation of the IEX profiles coming from the control and sample preparations. Specifically, elution maxima corresponding to the free metabolites, based on the control separation, were removed for downstream analysis. Note that, however, as differential elution was primarily dictated by shift in the elution maxima, we also accepted elution maxima with the average fold-change intensity in the sample at least 3-time more than in the control.

Protein purification

Recombinant PyrE was purified from an *E. coli* expression strain received from ASKA collection[84] engineered to express a His-tagged PyrE. The process began with a bacterial preculture in LB medium, which was incubated at 37°C for 3 h. Protein expression was then induced by adding 100 μM of IPTG, followed by incubation for 4–5 h at 37°C. Frozen bacterial cells were used for subsequent protein purification. For small-scale purification, we utilized the His-Bind purification kit (Merck, USA) and followed the manufacturer's protocol. For large-scale purification, the supernatant was loaded onto a 1 mL HisTrap (Cytiva) column connected to NGC Quest 10 (Bio-Rad), previously equilibrated with Buffer A (25 mM Tris–HCl pH 8.0, 300 mM NaCl, 5% (v/v) glycerol, 10 mM imidazole). The column was washed with 10 mL of Buffer A and the recombinant protein was eluted with a linear gradient of imidazole (10–300 mM, 50 mL). The fractions containing the enzyme of interest were collected and concentrated and desalted using ultra-centrifugal filters. The purified protein was stored in 0.1 M Tris-HCl buffer pH 8.8 at 4°C for use in future experiments. FabF cloned into pET-28a expression vector was expressed in *E coli* BL21 (DE3) strains. Bacteria were grown in the LB medium to the OD of 0.8–1.0. Protein expression was then induced by adding 100 μM of IPTG, followed by 4 h incubation at RT. Expression, induction and purification of the recombinant protein were performed as described above. The purified protein was stored at 4°C in 50 mM sodium phosphate buffer pH 7.4 supplemented with 300 mM NaCl.

Microscale thermophoresis (MST)

A dilution series of PRPP and lumichrome was prepared by using 16-step, 2-fold serial dilution for each ligand. The concentration in this series started at 781 μM for PRPP and 200 μM for lumichrome, decreasing with each step. These dilutions were carried out in their respective binding buffers, which consisted of 50 mM Tris-HCl (pH 8.0 for PRPP and pH 9.5 for lumichrome), 150 mM NaCl, 10 mM $MgCl_2$, and 0.05% Tween 20. Each reaction mixture was created by mixing 6 μL of a protein working solution, containing 6.25 μM of His-tagged PyrE labeled with RED dye (Nanotemper Technologies, USA), with an equal volume of the respective ligand solution. The mixtures were loaded into standard capillaries. Subsequently, these capillaries were inserted into the Monolith device (Nanotemper Technologies, USA), and the measurement of MST was conducted following the manufacturer's provided instructions. Analysis of MST data involved standard MST traces and the plotting of changes in normalized fluorescence (F_{norm}) against the ligand concentration, enabling the determination of the dissociation constant (K_d). This analysis was performed using Monolith analysis software. FabF protein was labeled with the RED-MALEIMIDE or RED-NHS dye (Nanotemper Technologies, USA) according to the manufacturer's instructions. Binding was performed in premium capillaries in 25 m M PIPES pH 7.5 and 200 mM NaCl. MST power was set to high. Data analysis was performed using Monolith software.

Thermal shift assay (TSA)

To assess the shift in the unfolding temperature of PyrE in the presence of the ligands compared to the unfolding temperature obtained in the absence of any ligand, we conducted a thermal shift assay (TSA). Each reaction solution was prepared with 2.5 μg of purified protein, the ligand solution, and 1X protein thermal shift dye (Applied Biosystem, USA), following the manufacturer's instructions. The melting curve for each treatment was monitored using a real-time PCR instrument, as the temperature was incrementally raised from 25°C to 99 °C at a rate of 0.05°C per second.

Orotate phosphoribosyltransferase assay

The activity of orotate phosphoribosyltransferase (OPRTase) was measured as previously described with some modification.[89] The reaction mixture consisted of 100 mM Tris-HCl (pH 8.8), 5 mM $MgCl_2$, 0.2 mM orotic acid and PRPP ranging from 0.125 to 4 mM. To examine the effect of lumichrome on PyrE activity, 100 μM of lumichrome was added. The reaction is initiated by adding PyrE into the

mixture. The decrease in absorbance at 295 nm, resulting from the conversion of orotic acid (ε_{295} = 3950 M^{-1} cm^{-1}), was monitored after 1-min of incubation at 30°C. The activity was determined by calculating the amount of enzyme required to convert 1 μmol of orotic acid to orotidine monophosphate (OMP) per minute.

Congo-red binding assay

E. coli W3110S was cultured overnight at 37°C in LB medium. Next, 2 μL of the culture was inoculated onto LB 1/4 agar medium containing 0.004% Congo-red and 0.002% Coomassie blue (CR medium). To investigate the impact of lumichrome on Congo-red binding, 10 μM and 100 μM of lumichrome was added into the CR medium. The bacterial strain was grown at 30°C for 24 h, followed by incubation at 4°C for an additional 24 h. The bacterial cells were harvested and extracted using DMSO. After centrifugation at 13,500 × *g* for 1 min, the absorbance of the supernatant was measured at 497 nm.

Molecular docking

3D structural information files for the proteins of interest were downloaded from RCSB PDB.[90] Potential ligand binding sites for these structures were predicted using PrankWeb.[91] If more than one binding pocket was predicted with high probability, the binding pocket with highest probability was chosen for the subsequent molecular docking. The chosen binding sites were then used to perform molecular docking with the SeeSAR software v13.0.1 (BioSolveIT GmbH). Subsequently, 2D interaction diagrams were generated for the top docking pose for each compound-protein complex using PoseView.[92]

QUANTIFICATION AND STATISTICAL ANALYSIS

Figure 3B; Fisher's exact test was calculated using https://www.socscistatistics.com Website. Asterisk indicates *p*-value <0.01
Figures 3D–3I; Functional over-representation, including significance, of the protein targets of the different compound classes was calculated and visualized using ShinyGO 0.77 app.[36] Figure 5J; significance was estimated using unpaired two-tailed Student's t test assuming unequal variance embedded in the Excel; ***: $p < 0.001$, ****: $p < 0.0001$.