Active learning maps the emergent dynamics of enzymatic reaction networks.

Miglė Jakštaitė[1], Tao Zhou[1], Frank Nelissen[1], Wilhem T.S. Huck ✉,[1], Bob van Sluijs ✉,[1]

1   Institute for Molecules and Materials, Radboud University, Nijmegen, AJ, The Netherlands.

Abstract

The dynamic properties of enzymatic reaction networks (ERNs) are difficult to predict due to the emergence of allosteric interactions, product inhibitions and the competition for resources, that all only materialize once the networks have been assembled. Combining experimental kinetics studies with computational modelling allows us to extract information on these emergent dynamic properties and build predictive models. Here, we utilized the pentose phosphate pathway to demonstrate that previously reported approaches to construct maximally informative datasets can be significantly improved by pulsing both enzymes and substrates into microfluidic flow reactors (instead of substrates only). Our method augments information available from online databases, to map the emergent dynamic behaviours of a network.

Introduction

The emerging field of cell-free synthetic biology aims to construct complex enzymatic reaction networks (ERNs) *in vitro*, as they offer a route to scaling up the production of valuable compounds[1-10] or build functional systems[11-14]. This cell-free approach offers advantages over traditional *in vivo* metabolic engineering since it avoids difficult product purification steps, improves yields, allows for the production of compounds that are toxic to cells, and negates the sequestration of input substrates by other metabolic pathways[15-19]. Although purified enzymes and added co-factors are costly, cell-free networks offer much greater flexibility and control over the construction of multi-enzymatic cascades[20].

Several groups have recently reported major breakthroughs in the development of complex ERNs, starting from simple and cheap materials like glucose or fixating CO_2 to produce valuable compounds as monoterpenes, polyhydroxybutyrate, cannabinoids, malate or 6-deoxyerythronolide B[21-26]. It should be noted that complex networks can also be constructed using crude cell lysates for production of n-butanol, mevalonate or limonene[27-29]. These networks highlight the potential of cell-free biocatalysis.

However, optimizing ERNs *in vitro* can be laborious and time-intensive, and it is often not possible to identify input conditions that represent a global optimum[30-36]. Rondelez demonstrated that in dynamic environments, subtle activation and inhibition processes can emerge in ERNs, affecting enzyme efficiency[37]. Yet, mapping emergent dynamic processes, such as allosteric interactions, product inhibitions, and competition for resources, is challenging because these processes only become apparent once ERNs are assembled[38]. Unfortunately, public databases such as BRENDA cannot account for all possible cross reactions for an enzyme nor report the exact kinetics of an enzyme in a new environment.

In this context, we recently reported an active learning workflow utilizing an optimal experimental design (OED) algorithm to train a kinetic model that can subsequently be queried for input combinations that lead to any desired output[39-41]. The OED algorithm designs a sequence of out-of-equilibrium perturbations for species that flow into the continuously stirred tank reactor (CSTR) which maximizes the information about the reaction kinetics[42-47]. Within the active learning workflow this OED procedure occurs iteratively. The time course data of an optimally designed pulse experiment is added to a database. The last experiment added to this database serves as test data, the remainder as training data. A kinetic model is subsequently trained. If the model can predict the test data of the last pulse experiment, the cycle stops, if not, a new experiment is designed (**Figure 1**).

Previously, we used enzymes immobilized on beads[48-54] and applied this workflow to a system where the catalytic pathway was enclosed within an open CSTR [35-37,55-57]. However, the use of beads with immobilized enzymes in flow systems presents two distinct disadvantages. First, immobilization is variable and results in K_M and k_{cat} values that are different from those reported in online databases. Second, the potential input space is less informative (enzyme concentrations are fixed), which results

in less informative training data per experiment. Thus, multiple iterations of the OED cycle were needed to parameterize a model.

Here, we extend the active learning workflow by incorporating pulses of both enzymes and substrates into the CSTR, using the otherwise well-characterized pentose phosphate pathway (PPP) as a model ERN (**Figure 2**). The topology of the PPP network allowed us to gauge the effect of emergent nonlinear behaviour on the dynamics of the network and highlights the utility of the active learning workflow in this context. We show that the optimal perturbation of enzymes and substrates led to more informative data. This enabled a 'one and done' approach where we trained a quantitatively accurate model in a single experiment.
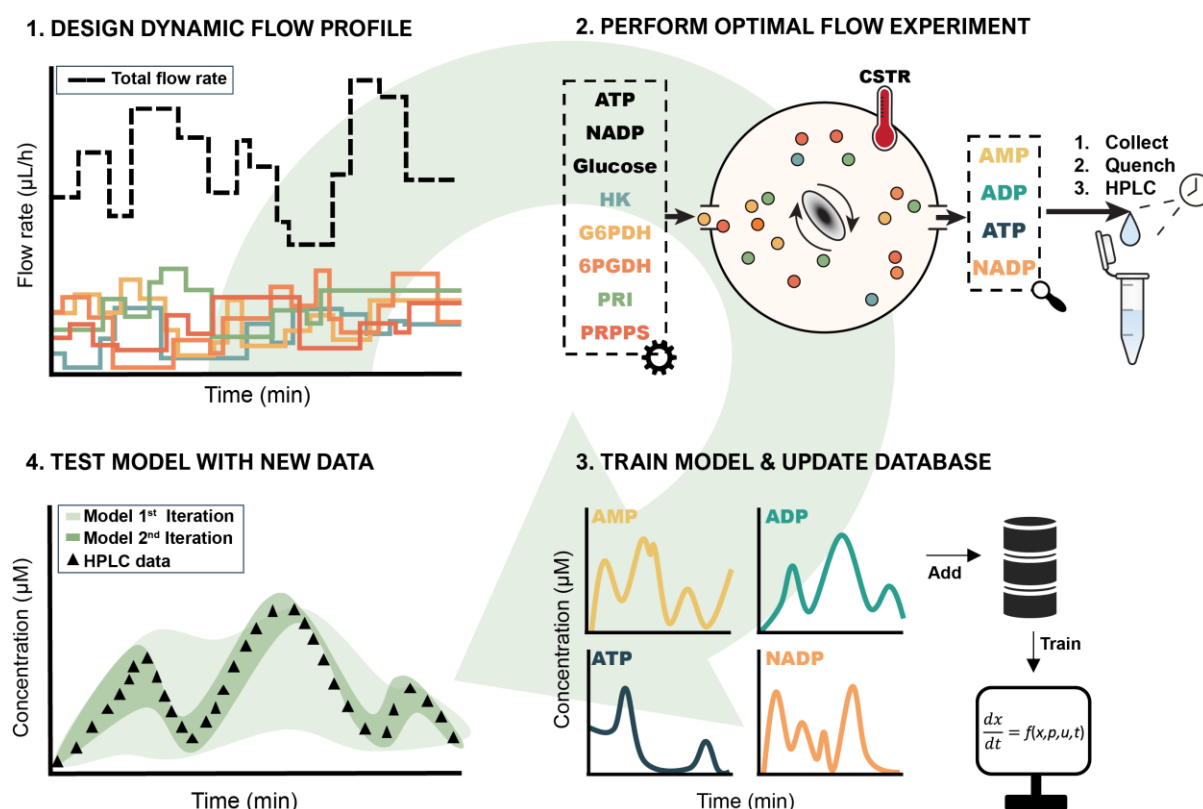
Figure 1. *An overview of the active learning procedure applied pentose phosphate pathway.* **1.** *The OED algorithm designs a sequence of pulsed inputs for each species that flow into a temperature controlled continuously stirred tank reactor (CSTR volume ~150 µL).* **2.** *These species include five enzymes in free form; HK, G6PDH, 6PGDH, PRI and PRPPS and 3 substrates; ATP NADP and Glucose. Eight individual syringes control the unique inflow profile of each species. Droplets formed at the end of the CSTR reactor are collected, quenched, and measured offline by HPLC. We observe AMP, ADP, ATP and NADP.* **3.** *The first experiment is added to a database and used to train a model, from thereon every new experiment added to the database serves as test data as the pipeline uses the other experiments as training data.* **4.** *When the model has sufficient predictive power to control the catalytic processes inside the reactor, the active learning cycle stops.*

Results and Discussion

Model of pentose phosphate pathway.
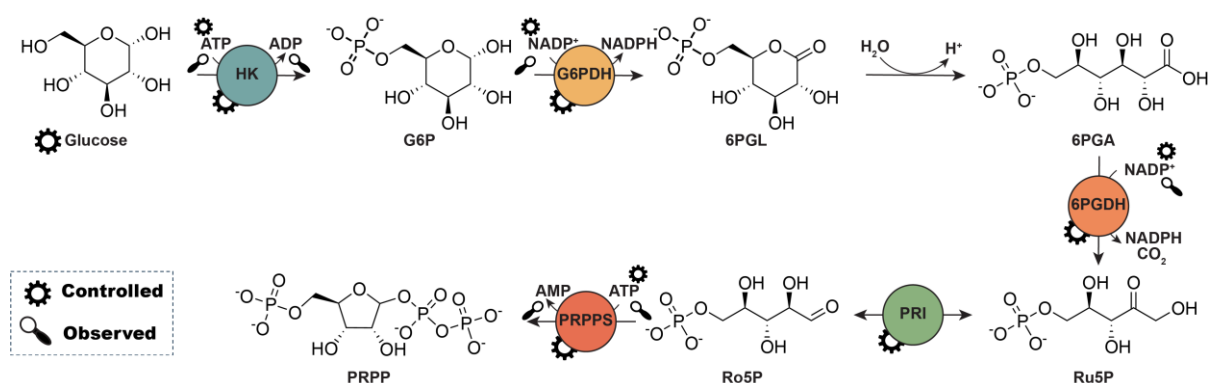
Figure 2. *An overview of the reactions taking place inside the CSTR shown in* **Figure 1**. *The gears denote the species that flow into the reactor, the magnifying glass the species which are observed in the quenched solution. Among these observables, AMP and ADP are not part of the input and AMP is the final product in the pathway. ADP is known to inhibit PRPPS.*

Figure 2 shows the PPP ERN cascade used in this study. The system contains five enzymes – hexokinase (HK), glucose-6-phosphate dehydrogenase (G6PDH), 6-phosphogluconic dehydrogenase (6PGDH), phosphoriboisomerase (PRI) and phosphoribosyl pyrophosphate synthetase (PRPPS), two cofactors – NADP and ATP, and one substrate – glucose. At the start, glucose is converted to glucose-6-phospate (G6P) by HK using ATP as a cofactor. G6PDH using NADP as a cofactor converts G6P to 6-phosphogluco-δ-lactone (6PGL), which undergoes spontaneous hydrolysis to 6-phosphogluconate (6PGA). Subsequently, 6PGA is converted to ribulose-5-phosphate (Ru5P) by 6PGDH using NADP as a cofactor, which is reversibly isomerised to ribose-5-phosphate (Ro5P) by PRI. In the last step, Ro5P is converted to PRPP by PRPPS using ATP as a cofactor (PRPPS and PRI were purified according to Arthur et al. [58], for more information see **Supporting Information 1.5**). The final product PRPP is the precursor for the nucleotide synthesis pathway ERN described earlier[40] as well as a building block for DNA and RNA. The Figure shows the compounds and enzymes that are controlled by external inputs and the intermediates (observables) that were quantified using HPLC (**Supporting information 1.4, Figure S5**)[59].

The cofactors in these catalytic steps are reported to induce a conformational change that allows the substrate to bind (an ordered sequential bi-bi reaction)[60-64]. Reactions are also assumed to be irreversible under physiological conditions, however because we operate in a very broad regime of both substrate and enzyme concentrations within the CSTR, we assumed reactions were reversible. Finally, BRENDA reports that the inhibition of PRPPS by ADP is conserved for enzymes in most organisms[65].

On this basis, we constructed an initial, kinetic model of ordinary differential equation that maps these processes, consisting of 10 reactions and 32 kinetic parameters (see **Supporting Information 2.1**). To test the activity of each of the enzymes *in vitro*, we performed separate batch experiments (**Supporting Information 1.1, Figure S1-2**). This data was subsequently used to constrain an initial model, required for informative OED (**Supporting Information 2.1-2.2**)[66]. However, before any flow experiments were performed, we probed the ERN *in silico* and compared the efficacy of the active learning procedure for different experimental scenarios, and subsequently learned what could be expected. From this study we were able to extract the experimental requirements and define a simple hypothesis regarding potential hidden dynamics present within the ERN.

### *In silico* kinetic study of PPP network.

We explored different scenarios for the active learning workflow by training models on *in silico* data. **Figure 3a** outlines two scenarios: in scenario 1, the PPP pathway is coupled to beads, and three substrates can be pulsed into the CSTR. In scenario 2, enzymes are free and flow into the reactor. For each scenario, we created two optimally designed training datasets (**Supporting information 3.1-3.2, Figure S6**): one where transients are measured and another where only steady states are measured. The model was trained on each of the four datasets separately and subsequently tested with a new pulse experiment (**Figure S7**), consistent with the active learning workflow (**Figure 1**).

Figure 3b subsequently shows the relative prediction error of each scenario. As expected, utilizing free enzymes within the active learning pipeline is more informative, resulting in a >4 fold lower prediction error[40]. We also probed the effect of ADP inhibiting PRPPS by training a model that did not include this interaction. If ADP does not inhibit PRPPS it becomes easier to train a model, since less informative data from other scenarios could be used to train an equally predictive model.

In **Figure 3c** we opted to further explore the information gained within the transient free enzyme dataset since that mirrors our experimental *in vitro* set-up. We trained the model - including PRPPS inhibition - with different sizes of training data comprising short and long experiment runtimes. The time series plots show the prediction of AMP and ADP, ATP and NADP test data (shaded area). As experimental runtimes increase, the prediction error of the model declines, this decline levels off after 240 minutes.

With this *in silico* exploration, we could *a priori* set an experimental requirement and define a simple hypothesis. First, this analysis indicates an experimental runtime of at least 240 minutes is preferred to train the model in a single iteration. Second, **Figure 3c** shows that this preferred runtime holds true when ADP inhibits PRPPS, however, less informative data can also be used if there is no emergent network behaviour. From this we can assume the opposite holds true as well, that any model inaccuracies after including 240 minutes of actual *in vitro* data in the training dataset are likely due to

additional nonlinear effects. Effects stemming from either emergent network behaviour or environmental factors which cannot be easily gleaned from public databases[37,67-69].
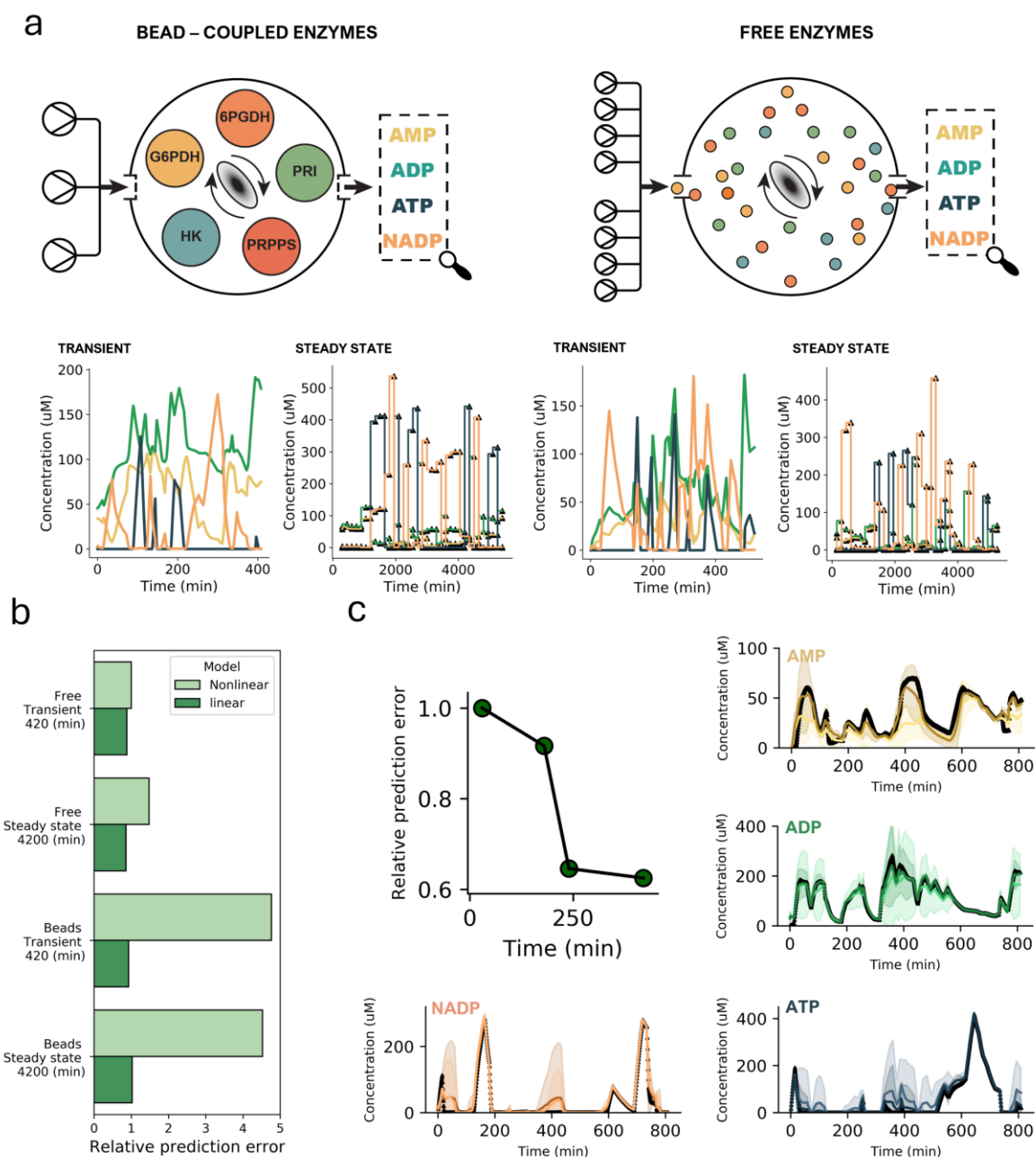

Figure 3. *An in silico study exploring the difference between two strategies to control catalysis inside reactors.* ***a)*** *in scenario 1 the PPP pathway enzymes are coupled to beads; three substrates can be pulsed into the reactor. In scenario 2 enzymes are free and flow into the reactor. For each scenario we subsequently created two training datasets in silico, one where transients are measured for an optimally designed experiment and one where only its steady state is measured. For each of the four datasets the inputs were altered 35 times.* ***b)*** *Shows the difference in the prediction error for the scenarios after 420 minutes (transient) or 4200 minutes (steady state) of training, for a model that includes an inhibition reaction between ADP and PRPPS (nonlinear) and one that does not (linear), simulations of best fits to training data, n = 10.* ***c)*** *Focusses on the transient free enzyme OED dataset,*

by plotting the relative prediction error for different lengths of training data. After the training dataset contains 240 minutes of data this prediction error no longer becomes smaller. The time course test data (black diamonds) and the model predictions serve as a visual reference to the relative prediction error. The shaded areas represent the uncertainties (3 standard deviations). The lighter time course prediction in each time series plot is the trained on 30 minutes of in silico data, (more uncertainty), the darker time course prediction is trained on 240 minutes of in silico data, for both simulations based on best fit to training data, n = 10.

Optimally designed *in vitro* study of the PPP network.

To confirm free enzymes could be used within the active learning workflow and test the hypothesis, we conducted two *in vitro* experiments. Eight syringe pumps, each controlling an individual species—HK, G6PDH, 6PGDH, PRI, PRPPS, glucose, ATP, and NADP—were connected to a CSTR (volume 121 µL) and a single outlet tubing (**Supporting Information 1.2-1.3, Figure S3-4**). Each drop was collected, quenched and analysed offline by HPLC (**Supporting Information 1.4, Figure S5**). **Figures 4a** and **4b** summarize step two of the workflow, adding two *in vitro* experiments to the database. Each panel shows both the inflow profile (left) and the raw data (right). To explore a larger input space and generate sufficiently complex test data, the stock concentrations of enzymes and substrates in each syringe differed between iterations one and two (**Supporting Information 3.3,** concentrations for enzymes inside reactor shown in **Figure S9**).
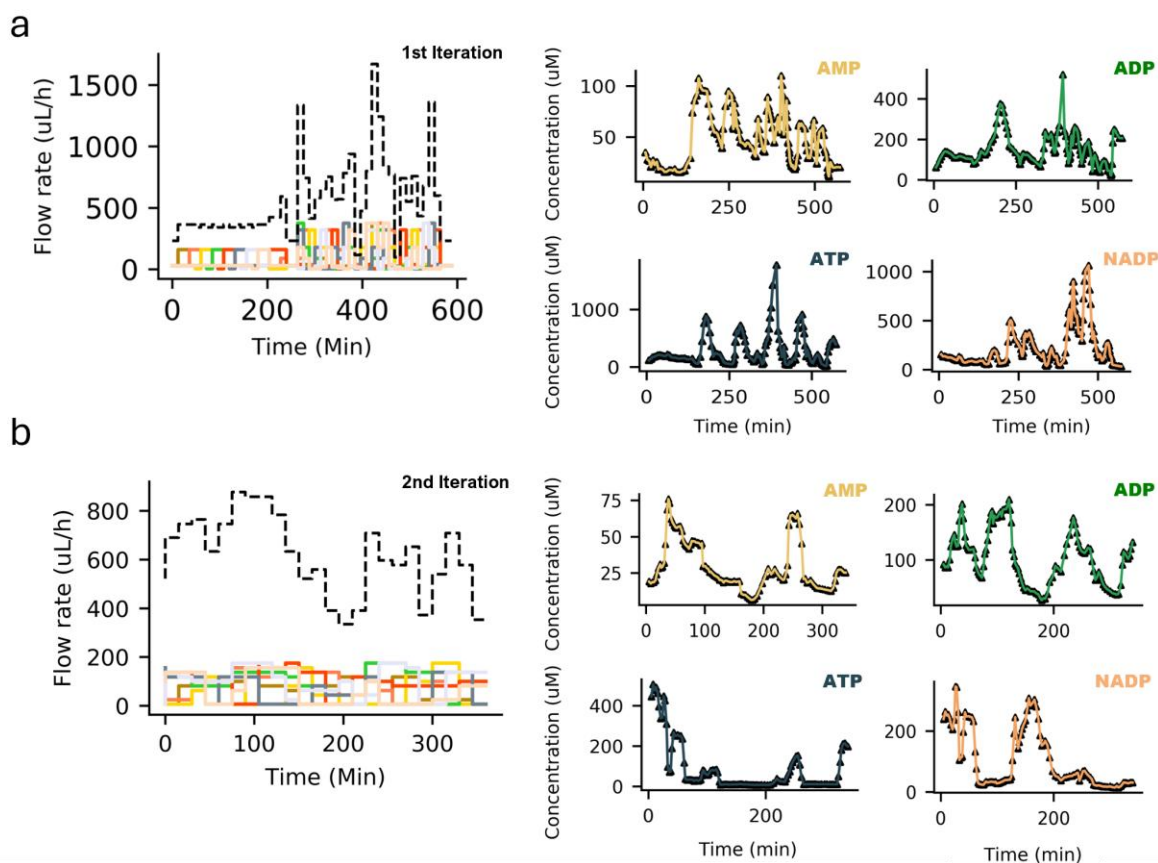
Figure 4. *Step two of the active learning workflow, this figure summarizes two iterations of the active learning cycle for the PPP network, syringe stock concentrations can be found in* **Table S3**. ***a)*** *Shows the first iteration, with the flow profile shown on the left and the HPLC data on the right, the diamonds represent the data as measured on the HPLC. The trace through the data points was added for clarity highlighting the dynamic nature of the data. The first 200 minutes of the first experiment are manually designed to ensure the experimental setup functions (***Figure S8***).* ***b)*** *Shows the second optimally designed iteration with the flow profile shown on left and HPLC data shown on the right.* **Figure S9** *shows the changes in enzyme concentrations for both experiments.*

Figure 5 subsequently shows the results of step three in the workflow, training the model OED data from the first iteration. The results indicate that the model has sufficient degrees of freedom to approximate the data (**Figure S10 & S11**). Additionally, we were able to confirm ADP indeed inhibits PRPPS, noting that models that do not contain this interaction do not converge as well (**Figure S12**).

Finally**, Figure 6** shows the last step in the active learning workflow, testing the model. We plot the HPLC test data versus the model predictions for each species and compute the R² value (R² = 0.891). The subsequent time-course HPLC data of all measured species are represented by the black diamonds, the predictions and their uncertainty (three standard deviations) are represented by the coloured line (mean) and shaded are (uncertainty).

These results highlight two key factors. First, the active learning workflow is more efficient when free enzymes are used: in previous work where bead-coupled enzymes were used in the workflow, 4 iterations were needed to achieve similar results[40]. Second, although the model is remarkably close for such a complex experiment, it does not capture some of the smaller dynamic changes observed in the data for the duration of the entire experiment, something we observed in previous work as well[11,38-40]. AMP and ADP are uncontrolled species within the experiment, thus the most difficult to predict accurately. For ADP, the model was less certain (shaded area) than ATP and NADP for most of the experimental runtime, but quantitatively accurate. For AMP, the model misses some smaller transients quantitatively (e.g., at 90 and 260 minutes). It appears as if the model response to the change in conditions inside the CSTR is not fast enough. This is also observed for the convergence of the model to the training data (**Figure 4**), where it appears the model could not 'keep the same pace'. Although technical noise can reduce the quality of the test data, it is unlikely that measurement errors sampled from normal distribution results in a transient curve of consecutive data points. Thus, some additional nonlinear interactions or environmental factors must be present (as hypothesized in **Figure 3**).

To explore if nonlinear interactions were missing, we added substrate and product inhibition terms to the model and used Hill type equations and subsequently extrapolated where each model performs best (**Supporting Information 3.4**). **Figure S13** and **Figure S14** summarize these efforts. We observe that adding product inhibition terms to the model causes better convergence for the initial part of experiment for AMP, whereas adding activation terms does so for the latter part of the experiments, but none led to a significant mean improvement. Nonetheless, with an $R^2 = 0.891$ we can state that the trained model has mapped the dynamics of the ERN accurately.
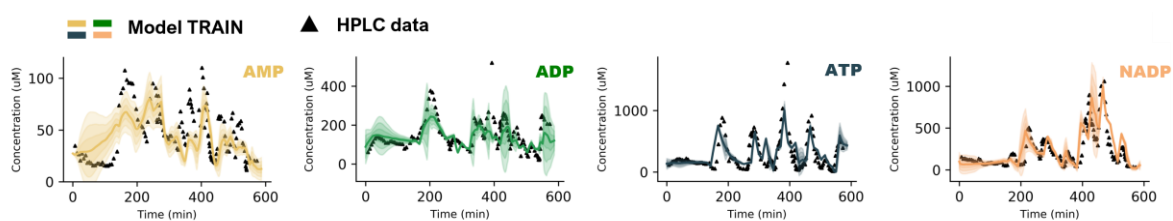
Figure 5. *Shows the result of the third step in the active learning cycle, training the model. The black diamonds represent the HPLC data of the first iteration, the line represent the mean prediction value and the shaded area the 1st, 2nd and 3rd standard deviation of that mean, 10 bet fits shown. More fit metric information shown in* ***Figure S10*** *and* ***Figure S11.***
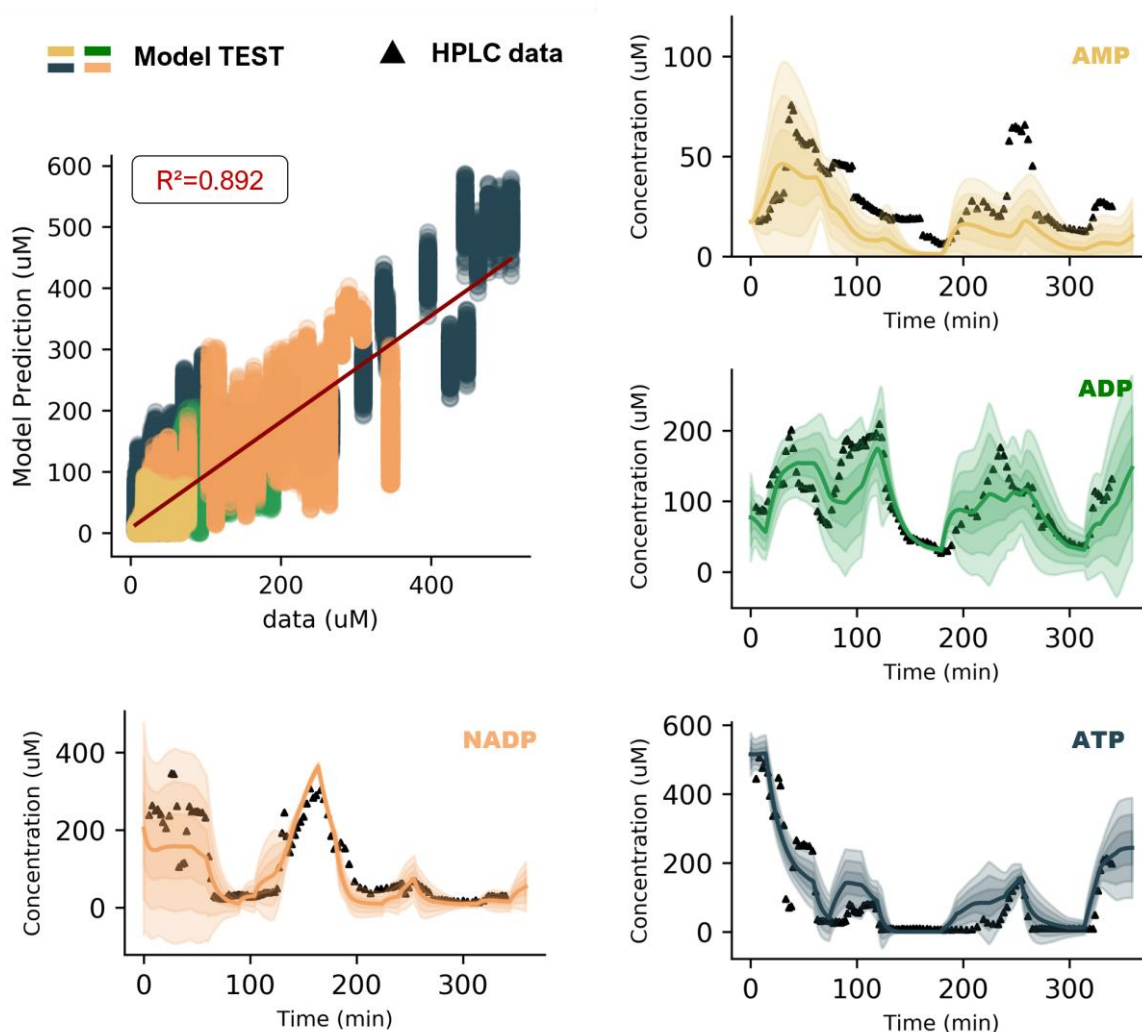
Figure 6. *Shows the result of the fourth step in the active learning cycle, testing the model. Utilizing the model trained on the first iteration of the database to predict the experiment performed in the second iteration. Top left figure shows a scatter plot with the model prediction values for each time point on the y-axis and the corresponding HPLC data points on the x-axis (R^2 = 0.891). The time-course is subsequently shown as well, with data as measured on the HPLC (diamond) and the model predictions for AMP and ADP, ATP and NADP plotted separately. The line represents the mean prediction value and the shaded area the 1st, 2nd and 3rd standard deviation of that mean. For these simulation the top 10 fits for the data produced in iteration 1 were chosen (see **Supporting Information 3.4**).*

Conclusion

In this work, we applied an active learning workflow (**Figure 1**) to study the dynamics of biocatalytic networks of free enzymes, specifically the PPP network (**Figure 2**). We showed that this workflow is especially useful when emergent interactions (e.g., PRPPS inhibition) are present, a persistent challenge in network assembly (**Figure 3** & **Figure 4**). Second, we showed that data from a single OED experiment were sufficiently complex to parameterize a kinetic model that captures all key sensitivities

governing the dynamics of the ERN in flow. Even though the addition of enzymes to the input space results in a more complex dynamic landscape (**Figure S9**), optimally perturbing enzymes and substrates enabled a 'one and done' approach, where a single OED iteration was sufficient to train a predictive model of 32 parameters (**Figure 5** & **Figure 6**).

Overall, the active learning workflow balances the complexity of the network (cross reactions included), the parameterization of the model, and the available information in the data (e.g., the number of observables or states perturbed). A change in each can result in longer or shorter optimization cycles[39-40]. Going forward, we can explore the utility of a single mechanistic model to capture the dynamics for all regions of the input space. In this work the model had sufficient predictive power to control biocatalysis, but we observed small deviations in the predictions as the environment changed rapidly, indicating that the assumptions embedded within the model's mechanistic description were not quantitatively accurate for the duration of the entire experiment [70-71]. These errors might propagate as networks become larger[24-25,71] or might hinder the forward design of functional networks, such as biosensors, which rely on the precise dynamic control for time or dose dependent responses[72]. At the same time, these errors shed light on the interactions that emerge in fully assembled networks. In future work, one could use the deviations from the model to design OED experiments that identify hitherto unknown interactions.

Since we do not want a return to protracted optimization processes, validating cross reactions or mapping the conformational state of each enzyme in different environments is not an option. Instead, we should aim to leverage ability of the active learning cycle to efficiently probe the sensitivities of the system and use this to either rapidly screen different topologies for different parts of the input space or train other types of models that can support the kinetic model. Recently, new generative AI tools using generative adversarial networks have been developed to rapidly identify ERN topologies and kinetics[73-76]. Additionally, machine learning has been introduced to control complex systems[77], including catalytic networks, which is promising[78]. Overall, we propose that any general solution to map ERN dynamics and enable 'plug and play' design requires an active learning workflow that probes input sensitivities as efficiently as possible akin to the method presented here.

Methods

A CSTR (volume 121 µl) made of poly(methyl methacrylate) equipped with four inlets, which were doubled to eight using Y connectors. The outlet tubes differed between experiments and had their own volume (109 µl and 69 µl for the first and second experiment, respectively, see **Figure S4**. Enzyme and substrate input flow rates were controlled by Cetoni neMESYS syringe pumps and Hamilton syringes and changed every 12 or 15 min. For more information about the flow set-up, including software for drop detector see **Supporting information 1.2-1.3**. HPLC analyses were performed on Shimadzu

Nexera X3 instrument. Conditions: Inertsil ODS-4 C18 column (3 μm pore size, 150 × 4.6 mm; GL Science) and a guard column (3 μm pore size; 10 × 4.6 mm) at 40 °C. Analysis was based off[59]. For more information on the buffer solutions used see **Supporting information 1.4.** Protocol for plasmid isolation of enzymes can be found in **Supporting information 1.5**[58]. The custom optimal experimental design software was written in Python 3.7, Delaware, USA. The code can be found at http://github.com/huckgroup/OED. Some elements of the software are present in previous work[79-80]. ODE solvers utilize AMICI, a C++ compiler for differentiation[82]. For more information on experimental design see **Supporting information 2**, previous applications[39-40] or excellent work by [42-47].

Data Availability

Raw data and notebooks used to generate the figures can be found at Huckgroup Github at https://github.com/huckgroup/OED/PPP. The estimated parameters for each of the modelling steps, the input flow profiles, and raw data can be found in the excel sheet available in the Github repository.

Acknowledgements

References

1. Ghosh, Souvik, et al. "Exploring Emergent Properties in Enzymatic Reaction Networks: Design and Control of Dynamic Functional Systems." *Chemical Reviews* 124.5 (2024): 2553-2582.
2. Burgener, S., Luo, S., McLean, R., Miller, T. E. & Erb, T. J. A roadmap towards integrated catalytic systems of the future. Nat. Catal. 3, 186–192 (2020).
3. Rasor, B. J. et al. Toward sustainable, cell-free biomanufacturing. Curr. Opin. Biotechnol. 69,136–144 (2021).
4. Cai, Tao, et al. "Cell-free chemoenzymatic starch synthesis from carbon dioxide." *Science* 373.6562 (2021): 1523-1527.
5. Zhang, Yi-Heng Percival, Jibin Sun, and Yanhe Ma. "Biomanufacturing: history and perspective." *Journal of Industrial Microbiology and Biotechnology* 44.4-5 (2017): 773-784.
6. Winkler, Christoph K., et al. "Accelerated reaction engineering of photo (bio) catalytic reactions through parallelization with an open-source photoreactor." *ChemPhotoChem* 5.10 (2021): 957-965.
7. Vázquez-González, Margarita, Chen Wang, and Itamar Willner. "Biocatalytic cascades operating on macromolecular scaffolds and in confined environments." *Nature Catalysis* 3.3 (2020): 256-273.
8. Benítez-Mateos, Ana I., David Roura Padrosa, and Francesca Paradisi. "Multistep enzyme cascades as a route towards green and sustainable pharmaceutical syntheses." *Nature chemistry* 14.5 (2022): 489-499.
9. Markin, C. J., et al. "Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics." *Science* 373.6553 (2021): eabf8761.
10. Jain, Shubhanshu, Felipe Ospina, and Stephan C. Hammer. "A New Age of Biocatalysis Enabled by Generic Activation Modes." *JACS Au* (2024).
11. Helwig, Britta, et al. "Bottom-Up Construction of an Adaptive Enzymatic Reaction Network." *Angewandte Chemie* 130.43 (2018): 14261-14265.
12. van Roekel, Hendrik WH, et al. "Programmable chemical reaction networks: emulating regulatory functions in living cells using a bottom-up approach." *Chemical Society Reviews* 44.21 (2015): 7465-7483.
13. Yue, Liang, et al. "Nucleic acid based constitutional dynamic networks: From basic principles to applications." *Journal of the American Chemical Society* 142.52 (2020): 21577-21594.
14. Teders, Michael, et al. "Reversible photoswitchable inhibitors generate ultrasensitivity in out-of-equilibrium enzymatic reactions." *Journal of the American Chemical Society* 143.15 (2021): 5709-5716.

15. Grubbe, W. S., Rasor, B. J., Krüger, A., Jewett, M. C., & Karim, A. S. (2020). Cell-free styrene biosynthesis at high titers. *Metabolic Engineering*, *61*, 89-95.
16. Korman, Tyler P., et al. "A synthetic biochemistry system for the in vitro production of isoprene from glycolysis intermediates." *Protein Science* 23.5 (2014): 576-585.
17. Ye, Xinhao, et al. "Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails." *ChemSusChem: Chemistry & Sustainability Energy & Materials* 2.2 (2009): 149-152.
18. Lin, Baixue, and Yong Tao. "Whole-cell biocatalysts by design." *Microbial cell factories* 16 (2017): 1-12.
19. Rollin, Joseph A., Tsz Kin Tam, and Y-H. Percival Zhang. "New biotechnology paradigm: cell-free biosystems for biomanufacturing." *Green chemistry* 15.7 (2013): 1708-1719.
20. Kwon, Yong-Chan, and Michael C. Jewett. "High-throughput preparation methods of crude extract for robust cell-free protein synthesis." *Scientific reports* 5.1 (2015): 8663.
21. Korman, Tyler P., Paul H. Opgenorth, and James U. Bowie. "A synthetic biochemistry platform for cell free production of monoterpenes from glucose." *Nature communications* 8.1 (2017): 15526.
22. Opgenorth, Paul H., Tyler P. Korman, and James U. Bowie. "A synthetic biochemistry module for production of bio-based chemicals from glucose." *Nature chemical biology* 12.6 (2016): 393-395.
23. Valliere, Meaghan A., et al. "A cell-free platform for the prenylation of natural products and application to cannabinoid production." *Nature communications* 10.1 (2019): 565.
24. Schwander, Thomas, et al. "A synthetic pathway for the fixation of carbon dioxide in vitro." *Science* 354.6314 (2016): 900-904.
25. Diehl, Christoph, et al. "Synthetic anaplerotic modules for the direct synthesis of complex molecules from CO2." *Nature chemical biology* 19.2 (2023): 168-175.
26. Giaveri, Simone, et al. "Integrated translation and metabolism in a partially self-synthesizing biochemical network." *Science* 385.6705 (2024): 174-178.
27. Karim, Ashty S., and Michael C. Jewett. "A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery." *Metabolic engineering* 36 (2016): 116-126.
28. Dudley, Quentin M., Kim C. Anderson, and Michael C. Jewett. "Cell-free mixing of Escherichia coli crude extracts to prototype and rationally engineer high-titer mevalonate synthesis." ACS synthetic biology 5.12 (2016): 1578-1588.
29. Dudley, Q. M., C. J. Nash, and M. C. Jewett. "Cell-free biosynthesis of limonene using enzyme-enriched Escherichia coli lysates. Synth. Biol. 4 (1), ysz003." (2019).
30. Morgado, G., Gerngross, D., Roberts, T. M. & Panke, S. Synthetic biology for cell-free biosynthesis: fundamentals of designing novel in vitro multi-enzyme reaction networks. Adv. Biochem. Eng. Biotechnol. 162, 117–146 (2018).
31. Hold, C., Billerbeck, S. & Panke, S. Forward design of a complex enzyme cascade reaction. Nat. Commun. 7, 12971 (2016).
32. Taylor, C. J. et al. Flow chemistry for process optimisation using design of experiments. J. Flow. Chem. 11, 75–86 (2021).
33. Bujara, M., Schumperli, M., Pellaux, R., Heinemann, M. & Panke, S. Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. Nat. Chem. Biol. 7, 271–277 (2011).
34. Shen, L. et al. A combined experimental and modelling approach for the Weimberg pathway optimisation. Nat. Commun. 11, 1098 (2020).
35. Pandi, A., Diehl, C., Yazdizadeh Kharrazi, A., Scholz, S. A., kova, E., Faure, L., ... & Erb, T. J. (2022). A versatile active learning workflow for optimization of genetic and metabolic networks. *Nature Communications*, *13*(1), 3876.
36. Borkowski, O., Koch, M., Zettor, A., Pandi, A., Batista, A. C., Soudier, P., & Faulon, J. L. (2020). Large scale active-learning-guided exploration for in vitro protein production optimization. *Nature communications*, *11*(1), 1872.
37. Rondelez, Yannick. "Competition for catalytic resources alters biological network dynamics." *Physical review letters* 108.1 (2012): 018102.
38. Baltussen, M. G., van de Wiel, J., Fernández Regueiro, C. L., Jakštaitė, M., & Huck, W. T. (2022). A Bayesian approach to extracting kinetic information from artificial enzymatic networks. *Analytical Chemistry*, *94*(20), 7311-7318.
39. Duez, Quentin, et al. "Quantitative Online Monitoring of an Immobilized Enzymatic Network by Ion Mobility–Mass Spectrometry." Journal of the American Chemical Society (2024).
40. van Sluijs, B., Zhou, T., Helwig, B., Baltussen, M. G., Nelissen, F. H., Heus, H. A., & Huck, W. T. (2024). Iterative design of training data to control intricate enzymatic reaction networks. *Nature Communications*, *15*(1), 1602.
41. van Sluijs, B., Maas, R. J., van der Linden, A. J., de Greef, T. F., & Huck, W. T. (2022). A microfluidic optimal experimental design platform for forward design of cell-free genetic networks. *Nature Communications*, *13*(1), 3626.
42. de Aguiar, P. Fernandes, et al. "D-optimal designs." Chemometrics and intelligent laboratory systems 30.2 (1995): 199-210.
43. Sinkoe, A. & Hahn, J. Optimal experimental design for parameter estimation of an IL-6 signaling model. Processes 5, 49 (2011)
44. Raue, A. et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25, 1923–1929 (2009).
45. Villaverde, Alejandro F., et al. "A protocol for dynamic model calibration." Briefings in bioinformatics 23.1 (2022): bbab387.
46. Fernandez Villaverde, Alejandro, et al. "Assessment of prediction uncertainty quantification methods in systems Biology." IEEE/ACM Transactions on Computational Biology and Bioinformatics: 15455963 (2023).
47. Villaverde, Alejandro F., et al. "Benchmarking optimization methods for parameter estimation in large kinetic models." Bioinformatics 35.5 (2019): 830-838.
48. Velasco-Lozano, Susana, et al. "Co-immobilization and colocalization of multi-enzyme systems for the cell-free biosynthesis of aminoalcohols." *ChemCatChem* 12.11 (2020): 3030-3041.
49. Dubey, Nidhi C., and Bijay P. Tripathi. "Nature inspired multienzyme immobilization: Strategies and concepts." *ACS Applied Bio Materials* 4.2 (2021): 1077-1114.
50. Tang, Zhongyao, Yuri Oku, and Tomoko Matsuda. "Application of Immobilized Enzymes in Flow Biocatalysis for Efficient Synthesis." *Organic Process Research & Development* 28.5 (2024): 1308-1326.
51. Reus, Bente, Matteo Damian, and Francesco G. Mutti. "Advances in cofactor immobilization for enhanced continuous-flow biocatalysis." *Journal of Flow Chemistry* 14.1 (2024): 219-238.
52. Ruscoe, Rebecca E., and Sebastian C. Cosgrove. "Future directions in flow biocatalysis: the impact of new technology on sustainability." *Current Opinion in Green and Sustainable Chemistry* (2024): 100954.
53. Paschalidis, Leandros, et al. "Modeling Enzymatic Cascade Reactions Immobilized in Plug-Flow Reactors for Flow Biocatalysis." *Chemie Ingenieur Technik* 96.6 (2024): 741-748.
54. Ma, Huan, et al. "Protocell Flow Reactors for Enzyme and Whole-cell Mediated Biocatalysis." *Advanced Materials* (2024): 2404607.
55. Pogodaev, Aleksandr A., et al. "Modular design of small enzymatic reaction networks based on reversible and cleavable inhibitors." Angewandte Chemie 131.41 (2019): 14681-14685.

56. Teders, Michael, et al. "Reversible photoswitchable inhibitors generate ultrasensitivity in out-of-equilibrium enzymatic reactions." Journal of the American Chemical Society 143.15 (2021): 5709-5716.
57. Bajić, Marijan, et al. "A paradigm shift for biocatalytic microreactors: Decoupling application from reactor design." *Biochemical Engineering Journal* 205 (2024): 109260.
58. Arthur, Patrick K., Luigi J. Alvarado, and T. Kwaku Dayie. "Expression, purification and analysis of the activity of enzymes from the pentose phosphate pathway." Protein expression and purification 76.2 (2011): 229-237.
59. Nakajima, Kazuki, et al. "Simultaneous determination of nucleotide sugars with ion-pair reversed-phase HPLC." *Glycobiology* 20.7 (2010): 865-871.
60. Cook, Paul F., and William Wallace Cleland. *Enzyme kinetics and mechanism*. Garland Science, 2007.
61. Mulcahy, Patricia, et al. "Application of kinetic-based biospecific affinity chromatographic systems to ATP-dependent enzymes: studies with yeast hexokinase." *Analytical biochemistry* 309.2 (2002): 279-292.
62. Velasco, Pilar, et al. "Purification, characterization and kinetic mechanism of glucose-6-phosphate dehydrogenase from mouse liver." *The International journal of biochemistry* 26.2 (1994): 195-200.
63. Hanau, Stefania, et al. "6-Phosphogluconate dehydrogenase mechanism: evidence for allosteric modulation by substrate." *Journal of Biological Chemistry* 285.28 (2010): 21366-21371.
64. Switzer, Robert L. "19. Phosphoribosylpyrophosphate Synthetase and Related Pyrophosphokinases." *The enzymes*. Vol. 10. Academic Press, 1974. 607-629.
65. Willemoës, Martin, Bjarne Hove-Jensen, and Sine Larsen. "Steady State Kinetic Model for the Binding of Substrates and Allosteric Effectors to Escherichia coli Phosphoribosyl-diphosphate Synthase." *Journal of Biological Chemistry* 275.45 (2000): 35408-35412.
66. Ruess, Jakob, et al. "Iterative experiment design guides the characterization of a light-inducible gene expression circuit." *Proceedings of the National Academy of Sciences* 112.26 (2015): 8148-8153.
67. Donzella, Silvia, and Martina Letizia Contente. "The joint effort of enzyme technology and flow chemistry to bring biocatalytic processes to the next level of sustainability, efficiency and productivity." *Journal of Flow Chemistry* 14.1 (2024): 85-96.
68. De Santis, Piera, Lars-Erik Meyer, and Selin Kara. "The rise of continuous flow biocatalysis–fundamentals, very recent developments and future perspectives." *Reaction Chemistry & Engineering* 5.12 (2020): 2155-2184.
69. Wilding, Kristen M., et al. "The emerging impact of cell-free chemical biosynthesis." *Current opinion in biotechnology* 53 (2018): 115-121.
70. Gunawardena, Jeremy. "Models in biology:'accurate descriptions of our pathetic thinking'." *BMC biology* 12 (2014): 1-11.
71. Rohwer, Johann M., Arno J. Hanekom, and Jan-Hendrik S. Hofmeyr. "A universal rate equation for systems biology." *Proc. 2nd Int. Symp. on Experimental Standard Conditions of Enzyme Characterizations (ESEC 2006), Beilstein Institute, Frankfurt am Main, Germany*. 2007.
72. Kurbanoglu, Sevinc, Cem Erkmen, and Bengi Uslu. "Frontiers in electrochemical enzyme based biosensors for food and drug analysis." *TrAC Trends in Analytical Chemistry* 124 (2020): 115809.
73. Choudhury, Subham, et al. "Reconstructing kinetic models for dynamical studies of metabolism using generative adversarial networks." *Nature Machine Intelligence* 4.8 (2022): 710-719.
74. Choudhury, Subham, et al. "Generative machine learning produces kinetic models that accurately characterize intracellular metabolic states." *bioRxiv* (2023): 2023-02.
75. Keyl, Philipp, et al. "Single-cell gene regulatory network prediction by explainable AI." *Nucleic Acids Research* 51.4 (2023): e20-e20.
76. Helleckes, Laura M., et al. "Machine learning in bioprocess development: from promise to practice." *Trends in biotechnology* 41.6 (2023): 817-835.
77. Dijkman, Jacobus, et al. "Learning Neural Free-Energy Functionals with Pair-Correlation Matching." arXiv preprint arXiv:2403.15007 (2024).
78. Owoyele, Opeoluwa, and Pinaki Pal. "ChemNODE: A neural ordinary differential equations framework for efficient chemical kinetic solvers." Energy and AI 7 (2022): 100118.
79. Sakai, Andrei, et al. "Cell-Free Expression System Derived from a Near-Minimal Synthetic Bacterium." ACS Synthetic Biology 12.6 (2023): 1616-1623.
80. Hu, Xinyu, et al. "ARTseq-FISH reveals position-dependent differences in gene expression of micropatterned mESCs." Nature Communications 15.1 (2024): 3918.
81. Fröhlich, Fabian, et al. "AMICI: high-performance sensitivity analysis for large ordinary differential equation models." Bioinformatics 37.20 (2021): 3676-3677.