

## CBS LNBR 2021

### Modelos metabólicos em escala genômica no *design* de microrganismos

#### 1. O que são modelos metabólicos em escala genômica?

**Resposta:** Os modelos metabólicos em escala genômica (GEMs) são uma representação matemática das relações gene-proteína-reação caracterizadas pelo genoma de um dado microrganismo ou comunidade microbiana. Estes modelos consistem em centenas ou milhares de reações e são baseados na estequiometria de todas as possíveis reações metabólicas de uma célula, considerando a reversibilidade destas reações. Através dos GEMs e de ferramentas matemáticas é possível simular o comportamento de microrganismos.

#### 2. Quais as possíveis aplicações dos GEMs?

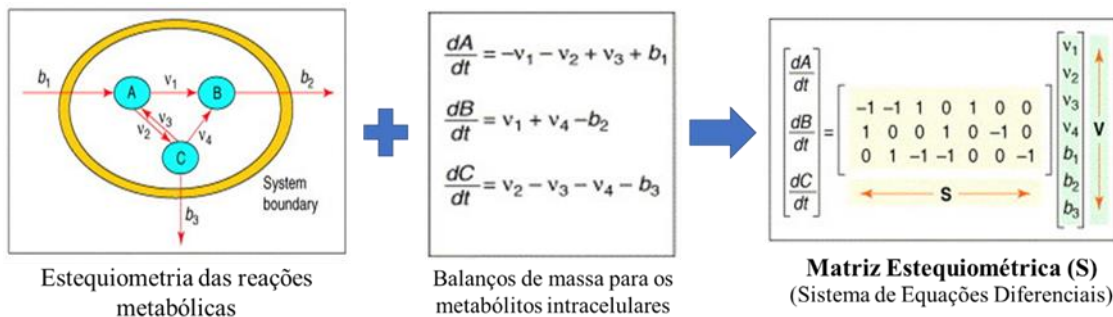
**Resposta:** Desde a divulgação da primeira sequência completa do genoma da bactéria *Haemophilus influenzae* na década 90, as pesquisas sobre modelos metabólicos em escala genômica (GEMs) abriram uma nova era em engenharia metabólica. Hoje em dia, é crescente o uso de GEMs para analisar e simular o metabolismo celular de diferentes organismos, incluindo microrganismos, plantas e mamíferos. Os GEMs vêm sendo empregados para: (i) etapa de design e otimização de microrganismos para produção de bioquímicos / biocombustíveis / bioplásticos no contexto do ciclo de *Design-Build-Test-Learn* (DBTL) da Biologia Sintética; (ii) prever a viabilidade de um organismo sob uma determinada condição, possibilitando analisar e sugerir drogas capazes de inibir efetivamente um dado patógeno; (iii) elucidar disfunções metabólicas em células para compreender doenças humanas; (iv) prever os produtos metabólicos de uma comunidade microbiana sem haver a necessidade de isolar os microrganismos.

#### 3. Como os GEMs são reconstruídos? E quais softwares estão sendo utilizados?

**Resposta:** O processo de reconstrução dos GEMs compreende quatro fases fundamentais: sequenciamento e anotação do genoma do organismo de interesse, desenvolvimento da rede metabólica, conversão da rede num modelo estequiométrico e, finalmente, a validação do modelo metabólico. Mais recentemente, ferramentas automatizadas estão disponíveis e possibilitam a reconstrução de “rascunhos” de GEMs (*draft*). Dentre estas ferramentas destaca-se: ModelSEED, RAVEN Toolbox, CarveMe, GapSeq, ambas livres, de código aberto e fácil de usar. Após a reconstrução do *draft*, uma etapa de refinamento manual é necessária para melhorar a acurácia deste *draft*. Para isso, diferentes bancos de dados podem ser empregados, como: KEGG, MetaCyc, KBase, Brenda, entre outros.

**4. Como transformar GEMs em uma ferramenta eficiente para o design e melhoria de microrganismos? (No curso relacionamos os GEMs ao “Google Maps” da Biologia Sintética)**

**Resposta:** Primeiramente é necessário adicionar uma equação referente a produção de biomassa, incluindo macromoléculas e precursores necessários para crescimento e divisão celular. Em seguida, é preciso converter as informações presentes no modelo estequiométrico curado em um modelo matemático. Para isso, balanço de massa envolvendo todos os metabólitos extracelulares e intracelulares deve ser aplicado, levando em consideração os diferentes compartimentos da célula (espaço extracelular, periplasma, citoplasma, mitocôndria, vacúolo...) e os transportes dos metabólitos entre os diferentes compartimentos, gerando assim, uma matriz estequiométrica  $S$  (Figura 1) que representa um Sistema de Equações Diferenciais complexo de ser resolvido. A fim de solucionar esse sistema, métodos de otimização baseados em restrições (*constraint-based modeling* – CBM) são comumente empregados para simular o comportamento de organismos. Dentre os métodos mais conhecidos, está a Análise de Balanço de Fluxos (*Flux Balance Analysis* – FBA), uma abordagem que assume um estado (pseudo) estacionário (ou seja, não há variação da concentração dos metabólitos ao longo do tempo) e requer uma função objetivo de modo a obter um problema de programação linear facilmente resolvível. A função objetivo mais empregada é de maximização da biomassa, uma vez que na natureza os microrganismos se adaptam e evoluem para maximizar o crescimento. Durante o processo de otimização de um microrganismo mais de uma função objetivo pode ser adotada, por exemplo: maximizar produção de biomassa e minimizar a produção de um sub-produto indesejado.



**Figura 1** - Representação matemática de um modelo metabólico em escala genômica, em que  $dA/dt$ ,  $dB/dt$  e  $dC/dt$  representam a variação dos componentes A, B e C ao longo do tempo;  $S$  a matriz estequiométrica na qual as linhas representam os metabólitos e as colunas representam as reações e  $v$  o vetor que representa os fluxos de cada reação.

**Fonte:** Adaptado de Kauffman, Prakash e Edwards, 2003

**5. Quais algoritmos são mais usados e eficazes na identificação de alvos para melhorar a produção do produto de interesse?**

**Resposta:** Vários algoritmos sofisticados estão sendo desenvolvidos para sugerir diferentes tipos de modificações genéticas. Podemos citar: (i) OptKnock, uma abordagem de programação em dois níveis amplamente empregada para identificar estratégias de deleção de genes, Varredura de fluxo baseada em fluxo objetivo imposto; (ii) Fluxo baseada em Fluxo Objetivo Imposto (*Flux Scanning based on Enforced Objective Flux - FSEOF*) que sugere super (amplificação) e sub-expressão (atenuação) de genes com o objetivo de intensificar a produção do produto de interesse; (iii) OptStrain que pesquisa em bancos de dados de reações a fim de encontrar novas possíveis reações a serem incluídas à rede metabólica para melhorar o rendimento teórico máximo do produto de interesse, entre outros (Maia, Rocha e Rocha, 2016).

#### Literatura sugerida sobre GEMs:

- Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome biology*, 20(1), 1-18.
- Kauffman, K. J., Prakash, P., & Edwards, J. S. (2003). Advances in flux balance analysis. *Current opinion in biotechnology*, 14(5), 491-496.
- Maia, P., Rocha, M., Rocha, I. (2016). In silico constraint-based strain optimization methods: the quest for optimal cell factories. *Microbiology and Molecular Biology Reviews*, 80(1), 45-67.
- Orth, J. D., Thiele, I., Palsson, B. Ø. (2010). What is flux balance analysis? *Nature biotechnology*, 28(3), 245-248.

## **Ferramentas moleculares para edição de genomas em microrganismos**

**1.** *Cite exemplos de métodos que podem ser utilizados para ligação dos blocos de DNA e montagem de vetores. Explique o princípio de cada método.*

**Resposta:** Os principais métodos utilizados para ligação de blocos de DNA (biobricks) ou montagem de vetores são:

a) Golden Gate: Esse método é bastante utilizado na ligação de blocos de DNA, também chamados de *biobricks*, que podem ser adquiridos através de um banco de *biobricks* organizado pela iGEM. Esse banco fornece diferentes blocos de DNA como promotores, sítio de ligação ao ribossomo (RBS) e terminadores que podem ser ligados a uma sequência codificadora (CDS = coding sequence) específica para expressão na célula. Esse método de ligação utiliza as enzimas endonucleases do tipo IIS que clivam o DNA formando extremidades coesivas de 4 nucleotídeos. As diferentes extremidades coesivas formadas podem parear com blocos de DNA que possuem uma sequência

complementar e em seguida uma DNA ligase promove a ligação das partes pareadas resultando na montagem completa da construção.

b) Os métodos denominados de *Gibson assembly*, *In-fusion cloning* e SLIC (*sequence and ligation independent cloning*) possuem o mesmo princípio. Inicialmente, fragmentos de DNA são amplificados por PCR (Polymerase Chain Reaction), sendo que esses blocos de DNA apresentam uma sequência de pareamento (ou homologia) entre eles que pode variar entre 16 a 20 nucleotídeos dependendo do método utilizado. Em seguida, os diferentes blocos de DNA produzidos são tratados com uma enzima que possui atividade de exonuclease para formar extremidades de DNA com fita simples que podem parear. Após o pareamento dos diferentes blocos de DNA, ocorre o reparo das regiões de pareamento (ou homologia) através da enzima DNA polimerase e a ligação dos blocos pela DNA ligase.

c) *Outro método que também pode ser utilizado na montagem de cassetes de expressão e vetores é o TAR (Transformation-associated recombination)*, o qual utiliza enzimas envolvidas no processo de recombinação em células de leveduras. Por tanto, esse método de montagem de blocos de DNA, diferentemente dos outros métodos mencionados acima, ocorre *in vivo*. Para que o processo de recombinação ocorra, os diferentes blocos de DNA amplificados por PCR devem apresentar uma homologia de no mínimo 40 pares de bases entre eles. Células de leveduras são transformadas com os diferentes fragmentos de DNA produzidos e, através do processo de recombinação nas células, os blocos de DNA são ligados formando a construção desejada.

**2. Qual a diferença entre as classes I e II dos sistemas CRISPR/Cas? Cite as características comuns dos membros da Classe II e tipo 2.**

**Resposta:** O efetor da classe I é formado por um complexo de múltiplas proteínas que interagem entre si e com o crRNA (CRISPR RNA), para ser direcionado para a sequência alvo no DNA. Por outro lado, o efetor da classe II é formado somente por uma única proteína (Cas9, Cas12 ou Cas13, por exemplo) que interage com crRNA para ser guiada até a sequência alvo no DNA. Por isso a classe II é mais fácil de ser manipulada para edição de genomas.

O sistema CRISPR da classe II e tipo 2 mais conhecido e estudado é o sistema CRISPR/Cas9. Neste sistema, a proteína Cas9 é o efetor que interage com o crRNA processado na célula que contém a sequência guia para parear com a sequência alvo no DNA. A interação da Cas9 com o crRNA depende da formação de um complexo entre crRNA e outro RNA denominado tracrRNA (RNA transativador do crRNA). Portanto, o sistema CRISPR/Cas9 é formado pela enzima Cas9 e o complexo crRNA-tracrRNA. O crRNA contém 20 nucleotídeos que pareiam com o sítio alvo no DNA que será clivado pela Cas9. O sítio alvo no DNA possui uma sequência PAM (*Protospacer adjacent motif*) que é reconhecida pela Cas9 (5'-NGG-3') e posiciona os domínios de nucleases HNH e RuvC para clivarem a fita pareada com o sgRNA e a fita não pareada, respectivamente. Portanto, Cas9 cliva as duas fitas de DNA, a pareada com o sgRNA e a não pareada, geralmente entre o terceiro e quarto nucleotídeo a partir da sequência

PAM. O domínio de Cas9 que reconhece a sequência PAM (5'-NGG-3') é denominado de PI (*PAM Interacting*).

**3. Qual a diferença entre crRNA, TracrRNA e sgRNA dentro do sistema CRISPR/Cas9? Quais as funções que as proteínas Cas9 modificadas Nickase e Dead-Cas9 podem desenvolver?**

**Resposta:** O crRNA contém a sequência guia que pareia com a sequência alvo do DNA que será clivado pela Cas9, sendo produzido através do processamento do pré-crRNA pela RNase III na célula bacteriana. O crRNA processado liga-se e forma um complexo com o tracrRNA. O tracrRNA contém uma região com estruturas secundárias tipo grampos (denominada de *scaffold*) que são reconhecidas pela Cas9 e auxilia no posicionamento correto da sequência guia (20 nucleotídeos) do crRNA dentro da enzima.

Para facilitar a edição de genomas, a biologia sintética criou o sgRNA (*single guide RNA*) que corresponde a um único RNA com características conservadas do complexo crRNA-tracrRNA. O sgRNA carrega a sequência guia de 20 nucleotídeos que pareia com o DNA alvo e direciona a Cas9 para clivar a fitas após o reconhecimento da sequência PAM. A sequência guia de 20 nucleotídeos pode ser mudada para parear com uma sequência específica na região do DNA que deseja fazer uma edição.

Versões mutantes da enzima Cas9 foram geradas para diversificar o processo de edição de genomas e controle da expressão gênica em diferentes organismos. A enzima Cas9 cliva as duas fitas de DNA, a pareada com o sgRNA e a não pareada através dos seus domínios de nuclease HNH e RuvC, respectivamente. A mutação no sítio catalítico em um desses domínios HNH (H840A) ou RuvC (D10A) produz a enzima Cas9 modificada que pode clivar somente uma das fitas de DNA, sendo denominada de *Nickase* (nCas9). Quando a enzima Cas9 carrega ambas mutações nos dois domínios de nuclease HNH (H840A) e RuvC (D10A), tem-se uma versão inativa de Cas9, ou seja, a enzima não consegue clivar nenhuma das fitas do DNA alvo. A enzima inativa Cas9 que carrega as duas mutações HNH (H840A) e RuvC (D10A) é denominada de *Dead Cas9* (dCas9). A dCas9 não cliva as fitas de DNA, mas consegue interagir com sequência alvo do DNA guiada pelo sgRNA. Por isso, dCas9 tem sido usada em função com outros domínios para edição de genomas através da desaminação de bases nucleotídeos ou para ativar (CRISPRa) /reprimir (CRISPRi) a expressão de genes na célula.

**4. Durante a edição gênica por CRISPR/Cas9, as duas fitas de DNA são clivadas e ativam o sistema de reparo da célula, que pode ser NHEJ ou HR. Quando é ativado o sistema de reparo HR?**

**Resposta:** O reparo por recombinação homóloga (*HR = homologous recombination*) ocorre quando um fragmento de DNA denominado de doador está presente na célula durante a clivagem das fitas de DNA por Cas9. O DNA doador precisa conter uma sequência homóloga para parear próximo da região clivada. Na ausência do DNA doador a célula realiza o reparo da região do DNA clivado pelo sistema NHEJ (*Non-*

*homologous end joining*) ou união das extremidades não-homólogas. O reparo tipo NHEJ é impreciso e pode gerar inserção ou deleção de nucleotídeos na região clivada do DNA. O conjunto de modificações ocasionados pelo sistema NHEJ é denominado de Indels (Inserções e deleções).

**5. Cite alguns procedimentos que podem ajudar a evitar off-targets no processo de edição gênica por CRISPR/Cas9.**

**Resposta:** Entre os procedimentos que podem auxiliar a redução de *off-targets* por Cas9 nas células têm-se:

- Escolha de promotores adequados para controlar a expressão de Cas9 nas células. Um promotor constitutivo que produzirá uma grande concentração de Cas9 na célula pode resultar na ligação de Cas9 em outros sítios de DNA inespecíficos e culminar na clivagem dos mesmos.
- A sequência guia no sgRNA deve conter 40% a 60% de GC e evitar o não pareamento de bases próximo a sequência PAM no DNA alvo.
- Usar variantes mutantes da enzima Cas9 com alta precisão para edição de genomas como, por exemplo, a Cas9-HiFi.
- Utilizar programas como o CRISPOR para desenhar o gRNA com um sequência guia mais eficiente e evitar o pareamento com sítios inespecíficos no genoma.

**6. Como INDELS (Inserções e deleções) ocasionados pelo sistema CRISPR/Cas9 podem ser detectados?**

**Resposta:** Os Indels gerados durante o processo de edição do genoma podem ser detectados por sequenciamento do genoma (procedimento que exige um grande investimento de recursos) ou PCR de regiões do genoma que são previstas como alvos através do estudo *in silico* para identificar *on-targets* e *off-targets*. A análise *in silico* busca por regiões do genoma a ser editado onde a sequência guia do sgRNA pode parear. Além disso, enzimas específicas que detectam o mal-pareamento de bases em regiões do DNA alvo, como a T7 endonuclease, pode ajudar a confirmar a presença de Indels após a edição do genoma.

Literatura sugerida sobre ferramentas de edição de genomas:

- Fels U, Gevaert K, Van Damme P. Bacterial Genetic Engineering by Means of Recombineering for Reverse Genetics. *Front Microbiol.* 2020 Sep 11;11:548410. doi: 10.3389/fmicb.2020.548410.
- Noskov VN, Kouprina N, Leem SH, Ouspenski I, Barrett JC, Larionov V. A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic

region in yeast. BMC Genomics. 2003 Apr 29;4(1):16. doi: 10.1186/1471-2164-4-16.

- Marraffini LA. CRISPR-Cas immunity in prokaryotes. Nature. 2015 Oct 1;526(7571):55-61. doi: 10.1038/nature15386.
- Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, Charpentier E, Cheng D, Haft DH, Horvath P, Moineau S, Mojica FJM, Scott D, Shah SA, Siksnyš V, Terns MP, Venclovas Č, White MF, Yakunin AF, Yan W, Zhang F, Garrett RA, Backofen R, van der Oost J, Barrangou R, Koonin EV. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol. 2020 Feb;18(2):67-83. doi: 10.1038/s41579-019-0299-x.
- Jiang F, Doudna JA. CRISPR-Cas9 Structures and Mechanisms. Annu Rev Biophys. 2017 May 22;46:505-529. doi: 10.1146/annurev-biophys-062215-010822.
- Nidhi S, Anand U, Oleksak P, Tripathi P, Lal JA, Thomas G, Kuca K, Tripathi V. Novel CRISPR-Cas Systems: An Updated Review of the Current Achievements, Applications, and Future Research Perspectives. Int J Mol Sci. 2021 Mar 24;22(7):3327. doi: 10.3390/ijms22073327.
- Naeem M, Majeed S, Hoque MZ, Ahmad I. Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. Cells. 2020 Jul 2;9(7):1608. doi: 10.3390/cells9071608.
- Bennett EP, Petersen BL, Johansen IE, Niu Y, Yang Z, Chamberlain CA, Met Ö, Wandall HH, Frödin M. INDEL detection, the 'Achilles heel' of precise genome

editing: a survey of methods for accurate profiling of gene editing induced indels. *Nucleic Acids Res.* 2020 Dec 2;48(21):11958-11981. doi: 10.1093/nar/gkaa975.

## **Engenharia metabólica – ferramentas e abordagens na construção de microrganismos**

*1. Quais são os principais desafios na obtenção de uma microbial factory pela introdução de uma via metabólica heteróloga?*

**Resposta:** Os principais desafios de um sistema microbiano de produção heteróloga são (i) a falta de precursores biossintéticos ou de uma função enzimática chave no metabolismo endógeno do chassis microbiano para a manutenção da via heteróloga, (ii) incompatibilidade de sistemas regulatórios, fazendo-se necessária a introdução de reguladores ou substituição dos promotores nativos, e (iii) expressão não coordenada dos genes biossintéticos que pode resultar em sobrecarga metabólica e/ou acúmulo de subprodutos tóxicos, limitando o rendimento e produtividade do bioproduto de interesse. Na expressão heteróloga de proteínas, a plataforma microbiana deve ser compatível à proteína a ser produzida, uma vez que proteínas eucarióticas, por exemplo, passam por diversas modificações pós-traducionais, como glicosilação e formação de pontes de disulfeto, as quais são realizadas a depender do chassis microbiano escolhido.

*2. De que forma abordagens multi-ômicas vem sendo incorporadas na engenharia metabólica de microrganismos?*

**Resposta:** Análises ômicas, incluindo genômica, transcriptômica, proteômica, metabolômica e fluxômica, fornecem informações sistêmicas, ou seja, globais, das características celulares e metabólicas do microrganismo engenheirado sob diferentes condições genotípicas (diferentes plataformas microbianas) e ambientais (por exemplo, diferentes meios de cultura/fontes de carbono e modos de cultivo). Utilizando-se dessa ampla e rica gama de informações pode-se identificar os gargalos e potenciais modificações que possam otimizar a microbial factory. A fluxômica consiste na determinação dos fluxos metabólicos em um sistema biológico. Para se aferir o fluxo de carbono, por exemplo, é feita a quantificação da incorporação do isótopo <sup>13</sup>C em metabólitos seguida da análise por espectrometria de massas e ressonância magnética nuclear. Por isso, recentemente tem-se aumentado o interesse por análises de fluxômica a fim de se avaliar o metabolismo do microrganismo e direcionar o fluxo para a produção do composto de interesse. Além disso, dados de análises multi-ômicas podem ser integrados a modelos metabólicos em escala genômica para o desenvolvimento de modelos condição-específica mais acurados. Também, dados de genômica, transcriptômica e proteômica, depositados em banco de dados, podem ser



utilizados para a predição *in silico* de vias metabólicas através de ferramentas computacionais.

**3. Quais são as estratégias de engenharia metabólica utilizadas para acoplar a produção do composto de interesse ao crescimento microbiano?**

**Resposta:** A síntese de um bioproduto pela introdução de uma via metabólica heteróloga pode resultar em uma sobrecarga metabólica. Como consequência, uma cepa que perde a capacidade de produção é selecionada e torna-se predominante no cultivo perante as cepas que mantiveram a capacidade de síntese do bioproduto. Dessa forma, estratégias para acoplar a produção ao crescimento microbiano faz com que o microrganismo cresça somente se houver fluxo através da via heteróloga, resultando, portanto, em uma maior e mais estável produção do composto de interesse. Através de ferramentas computacionais de otimização de *microbial factories* (e.g., OptKnock e RobustKnock), deleções estratégicas de genes são encontradas para que o fluxo passe necessariamente pela via heteróloga para o balanço de energia e cofatores, ou, para que o fluxo passe necessariamente pela via heteróloga para que precursores fundamentais para o crescimento celular sejam sintetizados. Como material de estudo, Klamt & Mahadevan publicaram um artigo bastante completo sobre este tema.

**4. Quais são as principais contribuições do machine learning para a engenharia metabólica de microrganismos?**

**Resposta:** O objetivo do machine learning (ML) é aprender uma função capaz de, a partir de um conjunto de dados pareados de entrada e saída, chamados de *training data*, mapear um determinado dado de entrada a um valor de saída desejado. O ML opera com base na suposição de que existe uma correlação entre os dados de entrada e os dados de saída. Quanto mais *training data* for fornecido, mais precisa será a função obtida do aprendizado. Dessa forma, o ML vem sendo utilizado e demonstrado em diversos trabalhos na literatura em todos os estágios do design de um microrganismo com o objetivo de ajudar a encontrar soluções não óbvias de engenharia e reduzindo as etapas necessárias de otimização. Seguem alguns exemplos de aplicação do ML na engenharia metabólica. No sistema CRISPR/Cas para edição de genomas, diferentes designs de *guide RNA* (gRNA) apresentam eficiências distintas de edição, o que é crucial para que a edição, deleção ou introdução de uma via metabólica, seja realizada com sucesso. Utilizando-se de bibliotecas de gRNA e dados de eficiência e especificidade de cada design, modelos de ML vem desempenhando o importante papel de descobrir regras e para aumentar a previsibilidade de novos designs de gRNAs. Existem também estudos que se utilizam das sequências de promotores *de ribosome binding sites* (RBS) a fim de prever a força funcional desses elementos. Dessa forma, coleções de promotores, RBS, e outras partes relevantes para o design ótimo de uma via metabólica são obtidas com maior confiança e previsibilidade. Na engenharia de bioprocessos, o ML colabora principalmente na

geração de *soft sensors*. Estes sensores são obtidos da correlação de variáveis facilmente medidas online (e.g., oxigênio dissolvido) e offline (e.g., densidade ótica do microrganismo no meio de cultivo), para a obtenção de uma medida de extrema relevância para o bioprocessamento, no caso o *soft sensor*, mas que não é trivial de ser obtida (e.g., bioplástico como o produto de interesse). Para maior aprofundamento, a excelente revisão de Volk *et al.* aborda o tema dessa pergunta de forma compreensiva e abrangente.

**5. Existem empresas nacionais que atuam no desenvolvimento de microbial factories por estratégias de engenharia metabólica? Qual é o cenário brasileiro atual nesta área?**

**Resposta:** Sim, existem empresas brasileiras que atuam no design de microrganismos utilizando estratégias de engenharia metabólica, mas o número é ainda bastante escasso. Como discutido na palestra, um alto aporte financeiro e muitos anos de desenvolvimento são geralmente necessários para que se alcance os níveis alvo de título, produtividade e rendimento do bioquímico de interesse. A engenharia metabólica de microrganismos está em plena expansão e o avanço de diversas tecnologias vem contribuindo para que essa ciência conquiste mais espaço na indústria. A OCDE prevê que em 2030 2,7% do PIB dos países membros será proveniente da manufatura de químicos por rotas biológicas. Entre os exemplos de empresas de biotecnologia brasileiras de maior porte, destacam-se a Braskem, do ramo petroquímico, e a GranBio, empresa de biotecnologia industrial produtora de etanol celulósico, ambas fazendo uso da engenharia metabólica de microrganismos para produção de seus bioquímicos de interesse. Outro exemplo é a Integra Bioprocessos, a qual realiza a engenharia de microrganismos para converter resíduos agroindustriais em produtos de alto valor agregado. Um dos seus produtos é o ácido láctico produzido por *Pichia pastoris* modificada geneticamente utilizando glicerina, um resíduo da indústria do biodiesel, como substrato. O LNBR, instituição sede do curso de biologia sintética, possui diversos projetos de desenvolvimento de microrganismos, em parceria com empresas ou por iniciativa institucional, para a produção de químicos diversos usando leveduras e bactérias como chassis. O desenvolvimento de um fungo filamentoso para a produção de um coquetel enzimático de desconstrução da biomassa foi recentemente publicado por pesquisadores da instituição.

Literatura sugerida sobre engenharia metabólica:

- Biz, A., Proulx, S., Xu, Z., Siddartha, K., Indrayanti, A.M. and Mahadevan, R., 2019. Systems biology based metabolic engineering for non-natural chemicals. *Biotechnology Advances*.

- Calero, P. and Nikel, P.I., 2019. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microbial Biotechnology*.
- Chae, T.U., Choi, S.Y., Kim, J.W., Ko, Y.S. and Lee, S.Y., 2017. Recent advances in systems metabolic engineering tools and strategies. *Current Opinion in Biotechnology*.
- Chubukov, V., Mukhopadhyay, A., Petzold, C.J., Keasling, J.D. and Martín, H.G., 2016. Synthetic and systems biology for microbial production of commodity chemicals. *NPJ Systems Biology and Applications*.
- Klamt, S., & Mahadevan, R., 2015. On the feasibility of growth-coupled product synthesis in microbial strains. *Metabolic engineering*.
- Löbs, A.K., Schwartz, C. and Wheeldon, I., 2017. Genome and metabolic engineering in non-conventional yeasts: current advances and applications. *Synthetic and Systems Biotechnology*.
- Meyer, V., 2021. Metabolic engineering of filamentous fungi. *Metabolic Engineering: Concepts and Applications*.
- Nielsen, J. and Keasling, J.D., 2016. Engineering cellular metabolism. *Cell*.
- Stephanopoulos, G., 2012. Synthetic biology and metabolic engineering. *ACS Synthetic Biology*.
- Volk, M. J., Lourentzou, I., Mishra, S., Vo, L. T., Zhai, C., & Zhao, H, 2020. Biosystems Design by Machine Learning. *ACS Synthetic Biology*.

## **Evolução adaptativa de plataformas microbianas industriais**

### **1. Qual é o princípio da evolução adaptativa e quais as suas principais aplicações?**

**Resposta:** Evolução adaptativa (ou engenharia evolutiva ou evolução adaptativa laboratorial (ALE) – termos sinônimos) é uma estratégia de melhoramento de microrganismos baseada no princípio de variação e seleção natural e amplamente utilizada para otimizar características de microrganismos de interesse biotecnológico. Nesse processo, pressões seletivas são impostas em uma célula isogênica, ou em uma população de células, visando a seleção de um fenótipo desejado. Mutações benéficas, ou seja, mutações que favorecerem o fenótipo na condição da pressão seletiva imposta, tendem a ser fixadas na população por fornecer vantagens adaptativas ao microrganismo. Entre os casos mais comuns de aplicação da tecnologia, destaca-se o uso do processo de evolução para selecionar microrganismos mais tolerantes a condições de estresse de um processo industrial, otimização de métricas de produção (TRY – Titer/Rate/Yield) e aumento de taxas de velocidade de consumo de substratos não convencionais.

**2. Por que abordagens multi-ômicas podem ser aplicadas na compreensão das alterações fisiológicas obtidas pelo microrganismo evoluído?**

**Resposta:** Uma das maiores vantagens ao se utilizar procedimentos de evolução adaptativa no melhoramento de microrganismos é que essa abordagem permite a identificação de novos alvos que não podem ser preditos racionalmente e expande as possibilidades de modificações que podem ser inseridas no seu design. É uma abordagem complementar as ferramentas racionais de modificação. A identificação do conjunto de modificações benéficas fixadas no genoma do microrganismo pode ser acessada pelo sequenciamento do genoma e posterior uso de abordagens de bioinformática. Ao comparar o genoma das cepas evoluídas com o parental selvagem, é possível buscar por modificações do tipo SNPs (Single Nucleotide Polymorphisms), CNVs (Copy Number Variation), Indels (INsertion/DELetion), entre outras. Procedimentos de engenharia reversa com a introdução das mutações identificadas no microrganismo parental utilizando CRISPR/Cas ou outras ferramentas de edição de genomas visam a confirmação de fenótipo obtido e a sua associação com as mutações identificadas. Frequentemente, a elucidação dos mecanismos responsáveis pela mudança fenotípica é uma tarefa complexa e nem sempre a identificação do conjunto de mutações no DNA da cepa evoluída permite a criação de hipóteses completas. Nesse sentido, a utilização de abordagens multi-ômicas complementares permite uma compreensão expandida e sistêmica, analisando as mudanças em níveis de expressão gênica (transcriptoma), quantidade de proteínas produzidas (proteoma), concentração de metabólitos e possíveis gargalos na via metabólica (metaboloma), entre outras. A integração desses dados permite uma compreensão mais ampla das alterações fisiológicas fixadas nas cepas evoluídas e fornece dados que podem ser utilizados para modificar racionalmente o microrganismo de interesse.

**3. As pressões seletivas impostas durante o procedimento de evolução adaptativa sempre produzem microrganismos com mutações estáveis?**

**Resposta:** Nem sempre. Na maior parte dos procedimentos evolução, um fenótipo superior selecionado está associado a mutações fixadas no genoma. Se esse conjunto de mutações não resultar em algum efeito que afete drasticamente o *fitness* do microrganismo, essa mutação tende a ser mantida. Mas existem algumas exceções onde as mutações adquiridas podem ser menos estáveis. O aumento do número de cópias resultando em duplicações *in tandem* de sequências gênicas podem ser sítios prováveis de recombinação e perda da região se a pressão seletiva não estiver sendo mantida para favorecer aquela característica. Outra possibilidade interessante e que já foi descrita em alguns experimentos de evolução, é quando o fenótipo superior 'adquirido' perde a estabilidade ao se remover a pressão seletiva. Diversas possibilidades podem ser responsáveis por esse evento. Uma das possíveis causas são mecanismos epigenéticos, onde alterações no fenótipo são causadas por variações em níveis de expressão gênica causadas por modificações em histonas. Mas tais modificações não são estáveis por não serem alterações fixadas no genoma e, portanto, não são características herdáveis estavelmente para gerações futuras. Existem procedimentos de evolução específicos que visam minimizar eventos como esse.

#### **4. Como o uso de biosensores pode ajudar na seleção de células superprodutoras do bioquímico de interesse?**

**Resposta:** Nem sempre uma pressão seletiva pode ser imposta no cultivo de um microrganismo para a obtenção de um fenótipo de interesse. Em um experimento de evolução, é possível otimizar o consumo de um substrato de maneira mais eficiente, melhorar parâmetros de cinética fermentativa de um processo industrial ou mesmo aumentar a resistência de um microrganismo a diferentes tipos de estresse. Porém, qual estratégia pode ser empregada para aumentar a produção de uma molécula de interesse utilizando-se dos princípios de variabilidade gênica e seleção em larga escala de um processo de evolução? Nesse contexto, biosensores podem ser construídos e usados para monitorar e regular a biossíntese de produtos de interesse. O termo biossensor é amplo e diferentes abordagens podem ser empregadas. Mas no contexto do curso CBS, um biossensor pode ser um elemento construído dentro de um circuito genético que é responsivo a presença do metabólito de interesse e regula diretamente um gene repórter. O uso de fatores de transcrição responsivos a moléculas específicas de interesse é um dos elementos mais empregados no design de biossensores. A diversidade gênica pode ser introduzida artificialmente através de agentes mutagênicos e as mutações benéficas resultantes, que aumentam a produção do composto de interesse, geram um maior sinal através do gene repórter induzido pelo biossensor construído no circuito genético. Esse sinal pode ser a emissão de fluorescência que é detectada por um citômetro de fluxo acoplado a um FACS (*Fluorescence-Activated Cell Sorting*) e leva a separação de células com maiores níveis de produção do composto de interesse. Associar abordagens que induzem o aumento da diversidade genética com o uso de biossensores específicos e triagem em escala

*high-throughput* utilizando o FACS é uma das estratégias mais poderosas atualmente na seleção de microrganismos superprodutores do composto de interesse.

**5. Como interações gênicas podem influenciar no fenótipo observado no experimento de evolução adaptativa?**

**Resposta:** Epistasia é um evento de interação gênica onde a ação de mais de um gene é responsável pela característica fenotípica observada. Durante um experimento de evolução adaptativa, é possível que o conjunto de mutações identificadas durante o sequenciamento do genoma das células evoluídas tenha um efeito aditivo e seja responsável pelo microrganismo superior selecionado. A sinergia da combinação das mutações pode ser verificada durante a engenharia reversa com a introdução das mutações no microrganismo parental isoladamente e em combinação. Procedimentos de evolução de longa duração geralmente selecionam esse tipo de evento. O mecanismo molecular responsável pela interação gênica normalmente não óbvio e pode ser explorado com o auxílio de abordagens multi-ômicas, como explicado anteriormente.

Literatura sugerida sobre evolução adaptativa:

- Van den Bergh B, Swings T, Fauvart M, Michiels J. Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution. *Microbiol Mol Biol Rev.* 2018 Jul 25;82(3):e00008-18.
- Horinouchi T, Furusawa C. Understanding metabolic adaptation by using bacterial laboratory evolution and trans-omics analysis. *Biophys Rev.* 2020 Jun;12(3):677-682.
- Long, A., Liti, G., Luptak, A. et al. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat Rev Genet* 16, 567–582 (2015).
- Sandberg TE, Salazar MJ, Weng LL, Palsson BO, Feist AM. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab Eng.* 2019 Dec;56:1-16.
- McDonald, M., Rice, D. & Desai, M. Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature* **531**, 233–236 (2016).

- Long, C.P., Antoniewicz, M.R. How adaptive evolution reshapes metabolism to improve fitness: recent advances and future outlook. *Current Opinion in Chemical Engineering*, 22, 209-215 (2018)
- Hossain GS, Saini M, Miyake R, Ling H, Chang MW. Genetic Biosensor Design for Natural Product Biosynthesis in Microorganisms. *Trends Biotechnol.* 2020 Jul;38(7):797-810.
- Caspeta L., Chen Y., Ghiaci P., Feizi A., Buskov S., Hallström B.M., Petranovic D., Nielsen J. Altered sterol composition renders yeast thermotolerant. *Science*. 2014; 346: 75-78

## **Ferramentas de biologia sintética para redução de sobrecarga metabólica e aumento de produção**

*1. Por que a introdução de vias de produção heterólogas em microrganismos podem causar uma sobrecarga metabólica?*

**Resposta:** A introdução de vias de produção pode afetar negativamente a rede regulatória nativa de microrganismos. A expressão de genes necessários para produção de um produto de interesse pode interferir, por exemplo, na disponibilidade de fatores necessários para as maquinarias de replicação, transcrição e tradução. A introdução de novas vias metabólicas também pode causar o acúmulo de intermediários ou produtos finais tóxicos que afetam o crescimento celular. Além disso, a geração de produtos de interesse pode requerer precursores que também atuam como precursores de vias nativas essenciais. Coletivamente, esses fatores podem levar a uma sobrecarga metabólica, que pode levar a uma redução tanto do fitness quanto da capacidade produtiva dos microrganismos engenheirados.

*2. Existe uma estratégia para reduzir a sobrecarga metabólica ou é dependente do tipo de processo e microrganismos utilizado?*

**Resposta:** Diversas estratégias têm sido empregadas no intuito de minimizar a sobrecarga metabólica causada pela introdução de vias de produção em microrganismos. Por exemplo, podemos citar estratégias que permitem o balanço da expressão de enzimas, a regulação dinâmica da via de produção, a separação de passos da via em diferentes cepas que juntas formam um consórcio microbiano, o emprego de sistemas cell-free, entre outros. O nível de sobrecarga metabólica pode sim variar de acordo com a via biossintética em estudo e o microrganismo utilizado.

*3. Como evitar esta superpopulação de mutantes?*

**Resposta:** A redução do enriquecimento de mutantes durante o processo de escalonamento pode ser alcançada por meio de estratégias que reduzam a vantagem competitiva ou viabilidade de mutantes. Por exemplo, as estratégias citadas na resposta acima podem reduzir a sobrecarga metabólica causada pela geração do produto de interesse, com isso a taxa de crescimento entre células produtoras e mutantes fica mais próxima. Com isso, mesmo com a emergência de alguns mutantes na população estas células não vão enriquecer tão rapidamente e dominar a fermentação. Outras estratégias incluem a diminuição do número de sequências repetitivas no genoma, as quais podem favorecer eventos de recombinação; o aumento do número de cópias de genes de interesse; o desenvolvimento de circuitos genéticos que acoplam a capacidade produtiva das células à capacidade de crescimento, entre outros.

**4. É possível utilizar DCCR ou outro tipo de técnica de planejamento experimental para diminuir a quantidade de combinações [de níveis de expressão de enzimas] que devem ser feitas? Como é a viabilidade de se fazer e validar tantas construções pensando em tempo/custos no laboratório?**

**Resposta:** Sim, o uso de metodologias de planejamento experimental, como DCCR (delineamento composto central rotacional), é uma excelente maneira para minimizar o número de ensaios. Essas estratégias facilitam a aquisição de informações fundamentais para a etapa de LEARN do ciclo de DESIGN-BUILD-TEST-LEARN, e podem otimizar o tempo e esforço dedicado em rounds de modificações genéticas. De qualquer forma, quando muitas combinações ainda se fazem necessárias, o uso de biossensores e automatização laboratorial também aumentam a viabilidades destes ensaios em laboratório.

**5. Quais são as principais dificuldades no momento de transferir essas cepas modificadas do laboratório para uma planta piloto?**

**Resposta:** Muitas vezes as condições de cultivo utilizadas no laboratório não representam as condições presentes em cultivos de maior escala. A presença de outros tipos de estresses, bióticos e abióticos, podem influenciar o comportamento dos microrganismos desenvolvidos e a performance desejada de produção e tolerância à estresses. Além disso, experimentos em escala de laboratório não requerem a propagação das células por muitas gerações. Fermentações em escalas maiores, requerem a propagação de células por muitas gerações. Durante este período, a população de células pode começar a se tornar muito heterogênea devido ao aparecimento e enriquecimento de mutantes que crescem mais rápido na população, mas produzem menos.

**6. Em qual banco de dados essas bibliotecas de combinações de vias são submetidas?**

**Resposta:** Um exemplo de repositório contendo partes genéticas, como promotores e



terminadores, em formato modular é o Registry of Standard Biological Parts ([http://parts.igem.org/Main\\_Page](http://parts.igem.org/Main_Page)). Neste banco de dados é possível adquirir, por exemplo, uma extensa coleção de partes genéticas de diversos organismos utilizadas pelos times participantes da competição iGEM. Além disso, outros repositórios como Addgene (<http://www.addgene.org>) são usados para depósito de outras coleções publicadas independentemente.

#### Literatura sugerida:

- Rugbjerg, P., & Sommer, M. O. A. (2019). Overcoming genetic heterogeneity in industrial fermentations. *Nature Biotechnology*, 37(8), 869–876. <https://doi.org/10.1038/s41587-019-0171-6>
- Wu, G., Yan, Q., Jones, J. A., Tang, Y. J., Fong, S. S., & Koffas, M. A. G. (2016). Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends in Biotechnology*, 34(8), 652–664. <https://doi.org/10.1016/j.tibtech.2016.02.010>
- Ceroni, F., Boo, A., Furini, S., Goroehowski, T. E., Borkowski, O., Ladak, Y. N., Awan, A. R., Gilbert, C., Stan, G. B., & Ellis, T. (2018). Burden-driven feedback control of gene expression. *Nature Methods*, 15(5), 387–393. <https://doi.org/10.1038/nmeth.4635>
- Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K., & Sommer, M. O. A. (2018). Diverse genetic error modes constrain large-scale bio-based production. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-03232-w>
- Moser, F., Espah Borujeni, A., Ghodasara, A. N., Cameron, E., Park, Y., & Voigt, C. A. (2018). Dynamic control of endogenous metabolism with combinatorial logic circuits. *Molecular Systems Biology*, 14(11), 1–18. <https://doi.org/10.15252/msb.20188605>
- Young, E. M., Zhao, Z., Gielesen, B. E. M., Wu, L., Benjamin Gordon, D., Roubos, J. A., & Voigt, C. A. (2018). Iterative algorithm-guided design of massive strain

- libraries, applied to itaconic acid production in yeast. *Metabolic Engineering*, 48(May), 33–43. <https://doi.org/10.1016/j.ymben.2018.05.002>
- Naseri, G., & Koffas, M. A. G. (2020). Application of combinatorial optimization strategies in synthetic biology. *Nature Communications*, 11(2446), 1–14. <https://doi.org/10.1038/s41467-020-16175-y>
  - Young, R., Haines, M., Storch, M., & Freemont, P. S. (2020). Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. *Metabolic Engineering*, 63(November 2020), 81–101. <https://doi.org/10.1016/j.ymben.2020.12.001>
  - Tan, S. Z., & Prather, K. L. (2017). Dynamic pathway regulation: recent advances and methods of construction. *Current Opinion in Chemical Biology*, 41, 28–35. <https://doi.org/10.1016/j.cbpa.2017.10.004>
  - Rugbjerg, P., Sarup-Lytzen, K., Nagy, M., & Sommer, M. O. A. (2018). Synthetic addiction extends the productive life time of engineered Escherichia coli populations. *Proceedings of the National Academy of Sciences of the United States of America*, 115(10), 2347–2352. <https://doi.org/10.1073/pnas.1718622115>
  - Gupta, I. M. B. Reizman, C. R. Reisch, K. L. J. Prather, Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat. Biotechnol.* 35, 273–279 (2017).
  - Doong, S. J., Gupta, A. & Prather, K. L. J. Layered dynamic regulation for improving metabolic pathway productivity in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* 115, 2964–2969 (2018).
  - Corrêa, G. G. et al. A modular autoinduction device for control of gene expression in *Bacillus subtilis*. *Metab. Eng.* 61, 326–334 (2020).
  - Deaner, M. & Alper, H. S. Systematic testing of enzyme perturbation sensitivities via graded dCas9 modulation in *Saccharomyces cerevisiae*. *Metab. Eng.* 40, 14–22 (2017).

## Engenharia de proteínas na construção de rotas metabólicas

### 1. Como definir qual reação dar enfoque em uma via metabólica?

**Resposta:** Muitas pesquisas se baseiam nos parâmetros termodinâmicos das enzimas que compõem uma rota metabólica para determinar qual será o ponto de “problema” na produção de algum produto de interesse. Entretanto, muitos destes parâmetros são obtidos por análises *in vitro* (fora da célula, em ambientes controlados), que, assim, não se correlacionam ao que acontecerá *in vivo* (dentro da célula, ambiente real). Desta forma, a principal maneira de encontrar um gargalo em uma rota é experimentalmente. Primeiramente, é necessário incorporar a rota metabólica biossintética no seu organismo hospedeiro e, a partir daí, realizar experimentos e análises para mensurar rendimentos de todos os intermediários, produto e subprodutos em relação à quantidade de substrato consumido. Normalmente, intermediários com alto rendimento indicam que a reação subjacente possui um problema a ser investigado. Estas quantificações são realizadas por HPLC (Cromatografia Líquida) ou GC (Cromatografia Gasosa), principalmente.

De forma concomitante, a literatura pode ser uma forte aliada para prever gargalos nas rotas biossintéticas. Como exemplo, já se sabe que vias metabólicas com desbalanço redox (NADH produzido não é totalmente consumido, por exemplo) tendem a limitar o rendimento do produto; ou enzimas que utilizem cofatores escassos na célula também tendem a ser limitantes (como enzimas que utilizam ferro como cofator).

### 2. Quando eu opto por uma mutagênese aleatória ou uma mutagênese racional?

**Resposta:** Uma das principais características para optar por um método ou outro é a existência de informações estruturais desta proteína. Algumas abordagens não podem ser resolvidas por metodologias de mutagênese aleatória, o que torna obrigatória a utilização de um desenho racional para as mutações. Isso acontece quando, por exemplo, se necessita mudar o substrato preferencial de uma enzima e não se possui um método simples para seleção do melhor mutante, sendo assim necessário testar todos variantes isoladamente. Para este trabalho de mutagênese racional é necessária a informação estrutural da proteína, que às vezes, mesmo com muito esforço, não é alcançada. Uma forma de se apontar mutações é pela utilização indireta de proteínas ortólogas com informação estrutural, ou seja, a partir de alinhamentos de sequência de várias proteínas “parecidas” e comparação com as estruturas existentes é possível prever quais aminoácidos estão envolvidos na ligação do substrato e, assim, limitar as mutações.

Por outro lado, algumas abordagens são resolvidas por mutagênese aleatória mesmo com a existência de informações estruturais, como por exemplo, quando se deseja modificar pH e temperatura ótimos de atividade de uma enzima. Estes parâmetros são muito difíceis de serem previstos a partir da composição da proteína. No entanto, o

uso de métodos de mutagênese aleatória requer um grande esforço para encontrar o melhor mutante entre todas as variantes geradas. Os principais métodos para mutagênese aleatória utilizados são o *error-prone PCR* e o *DNA shuffling*.

### 3. Como obter a proteína pura para ensaios de cristalografia?

**Resposta:** A determinação da estrutura terciária é realizada a partir de ensaios de difração de cristais gerados com uma proteína pura, para que esta possa se organizar e cristalizar com maiores chances de sucesso. Normalmente, esta proteína em estudo é expressa heterologicamente em organismos modelos que permitem alta produção, como a bactéria *Escherichia coli* ou a levedura *Pichia pastoris*. Na construção do vetor de expressão, adiciona-se uma sequência rica em histidinas (6xHis) na extremidade N- ou C-terminal da proteína, que permite que esta seja purificada por cromatografia de afinidade, uma vez que a histidina possui afinidade por níquel ou cobalto. Esta cromatografia de afinidade é um processo fácil e barato para purificação apenas da proteína de interesse, pois é necessário apenas passar o extrato celular por uma resina com níquel e apenas esta proteína de interesse permanecerá ligada. Embora este passo já resulte em uma proteína bastante pura, passos adicionais podem ser utilizados para aumentar a pureza da proteína, como cromatografia de exclusão molecular, em que se utiliza uma coluna porosa que separa as proteínas por tamanho molecular e assim é possível recuperar apenas a proteína de interesse quando há proteínas contaminantes com pesos moleculares diferentes. Há alguns métodos biofísicos para se avaliar a homogeneidade da sua proteína, o principal deles é o DLS (*Dynamic light scattering*), que se baseia na propriedade física de espalhamento de luz que toda partícula causa; como este espalhamento é relativo ao tamanho das partículas na solução, é possível estipular se a maioria das partículas tem o mesmo tamanho e, assim, estipular o estado oligomérico da proteína em solução. Dessa forma, para maiores chances de sucesso na etapa de cristalização, é crucial obter uma amostra de proteína pura e homogênea (apenas um estado oligomérico).

### 4. Como se obter a estrutura tridimensional quando a cristalografia falha?

**Resposta:** Ao empregarmos a técnica de Cristalografia de raios-X para resolução de estruturas tridimensionais de proteínas, podemos encontrar problemas em diversas etapas, desde a não obtenção dos cristais de proteína ou até mesmo na etapa de difração. Entretanto, existem alternativas que também possibilitam os estudos estruturais de proteínas. Atualmente, há a utilização de softwares de predição de estrutura terciária que se baseiam em inteligência artificial e estruturas homólogas, como o RoseTTAFold e o AlphaFold2. Há também outras técnicas que podem ser utilizadas para se obter informações estruturais, como *crosslinking* com análise de espectroscopia de massa (XLMS), ressonância nuclear magnética (RNM) e criomicroscopia (Cryo-EM); no entanto, todas estas técnicas têm suas peculiaridades que ainda não podem ser empregadas em todos os casos, como por exemplo, RMN

funciona melhor com proteínas menores, enquanto Cryo-EM funciona melhor para proteínas grandes ou complexos.

**5. Como saber se minha proteína modificada foi expressa com sucesso e como checar sua localização na célula?**

**Resposta:** Após terminar toda a experimentação dos seus mutantes *in vitro*, ou mesmo em um organismo modelo, é necessário reincluir esta proteína no “organismo produtor”. É aconselhável avaliar se as mutações não geraram algum problema relacionado a expressão da proteína ou mesmo uma inatividade total. Isso pode ser observado diretamente pelos rendimentos dos intermediários e produto, ou então por abordagens que provam que sua proteína está presente naquele sistema. Uma destas abordagens é a adição de uma sequência sintética na sua proteína de estudo para realizar um ensaio de imuno-localização (WesternBlot), o que prova que sua proteína foi expressa com sucesso. Outra forma é a utilização do extrato celular bruto do seu “organismo produtor” para aferição da atividade *in vitro*, pois mesmo que esta atividade seja observada *in vitro*, a proteína provavelmente estará ativa *in vivo*.

Uma outra questão, mais importante em trabalhos utilizando hospedeiros eucarióticos, é quanto a localização da proteína de estudo. Muitas vezes é interessante direcionar as rotas metabólicas a uma ou outra organela dentro da célula; e isto pode ser feito por meio da adição de sequências direcionadoras, denominadas peptídeo sinal. A técnica mais prática para checar a localização da proteína heteróloga é expressá-la *in frame* com um gene repórter, como *Green Fluorescent Protein* (GFP), por exemplo. Desta forma, você consegue identificar a localização por visualização com microscopia de fluorescência. Uma outra forma é a utilização da técnica de centrifugação diferencial, em que se separa o conteúdo do citoplasma e o que está dentro de alguma organela.

Literatura sugerida sobre evolução dirigida de proteínas:

Santos CA et al. *Sci Rep*, 2019. Doi: 10.1038/s41598-019-41300-3.

Zanphorlin LM et al. *Biotechnol Bioeng*, 2019. Doi: 10.1002/bit.26899

Farwick A et al. *Proc Natl Acad Sci USA*, 2014. Doi: 10.1073/pnas.1323464111

Schmidt-Dannert C et al. *Nat Biotechnol*, 2000. Doi: 10.1038/77319

Thomik T et al. *Nat Chem Biol*, 2017. Doi: 10.1038/nchembio.2457

Zhao EM et al. *Nature*, 2018. Doi: 10.1038/nature26141

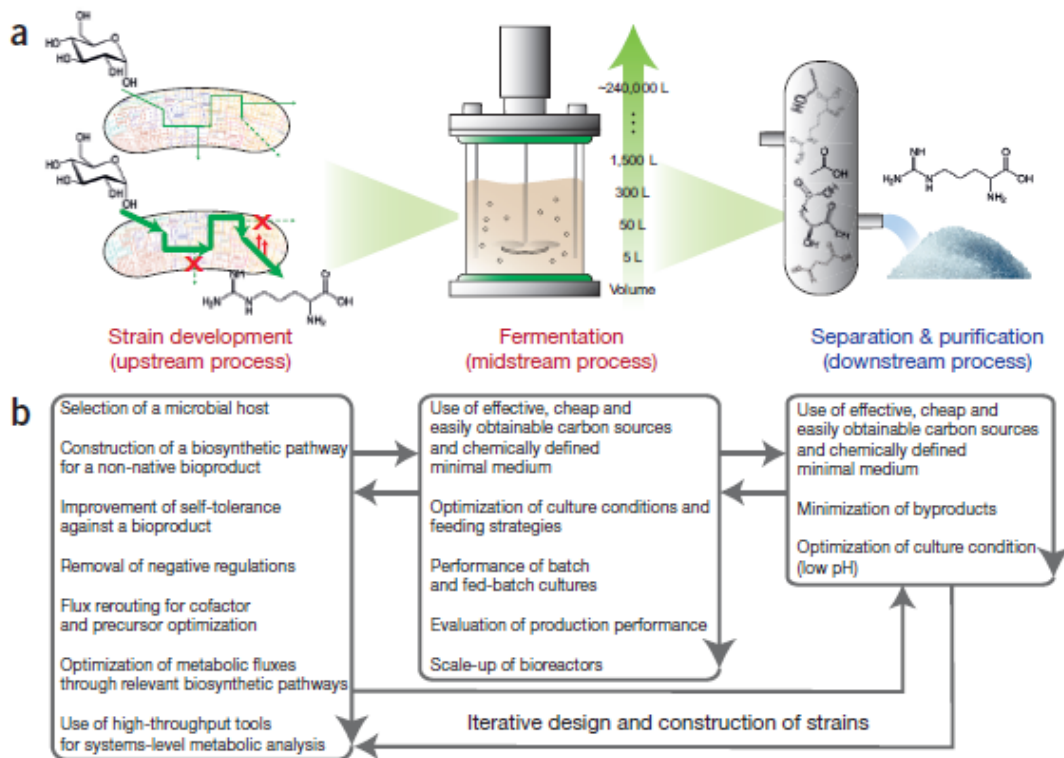
## **Desenvolvimento de bioprocessos industriais**

Recentemente o conceito de bioprocessos foi ampliado por vários autores, sendo considerado a escolha da matéria-prima, a construção de microrganismos robustos além do desenvolvimento de processos fermentativos e de separação e purificação eficientes (Figura 2).

- Processo *upstream*: nesta etapa considera-se a escolha da matéria-prima (açúcar, material lignocelulósico, óleo vegetal, gás) ideal para o processo idealizado, assim como a escolha do chassi e construção do microrganismo utilizando as estratégias de Biologia Sintética apresentadas no curso.

- Processo *midstream*: nesta etapa considera-se o desenvolvimento de um processo fermentativo com o objetivo de aumentar a concentração, produtividade e rendimento do produto de interesse (TRY - Titer, Rate, Yield). Em escala de laboratório, normalmente utiliza-se biorreatores automatizados de 0,5 a 5,0 L para (i) selecionar o meio de cultura mais adequado para o processo idealizado; (ii) definir estratégia de alimentação (batelada / batelada-alimentada / contínuo / com reciclo de célula); (iii) identificar a produção de sub-produtos e possíveis efeitos de inibição do crescimento celular e/ou do produto de interesse que possam ser contornados em novos rounds de construção do microrganismo; (iv) construir modelos cinéticos e/ou GEMs fundamentais para novos rounds de construção do microrganismo.

- Processo *downstream*: nesta etapa considera-se o desenvolvimento de um processo para separação e purificação do produto de interesse. É uma etapa de grande relevância a nível industrial já que pode representar de 20–50% (ou mais) do custo total de um bioprocessos, dependendo da quantidade de produto produzida no processo fermentativo, número de etapas e do tipo de equipamentos necessários, além do grau de pureza necessário para o processo idealizado. O *downstream* segue uma sequência de operações unitárias que geralmente consiste em separar as células (sólido) do meio líquido (caldo) através de operações como centrifugação e/ou filtração (vale ressaltar que para produtos intracelulares, como algumas enzimas, constituintes da parede celular, uma etapa de rompimento celular é requerida) e purificar o produto de interesse através de operações como precipitação, ultrafiltração, cromatografia, entre outras, dependendo da pureza da produto-alvo.



**Figura 2** – As três principais etapas envolvidas no desenvolvimento de um bioprocesso: Desenvolvimento de cepas eficientes (*upstream*); processo fermentativo (*midstream*); processo de separação e purificação (*downstream*). **Fonte:** Lee e Kim (2015)

Literatura sugerida sobre bioprocessos:

- Campbell, K., Xia, J., Nielsen, J. (2017). The impact of systems biology on bioprocessing. *Trends in biotechnology*, 35(12), 1156-1168.
- Lee, S. Y., Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nature biotechnology*, 33(10), 1061-1072.
- Rangel, A. E., Gómez Ramírez, J. M., Gonzalez Barrios, A. F. (2020). From industrial by-products to value-added compounds: the design of efficient microbial cell factories by coupling systems metabolic engineering and bioprocesses. *Biofuels, Bioproducts and Biorefining*, 14(6), 1228-1238.

## Sobre o CNPEM

Ambiente de pesquisa e desenvolvimento sofisticado e efervescente, único no País e presente em poucos polos científicos no mundo, o Centro Nacional de Pesquisa em Energia e Materiais (CNPEM) é uma organização social supervisionada pelo Ministério da Ciência, Tecnologia, Inovações (MCTI). O Centro é constituído por quatro Laboratórios Nacionais e é berço do mais complexo projeto da ciência brasileira – o Sirius – uma das mais avançadas fontes de luz síncrotron do mundo. O CNPEM reúne equipes multitemáticas altamente especializadas, infraestruturas laboratoriais mundialmente competitivas e abertas à comunidade científica, linhas de pesquisa em áreas estratégicas, projetos inovadores em parcerias com o setor produtivo e ações de treinamento para pesquisadores e estudantes.

O Centro constitui um ambiente movido pela busca de soluções com impacto nas áreas de Saúde, Energia, Meio Ambiente, Novos Materiais, entre outras. As competências singulares e complementares presentes no CNPEM impulsionam pesquisas e desenvolvimentos inovadores nas áreas de luz síncrotron; engenharia de aceleradores; descoberta de novos medicamentos, inclusive a partir de espécies vegetais da biodiversidade brasileira; mecanismos moleculares envolvidos no início e progressão do câncer; doenças cardíacas e neurodesenvolvimento; nanopartículas funcionalizadas para combater bactérias, vírus, câncer; novos sensores e dispositivos nanoestruturados para os setores de petróleo e gás e saúde; soluções biotecnológicas para o desenvolvimento sustentável de biocombustíveis avançados, bioquímicos e biomateriais.