Electrum: Visualization, analysis, and contextualization of high-throughput protein-metabolite interaction datasets

Jordan A. Berg[1], Ian George[1], Youjia Zhou[2,3], Kevin G. Hicks[1], Bei Wang[2,3], Jared Rutter[1,4,5]

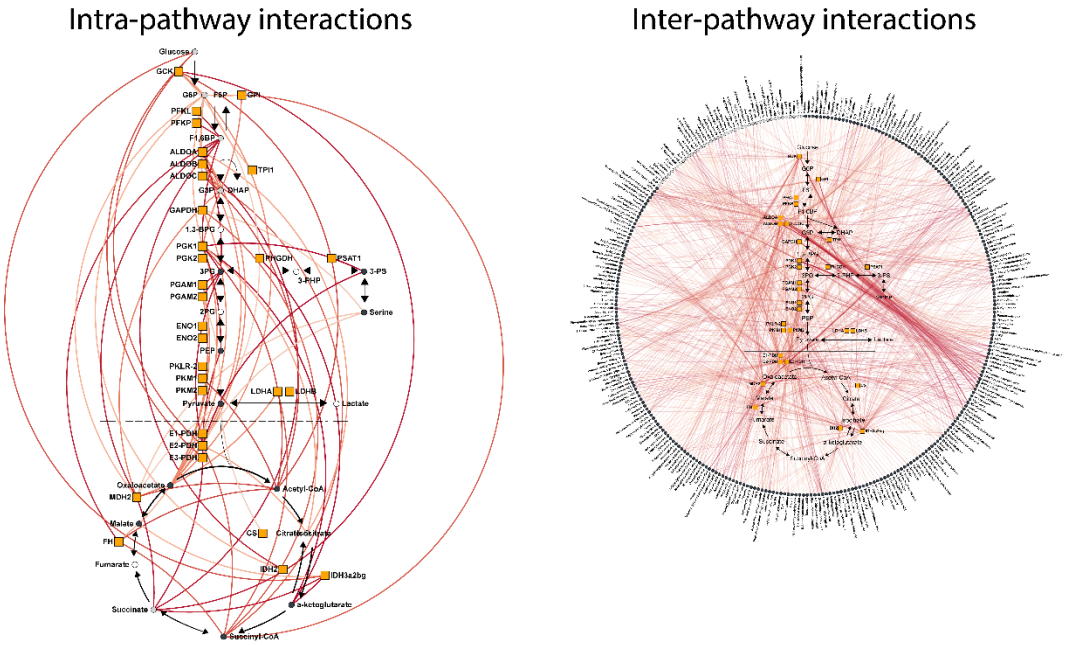
1 Department of Biochemistry, University of Utah, Salt Lake City, Utah, USA, 84112
2 School of Computing, University of Utah, Salt Lake City, Utah, USA, 84112
3 Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, Utah, USA, 84112
4 Diabetes & Metabolism Research Center, University of Utah, Salt Lake City, Utah, USA, 84112
5 Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah, USA, 84112

Long Abstract:

Enzymatic reactions form the basis of metabolism's generation of the molecular resources required for life. Several factors can direct the actions of these enzymes; however, and one primary mechanism involves the allosteric interaction of a metabolite with an enzyme. Over the past century, many such protein-metabolite interactions have been annotated and described within cellular and disease biology. However, the identification of these interactions is laborious and thus many biologically relevant interactions are yet to be discovered. Yet, the discovery of these interactions is important, and may lead to valuable insights into the biochemical mechanisms different processes and diseases.

To facilitate the rapid and unbiased discovery of novel regulatory interactions, in other works we developed “MIDAS”, or “mass spectrometry integrated with equilibrium dialysis for the discovery of allostery systematically” [1-2]. MIDAS detects protein-metabolite interactions by using a dual-chamber system separated by a dialysis film. Initially, the purified protein is in one chamber, and pools of various metabolites are in the second chamber. Metabolites are then allowed to dialyze between the two chambers until equilibrium is reached, while the dialysis film matrix is such that the purified protein is unable to pass between the chambers. Based on this methodology, metabolites interacting with the given protein will either be enriched (interaction) or depleted (enzymatic consumption). Once equilibrium has been reached, the two chambers are processed using mass spectrometry and fold changes between the initial metabolite pool concentrations and the dialyzed concentrations are calculated to determine the exchange of metabolites. Due to the quantity of data produced through this methodology and the absence of analytical methods for interpreting this unique data type, we built an interactive tool, Electrum, to aid in the analysis and contextualization of these interaction datasets.

Fig. 1: Electrum enables exploration of high-throughput protein-metabolite mass spectrometry data. *(left) Central carbon metabolism enzymes assayed by MIDAS and their interactions with MIDAS metabolites that have been canonically annotated within the central carbon metabolism pathway. (right) Central carbon metabolism enzymes assayed by MIDAS and their interactions with the total MIDAS metabolite library. Squares indicate enzymes, circles indicate metabolites, and edges represent enzyme-metabolite interactions with the color hue scaled by the q-value for that interaction. q-value threshold set at > 0.1.*



To standardize the output data from MIDAS for visualization in Electrum, interaction database files require each row to summarize an interaction with a metabolite and a target protein, along with standard MIDAS enrichment and statistical values. Electrum contains two main modes of visualizing the data housed within the MIDAS database. The first is to view an individual enzyme and its interacting metabolites. The second visualization mode within Electrum partitions assayed proteins into canonical biological pathways (Fig. 1). Interactions between proteins and metabolites can be denoted one of two ways: 1) Edge weights are unscaled and edge color intensities are scaled by the interactions q-value percentile among all other interaction q-values for protein passing the set threshold. 2) Edge weights are scaled based on the interaction's q-value and are color-scaled based on the fold-change value of the metabolite concentration between the MIDAS dialysis chambers. Hovering over a protein-metabolite edge will display the protein and metabolite IDs and names, fold change values, and q-values and will highlight that particular edge. Hovering over a metabolite will display metadata about that given entity and highlight all edges involving that metabolite and its protein interaction(s). Hovering over a protein node will highlight all metabolites with which that protein interacts. At the time of writing, most of the central carbon metabolism pathway has been interrogated using MIDAS [2]; however, the structure of Electrum and the data are such that as additional pathways are assayed, these can be facilely visualized within this framework.

To shed additional light on the biological context of these protein-metabolite interactions, Electrum is integrated with the Metaboverse human reaction database [3]. Metaboverse generates a representation of all pathways within human metabolism and other biological pathways based on the Reactome knowledgebase, and additionally integrates additional metadata from the KEGG, HMDB, and other databases [4-9]. Using this information, Electrum can then show biological reactions and pathways for a selected metabolite to provide insights on how this canonical or unknown protein-metabolite interaction may be operating biologically.

For a metabolite to interact with a protein, the structures of both the protein and the metabolite must be complementary insomuch as to enable binding or interaction between the two components. To confirm such an interaction, often an experimental technique, such as X-ray crystallography of the protein in the presence of the metabolite, is required. We thus set out to develop a metabolite substructure enrichment analysis that would enable the enrichment of conserved substructures of metabolites interacting with a given target protein. Metabolite substructure ontology is derived from ClassyFire [10]. These enrichments can serve as valuable starting points for identifying shared structures between metabolites that could indicate a possible interaction mechanism with their target protein. Paired with demarcation between enriching and depleting interactions, this can be used to additionally differentiate substructures that may correlate with either interaction or enzymatic consumption. To our knowledge, this is the first implementation of metabolite substructure enrichment analysis and has the potential to identify putative interaction mechanisms between proteins and metabolites.

References

1. Orsak T, et al. Revealing the allosterome: systematic identification of metabolite-protein interaction. Biochemistry 2012 Jan 10;51(1):225-32. https://doi.org/10.1021/bi201313s.
2. Hicks KG, et al. In progress
3. Berg JA, et al. Metaboverse: Metabolic network pattern recognition toolscontextualize multi-omic data and reveal disease-relevant signatures. bioRxiv 2020.06.25.171850; https://doi.org/10.1101/2020.06.25.171850
4. Joshi-Tope G, et al. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res. 2005 Jan 1;33(Database issue):D428-32. https://doi.org/10.1093/nar/gki072.
5. Jassal B, et al. The reactome pathway knowledgebase. Nucleic Acids Res.2020 Jan 8;48(D1):D498-D503. https://doi.org/10.1093/nar/gkz1031.
6. Fabregat A, et al. The Reactome Pathway Knowledgebase. Nucleic AcidsRes. 2018 Jan 4;46(D1):D649-D655. https://doi.org/10.1093/nar/gkx1132.
7. Wishart DS, et al. HMDB 4.0: the human metabolome database for 2018.Nucleic Acids Res. 2018 Jan 4;46(D1):D608-D617. https://doi.org/10.1093/nar/gkx1089.
8. Kanehisaa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000 Jan 1; 28(1): 27–30. https://doi.org/10.1093/nar/28.1.27.
9. Kanehisa M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017 Jan 4;45(D1):D353-D361. https://doi.org/10.1093/nar/gkw1092.
10. Djoumbou Feunang Y, et al. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J Cheminform. 2016 Nov 4;8:61. https://doi.org/10.1186/s13321-016-0174-y.