Biological Pathways Analysis and Engineering

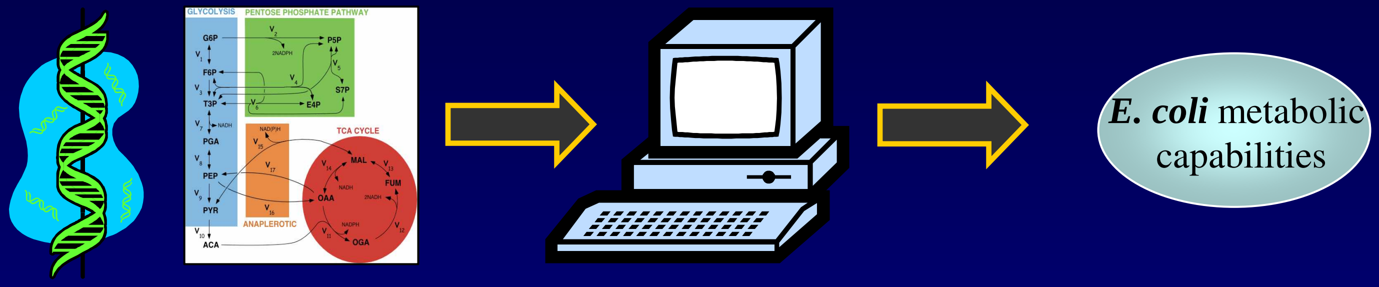
Costas D. Maranas

Penn State University
University Park, PA 16802

E-mail: costas@psu.edu
Web page: fenske.che.psu.edu/faculty/cmaranas

Presentation Outline

- □ Systems biology and the constraints-based modeling approach



E. coli metabolic capabilities

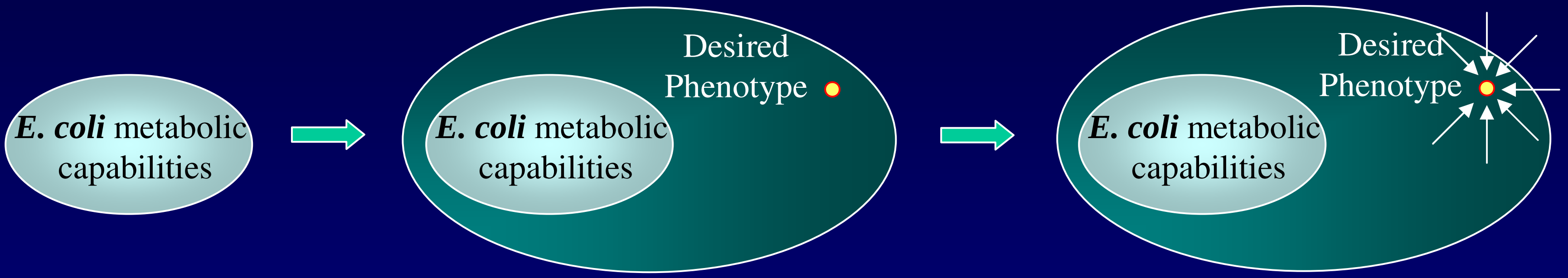
- □ Pathway discovery and optimization

- □ Constraining allowable cellular behavior

- □ Metabolic network structural and topological analysis

Presentation Outline

- Systems biology and the constraints-based modeling approach


E. coli metabolic capabilities
E. coli metabolic capabilities
Desired Phenotype
E. coli metabolic capabilities
Desired Phenotype

- Pathway discovery and optimization

  How can we systematically select the appropriate set of pathways/genes to recombine into existing production systems?

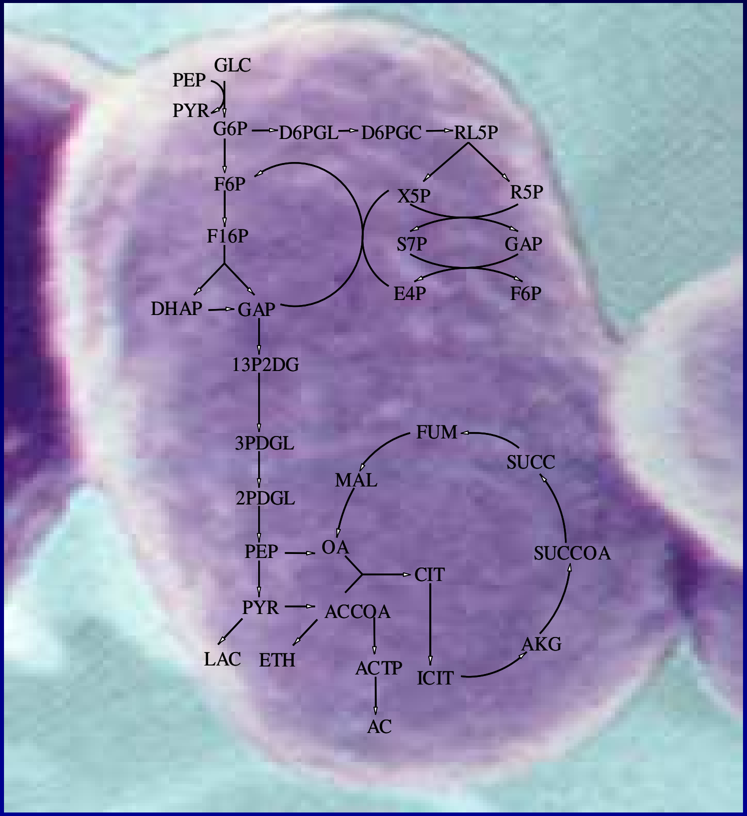
- Constraining allowable cellular behavior

  How can we identify gene knockouts that will force biochemical overproduction by coupling it with cell growth?

- Metabolic network structural and topological analysis

  How can we identify multiple metabolic manipulations for producing a desired product and also computationally evaluate of the consequences of potential network modifications ?

Chemical factory on the μm scale

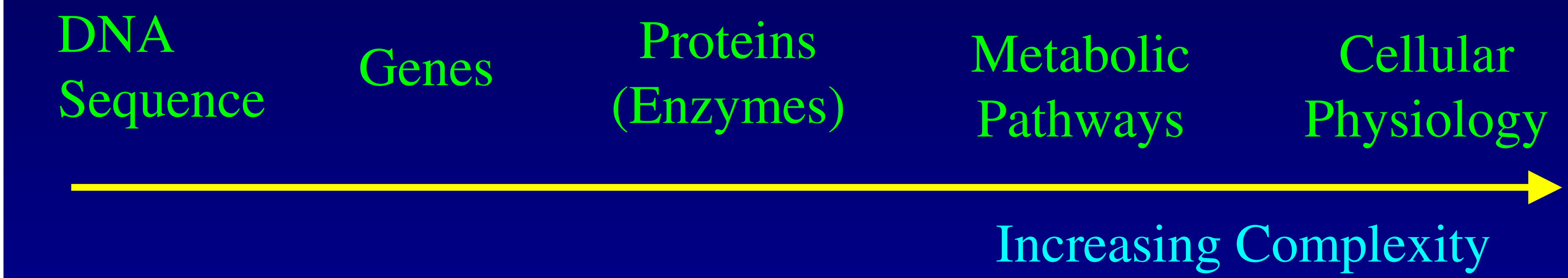


Escherichia coli

Chemical Process Plant

What is Metabolism?

- Metabolism is the totality of chemical reactions that occur in an organism



DNA Sequence

Genes

Proteins (Enzymes)

Metabolic Pathways

Cellular Physiology

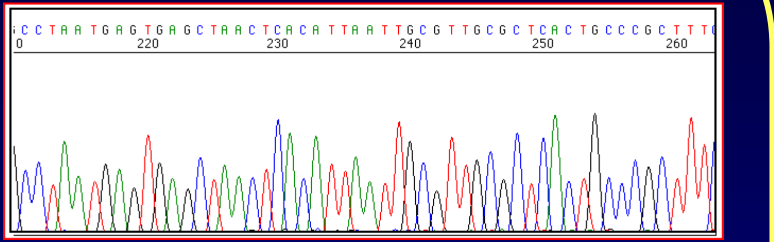
Increasing Complexity

- Metabolic Engineering is the analysis and modification of metabolic pathways
    - Applications include biochemical production, bioremediation, and drug discovery

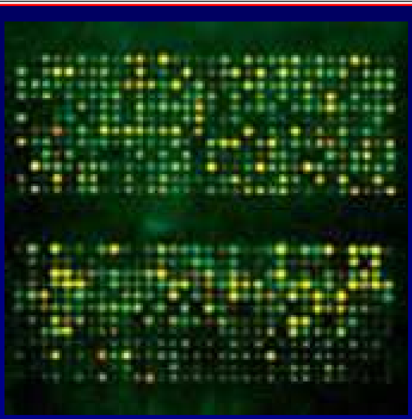
HT Technologies for Systems Biology

DNA Sequence

Automated Sequencing
Genomics



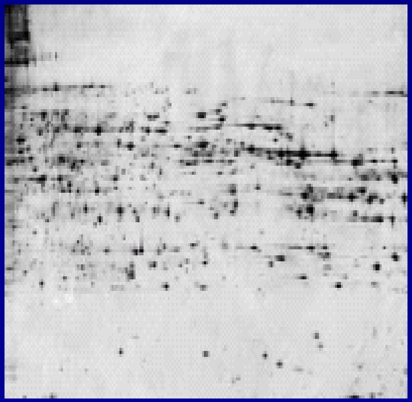
Genes



DNA Microarrays
Transcriptomics

Enzymes
(Gene Products)

2-D Protein Arrays
Proteomics



Metabolism and
Cellular Physiology

Metabolic Reconstruction Technology

Genome Annotation:

DNA sequence

⬇ ORFs identification

Genes

⬇ ORFs assignment

Genes Products

⬇

Function

Genome Database

ORF = open reading frame, a short fragment of DNA that is translated into RNA message

⬇

Metabolic Reconstruction:

List of reactions

⬇

Pathway Database ➡ Organism’s metabolism

➡

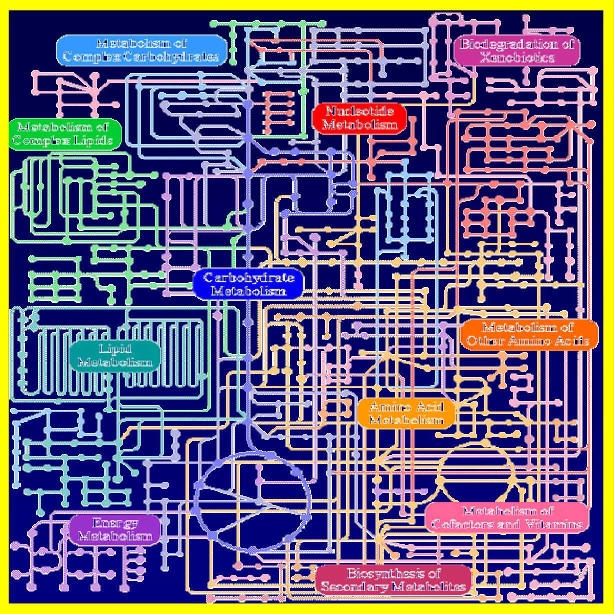
Organism-Specific Model Construction:

Literature Review ↔ Wet Lab

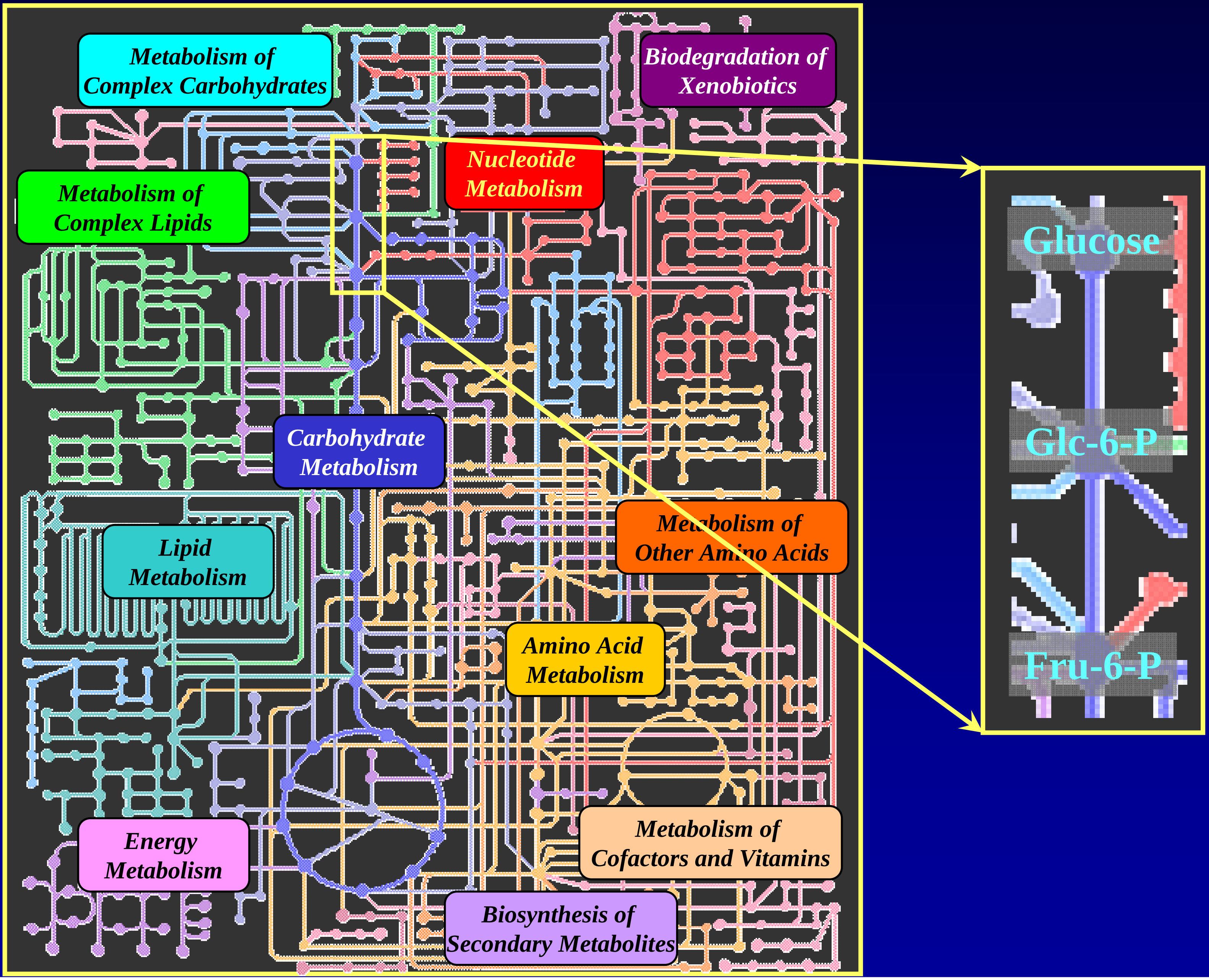
Literature Review ↕ Manual curation

Manual curation ↔ Wet Lab

⬆

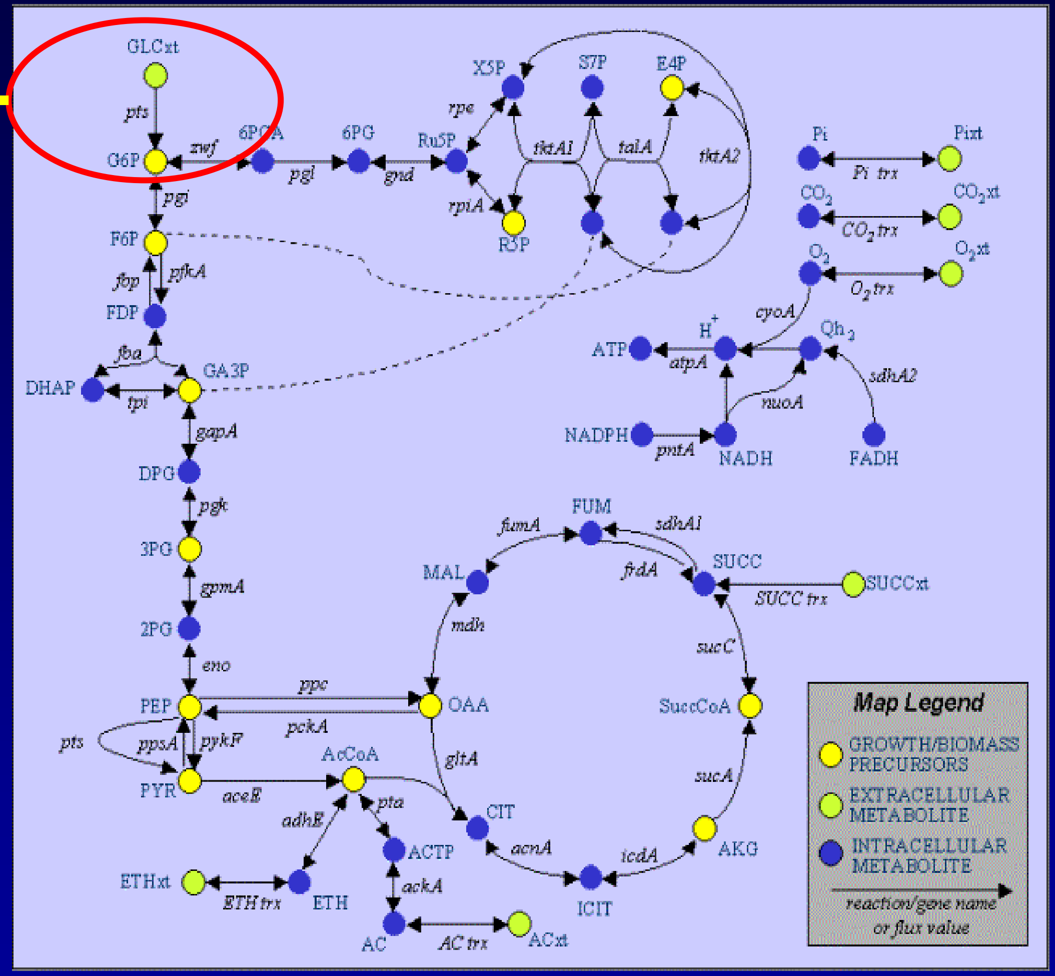

Metabolism of Complex Carbohydrates
Biodegradation of Xenobiotics
Nucleotide Metabolism
Metabolism of Complex Lipids
Carbohydrate Metabolism
Metabolism of Other Amino Acids
Lipid Metabolism
Amino Acid Metabolism
Metabolism of Cofactors and Vitamins
Energy Metabolism
Biosynthesis of Secondary Metabolites

Complexity of Metabolic Networks


Metabolism of Complex Carbohydrates
Biodegradation of Xenobiotics
Nucleotide Metabolism
Metabolism of Complex Lipids
Carbohydrate Metabolism
Metabolism of Other Amino Acids
Lipid Metabolism
Amino Acid Metabolism
Energy Metabolism
Metabolism of Cofactors and Vitamins
Biosynthesis of Secondary Metabolites
Glucose
Glc-6-P
Fru-6-P

“Small” E. coli Model (< 100 reactions)

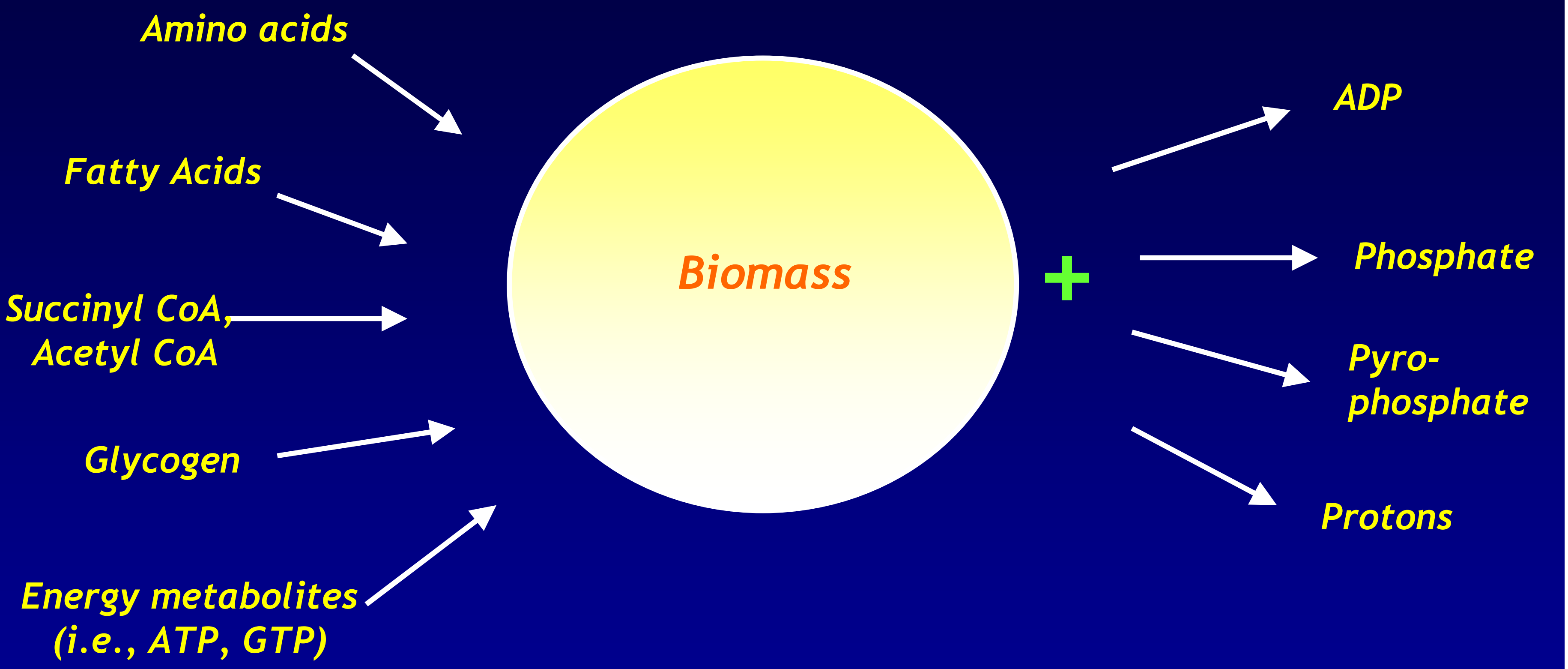
Glucose → G6P



Central Metabolism in *E. coli*

(http://gcrg.ucsd.edu/downloads/PathwayFBA/default.htm)

Biomass Composition



Biomass formation modeled as:

$\Sigma(x)(\text{metabolite}) \rightarrow \text{Biomass} + \Sigma\ (y)(\text{metabolite})$

Mass Balances in Metabolic Networks

Reaction network:

$$A \xleftrightarrow{v_1} B$$

$$2B \xrightarrow{v_2} C$$

Mass balance:

$$\frac{d[B]}{dt} = v_1 - 2\,v_2$$

Metabolic Flux:

$$v \equiv \frac{\text{mmol B}}{\text{gDW}\cdot\text{hr}}$$

Stoichiometric matrix:

$$\begin{array}{c|cc} & v_1 & v_2 \\ \hline A & -1 & 0 \\ B & 1 & -2 \\ C & 0 & 1 \end{array}$$

Steady-state assumption:

$$0 = \sum_j S_{ij} v_j$$

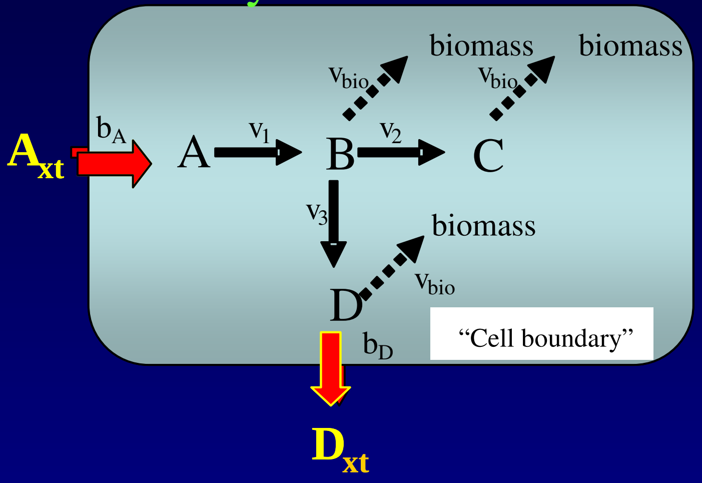
***i* = set of metabolites (chemical species)**

***j* = set of reactions**

Flux Balance Analysis

Given:

(1) Stoichiometry of the network



(2) Cellular composition information

Biomass Composition (mmol /gDW)	
B	**2**
C	**3**
D	**4**

(3) Substrate uptake rate

$$b_A = -10\frac{\text{mmol}}{\text{gDW*hr}}$$

Optimization Model:

5 unknowns > 4 equations

Maximize v_{bio}

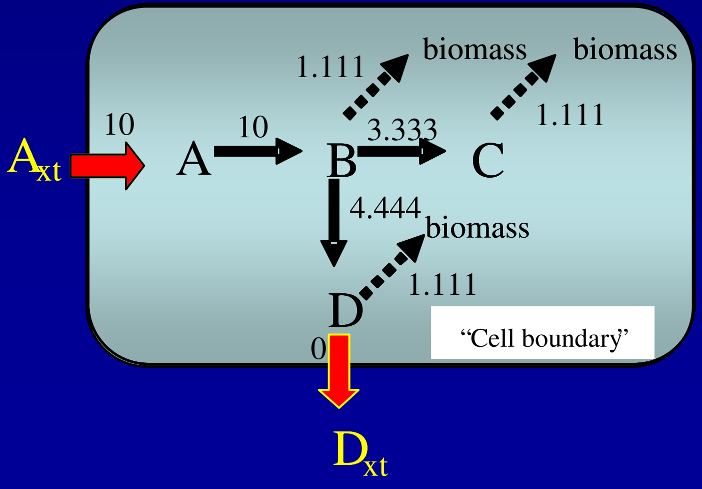
subject to

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & -2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_{bio} \end{pmatrix} = \begin{pmatrix} -10 \\ 0 \\ 0 \\ b_D \end{pmatrix}$$

$$v_1, v_2, v_3, v_{bio}, b_D > 0$$

Model Predictions:

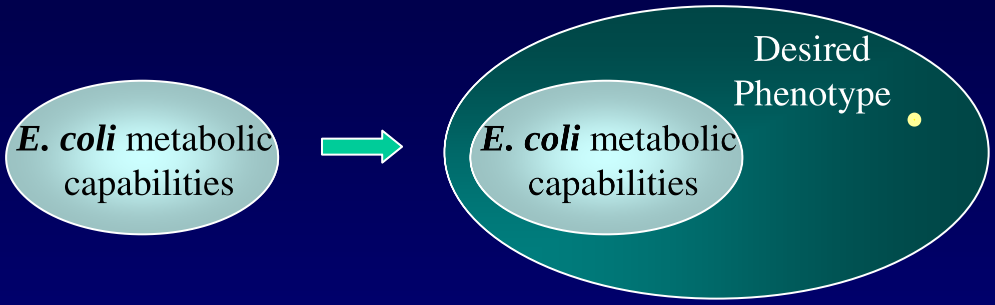
Growth rate: $v_{bio} = 1.111\ \text{hr}^{-1}$



Units of flux: mmol/(gDW*hr)

Presentation Outline

- □ Systems biology and the constraints-based modeling approach



E. coli metabolic capabilities

Desired Phenotype

E. coli metabolic capabilities

- □ Pathway discovery and optimization

  How can we systematically select the appropriate set of pathways/genes to recombine into existing production systems?

- □ Constraining allowable cellular behavior

  How can we identify gene knockouts that will force biochemical overproduction by coupling it with cell growth?

- □ Metabolic network structural and topological analysis

  How can we identify multiple metabolic manipulations for producing a desired product and also computationally evaluate of the consequences of potential network modifications ?

Expanding the Capabilities

E. coli Stoichiometric Models:

- **Pramanik & Keasling (1997)**

  (300 reactions, 289 metabolites)
- **Edwards & Palsson (2000)**

  (627 reactions, 438 metabolites)
- **Reed, Vo, Schilling & Palsson (2003)**

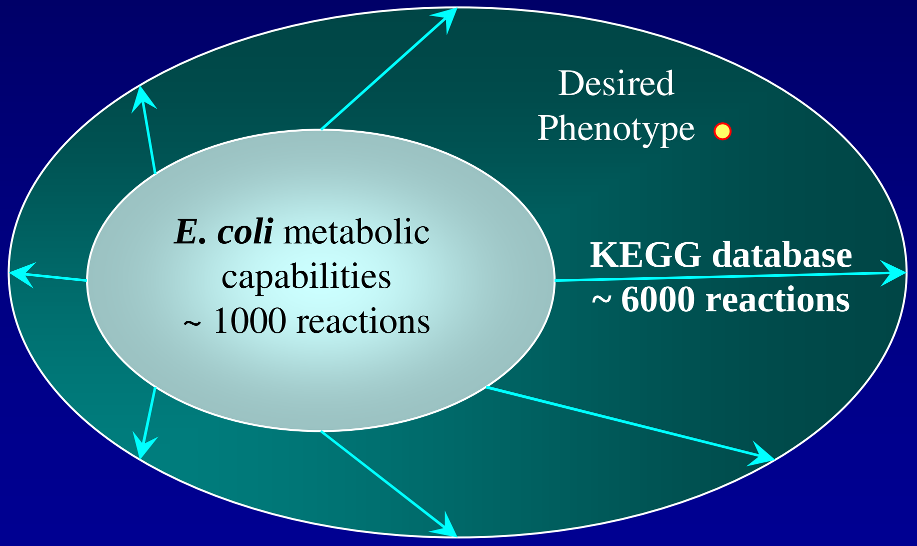
  (931 reactions, 625 metabolites)

5,700
reactions

from KEGG
and MetaCyc
databases

⇨

Multi-organism
Reaction
Network



Desired
Phenotype ●

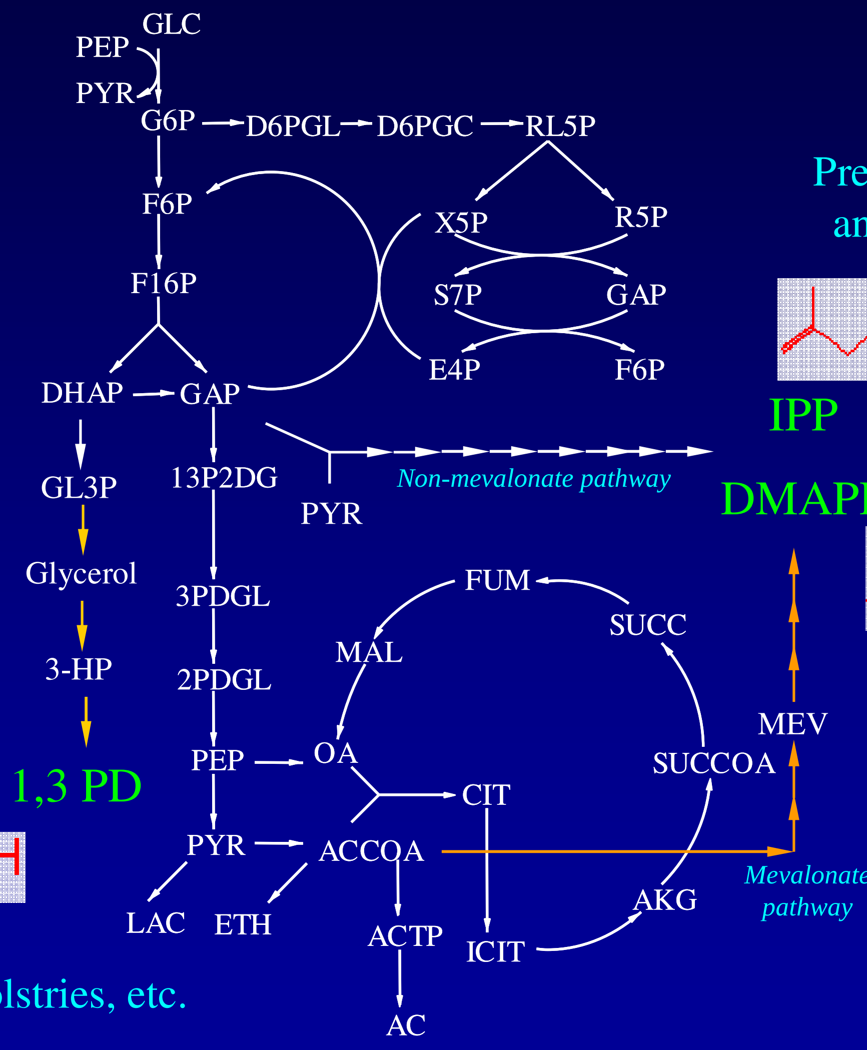
E. coli metabolic
capabilities
~ 1000 reactions

KEGG database
~ 6000 reactions

- Burgard, A.P. and C.D. Maranas (2001), “Probing the Performance Limits of the *Escherichia coli* Metabolic Network Subject to Gene Additions or Deletions,” *Biotechnology and Bioengineering*, **74**, 364-375.

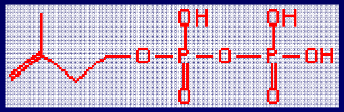
Pharkya, P., Burgard, A.P and C.D. Maranas (2004), “OptStrain: A Computational Framewrok for Redesign of Microbial Production Systems,” *Genome Research, 14,* 2367-2376.

Gene Addition Study



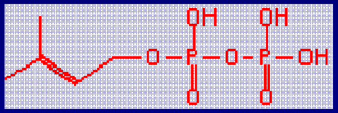
1,3 PD

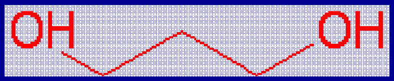
Precursors to anticancer and antimalarial drugs



IPP

DMAPP





Useful in Apparels, Upholstries, etc.

Mathematical Framework

Maximize $\sum_j c_j v_j$

Subject to $\sum_j S_{ij} v_j = 0$

$$y_j = \begin{cases} 1 & \text{if reaction flux switched “on”} \\ 0 & \text{if reaction flux switched “off”} \end{cases} \qquad 0 \le v_j \le v_j^{\max} \cdot y_j$$

Procedure:

1) **Find maximum theoretical yield using all reactions in multi-organism reaction network**

2) **Find minimum number of non-E. coli reactions necessary to achieve maximum yield**

Minimize $\sum_{\substack{j = nonEcoli \\ reactions}} y_j$

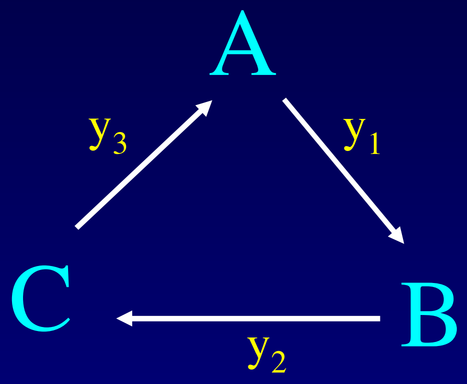
Subject to $\sum_j S_{ij} v_j = 0$

$$0 \le v_j \le v_j^{\max} \cdot y_j$$

Yield = maximum theoretical

Preprocessing Techniques

(1) Futile Cycle Exclusion

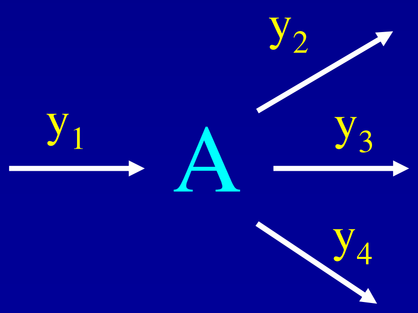


$$y_1 + y_2 + y_3 \le 2$$

(~ 350 futile cycles eliminated)

(2) Connectivity Constraints

If an internal metabolite is produced…
at least one reaction consuming this metabolite must be active



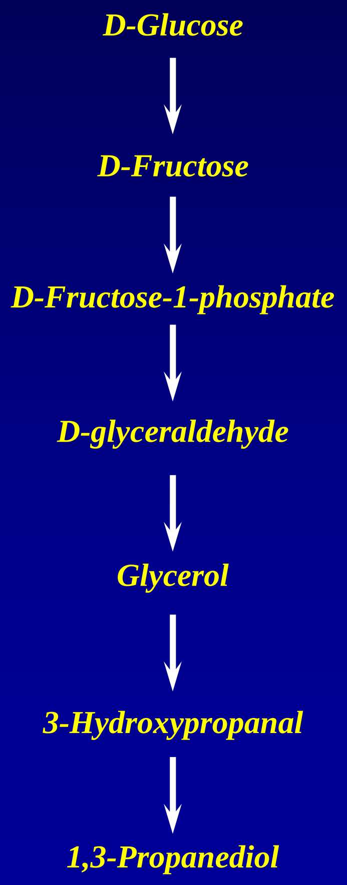
$$y_1 \le y_2 + y_3 + y_4$$

(~ 700 connectivity constraints added)

1) Pathway Discovery

Example: Glucose to 1,3-Propanediol using KEGG database

Typical Shortest Path Result



D-Glucose

D-Fructose

D-Fructose-1-phosphate

D-glyceraldehyde

Glycerol

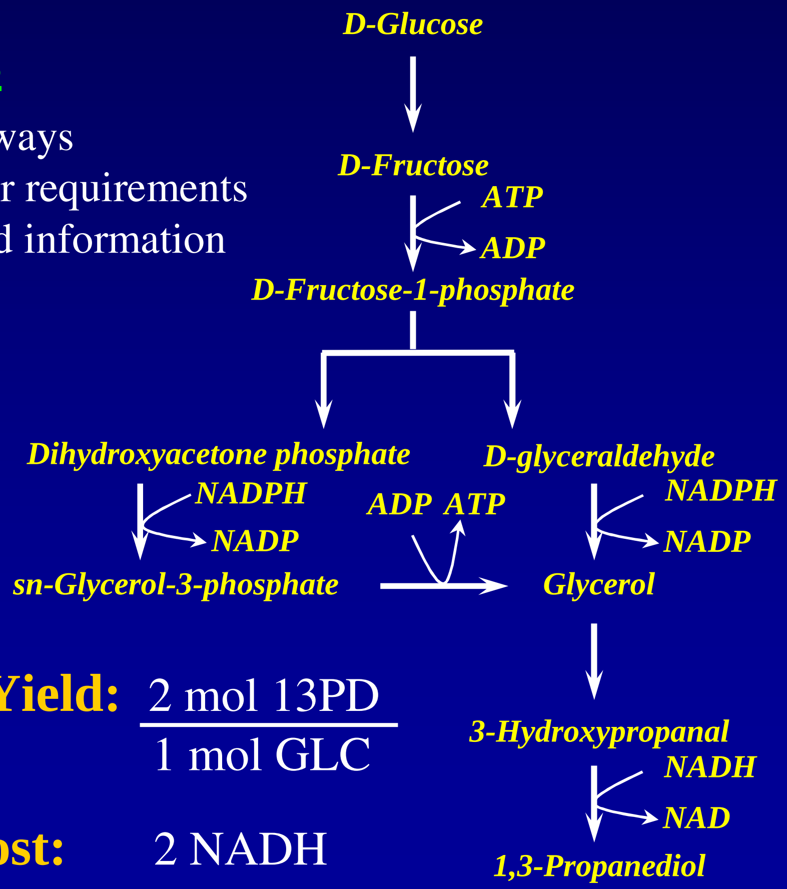
3-Hydroxypropanal

1,3-Propanediol

Key Advantages

- ❑ Branched pathways
- ❑ Energy/cofactor requirements
- ❑ Maximum yield information

Balanced Shortest Path Result at Max. Yield



D-Glucose

D-Fructose

ATP

ADP

D-Fructose-1-phosphate

Dihydroxyacetone phosphate

D-glyceraldehyde

NADPH

NADP

sn-Glycerol-3-phosphate

ADP ATP

NADPH

NADP

Glycerol

3-Hydroxypropanal

NADH

NAD

1,3-Propanediol

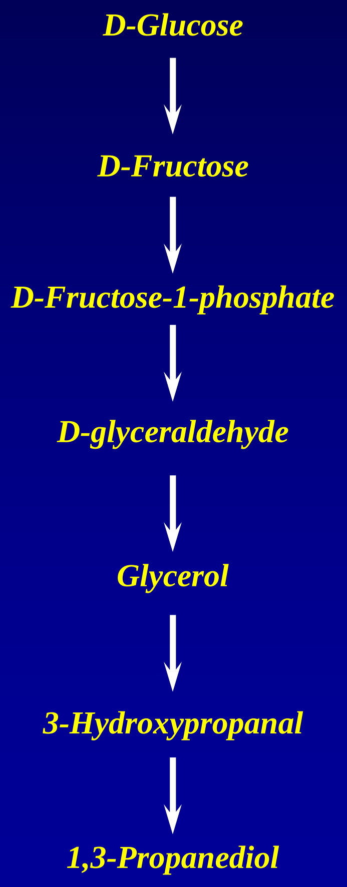
Max. Yield: $\frac{\text{2 mol 13PD}}{\text{1 mol GLC}}$

Net Cost: 2 NADH
2 NADPH

1) Pathway Discovery

Example: Glucose to 1,3-Propanediol using KEGG database

Typical Shortest Path Result



D-Glucose

D-Fructose

D-Fructose-1-phosphate

D-glyceraldehyde

Glycerol

3-Hydroxypropanal

1,3-Propanediol

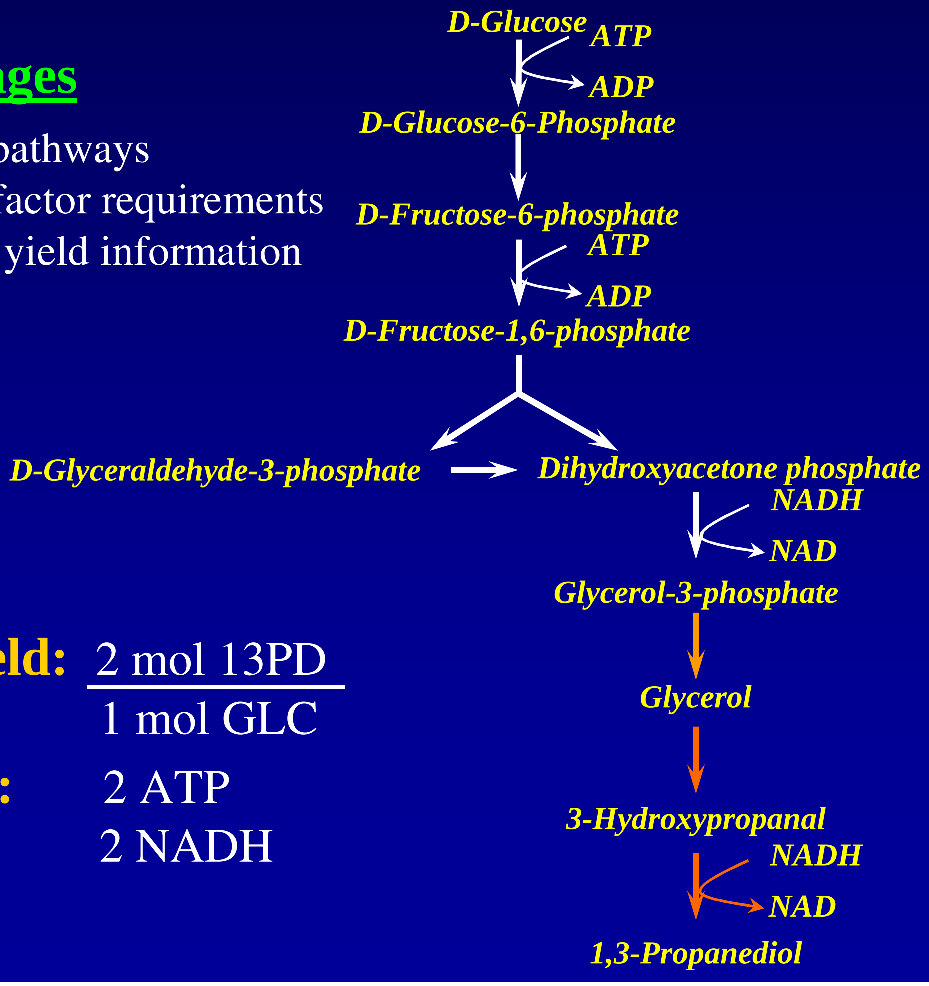
Key Advantages

- ❑ Branched pathways
- ❑ Energy/cofactor requirements
- ❑ Maximum yield information

Max. Yield: $\frac{\text{2 mol 13PD}}{\text{1 mol GLC}}$

Net Cost: 2 ATP
2 NADH

Alternative Pathway



D-Glucose

ATP

ADP

D-Glucose-6-Phosphate

D-Fructose-6-phosphate

ATP

ADP

D-Fructose-1,6-phosphate

D-Glyceraldehyde-3-phosphate

Dihydroxyacetone phosphate

NADH

NAD

Glycerol-3-phosphate

Glycerol

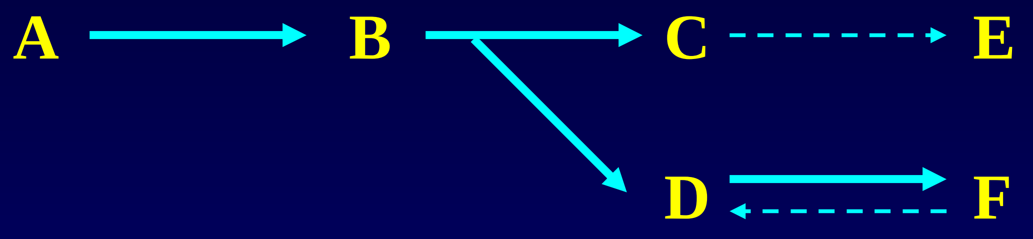
3-Hydroxypropanal

NADH

NAD

1,3-Propanediol

2) Minimal Reaction Network Study



Objective: **Identify the minimal sets of reactions capable of supporting various growth rates on different substrates**

Flux balance model: Edwards & Palsson, 2000

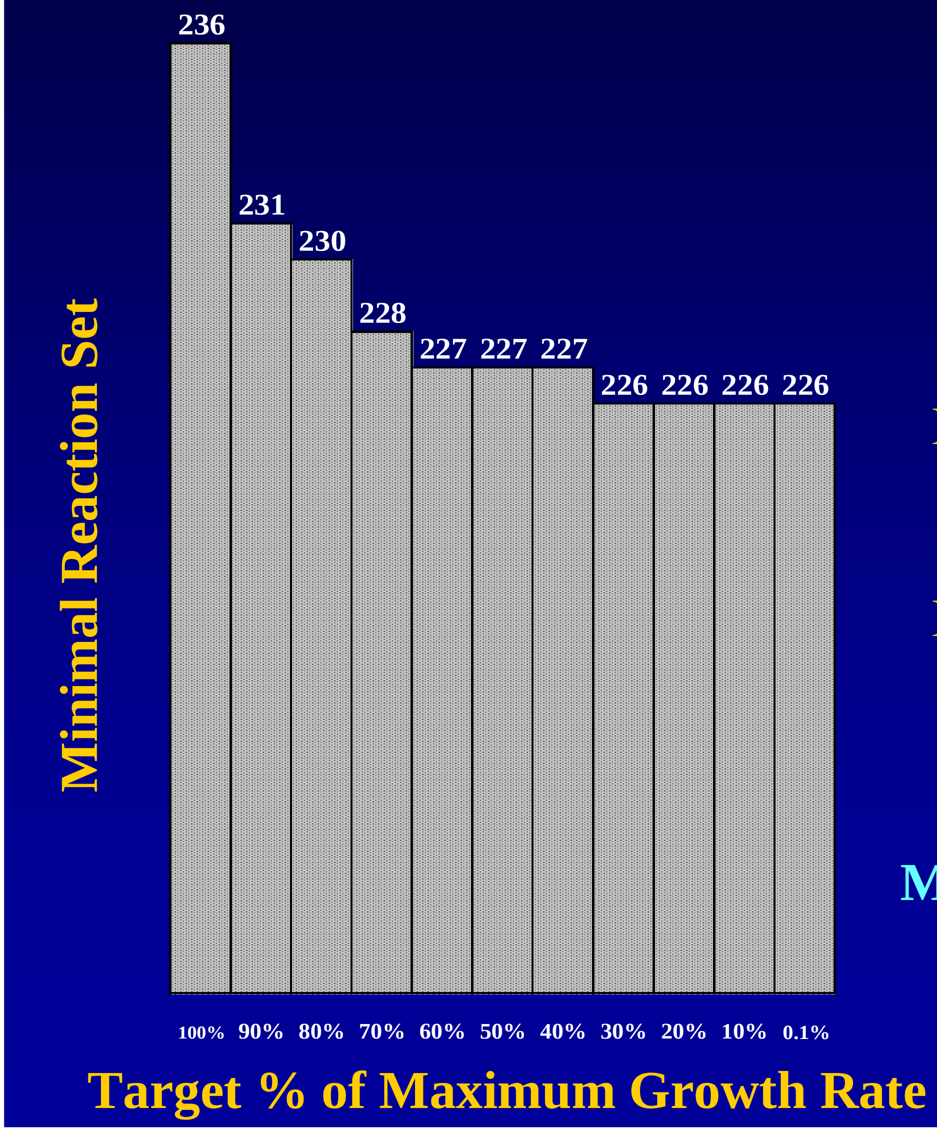
Biomass composition: Neidhardt, 1996

Two uptake environments: (i) Glucose only
(ii) Multiple organic uptake

❑ **Burgard & Maranas, *Biotechnol. Prog.*, 2001**

2) Minimum Network Results – Case 1

(i) Glucose-only uptake



Reaction set reductions are attained by successively eliminating energy producing reactions occurring in

(i) glycolysis

(ii) the TCA cycle

(iii) the pentose phosphate pathway

Proposed MILP Framework:

Contains 11 of 12 lethal gene deletions

FBA Single Gene Deletions:

Identifies 7 of 12 lethal gene deletions

Minimal Reaction Sets Are Not Unique!!

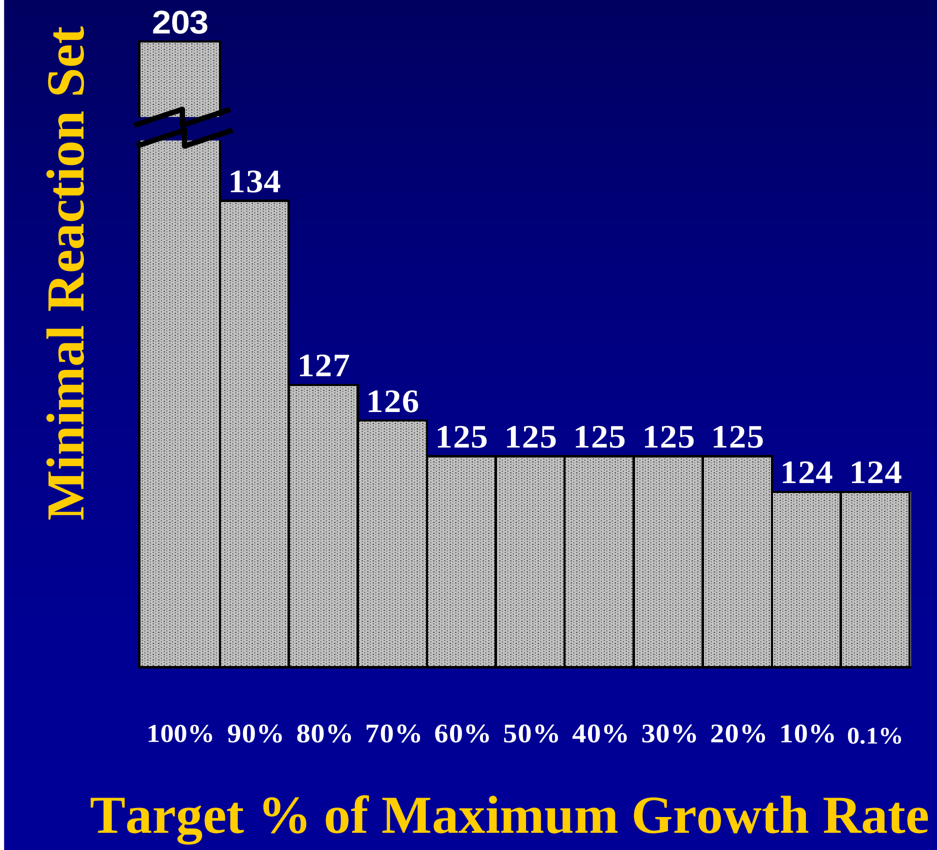
2) Minimum Network Results – Case 2

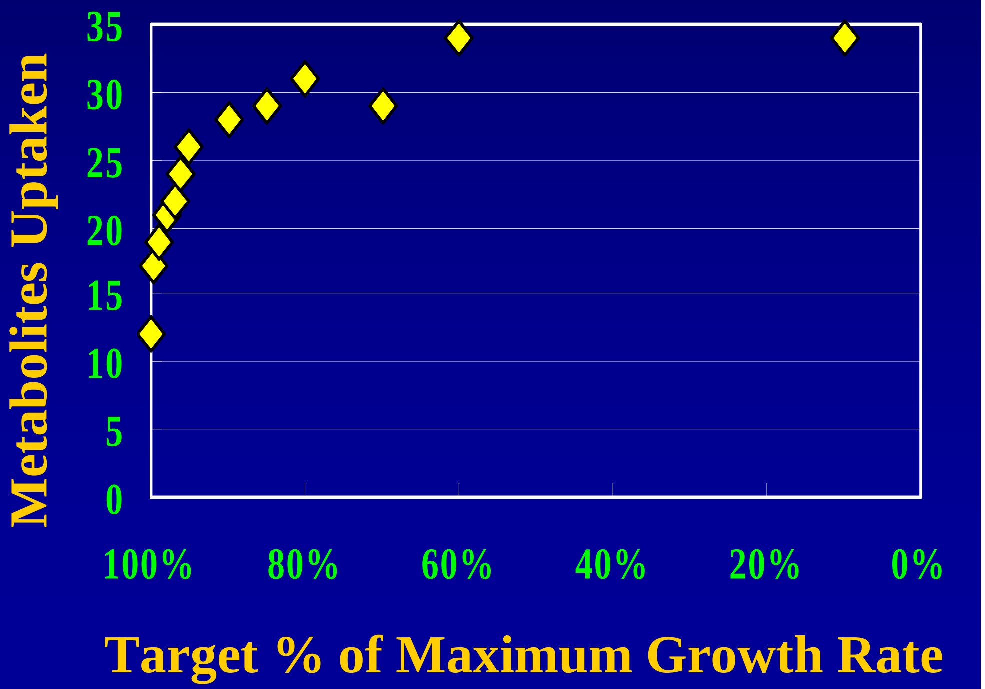
(ii) Multiple organic uptake

Reaction set reductions are generally attained by

(i) importing additional metabolites at successively lower growth rates

(ii) eliminating energy producing reactions from the core pathways

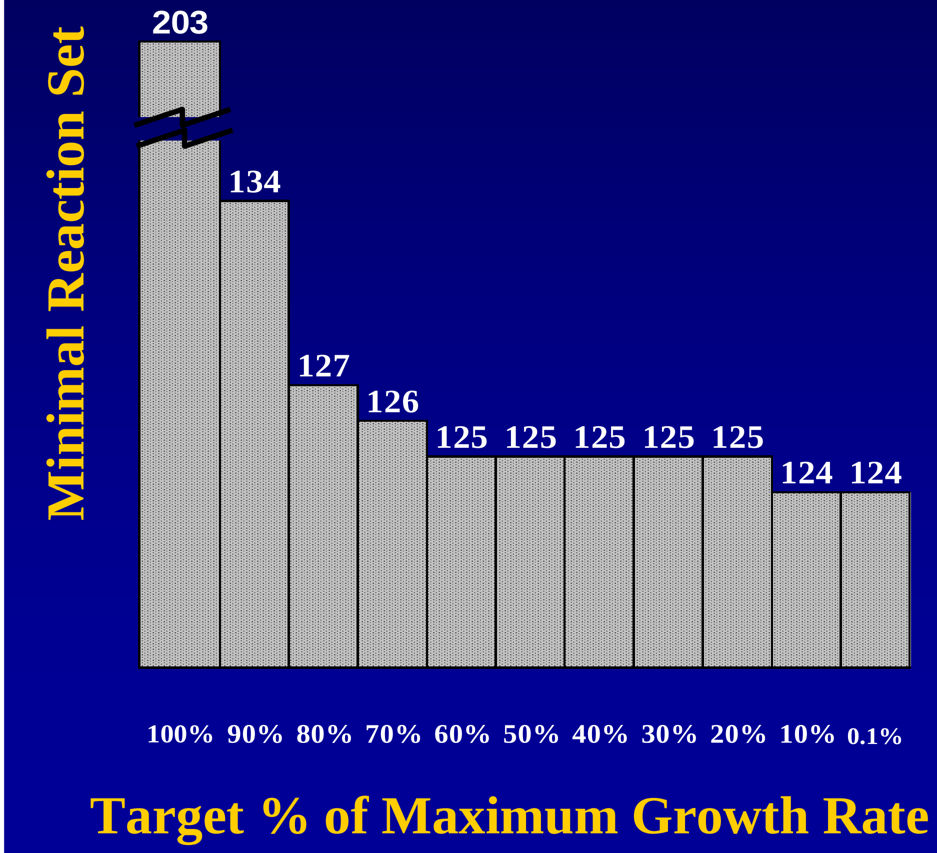




2) *Minimum Network Results – Case 2*

(ii) Multiple organic uptake

Mushegian and Koonin (1996) ⇨ 94 metabolic genes in minimal gene set

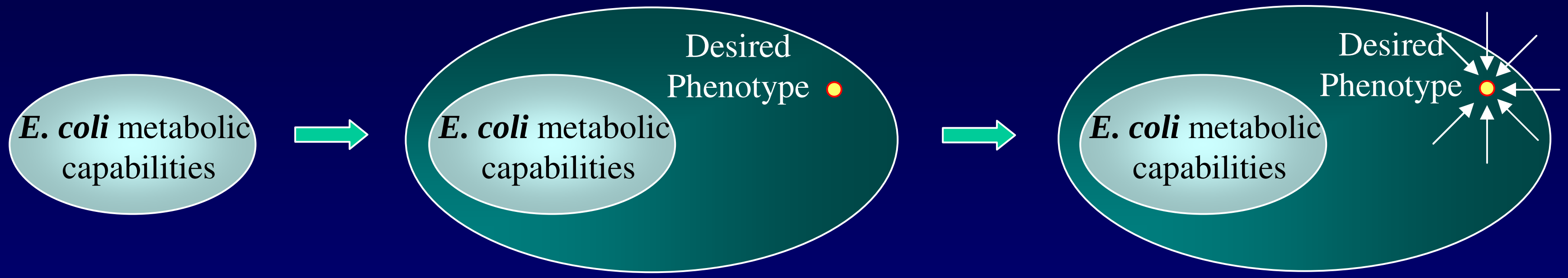


Minimal Reaction Set:

Functional Classification	# rxns
ALA Isomerization	1
Alternative Carbon Source	7
Anaplerotic Reactions	1
ATP Maintenance	1
Biomass Synthesis	1
Cell Envelope Biosynthesis	3
EMP Pathway	5
Lipid A Biosynthesis	9
LPS Sugar Biosynthesis	7
Membrane Lipid Biosynthesis	16
Murein Biosynthesis	10
Pentose Phosphate Pathway	4
Pyrimidine Biosynthesis	1
Respiration	5
Salvage Pathways	17
Transport	36
	124

Presentation Outline

- ❑ Systems biology and the constraints-based modeling approach


E. coli metabolic capabilities
Desired Phenotype
E. coli metabolic capabilities
Desired Phenotype
E. coli metabolic capabilities

- ❑ Pathway discovery and optimization

  How can we systematically select the appropriate set of pathways/genes to recombine into existing production systems?

- ❑ Constraining allowable cellular behavior

  How can we identify gene knockouts that will force biochemical overproduction by coupling it with cell growth?

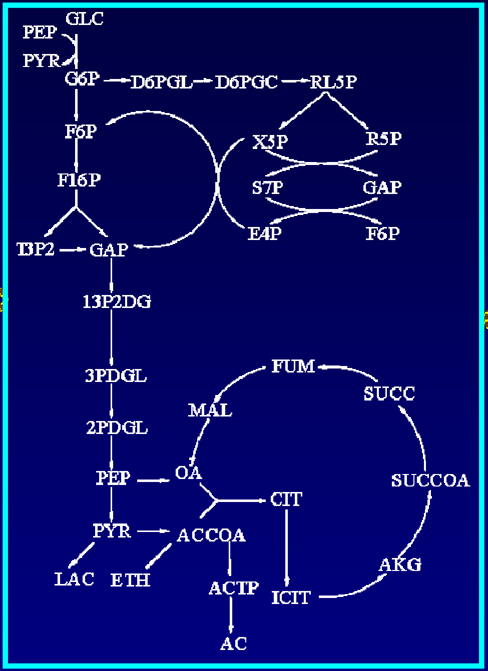
- ❑ Metabolic network structural and topological analysis

  How can we identify multiple metabolic manipulations for producing a desired product and also computationally evaluate of the consequences of potential network modifications ?

Motivating Challenge

Cellular Objective

(*e.g.*,
Max. biomass production
Min. metabolic adjustment
Max. ATP production, etc.)



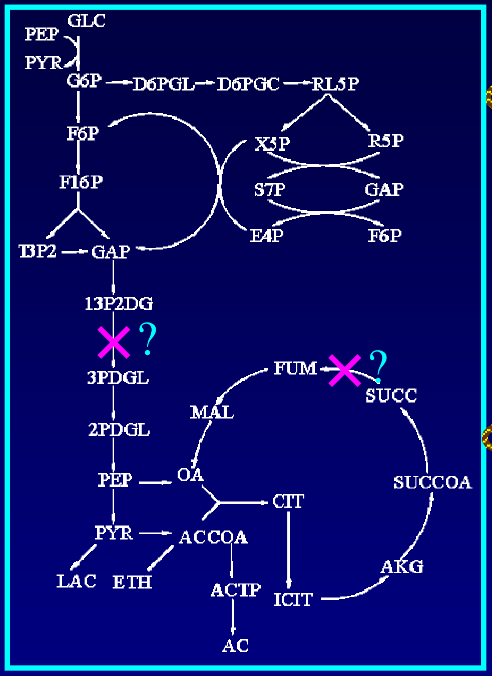
Bioengineering Objective

(*e.g.*,
Max. 1,3-propanediol yield
Max. lactate yield, etc.)

Maximum Theoretical Yields (mol/mol glucose)		Experimental Yields (mol/mol glucose)	
Acetate	2.40	Acetate	0.81
Ethanol	2.00	Ethanol	0.84
Formate	2.78	Formate	1.16
Lactate	2.00	Lactate	0.10
Succinate	1.71	Succinate	0.34

(Stokes, *J. Bacteriology*, 1949)

Motivating Challenge



Bioengineering Objective

(*e.g.*,
Max. biomass production
Min. metabolic adjustment
Max. ATP production, etc.)

Cellular Objective

(*e.g.*,
Max. 1,3-propanediol yield
Max. lactate yield, etc.)

Maximum Theoretical Yields (mol/mol glucose)

Acetate	2.40
Ethanol	2.00
Formate	2.78
Lactate	2.00
Succinate	1.71

Experimental Yields (mol/mol glucose)

Acetate	0.81
Ethanol	0.84
Formate	1.16
Lactate	0.10
Succinate	0.34

(Stokes, *J. Bacteriology*, 1949)

OptKnock Bilevel Optimization Framework

Outer Problem:

adjust ***knockouts***

→ optimize bioeng. objective

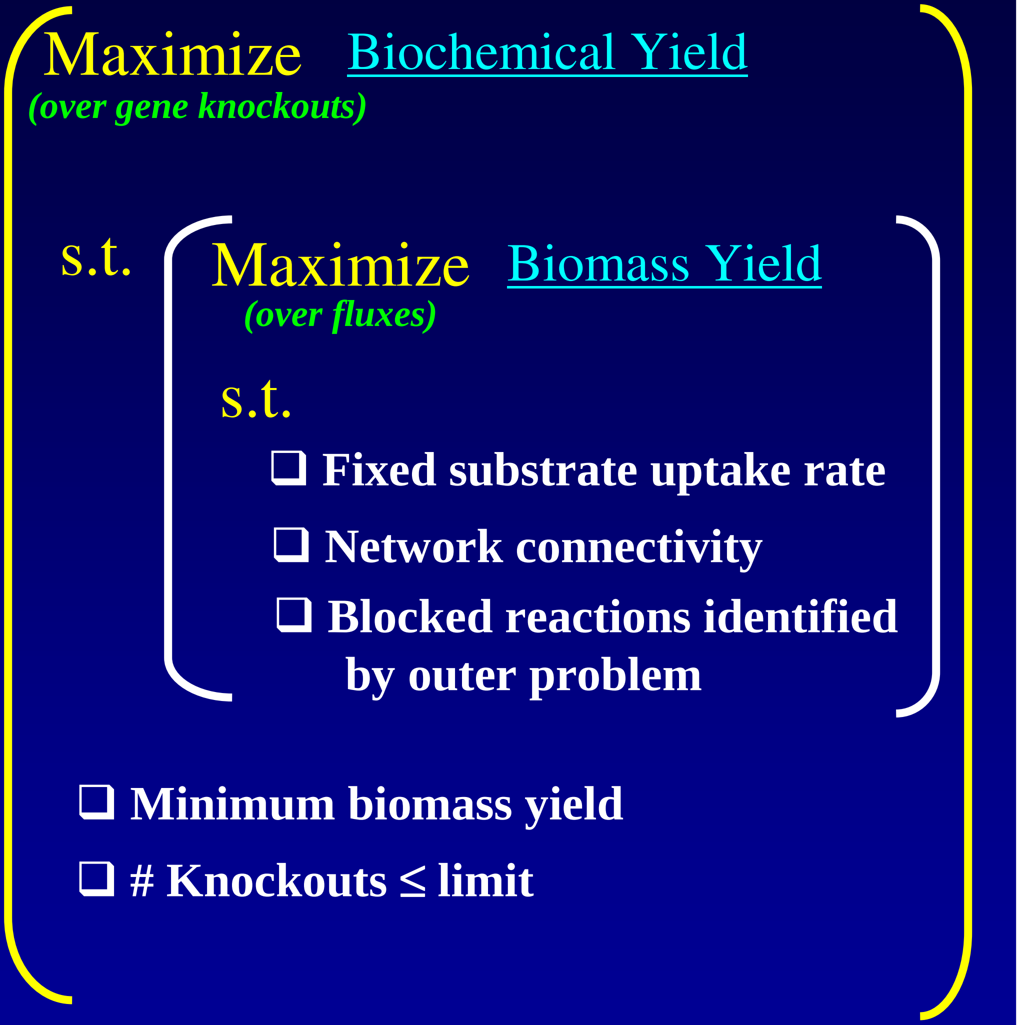
- Max. 1,3-propanediol yield
- Max. lactate yield

Inner Problem:

adjust ***reaction fluxes***

→ optimize cellular objective

- Max. biomass yield
- Min. metabolic adjustment
- Max. ATP yield



Maximize Biochemical Yield
(over gene knockouts)

s.t.

Maximize Biomass Yield
(over fluxes)

s.t.

❑ **Fixed substrate uptake rate**

❑ **Network connectivity**

❑ **Blocked reactions identified by outer problem**

❑ **Minimum biomass yield**

❑ **# Knockouts ≤ limit**

- Burgard, A.P., Pharkya, P., and C.D. Maranas (2003), “OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization,” *Biotechnology and Bioengineering*, 84, 647-657.
- Pharkya, P., Burgard, A.P., and C.D. Maranas (2003), “Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock,” *Biotechnology and Bioengineering*, 84, 887-899.

LP Duality Theory

Z_{DUAL}

DUAL

Minimize $_{u_i,\, g}$ $\quad Z_{DUAL} = uptake \cdot g$

subject to $\quad \sum_i u_i S_{i,GLC} + g = 0$

$$\sum_i u_i S_{i,Biomass} \geq 1$$

$$\sum_i u_i S_{ij} \geq 0$$

Optimal solution if and only if:

$$Z_{DUAL} = Z_{PRIMAL}$$

PRIMAL (Inner problem)

Maximize $_{v_j}$ $\quad Z_{PRIMAL} = v_{Biomass}$

subject to $\quad \sum_j S_{ij} v_j = 0 \longleftarrow u_i$

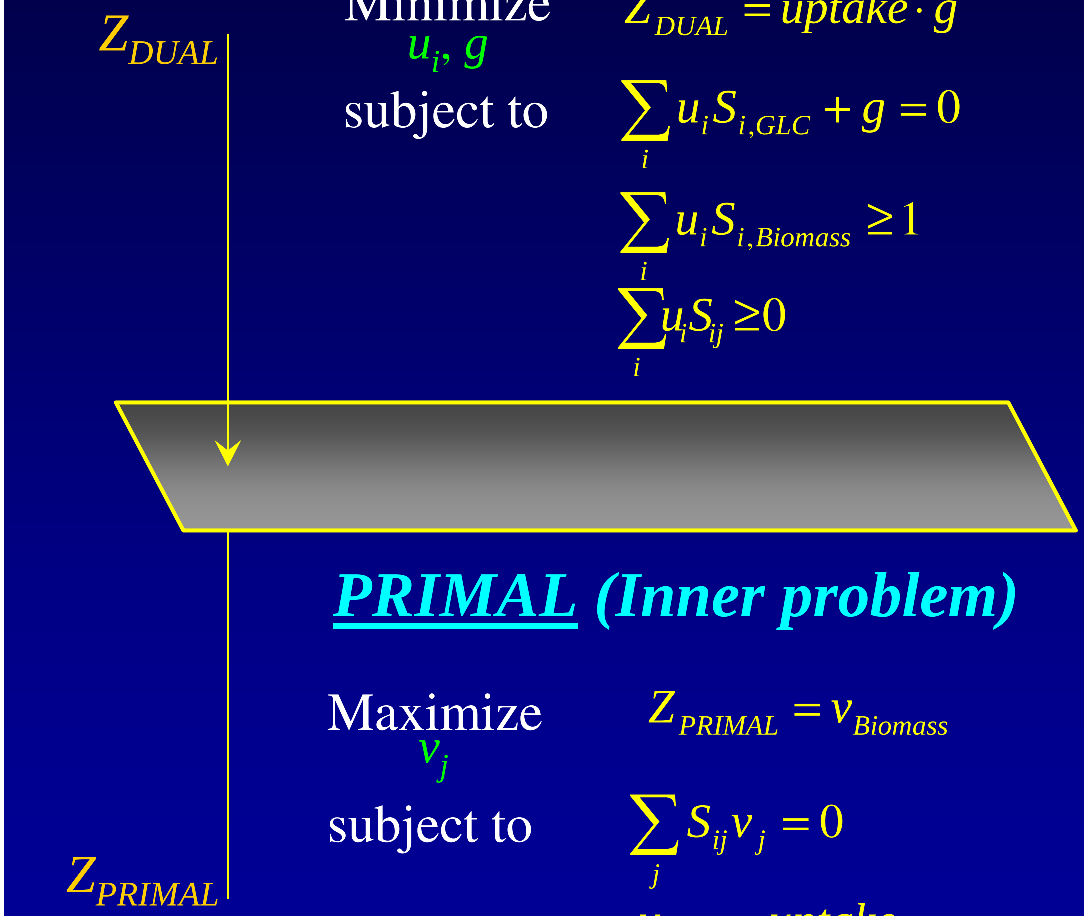
$$v_{GLC} = uptake \longleftarrow g$$

$$v_j \geq 0$$

u_i, g: multipliers

Z_{PRIMAL}

LP Duality Theory



DUAL

Minimize $\underset{u_i,\, g}{}$ $\quad Z_{DUAL} = uptake \cdot g$

subject to

$$\sum_i u_i S_{i,GLC} + g = 0$$

$$\sum_i u_i S_{i,Biomass} \ge 1$$

$$\sum_i u_i S_{ij} \ge 0$$

PRIMAL (Inner problem)

Maximize $\underset{v_j}{}$ $\quad Z_{PRIMAL} = v_{Biomass}$

subject to

$$\sum_j S_{ij} v_j = 0$$

$$v_{GLC} = uptake$$

$$v_j \ge 0$$

MILP problem

Maximize $\underset{y_j,\, u_i,\, g,\, v_j}{}$ $\quad v_{Product}$

subject to

$$uptake \cdot g = v_{Biomass}$$

Dual

$$\sum_i u_i S_{i,Biomass} \ge 1$$

$$\sum_i u_i S_{i,GLC} + g = 0$$

$$\sum_i u_i S_{ij} \ge 0$$

Primal

$$\sum_j S_{ij} v_j = 0$$

$$v_{GLC} = uptake$$

$$0 \le v_j \le v_j^{\max} \cdot y_j$$

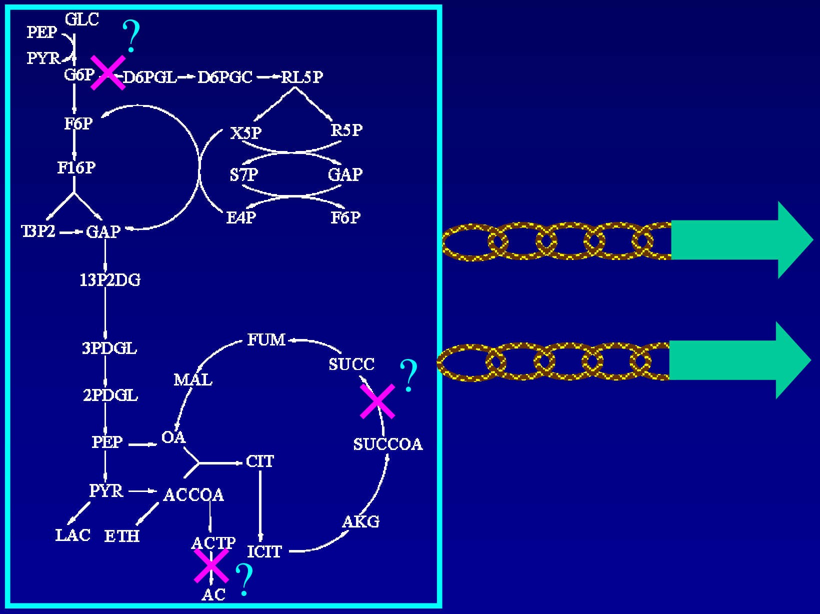
$$\sum_j (1 - y_j) \le \#\ of\ knockouts$$

$$y_j \in \{0,1\}$$

Optimal Gene Knockout Identification

Questions:

- ❑ Identify single, double, triple, and quadruple knockout strategies?
- ❑ Characterize allowable envelope of biomass vs. biochemical production?



Growth

Biochemical Production

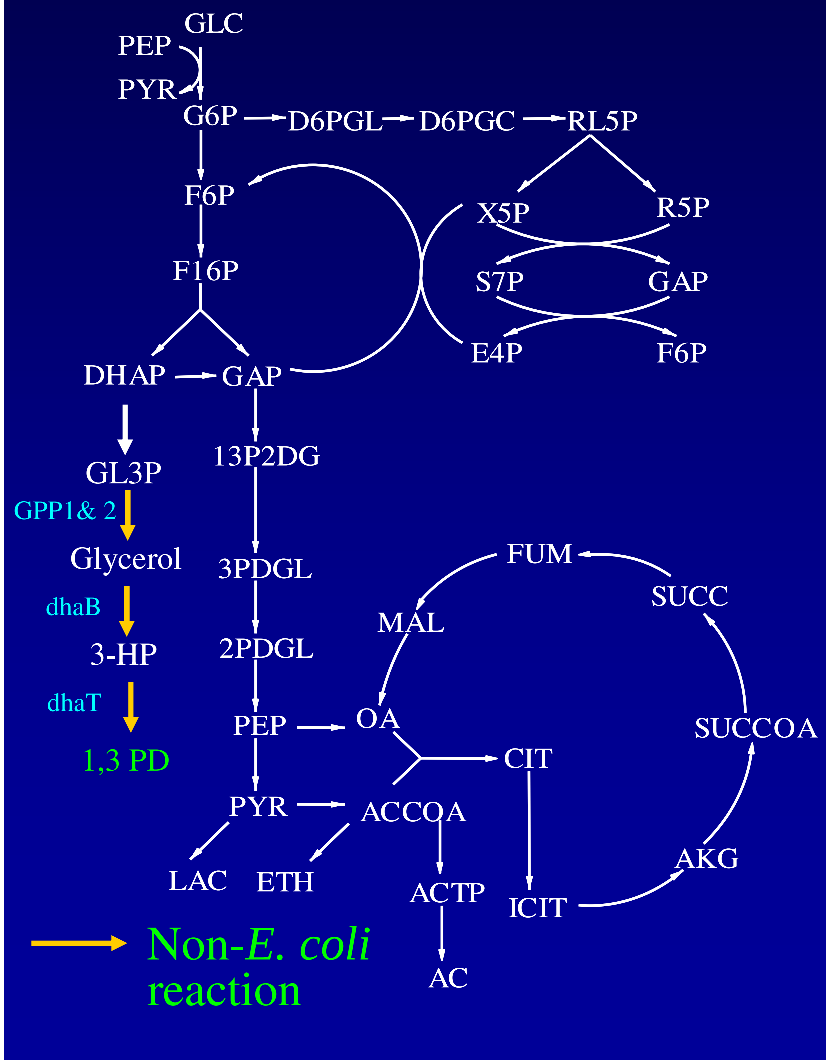
Case Studies:

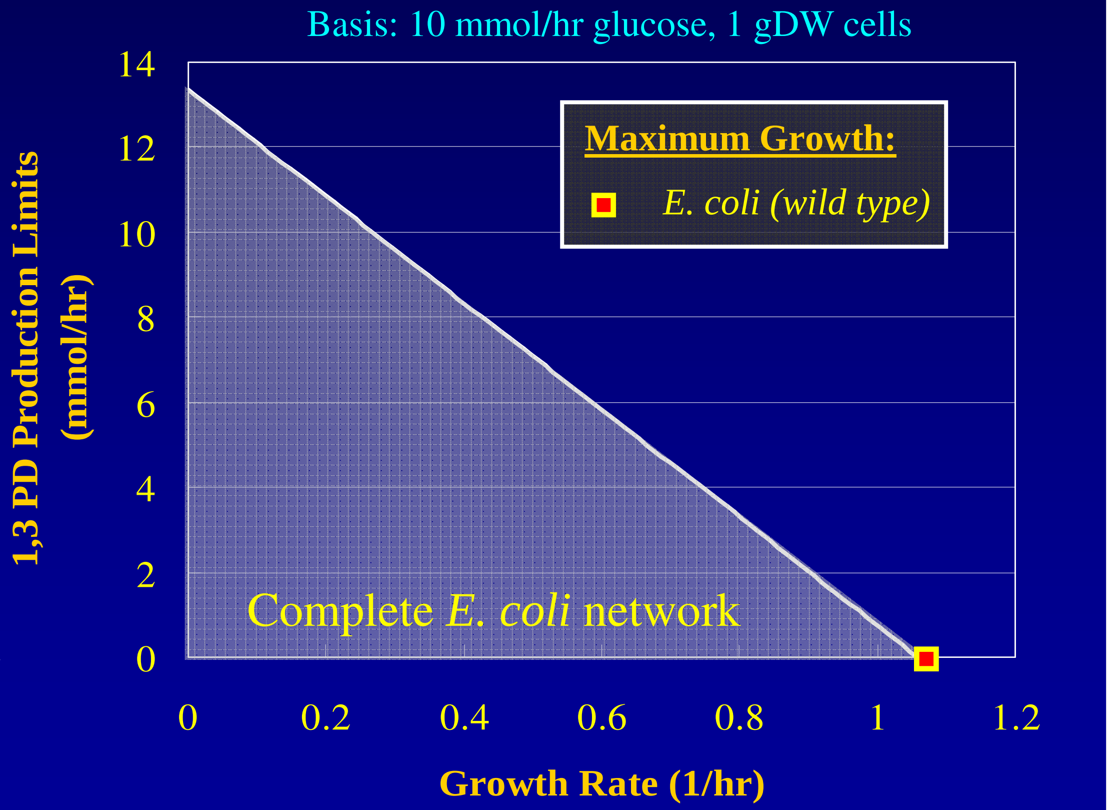
(1) 1,3 Propanediol (2) Lactate

1,3 PD Overproduction

Non-*E. coli* genes/enzymes:

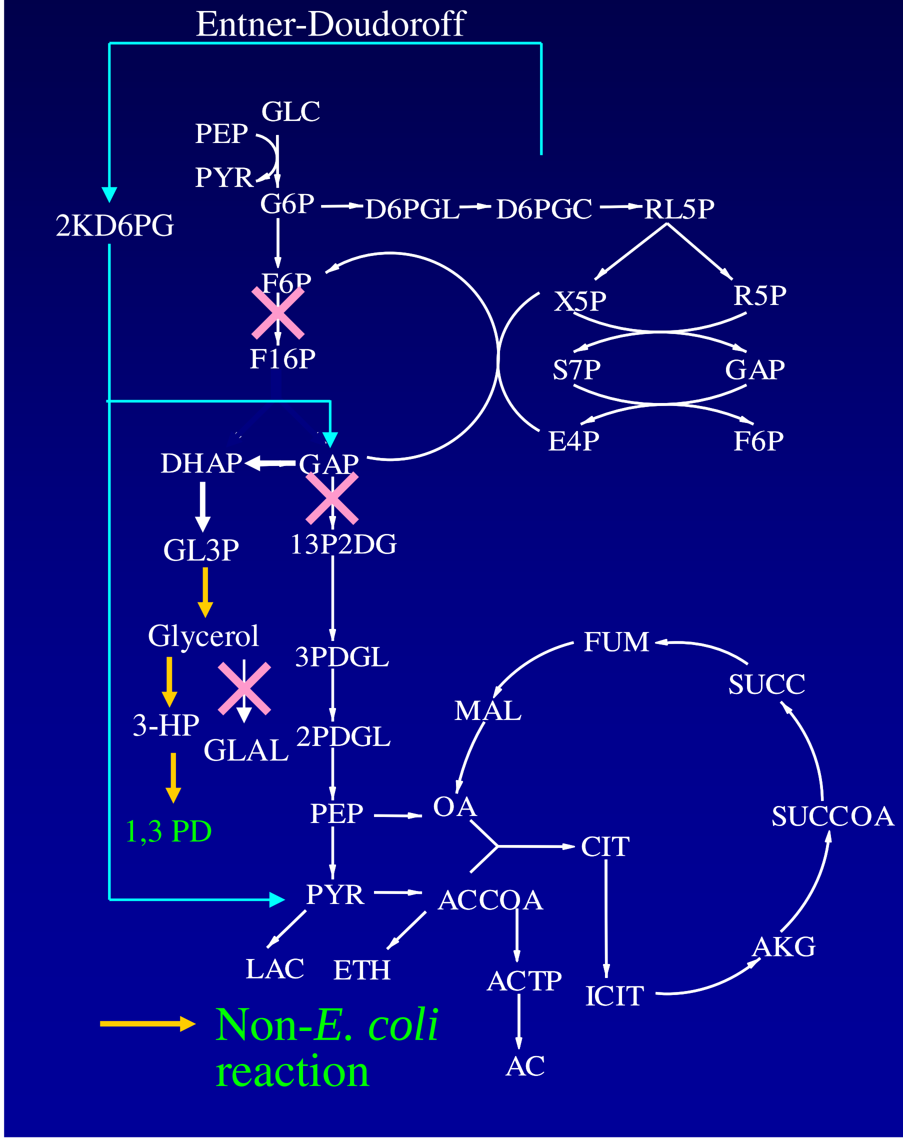
GPP1&2:	glycerol-3-phosphatase	*Saccharomyces cerevisiae*
dhaB:	glycerol dehydratase	*Klebsiella pneumoniae*
dhaT:	1,3 PD oxidoreductase	*Klebsiella pneumoniae*





1,3 PD Overproducing Mutants

Mutant A:
(1) Aldehyde dehydrogenase (adhC)
(2) Phosphoglycerate kinase (pgk) or Glyceraldehyde-3-phosphate dehydrogenase (gapA, gapC1C2)
(3) Fructose-1,6-bisphosphatase (fbp) or Fructose-1,6-bisphosphate aldolase (fba)

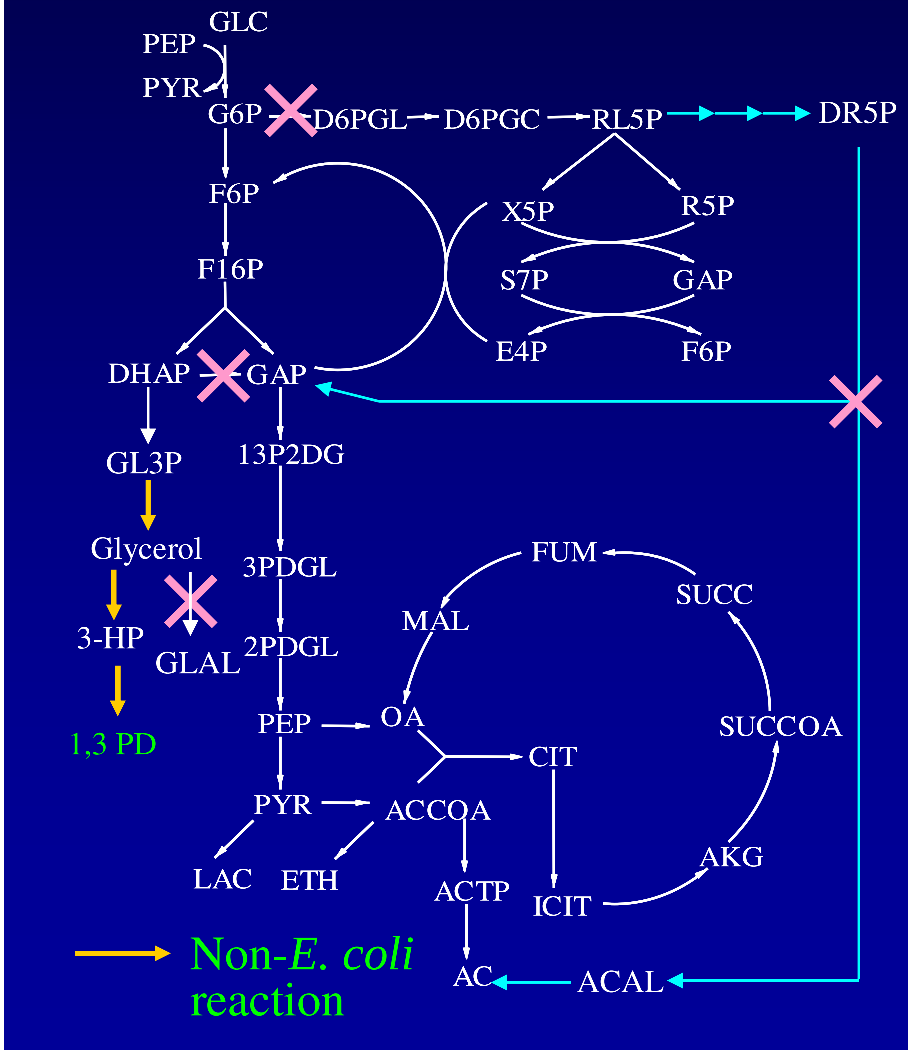


Basis: 10 mmol/hr glucose, 1 gDW cells

"Wild" type:	Maximum Biomass :	1.05 hr^{-1}
	1,3 PD:	0.00 mmol/hr
Mutant A:	Maximum Biomass :	0.21 hr^{-1}
	1,3 PD:	9.66 mmol/hr

1,3 PD Overproducing Mutants

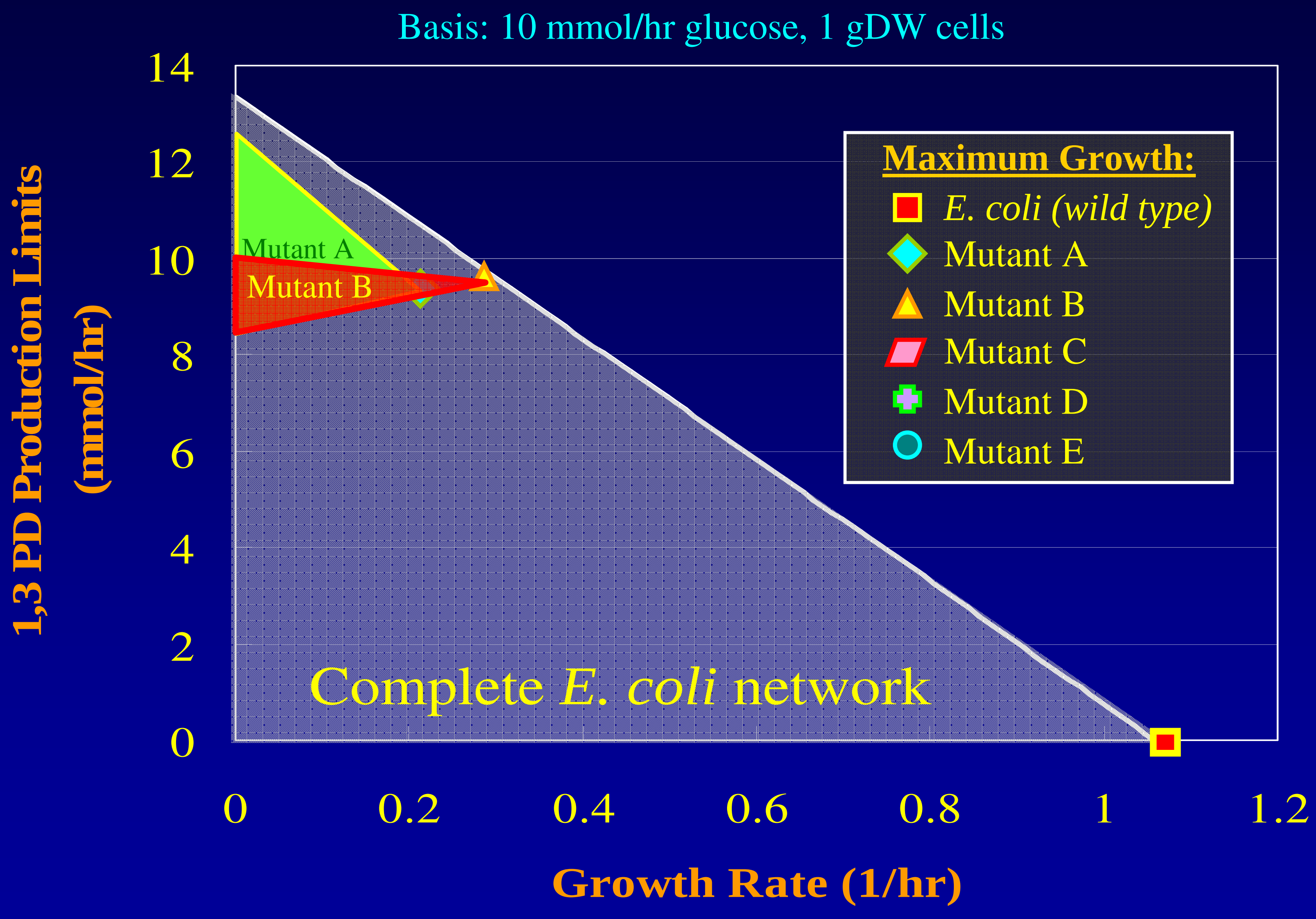
Mutant B:
(1) Aldehyde dehydrogenase (adhC)
(2) Triose phosphate isomerase (tpiA)
(3) Glucose 6-phosphate-1-dehydrogenase (zwf) or 6-Phosphogluconolactonase (pgl)
(4) Deoxyribose-phosphate aldolase (deoC)



Basis: 10 mmol/hr glucose, 1 gDW cells

"Wild" type:	Maximum Biomass :	$1.05\ \mathrm{hr}^{-1}$
	1,3 PD :	0.00 mmol/hr
Mutant A:	Maximum Biomass :	$0.21\ \mathrm{hr}^{-1}$
	1,3 PD :	9.66 mmol/hr
Mutant B:	Maximum Biomass :	$0.29\ \mathrm{hr}^{-1}$
	1,3 PD :	9.67 mmol/hr
Mutant C:	Maximum Biomass :	$0.11\ \mathrm{hr}^{-1}$
	1,3 PD :	9.84 mmol/hr
Mutant D:	Maximum Biomass :	$0.14\ \mathrm{hr}^{-1}$
	1,3 PD :	9.78 mmol/hr
Mutant E:	Maximum Biomass :	$0.16\ \mathrm{hr}^{-1}$
	1,3 PD :	9.75 mmol/hr

1,3 PD Mutant Characterization



Driving Microbial Strain Design

Bioengineering Research Partnership (BRP) team

OptKnock Predictions:	Costas Maranas	***The Pennsylvania State University***
Strain Construction:	Fred Blattner	***University of Wisconsin***
Adaptive Evolution:	Bernhard Palsson	***University of California, San Diego***
HT Characterization:	Jay Keasling	***University of California, Berkeley***
Data/Software Integration:	Christophe Schilling	***Genomatica, Inc.***

Overall Objective

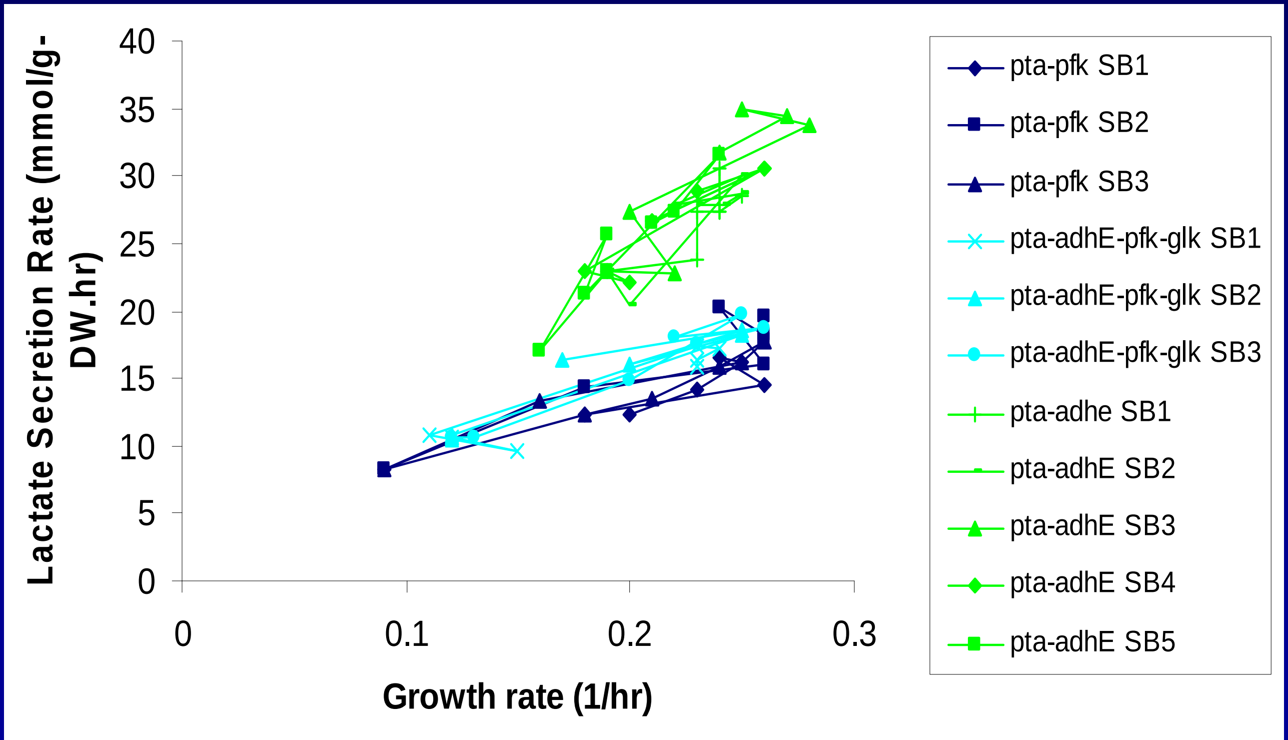
Improved production of lactate, succinate, and terpenes in *E. coli*

Lactate Mutant Experimentation

***Blattner Lab:* Strain Construction**
***Palsson Lab:* Adaptive Evolution**

60 days ~ 500 generations

Three designs constructed:
1) pta-adhE (5 strains evolved)
2) pta-pfk (3 strains evolved)
3) pta-adhE-pfk-glk (3 strains evolved)

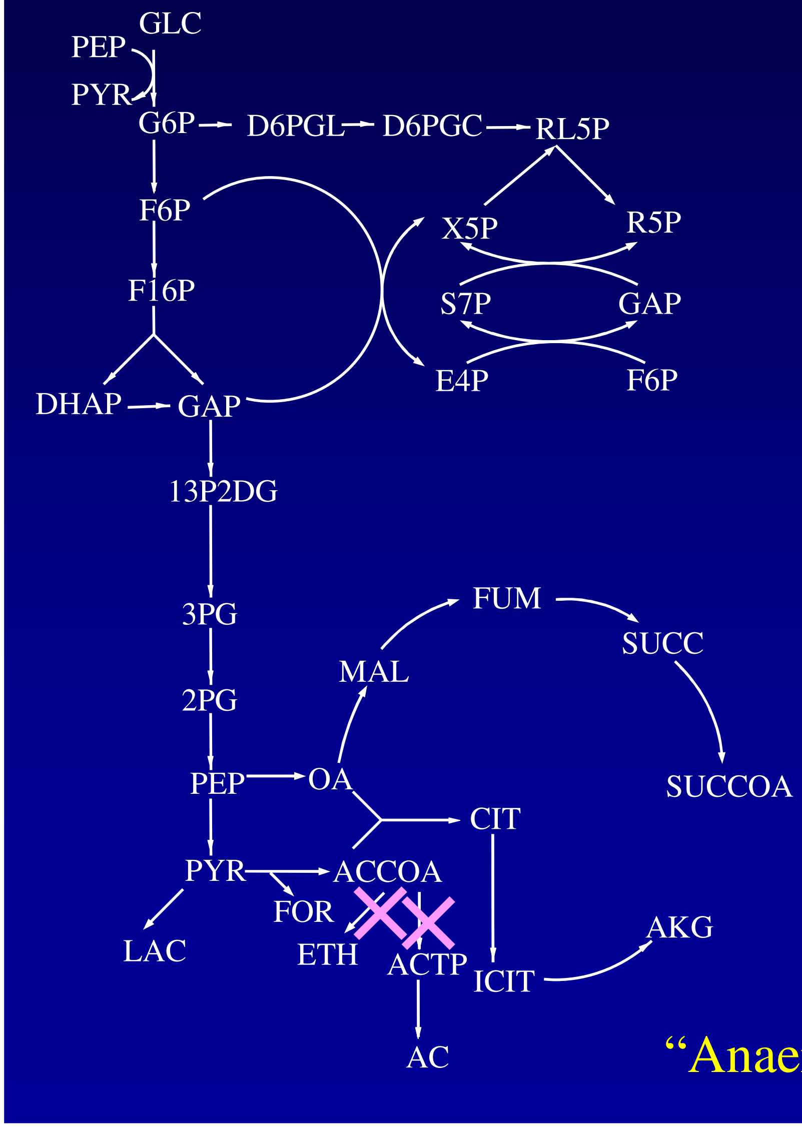


Increased glucose uptake rates or increased lactate yields ?

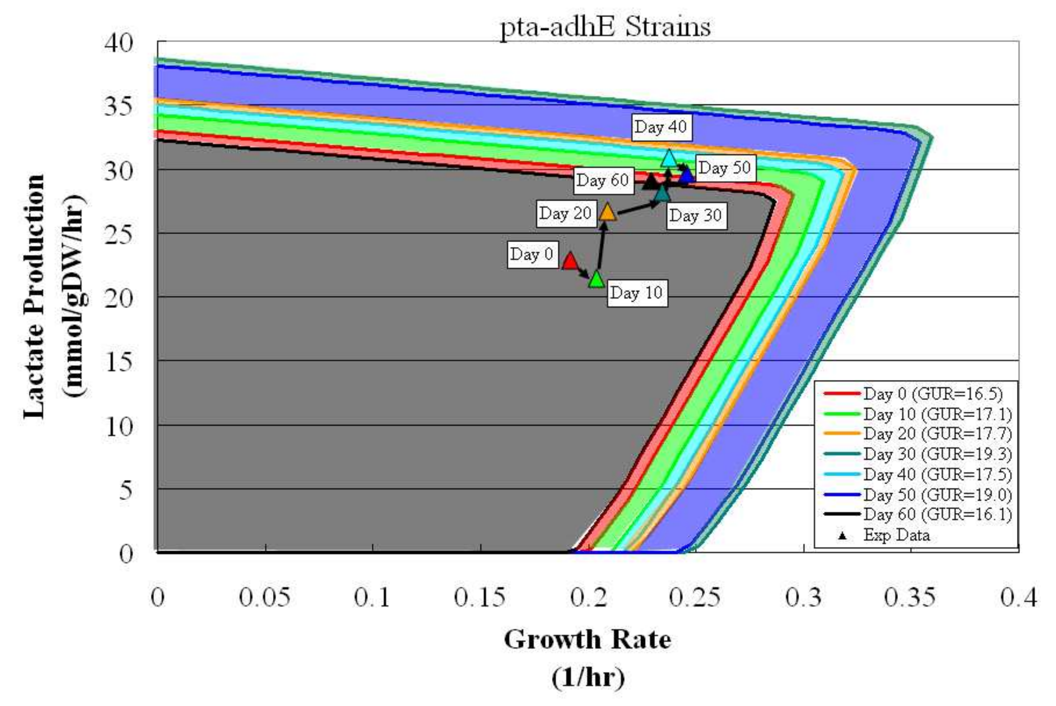
Lactate Overproducing Mutant

Knockouts: (1) Acetaldehyde dehydrogenase (adhE)

(2) Phosphotransacetylase (pta)



***Blattner Lab:** Strain Construction*
***Palsson Lab:** Adaptive Evolution*



"Anaerobic Conditions"

The OptStrain Procedure

Step 1: Compilation and curation of the Universal database.

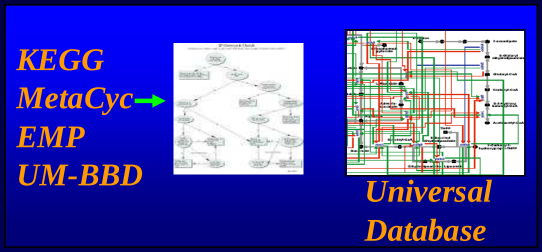
Step 2: Determination of the maximum yield of the desired product from an optimal substrate choice.

Step 3: Minimizing reliance on non-native reactions while satisfying optimal performance criteria.

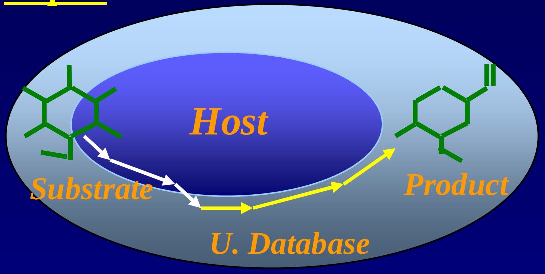
Step 4: Optimal gene deletion determination for coupling biomass production to biochemical formation.

Pharkya et al. (2004), Genome Research, 14(11), 2367-2376.

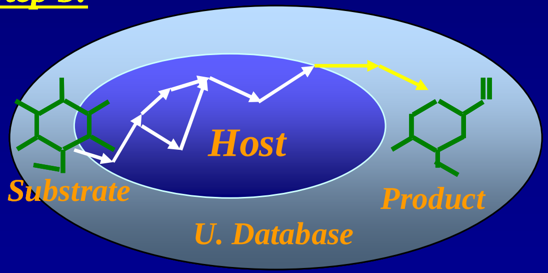
Step 1:



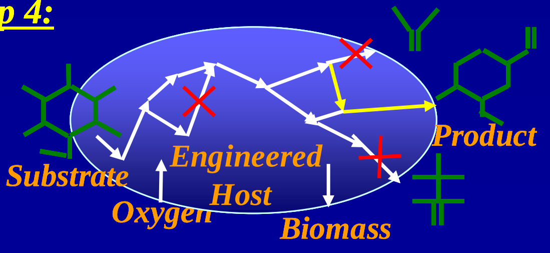
Step 2:



Step 3:



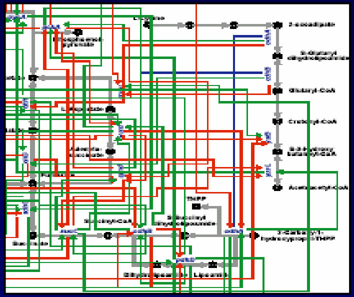
Step 4:



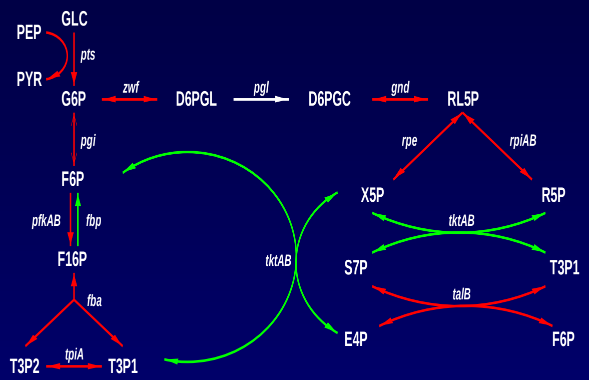
OptStrain - Step 1 : Curation of the database

(a) Database compilation

KEGG
MetaCyc
EMP
UM-BBD



Universal database
(~ 5,700 reactions)



(b) Database curation

- *Perl scripts convert the reactions to a format readable by the GAMS optimization environment.*

```
ENTRY       R00013
NAME        Glyoxylate carboxy-lyase (dimerizing)
DEFINITION  2 Glyoxylate <=> 2-Hydroxy-3-oxopropanoate + CO2
EQUATION    2 C00048 <=> C01146 + C00011
PATHWAY     PATH: RN00630  Glyoxylate and dicarboxylate metabolism
ENZYME      4.1.1.47
///
```

```
S('47','13')= -2;
S('921','13')= 1;
S('11','13')= 1;
```

OptStrain - Step 1 : Curation of the database

- *Parse the number of atoms of each element in each compound.*

e.g. for glyoxylate C(‘47’) = 6, H(‘47’) = 14, N(‘47’) = 2, O(‘47’) = 2.

- *Elimination of elementally unbalanced reactions.*

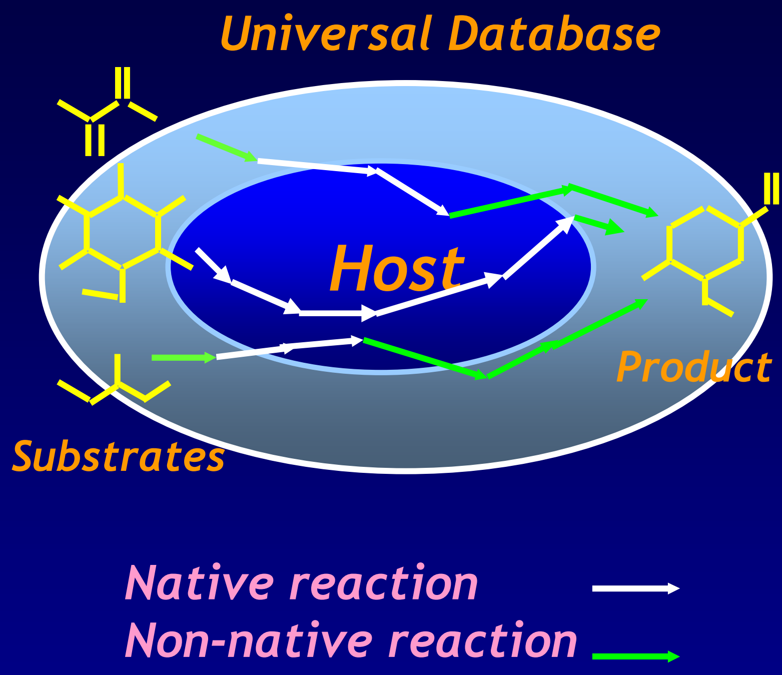
sn-Glycerol 3-phosphate + Acyl-CoA ↔ 1-Acyl-sn-glycerol 3-phosphate + CoA.

- *Exclusion of compounds with ambiguous (e.g. compounds with unspecified number of alkyl units R) or unspecified number of repeat units. e.g.* trans-2-Enoyl-CoA – $C_{25}H_{39}N_7O_{17}P_3S(CH_2)_n$.

- *Elimination of compounds with no chemical formulae.*

Database available as **Supplementary information on Genome Research** website :
http://www.genome.org/current.shtml
and on our webpage: **http://fenske.che.psu.edu/Faculty/CMaranas/pubs.html**

Step 2 : Determination of the maximum yield



Max (v_j)

$$MW_i \cdot \sum_j S_{ij} v_j \quad , \quad i = P$$

Maximize product yield

s.t.

$$\sum_{j=1}^{M} S_{ij} v_j \geq 0,$$

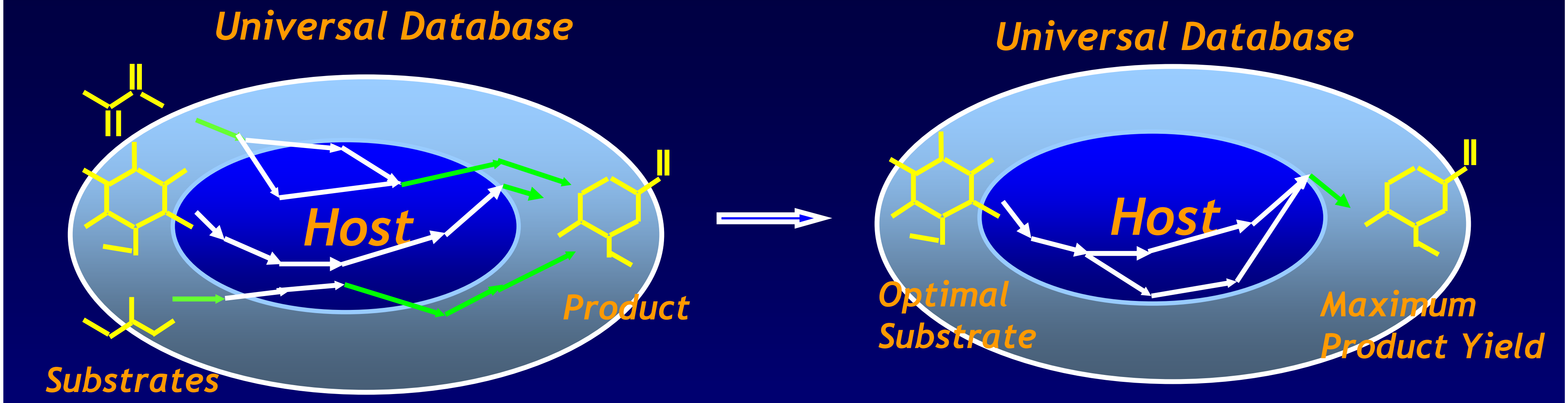
Allows secretion of metabolites

$$\sum_{i \in \Re} \left(MW_i \cdot \sum_{j=1}^{M} S_{ij} v_j \right) = -1$$

Total substrate uptake scaled to 1 unit

- Universal database of reactions
- Different substrates evaluated
- Maximum yield determination

Step 3 : Minimizing non-native reactions in host



Minimize ∑ Non-native reactions (MILP)

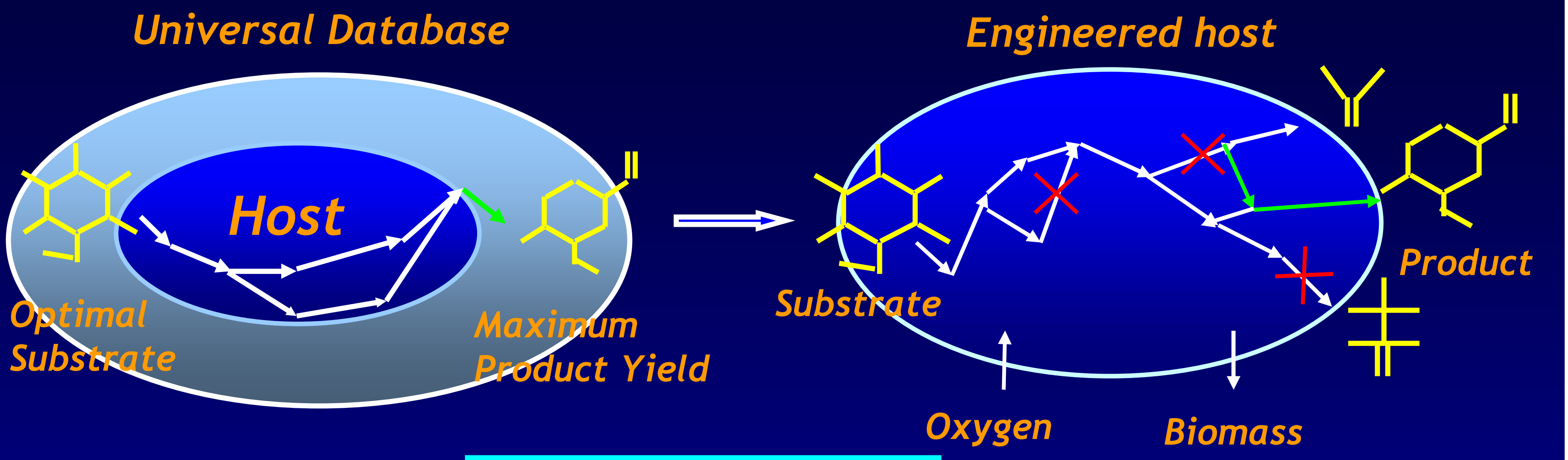
subject to

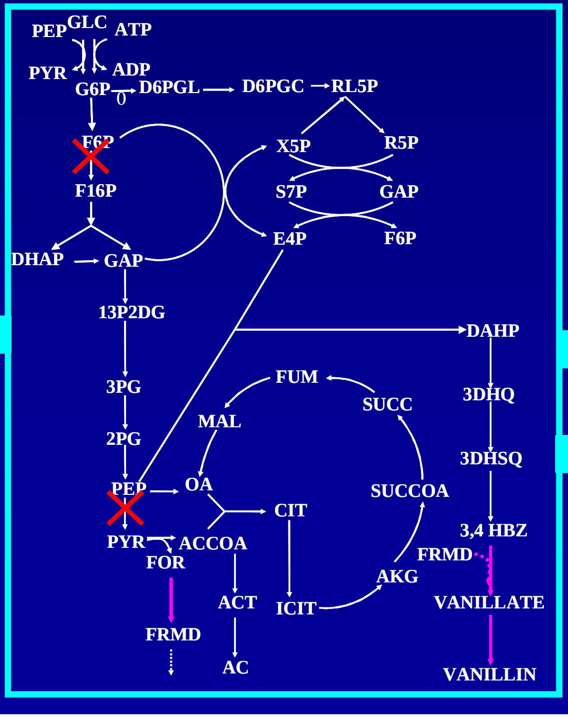
Stoichiometric mass balances

Optimal substrate uptake

Max. product yield formation

Step 4 : Optimum gene deletion determination





Bilevel Optimization framework OptKnock

Cellular Objective

(*e.g.*,
Max. biomass production
Min. metabolic adjustment
Max. ATP production, etc.)

Bioengineering Objective

(*e.g.*,
Max. lycopene yield
Max. lactate yield, etc.)

Cellular Objective

Vanillin overproduction in E. coli

Step 2:

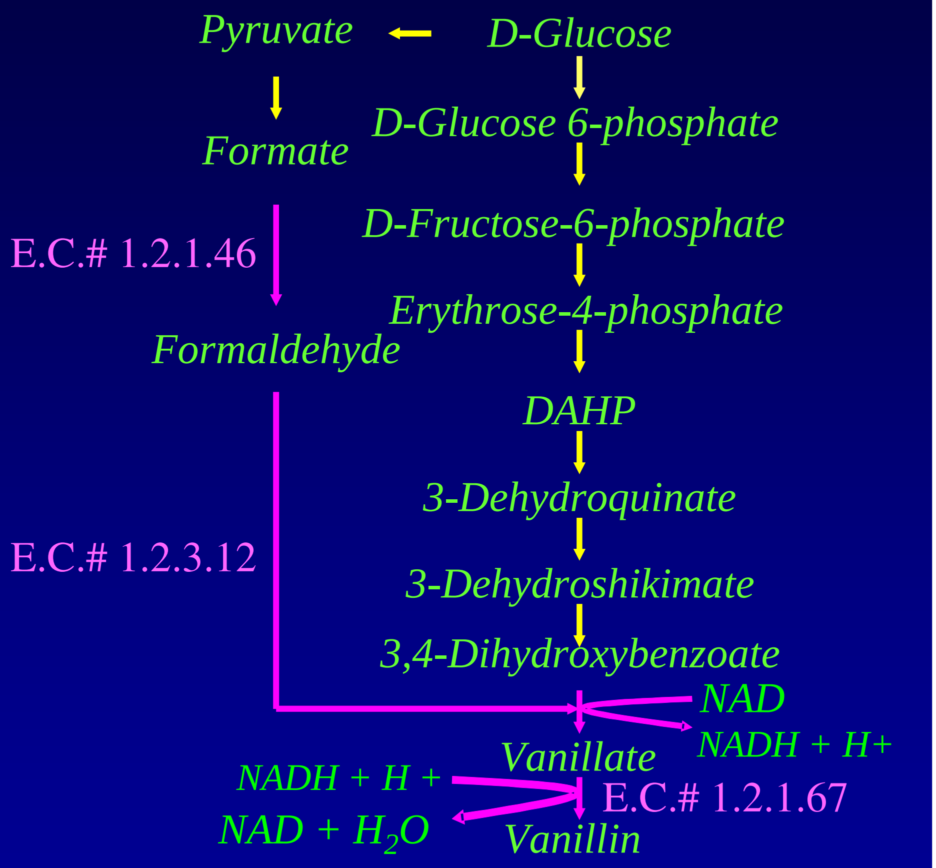
- Maximum theoretical vanillin yield : 0.63 g/g glucose.

Step 3:

No. of non-native functionalities : 3

- Alternative pathways found.
- Theoretical yields in the augmented *E. coli* network almost identical for these gene addition strategies.
- One strategy :

(i) E.C.# 1.2.1.46
(ii) E.C.# 1.2.3.12
(iii) E.C.# 1.2.1.67

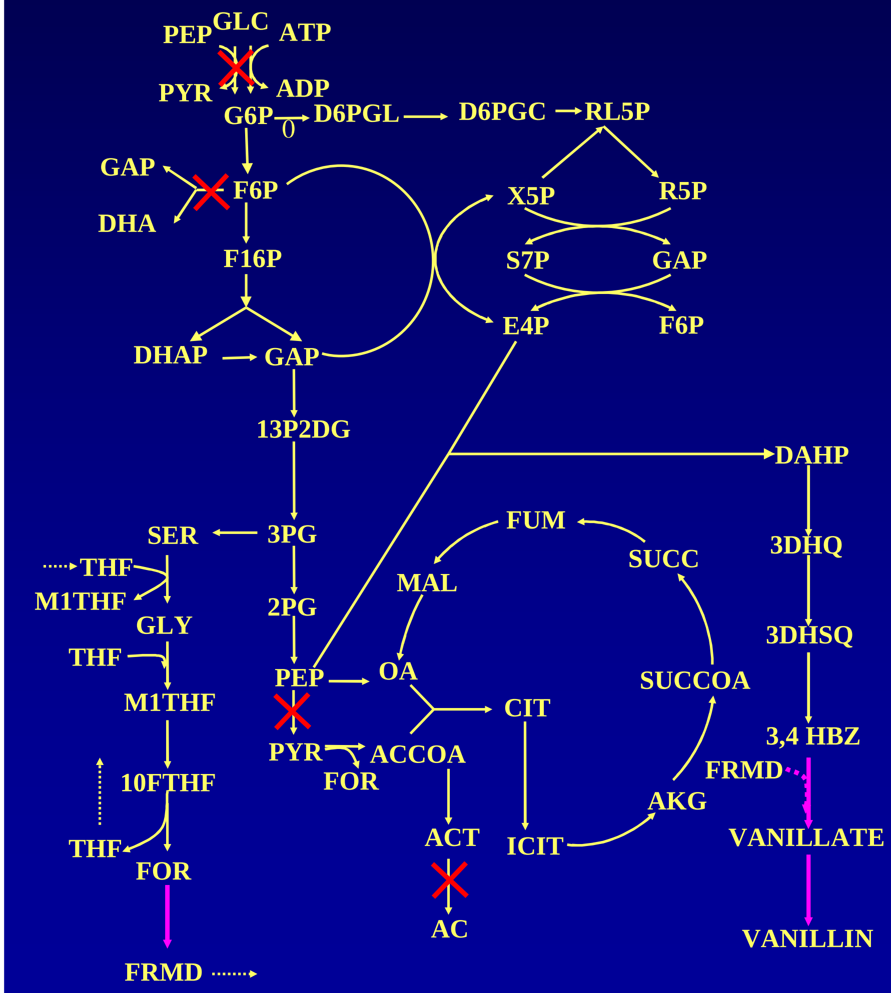


Reaction addition steps same as those identified in Li and Frost, 1998.

Vanillin overproduction in E. coli

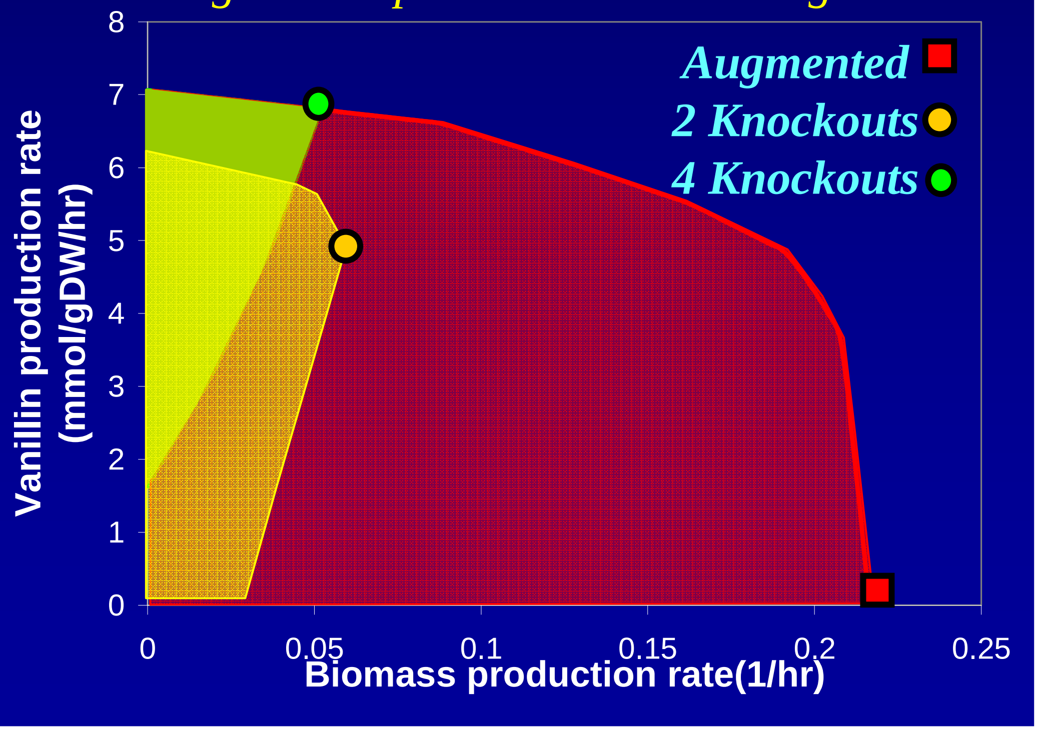
Step 4: Identification of knockout strategies – Quadruple strategy

(i) Acetate kinase
(ii) Pyruvate kinase
(iii) PTS transport
(iv) Fructose-6-phosphate aldolase



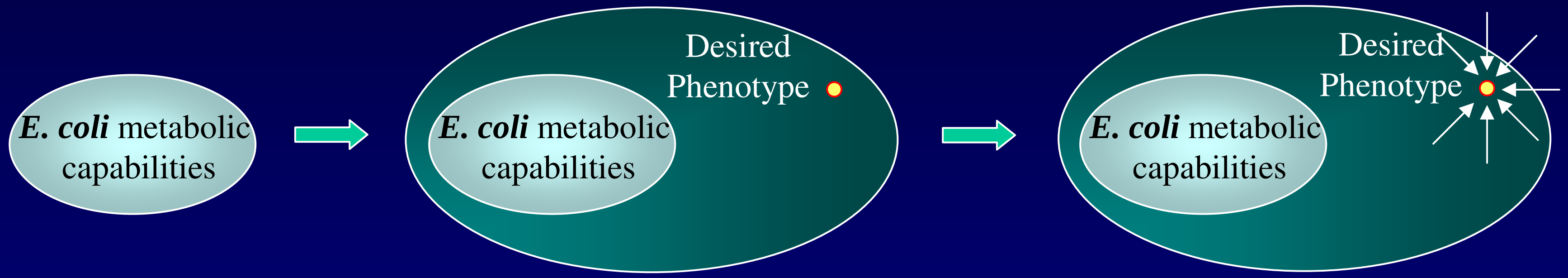
Mode of growth is anaerobic in all cases!

Basis glucose uptake rate: 10 mmol/ gDW/ hr



Presentation Outline

❑ Systems biology and the constraints-based modeling approach



E. coli metabolic capabilities

E. coli metabolic capabilities

Desired Phenotype

E. coli metabolic capabilities

Desired Phenotype

❑ Pathway discovery and optimization

How can we systematically select the appropriate set of pathways/genes to recombine into existing production systems?

❑ Constraining allowable cellular behavior

How can we identify gene knockouts that will force biochemical overproduction by coupling it with cell growth?

❑ Metabolic network structural and topological analysis

How can we identify multiple metabolic manipulations for producing a desired product and also computationally evaluate of the consequences of potential network modifications ?

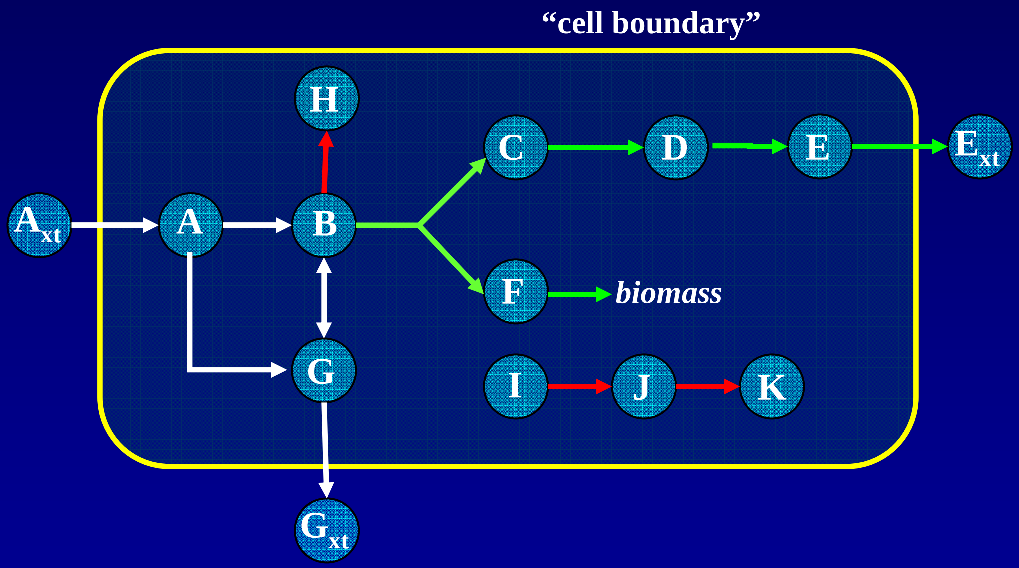
Topological Network Properties

Blocked Reactions

All reactions that cannot carry flux under SS conditions

Enzyme Subsets

Reactions that are always simultaneously utilized in the same ratio under SS conditions



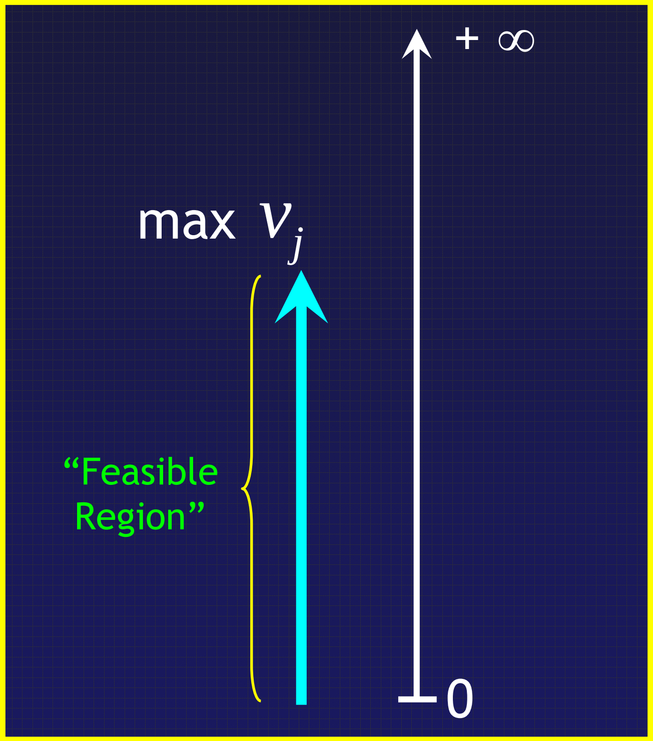
Convex Analysis-based Methods

- Elementary Modes
- Extreme Pathways

Application to Complete Genome-scale Networks

COMBINATORIAL EXPLOSION

Identifying Blocked Reactions



i = *metabolites*
j = *reactions*

maximize v_j
subject to

$$\sum_{j=1}^{M} S_{ij} v_j = 0$$

$$v_{uptake} \leq v_{uptake\_max}$$

$$v_j \geq 0$$

Reversible reactions: $v = v^f - v^b$ where $v^f, v^b \geq 0$

Solve series of linear programs for all reactions j.

If the maximum value of a particular flux is equal to zero, then the reaction is blocked.

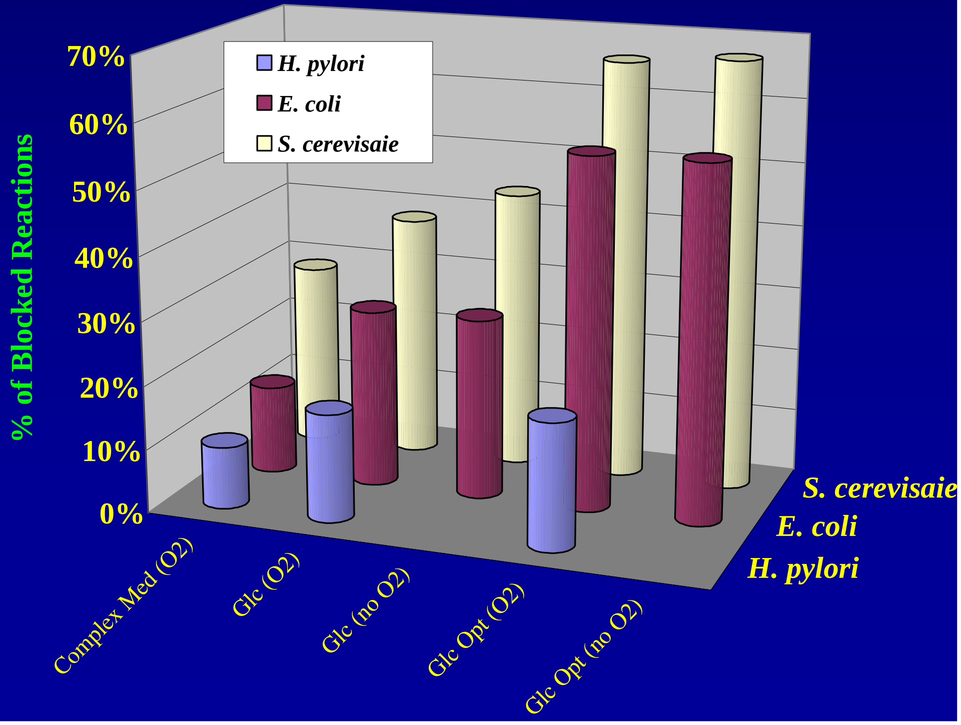
Blocked Reaction Results

Growth Condition	*H. pylori* 389 rxns	*E. coli* 740 rxns	*S. cerevisiae* 1173 rxns
	number of blocked reactions		
Universal Media (Aerobic)	38	103	338
Glucose (Aerobic)	66	207	450
Glucose (Anaerobic)		210	515
Glucose Optimal (Aerobic)	77	408	774
Glucose Optimal (Anaerobic)		410	789

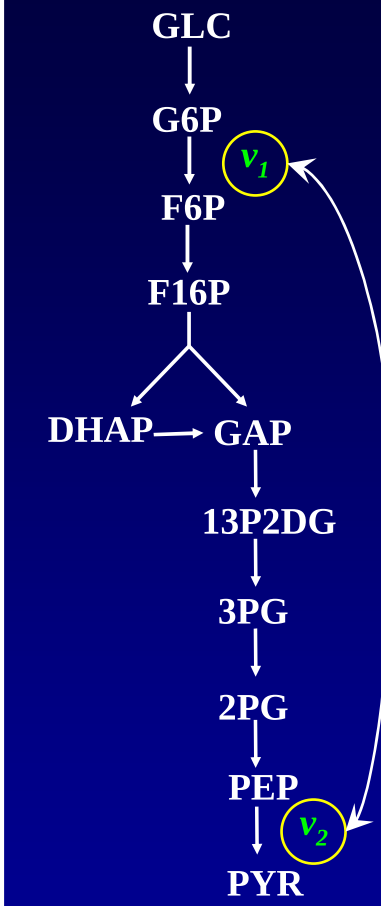
Network size ↑

Environmental constraints ↑

% Blocked reactions ↑



Steady-state Flux Coupling Analysis



Objective:

Identify sets of coupled reactions, equivalent knockouts, and affected reactions in genome-scale stoichiometric models.

Questions:

Coupled Reactions

How are the two fluxes v_1 and v_2 related?

1. Does v_1 imply v_2?
2. Do v_1 and v_2 imply each other?
3. Are v_1 and v_2 completely uncoupled?

Coupled Reaction Sets

Sets of reactions that have to be simultaneously utilized.

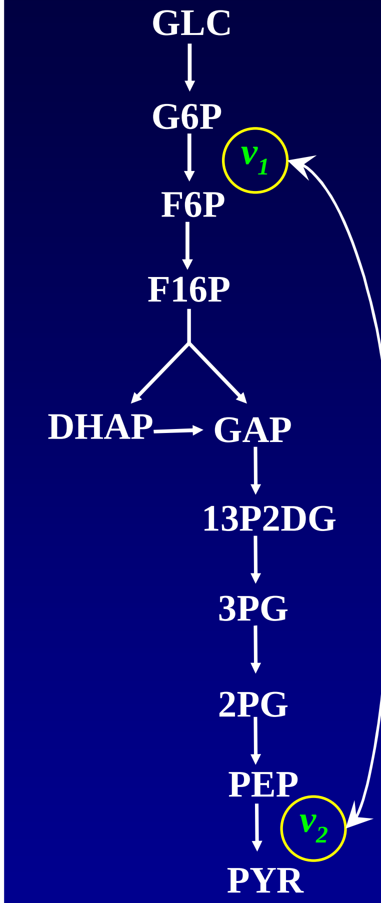
(All reactions coupled reaction set must be “on” or “off”)

Application to genome-scale models:

1.	*Helicobacter pylori*	389 reactions	(Schilling et al., *J Bacteriol*, 2002)
2.	*Escherichia coli*	627 reactions	(Edwards and Palsson, *PNAS*, 2000)
3.	*Saccharomyces cerevisiae*	1173 reactions	(Forster et al., *Genome Res*, 2003)

- Burgard, A.P., Nikolaev, E.V., and C.D. Maranas (2004), “Flux coupling analysis of genome-scale metabolic network reconstructions ,” *Genome Research*, **14**, 301-312.

Steady-state Flux Coupling Analysis



Objective:

Identify sets of coupled reactions, equivalent knockouts, and affected reactions in genome-scale stoichiometric models.

Questions:

Coupled Reactions

How are the two fluxes v_1 and v_2 related?

1. Does v_1 imply v_2?
2. Do v_1 and v_2 imply each other?
3. Are v_1 and v_2 completely uncoupled?

Equivalent Knockouts

Which reactions could alternatively be deleted to force the flux through a particular reaction to zero?

Affected Reactions

Which reaction fluxes will be forced to zero if a particular reaction is removed from the network?

Application to genome-scale models:

1. *Helicobacter pylori*	389 reactions	(Schilling et al., *J Bacteriol*, 2002)
2. *Escherichia coli*	627 reactions	(Edwards and Palsson, *PNAS*, 2000)
3. *Saccharomyces cerevisiae*	1173 reactions	(Forster et al., *Genome Res*, 2003)

- Burgard, A.P., Nikolaev, E.V., and C.D. Maranas (2004), "Flux coupling analysis of genome-scale metabolic network reconstructions ," *Genome Research*, **14**, 301-312.

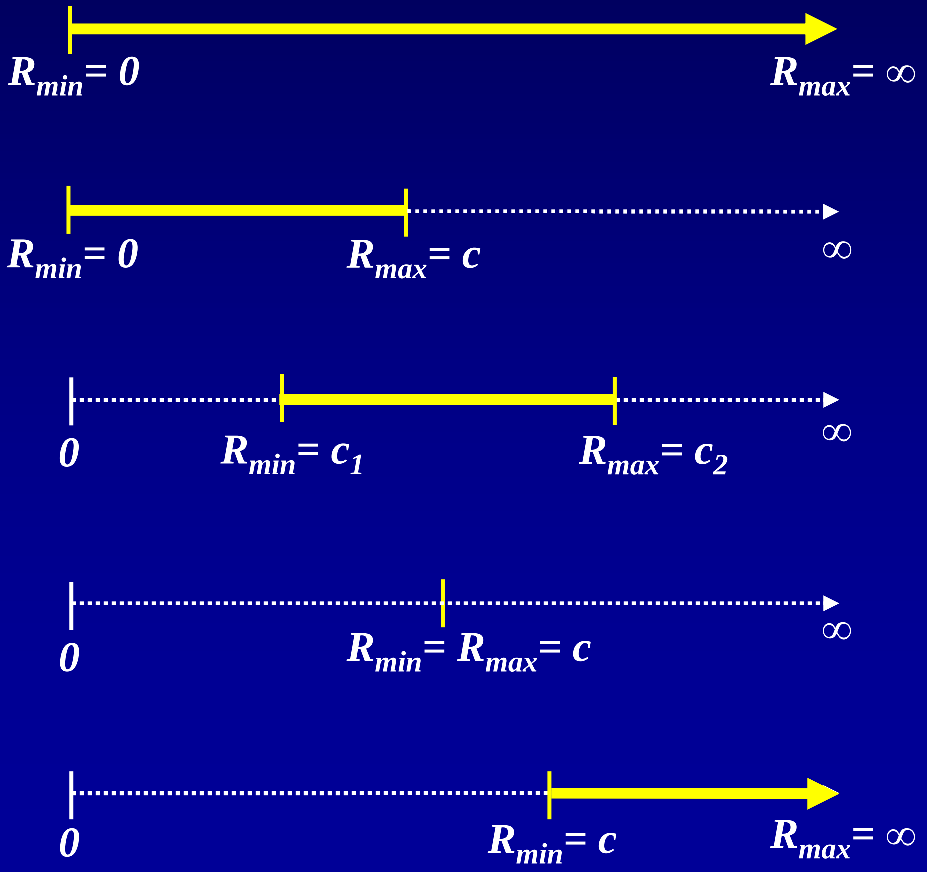
Identifying Coupled Reactions

max $\quad v_1 / v_2 = \boldsymbol{R_{max}}$

min $\quad v_1 / v_2 = \boldsymbol{R_{min}}$

$$\boxed{R_{min} \le \frac{v_1}{v_2} \le R_{max}}$$

Potential Flux Ratio Outcomes:



Uncoupled:

$R_{min} = 0$ $\quad$ $R_{max} = \infty$

Directionally Coupled: $\boxed{v_1 \rightarrow v_2}$

$R_{min} = 0$ $\quad$ $R_{max} = c$ $\quad$ ∞

Partially Coupled: $\boxed{v_1 \leftrightarrow v_2}$

0 $\quad$ $R_{min} = c_1$ $\quad$ $R_{max} = c_2$ $\quad$ ∞

Fully Coupled: $\boxed{v_1 \Leftrightarrow v_2}$

0 $\quad$ $R_{min} = R_{max} = c$ $\quad$ ∞

Directionally Coupled: $\boxed{v_2 \rightarrow v_1}$

0 $\quad$ $R_{min} = c$ $\quad$ $R_{max} = \infty$

Identifying Coupled Reactions

To find the presence and directionality of coupling between two fluxes v_1 and v_2:

Fractional programming formulation:

maximize (or minimize) v_1 / v_2

subject to

$$\sum_{j=1}^{M} S_{ij} v_j = 0$$

$$v_{uptake} \leq v_{uptake\_max}$$

$$v_j \geq 0$$

transformation

Linear formulation: where $\widehat{v}_j = v_j / v_2$

maximize (or minimize) $\widehat{v}_1$

subject to

$$\sum_{j=1}^{M} S_{ij} \widehat{v}_j = 0$$

$$\widehat{v}_{uptake} \leq v_{uptake\_max} \cdot t$$

$$\widehat{v}_j \geq 0$$

$$\widehat{v}_2 = 1$$

$$t \geq 0$$

Flux Coupling Finder (FCF) Procedure

STEP 1 Identify and aggregate all isozymes

STEP 2 Identify and eliminate all blocked reactions

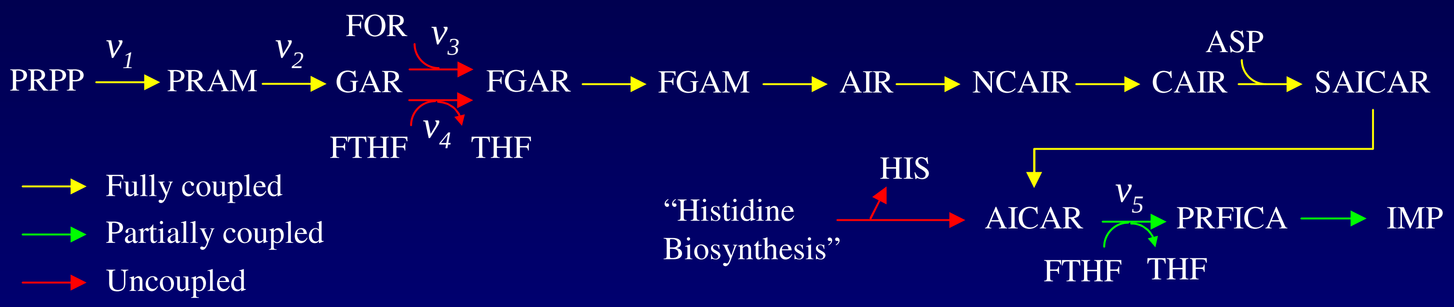
STEP 3 Loop over all reactions *j*

- Loop over all reactions *(j' > j)* and *(j'* not already part of coupled reaction set)
  - Calculate max/min $v_j/v_{j'}$ (Solve 2 LP's)
  - Classify reaction pair *(j, j'*):
    1. Uncoupled
    2. Directionally Coupled $\boxed{v_j \rightarrow v_{j'}}$ or $\boxed{v_{j'} \rightarrow v_j}$
    3. Partially Coupled $\boxed{v_j \leftrightarrow v_{j'}}$
    4. Fully Coupled $\boxed{v_j \Leftrightarrow v_{j'}}$
  - If $v_j \Leftrightarrow v_{j'}$, reactions belong to same enzyme subset.
  - If $v_j \leftrightarrow v_{j'}$, reactions belong to same coupled reaction set.

Property: $\boxed{\text{If } v_1 \Leftrightarrow v_2 \text{ and } v_2 \Leftrightarrow v_3 \text{ then } v_1 \Leftrightarrow v_3}$

Example of Coupled Reaction Set in E. coli

Nucleotide Biosynthesis:



PRPP $\xrightarrow{v_1}$ PRAM $\xrightarrow{v_2}$ GAR

FOR v_3

FGAR

FTHF v_4 THF

FGAR → FGAM → AIR → NCAIR → CAIR → SAICAR

ASP

→ Fully coupled

→ Partially coupled

→ Uncoupled

HIS

"Histidine Biosynthesis" → AICAR

AICAR $\xrightarrow{v_5}$ PRFICA → IMP

FTHF THF

- ❑ Fully coupled fluxes $\boxed{v_j \Leftrightarrow v_{j'}}$ (*e.g.*, $v_1/v_2 = 1$)

- ❑ Partially coupled fluxes $\boxed{v_j \leftrightarrow v_{j'}}$ (*e.g.*, $1 \le v_1/v_5 \le 1.2$)

- ❑ Fluxes v_3 and v_4 uncouple each other

Reaction Coupling Statistics

19 sets of 2 reactions

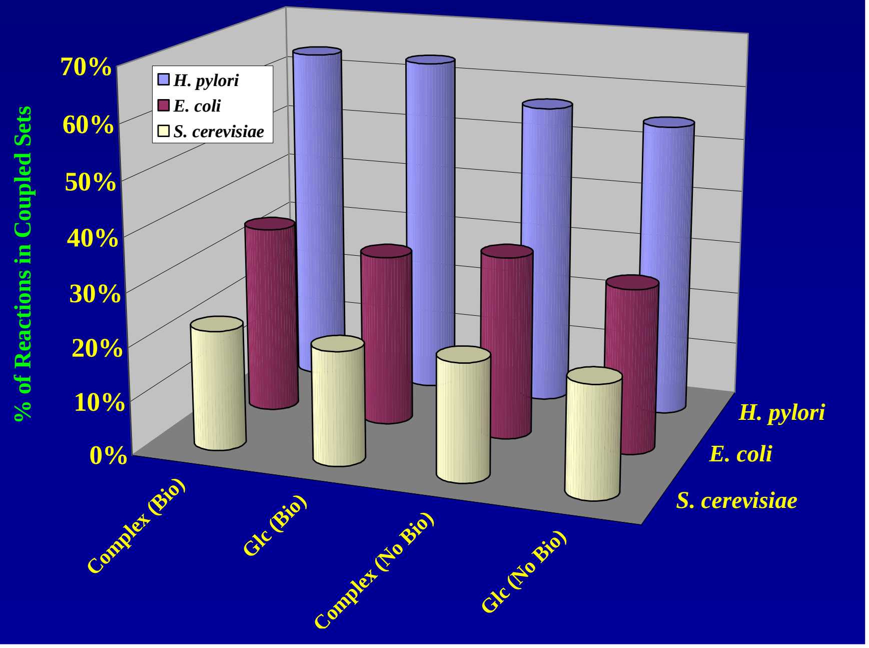
Coupled to biomass formation

	H. pylori				***E. coli***				***S. cerevisiae***			
	biomass reaction		**no biomass reaction**		**biomass reaction**		**no biomass reaction**		**biomass reaction**		**no biomass reaction**	
	complex media	**glucose minimal**	**complex media**	**glucose minimal**	**complex media**	**glucose minimal**	**complex media**	**glucose minimal**	**complex media**	**glucose minimal**	**complex media**	**glucose minimal**
	19 (2)	15 (2)	25 (2)	23 (2)	45 (2)	33 (2)	48 (2)	37 (2)	51 (2)	45 (2)	51 (2)	46 (2)
	8 (3)	7 (3)	10 (3)	9 (3)	9 (3)	6 (3)	9 (3)	9 (3)	13 (3)	11 (3)	14 (3)	14 (3)
	2 (4)	1 (5)	3 (4)	2 (4)	4 (4)	1 (4)	4 (4)	3 (4)	6 (4)	5 (4)	7 (4)	5 (4)
	1 (6)	1 (7)	1 (5)	2 (5)	3 (5)	1 (5)	5 (5)	5 (5)	2 (5)	3 (5)	2 (5)	3 (5)
	2 (7)	1 (10)	2 (6)	3 (6)	2 (6)	3 (7)	2 (6)	2 (6)	2 (6)	2 (6)	2 (6)	2 (6)
	1 (10)	**1 (174)**	3 (7)	2 (7)	1 (7)	1 (10)	2 (7)	2 (7)	1 (7)	1 (7)	1 (7)	1 (7)
	1 (148)		1 (8)	1 (8)	1 (8)	**1 (112)**	1 (8)	1 (8)	2 (8)	2 (8)	2 (8)	2 (8)
			1 (9)	1 (9)	2 (9)		3 (9)	3 (9)	1 (9)	1 (9)	1 (9)	1 (9)
			4 (10)	4 (10)	**1 (66)**		1 (10)	1 (10)	1 (12)	1 (12)	1 (12)	1 (12)
			1 (13)	1 (13)			1 (17)	1 (17)	**1 (30)**	**1 (34)**	1 (17)	1 (17)
			1 (20)	1 (20)								
total reactions in subsets:	248	247	220	213	259	236	252	226	261	248	255	242
total subsets:	34	26	52	49	68	46	76	64	80	72	82	76

- ❑ Network size ↑ % Coupled reactions ↓

- ❑ Conditions only slightly affect coupling.

- ❑ Presence of biomass reaction → Large coupled reaction sets



Enzyme Subset Identification in H. pylori

(Schilling et al., *J Bacteriol*, 2002)

Amino acid metabolism

- DHS1, AROB, AROQ, AROE, AROK, AROA, AROC
- TRPD, TRPC1, TRPC2, TRPAB
- TYRA1, TYRA2, ASPB2
- METL2, THRB, THRC
- DAPA, DAPB, DAPD, DAPC, DAPE, DAPF
- ADCSASE_r, METH
- SERA, SERC, SERB
- SPEA, SPEB
- SPED, SPEE, MTHAKN, MTHRKN, MTHIPIS, NE1PH, NE3UNK, TNSUNK
- CYSDN, CYSC, CYSH, CYSU, CYSE, CYSK

Central metabolism

- FBP, FBA_r
- GAP, PGK
- PGM, ENO
- PGL, EDD, EDA
- GLTA, ACNB, ICD
- SCOT, ATOB

Lipid and Cell Envelope

- ACCABCD, FABD
- FABH1, FABF
- C12OSN, DGKA, LPXA, ENVA, LPXD, USHA12, LPXB, LPXK, KDTAI, KDOLIPH, ASPISO, KDSA, KDOPH, KDSB, PAPHTSE, GMHA, LPSSYN
- PGSA2, PGPP
- GLMS, GLMM, GLMU
- MURZ, MURB, MURC, MURD, MURE, MURF, GLR, DDLA, MRAY, MURG

Nucleotide Metabolism

- PYRA, PYRB, PYRC, PYRD
- PYRE, PYRF
- PURF, PURD, PURL, PURM, PURK, PURE, PURC, PURB1, PURH1, PURH2
- PURA, PURB2, GUAB, GUAA
- NDK4, KRDAB4
- NDK6, NRDAB1
- NDK7, NRDAB3
- NDK8, PRM1
- DEOD2, DEOD8_r

Transport

- ADHE2, ETHTP_r
- PTA, ACKA
- GALU, ALGC1
- GLCTP, GLK1
- PROTPI, NATP_r
- LACTP, DLD
- BCRBTP_r, ICFA
- GLCD, GLLDHR, KATA

Vitamin and cofactor

- FOLE, DNTPH, DHPPH, FOLB, FOLK, PABB, PABC, FOLP, FOLC
- FOLD1, FOLD2
- GLTX, HEMA, HEML, HEMB, HEMC, HEMD, HEME, HEMF, HEMG, HEMH
- RIBA, RIBD1, RIBD2, PMDPHT, RIBB, RIBE, RIBC, RIBF1, FIBF2
- PANB, ILVC3, PAND, PANC, COAA, PCLIG, PCDCL, PATRAN, DPHCOAK
- IPPPISO, ISPA1, ISPA2
- NADB, NADA, NADC, NADD, NADE
- MENF, MEND1, MEND2, MENC, MENE, MENB, MENA

Enzyme Subset Identification in H. pylori

(Schilling et al., *J Bacteriol*, 2002)

Amino acid metabolism

- DHS1, AROB, AROQ, AROE, AROK, AROA, AROC
- TRPD, TRPC1, TRPC2, TRPAB
- TYRA1, TYRA2, ASPB2
- METL2, THRB, THRC
- DAPA, DAPB, DAPD, DAPC, DAPE, DAPF
- ADCSASE_r, METH, MENG
- SERA, SERC, SERB
- SPEA, SPEB
- SPED, SPEE, MTHAKN, MTHRKN, MTHIPIS, NE1PH, NE3UNK, TNSUNK
- CYSDN, CYSC, CYSH, CYSU, CYSE, CYSK, SLFTP

Central metabolism

- FBP, FBA_r
- GAP, PGK
- PGM, ENO
- PGL, EDD, EDA
- GLTA, ACNB, ICD
- SCOT, ATOB, ACCTP

Lipid and Cell Envelope

- ACCABCD, FABD
- FABH1, FABF
- C12OSN, DGKA, LPXA, ENVA, LPXD, USHA12, LPXB, LPXK, KDTAI, KDOLIPH, ASPISO, KDSA, KDOPH, KDSB, PAPHTSE, GMHA, LPSSYN
- PGSA2, PGPP
- GLMS, GLMM, GLMU
- MURZ, MURB, MURC, MURD, MURE, MURF, GLR, DDLA, MRAY, MURG

Nucleotide Metabolism

- PYRA, PYRB, PYRC, PYRD
- PYRE, PYRF
- PURF, PURD, PURL, PURM, PURK, PURE, PURC, PURB1, PURH1, PURH2
- PURA, PURB2, GUAB, GUAA
- NDK4, KRDAB4
- NDK6, NRDAB1
- NDK7, NRDAB3
- NDK8, PRM1
- DEOD2, DEOD8_r, NUPCTPS

Transport

- ADHE2, ETHTP_r
- PTA, ACKA
- GALU, ALGC1
- GLCTP, GLK1
- PROTPI, NATP_r
- LACTP, DLD
- BCRBTP_r, ICFA
- GLCD, GLLDHR, KATA

Missed Subsets

- OOR, FRDO
- POR, FLDO
- TDK1, NUPCTP4
- DEODG, GSNTP

Vitamin and cofactor

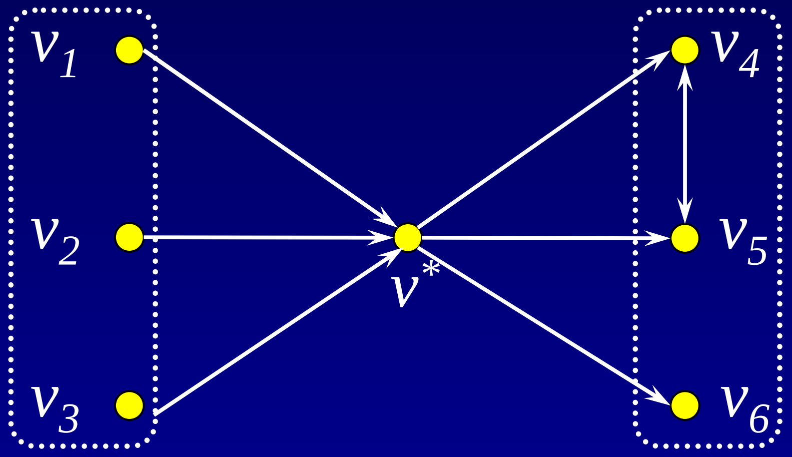
- FOLE, DNTPH, DHPPH, FOLB, FOLK, PABB, PABC, FOLP, FOLC, ACEB
- FOLD1, FOLD2
- GLTX, HEMA, HEML, HEMB, HEMC, HEMD, HEME, HEMF, HEMG, HEMH
- RIBA, RIBD1, RIBD2, PMDPHT, RIBB, RIBE, RIBC, RIBF1, FIBF2
- PANB, ILVC3, PAND, PANC, COAA, PCLIG, PCDCL, PATRAN, DPHCOAK, ILVE2
- IPPPISO, ISPA1, ISPA2
- NADB, NADA, NADC, NADD, NADE
- MENF, MEND1, MEND2, MENC, MENE, MENB, MENA

Breaking network into subsystems does miss various couplings.

Directional Coupling

Equivalent Knockouts vs. Affected Reactions

Reactions "affected" by *v**   "Equivalent KO's" for *v**

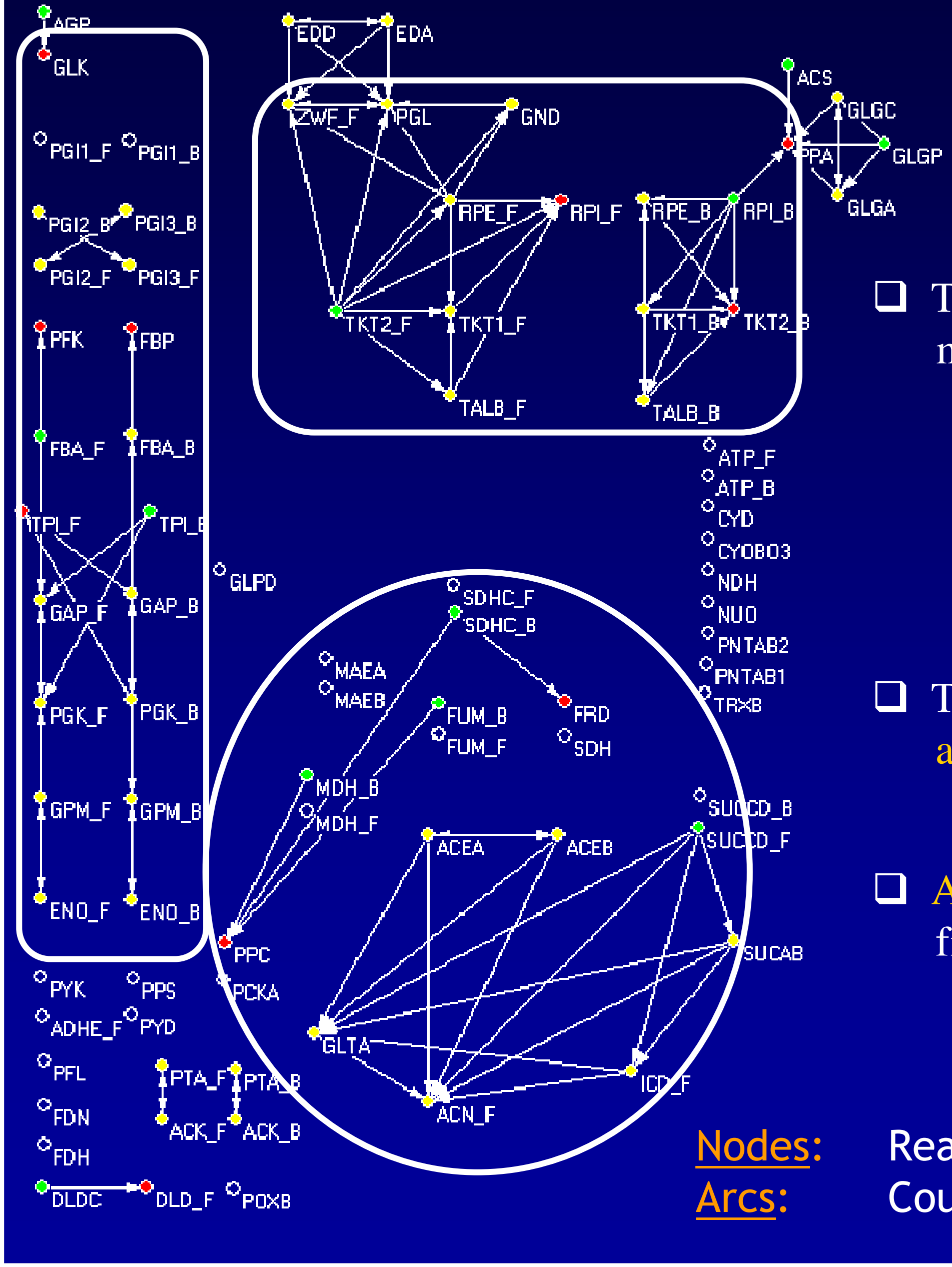


Corresponding flux ratio outcomes,

$$v_{1,2,\text{ or }3} \leq c \cdot v^*$$

$$v^* \leq c \cdot v_{4,5,\text{ or }6}$$

E. coli Central Metabolic Network Coupling



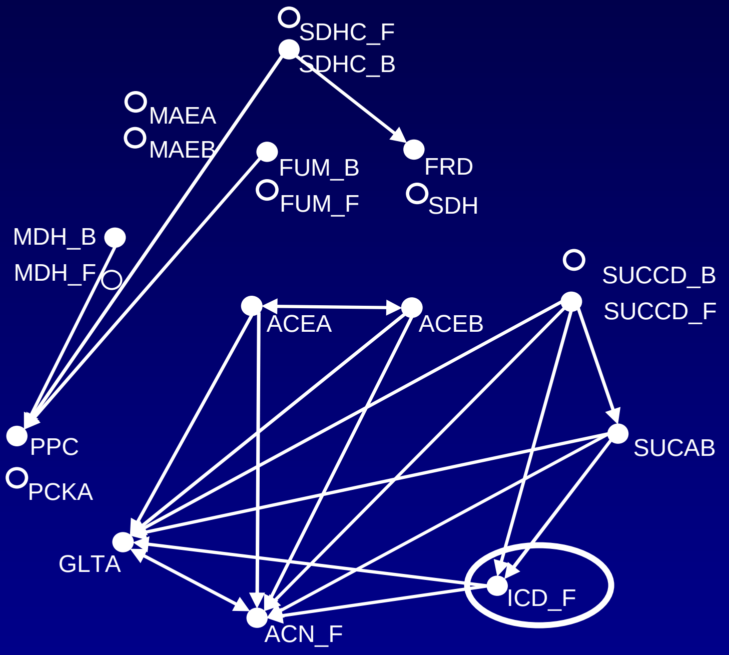
Growth on glucose minimal media
Steady-state conditions

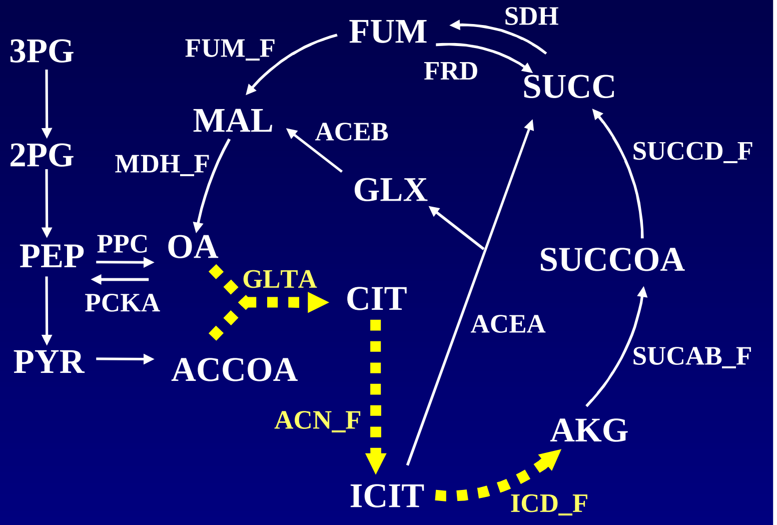
- ❑ The forward and backward directions of the major metabolic pathways show significant internal coupling.
  - (1) glycolysis
  - (2) pentose phosphate pathway
  - (3) TCA cycle
- ❑ The major pathways above are uncoupled from one another.
- ❑ Anaplerotic and respiration reactions are uncoupled from the rest of central metabolism.

Nodes: Reactions
Arcs: Couplings

E. coli Central Metabolic Network Coupling

Growth on glucose minimal media
Steady-state conditions





Removing either GLTA or ACN forces the flux thru ICD to zero.

Equivalent knockouts

ICD_F: Isocitrate dehyrogenase

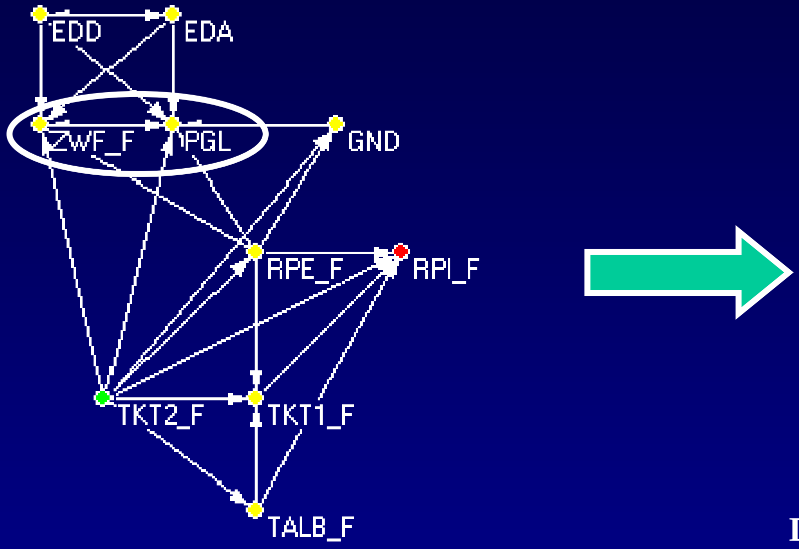
1) **GLTA:** Citrate synthase
2) **ACN_F:** Aconitase

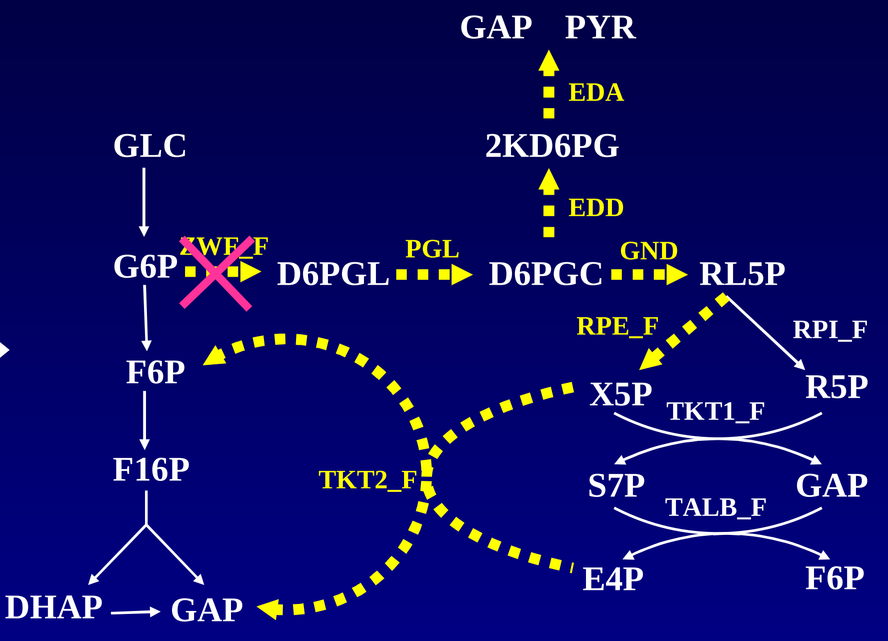
Lethal Mutations for E. coli Growth on Glucose Minimal Media

icdA	(Helling and Kukora, *J. Bacteriol.*, 1971)
gltA	(Lakshmi and Helling, *J. Bacteriol.*, 1976)
acnAB	(Gruer et al., *Microbiology*, 1997)

E. coli Central Metabolic Network Coupling

Growth on glucose minimal media
Steady-state conditions





Sets of affected reactions

ZWF:	Glucose 6-phosphate-1-dehydrogenase
PGL:	6-Phosphogluconolactonase

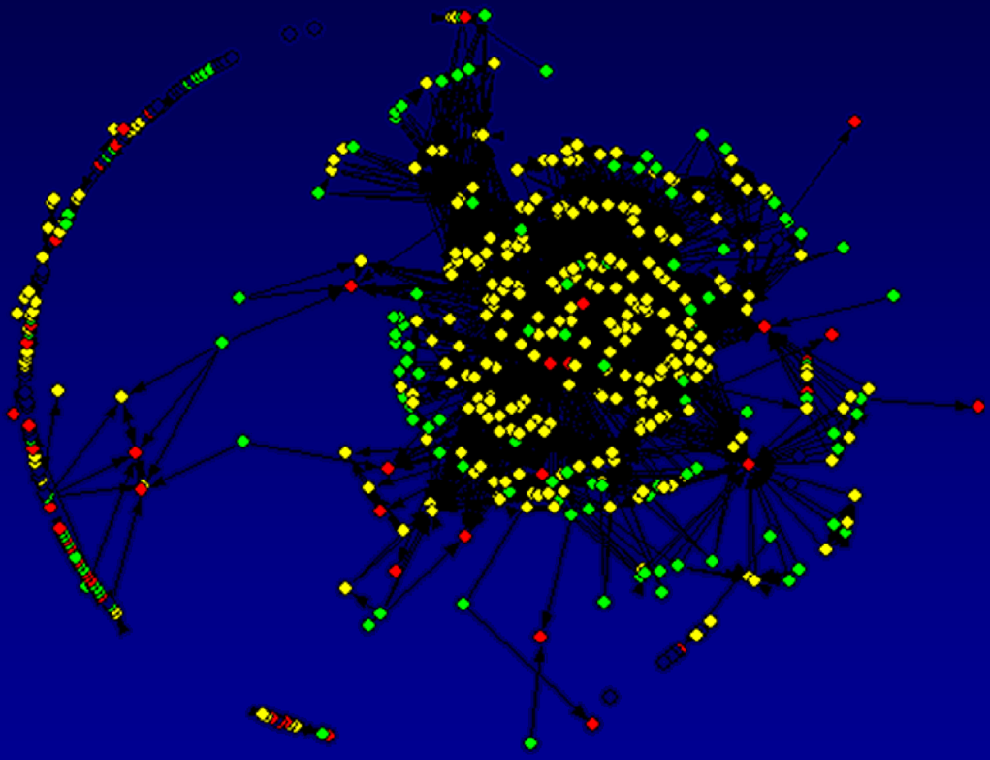
1) **EDD:**	Phosphogluconate dehydratase
2) **EDA:**	2-Keto-3-deoxy-6-phosphogluconate aldolase
3) **GND:**	6-Phosphogluconate dehydrogenase
4) **RPE_F:**	Ribulose phosphate 3-epimerase
5) **TKT2_F:**	Transketolase

Removing ZWF or PGL forces flux thru the five affected reactions to zero.

Scale-free Nature of Directional Coupling

Genome-wide metabolic coupling

E. coli **growth on a glucose minimal medium**



Nodes: Reactions

Arcs: Directional coupling

Scale Free Architectures

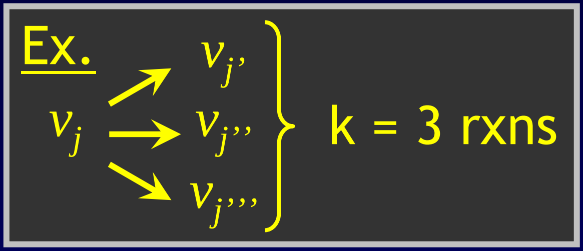
Metabolite Centered Graphs (Jeong et al., *Nature*, 2000)
(Wagner and Fell, *Proc. R. Soc. Lond. B. Biol. Sci.*, 2001)

Reaction Flux Centered Graphs (Burgard, et. al., *Genome Res.*, 2004)

Scale-free Nature of Directional Coupling

k = Number of reactions

N(k) = Number of reactions implying k reactions



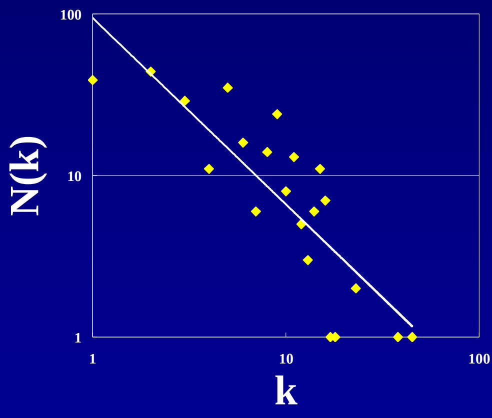
Ex.

$v_j \rightarrow \left. \begin{matrix} v_{j'} \\ v_{j''} \\ v_{j'''} \end{matrix} \right\}$ k = 3 rxns

Helicobacter pylori

N(k) = 94.6 $k^{-1.154}$

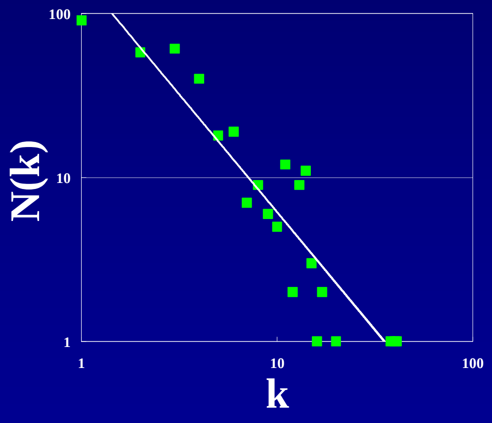
R^2 = 0.707



Escherichia coli

N(k) = 167.7 $k^{-1.436}$

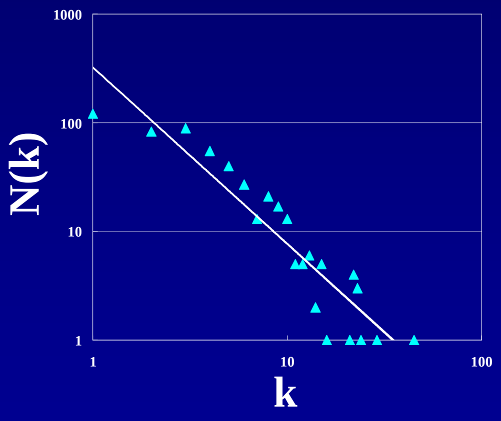
R^2 = 0.848



Saccharomyces cerevisiae

N(k) = 323.7 $k^{-1.624}$

R^2 = 0.864



Scale Free Architectures

Metabolite Centered Graphs

(Jeong et al., *Nature*, 2000)
(Wagner and Fell, *Proc. R. Soc. Lond. B. Biol. Sci.*, 2001)

Reaction Flux Centered Graphs

(Burgard, et. al., *Genome Res.*, 2004)

Summary

Flux Coupling Finder (FCF) Procedure:

❑ Identification of reaction coupling in genome-scale models

❑ Locate equivalent knockouts and sets of affected reactions

❑ Aid in model consistency testing and guiding/interpreting genetic manipulations

❑ Implementation in C++ using LINDO

CPU times ~ minutes (Intel Pentium IV, 2.4 GHz, 512 MB RAM PC)

Publications (fenske.che.psu.edu/faculty/cmaranas)

Strain design

- Pharkya, P. and C.D. Maranas (2005), "An Optimization Framework for Identifying Reaction Activation/Inhibition or Elimination Candidates for Overproduction in Microbial Systems", submitted.
- Pharkya, P., A.P. Burgard and C.D. Maranas (2004), "OptStrain: A Computational Framework for Redesign of Microbial Production Systems", *Genome Research*, 14, 2367-2376.
- Pharkya, P., A.P. Burgard and C.D. Maranas (2003), "Exploring the Overproduction of Amino Acid Using the Bilevel Optimization Framework OptKnock," *Biotechnology and Bioengineering*, 84, 887-899.
- Burgard, A.P., P. Pharkya and C.D. Maranas (2003), "OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization," *Biotechnology and Bioengineering*, 84, 647-657

Metabolite flux and concentration coupling analysis

- Nikolaev, E.V., A.P. Burgard, and C.D. Maranas (2005), "Elucidation and Structural Analysis of Conserved Pools for Genome-Scale Metabolic Reconstructions," *Biophysical Journal*, 88, 37-49.
- Burgard, A.P.†, E.V. Nikolaev†, C.H. Schilling and C.D. Maranas (2004), "Flux Coupling Analysis of Genome-scale Metabolic Network Reconstructions," *Genome Research*, 14, 301-312

Inferring and Testing Metabolic Objective Functions

- Burgard, A.P. and C.D. Maranas (2002), "An Optimization-based Framework for Inferring and Testing Hypothesized Metabolic Objective Functions," *Biotechnology and Bioengineering*, 82, 670-677.

Minimal Reaction Set Identification

- Burgard, A.P., S. Vaidyaraman and C.D. Maranas (2001), "Minimal Reaction Sets for *Escherichia coli* Metabolism under Different Growth Requirements and Uptake Environments," *Biotech. Progress*,17, 791-797.

Pathway discovery and optimization

- Burgard, A.P. and C.D. Maranas (2001), "Probing the Performance Limits of the *Escherichia coli* Metabolic Network Subject to Gene Additions or Deletions," *Biotechnology and Bioengineering*, 74, 364-375.

Signaling networks

- Dasika, M., Burgard, A.P. and C.D. Maranas (2005), "A computational framework for topological analysis and targeted disruption of signal transduction networks" *submitted*.

Acknowledgements

Graduate Researchers:

Tony Burgard
Priti Pharkya
Jon Chin

Post-Doctoral Researcher:

Evgeni Nikolaev

Collaborators:

Fred Blattner	***University of Wisconsin***
Jay Keasling	***University of California, Berkeley***
Bernhard Palsson	***University of California, San Diego***
Christophe Schilling	***Genomatica, Inc.***
Patrick Cirino	***Penn State***

Funding Sources:

National Science Foundation and Department of Energy