观点

DOI: 10.12211/2096-8280.2026-069

虚拟细胞赋能生物制造：从菌株设计到生产过程优化

戴振武[1,2,3]，钱鎏佳[1,2,3]，蒋恒[1,2,3]，孙瑞[1,2,3]，周梓卓[1,2,3]，郭天南[1,2,3]
([1] 西湖大学医学院附属杭州市第一人民医院，医学蛋白质组学全国重点实验室，西湖大学医学院，西湖大学未来生物医药学院，浙江 杭州 310024；[2] 西湖实验室（生命科学和生物医学浙江省实验室），西湖智能蛋白质组中心，浙江 杭州 310024；[3] 西湖大学生命科学学院，西湖大学未来产业研究中心，浙江 杭州 310024)

摘要：人工智能虚拟细胞通过整合生物学先验知识、机制模型与多模态实验数据，为细胞状态表征、扰动响应预测和机制分析提供了新的计算框架。用于生物制造时，虚拟细胞具有将细胞响应与生产性能联系、服务菌株设计与培养优化的潜力，但模型输出与过程速率及生产性能的关系尚未建立，培养环境和放大过程中的设备与工艺约束也难以纳入。本文在回顾虚拟细胞的发展与主要建模方向的基础上，提出了面向生物制造的虚拟细胞应用框架，讨论其在底盘选择、菌株改造、培养优化和生产放大中的应用。虚拟细胞模型与生物过程模型可通过底物摄取、耗氧和产物生成等速率耦合，并借助缩小尺度实验检验细胞响应与生产性能之间的关系。在本文提出的虚拟细胞支持的生物制造平台中，模型预测候选方案的影响，智能体组织证据检索、模型调用与实验验证。实验记忆保存运行记录与结果，为跨轮次的设计—构建—测试—学习循环和单批次培养过程调控提供依据。分析表明，当前的生物制造可为虚拟细胞提供具有明确目标和约束的验证场景，而随着模型与实验证据的积累，虚拟细胞有望支持工程规律的归纳，并拓展至新底盘、多菌株协同、多产物共生产及质量与安全控制。

关键词：人工智能；虚拟细胞；生物制造；微生物细胞工厂；生物过程工程

中图分类号：Q81；TP18　　**文献标志码：**A

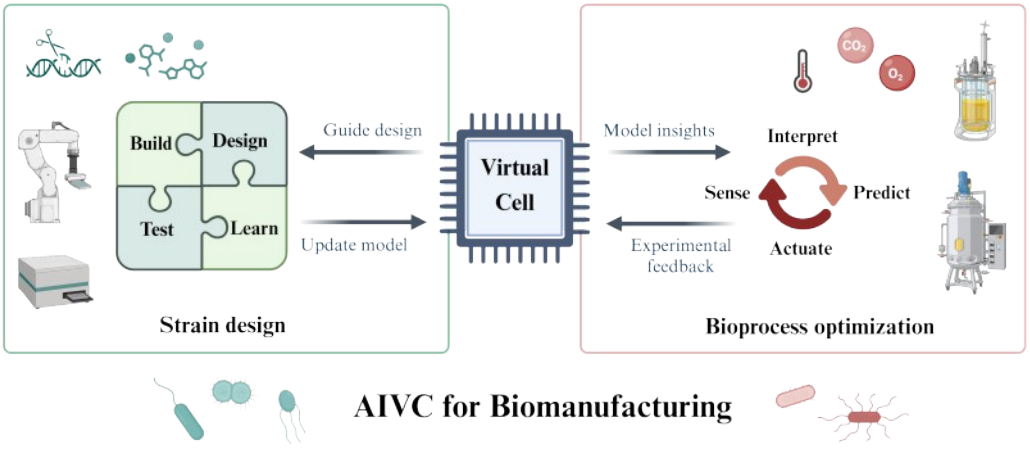
Virtual cells for biomanufacturing: from strain design to bioprocess optimization

DAI Zhenwu[1,2,3], QIAN Liujia[1,2,3], JIANG Heng[1,2,3], SUN Rui[1,2,3], ZHOU Zizhuo[1,2,3], GUO Tiannan[1,2,3]
([1]*Affiliated Hangzhou First People's Hospital, State Key Laboratory of Medical Proteomics, School of Medicine, School of Future Biomedicine, Westlake University, Hangzhou 310024, Zhejiang, China*; [2]*Westlake Center for Intelligent Proteomics, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou 310024, Zhejiang, China*; [3]*Research Center for Industries of the Future, School of Life Sciences, Westlake University, Hangzhou 310024, Zhejiang, China*)

Abstract: Artificial intelligence virtual cells (AIVCs) integrate biological prior knowledge, mechanistic models, and multimodal experimental data to provide a computational framework for representing cellular states, predicting

收稿日期：2026-07-30 修回日期：2026-11-12
引用本文：戴振武，钱鎏佳，蒋恒，孙瑞，周梓卓，郭天南. 虚拟细胞赋能生物制造:从菌株设计到生产过程优化[J]. 合成生物学，2026，7. DOI：10.12211/2096-8280.2026-069
Citation：DAI Zhenwu，QIAN Liujia，JIANG Heng，SUN Rui，ZHOU Zizhuo，GUO Tiannan. Virtual cells for biomanufacturing: from strain design to bioprocess optimization[J]. Synthetic Biology Journal，2026，7. DOI：10.12211/2096-8280.2026-069

perturbation responses, and interpreting cellular mechanisms. Their application to biomanufacturing requires linking cellular responses to production performance while accounting for cultivation conditions, equipment and process constraints. Strain engineering often combines multiple modifications whose effects depend on host background and cultivation conditions, while scale-up changes mixing, mass transfer, and the local environments experienced by cells. Quantitative predictions of these responses could help compare candidate designs and evaluate whether laboratory performance can be maintained at production scale. This article examines the development of AIVCs and major modeling approaches, focusing on microbial cell factories and their applications in chassis selection, strain engineering, cultivation optimization, and scale-up. AIVC and bioprocess models may be coupled through substrate uptake, oxygen consumption, and product formation rates to describe interactions between cellular states and cultivation environments. Scale-down experiments can reproduce relevant environmental fluctuations and test these relationships in candidate strains. In the envisioned AIVC-supported biomanufacturing platform, models predict the effects of candidate designs, while AI agents coordinate evidence retrieval, model use, and experimental validation. Experimental memory retains strain backgrounds, cultivation conditions, operational records, and results to support successive design-build-test-learn cycles. During individual cultivation runs, observations and model predictions may inform a sense-interpret-predict-actuate cycle, with feedback used to adjust operations under predefined constraints. Engineering applications require attention to measurement timing, spatial representativeness, and prediction reliability across conditions. Explanations for abnormal behavior need to be assessed against cultivation histories as well as equipment and operational records. Validation should establish whether model-guided decisions improve production performance, reduce experimental effort, and remain useful across independent batches and relevant scales. Biomanufacturing provides AIVCs with experimental tasks, explicit objectives, and engineering constraints through which these capabilities can be tested and refined. As models and experimental evidence accumulate, AIVCs may help derive engineering principles with defined ranges of applicability, supporting new chassis development, functional module adaptation, multi-strain production, multi-product co-production, and quality and safety control.



Created in BioRender. lab, G. (2026) https://BioRender.com/i8hqktm

Keywords: artificial intelligence; virtual cell; biomanufacturing; microbial cell factory; bioprocess engineering

人工智能（artificial intelligence，AI）正在从分子水平的结构和功能研究扩展到细胞及组织等尺度的状态表征和扰动响应预测。全细胞机制模型、高通量实验和多组学测量技术的进步提升了

人们对于不同物种和类型的细胞组成、功能及动态变化的系统描述能力[1-3]。在此基础上发展的人工智能虚拟细胞（AI virtual cell，AIVC）结合了生物学先验知识与数据驱动方法，表征细胞状态并预测其对扰动的响应，进而指导实验设计，为探索细胞功能及疾病机制提供支持[4]。

面向生物制造，人工智能已应用于酶挖掘[5]、调控序列设计[6]、功能蛋白质设计[7]和发酵过程数字化建模[8]。相较于面向特定分子、工程环节或过程变量的应用，虚拟细胞需要与菌株改造和生产过程相结合，通过模型预测与实验反馈支持工程优化。本文简要回顾虚拟细胞的发展历史及近期进展，以微生物细胞工厂和生产工艺为主要对象，讨论其在底盘筛选、工艺开发和生产放大中的应用，以及与设计—构建—测试—学习（Design-Build-Test-Learn，DBTL）循环的结合。

1 虚拟细胞研究现状

细胞建模早期主要依据已知生化反应及调控关系模拟细胞行为。1997年，E-Cell系统通过统一的反应规则，将基因表达、能量代谢和磷脂合成等过程纳入细胞仿真[9]。2012年Karr等构建的生殖支原体全细胞模型进一步整合DNA复制等过程，展示了由基因型预测多类细胞表型的可能性[1]。在酿酒酵母中，Yeast-GEM等共识代谢模型的扩展研究引入了酶容量、热力学和蛋白质资源分配等约束，将酶表达与资源分配对代谢通量和细胞生长的影响纳入预测[10-14]。机制模型依据已知反应网络和约束条件推演细胞表型，其覆盖范围与动态预测能力仍取决于机制知识和参数表征是否充分。

高通量扰动筛选与多维测量技术的发展，使细胞状态及其对遗传和环境扰动的响应得以大规模表征。以Perturb-seq[15]，Perturb-ATAC[16]为代表的单细胞扰动筛选将混合遗传扰动与转录组、染色质可及性等分子测量结合；光学混合筛选（optical pooled screening，OPS）通过显微成像等技术将扰动与细胞形态、亚细胞定位及动态表型关联[17]。蛋白质组[3]、代谢组[18]等测量手段也应用于大规模扰动实验，形成了多层次的细胞响应图谱。

随着实验数据积累与机器学习方法发展，从大规模测量中表征细胞状态及计算扰动响应成为生物学的重要研究方向。2024年Bunne等提出了较完整的人工智能虚拟细胞建设框架，通过结合多模态数据与机器学习模型跨尺度表征和模拟细胞行为，支持扰动响应预测与虚拟实验[4]。围绕这一目标，Biohub通过虚拟生物学计划推动大规模多模态数据建设[19]；Arc Institute开发State[20]等扰动响应预测模型，并组织虚拟细胞挑战赛以推动模型评测[21]；SciLifeLab的Alpha Cell计划分阶段整合蛋白质分布、空间结构与时间动态信息，逐步推进细胞状态预测[22]。这些工作共同推动虚拟细胞从概念框架走向数据建设、模型开发和标准化等相关评测。

2 状态表征与扰动响应建模

细胞状态可通过基因表达、蛋白质丰度和代谢活动等特征描述；状态表征模型则从这些测量中学习可用于比较与预测的数值表示[4]。以Geneformer[23]、scGPT[24]和scFoundation[25]为代表的单细胞基础模型从大规模转录组数据中学习细胞表示。SATURN[26]和UCE[27]支持跨物种的细胞状态比较与知识迁移，CAPTAIN[28]则联合转录组和细胞表面蛋白测量，支持多模态细胞状态表征与跨模态预测（表1）。

细胞状态的表征还需要考虑分子间关系、空间组织及功能模块。D-SPIN[31]从大规模单细胞扰动数据中推断基因或基因程序之间的相互作用，建立可解释的调控网络模型，并描述不同扰动条件下细胞状态分布的变化。Nicheformer[32]和VirTues[33]分别利用空间转录组和多重蛋白质成像，将细胞状态与邻域组成及组织结构联系起来。虚拟酵母蓝图提出以机制知识、亚细胞结构和动态状态为数据基础，将酿酒酵母的膜系统、遗传信息处理和线粒体能量代谢等细胞过程组织为相互关联的功能模块[42-43]。

在状态表征基础上，扰动响应模型进一步预测遗传或环境变化引起的状态转换。State[20]结合细胞状态表征与转换模型，预测不同细胞背景下

表1 虚拟细胞状态表征与扰动响应的主要建模方向

Table 1 Major modeling approaches for cell-state representation and perturbation response in virtual cells

建模方向 Modeling focus	主要能力 Key capabilities	代表性模型 Representative models
大规模细胞状态表征	从大规模转录组数据中学习基因与细胞表示，通过任务适配支持细胞类型注释、状态识别或跨数据集整合	Geneformer[23]、scGPT[24]、scFoundation[25]、Tahoe-x1[29]
跨物种表征	结合分子序列信息与基因表达数据建立跨物种共享细胞表示，支持细胞状态比较与注释迁移	SATURN[26]、UCE[27]
多模态整合	联合转录组、细胞表面蛋白等测量学习多模态细胞表示，支持跨模态预测或数据整合	CAPTAIN[28]、scGPT[24]
调控网络推断	推断基因或基因程序间的网络关系，提出候选调控因子，解释细胞状态及扰动响应	scPRINT[30]、D-SPIN[31]
组织与微环境表征	结合空间测量表征细胞邻域与组织结构，关联细胞状态和局部微环境	Nicheformer[32]、VirTues[33]
扰动响应预测	根据细胞背景与扰动条件预测表达响应或群体状态分布，并预测未直接测量条件下的响应	State[20]、CellFlow[34]、PerturbDiff[35]
因果效应估计	在明确的模型假设下区分混杂变异与扰动效应，估计干预效应并开展反事实分析	CINEMA-OT[36]、scCausalVI[37]
时间动态建模	从时间序列或单细胞状态分布学习分子响应动态，预测扰动响应或推断细胞状态转换路径	ProteinTalks[38]、DynNet[39]
知识约束与机制解释	结合知识图谱与检索证据预测扰动后的基因表达变化，生成可供检验的作用路径和机制解释	VCWorld[40]、AROMA[41]

的遗传和药物扰动响应。CellFlow[34]和PerturbDiff[35]通过生成式建模预测扰动后的细胞状态分布。CINEMA-OT[36]等因果推断方法通过区分潜在混杂因素与扰动相关变化，并在已观测条件之间构建反事实细胞匹配，推断单细胞水平的处理效应。

相较于预测扰动后终点状态，时间动态建模进一步为扰动后状态演化规律的推断提供了基础。ProteinTalks[38]基于大规模时间分辨的扰动蛋白质组数据，利用神经常微分方程（neural ordinary differential equations，Neural ODE）学习扰动相关的动态表征，支持药物响应及组合效应预测。DynNet[39]等模型也利用动态信息描述细胞状态转换。多个遗传或环境扰动的联合建模有望进一步用于分析单项扰动难以揭示的协同、拮抗和非线性相互作用。

在菌株开发与工艺优化中，跨物种与多模态建模有望用于底盘比较和改造靶点筛选，空间与时间建模可用于探索区室分工、产物富集和细胞状态转换。VCWorld[40]和AROMA[41]还分别针对药物与单基因扰动，结合知识图谱、证据检索和语言模型推理，预测基因表达变化并提出可能的作用路径。将这些方法用于工程微生物时，可围绕具体改造靶点或培养条件，结合目标宿主的实验评价其对方案筛选和结果解释的帮助。现有模型主要输出分子响应、细胞状态表示或状态分布；要用于菌株设计和培养优化，还需进一步建立这些输出与底物摄取、耗氧、产物生成等过程速率及生产性能之间的关系。

3 从细胞工厂开发到生产过程优化

细胞工厂开发涉及底盘选择、代谢路径设计、表达水平调节、转运与区室化、耐受性改良以及培养条件优化等相互关联的工程决策（表2）。其中，基于基因组尺度代谢模型的计算方法已广泛用于代谢改造靶点筛选。FSEOF（flux scanning based on enforced objective flux）[60]通过逐步提高目标产物通量约束，识别随目标通量增加而发生系统性变化的反应通量，据此提出过表达等候选干预靶点。ecFactory[61]进一步结合酶容量约束，预测过表达、下调和敲除等候选干预，并筛选组合方案。这类方法主要依据代谢通量、化学计量约束和酶需求提出改造建议，难以描述遗传干预

表2 常见菌株改造与培养方案决策

Table 2 Common decisions in strain engineering and cultivation control

类别 Category	决策方向 Decision area	具体工程决策 Specific engineering decisions	代表性文献 Representative references
底盘选择与菌株改造	底盘选择	比较候选宿主的合成能力、底物利用能力、耐受性及遗传操作条件，选择适合目标产物与生产环境的底盘	[44-45]
	代谢路径与通量优化	比较候选合成路径、酶及其变体，确定需增强、削弱或阻断的反应，协调前体、能量与还原力供给	[46-47]
	表达水平与资源分配	确定各基因的表达强度与多基因表达比例，选择启动子及基因拷贝数，协调酶需求与细胞资源供给	[48-50]
	转运与区室化设计	选择转运体和亚细胞定位方案，调整底物摄取、产物外排及区室间的物质交换	[47,51-52]
	耐受性与生产稳定性改良	比较耐受性、遗传稳定性及生产者群体稳定性改造方案，确定与生产目标相匹配的选择压力和筛选条件	[53-55]
培养与调控策略	培养基与操作条件优化	确定培养基组成、补料方式及供氧、pH和温度的操作范围，并制定不同培养阶段的调整方案	[45,56-57]
	生物传感与动态调控	选择传感信号与调控元件，确定响应阈值和切换时机，设计协调细胞生长与产物合成的调控策略	[58-59]

引起的调控网络响应及其随时间变化的细胞状态。

菌株改造常涉及多个靶点，其组合效果还受到宿主背景、改造顺序和培养条件的影响。面对这一需求，虚拟细胞建模的目标是进一步整合生物学先验、扰动实验和多模态测量，预测组合改造引起的细胞状态变化，并评估这些变化在不同培养条件下对生产性能的影响[4, 20, 38]，这要求模型输出从分子响应进一步延伸至与产量、产率、生产强度和稳定性相关的工程表型。高通量文库筛选[53, 62]、靶点验证和适应性实验室进化[63]可持续产生新的基因型—表型与环境响应数据；研究者可在同一实验条件下配对获取分子测量与生产性能数据，用于筛选候选菌株、检验模型预测并更新后续改造策略[43]（图1(a)）。

实验室中获得的生产优势能否在规模放大后仍然保持，还取决于菌株对培养环境梯度变化的响应。在实验室规模的反应器中，补料、供氧、pH和温度可按照设定范围调节以协调细胞生长与产物形成，罐内的底物浓度与溶氧基本均匀。随着培养体积增大，混合时间延长，传质条件变化，两者在罐内出现空间梯度。细胞随液体循环反复穿过补料口附近的高底物区与远离补料口的低底物区，并可能在高底物区经历局部氧限制[64-65]，这些环境波动可改变细胞生长、代谢及生产表现（表3）。因此，将实验室获得的菌株性能转化为稳定的生产表现，通常需要在生产相关条件下开展进一步验证，并根据放大过程中暴露出的限制调整培养工艺[69]。

微生物生产过程由实验室规模向工业生产规模转化，往往伴随着较长的开发周期和较高投入[69]。生物过程建模可用于分析培养过程中的关键限制并优化操作条件，为放大过程中的方案比较和工艺调整提供依据（表4）。物料衡算与动力学模型描述细胞生长、底物消耗和产物形成，并可与代谢通量模型耦合[70]。结合流动与传质模型，还可分析反应器中的环境分布及细胞暴露轨迹，将混合与传质条件对这些过程的影响纳入预测[65, 71]。细胞适应环境变化需要时间，不同的环境经历可能形成群体内的状态差异，群体平衡模型可进一步描述细胞状态分布及其演化[72]。

上述模型与实际培养过程的在线或离线测量结合后，可进一步用于过程数字孪生，实现状态估计、过程预测和操作方案比较[77-78]。混合建模和迁移学习也为模型在不同培养条件和生产过程之间的适配提供了方法[79]。然而，准确预测溶氧、底物浓度、生物量等过程变量，并不意味着模型已充分描述细胞内部状态及其调控响应。原料组成与培养环境的变化可能影响生产表现[80]；相同过程条件下，菌株背景与培养历史的差异也可能对应不同的细胞状态和生产表现。

虚拟细胞模型与生物过程模型的耦合有望通过物质交换与代谢速率，将细胞内部状态与外部

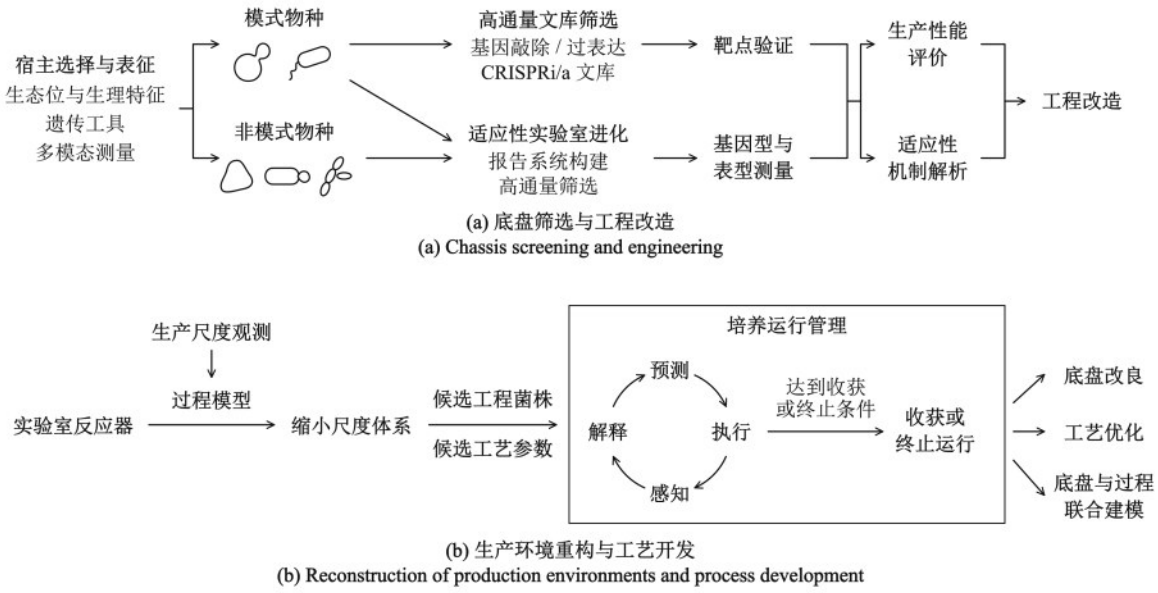


图1 细胞工厂开发与生产过程优化流程

（a）根据宿主生态、生理特征、遗传工具及多模态测量选择并表征候选底盘。（b）生产尺度观测与过程模型用于构建缩小尺度体系，评价候选菌株和工艺参数在生产相关环境波动下的表现。

Fig.1 Workflows for cell factory development and bioprocess optimization

(a) Chassis selection and characterization integrate information on ecological niches, physiological traits, genetic tools, and multimodal measurements. (b) Production-scale observations and process models guide the construction of scale-down systems that reproduce production-relevant environmental fluctuations and enable comparison of candidate engineered strains and process parameters.

表3 生产相关环境波动的实验模拟

Table 3 Experimental simulation of production-relevant environmental fluctuations

宿主 Host	研究目标 Research objective	实验体系 Experimental system	模拟的环境波动 Simulated environmental fluctuations	主要结果 Key findings	参考文献 References
大肠杆菌（*E. coli*）	生长与代谢响应	搅拌罐-活塞流反应器双区室体系（STR-PFR）	高葡萄糖区与低葡萄糖区交替暴露，并考察局部缺氧	配套的22 m³培养出现甲酸积累和生物量得率下降；缩小尺度实验诱发混合酸发酵及相关应激响应	[64]
谷氨酸棒杆菌（*C. glutamicum*）	单细胞生长响应	动态微流控单细胞培养体系	按CFD计算的环境轨迹，交替施加葡萄糖充足与饥饿条件	相比平均葡萄糖可用量相同的恒定供给条件，波动条件下生长速率降低约40%	[66]
酿酒酵母（*S. cerevisiae*，工业菌株Ethanol Red）	代谢与转录响应	单罐间歇补料恒化培养体系（工作体积1.7 L）	厌氧条件下周期性底物短缺，每周期停料2 min、恢复补料7 min	引发短时转录应激及长期资源分配调整；乙醇得率变化不显著	[67]
中国仓鼠卵巢细胞（CHO）	生长、代谢与产品质量响应	搅拌罐-活塞流反应器双区室体系（STR-PFR；搅拌罐规格5 L）	在不同溶氧区域间循环，通过改变PFR停留时间调节细胞的环境经历	相对参考补料培养，PFR停留时间超过90 s时乳酸明显积累，稳定期活细胞密度最多下降15%；培养液中宿主蛋白杂质增加，重组蛋白生产率未受影响	[68]

注：STR，搅拌罐反应器；PFR，活塞流反应器；CFD，计算流体力学；CHO，中国仓鼠卵巢细胞。

表4 生物过程建模的常用模型与方法

Table 4 Common models and methods for bioprocess modeling

技术路线 Modeling approach	主要方法 Main methods	主要功能 Main functions
生长与代谢动力学建模	物料衡算、反应动力学、动态通量平衡[70]	描述生长、底物消耗与产物形成，连接细胞代谢和培养过程
流动与传质建模	CFD、拉格朗日粒子追踪[65]；动态分区模型[71]	解析环境梯度与细胞暴露轨迹，支持放大分析和缩小尺度设计
群体平衡建模	群体平衡方程、矩方法[72]	描述细胞状态分布及其随生长、分裂和环境变化的演化
机理－数据混合建模	UDE[8]；半参数混合模型、PINN[73]	利用机制约束与过程数据学习未知动力学，预测培养过程
数据驱动方程辨识	SINDy[74]；符号回归[75]	从数据中筛选动力学关系，构建简洁、可解释的模型
概率辨识与状态估计	贝叶斯参数推断[76]；无迹卡尔曼滤波[77]	利用新测量更新参数分布与状态估计，预测过程状态并量化不确定性

注：CFD，计算流体力学；UDE，通用微分方程；PINN，物理信息神经网络；SINDy，非线性动力学稀疏辨识。

培养环境联系起来。过程模型根据设备结构与操作条件计算细胞所经历的底物、溶氧等环境轨迹[65，71]。若过程模型提示周期性低氧，虚拟细胞可以结合环境变化速率与代谢适应过程，预测应激和代谢状态，并检验这些状态能否解释底物摄取或产物生成速率的变化，辅助判断低氧限制能否通过供氧与补料调控缓解，或进一步通过菌株改造提高低氧适应能力。环境变化速率对代谢响应的影响已在酿酒酵母连续培养中被描述：突然提高稀释率可迅速诱发乙醇溢流，缓步提高则有利于呼吸适应并延迟发酵启动[81]。DynNet[39]和ProteinTalks[38]等工作表明动态模型可以从时间分辨测量中学习细胞状态的演化规律，为连接环境变化与细胞响应轨迹提供方法基础，但环境与细胞响应的具体关系仍需在工程宿主中建立和验证。

缩小尺度体系为验证上述细胞—过程耦合关系提供了可控实验环境。依据目标生产设备的过程观测及相应过程模型，可在实验室反应器中重现影响细胞状态与生产性能的关键环境波动，并比较候选工程菌株和培养策略[65]。缩小尺度设计还需考虑目标设备的供氧、混合及操作能力，以检验菌株在小试条件下获得的生产优势能否在目标生产环境中保持[69]。实验结果可进一步用于调整候选菌株与培养策略，并校正模型中的细胞响应与过程速率关系（图1(b)）。

4 虚拟细胞支持的生物制造平台：构建与挑战

菌株改良与工艺优化涉及相互依赖的多种策略，通常需要通过多轮DBTL循环开展设计、干预与验证。生物铸造厂将菌株构建、培养和检测组织为标准化、可复用的操作流程[82]，iBioFoundry及自动化编辑筛选平台进一步连接液体处理、培养、测量和数据回传，为多轮实验的执行与比较提供基础[83-84]。Automated Recommendation Tool（ART）利用机器学习与不确定性估计，根据已有实验数据推荐下一轮候选设计，为数据驱动的DBTL迭代提供了实例[85]。Coutant等结合计算模型与自动化实验平台，通过三轮实验迭代改进酿酒酵母双相生长转换模型，说明了实验反馈也可用于持续改进模型[86]。近年提出的自主研究（autonomous research）将假说生成、实验设计与结果解释组织为迭代过程，在研究者设定的目标与约束下，根据实验证据调整后续方案[87]。iCloudBiofoundry将研究规划与机器人实验相连接，并在酶与菌株工程中利用实验结果调整后续设计[88]。

面向具体生产任务，本文提出将虚拟细胞模型、智能体、自动化实验平台与实验记忆相结合，构建支持菌株改良与工艺优化的生物制造平台。虚拟细胞模型有望用于预测候选菌株改造与培养方案对细胞状态和生产性能的影响，智能体则围绕生产目标组织证据检索与模型调用，比较候选方案，识别预测中不确定或需要补充验证的问题，并据此规划实验任务（图2）[42]。自动化实验平台负责按照确定的方案执行菌株构建、培养和测量任务。在学习环节，智能体结合实际执行条件比较模型预测与测量结果，分析偏差并提出下一轮待验证假设。由此，实验反馈既可用于调整菌株改造与培养方案，也可用于修正模型，加深对于

细胞响应的理解。实验记忆既需保存当前周期的菌株背景、培养条件、操作记录和测量结果，也需跨周期保留方案选择依据及既往实验结果，使后续检索能够区分不同条件下获得的证据，并为方案比较和模型修正提供依据。例如，GPT-5驱动的无细胞蛋白合成研究将试剂批号、液体处理错误和实际孵育时长与蛋白滴度数据一并提供给模型，用于后续反应配方设计[89]。

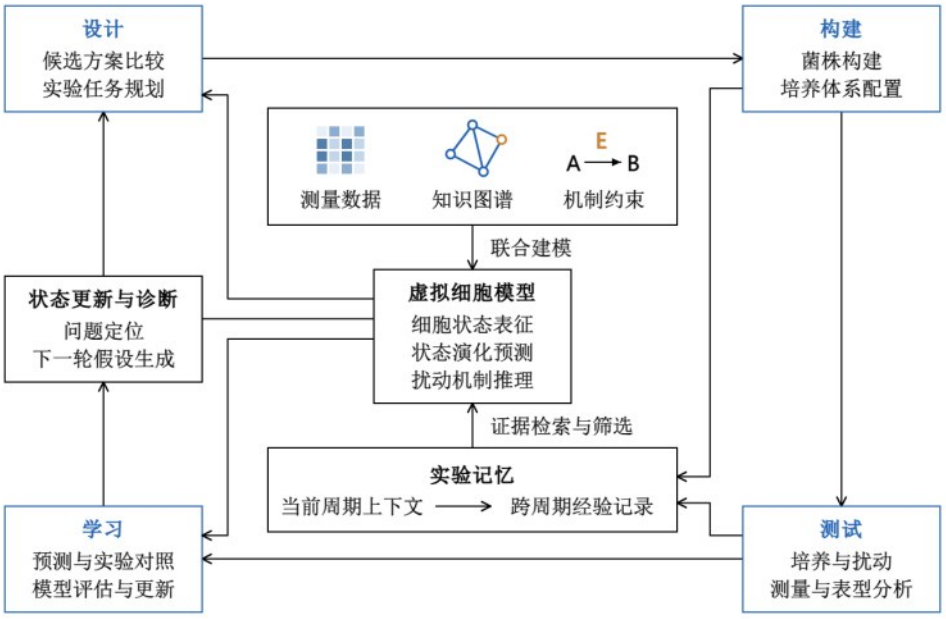


图2 虚拟细胞支持的生物制造设计—构建—测试—学习（DBTL）循环框架示意

Fig.2 Proposed framework for a virtual cell–supported design–build–test–learn (DBTL) cycle in biomanufacturing

在该平台中，跨轮次的DBTL迭代与单批次培养过程调控相互衔接。缩小尺度体系可用于验证单批次培养的感知—解释—预测—执行循环，运行数据则可反馈至后续DBTL迭代（图1(b)）。在过程分析技术（process analytical technology，PAT）框架下，溶氧与尾气监测、拉曼光谱及离线检测可与过程模型结合，支持细胞状态估计与培养过程调控[77, 90]。智能体可结合培养历史与当前观测调用模型，解释当前状态并比较可能的后续响应。控制系统则可采用模型预测控制（model predictive control，MPC）等方法，在预设操作约束下滚动优化补料[91]、供氧及搅拌策略[92]。该循环持续至达到收获或异常终止条件，运行结束后，智能体将观测数据与操作记录整理纳入实验记忆，支持下一轮DBTL设计。

然而，要通过这一循环持续改善生产性能，仍需克服多重挑战。首先，不同测量具有不同的时间分辨率和反馈延迟。在线过程信号与离线组学测量的采样频率和分析时长不同，当组学结果返回时，工程菌株可能已进入新的培养阶段[77]。平台需结合采样时间与培养记录，明确各项测量对应的培养状态并考虑分析与操作延迟，判断其能否用于当前调控。未能及时用于当前调控的组学结果，仍可用于解释本轮实验、修正模型和设计下一轮方案。

其次，现有测量只能部分观测细胞及其培养环境。例如，单点溶氧测量可能无法代表所有细胞经历的氧供给[93]。平台需考虑传感器与取样位置，结合流动与传质模型评价环境测量的代表性[71]，并通过多模态测量与受控扰动检验对细胞状态的判断。细胞状态与生产性能之间的联系还需在目标菌株和培养条件下通过独立批次验证，扩展至其他菌株或条件时则需重新评价其适用性。

再次，同一异常表型可能对应不同的生物学或工程原因。生产中的异常读数可能来源于细胞代谢或培养环境变化，也可能由设备运行或操作

问题引起，还可能是多种因素共同作用的结果，因此基于既往实验获得的解释不能直接套用于当前批次。智能体应结合培养历史、设备与操作记录核查异常，保留并评估可能的解释，先利用已有数据和必要检测缩小原因范围，再选择能够区分这些解释的测量或干预。验证方案需兼顾成本、耗时与设备条件，优先解决影响当前操作、后续菌株改造或放大决策的关键问题。

5 展 望

虚拟细胞建模的一个核心挑战是准确预测细胞对遗传改造和外界扰动的响应。生物制造则需要联合设计遗传改造与培养策略，提高生产性能，并检验这种优势能否在生产尺度的设备与工艺约束下保持，据此进一步调整菌株与工艺。因此，生物制造既对虚拟细胞提出明确的工程需求，也

表5 面向生物制造的开放问题与研究切入点

Table 5 Open questions and research entry points for biomanufacturing

研究方向 Research area	关键开放问题 Key open questions	研究切入点 Proposed research approaches	代表性研究 Representative studies
新物种底盘开发	在实验数据有限时，已有物种的知识能否支持新宿主的生产潜力评估，并预测遗传改造与培养条件对生产性能的影响？	结合序列与生理表型筛选候选宿主，通过虚拟扰动实验比较遗传改造与培养方案，以缩小实验筛选范围、减少测试轮次	非模式酵母 *I. orientalis* 的模型重建、表型验证与琥珀酸生产的代谢工程改造[97]；跨物种酶动力学参数预测及酶约束模型构建[98]
条件依赖的功能解析与模块适配	代谢、应激或调控相关的功能基因模块对生产的贡献，在多大程度上依赖宿主与培养条件？	结合多组学测量与扰动实验，解析功能基因模块的组成，比较其在不同宿主和培养条件下对生产的贡献，指导模块的选择、组合与改造	不同环境下的酵母遗传互作比较[99]；同一遗传逻辑模块在不同细菌宿主背景中的性能差异[100]
复杂原料条件下的预测可靠性	复杂原料的组成与可利用性变化，是否会改变候选菌株和培养方案的生产表现及优劣次序？人工智能模型能否预测这种变化，并识别模型不适用的原料条件？	交叉比较候选菌株在不同来源、批次原料中的生产表现，关联原料特征、细胞响应和生产结果；用留出的原料来源或批次数据检验预测及方案排序，并通过针对性实验识别造成差异的关键因素	评估不同玉米浆批次对大肠杆菌和枯草芽孢杆菌培养的影响，并通过针对性营养补充识别限制因素[101]
多菌株协同生产	在群体比例、空间分布和培养环境变化时，菌株间的营养互补与通路分工能否维持稳定生产？	从同种工程菌株组成的体系出发，联合预测代谢交换与群体动态，逐步拓展至跨物种组合，检验通路分工和培养条件对生产效率与稳定性的影响	以甘油为底物，将粘康酸合成途径拆分到两株大肠杆菌中，并调整两株的接种比例[102]；不同酵母物种间的通路分工与中间体转移[103]
多产物共生产	能否联合设计不同类型产物的合成路径与培养条件，通过前体共享和资源协调，使共生产获得优于分别生产的整体收益？	将产物组合与比例作为设计变量，预测前体、能量与还原力分配，比较通路配置和生产时序，并将分离纯化成本纳入收益评价	在同一酵母菌株中整合麦角硫因、红景天苷和 gadusol 的合成通路，实现三种产物共生产[104]
污染风险防控	宿主与培养条件的联合设计能否降低杂菌污染及噬菌体感染风险，同时兼顾生产性能与长期防控效果？	结合菌株竞争与噬菌体感染建模，比较选择性营养、抗噬菌体改造及培养调控方案，通过受控挑战实验评价防控效果与生产代价	建立非常规氮源（三聚氰胺、氰胺）和磷源（亚磷酸盐）的利用途径，使工程菌在非无菌培养中获得竞争优势[105]；将大肠杆菌的两个丝氨酸密码子转而编码亮氨酸，使受试噬菌体的蛋白质发生错译，从而阻断其感染[106]
质量控制与安全评价	菌株改造与培养条件对产物质量、杂质形成和生物安全风险的影响能否被预测，并用于指导生产设计与质量控制？	结合菌株状态、多组学测量和产品检测，建立菌株与培养条件对关键质量属性、宿主来源杂质及生物安全指标的影响关系；根据质量属性与产品功能和生产风险的关联确定控制优先级，并验证不同批次和生产条件下模型预测的适用范围	糖基化工程调整重组糖蛋白的糖链组成[107]；改造大肠杆菌脂多糖合成，以不激活人 TLR4 的脂质 IVA 替代常规脂多糖，降低重组蛋白制备物的内毒素活性[108]；通过改造必需酶建立宿主对非天然氨基酸的依赖，限制其脱离供给后的生长，用于生物遏制[109]

注：研究切入点为本文提出的方向；代表性研究对应已有的工程研究。

为其预测能力、适用范围和跨条件迁移能力提供了可检验的应用场景。面向生物制造的虚拟细胞需要预测的不仅是平均生产性能，还包括不同批次和生产条件下性能与质量属性的分布及其变化范围。

合成生物学通过表征和标准化生物元件支持模块化设计与生物系统构建，但有关宿主选择、表达剂量、通路负担和培养适配的工程经验，仍有待通过跨项目比较归纳为可量化的设计依据[94]。药物发现中的利平斯基五规则（Lipinski’s rule of five）表明，大量实验经验可以被归纳为简单而具有实用价值的筛选判据[95]。细胞工程中，Ceroni等在大肠杆菌中测量细胞的剩余表达容量（即细胞承担现有表达任务后还能支持多少额外基因表达），据此在输出相近的表达设计中筛选出细胞负担更低的方案[49]。生物过程工程通过传质系数等指标将反应器操作与传质能力联系起来，为设备设计和过程放大提供依据[69]。随着模型与实验证据的积累，虚拟细胞有望连接机制知识、扰动数据与工程经验，支持工程规律的归纳，并检验这些规律在不同宿主和培养条件下的适用范围。

依托虚拟细胞模型，研究者可进一步比较同一工程策略在不同遗传背景、培养条件与生产尺度下的效果，并分析造成差异的因素。这些知识可用于指导功能模块在新宿主中的适配及连续定向进化的选择条件设计[96]，并支持多菌株协同、多产物共生产等复杂生物制造体系的探索（表5）。近期提出的虚拟细胞世界模型（virtual cell world model，VCWM）[110]和生物医学世界模型[111]进一步强调持续更新的内部细胞状态、条件化干预模拟与长期动态预测，并将研究对象从单细胞拓展至群落、类器官和组织等复杂系统，为研究复杂生物制造体系中的多细胞相互作用、环境耦合和长期生产状态变化提供了参考。

虚拟细胞在生物制造中的价值不应仅以分子层面的预测精度衡量，还应通过其能否改善工程决策来检验，包括提高候选改造方案的有效性，以及减少实验轮次和试错成本。这种增益需要与现有代谢模型和过程模型比较，以识别虚拟细胞的新增价值。虚拟细胞还需要协助判断候选菌株和工艺能够达到怎样的生产水平，以及生产性能和关键质量属性在不同批次的生产条件下能否保持稳定[112-113]。面向产业应用，研究者还需进一步考虑原料与能源消耗、工程放大能力及生产成本，评价候选方案的系统可行性[114]。当前虚拟细胞模型、智能体、自动化实验与生产过程建模等组成部分的成熟度仍不一致，需要计算生物学、细胞生物学、合成生物学与过程工程持续协作，在明确模型适用范围和实际工程收益的基础上，逐步拓展虚拟细胞在生物制造中的应用。

参 考 文 献

[1] KARR JR, SANGHVI JC, MACKLIN DN, et al. A Whole-Cell Computational Model Predicts Phenotype from Genotype[J]. Cell, 2012, 150(2): 389-401. DOI: 10.1016/j.cell.2012.05.044.

[2] REPLOGLE JM, SAUNDERS RA, POGSON AN, et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq[J]. Cell, 2022, 185(14): 2559-2575.e28. DOI: 10.1016/j.cell.2022.05.013.

[3] MESSNER CB, DEMICHEV V, MUENZNER J, et al. The proteomic landscape of genome-wide genetic perturbations[J]. Cell, 2023, 186(9): 2018-2034. e21. DOI: 10.1016/j.cell.2023.03.026.

[4] BUNNE C, ROOHANI Y, ROSEN Y, et al. How to build the virtual cell with artificial intelligence: Priorities and opportunities[J]. Cell, 2024, 187(25): 7045-7063. DOI: 10.1016/j.cell.2024.11.015.

[5] WANG Z, XIE D, WU D, et al. Robust enzyme discovery and engineering with deep learning using CataPro[J]. Nature Communications, 2025, 16(1): 2736. DOI: 10.1038/s41467-025-58038-4.

[6] KOTOPKA BJ, SMOLKE CD. Model-driven generation of artificial yeast promoters[J]. Nature Communications, 2020, 11(1): 2113. DOI: 10.1038/s41467-020-15977-4.

[7] MADANI A, KRAUSE B, GREENE ER, et al. Large language models generate functional protein sequences across diverse families[J]. Nature Biotechnology, 2023, 41(8): 1099-1106. DOI: 10.1038/s41587-022-01618-2.

[8] BANGI MSF, KAO K, KWON JSI. Physics-informed neural networks for hybrid modeling of lab-scale batch fermentation for β-carotene production using *Saccharomyces cerevisiae*[J]. Chemical Engineering Research and Design, 2022, 179: 415-423. DOI: 10.1016/j.cherd.2022.01.041.

[9] TOMITA M, SHIMIZU T, SAITO K, et al. E-CELL: Software Environment for Whole Cell Simulation[J]. Genome Informatics, 1997, 8: 147-155. DOI: 10.11234/gi1990.8.147.

[10] ZHANG C, SÁNCHEZ BJ, LI F, et al. Yeast9: a consensus genome-scale metabolic model for *S. cerevisiae* curated by the community[J]. Molecular Systems Biology, 2024, 20(10): 1134-1150. DOI: 10.1038/s44320-024-00060-7.

[11] SÁNCHEZ BJ, ZHANG C, NILSSON A, et al. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints[J]. Molecular Systems Biology, 2017, 13(8): 935. DOI: 10.15252/msb.20167411.

[12] OFTADEH O, SALVY P, MASID M, et al. A genome-scale metabolic model of *Saccharomyces cerevisiae* that integrates expression constraints and reaction thermodynamics[J]. Nature Communications, 2021, 12(1): 4790. DOI: 10.1038/s41467-021-25158-6.

[13] ELSEMMAN IE, RODRIGUEZ PRADO A, GRIGAITIS P, et al. Whole-cell modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies[J]. Nature Communications, 2022, 13(1): 801. DOI: 10.1038/s41467-022-28467-6.

[14] 李永珠, 陈禹 . 酵母基因组规模模型进展及发展趋势[J]. 合成生物学, 2025, 6(3): 585-602.
LI Y Z, CHEN Y. Advances and prospects in genome-scale models of yeast[J]. Synthetic Biology Journal, 2025, 6(3): 585-602. DOI: 10.12211/2096-8280.2024-084.

[15] DIXIT A, PARNAS O, LI B, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens[J]. Cell, 2016, 167(7): 1853-1866.e17. DOI: 10.1016/j.cell.2016.11.038.

[16] RUBIN AJ, PARKER KR, SATPATHY AT, et al. Coupled Single-Cell CRISPR Screening and Epigenomic Profiling Reveals Causal Gene Regulatory Networks[J]. Cell, 2019, 176 (1-2): 361-376.e17. DOI: 10.1016/j.cell.2018.11.022.

[17] FELDMAN D, SINGH A, SCHMID-BURGK JL, et al. Optical Pooled Screens in Human Cells[J]. Cell, 2019, 179(3): 787-799.e17. DOI: 10.1016/j.cell.2019.09.016.

[18] CAMPOS AI, ZAMPIERI M. Metabolomics-Driven Exploration of the Chemical Drug Space to Predict Combination Antimicrobial Therapies[J]. Molecular Cell, 2019, 74(6): 1291-1303.e6. DOI: 10.1016/j.molcel.2019.04.001.

[19] Biohub. Biohub Launches the Virtual Biology Initiative to Galvanize a Global Effort to Create the Open Data Foundation for AI-Accelerated Biology[EB/OL]. (2026) [2026-09-21]. https://biohub.org/news/virtual-biology-initiative/.

[20] ADDURI A K, GAUTAM D, BEVILACQUA B, et al. Predicting cellular responses to perturbation across diverse contexts with State[J]. Cell, 2026, 189(19): 5914-5931. e20. DOI: 10.1016/j.cell.2026.07.052.

[21] Arc Institute. Arc Institute launches its inaugural virtual cell competition using AI to solve one of biology's biggest challenges[EB/OL]. (2025) [2026-09-21]. https://arcinstitute.org/news/virtual-cell-challenge-2025.

[22] SciLifeLab. Alpha Cell[EB/OL]. (2026) [2026-09-21]. https://www.scilifelab.se/alpha-cell/.

[23] THEODORIS CV, XIAO L, CHOPRA A, et al. Transfer learning enables predictions in network biology[J]. Nature, 2023, 618(7965): 616-624. DOI: 10.1038/s41586-023-06139-9.

[24] CUI H, WANG C, MAAN H, et al. scGPT: toward building a foundation model for single-cell multi-omics using generative AI[J]. Nature Methods, 2024, 21(8): 1470-1480. DOI: 10.1038/s41592-024-02201-0.

[25] HAO M, GONG J, ZENG X, et al. Large-scale foundation model on single-cell transcriptomics[J]. Nature Methods, 2024, 21(8): 1481-1491. DOI: 10.1038/s41592-024-02305-7.

[26] ROSEN Y, BRBIĆ M, ROOHANI Y, et al. Toward universal cell embeddings: integrating single-cell RNA-seq datasets across species with SATURN[J]. Nature Methods, 2024, 21: 1492-1500. DOI: 10.1038/s41592-024-02191-z.

[27] ROSEN Y, ROOHANI Y, AGRAWAL A, et al. Universal cell embedding provides a foundation model for cell biology[J]. Nature, 2026, 656(8126): 183-191. DOI: 10.1038/s41586-026-10689-z.

[28] JI B, HU T, WANG J, et al. CAPTAIN: a multimodal foundation model pretrained on co-assayed single-cell RNA and protein[J]. Nature Communications, 2026, 17(1): 6161. DOI: 10.1038/s41467-026-72882-y.

[29] GANDHI S, JAVADI F, SVENSSON V, et al. Tahoe-x1: Scaling Perturbation-Trained Single-Cell Foundation Models to 3 Billion Parameters[PP/OL]. V1. bioRxiv(2025-10-23) [2026-10-27]. https://www. biorxiv. org/content/10.1101/2025.10.23.683759v1.

[30] KALFON J, SAMARAN J, PEYRÉ G, et al. scPRINT: pre-training on 50 million cells allows robust gene network predictions[J]. Nature Communications, 2025, 16(1): 3607. DOI: 10.1038/s41467-025-58699-1.

[31] JIANG J, CHEN S, TSOU T, et al. D-SPIN constructs regulatory network models from scRNA-seq that reveal organizing principles of perturbation response[J]. Cell, 2026, 189(13): 4159-4184.e35. DOI: 10.1016/j.cell.2026.04.028.

[32] TEJADA-LAPUERTA A, SCHAAR AC, GUTGESELL R, et al. Nicheformer: a foundation model for single-cell and spatial omics[J]. Nature Methods, 2025, 22(12): 2525-2538. DOI: 10.1038/s41592-025-02814-z.

[33] WENCKSTERN J, JAIN E, VON QUERFURTH B, et al. The Virtual Tissues foundation model resolves spatial proteomics across scales[J]. Nature, 2026, 657(8132): 744-754. DOI: 10.1038/s41586-026-10884-y.

[34] KLEIN D, FLECK JS, BOBROVSKIY D, et al. CellFlow enables generative single-cell phenotype modeling with flow

matching[PP/OL]. V1. bioRxiv(2025-04-17) [2026-10-27]. https://www.biorxiv.org/content/10.1101/2025.04.11.648220v1.

[35] YUAN X, LIU X, ZHANG Y S, et al. PerturbDiff: functional diffusion for single-cell perturbation modeling[PP/OL]. V1. arXiv(2026-02-23) [2026-10-27]. https://arxiv. org/abs/2602.19685v1.

[36] DONG M, WANG B, WEI J, et al. Causal identification of single-cell experimental perturbation effects with CINEMA-OT [J]. Nature Methods, 2023, 20(11): 1769-1779. DOI: 10.1038/s41592-023-02040-5.

[37] AN S, CHO JW, CAO K, et al. scCausalVI disentangles single-cell perturbation responses with causality-aware generative model[J]. Cell Systems, 2025, 16(11): 101443. DOI: 10.1016/j.cels.2025.101443.

[38] SUN R, QIAN L, LI Y, et al. An operational perturbation proteomics-based virtual cell model[J]. Nature, 2026, 658 (8137): 1066–1076. DOI: 10.1038/s41586-026-11001-9.

[39] DOU J, LYU W, CHEN F, et al. Inferring stochastic dynamics by biophysical Neural ODE using single-cell transcriptomics [J]. Nature Communications, 2026, 17: 6901. DOI: 10.1038/s41467-026-73257-z.

[40] WEI Z, MA R, WANG Z, et al. VCWorld: A Biological World Model for Virtual Cell Simulation[PP/OL]. V2. arXiv(2026-02-27)[2026-10-27]. https://arxiv.org/abs/2512.00306v2.

[41] WANG Z, YE G, LIU W, et al. AROMA: Augmented Reasoning Over a Multimodal Architecture for Virtual Cell Genetic Perturbation Modeling[PP/OL]. V1. arXiv(2026-04-22)[2026-10-27]. https://arxiv.org/abs/2604.20263v1.

[42] QIAN L, ZHOU Z, ZHOU P, et al. Towards the construction of a virtual yeast[J]. Nature, 2026, 655(8121): 59-70. DOI: 10.1038/s41586-026-10574-9.

[43] QIAN L, DONG Z, GUO T. Grow AI virtual cells: three data pillars and closed-loop learning[J]. Cell Research, 2025, 35(5): 319-321. DOI: 10.1038/s41422-025-01101-y.

[44] LONG Y, HAN X, MENG X, et al. A robust yeast chassis: comprehensive characterization of a fast-growing *Saccharomyces cerevisiae*[J]. mBio, 2024, 15(2): e03196-23. DOI: 10.1128/mbio.03196-23.

[45] SHEMESH P, FISHMAN A. Optimal fermentation conditions for growth and recombinant protein production in *Pichia pastoris*: Strain selection, ploidy level and carbon source[J]. Current Research in Food Science, 2024, 9: 100840. DOI: 10.1016/j.crfs.2024.100840.

[46] LI G, LIANG H, GAO R, et al. Yeast metabolism adaptation for efficient terpenoids synthesis via isopentenol utilization[J]. Nature Communications, 2024, 15(1): 9844. DOI: 10.1038/s41467-024-54298-8.

[47] XIAO Z, ZHAO Y, WANG Y, et al. Sucrose-driven carbon redox rebalancing eliminates the Crabtree effect and boosts energy metabolism in yeast[J]. Nature Communications, 2025, 16(1): 5211. DOI: 10.1038/s41467-025-60578-8.

[48] AJIKUMAR PK, XIAO WH, TYO KEJ, et al. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*[J]. Science, 2010, 330(6000): 70-74. DOI: 10.1126/science.1191652.

[49] CERONI F, ALGAR R, STAN GB, et al. Quantifying cellular capacity identifies gene expression designs with reduced burden[J]. Nature Methods, 2015, 12(5): 415-418. DOI: 10.1038/nmeth.3339.

[50] 田晓军, 张日新. 合成基因回路面临的细胞"经济学窘境"[J]. 合成生物学, 2025, 6(3): 532-546.
TIAN X J, ZHANG R X. "Economics paradox" with cells in synthetic gene circuits[J]. Synthetic Biology Journal, 2025, 6 (3): 532-546. DOI: 10.12211/2096-8280.2024-083.

[51] DUNLOP MJ, DOSSANI ZY, SZMIDT HL, et al. Engineering microbial biofuel tolerance and export using efflux pumps[J]. Molecular Systems Biology, 2011, 7(1): 487. DOI: 10.1038/msb.2011.21.

[52] SIU KH, LEE V, DUEBER JE. Functional targeting of membrane transporters and enzymes to peroxisomes[J]. Nature Chemical Biology, 2025, 21(10): 1544-1553. DOI: 10.1038/s41589-025-01948-7.

[53] FANG L, CHEN Y, HE Q, et al. Mining novel gene targets for improving tolerance to furfural and acetic acid in Yarrowia lipolytica using whole-genome CRISPRi library[J]. Bioresource Technology, 2024, 403: 130764. DOI: 10.1016/j.biortech.2024.130764.

[54] RUGBJERG P, SARUP-LYTZEN K, NAGY M, et al. Synthetic addiction extends the productive life time of engineered *Escherichia coli* populations[J]. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115(10): 2347-2352. DOI: 10.1073/pnas.1718622115.

[55] WU Z, GAO J, GAO N, et al. Engineering chronological lifespan toward a robust yeast cell factory[J]. Proceedings of the National Academy of Sciences of the United States of America, 2025, 122(46): e2515324122. DOI: 10.1073/pnas.2515324122.

[56] HOLMES WJ, DARBY RA, WILKS MD, et al. Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime[J]. Microbial Cell Factories, 2009, 8(1): 35. DOI: 10.1186/1475-2859-8-35.

[57] DIETZSCH C, SPADIUT O, HERWIG C. A fast approach to determine a fed batch feeding profile for recombinant *Pichia pastoris* strains[J]. Microbial Cell Factories, 2011, 10(1): 85. DOI: 10.1186/1475-2859-10-85.

[58] ZHANG F, CAROTHERS JM, KEASLING JD. Design of a

dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids[J]. Nature Biotechnology, 2012, 30(4): 354-359. DOI: 10.1038/nbt.2149.

[59] ZHAO E M, ZHANG Y, MEHL J, et al. Optogenetic regulation of engineered cellular metabolism for microbial chemical production[J]. Nature, 2018, 555(7698): 683-687. DOI: 10.1038/nature26141.

[60] CHOI HS, LEE SY, KIM TY, et al. In Silico Identification of Gene Amplification Targets for Improvement of Lycopene Production[J]. Applied and Environmental Microbiology, 2010, 76(10): 3097-3105. DOI: 10.1128/aem.00115-10.

[61] DOMENZAIN I, LU Y, WANG H, et al. Computational biology predicts metabolic engineering targets for increased production of 103 valuable chemicals in yeast[J]. Proceedings of the National Academy of Sciences of the United States of America, 2025, 122(9): e2417322122. DOI: 10.1073/pnas.2417322122.

[62] 章益蜻, 刘高雯. 合成生物学视角下的基因功能探索与酵母工程菌株文库构建[J]. 合成生物学, 2025, 6(3): 685-700.
ZHANG Y Q, LIU G W. Exploration of gene functions and library construction for engineering strains from a synthetic biology perspective[J]. Synthetic Biology Journal, 2025, 6(3): 685-700. DOI: 10.12211/2096-8280.2024-079.

[63] NEWELL W, HAPETA P, BELL D, et al. Engineering de novo formatotrophy in the non-model yeast Y. lipolytica[J]. Nature Synthesis, 2026, 5(9): 1360-1373. DOI: 10.1038/s44160-026-01076-7.

[64] ENFORS SO, JAHIC M, ROZKOV A, et al. Physiological responses to mixing in large scale bioreactors[J]. Journal of Biotechnology, 2001, 85(2): 175-185. DOI: 10.1016/s0168-1656(00)00365-5.

[65] HARINGA C, DESHMUKH AT, MUDDE RF, et al. Euler-Lagrange analysis towards representative down-scaling of a 22 m3 aerobic *S. cerevisiae* fermentation[J]. Chemical Engineering Science, 2017, 170: 653-669. DOI: 10.1016/j.ces.2017.01.014.

[66] HO P, TÄUBER S, STUTE B, et al. Microfluidic Reproduction of Dynamic Bioreactor Environment Based on Computational Lifelines[J]. Frontiers in Chemical Engineering, 2022, 4: 826485. DOI: 10.3389/fceng.2022.826485.

[67] MINDEN S, ANIOLEK M, NOORMAN H, et al. Mimicked Mixing-Induced Heterogeneities of Industrial Bioreactors Stimulate Long-Lasting Adaption Programs in Ethanol-Producing Yeasts[J]. Genes, 2023, 14(5): 997. DOI: 10.3390/genes14050997.

[68] ANANE E, KNUDSEN IM, WILSON GC. Scale-down cultivation in mammalian cell bioreactors—The effect of bioreactor mixing time on the response of CHO cells to dissolved oxygen gradients[J]. Biochemical Engineering Journal, 2021, 166: 107870. DOI: 10.1016/j.bej.2020.107870.

[69] CRATER JS, LIEVENSE JC. Scale-up of industrial microbial processes[J]. FEMS Microbiology Letters, 2018, 365(13): fny138. DOI: 10.1093/femsle/fny138.

[70] PROMMA I, AUCOIN MG, ABUKHDEIR NM, et al. A coupled metabolic flux/compartmental hydrodynamic model for large-scale aerated bioreactors[J]. Computers & Chemical Engineering, 2024, 189: 108806. DOI: 10.1016/j.compchemeng.2024.108806.

[71] BISGAARD J, ZAHN JA, TAJSOLEIMAN T, et al. Data-based dynamic compartment model: Modeling of *E. coli* fed-batch fermentation in a 600 m3 bubble column[J]. Journal of Industrial Microbiology and Biotechnology, 2022, 49(5): kuac021. DOI: 10.1093/jimb/kuac021.

[72] PROMMA I, AUCOIN MG, ABUKHDEIR NM, et al. Modeling heterogeneity in large - scale bioreactors using the method of moments with a truncated normal distribution[J]. AIChE Journal, 2025, 71(10): e18922. DOI: 10.1002/aic.18922.

[73] THIRUGNANASAMBANDAM MK, PINTO J, MOSKOVKINA E, et al. A Physics-Informed Neural Network (PINN) framework for generic bioreactor modelling[J]. Computers & Chemical Engineering, 2025, 203: 109354. DOI: 10.1016/j.compchemeng.2025.109354.

[74] ZHANG D, SAVAGE TR, CHO BA. Combining model structure identification and hybrid modelling for photo - production process predictive simulation and optimisation[J]. Biotechnology and Bioengineering, 2020, 117(11): 3356-3367. DOI: 10.1002/bit.27512.

[75] FORSTER T, VÁZQUEZ D, CRUZ-BOURNAZOU MN, et al. Modeling of bioprocesses via MINLP-based symbolic regression of S-system formalisms[J]. Computers & Chemical Engineering, 2023, 170: 108108. DOI: 10.1016/j.compchemeng.2022.108108.

[76] LUNA MF, MIONE FM, KASPERSETZ L, et al. Automated regression of bioreactor models using a Bayesian approach for parallel cultivations in robotic platforms[J]. Biochemical Engineering Journal, 2025, 219: 109729. DOI: 10.1016/j.bej.2025.109729.

[77] DÜRR R, DUVIGNEAU S, SEIDEL C, et al. Multi-Rate Data Fusion for State and Parameter Estimation in (Bio-)Chemical Process Engineering[J]. Processes, 2021, 9(11): 1990. DOI: 10.3390/pr9111990.

[78] HELGERS H, SCHMIDT A, STRUBE J. Towards Autonomous Process Control—Digital Twin for CHO Cell-Based Antibody Manufacturing Using a Dynamic Metabolic Model[J]. Processes, 2022, 10(2): 316. DOI: 10.3390/pr10020316.

[79] RIEZZO L, KAY H, FENG Y, et al. Accelerating bioprocess

digital twin development by integrating hybrid modelling with transfer learning[J]. Chemical Engineering Journal, 2025, 511: 162018. DOI: 10.1016/j.cej.2025.162018.

[80] SHUKLA K, NAGPAL S, VYAS A. Opportunities of digital twin in controlling and monitoring microbial processes in food fermentation[J]. Systems Microbiology and Biomanufacturing, 2026, 6(3): 56. DOI: 10.1007/s43393-026-00454-8.

[81] WANG G, ZHANG L, ZHUANG Y, et al. Timescale-dependent metabolic adaptation governs dynamic Crabtree regulation in Saccharomyces cerevisiae[J]. Systems Microbiology and Biomanufacturing, 2026, 6(5): 146. DOI: 10.1007/s43393-026-00549-2.

[82] 唐婷, 付立豪, 郭二鹏, 等. 自动化合成生物技术与工程化设施平台[J]. 科学通报, 2021, 66(3): 300-309.
TANG T, FU L H, GUO E P, et al. Automation in synthetic biology using biological foundries[J]. Chinese Science Bulletin, 2021, 66(3): 300-309. DOI: 10.1360/TB-2020-0498.

[83] 卢挥, 张芳丽, 黄磊. 合成生物学自动化装置iBioFoundry的构建与应用[J]. 合成生物学, 2023, 4(5): 877-891.
LU H, ZHANG F L, HUANG L. Establishment of iBioFoundry for synthetic biology applications[J]. Synthetic Biology Journal, 2023, 4(5): 877-891. DOI: 10.12211/2096-8280.2023-027.

[84] SI T, CHAO R, MIN Y, et al. Automated multiplex genome-scale engineering in yeast[J]. Nature Communications, 2017, 8 (1): 15187. DOI: 10.1038/ncomms15187.

[85] RADIVOJEVIĆ T, COSTELLO Z, WORKMAN K, et al. A machine learning Automated Recommendation Tool for synthetic biology[J]. Nature Communications, 2020, 11(1): 4879. DOI: 10.1038/s41467-020-18008-4.

[86] COUTANT A, ROPER K, TREJO-BANOS D, et al. Closed-loop cycles of experiment design, execution, and learning accelerate systems biology model development in yeast[J]. Proceedings of the National Academy of Sciences of the United States of America, 2019, 116(36): 18142-18147. DOI: 10.1073/pnas.1900548116.

[87] CHENG W, MA M, SHEN S, et al. Towards Human-Led, Agent-Driven Autonomous Laboratories for the Life Sciences [PP/OL]. V1. Preprints. org(2026-08-05) [2026-10-27]. https://doi.org/10.20944/preprints202608.0273.v1.

[88] CHEN J, VOLK M, SHAFAEI S, et al. An AI-powered cloud biofoundry for autonomous biological research[PP/OL]. V1. bioRxiv(2026-10-01) [2026-10-27]. https://www. biorxiv. org/content/10.64898/2026.09.30.750244v1.

[89] SMITH A A, WONG E L, DONOVAN R C, et al. Using a GPT-5-driven autonomous lab to optimize the cost and titer of cell-free protein synthesis[PP/OL]. V1. bioRxiv(2026-02-05) [2026-10-27]. https://www. biorxiv. org/content/10.64898/2026.02.05.703998v1.

[90] HIRSCH E, PATAKI H, DOMJÁN J, et al. Inline noninvasive Raman monitoring and feedback control of glucose concentration during ethanol fermentation[J]. Biotechnology Progress, 2019, 35(5): e2848. DOI: 10.1002/btpr.2848.

[91] SANTOS LO, DEWASME L, COUTINHO D, et al. Nonlinear model predictive control of fed-batch cultures of micro-organisms exhibiting overflow metabolism: assessment and robustness[J]. Computers & Chemical Engineering, 2012, 39: 143-151. DOI: 10.1016/j.compchemeng.2011.12.010.

[92] PABLY P, HUUSOM J K, KAGER J. Model Predictive Control of dissolved oxygen for pulsed feeding of microbial cultivations[J]. Biochemical Engineering Journal, 2026, 234: 110261. DOI: 10.1016/j.bej.2026.110261.

[93] XING Z, KENTY BM, LI ZJ, et al. Scale - up analysis for a CHO cell culture process in large - scale bioreactors[J]. Biotechnology and Bioengineering, 2009, 103(4): 733-746. DOI: 10.1002/bit.22287.

[94] 林一瀚, 傅雄飞, 刘陈立, 等. 定量合成生物学: 合成生物系统的理性设计之路[J]. 合成生物学, 2025, 6(3): 493-496.
LIN Y H, FU X F, LIU C L, et al. Quantitative synthetic biology: toward rational design of synthetic biological systems [J]. Synthetic Biology Journal, 2025, 6(3): 493-496. DOI: 10.12211/2096-8280.2025-063.

[95] LIPINSKI CA, LOMBARDO F, DOMINY BW, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings[J]. Advanced Drug Delivery Reviews, 1997, 23(1-3): 3-25. DOI: 10.1016/s0169-409x(96)00423-1.

[96] DEBENEDICTIS E A, CHORY E J, GRETTON D W, et al. Systematic molecular evolution enables robust biomolecule discovery[J]. Nature Methods, 2022, 19(1): 55-64. DOI: 10.1038/s41592-021-01348-4.

[97] SUTHERS PF, DINH HV, FATMA Z, et al. Genome-scale metabolic reconstruction of the non-model yeast *Issatchenkia orientalis* SD108 and its application to organic acids production [J]. Metabolic Engineering Communications, 2020, 11: e00148. DOI: 10.1016/j.mec.2020.e00148.

[98] LI F, YUAN L, LU H, et al. Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction[J]. Nature Catalysis, 2022, 5(8): 662-672. DOI: 10.1038/s41929-022-00798-z.

[99] COSTANZO M, HOU J, MESSIER V, et al. Environmental robustness of the global yeast genetic interaction network[J]. Science, 2021, 372(6542): eabf8424. DOI: 10.1126/science.abf8424.

[100] TAS H, GROZINGER L, STOOF R, et al. Contextual dependencies expand the re-usability of genetic inverters[J]. Nature Communications, 2021, 12(1): 355. DOI: 10.1038/s41467-020-20656-5.

[101] WAHJUDI S M W, PETRZIK T, OUDENNE F, et al. Unraveling the potential and constraints associated with corn steep liquor as a nutrient source for industrial fermentations[J]. Biotechnology Progress, 2023, 39(6): e3386. DOI: 10.1002/btpr.3386.

[102] ZHANG H, LI Z, PEREIRA B, et al. Engineering *E. coli* – *E. coli* cocultures for production of muconic acid from glycerol [J]. Microbial Cell Factories, 2015, 14(1): 134. DOI: 10.1186/s12934-015-0319-0.

[103] GAO M, ZHAO Y, YAO Z, et al. Xylose and shikimate transporters facilitates microbial consortium as a chassis for benzylisoquinoline alkaloid production[J]. Nature Communications, 2023, 14(1): 7797. DOI: 10.1038/s41467-023-43049-w.

[104] ZHANG L, XIAO S, WANG D, et al. Metabolic engineering of *Saccharomyces cerevisiae* for co-production of ergothioneine, salidroside and gadusol[J]. Synthetic and Systems Biotechnology, 2026, 14: 471-479. DOI: 10.1016/j.synbio.2026.04.010.

[105] SHAW AJ, LAM FH, HAMILTON M, et al. Metabolic engineering of microbial competitive advantage for industrial fermentation processes[J]. Science, 2016, 353(6299): 583-586. DOI: 10.1126/science.aaf6159.

[106] NYERGES A, VINKE S, FLYNN R, et al. A swapped genetic code prevents viral infections and gene transfer[J]. Nature, 2023, 615(7953): 720-727. DOI: 10.1038/s41586-023-05824-z.

[107] HAMILTON SR, BOBROWICZ P, BOBROWICZ B, et al. Production of Complex Human Glycoproteins in Yeast[J]. Science, 2003, 301(5637): 1244-1246. DOI: 10.1126/science.1088166.

[108] MAMAT U, WILKE K, BRAMHILL D, et al. Detoxifying *Escherichia coli* for endotoxin-free production of recombinant proteins[J]. Microbial Cell Factories, 2015, 14(1): 57. DOI: 10.1186/s12934-015-0241-5.

[109] MANDELL DJ, LAJOIE MJ, MEE MT, et al. Biocontainment of genetically modified organisms by synthetic protein design[J]. Nature, 2015, 518(7537): 55-60. DOI: 10.1038/nature14121.

[110] XING EP, SONG L. A world model of the virtual cell[J]. Cell, 2026, 189(19): 5831-5844. DOI: 10.1016/j.cell.2026.08.042.

[111] NOORI A, FISHMAN N, FANG A, et al. World models for biomedicine[J]. Cell, 2026, 189(19): 5845-5869. DOI: 10.1016/j.cell.2026.08.032.

[112] OLSSON L, RUGBJERG P, TORELLO PIANALE L, et al. Robustness: linking strain design to viable bioprocesses[J]. Trends in Biotechnology, 2022, 40(8): 918-931. DOI: 10.1016/j.tibtech.2022.01.004.

[113] DE BRABANDER P, UITTERHAEGEN E, DELMULLE T, et al. Challenges and progress towards industrial recombinant protein production in yeasts: A review[J]. Biotechnology Advances, 2023, 64: 108121. DOI: 10.1016/j.biotechadv.2023.108121.

[114] 云慧敏, 陈必强, 谭天伟. 生物制造产业的系统可行性框架、边界机制与实现路径[J]. 中国工程科学, 2026, 28(4): 28-42.
YUN H M, CHEN B Q, TAN T W. System Feasibility of Biomanufacturing Industry: An Analytical Framework, Boundary Mechanisms, and Enhancement Pathways[J]. Strategic Study of Chinese Academy of Engineering, 2026, 28 (4): 28-42. DOI: 10.15302/J-SSCAE-2026.06.016.



通讯作者：郭天南（1981—），男，长聘教授、博士生导师，博士，研究方向为AI蛋白质组学与虚拟细胞构建，担任中国生物化学与分子生物学会蛋白质组学专业委员会主任委员、国际蛋白质组学会理事。团队提出了人工智能虚拟细胞原创构建框架，并建立首个基于扰动蛋白质组的虚拟细胞基础模型ProteinTalks。
E-mail：guotiannan@westlake. edu. cn



第一作者：戴振武（1998—），男，博士研究生。研究方向为AI蛋白质组学与虚拟细胞构建。
E-mail：daizhenwu@westlake. edu. cn

作者贡献声明：

戴振武：文献调研与整理、论文框架设计、论文初稿撰写、图表设计与绘制；钱鎏佳、孙瑞、周梓卓：概念构思、论文审阅与修改；蒋恒：文献调研与整理、论文审阅与修改；郭天南：概念构思、研究指导、论文审阅与最终定稿。

所有作者均已阅读并同意论文的最终版本。

利益冲突：

郭天南为西湖欧米（杭州）生物科技有限公司创始人。其余作者声明不存在利益冲突。