



## PERSPECTIVE OPEN



# Towards a widespread adoption of metabolic modeling tools in biopharmaceutical industry: a process systems biology engineering perspective

Anne Richelle<sup>1</sup> , Blandine David<sup>1</sup> , Didier Demaegd<sup>1</sup>, Marianne Dewerchin<sup>1</sup>, Romain Kinet<sup>1</sup>, Angelo Morreale<sup>1</sup>, Rui Portela<sup>1</sup> , Quentin Zune<sup>1</sup> and Moritz von Stosch<sup>1</sup> 

In biotechnology, the emergence of high-throughput technologies challenges the interpretation of large datasets. One way to identify meaningful outcomes impacting process and product attributes from large datasets is using systems biology tools such as metabolic models. However, these tools are still not fully exploited for this purpose in industrial context due to gaps in our knowledge and technical limitations. In this paper, key aspects restraining the routine implementation of these tools are highlighted in three research fields: monitoring, network science and hybrid modeling. Advances in these fields could expand the current state of systems biology applications in biopharmaceutical industry to address existing challenges in bioprocess development and improvement.

*npj Systems Biology and Applications* (2020)6:6; <https://doi.org/10.1038/s41540-020-0127-y>

## INTRODUCTION

The emergence of high-throughput technologies is elevating the biotechnology field to the era of big data<sup>1</sup>. This shifting paradigm created considerable challenges for the interpretation of large datasets and the generation of meaningful outcomes impacting process and product attributes. Among the numerous computational methods developed in this context, metabolic modeling tools allowed the coherent organization of large datasets into biological networks providing nonintuitive insights on biological systems that in vivo experiments alone cannot provide<sup>2</sup>. These model-based approaches have been proven to be invaluable at the level of preclinical research. For example, these approaches have been used for the design of new drugs by informing target selection and for the engineering of cells by rewiring metabolism towards the production of a product of interest<sup>3,4</sup>.

Pharmaceutical companies are already investing substantially in computational approaches to guide drug discovery and cell design. While these important applications are out of the scope of this paper, we refer the interested readers to detailed reviews on the usage of metabolic engineering and synthetic biology in industrial context<sup>5–8</sup>. However, model-based methods can be used for much more than this. For instance, these methods could also be applied at the industrial level in the field of process design, monitoring and control to lower the experimental effort and increase the process robustness. Indeed, model-based methods would allow a more rational design of bioprocesses but also the objective identification of the variables to monitor and control. Therefore, the routine implementation of these methods can be expected to greatly facilitate the implementation of regulatory requirements related to the Quality by Design paradigm (i.e., systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control) and Process Analytical Technologies initiative (mechanism to design, analyze and control manufacturing processes through the measurement of Critical Process Parameters that affect Critical Quality Attributes<sup>9</sup>). Numerous modeling

methods exist to describe and quantify the reactions occurring in a cell under specific environments, from deterministic kinetic models to stochastic and statistical models. While kinetic modeling has been widely used for small-scale metabolic models, these mechanistic models are for the moment not scalable and suitable for genome-wide approaches due to challenges in acquiring the required kinetic parameters (i.e., rate constants, enzymes and intracellular metabolite concentrations) and computational problems associated with such complex parametric systems (i.e., model nonlinearity, parameter identifiability, and computational tractability)<sup>10</sup>. On the other hand, while statistical modeling is very useful for the manipulation of large complex datasets, these black-box approaches have limited capacity to generate biologically relevant information for bioprocess improvement. Therefore, the lack of clear workflows to set up specific process models based on genome-wide information have limited the potential scope of computational biology applications in industry.

However, the systems biology field is in continuous development and genome-scale metabolic networks are now available for most of the industrially relevant organisms. Therefore, we should be able to systematically predict cell metabolic behavior under different specific complex media compositions and process conditions. These types of predictions should be used to objectively determine the optimal operating conditions with respect to the desired cell phenotype and related production criteria. Unfortunately, while numerous studies have successfully used genome-scale models to rationally guide the culture process design<sup>11–13</sup>, their practical implementation in biopharmaceutical research and development facilities still seems in its infancy. Actually, bioprocess improvement performance is still mainly achieved by semi-empirical media and bioprocess optimization (i.e., screening different media and process conditions and analysis using statistical design-of-experiment strategies)<sup>14</sup>. In this paper, gaps in current knowledge and technical limitations restraining the widespread usage of metabolic modeling tools for process

<sup>1</sup>GSK, Rixensart, Belgium. ✉email: [anne.x.richelle@gsk.com](mailto:anne.x.richelle@gsk.com); [moritz.x.von-stosch@gsk.com](mailto:moritz.x.von-stosch@gsk.com)

development (i.e., design, optimization and control) in biopharmaceutical industry are highlighted. Specifically, research topics in three fields (real-time monitoring, biological network science and hybrid modeling) are identified as key drivers for evolving the current state of some systems biology tools in industrial process engineering applications. Advances in these fields will be invaluable to address existing challenges in bioprocess development and improvement.

### REAL-TIME BIOPROCESS MONITORING: TRACK TO BETTER CONTROL

Real-time monitoring is a key issue for effective bioprocess operation. Indeed, one needs first to recognize and analyze cellular physiological state in function of environmental conditions to be further able to adequately manipulate operating conditions. Therefore, real-time monitoring technologies development is a critical aspect of the bioprocess automatization. Previously, only a small number of so-called “process variables” were commonly measured online in bioreactors (e.g., pH, temperature, OD, stirring). The scope of available real-time probes has now been significantly expanded thanks to the recent advances in analytical technologies and the adoption of Process Analytical Technologies initiative<sup>15,16</sup>.

In the context of bioprocesses, real-time measurements of the metabolites and cell concentrations in the cell culture medium are key elements to enable the prediction of cell phenotype with respect to operating conditions (e.g., by using constraint-based modeling approaches)<sup>17</sup>. Unfortunately, some of these experimental measurements cannot currently be achieved online in situ and, doing so, limit our capacity to have a live snapshot of the cell metabolism during the culture process. However, continuous improvements in analytical technologies and the emergence of the single-cell experiments open new perspectives for gaining insights into cell metabolism.

Spectroscopic methods (e.g., NIR, FTIR, mid-IR, Raman, Fluorescence) have attracted an increasing interest for their potential capacity of simultaneous in situ measurement of nutrients, metabolites and cell concentrations. These techniques have been successfully integrated into real-time pharmaceutical manufacturing to provide a “molecular fingerprint” of samples enabling in-process correction needed to ensure the acquisition of a specific quality product<sup>18,19</sup>. However, the capacity to systematically extract accurate quantitative data is, for the moment, limited to a handful number of metabolites (typically, glucose, lactate, glutamine, glutamate, and ammonium). We refer the reader to the publications of Li et al.<sup>20</sup> and Ryder<sup>21</sup> for detailed comparison of the performance and limitations of existing spectroscopic methods. In this context, chemometric modeling has been a state-of-art technique to effectively extract the maximum of information from spectroscopic spectra. It is worth to highlight that spectroscopic sensors can in this sense be embedded in the concept of soft sensor<sup>22</sup> as they are used to compensate the lack of specific measurement by reconstructing the missing signals based on available measurements using data analysis models<sup>23</sup>. This type of hybrid approach will therefore continue to greatly benefit from the flourishing evolution of artificial intelligence tools and certainly allows to overcome the current limitations of spectroscopic monitoring methods<sup>24</sup>.

Parallely, the rise of online single-cell probing opens the door for the specific investigation of cell metabolism and bioenergetics at a definition level never captured before<sup>25</sup>. These “single-cell approaches” in combination with adequate analytical methods could help unravel the impact of operating conditions on the cellular physiology. Specifically, the new generation of high-throughput-omics technologies allow to envision in a close future a near real-time measurements of cell transcriptome and proteome<sup>26–28</sup>. Interestingly, very few examples of interfacing

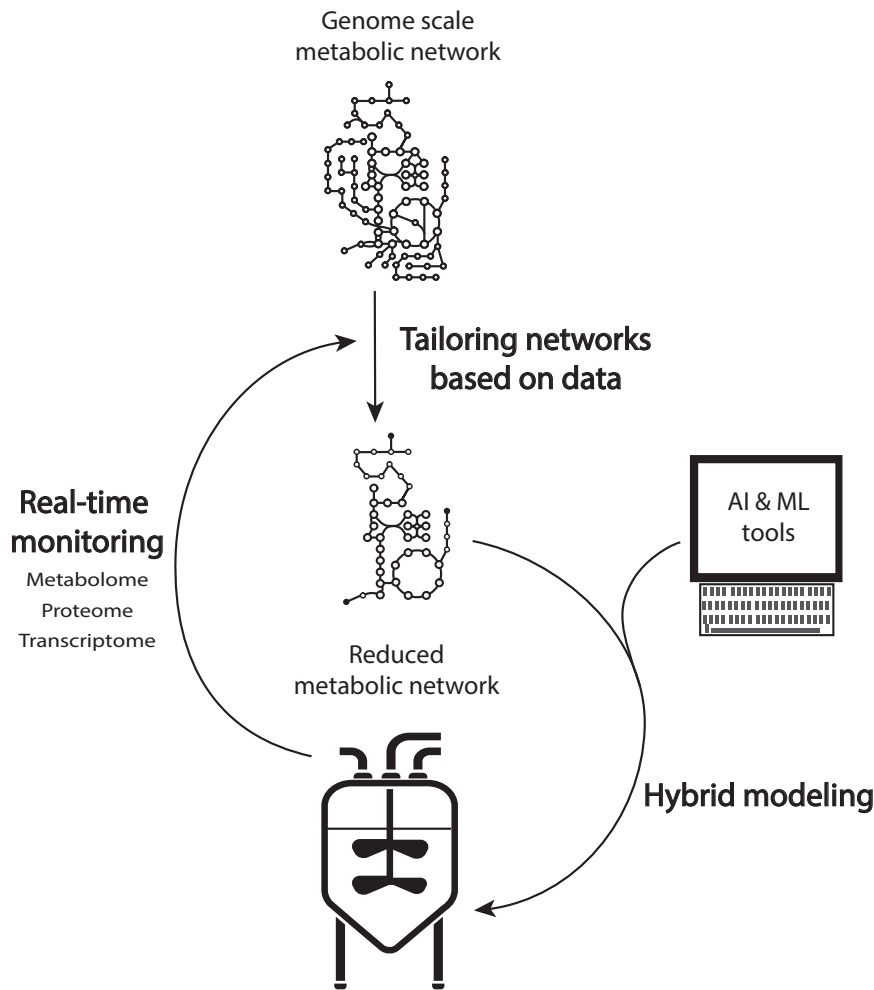
flow cytometry with a bioreactor and its use for effective bioreactor control can be found in the literature<sup>29</sup>.

### COMPLEXITY OF METABOLIC NETWORKS: REDUCE TO BETTER PREDICT

A genome-scale metabolic model is a network connecting all metabolic reactions that can occur in a specific organism with their associated metabolites, proteins and genes. While these models have been proven to be powerful tools for in silico simulation of cell metabolism<sup>12</sup>, the complexity of these large networks also hinders their utility in various practical applications. Actually, metabolic networks involving, most of the time, more reactions than metabolites, the associated system presents, generally, a multitude of solutions (i.e., underdetermined system). Therefore, metabolic network structures are also too complex to be handled for the development of utilitarian tools (e.g., optimization and control of a bioprocess based on cell phenotype prediction).

Different approaches have been proposed for tailoring these metabolic networks based on a priori knowledge and/or available experimental data. For example, algorithms like the one presented in Erdrich et al.<sup>30</sup> or Ataman et al.<sup>31</sup> allow an easy-to-implement systematic reduction of genome-scale networks into core models. However, these methods rely on the definition by the user of the part of the model that should be “protected” during the reduction (i.e., specific metabolites, reactions, phenotypes that need to be conserved) which is far from easy task. On the other hand, numerous algorithms have been developed to tailor genome-scale metabolic networks by integrating diverse types of -omics data, including transcriptomics, proteomics and metabolomics<sup>32–35</sup>. Actually, genome-scale models are used to link genes, enzymes, and metabolism through the use of gene-protein-reactions (GPR). Therefore, the knowledge of GPR is used by researchers to explain the metabolic state of a cell, based on the expression of metabolic genes and/or specific proteins<sup>36,37</sup>. The most common “integration algorithms” use gene expression data to recapitulate the metabolism of an organism under a specific condition (e.g., specific cell-line or operating conditions) by only extracting the subset of active reactions from the genome-scale model. These methods have been proven to be valuable for enhancing the accuracy of model-predicted growth rates and gene essentiality<sup>37</sup>. However, since currently no quantitative description of the GPR relationship exists in genome-scale metabolic models, the integration of gene expression data requires the use of strong assumptions to link the GPR expression and the metabolic reaction activity, which could lead to an oversimplification of the complex relation existing between fluxes, enzymes and genes<sup>36,38</sup>.

Finally, it is important to highlight that these “network tailoring” approaches generally do not completely leverage the problem of system underdetermination related to network complexity. Therefore, the choice of adequate strategy to solve the system (e.g., linear optimization techniques such as Flux Balance Analysis<sup>39</sup>) will always be required to achieve an instantaneous picture of the flux distributions in the cell. A recent research in this field presented a new approach using biochemical thermodynamic constraints (i.e., Gibbs energy) to shape the solution space of potential flux distributions<sup>40</sup>. The promising results obtained in this study pave a new way towards the understanding of the mechanisms that governs the definition of metabolic pathway usage across organisms and conditions. Further development of these computational methods will be crucial to obtain a holistic and integrated picture of the cell metabolism in function of specific environmental conditions while considering uncertainty associated with experiments, measurements and modeling.



**Fig. 1 From experimental data to bioprocess improvement.** Systematic workflow using data extracted from real-time monitoring to tailor genome-scale biological networks to core metabolic models that can be combined with artificial intelligence and machine learning tools for an effective implementation of control and optimization strategies.

### MODELING WITH HYBRID APPROACHES: COMBINE TO EFFECTIVELY IMPLEMENT

The application of artificial intelligence and machine learning in bioprocess engineering and systems biology has seen a significant increase in the last years thanks to the research advances in these fields<sup>41–43</sup> and their successful applications for e.g. in tumor detection<sup>44</sup>. In particular, advances in image recognition enable the automation of so far manual tasks, such as the possibility to automate the counting of cell colonies grown on petri-dishes, using digital imaging<sup>45</sup>. Surprisingly, the practical implementation of these hybrid approaches combining machine learning and the extended knowledge available about biological systems is still in its infancy in bioindustries.

Considering the availability of this prior knowledge, hybrid modeling and artificial intelligence approaches have the potential to significantly reduce the amount of experiments that need to be executed, while potentially increasing the region in which the model can reliably generate predictions. Actually, hybrid models seek to complement what is mechanistically known (e.g., about the metabolism and kinetics) with data-driven methods to describe the unknown parts. This type of models provides an attractive method for modeling biochemical processes as much of the underlying complexity can be lumped in the data-driven part.

Therefore, these approaches can be used to effectively establish the link between metabolism and operating conditions in a way that can be used to implement control and optimization strategies.

Despite their benefits and their first applications stemming from 1992<sup>46</sup>, the first application of hybrid modeling in systems biology only arose in 2010<sup>47</sup> and few have been reported ever since<sup>48–53</sup>. One reason might be the absence of dedicated software that would allow for a straightforward development of these type of models. Another reason might be the broad interdisciplinarity of competences required to implement these approaches (i.e., knowledge from engineering fields and data sciences that need ideally to be complemented with a biological background). Finally, there is still several major open questions that need to be investigated to allow a better combined usage of hybrid modeling and metabolic modeling tools<sup>54</sup>. From the effective generation of sufficiently informative experimental data to the joint parameter identification of the mechanistic and data-driven parts of the model, the assessment of these gaps will be critical for the development of a systematic workflow for integrating knowledge from systems biology into model structures enabling a next-generation technology for the life-sciences and biotechnology sectors, besides others (Fig. 1).

## BRIDGING THE GAPS TOWARDS EFFECTIVE CONTINUOUS PROCESSING

Process intensification through continuous processing has been adopted for decades in industries such as chemical and petrochemical. In the context of biopharmaceutical production, the move from batch to continuous or semi-continuous mode of production has been slowed down mainly due to the more conservative approach of this industry resulting from the stringent quality/regulatory requirements. Therefore, innovation in biotechnology has become more product- than process-centered<sup>55</sup>. While these old drivers of bioprocess development are still valid, new emerging challenges are forcing the change towards extended manufacturing flexibility and transferability. The commitment to deliver on them includes process development timeline decrease, productivity increase and product quality modulation (e.g., for biosimilarity).

The proper implementation of continuous manufacturing requires a thorough understanding of the process to achieve and maintain steady-state conditions (i.e., the causalities between physiological state, bioprocess parameters, productivity and product quality attributes). Furthermore, the in-process quality should be evaluated and ensured by the real-time analysis of process data. To this end, in-line process monitoring system linked to robust model-based control will be crucial technologies to support this real-time release testing<sup>56,57</sup>.

Despite their advantages, the potential of metabolic modeling tools is still not fully exploited to address these challenges in bioprocess development and improvement. This is mainly due to the upfront time and effort investment required to implement these approaches and gain in process understanding as critical gaps remain to set these methods up in clear workflows. Empirical process development techniques will continue to play an important role in addressing these challenges. However, systems biology-driven approaches are likely to be more impactful to provide insights into the cellular responses to process changes<sup>2,14</sup>.

Finally, the research topics, highlighted in this review, in the field of monitoring, network science and hybrid modeling, will also be cutting-edge technologies to improve new emergent biopharmaceutical production platforms such as cell-free systems. Actually, cell-free protein synthesis has been emerging as a flexible and powerful platform to address challenges in biomanufacturing by overcoming inherent limitations related to the use of living cells. Cell-free systems offer the ability to design metabolic pathways towards the production of desired products but also to synthesize complex proteins with unnatural amino acids and to build up artificial cells. Despite these promising features, challenges remain such as the proper control of post-translational modifications and the expansion of the genetic code for unnatural amino acids incorporation. To address these challenges, targeted gene editing and addition of adequate exogenous substances to control reaction conditions will be needed to optimally regulate the transcription and translation<sup>58,59</sup>. In this context, computational tools will be invaluable to systematically identify system limitations and areas of improvement for production efficiency, as demonstrated in the work of Vilkhovoy et al.<sup>60</sup>.

## DISCUSSION

We highlighted gaps in our knowledge and technical limitations restraining the effective application of metabolic modeling tools throughout the lifecycle of a biopharmaceutical process. Specifically, advances in real-time monitoring of bioprocesses, biological network modeling and their combination with data-driven approaches will be the key drivers to lower the time and cost associated with the development of new drugs<sup>4</sup>. Taken together, these technologies could facilitate the switch to continuous

processing in biopharmaceutical industry but also foresee needs for improvements in emergent biomanufacturing platforms such as cell-free systems.

Received: 10 October 2019; Accepted: 12 February 2020;

Published online: 13 March 2020

## REFERENCES

- Sadowski, M. I., Grant, C. & Fell, T. S. Harnessing QbD, programming languages, and automation for reproducible biology. *Trends Biotechnol.* **34**, 214–227 (2016).
- Chen, C., Le, H. & Goudar, C. T. Integration of systems biology in cell line and process development for biopharmaceutical manufacturing. *Biochem. Eng. J.* **107**, 11–17 (2016).
- Kubinyi, H. In *Computer Applications in Pharmaceutical Research and Development* (eds Wang, B. & Ekins, S.) 377–424 (John Wiley & Sons, 2006).
- Schmidt, B. J., Papin, J. A. & Musante, C. J. Mechanistic systems modeling to guide drug discovery and development. *Drug Discov. Today* **18**, 116–127 (2013).
- Choi, K. R. et al. Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* **37**, 817–837 (2019).
- Nielsen, J. & Keasling, J. D. Engineering cellular metabolism. *Cell* **164**, 1185–1197 (2016).
- Jain, K. K. Synthetic biology and personalized medicine. *Med. Princ. Pract.* **22**, 209–219 (2013).
- Trosset, J.-Y. & Carbonell, P. Synthetic biology for pharmaceutical drug discovery. *Drug Des. Devel. Ther.* **9**, 6285–6302 (2015).
- Rathore, A. S. & Winkle, H. Quality by design for biopharmaceuticals. *Nat. Biotechnol.* **27**, 26–34 (2009).
- Penas, D. R., González, P., Egea, J. A., Doallo, R. & Banga, J. R. Parameter estimation in large-scale systems biology models: a parallel and self-adaptive cooperative strategy. *BMC Bioinforma.* **18**, 52 (2017).
- Burgard, A. P., Vaidyaraman, S. & Maranas, C. D. Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake. *Environments* **17**, 791–797 (2001).
- Selvarasu, S. et al. Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture. *Biotechnol. Bioeng.* **109**, 1415–1429 (2012).
- Zhang, C. & Hua, Q. Applications of genome-scale metabolic models in biotechnology and systems medicine. *Front. Physiol.* <https://doi.org/10.3389/fphys.2015.00413> (2016).
- Kroll, P., Hofer, A., Ulonska, S., Kager, J. & Herwig, C. J. P. R. Model-based methods in the biopharmaceutical process lifecycle. *Farm Res.* **34**, 2596–2613 (2017).
- Gernaey, K. V. et al. Monitoring and control of microbioreactors: an expert opinion on development needs. *Biotechnol. J.* **7**, 1308–1314 (2012).
- Guerra, A., von Stosch, M. & Glassey, J. Toward biotherapeutic product real-time quality monitoring. *Crit. Rev. Biotechnol.* **39**, 289–305 (2019).
- Bordbar, A., Monk, J. M., King, Z. A. & Palsson, B. O. Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* **15**, 107–120 (2014).
- Baradez, M.-O., Bizziato, D., Hassan, E. & Marshall, D. Application of Raman spectroscopy and univariate modelling as a process analytical technology for cell therapy bioprocessing. *Front. Med. (Lausanne)* **5**, 47–47 (2018).
- Buckley, K. & Ryder, A. G. Applications of Raman spectroscopy in biopharmaceutical manufacturing: a short review. *Appl. Spectrosc.* **71**, 1085–1116 (2017).
- Li, M., Ebel, B., Chauchard, F., Guédon, E. & Marc, A. Parallel comparison of in situ Raman and NIR spectroscopies to simultaneously measure multiple variables toward real-time monitoring of CHO cell bioreactor cultures. *Biochem. Eng. J.* **137**, 205–213 (2018).
- Ryder, A. G. Cell culture media analysis using rapid spectroscopic methods. *Curr. Opin. Chem. Eng.* **22**, 11–17 (2018).
- Luttmann, R. et al. Soft sensors in bioprocessing: a status report and recommendations. *Biotechnol. J.* **7**, 1040–1048 (2012).
- Esmonde-White, K. A. et al. Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing. *Anal. Bioanal. Chem.* **409**, 637–649 (2017).
- Storey, E. E. & Helmy, A. S. Optimized preprocessing and machine learning quantitative Raman spectroscopy in biology. *J. Raman Spectrosc.* **50**, 958–968 (2019).
- Vasdekis, A. E. & Stephanopoulos, G. Review of methods to probe single cell metabolism and bioenergetics. *Metab. Eng.* **27**, 115–135 (2015).
- Erickson, B. K. et al. Active instrument engagement combined with a real-time database search for improved performance of sample multiplexing workflows. *J. Proteome Res.* **18**, 1299–1306 (2019).

27. Doerr, A. Single-cell proteomics. *Nat. Methods* **16**, 20–20 (2019).
28. Rhoads, A. & Au, K. F. PacBio Sequencing and its applications. *Genom. Proteom. Bioinforma.* **13**, 278–289 (2015).
29. Delvigne, F. & Goffin, P. Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnol. J.* **9**, 61–72 (2014).
30. Erdrich, P., Steuer, R. & Klamt, S. An algorithm for the reduction of genome-scale metabolic network models to meaningful core models. *BMC Syst. Biol.* **9**, 48 (2015).
31. Ataman, M., Hernandez Gardiol, D. F., Fengos, G. & Hatzimanikatis, V. redGEM: systematic reduction and analysis of genome-scale metabolic reconstructions for development of consistent core metabolic models. *PLoS Comput. Biol.* **13**, e1005444 (2017).
32. Aurich, M. K., Fleming, R. M. T. & Thiele, I. MetaboTools: a comprehensive toolbox for analysis of genome-scale metabolic models. *Front. Physiol.* **7**, 327–327 (2016).
33. Berger, B., Peng, J. & Singh, M. Computational solutions for omics data. *Nat. Rev. Genet.* **14**, 333–346 (2013).
34. Bersanelli, M. et al. Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinforma.* **17**, S15 (2016).
35. García Martín, H. et al. A method to constrain genome-scale models with <sup>13</sup>C labeling data. *PLoS Comput. Biol.* **11**, e1004363 (2015).
36. Machado, D. & Herrgård, M. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* **10**, e1003580 (2014).
37. Opdam, S. et al. A systematic evaluation of methods for tailoring genome-scale metabolic models. *Cell Syst.* **4**, 318–329.e316 (2017).
38. Richelle, A., Joshi, C. & Lewis, N. E. Assessing key decisions for transcriptomic data integration in biochemical networks. *PLoS Comput. Biol.* **15**, e1007185 (2019).
39. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
40. Niebel, B., Leupold, S. & Heinemann, M. An upper limit on Gibbs energy dissipation governs cellular metabolism. *Nat. Metab.* **1**, 125–132 (2019).
41. Angermueller, C., Pärnamaa, T., Parts, L. & Stegle, O. Deep learning for computational biology. *Mol. Syst. Biol.* **12**, 878 (2016).
42. Miotto, R., Wang, F., Wang, S., Jiang, X. & Dudley, J. T. Deep learning for healthcare: review, opportunities and challenges. *Brief. Bioinforma.* **19**, 1236–1246 (2018).
43. Oyetunde, T., Bao, F. S., Chen, J.-W., Martin, H. G. & Tang, Y. J. Leveraging knowledge engineering and machine learning for microbial bio-manufacturing. *Biotechnol. Adv.* **36**, 1308–1315 (2018).
44. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).
45. Ferrari, A., Lombardi, S. & Signoroni, A. Bacterial colony counting with convolutional neural networks in digital microbiology imaging. *Pattern Recogn.* **61**, 629–640 (2017).
46. Psychogios, D. C. & Ungar, L. H. A hybrid neural network-first principles approach to process modeling. *AIChE J.* **38**, 1499–1511 (1992).
47. von Stosch, M., Peres, J., de Azevedo, S. F. & Oliveira, R. Modelling biochemical networks with intrinsic time delays: a hybrid semi-parametric approach. *BMC Syst. Biol.* **4**, 131 (2010).
48. Carinhas, N. et al. Hybrid metabolic flux analysis: combining stoichiometric and statistical constraints to model the formation of complex recombinant products. *BMC Syst. Biol.* **5**, 34–34 (2011).
49. Ferreira, A. R. et al. Projection to latent pathways (PLP): a constrained projection to latent variables (PLS) method for elementary flux modes discrimination. *BMC Syst. Biol.* **5**, 181–181 (2011).
50. Hamilton, F., Lloyd, A. L. & Flores, K. B. Hybrid modeling and prediction of dynamical systems. *PLoS Comput. Biol.* **13**, e1005655 (2017).
51. Isidro, I. A. et al. Hybrid metabolic flux analysis and recombinant protein production in *Pichia pastoris* X-33 cultures expressing a single-chain antibody fragment. *Bioprocess. Biosyst. Eng.* **39**, 1351–1363 (2016).
52. Teixeira, A. P. et al. Cell functional enviromics: unravelling the function of environmental factors. *BMC Syst. Biol.* **5**, 92 (2011).
53. von Stosch, M., Carinhas, N. & Oliveira, R. In *Large-Scale Networks in Engineering and Life Sciences* (eds Benner, P. et al.) 367–388 (Springer International Publishing, 2014).
54. von Stosch, M., Oliveira, R., Peres, J. & Feyo de Azevedo, S. Hybrid semi-parametric modeling in process systems engineering: past, present and future. *Computers Chem. Eng.* **60**, 86–101 (2014).
55. Warikoo, V. et al. Integrated continuous production of recombinant therapeutic proteins. *Biotechnol. Bioeng.* **109**, 3018–3029 (2012).
56. Chaudary, R. S., Pazhayattil, A. & Spes, J. Continuous manufacturing: a generic industry perspective. *Pharmaceut. Technol.* <http://www.pharmtech.com/continuous-manufacturing-generic-industry-perspective> (2017).
57. Hernandez, R. Continuous manufacturing: a changing processing paradigm. *Bio-Pharm. Int.* **28**. <http://www.biopharminternational.com/continuous-manufacturing-changing-processing-paradigm> (2015).
58. Jiang, L., Zhao, J., Lian, J. & Xu, Z. Cell-free protein synthesis enabled rapid prototyping for metabolic engineering and synthetic biology. *Synth. Syst. Biotechnol.* **3**, 90–96 (2018).
59. Lu, Y. Cell-free synthetic biology: engineering in an open world. *Synth. Syst. Biotechnol.* **2**, 23–27 (2017).
60. Vilkhovoy, M. et al. Sequence specific modeling of *E. coli* cell-free protein synthesis. *ACS Synth. Biol.* **7**, 1844–1857 (2018).

## ACKNOWLEDGEMENTS

The development of this article was sponsored by GlaxoSmithKline Biologicals SA. All authors were employees of the GSK group of companies at the time of the development of this article.

## AUTHOR CONTRIBUTIONS

A.R. and M.v.S. wrote and edited the manuscript. All authors were involved in drafting the content and structure of the manuscript and approved the final version.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41540-020-0127-y>.

**Correspondence** and requests for materials should be addressed to A.R. or M.v.S.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020