Bioproduct Production

A Continually Evolving Process
Learning Objectives

• Explain the assumptions and limitations of strain design using constraint-based metabolic reconstructions.
• Explain the process of bioproduct identification.
• Explain the process of selecting a host strain.
• Explain the process of defining a bioproduct pathway.
• Explain the strain design process.
• Explain the purpose of understanding the bioproduct maximum production.
• Explain the purpose of carbon source selection.
• Explain the purpose of identifying nutrient & amino acid limitations.
• Explain the purpose of identifying undesired by-products.
• Explain the purpose of growth coupling.
• Explain the purpose of cofactor balancing.
• Explain the purpose of sampling analysis.
Focus of Bioproduction Modeling

Purpose: Identify phenotypes that can exist in exponential growth phase
Assumptions & Limitations

• Model cell steady-state phenotypes during exponential growth phase.
  ✓ Can model the different phenotypes that can exist during the exponential growth phase.
  ✓ Can understand the capabilities of each phenotype
  ✓ Can identify and modify cellular pathways to favor specific bioproduct producing phenotypes
  ✓ Constraint-based models do not model transitions between phenotypes
  ✓ Most genome-scale models do not include the genes required for the stationary phase (proteases, etc.)
  ✓ Most genome-scale models do not include the complete transcription and translation pathways

• The biomass function represents the average metabolic load required during exponential cell growth.
  ✓ The biomass function represents the average percentages of the component parts (amino acids, nucleotides, energy, etc.) that are included in 1 gm of cell biomass.

Purpose: Identify phenotypes that can exist in exponential growth phase
Host Strain Design Process

- Add Pathways & Plasmid
- Determine Production Upper Limits
- Identify & Manage Nutrient Limitations
- Identify & Limit Undesired Byproducts
- Identify & Add Knockouts
- Balance Cofactors
- Sampling Analysis
- Final Design

Strain A → Strain B → Strain C → Strain D → Strain E → Strain F → Strain G → Final Strain
Bioproduct Production

• Bioproduct Identification

• Select Host Strain

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
Categories of Chemicals Produced by Microbial Cell Factories

- **Natural-inherent chemicals** are endogenous metabolites in naturally isolated microorganisms and thus can be produced inherently through a native pathway.
- **Natural-noninherent chemicals** are those that are found in nature but are best produced in a heterologous host strain using noninherent pathways introduced from other hosts or metagenomes.
- **Non-natural–noninherent chemicals** are those that have not yet been found in nature but can be produced in a noninherent host strain by the establishment of heterologous pathways and enzymes, often using genes found from various sources in combination.
- **Nonnatural–created chemicals** are those that have not yet been found in nature and, owing to the lack of any known metabolic enzymes and pathways leading to their formation, can only be produced by creating synthetic enzymes and pathways with new functions.

Metabolite Databases

- Biocyc ✓ http://biocyc.org/
- KEGG ✓ http://www.genome.jp/kegg/
- OMIM ✓ http://omim.org/
- HGNC ✓ http://www.genenames.org/
- HPRD ✓ http://www.hprd.org/index_html
- e!Ensembl ✓ http://uswest.ensembl.org
- Vega ✓ http://vega.sanger.ac.uk/index.html
- UniProt ✓ http://www.uniprot.org/
Bioproduct Production

• Bioproduct Identification

• Select Host Strain

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
E. coli as a Host Organism

1. Unparalleled fast growth kinetics. In glucose-salts media and given the optimal environmental conditions, its doubling time is about 20 min. This means that a culture inoculated with a 1/100 dilution of a saturated starter culture may reach stationary phase in a few hours. However, it should be noted that the expression of a recombinant protein may impart a metabolic burden on the microorganism, causing a considerable decrease in generation time.

2. High cell density cultures are easily achieved. The theoretical density limit of an E. coli liquid culture is estimated to be about 200 gram dry cell weight/liter or roughly $10^{13}$ viable bacteria/ml. However, exponential growth in complex media leads to densities nowhere near that number.

3. Rich complex media can be made from readily available and inexpensive components.

4. Transformation with exogenous DNA is fast and easy. Plasmid transformation of E. coli can be performed in as little as 5 min.

Host Selection

• Select host that naturally produces the target bioproduct
• Select heterologous host that can be engineered
  ✓ Commonly used strains of E. coli
    ▪ [http://openwetware.org/wiki/E._coli_genotypes](http://openwetware.org/wiki/E._coli_genotypes)
• Commercial competent cell strains
  ✓ Invitrogen - Top10 -
  ✓ Promega -

<table>
<thead>
<tr>
<th>Some Commonly Used E. coli Strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG1</td>
</tr>
<tr>
<td>AB1157</td>
</tr>
<tr>
<td>B2155</td>
</tr>
<tr>
<td>BL21</td>
</tr>
<tr>
<td>BL21(AI)</td>
</tr>
<tr>
<td>BL21(DE3)</td>
</tr>
<tr>
<td>BL21(DE3) pLysS</td>
</tr>
<tr>
<td>BNN93</td>
</tr>
<tr>
<td>BNN97</td>
</tr>
<tr>
<td>BW26434, CGSC Strain # 7658</td>
</tr>
<tr>
<td>C600</td>
</tr>
<tr>
<td>C600 hflA150 (Y1073, BNN102)</td>
</tr>
<tr>
<td>CSH50</td>
</tr>
<tr>
<td>D1210</td>
</tr>
<tr>
<td>DB3.1</td>
</tr>
<tr>
<td>DH1</td>
</tr>
<tr>
<td>DH5α</td>
</tr>
<tr>
<td>DH5α Turbo (NEB)</td>
</tr>
<tr>
<td>DH10B (Invitrogen)</td>
</tr>
<tr>
<td>DH125 (Invitrogen)</td>
</tr>
<tr>
<td>DM1 (Invitrogen)</td>
</tr>
<tr>
<td>E. cloni(r) Salpha (Lucigen)</td>
</tr>
<tr>
<td>E. cloni(r) 10G (Lucigen)</td>
</tr>
<tr>
<td>E. cloni(r) 10G' (Lucigen)</td>
</tr>
<tr>
<td>E. coli K12 ER2738 (NEB)</td>
</tr>
<tr>
<td>ER2566 (NEB)</td>
</tr>
<tr>
<td>ER2267 (NEB)</td>
</tr>
<tr>
<td>HB101</td>
</tr>
<tr>
<td>HMS5174(DE3)</td>
</tr>
<tr>
<td>High-Control(tm) BL21(DE3) (Lucigen)</td>
</tr>
<tr>
<td>High-Control(tm) 10G (Lucigen)</td>
</tr>
</tbody>
</table>
### E. coli Genome-scale Reconstructions

- Escherichia coli 042
- Escherichia coli 536
- Escherichia coli 55989
- Escherichia coli ABU 83972
- Escherichia coli APEC O1
- Escherichia coli ATCC 8739
- Escherichia coli B str. REL606
- Escherichia coli BL21(DE3) AM946981
- Escherichia coli BL21(DE3) BL21-Gold(DE3)pLysS AG
- Escherichia coli BL21(DE3) CP001509
- Escherichia coli BW2952
- Escherichia coli CFT073
- Escherichia coli DH1
- Escherichia coli DH1 ME8569
- Escherichia coli E24377A
- Escherichia coli Ed1a
- Escherichia coli ETEC H10407
- Escherichia coli HS
- Escherichia coli IAI1
- Escherichia coli IAI39
- Escherichia coli IHE3034
- Escherichia coli KO11FL
- Escherichia coli LF82
- Escherichia coli NA114
- Escherichia coli O103:H2 str. 12009
- Escherichia coli O111:H- str. 11128
- Escherichia coli O127:H6 str. E2348/69
- Escherichia coli O157:H7 EDL933
- Escherichia coli O157:H7 str. EC4115
- Escherichia coli O157:H7 str. Sakai
- Escherichia coli O157:H7 str. TW14359
- Escherichia coli O26:H11 str. 11368
- Escherichia coli O55:H7 str. CB9615
- Escherichia coli O83:H1 str. NRG 857C
- Escherichia coli S88
- Escherichia coli SE11
- Escherichia coli SE15
- Escherichia coli SMS-3-5
- Escherichia coli str. K-12 substr. DH10B
- Escherichia coli str. K-12 substr. MG1655
- Escherichia coli str. K-12 substr. W3110
- Escherichia coli UM146
- Escherichia coli UMN026
- Escherichia coli UMNK88
- Escherichia coli UTI89
- Escherichia coli W
- Escherichia coli W CP002185
- Escherichia coli K-12 MG1655

Strain used at USU

---

Core and Pan Metabolic Capabilities of the E. coli Species.

For 55 unique E. coli strains

<table>
<thead>
<tr>
<th>Core:</th>
<th>Pan:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genes: 965</td>
<td>1,460</td>
</tr>
<tr>
<td>Reactions: 1,773</td>
<td>2,501</td>
</tr>
<tr>
<td>Metabolites: 1,665</td>
<td>2,043</td>
</tr>
</tbody>
</table>

B

Reaction Distribution by Subsystem

- Carbohydrate Metabolism
- Inner Membrane Transport
- Outer Membrane Transport
- Lipid Metabolism
- Cell Wall/Membrane/Envelope Metabolism
- Cofactor and Prosthetic Group Metabolism
- Amino Acid Metabolism
- Nucleotide Metabolism
- Inorganic Ion Transport and Metabolism
- Energy Production and Conversion
- Other

BL21(DE3) E.coli Strain

• BL21 cells are deficient in the Lon protease which degrades many foreign proteins.

• Another gene missing from the genome 21 is the one coding for the outer membrane protease OmpT, whose function is to degrade extracellular proteins. The liberated amino acids are then taken up by the cell. This is problematic in the expression of a recombinant protein as, after cell lysis, OmpT may digest it.

• Plasmid loss is prevented thanks to the hsdSB mutation which disrupts DNA methylation and degradation.

• When the genes for the desired recombinant protein placed under a T7 promoter, then T7 RNAP needs to be provided by the host cell. In the popular BL21(DE3) strain, the λDE3 prophage was inserted in the chromosome of BL21 and contains the T7 RNAP gene under the lacUV5 promoter.

• The BL21(DE3) and its derivatives are by far the most used strains for protein expression.

### Tested Cobra Toolbox Models

http://bigg.ucsd.edu/models

<table>
<thead>
<tr>
<th>Model Number</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>e_coli_core</td>
<td>Escherichia coli str. K-12 substr. MG1655</td>
</tr>
<tr>
<td>iAB_RCCB_283</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>iAF1260</td>
<td>Escherichia coli str. K-12 substr. MG1655</td>
</tr>
<tr>
<td>iAF692</td>
<td>Methanosaeta woesei sp. W1</td>
</tr>
<tr>
<td>iAF987</td>
<td>Geobacter metallireducens GS-5</td>
</tr>
<tr>
<td>iAPECO1_1312</td>
<td>Escherichia coli APEC O1</td>
</tr>
<tr>
<td>iAT_PLT_636</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>iB21_1397</td>
<td>Escherichia coli BL21(DE3)</td>
</tr>
<tr>
<td>iBWG_1329</td>
<td>Escherichia coli BW25113</td>
</tr>
<tr>
<td>ic_1306</td>
<td>Escherichia coli CFT073</td>
</tr>
<tr>
<td>iE2348C_1286</td>
<td>Escherichia coli O127:H6 str. E2348/69</td>
</tr>
<tr>
<td>iEC042_1314</td>
<td>Escherichia coli O42</td>
</tr>
<tr>
<td>iEC5989_1330</td>
<td>Escherichia coli 5989</td>
</tr>
<tr>
<td>iECABU_c1320</td>
<td>Escherichia coli ABU 83978</td>
</tr>
<tr>
<td>iECBH1_1328</td>
<td>Escherichia coli B str. REL606</td>
</tr>
<tr>
<td>iECBD_1354</td>
<td>Escherichia coli ‘BL21-Gold(DE3)pLysS AG’</td>
</tr>
<tr>
<td>iECDFH1_1363</td>
<td>Escherichia coli DH1</td>
</tr>
<tr>
<td>iECDFH1E_1363</td>
<td>Escherichia coli DH1</td>
</tr>
<tr>
<td>iECDFH1ME659_1439</td>
<td>Escherichia coli DH1</td>
</tr>
<tr>
<td>iE24377_1341</td>
<td>Escherichia coli E24377A</td>
</tr>
<tr>
<td>iE3D1_1282</td>
<td>Escherichia coli E3D1</td>
</tr>
<tr>
<td>iE471415_1262</td>
<td>Escherichia coli O157:H7 str. EC4115</td>
</tr>
<tr>
<td>iE5H3_1320</td>
<td>Escherichia coli H5</td>
</tr>
<tr>
<td>iECAI1_1343</td>
<td>Escherichia coli IA1</td>
</tr>
<tr>
<td>iECIA39_1322</td>
<td>Escherichia coli IA39</td>
</tr>
<tr>
<td>iECNA114_1301</td>
<td>Escherichia coli NA114</td>
</tr>
<tr>
<td>iECO103_1326</td>
<td>Escherichia coli O103:H2 str. 12009</td>
</tr>
<tr>
<td>iECO111_1330</td>
<td>Escherichia coli O111:H- str. 1128</td>
</tr>
<tr>
<td>iECO26_1355</td>
<td>Escherichia coli O26:H11 str. 11368</td>
</tr>
<tr>
<td>iECOL1_1307</td>
<td>Escherichia coli KE3034</td>
</tr>
<tr>
<td>iECOLC_1368</td>
<td>Escherichia coli ATCC 8739</td>
</tr>
<tr>
<td>iECPL_1309</td>
<td>Escherichia coli 536</td>
</tr>
<tr>
<td>iECX_1301</td>
<td>Escherichia coli O157:H7 str. Sakai</td>
</tr>
<tr>
<td>iEC588_1305</td>
<td>Escherichia coli 588</td>
</tr>
<tr>
<td>iECSE_1348</td>
<td>Escherichia coli SE11</td>
</tr>
<tr>
<td>iECSP_1327</td>
<td>Escherichia coli SE15</td>
</tr>
<tr>
<td>iEC5MS35_1347</td>
<td>Escherichia coli SMS-3-5</td>
</tr>
<tr>
<td>iECSP_1301</td>
<td>Escherichia coli O157:H7 str. TW14359</td>
</tr>
<tr>
<td>iECUMN_1333</td>
<td>Escherichia coli UMN026</td>
</tr>
<tr>
<td>iECW_1372</td>
<td>Escherichia coli W</td>
</tr>
<tr>
<td>iEKO11L_1354</td>
<td>Escherichia coli KO11FL</td>
</tr>
<tr>
<td>iETEC_1333</td>
<td>Escherichia coli ETEC H10407</td>
</tr>
<tr>
<td>iE2583_1286</td>
<td>Escherichia coli O55:H7 str. C89615</td>
</tr>
<tr>
<td>iEHN367</td>
<td>Clostridium ljungdahlii DSM 15258</td>
</tr>
<tr>
<td>iT341</td>
<td>Helicobacter pylori 26695</td>
</tr>
<tr>
<td>iJN678</td>
<td>Synechocystis sp. PCC 6803</td>
</tr>
<tr>
<td>iJN746</td>
<td>Pseudomonas putida KT2440</td>
</tr>
<tr>
<td>iJO1366</td>
<td>Escherichia coli K-12 substr. MG1655</td>
</tr>
<tr>
<td>iJR904</td>
<td>Escherichia coli K-12 substr. MG1655</td>
</tr>
<tr>
<td>iLF82_1304</td>
<td>Escherichia coli LF82</td>
</tr>
<tr>
<td>iLJ478</td>
<td>Thermotoga maritima MS8B</td>
</tr>
<tr>
<td>iMM1415</td>
<td>Mus musculus</td>
</tr>
<tr>
<td>iMM904</td>
<td>Saccharomyces cerevisiae S288c</td>
</tr>
<tr>
<td>iND750</td>
<td>Saccharomyces cerevisiae S288c</td>
</tr>
<tr>
<td>iNF661</td>
<td>Mycobacterium tuberculosis H37Rv</td>
</tr>
<tr>
<td>iNRG857_1313</td>
<td>Escherichia coli O83:H1 str. NR6 857C</td>
</tr>
<tr>
<td>iPC815</td>
<td>Yersinia pestis CO92</td>
</tr>
<tr>
<td>iRC1080</td>
<td>Chlamydomonas reinhardtii</td>
</tr>
<tr>
<td>iS_1188</td>
<td>Shigella flexneri 2a str. 24577</td>
</tr>
<tr>
<td>iSB619</td>
<td>Staphylococcus aureus subsp. aureus N315</td>
</tr>
<tr>
<td>iSbB5512_1146</td>
<td>Shigella boydii CDC 3083-94</td>
</tr>
<tr>
<td>iSO_1134</td>
<td>Shigella boydii Sb227</td>
</tr>
<tr>
<td>iSDY_1059</td>
<td>Shigella dysenteriae Sd197</td>
</tr>
<tr>
<td>iSF_1195</td>
<td>Shigella flexneri 2a str. 301</td>
</tr>
<tr>
<td>iSFV_1184</td>
<td>Shigella flexneri 5 str. 8401</td>
</tr>
<tr>
<td>iSFV_1172</td>
<td>Shigella flexneri 2002017</td>
</tr>
<tr>
<td>iSSON_1240</td>
<td>Shigella sonnei Ss046</td>
</tr>
<tr>
<td>iUMN146_1321</td>
<td>Escherichia coli UM146</td>
</tr>
<tr>
<td>iUMNK88_1353</td>
<td>Escherichia coli UMNK88</td>
</tr>
<tr>
<td>iUTI89_1310</td>
<td>Escherichia coli UTI89</td>
</tr>
<tr>
<td>iWFL_1372</td>
<td>Escherichia coli W</td>
</tr>
<tr>
<td>iY75_1357</td>
<td>Escherichia coli str. K-12 substr. W3110</td>
</tr>
<tr>
<td>iYL1228</td>
<td>Klebsiella pneumoniae subsp. pneumoniae MGH 78578</td>
</tr>
<tr>
<td>iYO844</td>
<td>Bacillus subtilis subsp. subtilis str. 168</td>
</tr>
<tr>
<td>iZ_1308</td>
<td>Escherichia coli O157:H7 str. EDL933</td>
</tr>
<tr>
<td>iRECONJ</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>iSTM_v1_0</td>
<td>Salmonella enterica subsp. enterica serovar Typhimurium str. LT2</td>
</tr>
</tbody>
</table>
Bioproduct Production

• Bioproduct Identification

• Select Host Strain

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
Defining the Pathway of the Desired Bioproduct

α-ketoglutarate

L-Glutamate

Secreted Bioproduct
Pathway Databases

- Biocyc
  - http://biocyc.org/
- KEGG
  - http://www.genome.jp/kegg/
- AmiGO 2
  - http://amigo.geneontology.org/amigo/landing
- Reactome
  - http://www.reactome.org/
- WikiPathways
- CPDB
  - http://cpdb.molgen.mpg.de/
OptStrain uses a multi-step process to first identify non-native reactions that would improve the host organism's maximum production capabilities. Reaction deletions can then be found which couple production and growth in the modified host metabolic network.

**OptStrain procedure.**

- **Step 1 -** Curation of database(s) of reactions to compile the Universal database, comprising only elementally balanced reactions.

- **Step 2 -** Identifies a maximum-yield path enabling the desired biotransformation from a substrate to product without any consideration for the origin of reactions. Note that the white arrows represent native reactions of the host and the yellow arrows denote non-native reactions.

- **Step 3 -** Minimizes the reliance on non-native reactions.

- **Step 4 -** Incorporates the non-native functionalities into the microbial host's stoichiometric model and applies the OptKnock procedure to identify and eliminate reactions competing with the targeted product. The red X's pinpoint the deleted reactions.

FMM (From Metabolite to Metabolite) is a tool for synthetic biology.

FMM can reconstruct metabolic pathways from one metabolite to the other one.

The different KEGG maps can be connected visually to show both local and global graphical views of the metabolic pathways.

Metabolic pathways can be comparative between several species by FMM (Comparative Analysis).

Post-translational modification (PTM) information of enzymes from numerous species is also supplied in FMM.

Open server access:

http://fmm.mbc.nctu.edu.tw/index.php

Bioproduct Production

• Bioproduct Identification

• Select Host Strain

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
Creating the Pathway in the Model: Adding Reactions

addReaction - Add a reaction to the model or modify an existing reaction

```
model = addReaction(model,rxnName,metaboliteList,stoichCoeffList,revFlag,lowerBound,upperBound,objCoeff,subSystem,grRule,geneNameList,systNameList,checkDuplicate)
model = addReaction(model,rxnName,rxnFormula)
```

**INPUTS**
- **model**: COBRA model structure
- **rxnName**: Reaction name abbreviation (i.e. 'ACALD')
  (Note: can also be a cell array {'abbr','name'})
- **metaboliteList**: Cell array of metabolite names or alternatively the reaction formula for the reaction
- **stoichCoeffList**: List of stoichiometric coefficients (reactants -ve, products +ve), empty if reaction formula is provided

**OPTIONAL INPUTS**
- **revFlag**: Reversibility flag (Default = true)
- **lowerBound**: Lower bound (Default = 0 or -vMax)
- **upperBound**: Upper bound (Default = vMax)
- **objCoeff**: Objective coefficient (Default = 0)
- **subSystem**: Subsystem (Default = '')
- **grRule**: Gene-reaction rule in boolean format (and/or allowed) (Default = '')
- **geneNameList**: List of gene names (used only for translation from common gene names to systematic gene names)
- **systNameList**: List of systematic names
- **checkDuplicate**: Check S matrix too see if a duplicate reaction is already in the model (Default true)

**OUTPUTS**
- **model**: COBRA model structure with new reaction
- **rxnIDexists**: Empty if the reaction did not exist previously, or if checkDuplicate is false. Otherwise it contains the ID of an identical reaction already present in the model.

**EXAMPLES**
1) Add a new irreversible reaction using the formula approach

```
model = addReaction(model,'newRxn1','A -> B + 2 C')
```

2) Add the same reaction using the list approach

```
model = addReaction(model,'newRxn1',{'A','B','C'},[-1 1 2],false);
```

http://opencobra.sourceforge.net/openCOBRA/opencobra_documentation/cobra_toolbox_2/index.html
Bioproduct Production

• Bioproduct Identification
• Select Host Strain
• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model
• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
Adding a Plasmid to a Cobra Model

• A plasmid is an expression vector designed for protein production in a cell. An expression vector must have elements necessary for protein expression. These may include a promoter, affinity tags for purification, the cloning site for the desired recombinant protein, a terminator and a selection marker (antibiotic resistance).

• The first option is to include all the plasmid components in the biomass function. This option is difficult to provide the flexibility needed to account for the variation in plasmid copy number and promoter strength.

• The second option is to create separate reactions that can be adjusted to model the plasmid load on the host cell. These separate reactions can include: Nucleotide precursors, antibiotic marker production, Recombinant protein production, and Lac repressor production (for Lac promoters).
Anatomy of an Expression Vector


Promoters

- **T5 Promoter**: The T5 promoter gives high protein expression levels; it is tightly repressible, and works in any *E. coli* strain. Silencing of the promoter prior to IPTG induction is achieved using symmetrical lac operators spaced around the promoter to maximize cooperativity. This operator pair ensures significantly tighter repression than regular lac operators.

- **T7 Promoter**: The T7 promoter results in high protein expression levels and is repressible. The T7 promoter only works in T7 pol expressing *E. coli* strains (e.g., BL21(DE3) or T7 Express). Inducible by IPTG.

- **PhoA Promoter**: The bacterial alkaline phosphatase (phoA) promoter is a strong promoter that is inexpensively induced when the culture is starved for inorganic phosphate. In the presence of phosphate, expression is tightly regulated making this system useful for expression of toxic proteins. Very high product yields have been obtained by use of the phoA system.

- **Arabinose Promoter**: The arabinose-inducible promoter *ara* is capable of high level recombinant protein expression in the presence of arabinose and is tightly regulated by glucose in the absence of arabinose. The *ara* promoter controls the genes *ara* organized in one operon. The promoter is flanked by a pair of lac operators that are recognized by the lac repressor also carried on the plasmid. IPTG binds to the repressor, inducing expression. The *ara* promoter is compatible with *E. coli* strains BL21 or DH5a. In *E. coli*, glucose shuts down expression very well, even in the presence of rhamnose. This is why it works for autoinduction. You add both rhamnose and glucose, once glucose is consumed rhamnose induces.

- **Rhamnose Promoter**: Rhamnose-inducible bacterial expression vectors with different strength ribosome binding sites and origins of replication provide an excellent range of induced and uninduced expression levels. Rhamnose Expression Vectors are tightly controlled by rhamnose, enabling high expression yields, even for toxic or challenging proteins and can be used in any *E. coli* strain or other Gram-negative bacteria.

Components Produced by Plasmid

- Lac Repressor (154 kDa)
- Beta Lactamase (29 kDa)
- RP Gene
- pET101
- pBR322 ori
- T7 promoter

Nucleotide precursors
Plasmid Production Load

- **Nucleotide precursors**
  - The plasmid requires four nucleotide precursors (dGTP, dCTP, dATP, dTTP) for its replication based on an approximated E. coli GC content of 50.6%.

- **Total antibiotic marker production**
  - Can be up to 3% of the total protein mass

- **Recombinant protein production**
  - With high copy, strong promoter vectors the recombinant protein concentration can be as high as 50% of the total protein mass

- **Lac repressor production**
  - There must be a low concentration of repressors or gene expression could not be induced
  - Typically less than 0.1% of the total protein; should be able to ignored
Creating the Nucleotide Precursors Reactions

The biosynthetic precursors and energetic requirements for plasmid DNA replication

- Each plasmid (number of base pairs) requires four nucleotide precursors (dGTP, dCTP, dATP, & dTTP) for its DNA replication based on an approximated hosts GC content.

- For the case of the pOri2, which is composed of 4,575 bp, the plasmid requires four nucleotide precursors for its replication based on an approximated E coli K-12 GC content of 50.6%:
  - dGTP coefficient = 4575 * 0.506 = 2315.0
  - dCTP coefficient = 4575 * 0.506 = 2315.0
  - dATP coefficient = 4575 * 0.494 = 2260.1
  - dTTP coefficient = 4575 * 0.494 = 2260.1

- Thus
  - model=addReaction(model,'PLASMID','2315.0 dgtp[c] + 2315.0 dctp[c] + 2260.1 datp[c] + 2260.1 dttp[c] -> plasmid[c]');

- A demand reaction for the plasmid is required to set the production rate of the nucleotide precursors .
  - model = addDemandReaction(model,'plasmid[c]');
  - model = changeRxnBounds(model,'DM_plasmid[c]',0.63e-9,'l'); % Measured flux rates 0.64 pmol/gDW-hr (copy number ~600)

Creating the Antibiotic Marker Reactions
The biosynthetic precursors and energetic requirements for antibiotic marker protein

- The biosynthetic precursor balance for the sole plasmid encoded antibiotic marker protein combined with energetic requirements of 4.306 mol ATP/mol amino acids.

- The reaction includes the number of each amino acids that makes up the protein plus the energetic requirements of ATP, ADP and Pi which is the sum of the number of amino acids multiplied by 4.306 mol ATP/mol amino acids.

- For the case of the pOri2, where β-lactamase is used for resistance to penicillins, cephalosporins, cephemycins and carbapenems, the total number of amino acids is 286 (286 * 4.306 = 1,231.52)


- A demand reaction for the β-lactamase gene assumes 3% of total cellular protein (copy number ~600).

  ✓ model = addDemandReaction(model,'b-lactamase[c]');

  ✓ model = changeRxnBounds(model,'DM_b-lactamase[c]',0.000569,'l');

- The model does not include reactions for the synthesis of stress-response metabolites and proteins.

Creating Recombinant Protein Enzyme Reactions

- The biosynthetic precursor balance for the sole plasmid encoded recombinant protein combined with energetic requirements of 4.306 mol ATP/mol amino acids.

- The reaction includes the number of each amino acids that makes up the protein plus the energetic requirements of ATP, ADP and Pi which is the sum of the number of amino acids multiplied by 4.306 mol ATP/mol amino acids.

- For example, the heme oxygenase gene is represented by:

  ```
  ```

- A demand reaction for the heme oxygenase gene is required to set can be as high as 50% of the total protein mass.

  ```
  ✓ model = addDemandReaction(model,'HO[c]');
  ✓ model = changeRxnBounds(model,'DM_HO[c]',0.0102,'l');
  ```

- Remember that Cobra biomass includes the required components (amino acids, etc.) of the host cell reactions in the biomass function, thus the need to add the separate load of the recombinant protein enzyme to the model.
Cobra Tools for Analyzing the Plasmid Impact

- **Robustness Analysis** - Determines the change in fluxes for any reaction in both the plasmid and plasmid-free strains.

- **Production Envelopes** - Determines the bioproduct production capabilities of a strain that includes the plasmid.

- **Phenotype Phase Plane Analysis** - Determines the phenotype phase planes for the plasmid included strain.

- **Dynamic Flux Balance Analysis** - Determines the upper limit of bioproduct product production (g/L) for the plasmid included strain.
Bioproduct Production

- Bioproduct Identification
- Select Host Strain
- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model
- Strain Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis
Cobra Tools for Analyzing Maximum Production

- Robustness Analysis - Determines range of fluxes for any reaction in the designed strain, including the bioproduct pathway reactions, as a function of growth rate.

- Production Envelopes - Determines the bioproduct production capabilities of a designed strain as a function of growth rate.

- Phenotype Phase Plane Analysis - Determines the phenotype phase planes for the designed strain.

- Dynamic Flux Balance Analysis - Determines the upper limit of bioproduct product production (g/L) over time.
Bioproduct Production

• Bioproduct Identification
• Select Host Strain
• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model
• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
Clustering of Species by Unique Growth-supporting Conditions

Carbon Sources for iJO1366 Model
180/285 are Growth Supporting
**Nitrogen Sources for iJO1366 Model**

94/178 are growth Supporting

<table>
<thead>
<tr>
<th>Source</th>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
<th>Metabolite</th>
</tr>
</thead>
<tbody>
<tr>
<td>15dap</td>
<td>ade</td>
<td>colipap</td>
<td>enlipa</td>
<td>gthox</td>
<td>nac</td>
<td>spmd</td>
<td>xan</td>
</tr>
<tr>
<td>23camp</td>
<td>adn</td>
<td>cpgn</td>
<td>enter</td>
<td>gthrd</td>
<td>nh4</td>
<td>taur</td>
<td>xmp</td>
</tr>
<tr>
<td>23ccmp</td>
<td>adocbl</td>
<td>cpgn-un</td>
<td>etha</td>
<td>gtp</td>
<td>nnn</td>
<td>tcynt</td>
<td>xtsn</td>
</tr>
<tr>
<td>23cgm</td>
<td>agm</td>
<td>crn</td>
<td>fe3dhbs</td>
<td>gua</td>
<td>no</td>
<td>thm</td>
<td></td>
</tr>
<tr>
<td>23cmm</td>
<td>ala-B</td>
<td>crn-D</td>
<td>fe3hox</td>
<td>hocolipa</td>
<td>no2</td>
<td>thr-L</td>
<td></td>
</tr>
<tr>
<td>23dap</td>
<td>ala-D</td>
<td>csn</td>
<td>fe3hox-un</td>
<td>halipa</td>
<td>no3</td>
<td>thrp</td>
<td></td>
</tr>
<tr>
<td>26dapM</td>
<td>ala-L</td>
<td>cyan</td>
<td>fecrm</td>
<td>his-L</td>
<td>novbcn</td>
<td>thym</td>
<td></td>
</tr>
<tr>
<td>3amp</td>
<td>alalaal</td>
<td>cynt</td>
<td>fecrm-un</td>
<td>hom-L</td>
<td>o16a4colipa</td>
<td>thymd</td>
<td></td>
</tr>
<tr>
<td>3cmp</td>
<td>altn</td>
<td>cys-D</td>
<td>feenter</td>
<td>hxn</td>
<td>orn</td>
<td>tma</td>
<td></td>
</tr>
<tr>
<td>3gmp</td>
<td>amp</td>
<td>cys-L</td>
<td>feoxam</td>
<td>ile-L</td>
<td>orot</td>
<td>tmao</td>
<td></td>
</tr>
<tr>
<td>3ump</td>
<td>anhgm</td>
<td>cytd</td>
<td>feoxam-un</td>
<td>imp</td>
<td>peamn</td>
<td>trp-L</td>
<td></td>
</tr>
<tr>
<td>4but</td>
<td>arbtn</td>
<td>dad-2</td>
<td>frulys</td>
<td>indole</td>
<td>phe-L</td>
<td>ttrcyc</td>
<td></td>
</tr>
<tr>
<td>LalaDglu</td>
<td>arbtn-fe3</td>
<td>damp</td>
<td>g3pc</td>
<td>ins</td>
<td>pheme</td>
<td>tym</td>
<td></td>
</tr>
<tr>
<td>LalaDgluMdap</td>
<td>arg-L</td>
<td>dcmp</td>
<td>g3pe</td>
<td>kdo2lipid4</td>
<td>pnto-R</td>
<td>tyr-L</td>
<td></td>
</tr>
<tr>
<td>LalaDgluMdapDala</td>
<td>asn-L</td>
<td>dcyt</td>
<td>g3ps</td>
<td>leu-L</td>
<td>pro-L</td>
<td>tyrp</td>
<td></td>
</tr>
<tr>
<td>LalaLglu</td>
<td>asp-L</td>
<td>dgmp</td>
<td>gam</td>
<td>lipa</td>
<td>progly</td>
<td>uacgam</td>
<td></td>
</tr>
<tr>
<td>acgal</td>
<td>btn</td>
<td>dgsn</td>
<td>gam6p</td>
<td>lipa_cold</td>
<td>psclys</td>
<td>udpca gal</td>
<td></td>
</tr>
<tr>
<td>acgal1p</td>
<td>cbi</td>
<td>dimp</td>
<td>gbbtn</td>
<td>lys-L</td>
<td>pser-L</td>
<td>udpg</td>
<td></td>
</tr>
<tr>
<td>acgam</td>
<td>cbi1</td>
<td>din</td>
<td>gdpe</td>
<td>met-D</td>
<td>ptrc</td>
<td>udpgal</td>
<td></td>
</tr>
<tr>
<td>acgam1p</td>
<td>cgly</td>
<td>dopa</td>
<td>gln-L</td>
<td>met-L</td>
<td>pydam</td>
<td>udpgcur</td>
<td></td>
</tr>
<tr>
<td>acnum</td>
<td>chol</td>
<td>doxrcbn</td>
<td>glu-L</td>
<td>metsox-R</td>
<td>pydx</td>
<td>ump</td>
<td></td>
</tr>
<tr>
<td>acnum</td>
<td>chhts</td>
<td>dtmp</td>
<td>gly</td>
<td>metsox-S</td>
<td>pydxn</td>
<td>ura</td>
<td></td>
</tr>
<tr>
<td>acnam</td>
<td>cm</td>
<td>dump</td>
<td>glyb</td>
<td>mincyc</td>
<td>rfpamp</td>
<td>urea</td>
<td></td>
</tr>
<tr>
<td>accola</td>
<td>cmp</td>
<td>dur</td>
<td>gmp</td>
<td>mmet</td>
<td>ser-D</td>
<td>uri</td>
<td></td>
</tr>
<tr>
<td>acser</td>
<td>colip</td>
<td>eca4colipa</td>
<td>gsn</td>
<td>n2o</td>
<td>ser-L</td>
<td>val-L</td>
<td></td>
</tr>
</tbody>
</table>

### Phosphorus & Sulfur Sources for iJO1366 Model

<table>
<thead>
<tr>
<th>Phosphorus Sources</th>
<th>Sulfur Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>acgal1p[e]</td>
<td>5mtr[e]</td>
</tr>
<tr>
<td>acgam1p[e]</td>
<td>tcynt[e]</td>
</tr>
<tr>
<td>acolipa[e]</td>
<td>btn[e]</td>
</tr>
<tr>
<td>adoebi[e]</td>
<td>thm[e]</td>
</tr>
<tr>
<td>amp[e]</td>
<td>butso3[e]</td>
</tr>
<tr>
<td>cbl1[e]</td>
<td>tsul[e]</td>
</tr>
<tr>
<td>cmp[e]</td>
<td>cog[e]</td>
</tr>
<tr>
<td>colipa[e]</td>
<td>cysts-D[e]</td>
</tr>
<tr>
<td>colipap[e]</td>
<td>cysts-L[e]</td>
</tr>
<tr>
<td>damp[e]</td>
<td>dms[e]</td>
</tr>
<tr>
<td>dcmp[e]</td>
<td>dms[e]</td>
</tr>
<tr>
<td>dgmp[e]</td>
<td>ethso3[e]</td>
</tr>
<tr>
<td>dmp[e]</td>
<td>gthox[e]</td>
</tr>
<tr>
<td>dtmp[e]</td>
<td>gthrd[e]</td>
</tr>
<tr>
<td>dump[e]</td>
<td>h2s[e]</td>
</tr>
<tr>
<td>eca4colipa[e]</td>
<td>isetac[e]</td>
</tr>
<tr>
<td>enlipa[e]</td>
<td>lipoate[e]</td>
</tr>
<tr>
<td>f6p[e]</td>
<td>met-D[e]</td>
</tr>
<tr>
<td>g1p[e]</td>
<td>met-L[e]</td>
</tr>
<tr>
<td>g3pc[e]</td>
<td>metsox-R-L[e]</td>
</tr>
<tr>
<td>g3pe[e]</td>
<td>metsox-S-L[e]</td>
</tr>
<tr>
<td>g3pg[e]</td>
<td>mmet[e]</td>
</tr>
<tr>
<td>g3pi[e]</td>
<td>mso3[e]</td>
</tr>
<tr>
<td>g3ps[e]</td>
<td>so2[e]</td>
</tr>
<tr>
<td>g36p[e]</td>
<td>so3[e]</td>
</tr>
<tr>
<td>g37p[e]</td>
<td>so4[e]</td>
</tr>
<tr>
<td>g6p[e]</td>
<td>sulfac[e]</td>
</tr>
<tr>
<td>g66p[e]</td>
<td>taur[e]</td>
</tr>
<tr>
<td>gam6p[e]</td>
<td></td>
</tr>
<tr>
<td>udpg[e]</td>
<td></td>
</tr>
<tr>
<td>udpgal[e]</td>
<td></td>
</tr>
<tr>
<td>udpgcur[e]</td>
<td></td>
</tr>
<tr>
<td>ump[e]</td>
<td></td>
</tr>
<tr>
<td>xmp[e]</td>
<td></td>
</tr>
</tbody>
</table>

Catabolite Repression Example

The relationship between the glucose uptake rate (mmol gDW^{-1} hr^{-1}) (bottom axes, the dependent variable) and the resulting 1) growth rate (hr^{-1}) (left axes) and 2) oxygen uptake rate (mmol gDW^{-1} hr^{-1}) (right axes) produced during the sensitivity analysis using iAF1260. Using FBA and iAF1260, optimal growth was simulated under glucose aerobic conditions while varying which biomass objective function (BOF) was used along with the number of reactions available to the network due to transcriptional regulation. Two different BOFs were used, a core biomass objective function and wild-type biomass objective function, and regulation was imposed by not allowing any flux through reactions unavailable to the network due to transcriptional regulation (Covert et al, 2004). The results show that the predicted optimal growth rate and O_2 uptake rate are insensitive to the BOF used or level of transcriptional regulation imposed under these conditions.

### Catabolite Repression

- If the carbon source is glucose in an aerobic environment a more accurate model would block the reactions that would naturally be turned off as a result of catabolite repression.

- For the iaf1260 model the reactions blocked are given to the right.

<table>
<thead>
<tr>
<th>iaf1260 reactions block due to catabolite repression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>14GLUCANabcpp</strong></td>
</tr>
<tr>
<td><strong>14GLUCANtexi</strong></td>
</tr>
<tr>
<td><strong>ACOAD1f</strong></td>
</tr>
<tr>
<td><strong>ACOAD2f</strong></td>
</tr>
<tr>
<td><strong>ACOAD3f</strong></td>
</tr>
<tr>
<td><strong>ACOAD4f</strong></td>
</tr>
<tr>
<td><strong>ACOAD5f</strong></td>
</tr>
<tr>
<td><strong>ACOAD6f</strong></td>
</tr>
<tr>
<td><strong>ACOAD7f</strong></td>
</tr>
<tr>
<td><strong>ACOAD8f</strong></td>
</tr>
<tr>
<td><strong>ACS</strong></td>
</tr>
<tr>
<td><strong>ALDD2y</strong></td>
</tr>
<tr>
<td><strong>ALDD3y</strong></td>
</tr>
<tr>
<td><strong>ARAI</strong></td>
</tr>
<tr>
<td><strong>ARBabcpp</strong></td>
</tr>
</tbody>
</table>

Cobra Tools for Analyzing Carbon Sources

- Robustness Analysis - Determines range of fluxes for any reaction in the designed strain, including the bioproduct pathway reactions, as a function of growth rate.

- Production Envelopes - Determines the bioproduct production capabilities of a designed strain as a function of growth rate.

- Phenotype Phase Plane Analysis - Determines the phenotype phase planes, shadow prices, and reduced costs for the designed strain.

- Dynamic Flux Balance Analysis - Determines the upper limit of bioproduct product production (g/L) over time.
Bioproduct Production

• Bioproduct Identification

• Select Host Strain

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
    ✓ Undesired By-products
    ✓ Growth Coupling
    ✓ Cofactor Balancing
    ✓ Sampling Analysis
Media Comparison

- **Minimal media** are those that contain the minimum nutrients possible for colony growth, generally without the presence of amino acids, and are often used to grow "wild type" microorganisms. Minimal medium typically contains: a carbon source for bacterial growth, which may be a sugar such as glucose, water, and various salts which may vary among bacteria species and growing conditions; these generally provide essential elements such as magnesium, nitrogen, phosphorus, and sulfur to allow the bacteria to synthesize protein and nucleic acids.

- **LB (Lysogeny Broth)** is the most commonly used medium for culturing *E. coli*. It is easy to make, it has rich nutrient contents and its osmolarity is optimal for growth at early log phase. All these features make it adequate for protein production and compensate for the fact that it is not the best option for achieving high cell density cultures. Despite being a rich broth, cell growth stops at a relatively low density.

- **An undefined medium** (also known as a basal or complex medium) includes a carbon source such as glucose for bacterial growth, water, various salts needed for bacterial growth, a source of amino acids and nitrogen (e.g., beef, yeast extract). This is an undefined medium because the amino acid source contains a variety of compounds with the exact composition being unknown.

- In autoinduction media, a mixture of glucose, lactose, and glycerol is used in an optimized blend. Glucose is the preferred carbon source and is metabolized preferentially during growth, which prevents uptake of lactose until glucose is depleted, usually in mid to late log phase. Consumption of glycerol and lactose follows, the latter being also the inducer of lac-controlled protein expression. In this way, biomass monitoring for timely inducer addition is avoided, as well as culture manipulation.

M9 Minimal Medium

• One liter of M9 medium (Sigma catalog no. 6030) contains:
  ✓ Na₂HPO₄ · 7H₂O (6.8g), KH₂PO₄ (3g), NaCl (0.5g), NH₄Cl (1g), MgSO₄ (2 mM), CaCl₂ (0.1 mM)

• Growth on minimal medium was simulated by maximizing flux through a defined biomass objective function and allowing the uptake of
  ✓ NH₄, SO₄, O₂, and Pᵢ and the free exchange of H⁺, H₂O, and CO₂

• All exchange reaction lower constraints, except the following, should be greater than zero
  ✓ -1000 ≤ NH₄, SO₄, O₂, and Pᵢ ≤ 0
  ✓ -1000 ≤ H⁺, H₂O, and CO₂ ≤ 1000
  ✓ -1000 ≤ Carbon source ≤ 0
  ✓ Use the following commands to change the constraints
    ▪ model = changeRxnBounds(model, ’EX_xxx(e)’, -1000, ’l’)  
    ▪ model = changeRxnBounds(model, ’EX_xxx(e)’, 1000, ’u’)

✓ Verify that no other metabolites are allowed to be uptaken
  ▪ No other metabolites should have a negative lower constraint
  ▪ Check using the “printConstraints(model, -1001, 1)” command
Recipe for M-9 Minimal Media

• 5X M9 basis
  • Na$_2$HPO$_4$.12 H$_2$O 85.7 g
  • KH$_2$PO$_4$ 15.0 g
  • NaCl 2.5 g
  • Dissolve above components in 1000 ml of milli-Q and autoclave
• 5 g (NH$_4$)$_2$SO$_4$ in 15 ml of H$_2$O
• Trace elements
  • 1 g EDTA
  • 29 mg ZnSO$_4$.7H$_2$O
  • 198 mg MnCl$_2$.4H$_2$O
  • 254 mg CoCl$_2$.6H$_2$O
  • 13.4 mg CuCl$_2$
  • 147 mg CaCl$_2$
  • Dissolve in 100 ml of milli-Q and autoclave

• 20% (w/v) glucose: 25 g in 100 ml of milliQ and filter with 0.22 micron filter
• 0.1 M CaCl$_2$.2H$_2$O: 1.47 g in 100 ml milliQ and filter with 0.22 micron filter
• 1M MgSO$_4$.7H$_2$O: 24.65 g in 100 ml milliQ and filter with 0.22 micron filter
• 10 mM FeSO$_4$.7H$_2$O: 140 mg in 50 ml of milliQ (prepare fresh)
• 1% thiamine: 500mg in 10 ml of milliQ (prepare fresh)

• Proportions for 1 liter M-9 media
  - 200 ml of M-9 basis; 3 ml of (NH$_4$)$_2$SO$_4$; 1 ml of CaCl$_2$.2H$_2$O; 1 ml trace elements; 20 ml glucose; 1ml MgSO$_4$.7H$_2$O; 1 ml FeSO$_4$.7H$_2$O
  - 2ml thiamine; 1ml antibiotic (standard conc.)

### M9 Minimal Media

This *in silico* media assumes the cell can uptake all the minerals wanted/needed from the media. It does not allow amino acid uptake.

<table>
<thead>
<tr>
<th>Reaction Abbreviation</th>
<th>Reaction Name</th>
<th>Formula</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>EX_ca2(e)</td>
<td>Calcium exchange</td>
<td>ca2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cl(e)</td>
<td>Chloride exchange</td>
<td>cl[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_co2(e)</td>
<td>CO2 exchange</td>
<td>co2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cobalt2(e)</td>
<td>Co2+ exchange</td>
<td>cobalt2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cu2(e)</td>
<td>Cu2+ exchange</td>
<td>cu2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_fe2(e)</td>
<td>Fe2+ exchange</td>
<td>fe2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_fe3(e)</td>
<td>Fe3+ exchange</td>
<td>fe3[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_h(e)</td>
<td>H+ exchange</td>
<td>h[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_h2o(e)</td>
<td>H2O exchange</td>
<td>h2o[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_k(e)</td>
<td>K+ exchange</td>
<td>k[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_mg2(e)</td>
<td>Mg exchange</td>
<td>mg2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_mn2(e)</td>
<td>Mn2+ exchange</td>
<td>mn2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_mobd(e)</td>
<td>Molybdate exchange</td>
<td>mobd[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_na1(e)</td>
<td>Sodium exchange</td>
<td>na1[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_tungs(e)</td>
<td>tungstate exchange</td>
<td>tungs[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_zn2(e)</td>
<td>Zinc exchange</td>
<td>zn2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_ni2(e)</td>
<td>Ni2+ exchange</td>
<td>ni2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_sel(e)</td>
<td>Selenate exchange</td>
<td>sel[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_slnt(e)</td>
<td>selenite exchange</td>
<td>slnt[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_so4(e)</td>
<td>Sulfate exchange</td>
<td>so4[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_nh4(e)</td>
<td>Ammonia exchange</td>
<td>nh4[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_pi(e)</td>
<td>Phosphate exchange</td>
<td>pi[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cbl1(e)</td>
<td>Cob(I)alamin exchange</td>
<td>cbl1[e] &lt;=&gt;</td>
<td>-0.01</td>
<td>1000</td>
</tr>
</tbody>
</table>

---

## Amino Acid Exchange Reactions

<table>
<thead>
<tr>
<th>Rxn name</th>
<th>Rxn description</th>
<th>Formula</th>
<th>LB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>EX_ala_L(e)</td>
<td>L Alanine exchange</td>
<td>ala-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_arg_L(e)</td>
<td>L Arginine exchange</td>
<td>arg-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_asn_L(e)</td>
<td>L Asparagine exchange</td>
<td>asn-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_asp_L(e)</td>
<td>L Aspartate exchange</td>
<td>asp-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cys_L(e)</td>
<td>L Cysteine exchange</td>
<td>cys-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_gln_L(e)</td>
<td>L Glutamine exchange</td>
<td>gln-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_glu_L(e)</td>
<td>L Glutamate exchange</td>
<td>glu-L[e] &lt;=</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_gly(e)</td>
<td>Glycine exchange</td>
<td>gly[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_his_L(e)</td>
<td>L Histidine exchange</td>
<td>his-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_ile_L(e)</td>
<td>L Isoleucine exchange</td>
<td>ile-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_leu_L(e)</td>
<td>L Leucine exchange</td>
<td>leu-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_lyr_L(e)</td>
<td>L Lysine exchange</td>
<td>lys-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_met_L(e)</td>
<td>L Methionine exchange</td>
<td>met-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_phe_L(e)</td>
<td>L Phenylalanine exchange</td>
<td>phe-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_pro_L(e)</td>
<td>L Proline exchange</td>
<td>pro-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_ser_L(e)</td>
<td>L Serine exchange</td>
<td>ser-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_thr_L(e)</td>
<td>L Threonine exchange</td>
<td>thr-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_trp_L(e)</td>
<td>L Tryptophan exchange</td>
<td>trp-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_tyr_L(e)</td>
<td>L Tyrosine exchange</td>
<td>tyr-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>EX_val_L(e)</td>
<td>L Valine exchange</td>
<td>val-L[e]</td>
<td>0</td>
<td>1000</td>
</tr>
</tbody>
</table>

Note: Amino acids are only allowed to be secreted in the basic model (LB = 0). The model can be modified to allow amino acid uptake.

No essential amino acids for *E. coli*
K-12 Undefined Media

- K-12 is an undefined media based on yeast extract and phosphates.
- Growth in K-12 media was simulated by adjusting lower bounds of exchange reactions to correspond to media conditions.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>K12 Medium:</td>
<td></td>
</tr>
<tr>
<td>KH₂PO₄</td>
<td>2 g/L</td>
</tr>
<tr>
<td>K₂HPO₄·3H₂O</td>
<td>4 g/L</td>
</tr>
<tr>
<td>(NH₄)₂HPO₄</td>
<td>5 g/L</td>
</tr>
<tr>
<td>Yeast Extract</td>
<td>5 g/L</td>
</tr>
<tr>
<td>Glucose</td>
<td>25 g/L</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>0.5 g/L</td>
</tr>
<tr>
<td>Thiamine</td>
<td>2.5 mg/L</td>
</tr>
<tr>
<td>K12 trace metal</td>
<td>5 ml/L</td>
</tr>
<tr>
<td>K12 trace metal solution:</td>
<td></td>
</tr>
<tr>
<td>NaCl</td>
<td>5 g/L</td>
</tr>
<tr>
<td>ZnSO₄·7H₂O</td>
<td>1 g/L</td>
</tr>
<tr>
<td>MnCl₂·4H₂O</td>
<td>4 g/L</td>
</tr>
<tr>
<td>FeCl₃·6H₂O</td>
<td>4.75 g/L</td>
</tr>
<tr>
<td>CuSO₄·5H₂O</td>
<td>0.4 g/L</td>
</tr>
<tr>
<td>H₃BO₃</td>
<td>0.575 g/L</td>
</tr>
<tr>
<td>NaMoO₄·2H₂O</td>
<td>0.5 g/L</td>
</tr>
<tr>
<td>6N H₂SO₄</td>
<td>12.5 ml/L</td>
</tr>
</tbody>
</table>
## Assigned Metabolite Uptake Rates for K-12 Media

(Metabolite lower bound determined by initial concentration of metabolite times the ratio of the initial glucose concentration/lower bound)

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>MW (g/mol)</th>
<th>g/L in media</th>
<th>mmol/L in media</th>
<th>Lower bound (mmol gDW⁻¹ h⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>180.16</td>
<td>25</td>
<td>138.7655417</td>
<td>-11</td>
</tr>
<tr>
<td>Ammonium</td>
<td>18.03851</td>
<td>1.365829484</td>
<td>75.71742258</td>
<td>-6.002150375</td>
</tr>
<tr>
<td>Phosphate</td>
<td>94.9714</td>
<td>6.655736264</td>
<td>70.08147994</td>
<td>-5.55386948</td>
</tr>
<tr>
<td>Potassium</td>
<td>39.0983</td>
<td>1.945099309</td>
<td>49.74894838</td>
<td>-3.943619038</td>
</tr>
<tr>
<td>Sulfate</td>
<td>96.07</td>
<td>0.203323529</td>
<td>2.11641021</td>
<td>-0.16776864</td>
</tr>
<tr>
<td>Sulfate</td>
<td>35.453</td>
<td>0.062050364</td>
<td>1.750214773</td>
<td>-0.138740225</td>
</tr>
<tr>
<td>Copper</td>
<td>63.546</td>
<td>0.00490684</td>
<td>0.08786535</td>
<td>-0.00696512</td>
</tr>
<tr>
<td>Iron (III)</td>
<td>55.845</td>
<td>0.009040203</td>
<td>1.62856215</td>
<td>-0.16084933</td>
</tr>
<tr>
<td>Magnesium</td>
<td>24.305</td>
<td>0.049304203</td>
<td>2.02856215</td>
<td>-0.16084933</td>
</tr>
<tr>
<td>Magnesium</td>
<td>54.938044</td>
<td>0.005551956</td>
<td>0.10105848</td>
<td>-0.00801094</td>
</tr>
<tr>
<td>Molybdate</td>
<td>95.95</td>
<td>0.01826052</td>
<td>0.19031286</td>
<td>-0.00150861</td>
</tr>
<tr>
<td>Sodium</td>
<td>22.9879628</td>
<td>0.029775544</td>
<td>1.29516497</td>
<td>-0.102668246</td>
</tr>
<tr>
<td>Thiamine</td>
<td>265.35</td>
<td>0.0025</td>
<td>0.00942159</td>
<td>-0.000746848</td>
</tr>
<tr>
<td>Zinc</td>
<td>65.38</td>
<td>0.011363847</td>
<td>0.01738830</td>
<td>-0.001378378</td>
</tr>
<tr>
<td>Alanine</td>
<td>89.09</td>
<td>0.225</td>
<td>2.52553597</td>
<td>-0.200200247</td>
</tr>
<tr>
<td>Arginine</td>
<td>174.2</td>
<td>0.0145</td>
<td>0.83237657</td>
<td>-0.065982824</td>
</tr>
<tr>
<td>Asparagine</td>
<td>132.12</td>
<td>0.16</td>
<td>1.211020285</td>
<td>-0.095998062</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>MW (g/mol)</th>
<th>g/L in media</th>
<th>mmol/L in media</th>
<th>Lower bound (mmol gDW⁻¹ h⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartic acid</td>
<td>133.1</td>
<td>0.16</td>
<td>1.202103681</td>
<td>-0.09529124</td>
</tr>
<tr>
<td>Cysteine</td>
<td>121.16</td>
<td>0.02</td>
<td>0.165070981</td>
<td>-0.013085243</td>
</tr>
<tr>
<td>Glutamine</td>
<td>146.14</td>
<td>0.345</td>
<td>2.360749966</td>
<td>-0.187137594</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>147.13</td>
<td>0.345</td>
<td>2.34865085</td>
<td>-0.185878939</td>
</tr>
<tr>
<td>Glycine</td>
<td>75.07</td>
<td>0.135</td>
<td>1.798321567</td>
<td>-0.14255367</td>
</tr>
<tr>
<td>Histidine</td>
<td>155.15</td>
<td>0.06</td>
<td>0.387622527</td>
<td>-0.03065649</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>131.17</td>
<td>0.145</td>
<td>1.105435694</td>
<td>-0.08762833</td>
</tr>
<tr>
<td>Leucine</td>
<td>131.17</td>
<td>0.21</td>
<td>1.600975833</td>
<td>-0.126909995</td>
</tr>
<tr>
<td>Lysine</td>
<td>146.19</td>
<td>0.22</td>
<td>1.504890895</td>
<td>-0.119293303</td>
</tr>
<tr>
<td>Methionine</td>
<td>149.21</td>
<td>0.045</td>
<td>0.301588365</td>
<td>-0.02390703</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>165.19</td>
<td>0.12</td>
<td>0.726436225</td>
<td>-0.05758489</td>
</tr>
<tr>
<td>Proline</td>
<td>115.13</td>
<td>0.115</td>
<td>0.998870842</td>
<td>-0.079180891</td>
</tr>
<tr>
<td>Serine</td>
<td>105.09</td>
<td>0.13</td>
<td>1.237034922</td>
<td>-0.098060253</td>
</tr>
<tr>
<td>Threonine</td>
<td>119.12</td>
<td>0.135</td>
<td>1.133310947</td>
<td>-0.08938012</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>181.19</td>
<td>0.09</td>
<td>0.496716154</td>
<td>-0.03974888</td>
</tr>
<tr>
<td>Valine</td>
<td>117.15</td>
<td>0.165</td>
<td>1.408450704</td>
<td>-0.11648451</td>
</tr>
</tbody>
</table>

For more accurate results these uptake rates need to be measured.

Cobra Tools for Analyzing Nutrient & Amino Acid Limitations

- Robustness Analysis - Determines range of fluxes for any reaction, including nutrients and amino acids, in the designed strain, including the bioproduct pathway reactions, as a function of growth rate.

- Production Envelopes - Determines the bioproduct production capabilities of a designed strain as a function of growth rate for different nutrient and amino acids compositions.

- Phenotype Phase Plane Analysis - Determines the phenotype phase planes for the designed strain. The potential impact of different nutrients and amino acids on the designed strain can be determined through shadow prices and reduced costs.

- Dynamic Flux Balance Analysis - Determines the impact of nutrients and amino acids on the upper limit of bioproduct product production (g/L) over time.
Bioproduct Production

- Bioproduct Identification
- Select Host Strain
- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model
- Strain Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis
Cobra Tools for Analyzing and Reducing Undesired Byproducts

- **Multi-Production Envelopes** - Determines the undesired byproducts produced by a designed strain in a defined media.

- **Phenotype Phase Plane Analysis** - Determines the phenotype phase planes for the designed strain. The different planes are the result of phenotypes that could be the result of secreted byproducts.

- **Dynamic Flux Balance Analysis** - Determines the impact of byproducts on the upper limit of bioproduct product production (g/L) since some of the secreted byproducts could become carbon sources for the design strain.

- **Flux Balance Analysis** - Determines value of fluxes for given environmental conditions of the designed strain. These flux values can be used to understand the magnitude of all the secreted byproducts.
Bioproduct Production

• Bioproduct Identification
• Select Host Strain
• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model
• Strain Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis
Cobra Tools for Determining Knockouts

• With growth-coupling, the desired phenotype will show an increase in biomass yield coupled to an increase in the production rate of a desired by-product.

• OptKnock - Though a global search process identifies and subsequently removes metabolic reactions that are capable of uncoupling cellular growth from chemical production.

• GDLS - Using a genetic design local search algorithm to run faster than the global search performed by OptKnock, however, it is not guaranteed to identify the global optima.

• OptGene - An evolutionary programming-based method to determine gene knockout strategies for overproduction of a specific product (requires the Tomlab CPLEX solver).

• RobustKnock - A constraint-based method that predicts gene deletion strategies that lead to the over-production of chemicals of interest, by accounting for the presence of competing pathways in the network. RobustKnock provides more robust predictions than those obtained via current state-of-the-art methods (requires the Tomlab CPLEX solver).
Feist (2010)—Production Potential for Growth-Coupled Bioproducts

- An integrated approach through a systematic model-driven evaluation of the production potential for *E. coli* to produce multiple native products from different representative feedstocks through coupling metabolite production to growth rate.

- Optimal strain designs were based on designs which possess maximum yield, substrate-specific productivity, and strength of growth-coupling for up to 10 reaction eliminations (knockouts).

- The method introduced a new concept of objective function tilting for strain design.

---

Desirable Production Metrics

- All designs had to be growth coupled. Each equation examines a different desirable production phenotype.

- Product yield \((Y_{p/s})\): Maximum amount of product that can be generated per unit of substrate.

\[
Y_{p/s} = \frac{\text{production rate}_{product}}{\text{consumption rate}_{substrate}} \left( \frac{\text{mmol product}}{\text{mmol substrate}} \right)
\]

- Substrate-specific productivity (SSP): Product yield per unit substrate multiplied by the growth rate

\[
SSP = \frac{\text{production rate}_{product} \cdot \text{growth rate}}{\text{consumption rate}_{substrate}} \left( \frac{\text{mmol product}}{\text{mmol substrate} \cdot \text{hr}} \right)
\]

- Strength of growth coupling (SOC): Product yield per unit substrate divided by the slope of the lower edge of the production curve

\[
SOC = \left( \frac{\text{production rate}}{\text{slope}} \right)^2 \left( \frac{\text{mmol product}}{\text{mmol substrate} \cdot \text{hr}} \right)
\]

- The slope in this function is the slope of the line between the point of minimum production rate at maximum growth and the point of maximum growth at zero production on a production envelope plot. When this slope is high, it is possible for a strain to grow at very close to the maximum growth rate with only a small production rate, which is undesirable. Therefore, optimizing for maximum production rate is the same as optimizing for maximum product yield. Maximizing for substrate specific productivity introduces a non-linear objective function, which can be handled by OptGene but not OptKnock. Similarly, the strength of coupling is also a non-linear objective function and can only be handled by OptGene. Additionally, a penalty can be added to the scoring function in OptGene by multiplying the objective function with the following penalty function:

\[
\text{objective}_\text{new} = \text{objective}_\text{original} \ast \text{delPenalty}^{\text{numDels}}
\]

where \(\text{objective}_\text{new}\) is the new score of the objective function, \(\text{objective}_\text{original}\) is the original objective function (e.g., product yield), \(\text{delPenalty}\) is the deletion penalty, and \(\text{numDels}\) is the number of knockout reactions. This penalty ensures that designs with fewer knockouts will be selected over designs with similar phenotypes, but more knockouts. Fewer knockouts are desirable for ease of strain construction.

Constraint-based Metabolic Reconstructions & Analysis

H. Scott Hinton, 2016

Lesson: Bioproduct Production

Strain Design

Computation for strain design
Definition
Analysis of results / filtering

395 total designs
154 homofermentors
26 sub. / target pairs with > 0 yield
20 sub. / target pairs with 3 & 5 KO designs

407 total designs
137 homofermentors
22 sub. / target pairs with > 0 yield
Average 18.5 designs per pair

679 total designs
284 homofermentors
26 sub. / target pairs with > 0 yield
Average 18.9 designs per pair

Deletion vs. Yield analysis
Final designs for each sub. / target
- Max. yield, SSP, and SOC
36 sub. / target pairs with > 0 yield

Problem Formulation: Reduction of Model and Selection of Targeted Reactions

Cobra Model

- Remove dead-ends & minimize flux ranges
- Remove transport reactions and periphery metabolic pathways
- Remove reactions that act on high-carbon containing molecules
- Determine co-sets: only consider one reaction in a correlated set

Target reactions

## Substrate Conditions

<table>
<thead>
<tr>
<th>Carbon substrate(s)</th>
<th>Aerobicity</th>
<th>Wild type growth rate</th>
<th>Target reactions after reduction of scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>Anaerobic</td>
<td>0.459</td>
<td>142</td>
</tr>
<tr>
<td>Xylose</td>
<td>Anaerobic</td>
<td>0.319</td>
<td>141</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Anaerobic</td>
<td>0.119</td>
<td>140</td>
</tr>
<tr>
<td>Glucose</td>
<td>Aerobic</td>
<td>1.276</td>
<td>170</td>
</tr>
<tr>
<td>Xylose</td>
<td>Aerobic</td>
<td>1.131</td>
<td>165</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Aerobic</td>
<td>0.983</td>
<td>166</td>
</tr>
</tbody>
</table>

Maximum uptake rates for primary carbon sources were set to 20 mmol gDW\(^{-1}\) h\(^{-1}\). In aerobic simulations, maximum oxygen uptake rate was set to 20 mmol gDW\(^{-1}\) h\(^{-1}\).

Theoretical maximum production $Y_{p/s}$ (%) analysis.

<table>
<thead>
<tr>
<th>Product</th>
<th>No. of carbons</th>
<th>Anaerobic</th>
<th>Aerobic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Glucose</td>
<td>Xylose</td>
</tr>
<tr>
<td>Ethanol</td>
<td>2</td>
<td>49(^a)</td>
<td>49(^a)</td>
</tr>
<tr>
<td>α-Lactate</td>
<td>3</td>
<td>95(^a)</td>
<td>95(^a)</td>
</tr>
<tr>
<td>Glycerol</td>
<td>3</td>
<td>37</td>
<td>27</td>
</tr>
<tr>
<td>L-Alanine</td>
<td>3</td>
<td>95(^a)</td>
<td>76</td>
</tr>
<tr>
<td>L-Serine</td>
<td>3</td>
<td>47</td>
<td>35</td>
</tr>
<tr>
<td>Pyruvate</td>
<td>3</td>
<td>71</td>
<td>60</td>
</tr>
<tr>
<td>Fumarate</td>
<td>4</td>
<td>54</td>
<td>40</td>
</tr>
<tr>
<td>L-Malate</td>
<td>4</td>
<td>63</td>
<td>46</td>
</tr>
<tr>
<td>Succinate</td>
<td>4</td>
<td>93</td>
<td>81</td>
</tr>
<tr>
<td>2-Oxoglutarate</td>
<td>5</td>
<td>40</td>
<td>32</td>
</tr>
<tr>
<td>L-Glutamate</td>
<td>5</td>
<td>44</td>
<td>36</td>
</tr>
</tbody>
</table>

\(^a\) Indicates anaerobic condition where homofermentation of product is possible (<2wt% other carbon products, CO\(_2\) exempt), all aerobic conditions have homofermentation potential.

The Strain Designs Generated for Five Different Targets from Glucose and Xylose Anaerobically

The different target production rates (mmol gDW\(^{-1}\) hr\(^{-1}\)) are shown on the y-axis and the growth rate (hr\(^{-1}\)) is given on the x-axis.

Uncoupled Bioproducts

• In some cases it is desirable to avoid coupling the desired bioproduct to the growth of the cell.

• Through the use of an unregulated plasmid the bioproduct pathway can use nearly all of the cell's resources to create the desired bioproduct. The result is virtually no cell growth but with large bioproduct production.

• The key enzymes in the desired pathway are included in the plasmid.

• The flux supported by the key enzymes is a function of the plasmids copy number and promoter strength. The more copies of the plasmid and stronger promoter strength lead to larger flux.

Assumes the unregulated HEMEOX enzyme can produce 1.2 mmol/gDW-hr

These values equal the production rate of the HEMOX enzyme

Biliverdin_ProductionEnvelope_iJO1366.m
Maximum Production: Post-induction

E. coli BL21 (DE3) with pJ401-HO, 2012-09-05

Data provided by Dr. Dong Chen at Utah State University
Bioproduct Production

- Bioproduct Identification
- Select Host Strain
- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model
- Strain Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis
Redox Trafficking in the Core Metabolic Pathways: Cofactor View

Cofactor Balance

- Cofactor imbalance - Changes that will cause an imbalance in cofactor utilization that would need to be compensated for elsewhere in the network.

- The primary role of the reduced respiratory cofactor NADH is to transfer electrons to oxygen via the electron transport chain, generating the proton gradient that is used for oxidative phosphorylation of adenosine diphosphate (ADP) to adenosine triphosphate (ATP). Concurrently, the reduced cofactor NADPH donates electrons to anabolic reactions and drives the biosynthetic pathways in the cell.

- Solutions include
  - Altering genes/reaction cofactor specificity to balance cofactors (NADH vs NADPH)
  - Add heterogeneous genes (e.g. ferredoxins) that recycle cofactors

\[
2 \text{ reduced ferredoxin} + \text{ NADP} + H \leftrightarrow 2 \text{ oxidized ferredoxin} + \text{ NADPH}
\]

(a) Engineered cofactor imbalanced L-arabinose and D-xylose pathways, (b) Engineered cofactor balanced L-arabinose and D-xylose pathways. XR: Xylose reductase; LAD: L-arabitol dehydrogenase; LXR: L-xylulose reductase; XDH: xylitol dehydrogenase; XK: xylulokinase.


Cobra Tools for Balancing Cofactors

- Robustness Analysis - Used to determine pathway reaction improvements resulting from cofactor balancing.
- Production Envelopes - Determines the bioproduct production capabilities from the cofactor balanced strain.
- Phenotype Phase Plane Analysis - Determines the phenotype phase planes for the cofactor balanced strain.
- Dynamic Flux Balance Analysis - Determines the upper limit of bioproduct product production (g/L) over time for the cofactor balanced strain.
- OptSwap - Predicts bioprocessing strain designs by identifying optimal modifications of the cofactor binding specificities of oxidoreductase enzymes and complementary reaction knockouts (requires the Tomlab CPLEX solver)
OptSwap

- Central oxidoreductase enzymes (e.g., dehydrogenases, reductases) in microbial metabolism often have preferential binding specificity for one of the two major currency metabolites NAD(H) and NADP(H). These enzyme specificities result in a division of the metabolic functionality of the currency metabolites: enzymes reducing NAD+ to NADH drive oxidative phosphorylation, and enzymes reducing NADP+ to NADPH drive anabolic reactions.

- OptSwap predicts bioprocessing strain designs by identifying optimal modifications of the cofactor binding specificities of oxidoreductase enzymes and complementary reaction knockouts.

- Independent of the Cobra Toolbox and requires the Tomlab CPLEX solver

Z. A. King and A. M. Feist, Optimizing cofactor specificity of oxidoreductase enzymes for the generation of microbial production strains—OptSwap. Ind. Biotechnol. 9, 236-246
Bioproduct Production

- Bioproduct Identification
- Select Host Strain
- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model
- Strain Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis
Cobra Tools for Sampling Analysis

- gpSampler - Samples the solution space. Identifies the distribution of flux activity possible through a reaction for a given solution space.

- plotSampleHist - Compare flux histograms for one or more samples for one or more reactions

- sampleScatterMatrix - A scatter plot allows the visualization of the interaction between two network reactions. It can also display correlation coefficients associated with two reactions.

- identifyingCorrelSets - Two reactions are part of the same "correlated reaction set" if their fluxes are linearly correlated. If a knockout reaction has been identified then any of the correlated reaction set can be removed with the same impact.
Bioproduct Production

- Bioproduct Identification
- Select Host Strain
- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model
- Strain Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis
Biliverdin Production
A Continually Evolving Process

Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Biliverdin

Potential Biliverdin protective effects:

- endotoxic shock
- brain infarction
- vascular intimal hyperplasia
- vascular endothelial dysfunction
- lung injury
- airway inflammation
- ischemia reperfusion injury
- tissue graft rejection
- corneal epithelial injury
- hepatitis C infection
- arteriosclerosis
- colitis
- cancer cell proliferation
- type 2 diabetes
- type 1 diabetes (pancreatic Islet cell protection)

Jon Takemoto, USU SBI Bioproducts Summit, February 2012
Porphyrin and chlorophyll metabolism - Synechocystis sp. PCC 6803

Biliverdin Gene

Dong Chen, USU SBI Bioproducts Summit, February 2012
Biliverdin Production

- Bioproduct Identification

- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model

- Host Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis

- Final Design
Location of the Heme Biosynthesis Pathway in iJO1366 Model

Biliverdin Pathway in iJO1366 Model

Constraint-based Metabolic Reconstructions & Analysis

H. Scott Hinton, 2016

Utah State University

Lesson: Bioproduct Production

Biliverdin Pathway in iJO1366 Model

Heme Biosynthesis

Biliverdin Pathway

Biliverdin Pathway in iJO1366 Model
1. The HEMOX reaction was added to the cell with no regulatory constraints so it will automatically produce biliverdin based on the copy number and promoter strength of the HEMOX enzymes created by the plasmid.

2. To use the model, set an estimated lower limit for the HEMOX reaction.
Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Adding Biliverdin Pathway to the Model

% Input the E.coli core model
• model=readCbModel('iECD_1391');

% Add heme oxygenase enzyme

% Add biliverdin secretion reactions
model = addReaction(model, 'BILIVERDINtex', 'biliverdin[p] <=> biliverdin[e]');
model = addReaction(model, 'BILIVERDINtpp', 'biliverdin[c] <=> biliverdin[p]');
model = addReaction(model, 'EX_biliverdin(e)', 'biliverdin[e] <=>');

metID=findMetIDs(model, ['biliverdin[c]', 'biliverdin[p]', 'biliverdin[e]']);
model.metCharge(metID(1)) = -2;
model.metCharge(metID(2)) = -2;
model.metCharge(metID(3)) = -2;

% Add carbon monoxide secretion reactions
model = addReaction(model, 'COtpp', 'co[c] <=> co[p]');
model = addReaction(model, 'COtex', 'co[p] <=> co[e]');
model = addReaction(model, 'EX_co(e)', 'co[e] <=>');

Default Reaction Settings:
UB = 1000
LB = -1000

There also needs to be a reaction included for the HEMEOX reaction to set the flux rate based on the plasmid properties.
Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  → ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Components Produced by Plasmid

- **Lac Repressor** (154 kDa)
- **Nucleotide precursors**
- **Heme Oxygenase** (27 kDa)
- **Biliverdin** (582.646 Da)
- **Beta Lactamase** (29 kDa)
Plasmid Production Load

- Nucleotide precursors
  - The plasmid requires four nucleotide precursors (dGTP, dCTP, dATP, dTTP) for its replication based on an approximated E. coli GC content of 50.6%.

- Total antibiotic marker production
  - Can be up to 3% of the total protein mass

- Recombinant protein production
  - With high copy, strong promoter vectors the recombinant protein concentration can be as high as 50% of the total protein mass

- Lac repressor production
  - There must be a low concentration of repressors or gene expression could not be induced
  - Typically less than 0.1% of the total protein; should be able to ignored
Biliverdin Plasmid Reactions

- pET101-HO1 is composed of 6476 bp, assuming an E. coli GC content of 50.6%:
  - dGTP coefficient = \(6476 \times 0.506 = 3276.86\)
  - dCTP coefficient = \(6476 \times 0.506 = 3276.86\)
  - dATP coefficient = \(6476 \times 0.494 = 3199.14\)
  - dTTP coefficient = \(6476 \times 0.494 = 3199.14\)

- The reaction to produce the antibiotic marker protein beta-lactamase (ampicillin) includes amino acids plus the energetic requirements of 4.306 mol ATP/mol amino acid:

- The reaction to produce heme oxygenase enzyme:

- Each of the above reactions will require a demand function since these proteins needs to have a constant production.

Biliverdin Plasmid Enzyme Production

- Nucleotide precursors
  - Has an experimentally measured flux rate of 0.63 pmol/gDW·h, a good approximation but will vary with plasmids and hosts
- The antibiotic marker beta-lactamase for ampicillin (29,000 Da or mg/mmol) to produce the equivalent of 3% every hour of the total cellular protein (550 mg/gDW)
  - Marker concentration = \( \frac{0.03 \times 550 \text{ mg/gDW}}{29,000 \text{ mg/mmol}} \) = 0.000569 mmol/gDW
- For heme oxygenase (27.051 kDa) to produce the equivalent of 50% of the total cellular protein every hour
  - HO concentration = \( \frac{0.50 \times 550 \text{ mg/gDW}}{27,051 \text{ mg/mmol}} \) = 0.0102 mmol/gDW
- For the Lacl repressor (154.510 KDa) to produce the equivalent of 1% of the total cellular protein every hour
  - Repressor concentration = \( \frac{0.01 \times 550 \text{ mg/gDW}}{154,510 \text{ mg/mmol}} \) = 0.0000356 mmol/gDW - Small enough to ignore

Biliverdin Plasmid Cobra Code

% Add plasmid nucleotide precursors
model = changeRxnBounds(model,'PLASMID',0.63e-9,'l');
% Add demand reaction for plasmid DNA and set flux rate
model = addDemandReaction(model,'plasmid[c]');
% Add beta-lactamase protein
model = changeRxnBounds(model,'b-lactamase',0.000569,'l');
% Add demand reaction for beta-lactamase
model = addDemandReaction(model,'b-lactamase[c]');
% Add heme oxygenase protein
model = changeRxnBounds(model,'HOprotein',0.0102,'l');
% Add demand reaction for heme oxygenase
model = addDemandReaction(model,'HO[c]');
Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

➤ • Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Maximum bioproduct production is based on the cell's internal pathways plus the carbon sources, nutrients, etc. that can be accessed by the cell through the exchange reactions. The default minimal media assumes the following:

- Glucose is the only carbon source \((\text{EX}_{\text{glc}}(e) = -10)\)
  - Other carbon sources can be used
- The cell is in an aerobic condition
  - Unconstrained oxygen can enter/exit the cell
- No amino acids are allowed in enter the cell from the extracellular environment
  - All are allowed to transport out
- All minerals are allowed to enter/exit the cell unconstrained

The values of these exchanges between the extracellular space and the cell plus the internal pathways of the cell determine the maximum possible bioproduct production.
Plasmid Impact on Maximum Growth Rate and Biliverdin Production

No Plasmid

Fully Loaded Plasmid
Biliverdin Production

No Plasmid, Initial Biomass Concentration = 0.01
HEMEOX ≥ 0.01, EX_glc(e) ≥ -10, EX_o2(e) ≥ -20

Biliverdin_ProductionEnvelope_BL21.m
Biliverdin_Production_GrowthRate_PPP_BL21.m
DynamicBiliverdinProduction_Cont_BL21.m
Biliverdin Production

No Plasmid, Initial Biomass Concentration = 0.01
HEMEOX ≥ 1.0, EX_glc(e) ≥ -10, EX_o2(e) ≥ -20

Biliverdin_ProductionEnvelope_BL21.m
Biliverdin_Production_GrowthRate_PPP_BL21.m
DynamicBiliverdinProduction_Cont_BL21.m
Biliverdin Production

Plasmid, Initial Biomass Concentration = 0.01
HEMEOX ≥ 0.01, EX_glc(e) ≥ -10, EX_o2(e) ≥ -20
Biliverdin Production

Plasmid, Initial Biomass Concentration = 0.01
HEMEOX ≥ 0.8, \(EX_{\text{glc}}(e) \geq -10\), \(EX_{\text{o2}}(e) \geq -20\)

![Biliverdin Production Envelope with Plasmid](image)

![Biliverdin Production Growth Rate](image)

![Biomass Concentration vs Time](image)

![Substrate Concentrations](image)

Biliverdin_ProductionEnvelope_BL21.m  
Biliverdin_Production_GrowthRate_PPP_BL21_Plasmid.m  
DynamicBiliverdinProduction_Cont_BL21_Plasmid.m
Impact of HEMEOX Flux and Plasmid on Total Biliverdin Production

Initial Biomass Concentration = 0.01 g/L

DynamicBiliverdinProduction_Cont_Plasmid_BL21.m
Optimal HEMEOX Flux for Maximum Biliverdin Production

Based on Results from dynamicFBA, Initial Biomass Concentration = 0.01 g/L

Maximum Biliverdin Production

Production Time = 16 hours

- 1490 g/L
- HEMEOX = 0.12 mmol·gDW$^{-1}·h^{-1}$

Need to set upper limits on biomass concentration

DynamicBiliverdinProduction_Cont_BL21_maxProd.m

Maximum Biliverdin Production with Plasmid

Production Time = 16 hours

- 6.986 g/L
- HEMOX = 0.1 mmol·gDW$^{-1}·h^{-1}$

Plasmid has dramatic impact on biliverdin production

DynamicBiliverdinProduction_Cont_BL21_maxProd_Plasmid.m

- Production Time = 16 hours
- Plasmid has dramatic impact on biliverdin production
Optimal HEMEOX Flux for Maximum Biliverdin Production

Based on Results from dynamicFBA, Initial Biomass Concentration = 0.01 g/L

Maximum Biliverdin Production without Plasmid

- Production Time set at time when the biomass concentration reaches 200 g/L
- Production Time = 16 hours
- HEMOX = 0.5 mmol·gDW\(^{-1}\)·h\(^{-1}\)
- Biliverdin Concentration = 6.986 g/L

Maximum Biliverdin Production with Plasmid

- Production Time set at time when the biomass concentration reaches 200 g/L
- Production Time = 16 hours
- HEMOX = 0.1 mmol·gDW\(^{-1}\)·h\(^{-1}\)
- Biliverdin Concentration = 100.4 g/L

Plasmid has dramatic impact on biliverdin production

DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
Biliverdin Production Maximum
No Plasmid, Constraining Biomass Concentration < 200 g/L

\[ \text{EX}_{\text{glc}}(e) \geq -10; \text{EX}_{\text{o2}}(e) \geq -20; \text{initBiomass} = 0.01 \text{ gDW/L}; \text{timeStep} = 0.25 \text{ h}; \text{nSteps} = 64 \]

Region where biomass concentration is limited to less than 200 g/L

Processing time before the biomass concentration reaches 200 g/L. The processing is stopped before the biomass Concentration reaches 200 g/L

DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
Biliverdin Production with Uncoupled Plasmid

• For an uncoupled plasmid, the copy number and promoter strength determine the HEMEOX flux.

• To increase bioproduct production a biomass seed will grow at a high growth rate until it reaches a specific density (OD600 ≥ 30) and then the plasmid will be induced.

• Once the plasmid is induced the majority of the cells resources will be used to produce the bioproduct at the expense of slow growth rate.

• For converting OD600 measurements to gDW/L assume that when OD600 = 1 -> 0.47 gDW/L. Thus, OD600 = 30 -> 30*0.47 = 14.1 gDW/L
Biliverdin Production Maximum
Plasmid, Ferredoxin-NADPH, Biomass Concentration < 200 g/L

\[
\text{EX}_{\text{glc}}(e) \geq -10; \text{EX}_{\text{o2}}(e) \geq -20; \text{initBiomass} = 14.1 \text{ gDW/L} (\text{OD600} = 30); \ \text{timeStep} = 0.125 \text{ h}; \ n\text{Steps} = 128
\]

DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
Biliverdin Production Maximum
Plasmid, Ferredoxin-NADPH, Biomass Concentration < 200 g/L

EX_{glc}(e) \geq -10; EX_o2(e) \geq -20; \text{initBiomass} = 47 \text{ gDW/L (OD600 = 100)}; \text{timeStep} = 0.125 \text{ h}; n\text{Steps} = 128

DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
Maximum Biliverdin Production: Post-induction

\[ \text{EX}_{\text{glc}}(e) \geq -10; \text{EX}_{\text{o2}}(e) \geq -20; \text{initBiomass} = 14.1 \, \text{gDW/L (OD600 = 30)}; \text{timeStep} = 0.125 \, \text{h}; \text{nSteps} = 128 \]

Data provided by Dr. Dong Chen at Utah State University

DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
**Post-induction Maximum Biliverdin Production**

\[ EX_{\text{glc}}(e) \geq -10; \ EX_{\text{o2}}(e) \geq -20; \text{initBiomass} = 14.1 \text{ gDW/L (OD600 = 30)}; \ \text{timeStep} = 0.125 \text{ h}; \ nSteps = 128 \]

- **Biliverdin Production Envelope with Plasmid**
- **Maximum Biliverdin Production at 200g/L Biomass Limit**

*Annotations for figures:
- Biliverdin Production Envelope with Plasmid
- DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
## Carbon Source Maximum Biliverdin Production, Reduced Costs & Shadow Prices

<table>
<thead>
<tr>
<th>Name</th>
<th>Max Biliverdin Production</th>
<th>Reaction</th>
<th>Reduced Cost</th>
<th>Metabolite</th>
<th>Shadow Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Fructose</td>
<td>1.3211</td>
<td>EX_fru(e)</td>
<td>-0.134182</td>
<td>fru[c]</td>
<td>-0.135162</td>
</tr>
<tr>
<td>D-Glucose</td>
<td>1.3295</td>
<td>EX_glc(e)</td>
<td>-0.134182</td>
<td>glc-D[c]</td>
<td>-0.135162</td>
</tr>
<tr>
<td>Glycerol</td>
<td>0.7439</td>
<td>EX_glyc(e)</td>
<td>-0.076396</td>
<td>glyc[c]</td>
<td>-0.076396</td>
</tr>
<tr>
<td>Lactose</td>
<td>1.3295</td>
<td>EX_lcts(e)</td>
<td>-0.134182</td>
<td>lcts[c]</td>
<td>-0.135162</td>
</tr>
<tr>
<td>Maltose</td>
<td>2.5354</td>
<td>EX_malt(e)</td>
<td>-0.270323</td>
<td>malt[c]</td>
<td>-0.274241</td>
</tr>
<tr>
<td>D-Mannitol</td>
<td>1.3825</td>
<td>EX_mnl(e)</td>
<td>-0.141038</td>
<td>mnl[p]</td>
<td>-0.141038</td>
</tr>
<tr>
<td>D-Mannose</td>
<td>1.3211</td>
<td>EX_man(e)</td>
<td>-0.134182</td>
<td>man[c]</td>
<td>-0.139079</td>
</tr>
<tr>
<td>L-Rhamnose</td>
<td>1.3124</td>
<td>EX_rmn(e)</td>
<td>-0.133203</td>
<td>rmn[c]</td>
<td>-0.134182</td>
</tr>
<tr>
<td>D-Ribose</td>
<td>1.0677</td>
<td>EX_rib-D(e)</td>
<td>-0.112635</td>
<td>rib-D[c]</td>
<td>-0.116552</td>
</tr>
<tr>
<td>D-xylene</td>
<td>1.0983</td>
<td>EX_xyl-D(e)</td>
<td>-0.111655</td>
<td>xyl-D[c]</td>
<td>-0.112635</td>
</tr>
</tbody>
</table>

Biliverdin_Carbon_Source_ShadowPrices_BL21.m & Biliverdin_Carbon_Source_ReducedCosts_BL21.m & Biliverdin_Carbon_Source_BL21.m
Maximum Biliverdin Production by Carbon Source

EX_Carbon_Source = 20 mmol/gDW-h, EX_o2(e) = 20 mmol/gDW-h

Glucose  Glycerol  Mannitol

Biliverdin_Production_PPP_BL21.m
Biliverdin Production Phenotype Phase Planes
- Glucose With Plasmid -

EX_biliverdin(e) = 0.01

EX_biliverdin(e) = 0.1

EX_biliverdin(e) = 1.0
Biliverdin Production Phenotype Phase Planes
- Glycerol With Plasmid -

EX_biliverdin(e) = 0.01

EX_biliverdin(e) = 0.1

EX_biliverdin(e) = 0.8
Biliverdin Production

• Bioprocess Identification

• Bioprocess Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioprocess Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
### Constraint-based Metabolic Reconstructions & Analysis

**Lesson:** Bioproduct Production  
**BIE 5500/6500**  
**Utah State University**  
**H. Scott Hinton, 2016**

#### M9 Minimal Media

This *in silico* media assumes the cell can uptake all the minerals wanted/needed from the media. It does not allow amino acid uptake.

---


<table>
<thead>
<tr>
<th>Reaction Abbreviation</th>
<th>Reaction Name</th>
<th>Formula</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>EX_ca2(e)</td>
<td>Calcium exchange</td>
<td>ca2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cl(e)</td>
<td>Chloride exchange</td>
<td>cl[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_co2(e)</td>
<td>CO2 exchange</td>
<td>co2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cobalt2(e)</td>
<td>Co2+ exchange</td>
<td>cobalt2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cu2(e)</td>
<td>Cu2+ exchange</td>
<td>cu2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_fe2(e)</td>
<td>Fe2+ exchange</td>
<td>fe2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_fe3(e)</td>
<td>Fe3+ exchange</td>
<td>fe3[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_h(e)</td>
<td>H+ exchange</td>
<td>h[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_h2o(e)</td>
<td>H2O exchange</td>
<td>h2o[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_k(e)</td>
<td>K+ exchange</td>
<td>k[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_mg2(e)</td>
<td>Mg exchange</td>
<td>mg2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_mn2(e)</td>
<td>Mn2+ exchange</td>
<td>mn2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_mobd(e)</td>
<td>Molybdate exchange</td>
<td>mobd[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_na1(e)</td>
<td>Sodium exchange</td>
<td>na1[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_tungs(e)</td>
<td>Tungstate exchange</td>
<td>tungs[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_zn2(e)</td>
<td>Zinc exchange</td>
<td>zn2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_ni2(e)</td>
<td>Ni2+ exchange</td>
<td>ni2[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_sel(e)</td>
<td>Selenate exchange</td>
<td>sel[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_slnt(e)</td>
<td>Selenite exchange</td>
<td>slnt[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_so4(e)</td>
<td>Sulfate exchange</td>
<td>so4[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_nh4(e)</td>
<td>Ammonia exchange</td>
<td>nh4[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_pi(e)</td>
<td>Phosphate exchange</td>
<td>pi[e] &lt;=&gt;</td>
<td>-1000</td>
<td>1000</td>
</tr>
<tr>
<td>EX_cbl1(e)</td>
<td>Cob(I)alamin exchange</td>
<td>cbl1[e] &lt;= &gt;</td>
<td>-0.01</td>
<td>1000</td>
</tr>
</tbody>
</table>
Initial K-12 Mineral & Trace Metal Concentrations

<table>
<thead>
<tr>
<th>Chemical</th>
<th>MW (g/mol)</th>
<th>g/L in media</th>
<th>mmol/L in media</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>180.16</td>
<td>25</td>
<td>138.7655417</td>
</tr>
<tr>
<td>Ammonium</td>
<td>18.03851</td>
<td>1.365829484</td>
<td>75.71742258</td>
</tr>
<tr>
<td>Phosphate</td>
<td>94.9714</td>
<td>6.655736264</td>
<td>70.08147994</td>
</tr>
<tr>
<td>Potassium</td>
<td>39.0983</td>
<td>1.945099309</td>
<td>49.74894838</td>
</tr>
<tr>
<td>Sulfate</td>
<td>96.07</td>
<td>0.203323529</td>
<td>2.11641021</td>
</tr>
<tr>
<td>Chloride</td>
<td>35.453</td>
<td>0.062050364</td>
<td>1.750214773</td>
</tr>
<tr>
<td>Copper</td>
<td>63.546</td>
<td>0.000509009</td>
<td>0.008010093</td>
</tr>
<tr>
<td>Iron (III)</td>
<td>55.845</td>
<td>0.00490684</td>
<td>0.08786535</td>
</tr>
<tr>
<td>Magnesium</td>
<td>24.305</td>
<td>0.049304203</td>
<td>2.028562155</td>
</tr>
<tr>
<td>Manganese</td>
<td>54.938044</td>
<td>0.005551956</td>
<td>0.101058478</td>
</tr>
<tr>
<td>Molybdate</td>
<td>95.95</td>
<td>0.001826052</td>
<td>0.019031286</td>
</tr>
<tr>
<td>Sodium</td>
<td>22.9827976928</td>
<td>0.029775544</td>
<td>1.295164976</td>
</tr>
<tr>
<td>Thiamine</td>
<td>265.35</td>
<td>0.0025</td>
<td>0.009421519</td>
</tr>
<tr>
<td>Zinc</td>
<td>65.38</td>
<td>0.001136847</td>
<td>0.017388304</td>
</tr>
<tr>
<td>Alanine</td>
<td>89.09</td>
<td>0.225</td>
<td>2.525535975</td>
</tr>
<tr>
<td>Arginine</td>
<td>174.2</td>
<td>0.145</td>
<td>0.832376579</td>
</tr>
<tr>
<td>Asparagine</td>
<td>132.12</td>
<td>0.16</td>
<td>1.211020285</td>
</tr>
</tbody>
</table>

K12 Medium:
- KH₂PO₄: 2 g/L
- K₂HPO₄·3H₂O: 4 g/L
- (NH₄)₂HPO₄: 5 g/L
- Yeast Extract: 5 g/L
- Glucose: 25 g/L
- MgSO₄·7H₂O: 0.5 g/L
- Thiamine: 2.5 mg/L
- K12 trace metal: 5 ml/L

K12 trace metal solution:
- NaCl: 5 g/L
- ZnSO₄·7H₂O: 1 g/L
- MnCl₂·4H₂O: 4 g/L
- FeCl₃·6H₂O: 4.75 g/L
- CuSO₄·5H₂O: 0.4 g/L
- H₃BO₃: 0.575 g/L
- NaMoO₄·2H₂O: 0.5 g/L
- 6N H₂SO₄: 12.5 ml/L

Impact of Essential Nutrients in the Media on Cell Growth

How are the essential nutrients impact the production process? They include: 'EX_cl(e)', 'EX_cu2(e)', 'EX_fe3(e)', 'EX_k(e)', 'EX_mg2(e)', 'EX_sn2(e)', 'EX_snobd(e)', 'EX_ba1(e)', 'EX_nh4(e)', 'EX_pi(e)', 'EX_so4(e)', 'EX_zn2(e)'

[Sulfate & Ammonium Limited]
Identify the shadow prices for both minerals and amino acids when biliverdin production is optimized.

- Glutamate is a precursor to heme production
- Some of the amino acids can act as a carbon source.

### Biliverdin Mineral & Amino Acid Shadow Prices

<table>
<thead>
<tr>
<th>Exchange Reactions</th>
<th>Metabolite</th>
<th>Shadow Price</th>
<th>Exchange Reactions</th>
<th>Metabolite</th>
<th>Shadow Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>EX_ca2(e)</td>
<td>ca2[e]</td>
<td>0</td>
<td>EX_alan_L(e)</td>
<td>alan-L[e]</td>
<td>-0.06464</td>
</tr>
<tr>
<td>EX_cl(e)</td>
<td>cl[e]</td>
<td>0</td>
<td>EX_argin_L(e)*</td>
<td>arg-L[e]</td>
<td>-0.12341</td>
</tr>
<tr>
<td>EX_co2(e)</td>
<td>co2[e]</td>
<td>0</td>
<td>EX_asn_L(e)</td>
<td>asn-L[e]</td>
<td>-0.07052</td>
</tr>
<tr>
<td>EX_cobalt2(e)</td>
<td>cobalt2[e]</td>
<td>0</td>
<td>EX_asp_L(e)</td>
<td>asp-L[e]</td>
<td>-0.07052</td>
</tr>
<tr>
<td>EX_cu2(e)</td>
<td>cu2[e]</td>
<td>0</td>
<td>EX_cys_L(e)</td>
<td>cys-L[e]</td>
<td>-0.05289</td>
</tr>
<tr>
<td>EX_fe2(e)</td>
<td>fe2[e]</td>
<td>0</td>
<td>EX_gln_L(e)*</td>
<td>gln-L[e]</td>
<td>-0.11166</td>
</tr>
<tr>
<td>EX_fe3(e)</td>
<td>fe3[e]</td>
<td>0</td>
<td>EX_glu_L(e)*</td>
<td>glu-L[e]</td>
<td>-0.11068</td>
</tr>
<tr>
<td>EX_h(e)</td>
<td>h[e]</td>
<td>0</td>
<td>EX_gly(e)</td>
<td>gly[e]</td>
<td>-0.03624</td>
</tr>
<tr>
<td>EX_h2o(e)</td>
<td>h2o[e]</td>
<td>0</td>
<td>EX_his_L(e)</td>
<td>his-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_k(e)</td>
<td>k[e]</td>
<td>0</td>
<td>EX_ile_L(e)</td>
<td>ile-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_mg2(e)</td>
<td>mg2[e]</td>
<td>0</td>
<td>EX_leu_L(e)</td>
<td>leu-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_mn2(e)</td>
<td>mn2[e]</td>
<td>0</td>
<td>EX_lysin_L(e)</td>
<td>lys-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_mobd(e)</td>
<td>mobd[e]</td>
<td>0</td>
<td>EX_met_L(e)</td>
<td>met-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_na1(e)</td>
<td>na1[e]</td>
<td>0</td>
<td>EX_phe_L(e)</td>
<td>phe-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_tungs(e)</td>
<td>tungs[e]</td>
<td>0</td>
<td>EX_pro_L(e)</td>
<td>pro-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_zn2(e)</td>
<td>zn2[e]</td>
<td>0</td>
<td>EX_ser_L(e)</td>
<td>ser-L[e]</td>
<td>-0.1097</td>
</tr>
<tr>
<td>EX_ni2(e)</td>
<td>ni2[e]</td>
<td>0</td>
<td>EX_threon_L(e)</td>
<td>thr-L[e]</td>
<td>-0.05583</td>
</tr>
<tr>
<td>EX_sel(e)</td>
<td>sel[e]</td>
<td>0</td>
<td>EX_trp_L(e)</td>
<td>trp-L[e]</td>
<td>-0.09207</td>
</tr>
<tr>
<td>EX_sint(e)</td>
<td>sint[e]</td>
<td>0</td>
<td>EX_tyr_L(e)</td>
<td>tyr-L[e]</td>
<td>-0.05681</td>
</tr>
<tr>
<td>EX_so4(e)</td>
<td>so4[e]</td>
<td>0</td>
<td>EX_val_L(e)</td>
<td>val-L[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_nh4(e)</td>
<td>nh4[e]</td>
<td>0</td>
<td>EX_phen(e)</td>
<td>phen[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_pi(e)</td>
<td>pi[e]</td>
<td>0</td>
<td>EX_pheph(e)</td>
<td>phenph[e]</td>
<td>0</td>
</tr>
<tr>
<td>EX_cbl1(e)</td>
<td>cbl1[e]</td>
<td>0</td>
<td>EX_opi(e)</td>
<td>opi[e]</td>
<td>0</td>
</tr>
</tbody>
</table>

*Biliverdin_Nutrient_ShadowPrices_BL21.m
Biliverdin_Amino_Acid_ShadowPrices_BL21.m

*Potential media supplements based on regulatory constraints*
Potential Impact of L-Arginine, L-Glutamine, and L-Glutamate

\[ \text{EX}_{\text{glc}}(e) \geq -10; \text{EX}_o2(e) \geq -20 \]

Biliverdin Production Envelope with Plasmid

- **No Extra Amino Acids**: Small Improvement
- **L-Glutamate**: Small Improvement
- **L-Glutamine**: Small Improvement
- **L-Arginine**: Small Improvement

Biliverdin Production Envelope with Plasmid

- **EX_{glu}-L(e) \geq -0.1859**
- **EX_{gln}-L(e) \geq -0.1871**
- **EX_{arg}-L(e) \geq -0.0660**

Biliverdin_ProductionEnvelope_AminoAcids_BL21_Plasmid.m
Regulatory Issues

Regulatory constraints are not included in the constraint-based reconstructions.

Proline is self-limiting

Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Secreted Byproducts & Phenotypes

![Graph showing secreted byproducts and phenotypes]

- Acetate, Formate, Glycolate
- Acetate
- Formate
- Glycolate
- Ethanol
## Active Exchange Reactions as Oxygen Increases

<table>
<thead>
<tr>
<th>Reactions</th>
<th>$o_2$ = -1</th>
<th>$o_2$ = 2</th>
<th>$o_2$ = 3</th>
<th>$o_2$ = 4</th>
<th>$o_2$ = 5</th>
<th>$o_2$ = 6</th>
<th>$o_2$ = 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ec_biomass</td>
<td>0.28622</td>
<td>0.33618</td>
<td>0.38608</td>
<td>0.45398</td>
<td>0.48578</td>
<td>0.53521</td>
<td>0.57763</td>
</tr>
<tr>
<td>Ec_ac(e)</td>
<td>8.71606</td>
<td>9.34586</td>
<td>9.97566</td>
<td>10.6055</td>
<td>11.2353</td>
<td>11.7853</td>
<td>11.145</td>
</tr>
<tr>
<td>Ec_ca2(e)</td>
<td>-0.00149</td>
<td>-0.00175</td>
<td>-0.00201</td>
<td>-0.00227</td>
<td>-0.00253</td>
<td>-0.00279</td>
<td>-0.00301</td>
</tr>
<tr>
<td>Ec_cl(e)</td>
<td>-0.00203</td>
<td>-0.00230</td>
<td>-0.00257</td>
<td>-0.00284</td>
<td>-0.00311</td>
<td>-0.00338</td>
<td>-0.00365</td>
</tr>
<tr>
<td>Ec_co2(e)</td>
<td>0.049621</td>
<td>0.047812</td>
<td>0.046093</td>
<td>0.044194</td>
<td>0.042385</td>
<td>0.209667</td>
<td>2.85461</td>
</tr>
<tr>
<td>Ec_cobalt2(e)</td>
<td>-7.166E-06</td>
<td>-8.40E-06</td>
<td>-9.65E-06</td>
<td>-1.09E-05</td>
<td>-1.21E-05</td>
<td>-1.34E-05</td>
<td>-1.44E-05</td>
</tr>
<tr>
<td>Ec_h2o(e)</td>
<td>-1.55E-05</td>
<td>-1.66E-05</td>
<td>-1.76E-05</td>
<td>-1.86E-05</td>
<td>-1.94E-05</td>
<td>-2.02E-05</td>
<td>-2.11E-05</td>
</tr>
<tr>
<td>Ec_eno(e)</td>
<td>0.00002</td>
<td>-0.000002</td>
<td>-0.000002</td>
<td>-0.000002</td>
<td>-0.000002</td>
<td>-0.000002</td>
<td>-0.000002</td>
</tr>
<tr>
<td>Ec_eto(e)</td>
<td>8.61635</td>
<td>5.43574</td>
<td>4.05513</td>
<td>2.67452</td>
<td>1.29391</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ec_fe2(e)</td>
<td>-0.000236</td>
<td>-0.000277</td>
<td>-0.000319</td>
<td>-0.000369</td>
<td>-0.000403</td>
<td>-0.000442</td>
<td>-0.000477</td>
</tr>
<tr>
<td>Ec_fe3(e)</td>
<td>-0.000223</td>
<td>-0.000262</td>
<td>-0.000301</td>
<td>-0.00034</td>
<td>-0.000379</td>
<td>-0.000418</td>
<td>-0.000451</td>
</tr>
<tr>
<td>Ec_glc(e)</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
<td>-10</td>
</tr>
<tr>
<td>Ec_glyc(e)</td>
<td>0.000019</td>
<td>0.0000225</td>
<td>0.0000258</td>
<td>0.0000292</td>
<td>0.0000325</td>
<td>0.0000358</td>
<td>0.0000387</td>
</tr>
<tr>
<td>Ec_h(e)</td>
<td>28.2044</td>
<td>28.7485</td>
<td>29.2926</td>
<td>29.8367</td>
<td>30.3808</td>
<td>30.7401</td>
<td>31.2167</td>
</tr>
<tr>
<td>Ec_h2o(e)</td>
<td>1.02138</td>
<td>3.71756</td>
<td>6.41375</td>
<td>9.10993</td>
<td>11.8061</td>
<td>14.5698</td>
<td>18.3409</td>
</tr>
<tr>
<td>Ec_k(e)</td>
<td>-0.05587</td>
<td>-0.06561</td>
<td>-0.07553</td>
<td>-0.08508</td>
<td>-0.09492</td>
<td>-0.10447</td>
<td>-0.11247</td>
</tr>
<tr>
<td>Ec_meh(e)</td>
<td>5.7207</td>
<td>6.7207</td>
<td>7.7207</td>
<td>8.7207</td>
<td>9.7207</td>
<td>10.7207</td>
<td>11.7207</td>
</tr>
<tr>
<td>Ec_mg2(e)</td>
<td>-0.000248</td>
<td>-0.000292</td>
<td>-0.000335</td>
<td>-0.000378</td>
<td>-0.000421</td>
<td>-0.000465</td>
<td>-0.000504</td>
</tr>
<tr>
<td>Ec_mon2(e)</td>
<td>-0.00002</td>
<td>-0.000027</td>
<td>-0.000032</td>
<td>-0.000037</td>
<td>-0.000042</td>
<td>-0.000047</td>
<td>-0.000053</td>
</tr>
<tr>
<td>Ec_mod(e)</td>
<td>-3.69E-05</td>
<td>-4.34E-05</td>
<td>-4.98E-05</td>
<td>-5.62E-05</td>
<td>-6.27E-05</td>
<td>-6.90E-05</td>
<td>-7.45E-05</td>
</tr>
<tr>
<td>Ec_nh4(e)</td>
<td>-3.1315</td>
<td>-3.67035</td>
<td>-4.2092</td>
<td>-4.74805</td>
<td>-5.2869</td>
<td>-5.82600</td>
<td>-6.28600</td>
</tr>
<tr>
<td>Ec_n2(e)</td>
<td>-9.25E-05</td>
<td>-0.00011</td>
<td>-0.00012</td>
<td>-0.00014</td>
<td>-0.00016</td>
<td>-0.00017</td>
<td>-0.00019</td>
</tr>
<tr>
<td>Ec_o2(e)</td>
<td>-1</td>
<td>-2</td>
<td>-3</td>
<td>-4</td>
<td>-5</td>
<td>-6</td>
<td>-7</td>
</tr>
<tr>
<td>Ec_py(e)</td>
<td>-0.27611</td>
<td>-0.32423</td>
<td>-0.37236</td>
<td>-0.42048</td>
<td>-0.46861</td>
<td>-0.51629</td>
<td>-0.55733</td>
</tr>
<tr>
<td>Ec_so2(e)</td>
<td>-0.07219</td>
<td>-0.08477</td>
<td>-0.09736</td>
<td>-0.10994</td>
<td>-0.12252</td>
<td>-0.13499</td>
<td>-0.14572</td>
</tr>
<tr>
<td>Ec_zn2(e)</td>
<td>-9.76E-05</td>
<td>-0.00011</td>
<td>-0.00013</td>
<td>-0.00015</td>
<td>-0.00017</td>
<td>-0.00018</td>
<td>-0.00020</td>
</tr>
<tr>
<td>Ec_biliverdin(e)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Ec_co(e)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

*Based Metabolic Reconstructions & Analysis*
Undesired Secreted By-products

\[ \text{EX}_{\text{glc}}(e) > -10, \text{EX}_{\text{o2}}(e) > -5, \text{EX}_{\text{biliverdin}}(e) > 0.01 \]
Biliverdin Production

- Bioproduct Identification

- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model

- Host Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis

- Final Design
% Biliverdin_optKnock_BL21_PreprocessingReactions.m

clear; clc;

% Input the E.coli core model
model=readCbModel('iECD_1391');

Add biliverdin reactions

% Set key variables
model = changeRxnBounds(model, 'EX_glc(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_o2(e)', -20, 'l');
model = changeObjective(model, 'Ec_biomass_iJO1366_core_53p95M');

% Select potential knockout reaction
[GeneClasses RxnClasses modelIrrevFM] = pFBA(model, 'geneoption', 0);

undesiredReactions = model.rxns(ismember(model.subSystems,{'Cell Envelope Biosynthesis','Glycerophospholipid Metabolism', ...
'Inorganic Ion Transport and Metabolism','Lipopolysaccharide Biosynthesis / Recycling','Membrane Lipid Metabolism',...
'Murein Recycling','Transport, Inner Membrane','Transport, Outer Membrane Porin','Transport, Outer Membrane','tRNA Charging'}));

[transRxns,nonTransRxns] = findTransRxns(model,true);
[hiCarbonRxns,nCarbon] = findHiCarbonRxns(model,7);
[a,ids] = ismember([RxnClasses.Essential_Rxns; RxnClasses.ZeroFlux_Rxns; RxnClasses.Blocked_Rxns;...
undesiredReactions;hiCarbonRxns; {'ATPM'};{'Ec_biomass_iJO1366_core_53p95M'}], nonTransRxns);

id1=ids(a);
nonTransRxns(id1)=[];
selectedRxns=nonTransRxns;

save('Biliverdin_SelectedRxns_BL21.mat','selectedRxns');  \% Save the selected Reactions is a MAT file

Creating and Saving
Selected Reactions in BL21

Removing Undesired Subsystems

Save 'selectedRxns' to a "mat" file
% Biliverdin_OptKnock_iJO1366_WithPrecalculatedReactions.m

clear; clc;

% Input the E.coli core model
model=readCbModel('ecoli_iJO1366');

Add Biliverdin Reactions

% Set key variables
model = changeRxnBounds(model, 'EX_glc(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_o2(e)', -20, 'l');

% Set optimization objective
model = changeObjective(model, 'Ec_biomass_iJO1366_core_53p95M');

% Load list of selected reactions
load('Biliverdin_SelectedRxns_iJO1366');
target = 'EX_biliverdin(e)';
options.targetRxn = target;
options.vMax = 1000;
options.numDel = 2;
options.numDelSense = 'L';
constrOpt.rxnList = {'Ec_biomass_iJO1366_core_53p95M', 'ATPM'};
constrOpt.values = [0.1, 3.15];
constrOpt.sense = 'GE';

[optKnockSol,bilevelMILPproblem] = OptKnock(model, selectedRxns, options, constrOpt,{},true);
% [optKnockSol,bilevelMILPproblem] = OptKnock(model, selectedRxns, options, constrOpt,{},false);
[optKnockSol.rxnList']

[growthRate, minProd, maxProd] = testOptKnockSol(model, target, optKnockSol.rxnList)
**Biliverdin OptKnock Results**

**HEEMEOX**

1. **Knockout**
   - ans = 'SUCOAS'
   - growthRate = 0.9657
   - minProd = 0
   - maxProd = 0

2. **Knockouts**
   - ans = 'ADK3','RPI'
   - growthRate = 0.3345
   - minProd = 0
   - maxProd = 0

3. **Knockouts**
   - ans = 'GND', 'GNK', 'MECDPDH5'
   - growthRate = 0.9041
   - minProd = 0.1
   - maxProd = 0.1

4. **Knockouts**
   - ans = 'DHAPT','FUM','GLUDy','IPDDI'
   - growthRate = 0.8016
   - minProd = 0
   - maxProd = 0

**WHY?**
1. Can't find reactions to couple the output production to the biomass growth.

2. But, the HEMEOX reaction was added to the cell with no regulatory constraints so it will automatically produce biliverdin based on the copy number and promoter strength of the HEMEOX enzymes created by the plasmid.

3. To use the model, set a probable lower limit for the HEMOX reaction.
Biliverdin Production

• Bioprodct Identification

• Bioprodct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioprodct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Balancing Cofactors

- **ATP**
- **NADPH** ↔ **NADP**
- **L-Glutamate**
- **Ferredoxin**

**Heme Biosynthesis**

**Need to recycle**

Additional cofactors:
- **O₂**
- **Fe²⁺**
- **ATP**
- **NADPH** ↔ **NADP**

Other cofactors involved in the pathway include:
- **FNADPR**
- **flxso**
- **Heme**
Maximum Biliverdin Production: Ferredoxin NADPH

**Without Ferredoxin-NADPH**
- Max Production = 406.7 g/L
- 16 Hours Expression Time
- Unlimited Resources
- Seed = 14.1 g/L (OD600 = 30)
- Production Rate = 0.62 mmol·gDW⁻¹·h⁻¹

**With Ferredoxin-NADPH**
- Max Production = 409 g/L
- 16 Hours Expression Time
- Unlimited Resources
- Seed = 14.1 g/L (OD600 = 30)
- Production Rate = 0.64 mmol·gDW⁻¹·h⁻¹

Improved performance at high flux values
Maximum Biliverdin Production: Ferredoxin NADH

**Without Ferredoxin-NADH**

Max Production = 406.7 g/L

16 Hours Expression Time
Unlimited Resources
Seed = 14.1 g/L (OD600 = 30)

0.62 mmol·gDW⁻¹·h⁻¹

**With Ferredoxin-NADH**

Max Production = 406.7 g/L

16 Hours Expression Time
Unlimited Resources
Seed = 14.1 g/L (OD600 = 30)

No Change!

DynamicBiliverdinProduction_BL21_maxProd_200Limit_Plasmid.m
Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing

✓ Sampling Analysis

• Final Design
Constraint-based Metabolic Reconstructions & Analysis

H. Scott Hinton, 2016

Lesson: Bioproduct Production

Biliverdin Final Pathway

Glutamate and Glutamine

Secreted Biliverdin

Heme Biosynthesis

nadp
nadph
flxr
flxso

Utah State University
BIE 5500/6500

Lesson: Bioproduct Production
Biliverdin Sampling

Biliverdin_gpSampler_BL21_Three_Plots_Ferredoxin.m

**Graphs:**
- **EX-biliverdin(e)**
- **HEMOEX**
- **FCLT**
- **GLUDy**
- **ICDHyr**
- **NADTRHD**
- **FNADPR**
- **ATPS4rpp**
- **THD2pp**
- **GND**
- **GLNS**
- **NADH16pp**
- **Ec-biomass-UO1366-core-53pp95M**
- **EX-glc(e)**
- **EX-02(e)**

**Annotations:**
- $\text{Ex}_{\text{biliverdin}}(e) > 0.01$
- $\text{Ex}_{\text{biliverdin}}(e) > 0.1$
- $\text{Ex}_{\text{biliverdin}}(e) > 1.0$
Biliverdin Production

• Bioproduct Identification

• Bioproduct Pathway
  ✓ Defining pathway
  ✓ Adding pathway to the model
  ✓ Adding the plasmid to the model

• Host Design
  ✓ Bioproduct Maximum Production
  ✓ Carbon Sources
  ✓ Nutrient & Amino Acid Limitations
  ✓ Undesired By-products
  ✓ Growth Coupling
  ✓ Cofactor Balancing
  ✓ Sampling Analysis

• Final Design
Final Design

- **Carbon Source**
  - Glucose, Glycerol
- **Aerobic**
  - \( \text{EX}\_\text{o2(e)} \geq 20 \)
- **M9 Minimal Media**
  - Needs supplemental ammonium & sulfate
  - Ammonium hydroxide as the pH control
- **Byproducts**
  - None
- **Knockouts**
  - None
- **Cofactor Balancing**
  - Add Ferredoxin - NADPH

Max Production = 409 g/L
16 Hours Expression Time
Unlimited Resources
Seed = 14.1 g/L (OD600 = 30)

0.64 mmol gDW\(^{-1}\) h\(^{-1}\)
Heme Oxygenase (HO1) Protein Expression

- The vector with HO1 gene was transformed to *E. coli* BL21(DE3)
- LB medium with 100 \( \mu g/ml \) Kanamycin was used
- HO1 protein expression was induced by IPTG or lactose
- 37 °C, 225 rpm overnight

Dong Chen, USU SBI Bioproducts Summit, February 2012
Approximate Biliverdinin Growth Rates

The approximate growth rate can be determined from measured optical density measurements with the following equation

$$\mu = \frac{2.303 \cdot [\log_{10}(OD_2) - \log_{10}(OD_1)]}{t_2 - t_1}$$

For the pre-induction approximate growth rate set $OD_2 = 25$ at 7 hours and $OD_1 = 7$ at 4 hours

$$\mu_{pre} = \frac{2.303 \cdot [\log_{10}(25) - \log_{10}(7)]}{7 - 4} = 0.424 h^{-1}$$

For the post-induction approximate growth rate set $OD_2 = 33$ at 24 hours and $OD_1 = 28$ at 8 hours

$$\mu_{post} = \frac{2.303 \cdot [\log_{10}(33) - \log_{10}(28)]}{24 - 8} = 0.010 h^{-1}$$

Data provided by Dr. Dong Chen at Utah State University
Approximate Biliverdin Production Range

- In order to make an approximation of the biliverdin flux rate produced by heme oxygenase will require some very broad approximations to allow us to determine a potential range of operation.
  - Assumption #1: Due to the low growth rate during induction (0.010 hr\(^{-1}\)) we will assume that the number of cells remains the same which implies a constant mass during induction.
  - Assumption #2: An OD600 of 30 implies a dry cell weight of 14.1 gDW/L.
  - Assumption #3: To determine the upper and lower bounds of the flux rate we will assume that for the upper bound all biliverdin production will occur in the first hour and for the lower bound that the biliverdin production will be constant over the entire induction period.
  - Assumption #4: The purification process will only capture 10% of the produced biliverdin.

- The measure production of biliverdin at OD600 was a total of 30 mg/L over 16 hours

Upper Bound \(= \left( \frac{30\text{mg}}{L \cdot h} \right) \left( \frac{L}{14.1\text{gDW}} \right) = \frac{2.128\text{mg}}{g\text{DW} \cdot h} \)

Lower Bound \(= \left( \frac{30\text{mg}}{L \cdot 16h} \right) \left( \frac{L}{14.1\text{gDW}} \right) = \frac{0.1330\text{mg}}{g\text{DW} \cdot h} \)

- Converting to millimoles (mmol) using the molecular weight of biliverdin

Upper Bound \(= \left( \frac{2.128\text{mg}}{g\text{DW} \cdot h} \right) \left( \frac{\text{mmol}}{582.646\text{mg}} \right) = \frac{0.003652\text{mmol}}{g\text{DW} \cdot h} \)

Lower Bound \(= \left( \frac{0.1330\text{mg}}{g\text{DW} \cdot h} \right) \left( \frac{\text{mmol}}{582.646\text{mg}} \right) = \frac{0.000228\text{mmol}}{g\text{DW} \cdot h} \)

- For a purification process that only captures 10% of the produced biliverdin multiply these numbers by 10
Biliverdin Measured Results and Limitations

• For a purification process that only captures 10% of the produced biliverdin yields the following production range

\[
Upper \ Bound = \frac{0.03652 \ mmol}{gDW \cdot h}
\]

\[
Lower \ Bound = \frac{0.00228 \ mmol}{gDW \cdot h}
\]

• The growth rate during induction is 0.010 h\(^{-1}\)

• What is the maximum amount of biliverdin that could be produced per hour with a flux rate of 0.8 mmol\(\cdot gDW^{-1}\cdot hr^{-1}\)

\[
\left( \frac{0.8 \ mmol}{gDW \cdot h} \right) \left( \frac{582.646 \ mg}{mmol} \right) = \frac{466.12 \ mg}{gDW \cdot h}
\]

\[
\left( \frac{466.12 \ mg}{gDW \cdot h} \right) \left( \frac{14.1 \ gDW}{L} \right) = \frac{6.572 \ g}{L \cdot h}
\]
Constraint-based Metabolic Reconstructions & Analysis

H. Scott Hinton, 2016

The Production Limits Using Estimated HEMEOX Flux Rates

\[ \text{EX}_{\text{glc}}(e) \geq -10; \ \text{EX}_{\text{o2}}(e) \geq -20; \ \text{initBiomass} = 14.1 \ gDW/L \ (\text{OD600} = 30); \ \text{timeStep} = 0.125 \ h; \ nSteps = 128 \]
Biliverdin Production

- Bioproduct Identification
- Bioproduct Pathway
  - Defining pathway
  - Adding pathway to the model
  - Adding the plasmid to the model
- Host Design
  - Bioproduct Maximum Production
  - Carbon Sources
  - Nutrient & Amino Acid Limitations
  - Undesired By-products
  - Growth Coupling
  - Cofactor Balancing
  - Sampling Analysis
- Final Design