



# Transcriptional Regulatory Networks



# LEARNING OBJECTIVES

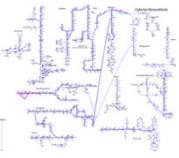
Each student should be able to:

- Explain the purpose of a transcriptional regulatory network.
- Explain the role of a transcription factor.
- Explain dynamic regulatory flux balance analysis.
- Explain the difference between dFBA, dRFBA and PROM regulatory approaches.
- Explain Boolean transcriptional regulation.
- Explain probabilistic transcriptional regulation.
- Describe the difference between the regular constraint-based FBA models and the regulatory FBA model.
- Describe the strengths and limitations of dRFBA analysis.

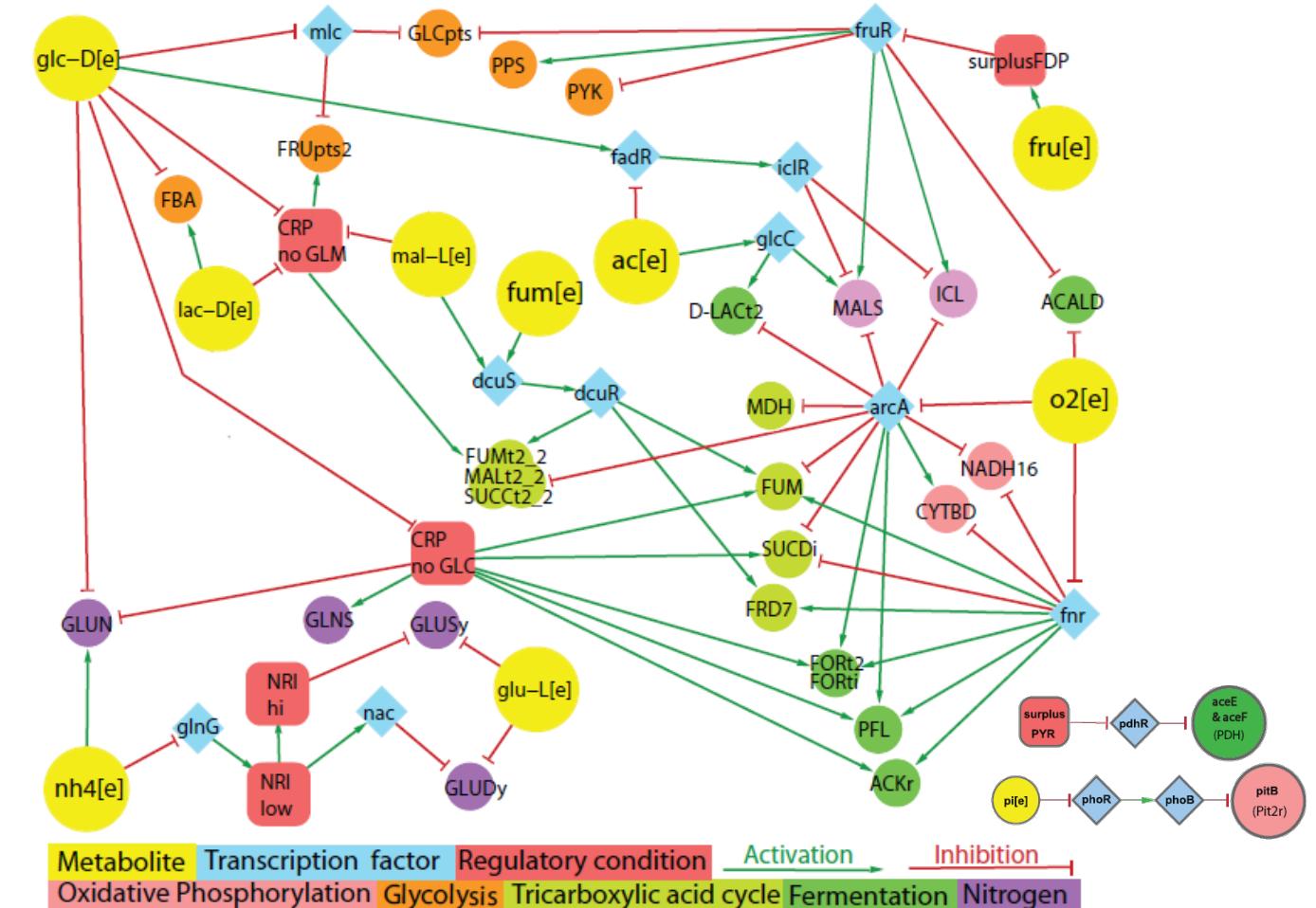
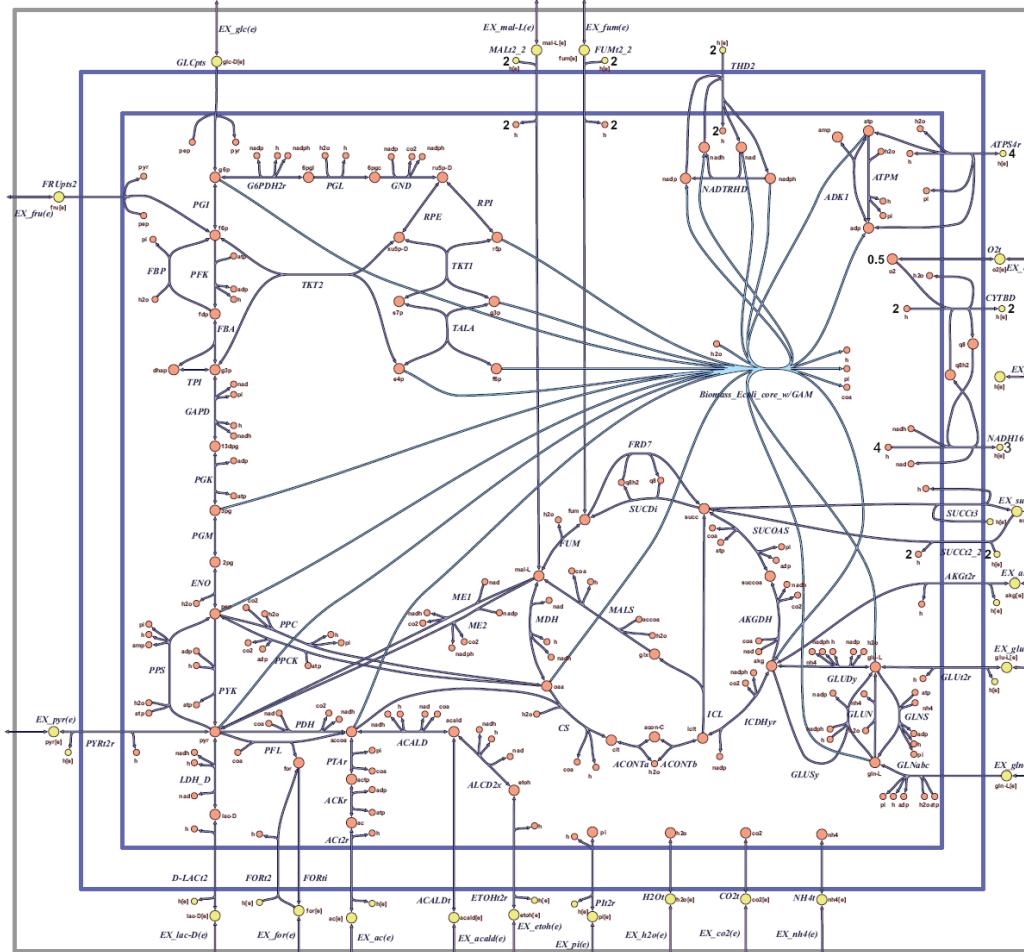


# Lesson Outline

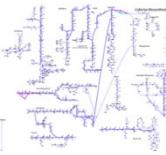
- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



# Metabolic and Transcriptional Regulatory Networks

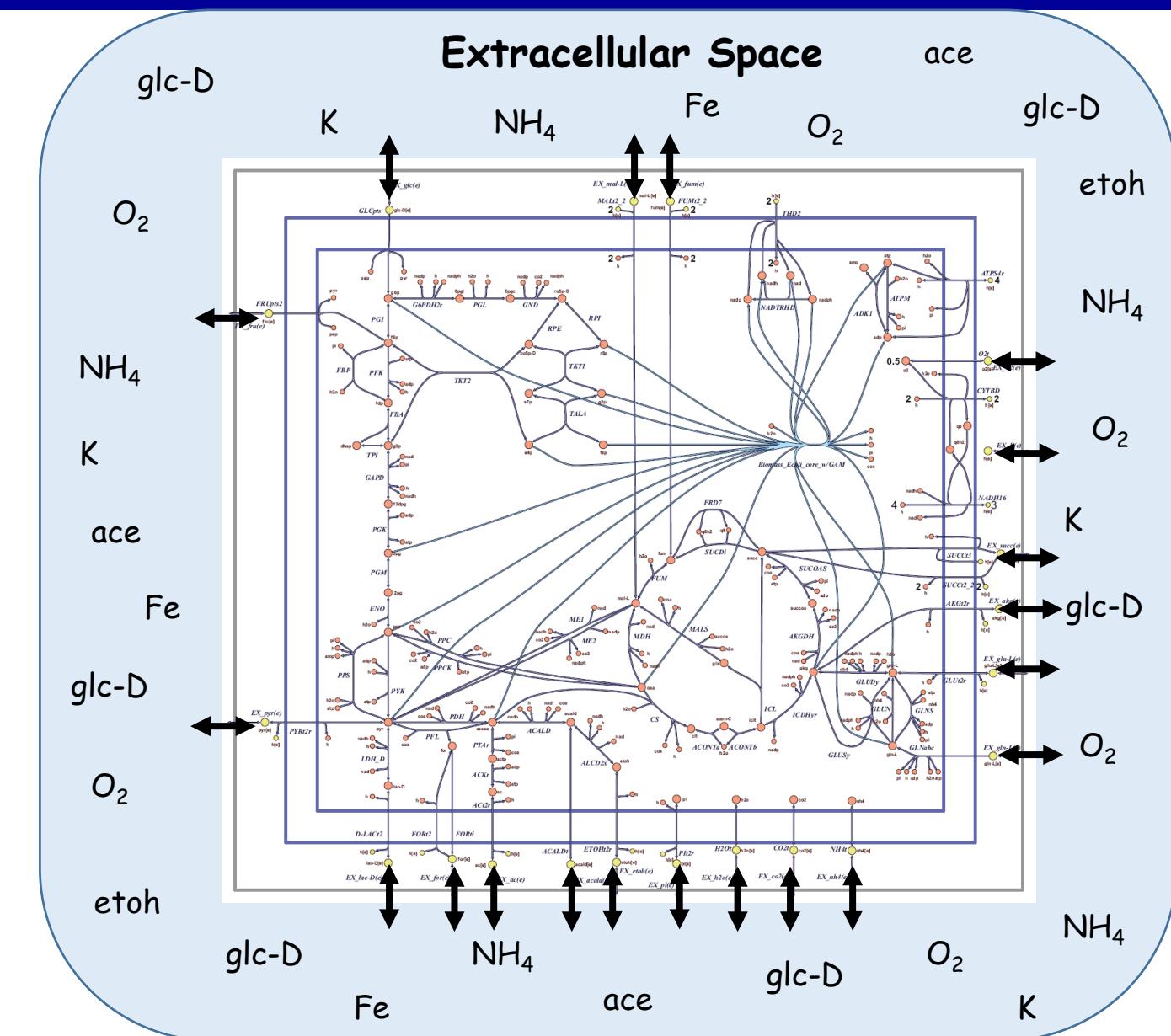


Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Regulatory Control

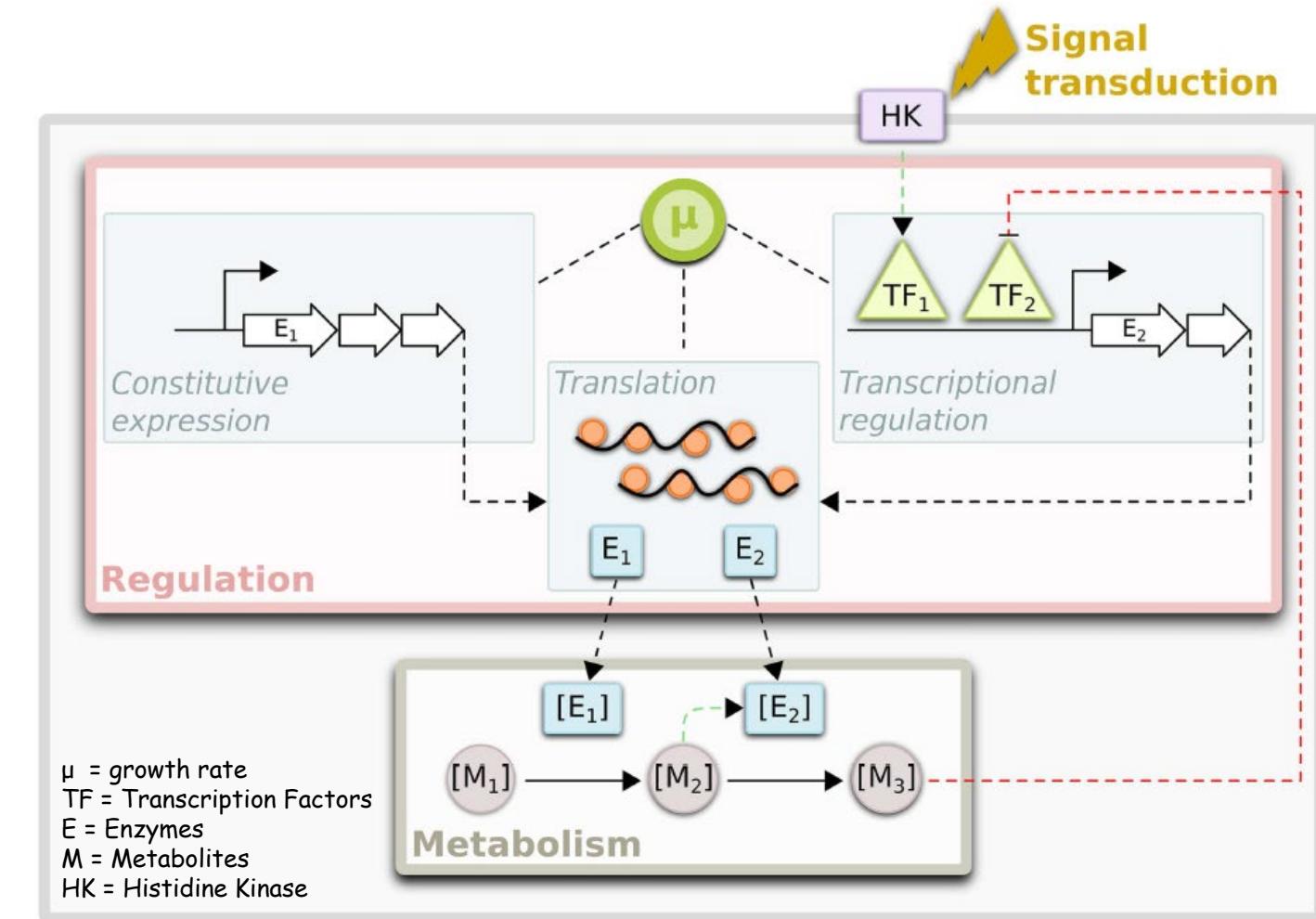
- Manage the uptake of metabolites into the cell.
  - Manage the secretion of metabolites into the extracellular space.
  - Adapt cell metabolism to external environment.
  - Optimize growth-rate
    - ✓ Catabolite repression
  - Protect the cell from environmental stresses
    - ✓ pH, osmolarity, high and low temperatures, starvation, oxidative stress.
  - Internal cell management and repair
    - ✓ DNA repair



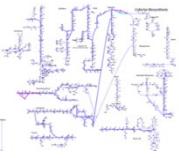


# Interconnections Between Regulation And Metabolism

Regulation of flux through metabolic networks is achieved by the control of enzyme levels ( $[E]$ ) and/or activities. Enzyme levels can be controlled transcriptionally via specific regulation of transcription factors (TFs) or via global mechanisms, which depend on factors such as growth rate ( $\mu$ ). The expression levels of constitutively expressed genes may be solely under control of these global mechanisms. In addition, growth rate also has a significant impact on translation rates. The activities of TFs can be modulated by specific metabolites ( $[M]$ ) or via post-translational modifications by histidine kinases (HK) that sense environmental cues, among other mechanisms. Enzyme activities can also be modulated via post-translational (allosteric) interactions with metabolites. All these networks are dynamic and in constant communication with one another to determine metabolic state of a cell.

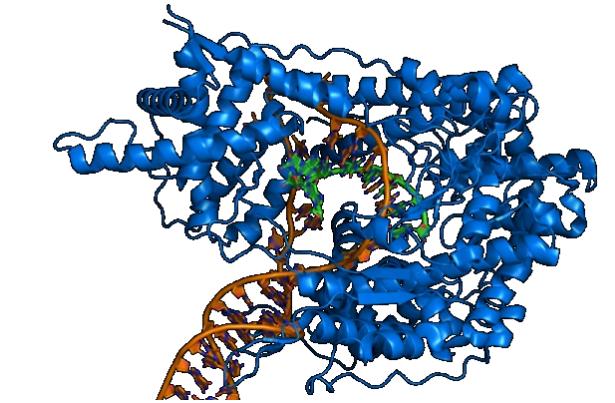
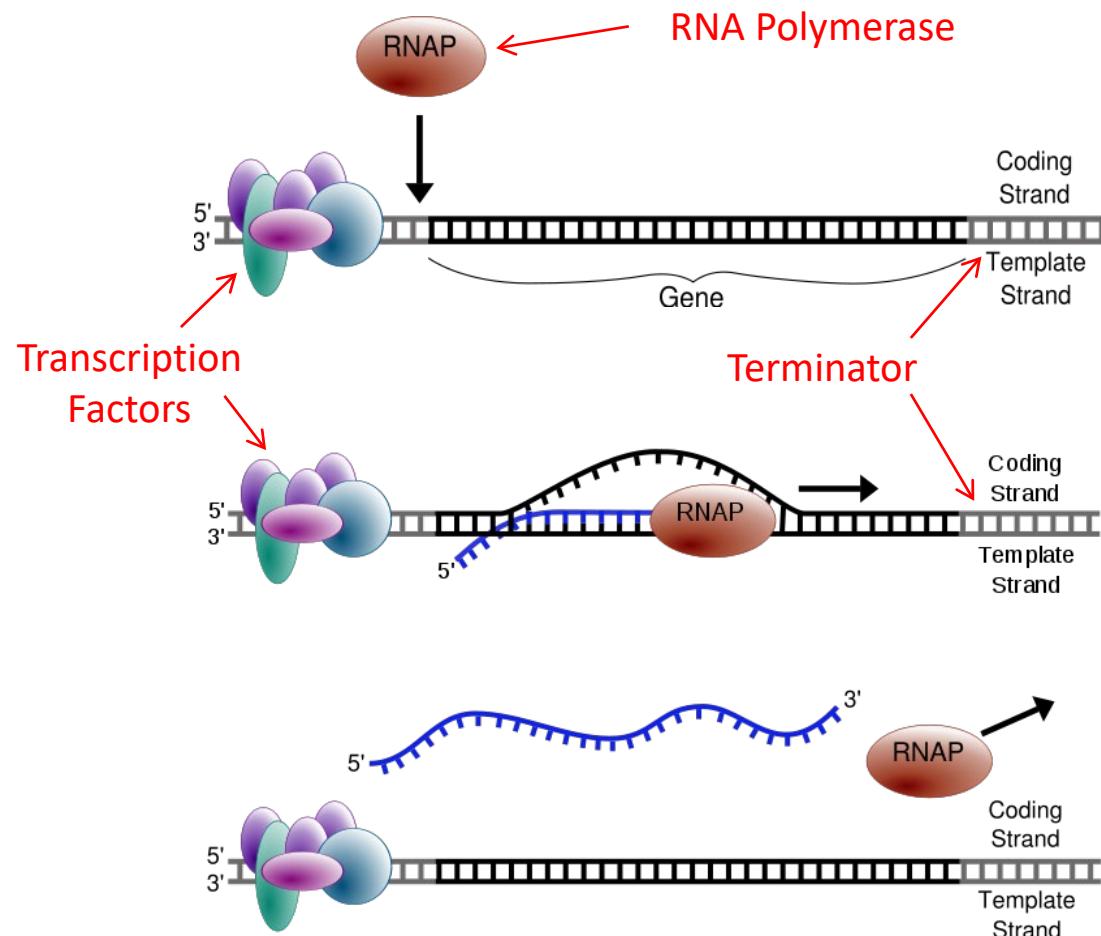


Imam, S., S. Schauble, et al. (2015). "Data-driven integration of genome-scale regulatory and metabolic network models." *Frontiers in microbiology* 6: 409.

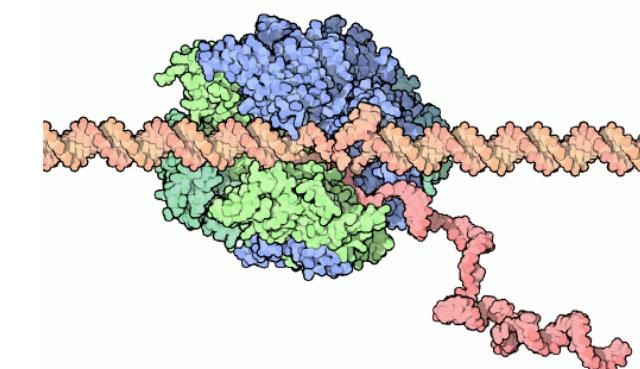


# TRANSCRIPTION

[http://en.wikipedia.org/wiki/Transcription\\_\(genetics\)](http://en.wikipedia.org/wiki/Transcription_(genetics))



T7 RNA Polymerase

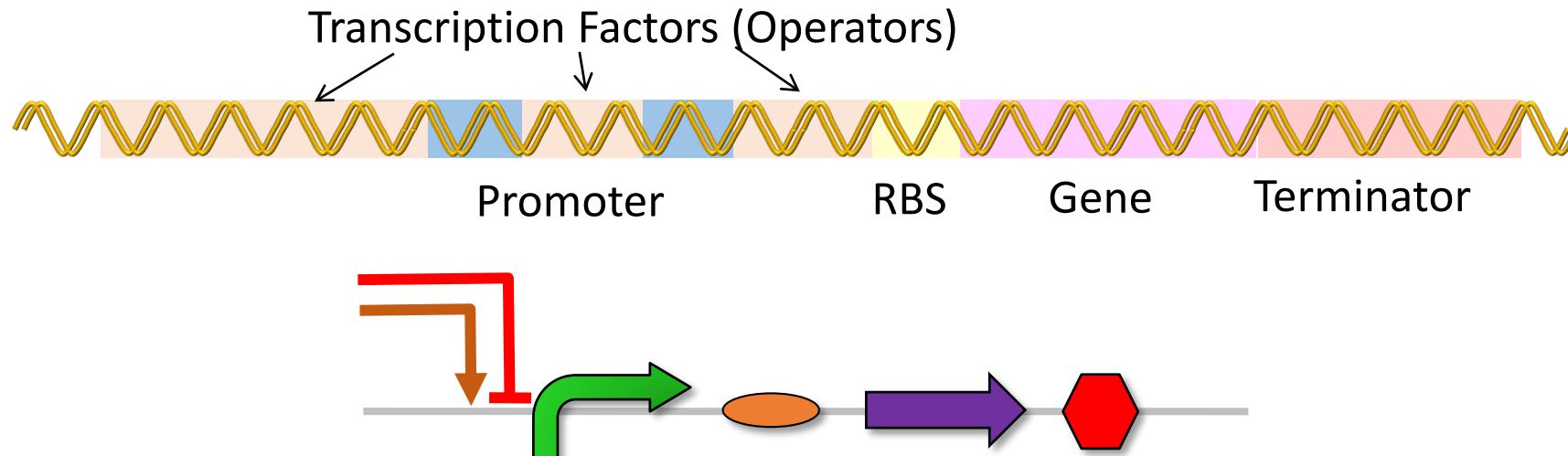


[http://creationwiki.org/pool/images/d/d1/RNA\\_polymerase.png](http://creationwiki.org/pool/images/d/d1/RNA_polymerase.png)



# OPERON

An **operon** contains one or more genes which can be transcribed into mRNA. Upstream of the genes lies a promoter sequence which provides a site for RNA polymerase to bind and initiate transcription. Surrounding the RNA Polymerase binding sites of the promoter lies sections of DNA called transcription factor binding sites which can be used to regulate the gene transcription. Finally, the terminator ends the transcription process by releasing the RNA Polymerase from the DNA.

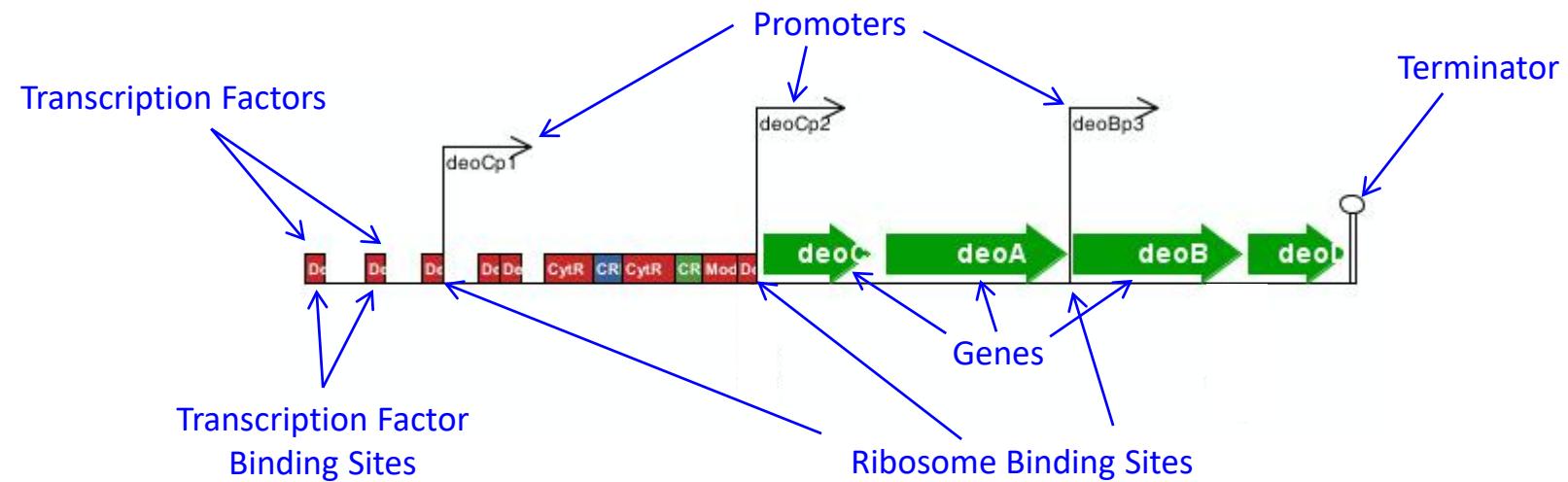




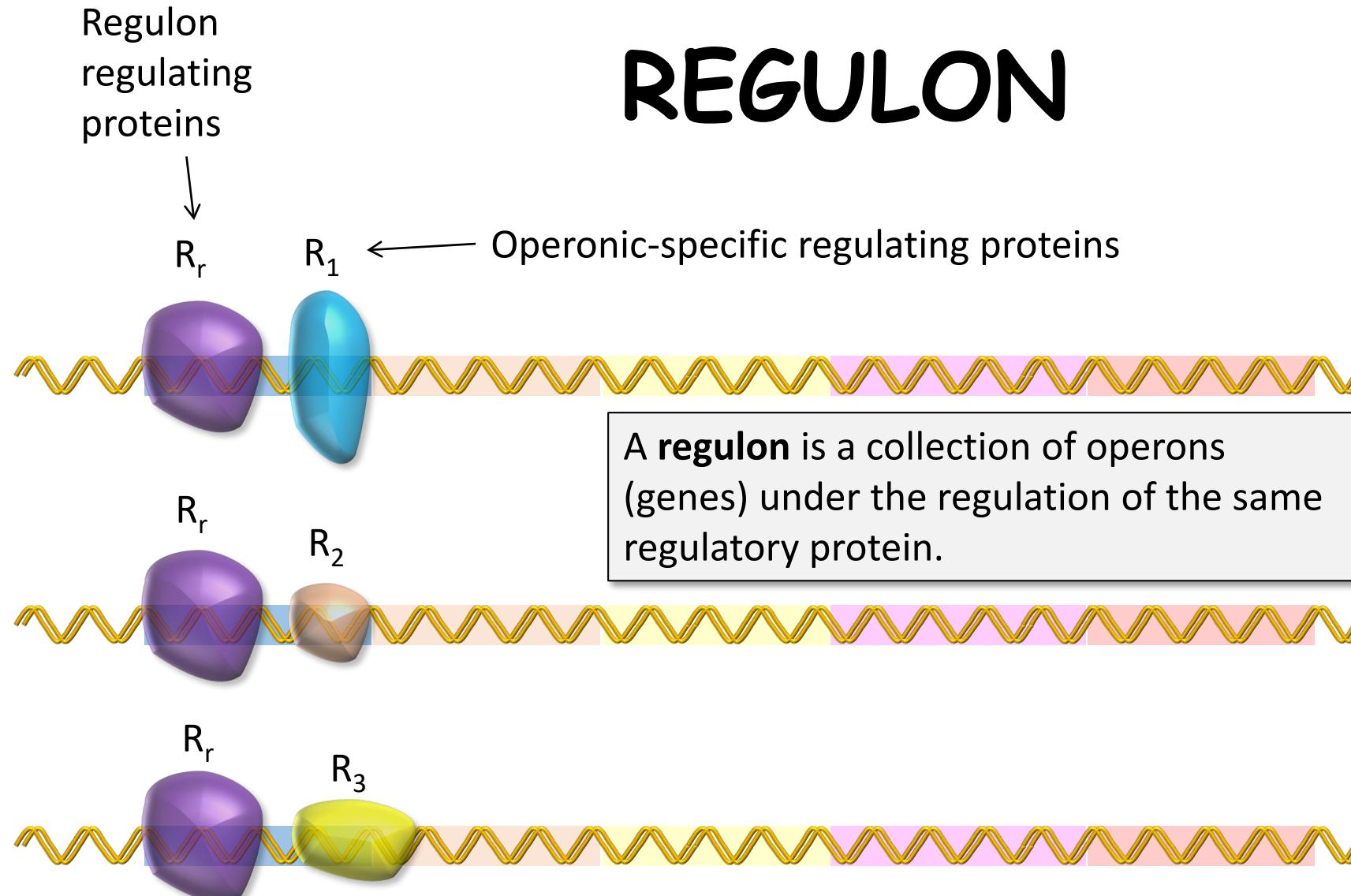
# OPERON

[http://regulondb.ccg.unam.mx/html/Project\\_glossary.jsp](http://regulondb.ccg.unam.mx/html/Project_glossary.jsp)

An **operon** is a set of one or several genes and their associated regulatory elements, which are transcribed as a single unit. An operon is a group of one or more genes transcribed in the same direction. It must contain a promoter upstream of all genes and a terminator downstream. It is also relatively common to find operons with several promoters, some of them internally located, thus, transcribing a partial group of genes.

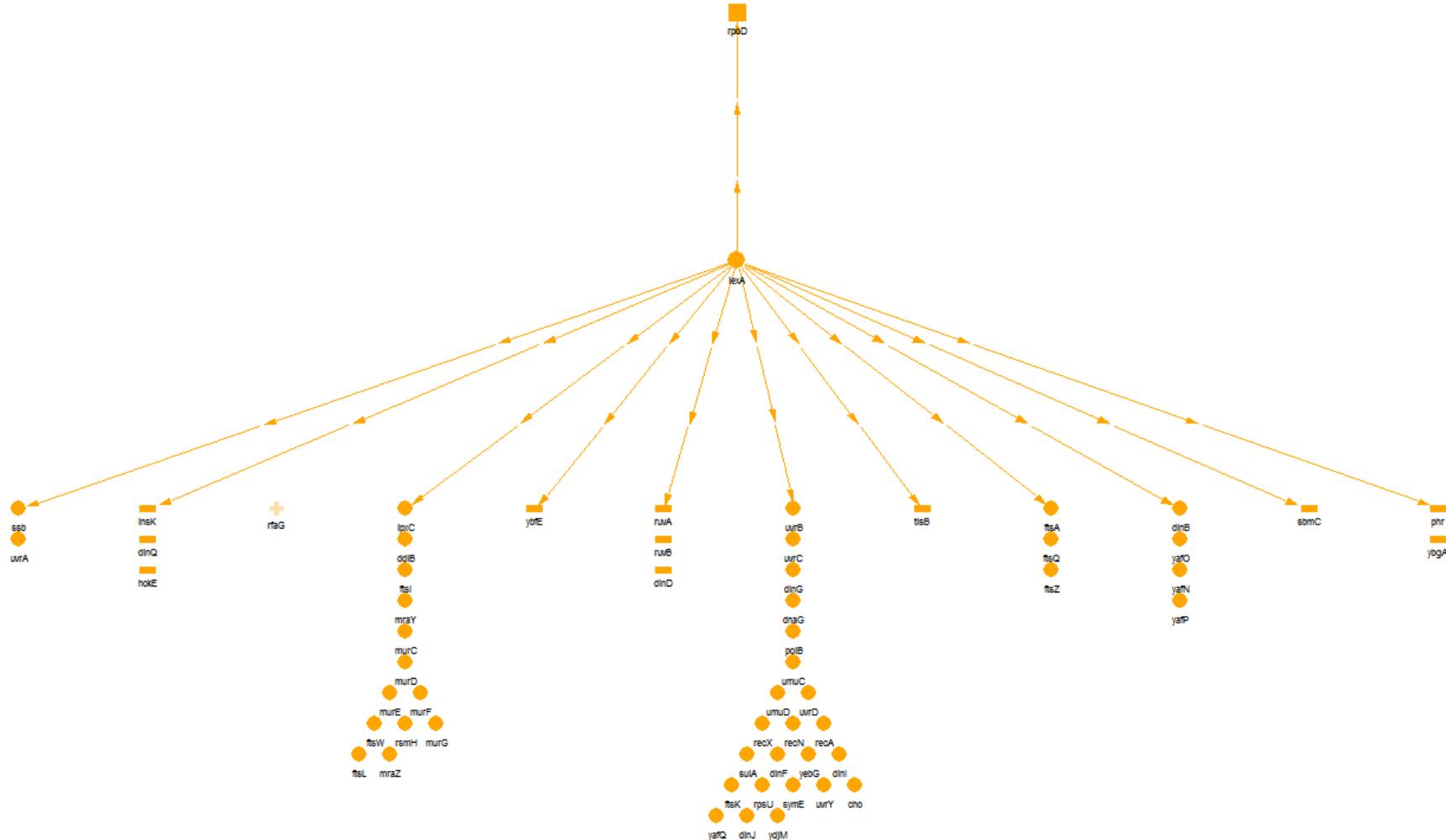


[http://regulondb.ccg.unam.mx/OperonControllerServlet?operon\\_id=ECK120014836&type=jsp](http://regulondb.ccg.unam.mx/OperonControllerServlet?operon_id=ECK120014836&type=jsp)





# SOS RESPONSE REGULON (LexA)





# Prominent Global Response Regulatory Systems

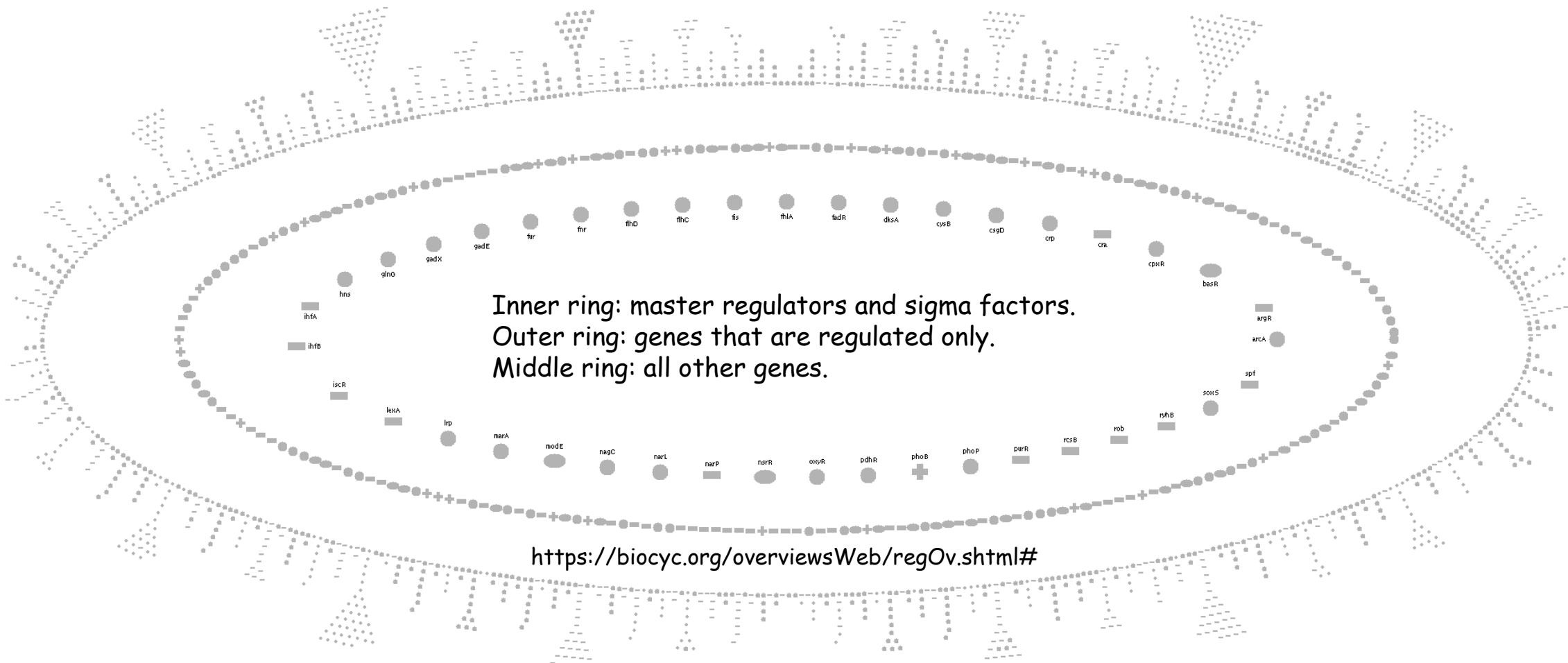
System	Function	Regulatory components	No. and kinds of genes regulated
Catabolite repression (CRP regulon)	Ensure priority use of premium substrates	In gram-negative bacteria, regulatory protein is CRP, nucleotide is cyclic AMP (cAMP); in gram-positive bacteria, regulatory protein is CcpA (no cAMP involvement)	Large numbers of genes encoding catabolic enzymes
Stringent response (RelA regulon)	Divert metabolism to fueling and biosynthesis	RelA and SpoT enzymes make nucleotide, ppGpp; ribosomal RelA acts when ribosome is "hungry" for aminoacyl-tRNA	Perhaps hundreds of genes in fueling and biosynthesis, as well as negative effects on genes encoding polymerization enzymes and RNA
Growth rate control	Adjust transcription and translation to match the growth-promoting ability of the medium	Components of RelA regulon; Fis protein; H-NS protein (see text and Table 13.4)	Hundreds of genes encoding ribosomal proteins, rRNA, and translation factors are adjusted to match the rate of protein synthesis possible in a given medium
Stationary phase (RpoS regulon)	Change cell structure and convert metabolism to new mode to optimize survival during non-growth	RpoS ( $\sigma^{38}$ , $\sigma^S$ ); cooperative with other global regulators, such as CRP, Lrp, H-NS	Direct or indirect effects on most genes of the cell

System	Function	Regulatory components	No. and kinds of genes regulated
Heat shock response (RpoH regulon)	Assist protein folding at high temperature; repair and/or degrade damaged proteins	RpoH ( $\sigma^{32}$ $\sigma^H$ )	Dozens of genes encoding proteases, protein chaperones, and other genes of uncertain function
Envelope stress response (RpoE regulon)	Assist protein folding or degrade unfolded membrane proteins; monitor outer membrane proteins, particularly porins	RpoE ( $\sigma^E$ , $\sigma^{24}$ ), RseA (anti- $\sigma^E$ ), and RseB and the protease DegS and YaeL	Many genes encoding proteases, periplasmic chaperones, and biosynthetic enzymes for the lipid A component of lipopolysaccharide in gram-negative bacteria
Envelope stress response (CpxAR regulon)	Assist protein folding or degrade unfolded membrane proteins; monitor envelope surface proteins, such as pili	CpxA (sensor kinase) and CpxR (response regulator); CpxP, an inhibitor of CpxA	Many genes involved in protein folding and trafficking within the envelope; partially overlapping with the RpoE regulon
Leucine response (Lrp regulon)	Assist cells to adjust to major changes in nutrient availability, particularly amino acids	Lrp (leucine response protein)	Many genes encoding proteins involved in amino acid biosynthesis, degradation, and transport; also genes encoding pilin proteins

M. Schaechter, J. L. Ingraham, F. C. Neidhardt, "Microbe", ASM Press, 2006, pp. 254-255.



# Regulatory Map of *E.coli* K-12 (EcoCyc)

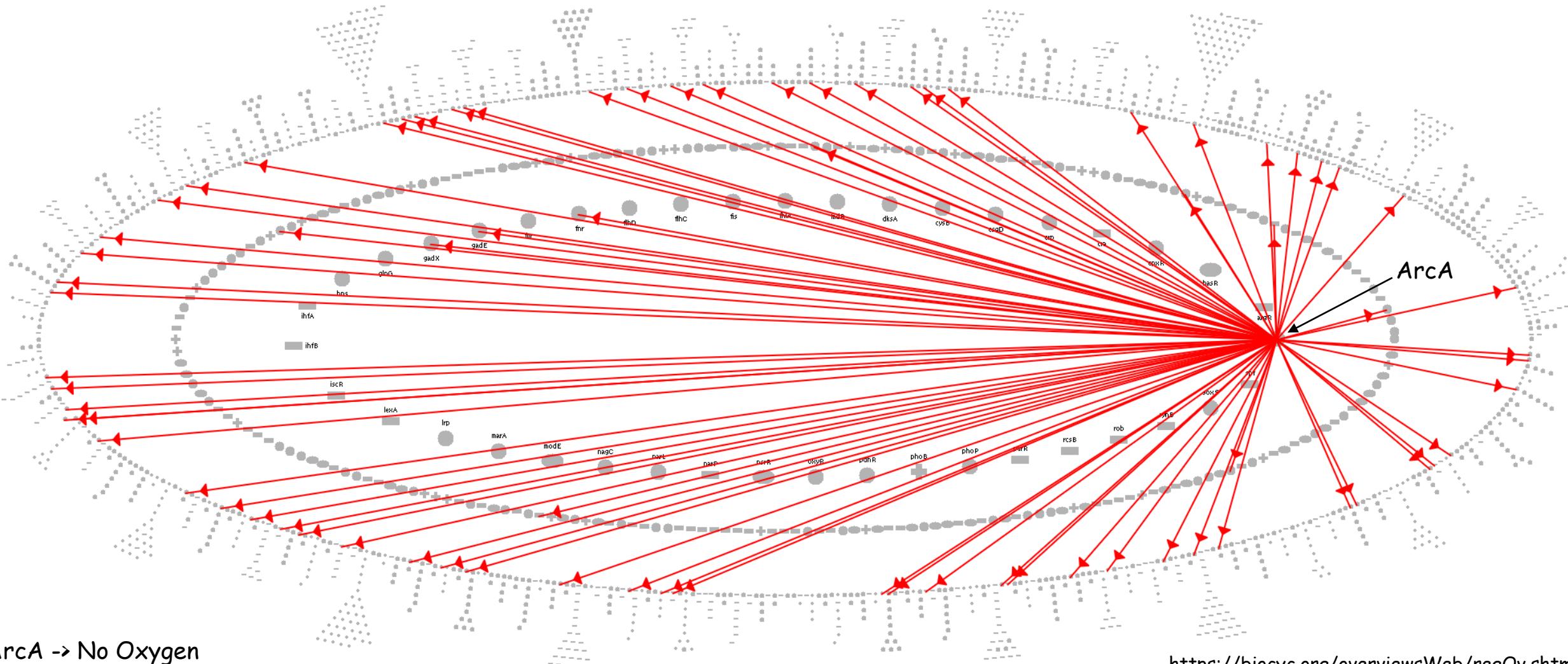


+ / - : a gene that has activators/inhibitors only  
o : a gene that has activators and inhibitors.

Oval : a gene that has all its regulators with an unknown mode of regulation.  
Square : a gene encoding a sigma factor.



# ArcA Impact in *E.coli* K-12 (EcoCyc)



ArcA → No Oxygen

<https://biocyc.org/overviewsWeb/regOv.shtml#>

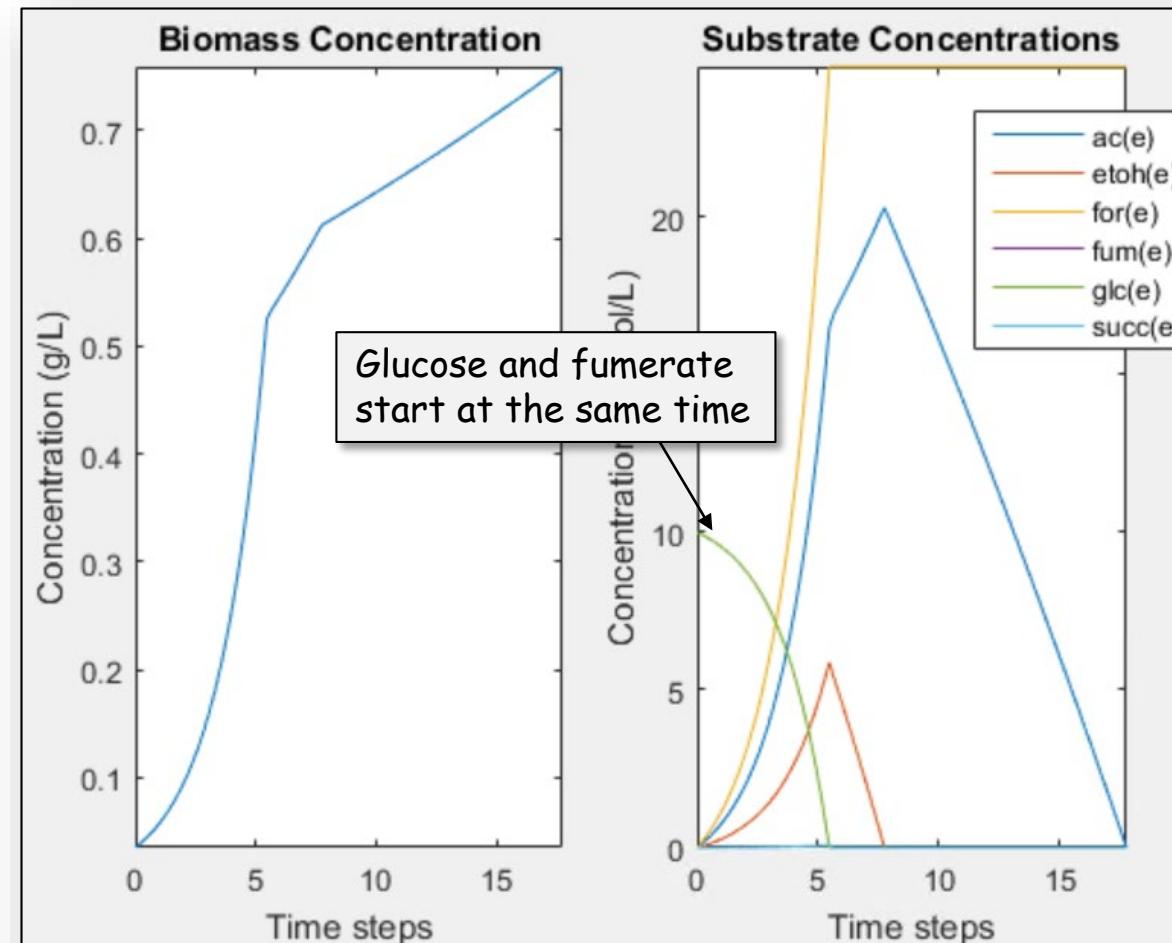


# Lesson Outline

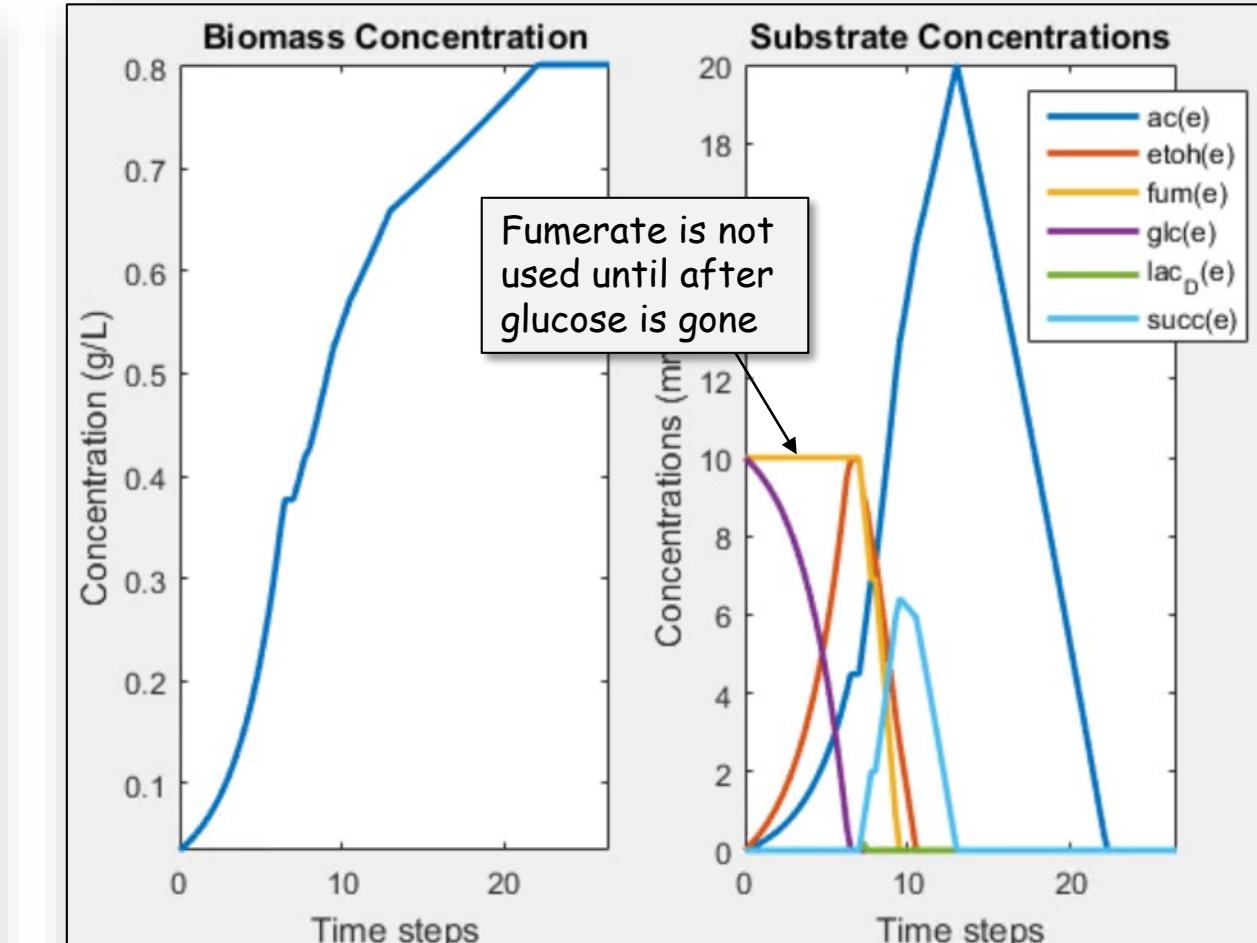
- Transcriptional Regulatory Networks
- • Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



# Comparing dynamicFBA and dynamicRFBA



dynamicGlucoseFumerate\_Core.m

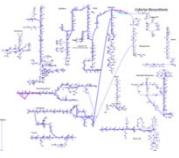


GlucoseFumerateCatabolite\_Repression.m



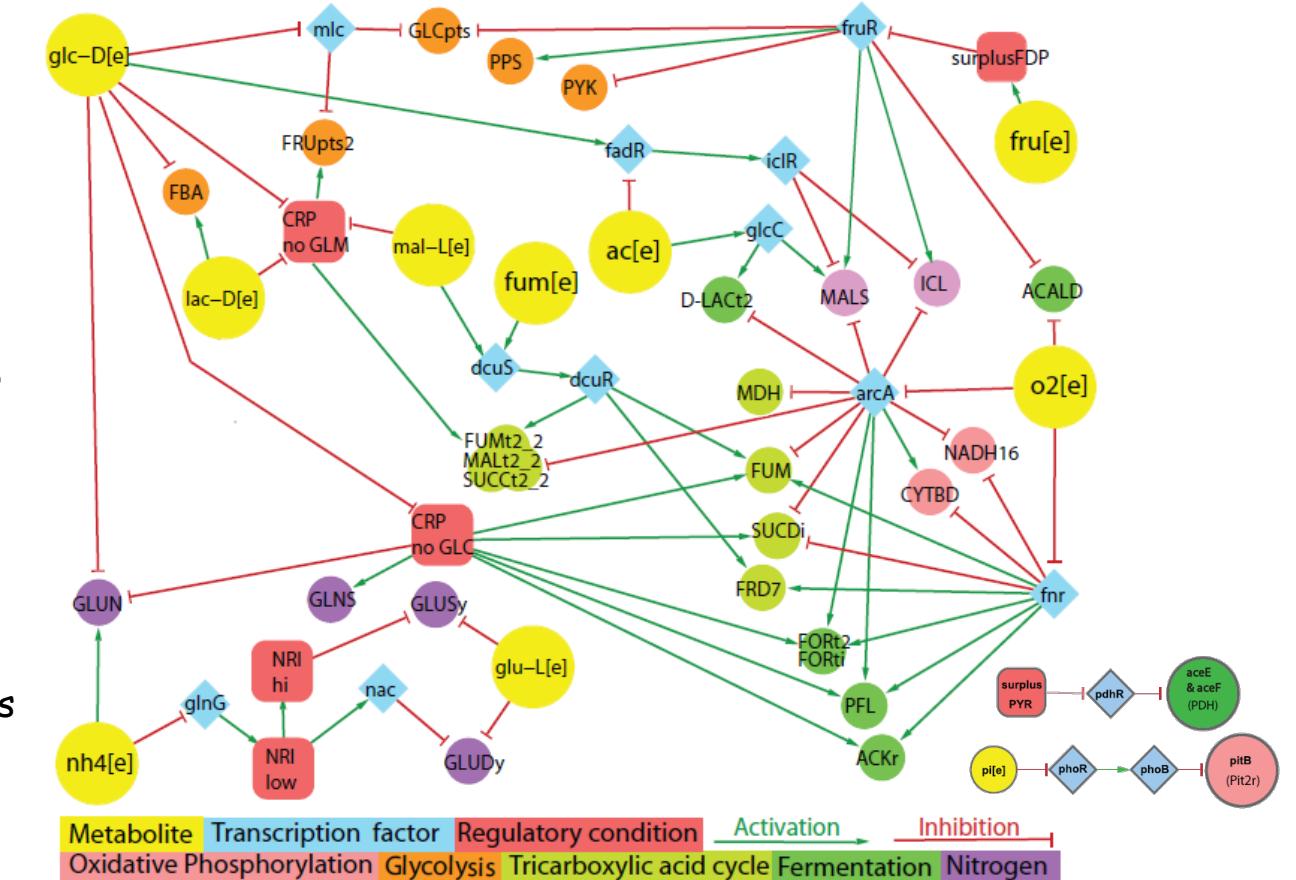
# Boolean Transcriptional Regulatory Networks

- In addition to the metabolic reconstruction, the core *E. coli* model also contains a Boolean representation of part of the associated transcriptional regulatory network.
- In response to external and internal stimuli, *in silico* transcription factors can either activate or repress genes associated with metabolic reactions. This regulation improves the predictive fidelity of the metabolic model by imposing additional context dependent constraints on certain reactions.
- The transcriptional regulatory reconstruction consists of a set of Boolean rules that dictate whether a gene is either fully induced or fully repressed.
- If the genes associated with an enzyme or transport protein/complex are repressed, then *in silico* flux is set to zero for the corresponding reaction. The solution space of the network shrinks when these additional constraints are imposed. Reactions that are not used due to regulatory effects are thus restricted, so when using flux balance analysis, the optimal flux distribution will be consistent with known regulation.
- This optimal flux distribution may be different from the flux distribution of an unregulated model. In this case, the flux distribution of the unregulated model violated at least one regulatory constraint, making it biologically unrealistic.
- The use of computationally implemented Boolean rules in a genome scale model has been shown to lead to more accurate flux balance analysis predictions



# Boolean Regulatory Networks

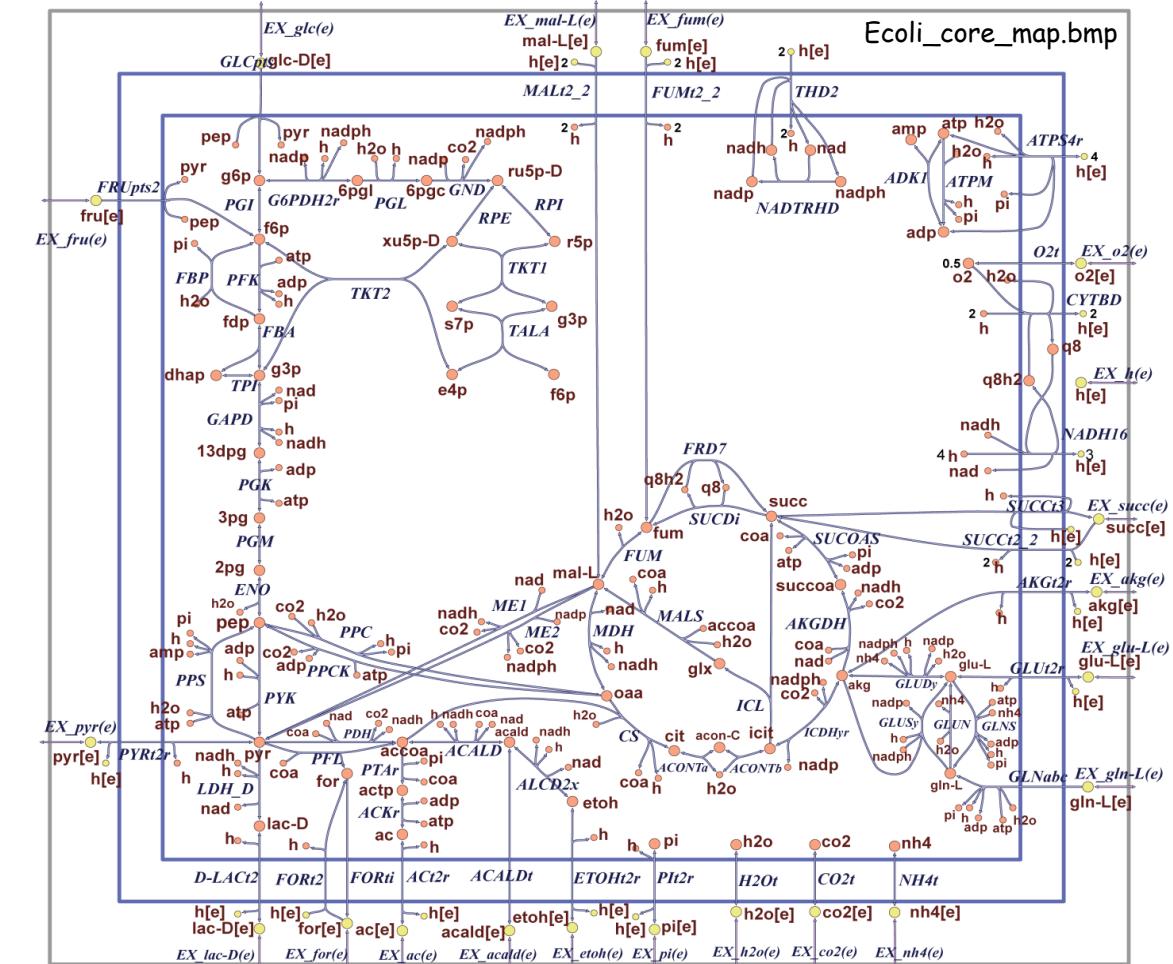
- A gene is considered to be induced when evaluation of the corresponding Boolean rule gives 'true'.
- In contrast, a gene is considered to be repressed, when evaluation of the corresponding Boolean rule gives 'false'.
- Boolean logic is used to evaluate each Boolean rule.
- Complex regulatory conditions will be represented with variables that represent a complex regulatory rule for a transcription factor that cannot be accurately represented with only one variable.
- By using Boolean logic, all rules in a regulatory network can be reduced to either 'true' or 'false', and ultimately this dictates whether each metabolic gene is induced or repressed.
- Not every gene in the metabolic network is controlled by the regulatory network, so the unregulated genes are assumed to always be active, and their fluxes are never constrained to zero.
- Regulatory control only applies to models that include regulatory information, that is, regulated models (modelReg)





# "modelReg" Regulatory Model

- Model name is "modelReg.mat"
- Built on the *E.coli* core model using a subset of the iMC1010 regulatory genes
- 153 total genes
- 52 regulatory genes
- 6 special regulatory variables
- 159 total genes & regulatory variables
- 11 external metabolites monitored
- 12 internal fluxes monitored
- Regulates the expression of 78 genes
- Works with the *E.coli* core map



Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# What is a dynamicRFBA Model?

A regulated model will include the following extra cells

- A list of the regulatory genes
  - ✓ b4014, b4015, etc.
- A list of the external metabolites that are regulatory inputs (regulatoryInputs1)
- A list of the internal reactions that are regulatory changes (regulatoryInputs2)
- A list of the regulatory rules
  - ✓ true
  - ✓ Crp
  - ✓ NOT ArcA
  - ✓ NOT PdhR OR Fis
  - ✓ CRPnoGLM AND (NOT ArcA) AND DcuR

Field	Value
rxns	95x1 cell
mets	72x1 cell
S	72x95 sparse dou...
rev	95x1 double
Ib	95x1 double
ub	95x1 double
c	95x1 double
metCharge	72x1 int32
rules	95x1 cell
genes	137x1 cell
rxnGeneMat	95x137 sparse do...
grRules	95x1 cell
subSystems	95x1 cell
confidenceScores	95x1 cell
rxnReferences	95x1 cell
rxnECNumbers	95x1 cell
rxnNotes	95x1 cell
rxnNames	95x1 cell
metNames	72x1 cell
metFormulas	72x1 cell
metChEBIID	72x1 cell
metKEGGID	72x1 cell
metPubChemID	72x1 cell
metInChIString	72x1 cell
b	72x1 double
description	'ecoli_textbook'

Field	Value
rxns	95x1 cell
mets	72x1 cell
S	72x95 sparse dou...
rev	95x1 double
Ib	95x1 double
ub	95x1 double
c	95x1 double
rules	95x1 cell
genes	137x1 cell
rxnGeneMat	95x137 sparse do...
grRules	95x1 cell
subSystems	95x1 cell
rxnNames	95x1 cell
metNames	72x1 cell
metFormulas	72x1 cell
b	72x1 double
description	'Ecoli core model'
regulatoryGenes	159x1 cell
regulatoryInputs1	11x1 cell
regulatoryInputs2	12x1 cell
regulatoryRules	159x1 cell

See core\_regulatory\_rules.xls



# Actively Monitored Regulatory Inputs

## External Metabolite Inputs

A list of the external metabolites that are regulatory inputs (regulatoryInputs1)

o2[e]

glu-L[e]

glc-D[e]

nh4[e]

succ[e]

fum[e]

mal-L[e]

ac[e]

pi[e]

lac-D[e]

fru[e]

## Internal Reaction Inputs

A list of the internal reactions that are regulatory changes (regulatoryInputs2)

FBP

TKT2

TALA

PGI

ME2

ME1

GLCpts

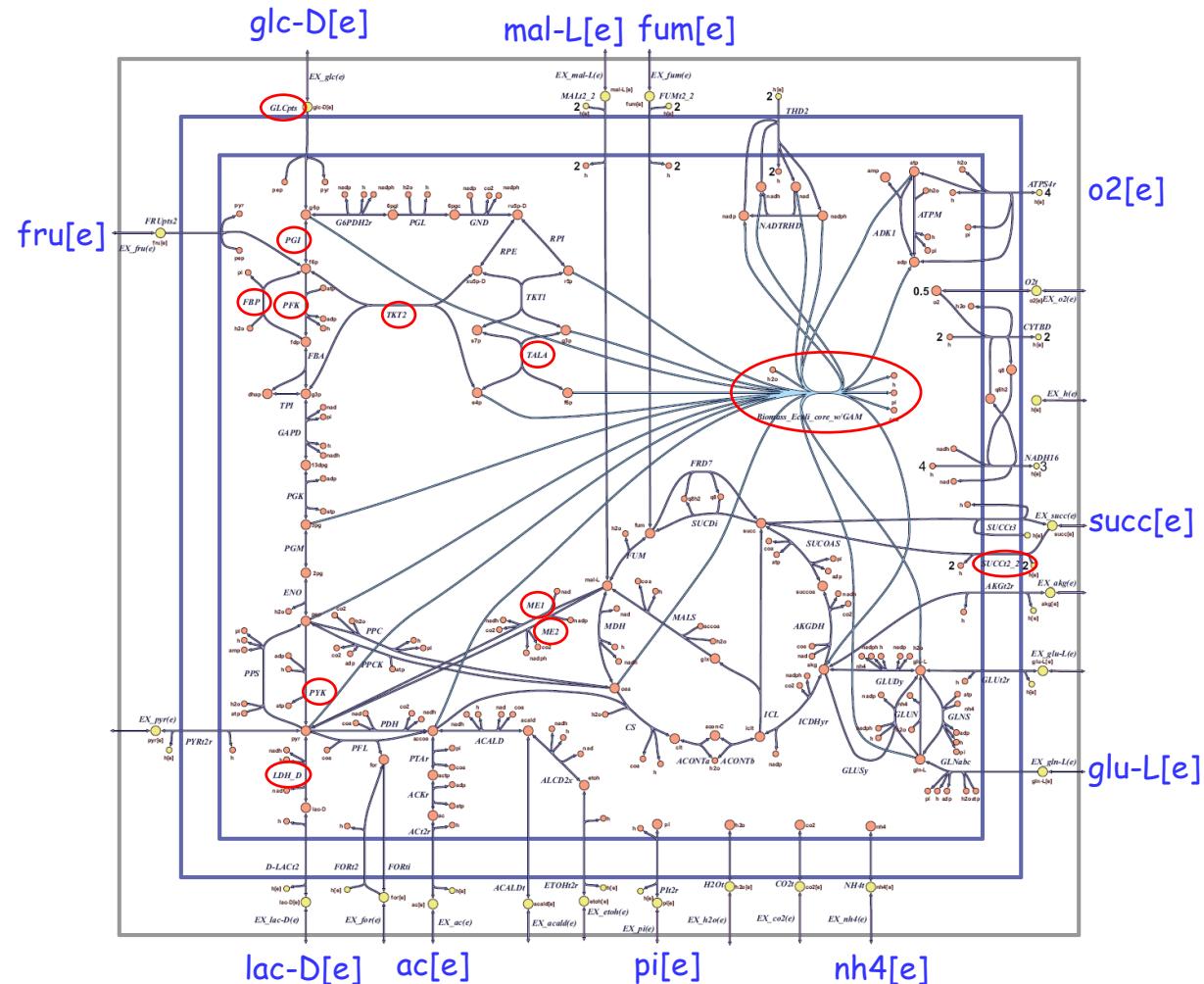
PYK

PFK

LDH\_D

SUCCt2\_2

Biomass\_Ecoli\_core\_w\_GAM





## Transcription Factors

bNum	Gene	Rule
b4401	ArcA	NOT o2[e]
b3357	Crp	CRPnoGLC
b4124	DcuR	DcuS
b4125	DcuS	succ[e] OR fum[e] OR mal-L[e]
b1187	FadR	glc-D[e] OR (NOT ac[e])
b3261	Fis	Biomass_Ecoli_core_w_GAM
b1334	Fnr	NOT o2[e]
b0080	FruR (cra)	NOT surplusFDP
b2980	GlcC	ac[e]
b3868	GlnG	NOT nh4[e]
b4018	IclR	FadR
b1594	Mlc	NOT glc-D[e]
b1988	Nac	NRI_low
b0113	PdhR	NOT surplusPYR
b0399	PhoB	PhoR
b0400	PhoR	NOT pi[e]
	CRPnoGLC	NOT glc-D[e]
	CRPnoGLM	NOT (glc-D[e] OR mal-L[e] OR lac-D[e])
	NRI_hi	NRI_low
	NRI_low	GlnG
	surplusFDP	((NOT FBP) AND (NOT (TKT2 OR TALA OR PGI))) OR fru[e]
	surplusPYR	(NOT (ME2 OR ME1)) AND (NOT (GLCpts OR PYK OR PFK OR LDH_D OR SUCCt2_2))

## Regulatory Genes in the Regulated Core Model

core\_regulatory\_rules.xls

Variables that simulate more complex regulatory operations



# Regulatory Rules in Matlab Boolean Logic

ModelReg Matlab Boolean Logic.xlsx

bNum	Gene	Rule	Matlab Expression
b4401	ArcA	NOT o2[e]	$\sim o2[e]$
b3357	Crp	CRPnoGLC	CRPnoGLC
b4124	DcuR	DcuS	b4125
b4125	DcuS	succ[e] OR fum[e] OR mal-L[e]	succ[e]    fum[e]    mal-L[e]
b1187	FadR	glc-D[e] OR (NOT ac[e])	glc-D[e]    ( $\sim ac[e]$ )
b3261	Fis	Biomass_Ecoli_core_w_GAM	Biomass_Ecoli_core_w_GAM
b1334	Fnr	NOT o2[e]	$\sim o2[e]$
b0080	FruR	NOT surplusFDP	$\sim surplusFDP$
b2980	GlcC	ac[e]	ac[e]
b3868	GlnG	NOT nh4[e]	$\sim nh4[e]$
b4018	IclR	FadR	b1187
b1594	Mlc	NOT glc-D[e]	$\sim glc-D[e]$
b1988	Nac	NRI_low	NRI_low
b0113	PdhR	NOT surplusPYR	NOT surplusPYR
b0399	PhoB	PhoR	b0400
b0400	PhoR	NOT pi[e]	$\sim pi[e]$
CRPnoGLC	CRPnoGLC	NOT glc-D[e]	$\sim glc-D[e]$
CRPnoGLM	CRPnoGLM	NOT (glc-D[e] OR mal-L[e] OR lac-D[e])	$\sim (glc-D[e]    mal-L[e]    lac-D[e])$
NRI_hi	NRI_hi	NRI_low	NRI_low
NRI_low	NRI_low	GlnG	b3868
surplusFDP	surplusFDP	((NOT FBP) AND (NOT (TKT2 OR TALA OR PGI))) OR fru[e]	(( $\sim FBP$ ) && ( $\sim (TKT2    TALA    PGI)$ ))    fru[e]
surplusPYR	surplusPYR	(NOT (ME2 OR ME1)) AND (NOT (GLCpts OR PYK OR PFK OR LDH_D OR SUCCt2_2))	(( $\sim ME2    ME1$ )) && (( $\sim (GLCpts    PYK    PFK    LDH_D    SUCCt2_2)$ ))



# Genes & Regulatory Rules for *E.coli* Core Regulatory Model

Core Genes and Regulatory Rules.xlsx

Gene	Regulatory Rule
'b0008'	'true'
'b0080'	'~surplusFDP'
'b0113'	'~surplusPYR'
'b0114'	'~b0113    b3261'
'b0115'	'~b0113    b3261'
'b0116'	'true'
'b0118'	'true'
'b0351'	'true'
'b0356'	'true'
'b0399'	'b0400'
'b0400'	'~pi[e]'
'b0451'	'true'
'b0474'	'true'
'b0485'	'true'
'b0720'	'true'
'b0721'	'(~(b4401    b1334))    b3357    b3261'
'b0722'	'(~(b4401    b1334))    b3357    b3261'
'b0723'	'(~(b4401    b1334))    b3357    b3261'
'b0724'	'(~(b4401    b1334))    b3357    b3261'
'b0726'	'true'
'b0727'	'true'
'b0728'	'true'
'b0729'	'true'
'b0733'	'(~b1334)    b4401'
'b0734'	'(~b1334)    b4401'
'b0755'	'true'
'b0767'	'true'
'b0809'	'true'
'b0810'	'true'
'b0811'	'true'
'b0875'	'true'
'b0902'	'b4401    (b1334 && b3357)'
'b0903'	'b4401    (b1334 && b3357)'
'b0904'	'b4401    (b1334 && b3357)'
'b0978'	'true'

Regulated Expression

Constitutive Expression

Gene	Regulatory Rule
'b3738'	'true'
'b3739'	'true'
'b3868'	'~nh4[e]'
'b3870'	'b3357'
'b3916'	'true'
'b3919'	'true'
'b3925'	'true'
'b3951'	'b4401    b1334'
'b3952'	'b4401    b1334'
'b3956'	'true'
'b3962'	'true'
'b4014'	'(~b4018) && ((~b4401)    b0080)'
'b4015'	'(~b4018) && ((~b4401)    b0080)'
'b4018'	'b1187'
'b4025'	'true'
'b4077'	'true'
'b4090'	'true'
'b4122'	'b1334    b3357    b4124'
'b4124'	'b4125'
'b4125'	'succ[e]    fum[e]    mal-L[e]'
'b4151'	'b1334    b4124'
'b4152'	'b1334    b4124'
'b4153'	'b1334    b4124'
'b4154'	'b1334    b4124'
'b4232'	'true'
'b4301'	'true'
'b4395'	'true'
'b4401'	'~o2[e]'
's0001'	'true'
'CRPnoGLC'	'~glc-D[e]'
'CRPnoGLM'	'~(glc-D[e]    mal-L[e]    lac-D[e])'
'NRI_hi'	'NRI_low'
'NRI_low'	'b3868'
'surplusFDP'	'((~FBP) && ((~TKT2    TALA    PGI)))    fru[e]'
'surplusPYR'	'(~(ME2    ME1)) && ((~GLCpts    PYK    PFK    LDH_D    SUCCt2_2))'

External Metabolite Inputs

Internal Reaction Inputs



## Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)

```
[concentrationMatrix,excRxnNames,timeVec,biomassVec,drGenes,constrainedRxns,states] = ...
    dynamicRFBA(model,substrateRxns,initConcentrations,initBiomass,timeStep,nSteps,plotRxns,exclUptakeRxns)
```

model	a regulatory COBRA model
substrateRxns	list of exchange reaction names for substrates initially in the media that may change (i.e. not h <sub>2</sub> O or co <sub>2</sub> )
initConcentrations	initial concentrations of substrates (in the same structure as substrateRxns)
initBiomass	initial biomass
timeStep	time step size
nSteps	maximum number of time steps
plotRxns	reactions to be plotted
exclUptakeRxns	list of uptake reactions whose substrate concentrations do not change (opt, default {'EX_co2(e)', 'EX_o2(e)', 'EX_h2o(e)', 'EX_h(e)'} )
concentrationMatrix	matrix of extracellular metabolite concentrations
excRxnNames	names of exchange reactions for the EC metabolites
timeVec	vector of time points
biomassVec	vector of biomass values
drGenes	vector of downregulated genes
constrainedRxns	vector of downregulated reactions
states	vector of regulatory network states

If no initial concentration is given for a substrate that has an open uptake in the model (i.e. model.lb < 0) the concentration is assumed to be high enough to not be limiting. If the uptake rate for a nutrient is calculated to exceed the maximum uptake rate for that nutrient specified in the model and the max uptake rate specified is > 0, the maximum uptake rate specified in the model is used instead of the calculated uptake rate.



# optimizeRegModel

```
% Test_modelReg.m
clear; clc;
model = readCbModel('modelReg.mat');

% Initial Conditions
model = changeRxnBounds(model, 'EX_glc(e)', -10, '1');
model = changeRxnBounds(model, 'EX_fum(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -30, '1');

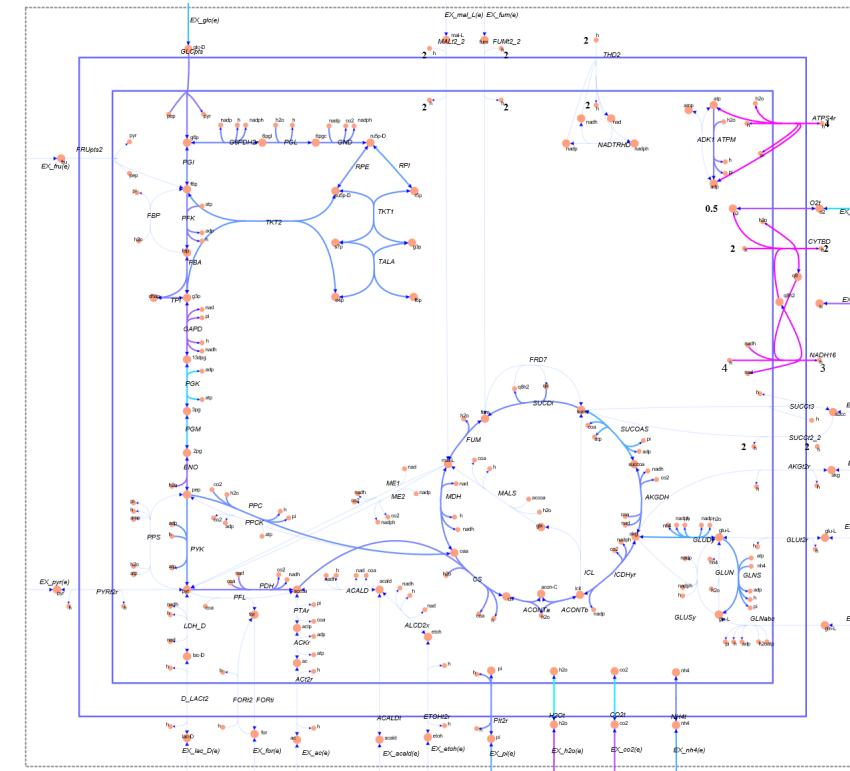
% FBA analysis using regulated constraints
[FBAsol, DRgenes, constrainedRxns, cycleStart, states] = optimizeRegModel(model);

% Create a statevector showing the states created in the regulatory process
stateVector = [model.regulatoryGenes; model.regulatoryInputs1; model.regulatoryInputs2];

% Print the regulatory states
printLabeledData(stateVector, [states{1,1}, states{1,2}, states{1,3}, states{1,4}, states{1,5}, states{1,6}, states{1,7}]);

% Print the regulated optimized FBA fluxes
printFluxVector(model, FBAsol{1,1}.x, true)

% Draw the flux values on the map "target.svg"
map=readCbMap('ecoli_Textbook_ExportMap');
options.zeroFluxWidth = 0.1;
options.rxnDirMultiplier = 10;
drawFlux(map, model, FBAsol{1,1}.x, options);
```



b0008	0	1	1	1	1	1	1
b0080	0	1	0	0	1	1	1
b0113	0	1	0	0	1	1	1
b0114	0	1	0	1	1	1	1
b0115	0	1	0	1	1	1	1
b0116	0	1	1	1	1	1	1
b0118	0	1	1	1	1	1	1
b0351	0	1	1	1	1	1	1



# Monitoring Calculation of Regulatory States

Gene	Iterations towards Anaerobic-Glucose Regulatory State						Final State
	State #1	State #2	State #3	State #4	State #5	State #6	
b0007	0	1	1	1	1	1	1
b0008	0	1	1	1	1	1	1
b0019	0	0	1	1	1	1	1
b0020	0	0	1	1	1	1	1
b0025	0	1	1	1	1	1	1
b0029	0	1	1	1	1	1	1
b0031	0	1	1	1	1	1	1
b0032	0	1	1	1	1	1	1
b0033	0	0	0	0	0	0	0
b0034	0	0	1	0	0	1	1
b0036	0	0	0	0	0	0	0
b0038	0	0	0	0	0	0	0
b0040	0	0	0	0	0	0	0
b0048	0	1	1	1	1	1	1
b0049	0	1	1	1	1	1	1

Gene is "on"

Gene is "off"

Previous  
Gene  
State

The environment changes the on/off status of a gene, that gene change can trigger another gene to change, which can lead to more changes. This iteration of states eventually allows the regulatory control to achieve a steady state for the final gene state.

Final  
Gene  
State



# Growth with Glucose Substrate in Low Aerobic Environment

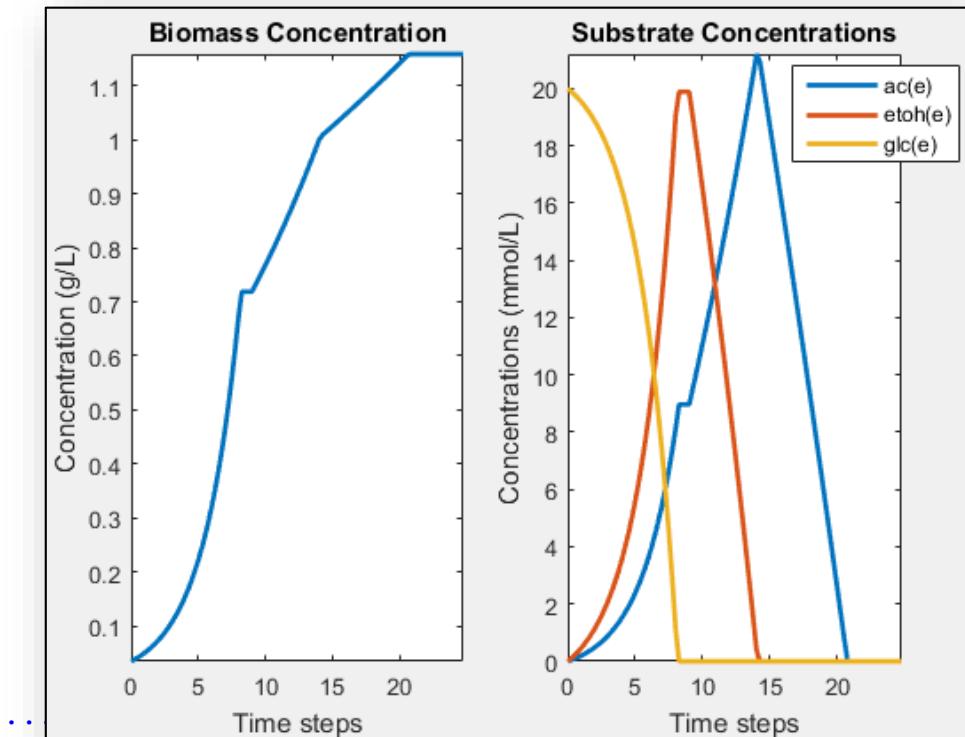
```
% Glucose_Low_Aerobic.m
clear;

model = readCbModel('modelReg.mat');
modelReg = changeRxnBounds(modelReg, 'EX_glc(e)', -10, '1');
modelReg = changeRxnBounds(modelReg, 'EX_o2(e)', -5, '1');

substrateRxns = {'EX_glc(e)'};
initConcentrations = [20];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoH(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', ...
    'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxn, states] = ...
    dynamicRFBA(modelReg, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

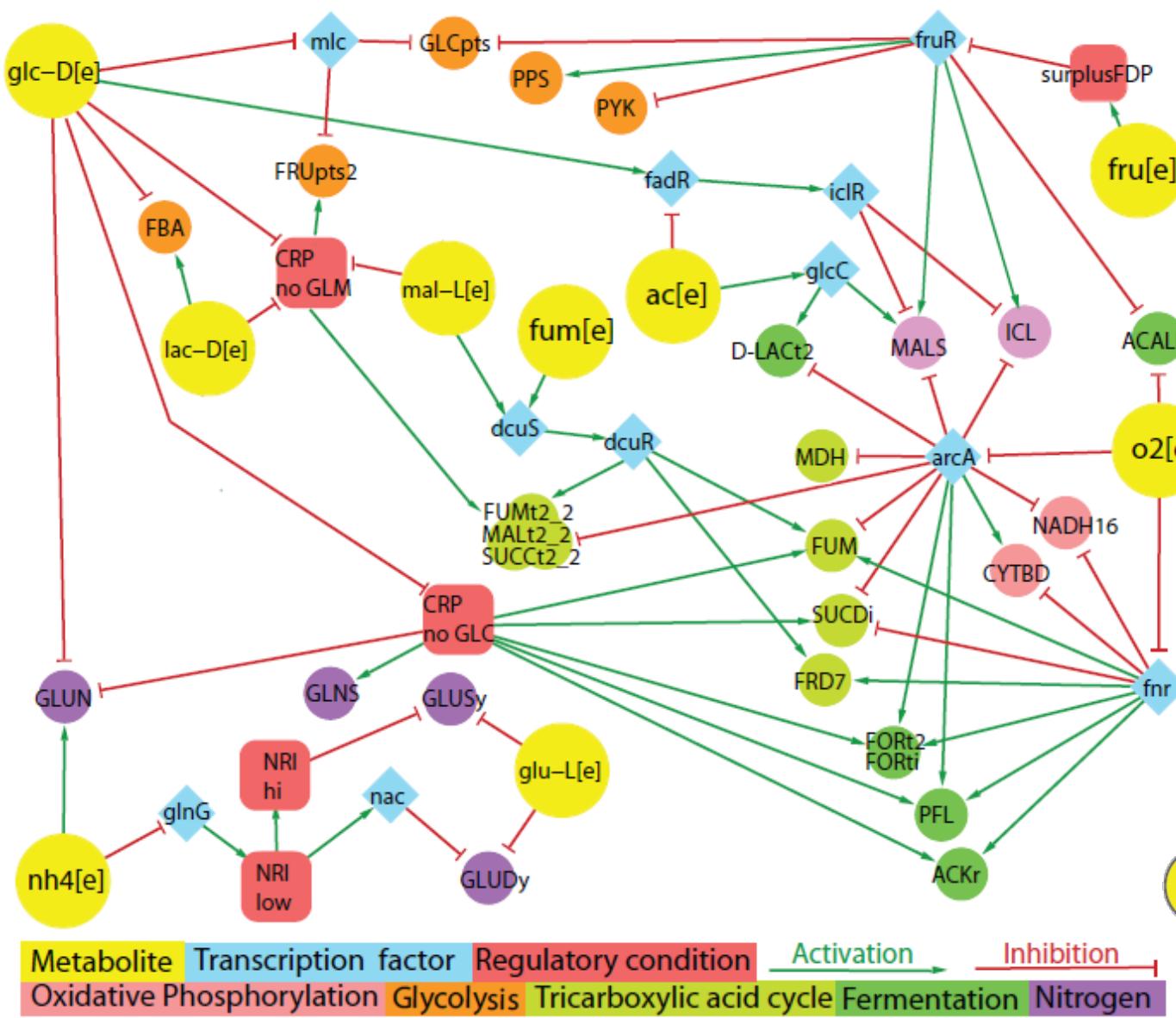
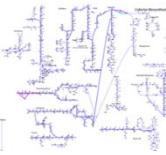
% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```



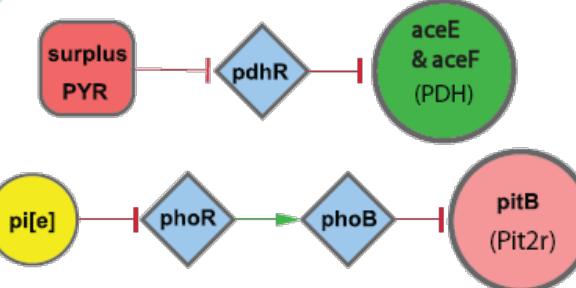


# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- • Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



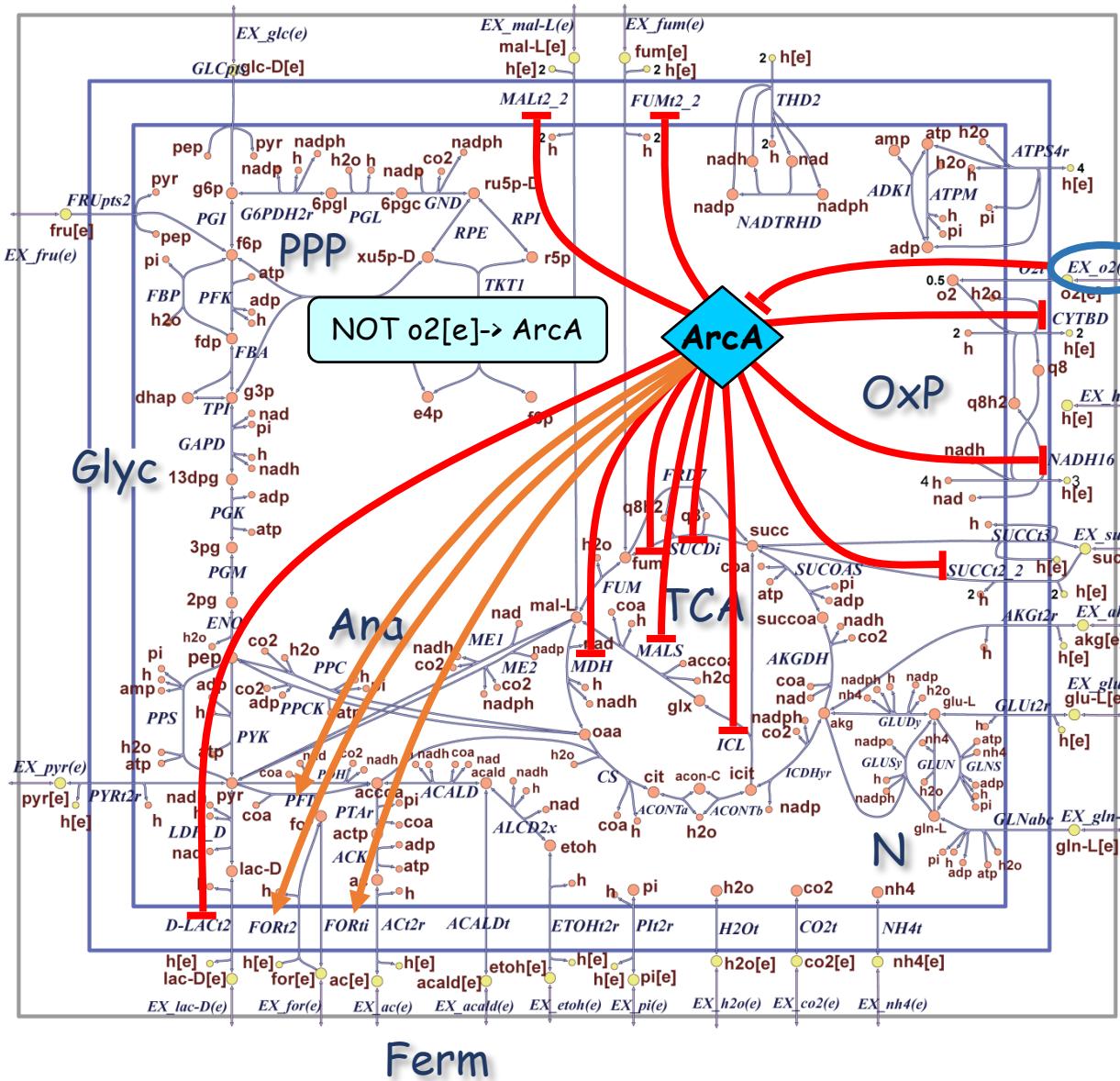
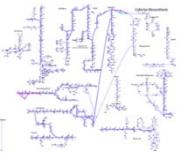
# Regulatory Map of the Regulated Core Model





# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
- ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



## Arca

- ARcA is a transcriptional regulator
- Low oxygen availability signals the activation of the global regulator ArcA.
- Represses the transporters for malate, fumarate, lactate, and succinate.
- Downregulates the glyoxylate cycle
- Downregulates the energy producing portion of the TCA cycle
- Upregulates the fermentation pathway for formate
- Downregulates oxidative phosphorylation

Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Constraint-based Metabolic Reconstructions & Analysis

2021 H. Scott Hinton

38

 EcoCyc
A member of the BioCyc database collection

Please take 3 minute BioCyc survey

[LOGIN | Utah State University sub](#)

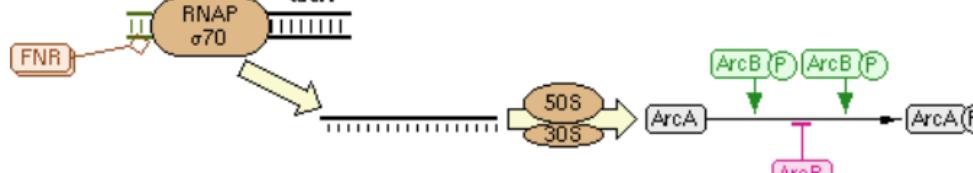
Searching *Escherichia coli* K-12 substr. MG1655 (EcoCyc)

[Sites](#) | [Search](#) | [Genome](#) | [Metabolism](#) | [Analysis](#) | [SmartTables](#) | [Help](#)

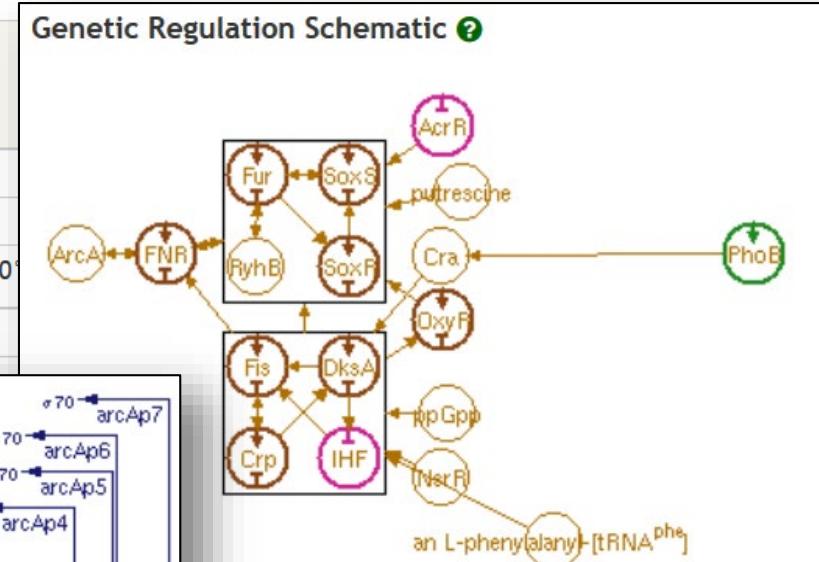
gene	polypeptide		
<b>arcA</b>	<b>ArcA transcriptional dual regulator</b>		
<i>Escherichia coli</i> K-12 substr. MG1655			
Synonyms	sfrA, cpxC, dye, fexA, msp, seg, dye resistance protein, ArcA response regulator		
Accession IDs	EG10061 (EcoCyc) b4401 ECK4393 P0A9Q1 (UniProt)	Length	717 bp / 238 aa
		Map Position	[4,639,590 <- 4,640,306] (99.96 centisomes, 360°)
		Location	cytosol
Reactions	$\text{ArcA} + \text{ArcB-Phis}^{717} \rightarrow \text{ArcA-Pasp}^{54} + \text{ArcB}$ $\text{ArcB-Phis}^{292} + \text{ArcA} \rightarrow \text{ArcA-Pasp}^{54} + \text{ArcB}$		
Pathway	ArcAB Two-Component Signal Transduction System, quinone		
Evidence	 Gene expression analysis [Cho06]  Assay of protein purified to homogeneity [Cho06]		

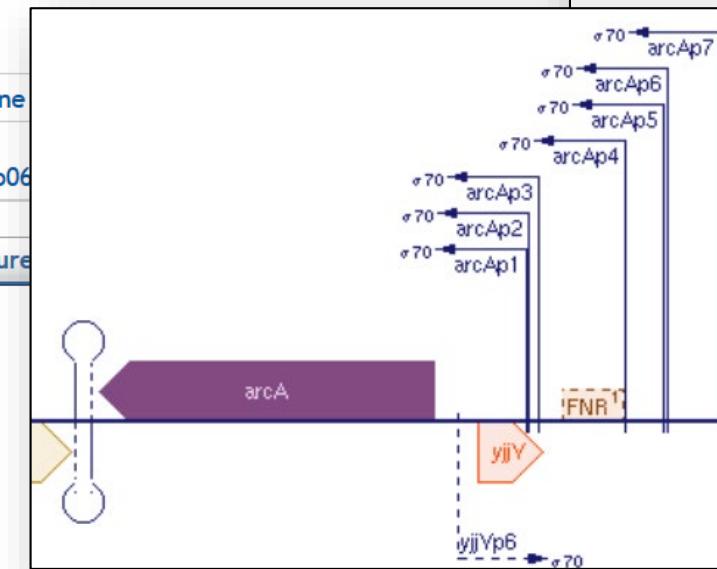
[Summary](#) [GO Terms \(12\)](#) [Essentiality](#) [Reactions \(2\)](#) [Regulon](#) [Protein Feature](#)

**Regulation Summary Diagram**



**Genetic Regulation Schematic**





<https://biocyc.org/gene?orgid=ECOLI&id=EG10061>

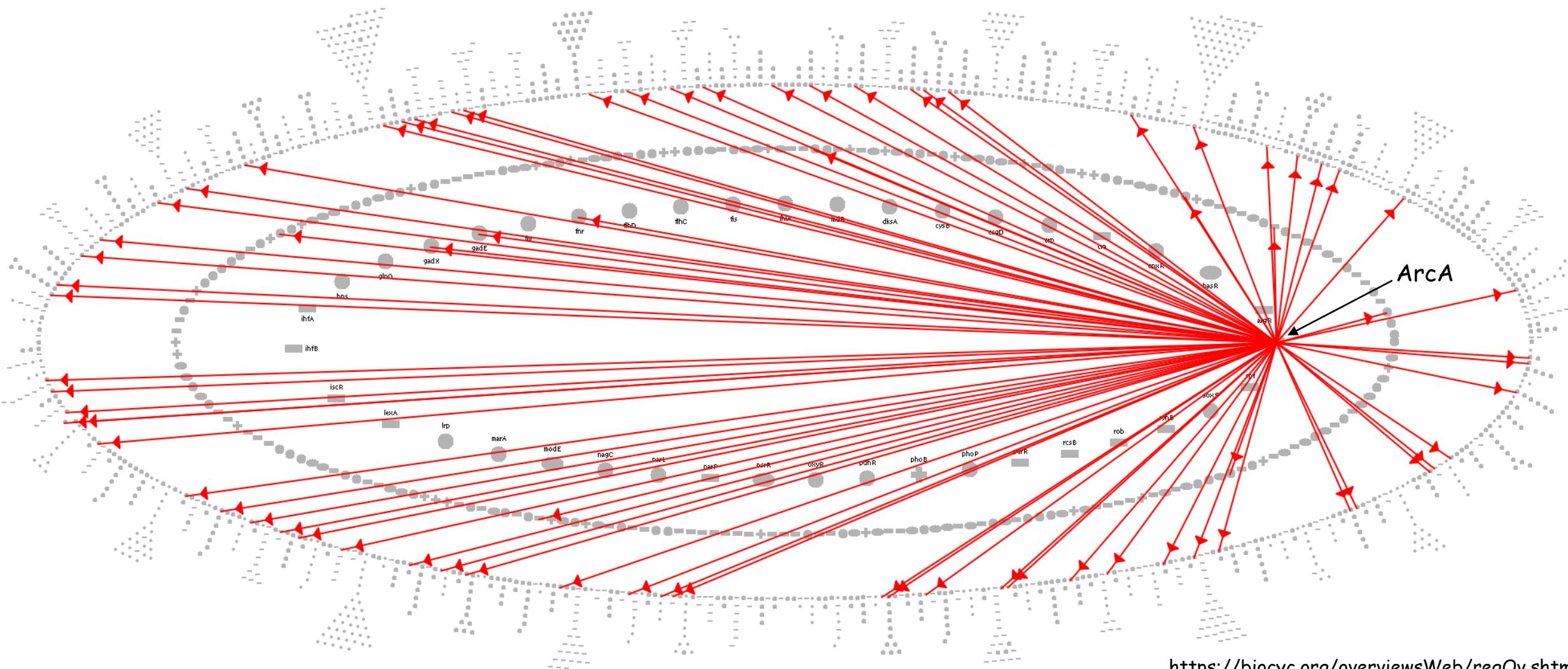
Utah State University

BIE 5500/6500

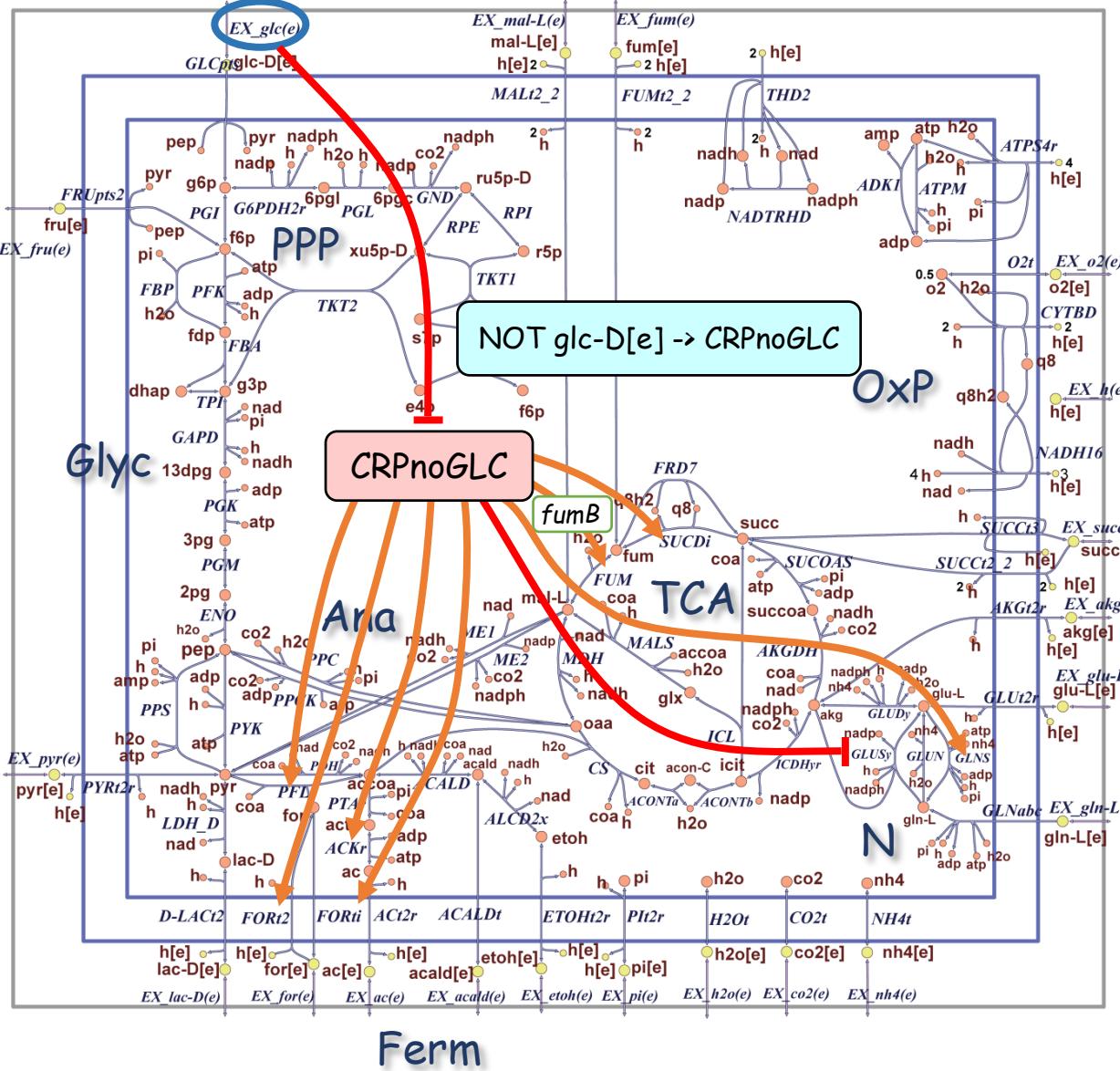
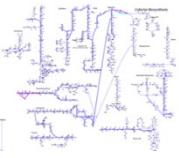
Lesson: Transcriptional Regulatory Networks



# ArcA Impact in *E.coli* K-12 (EcoCyc)



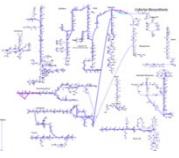
<https://biocyc.org/overviewsWeb/regOv.shtml#>



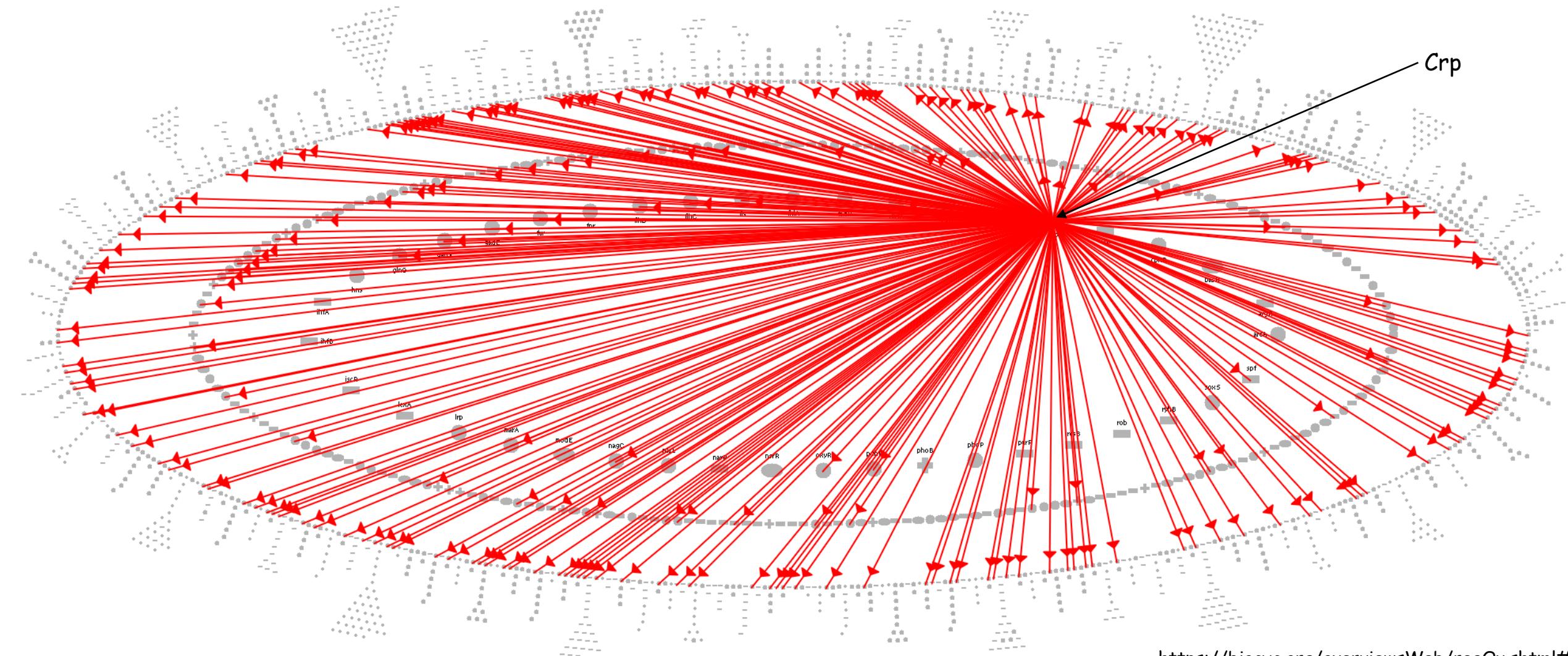
# CRPnoGLC

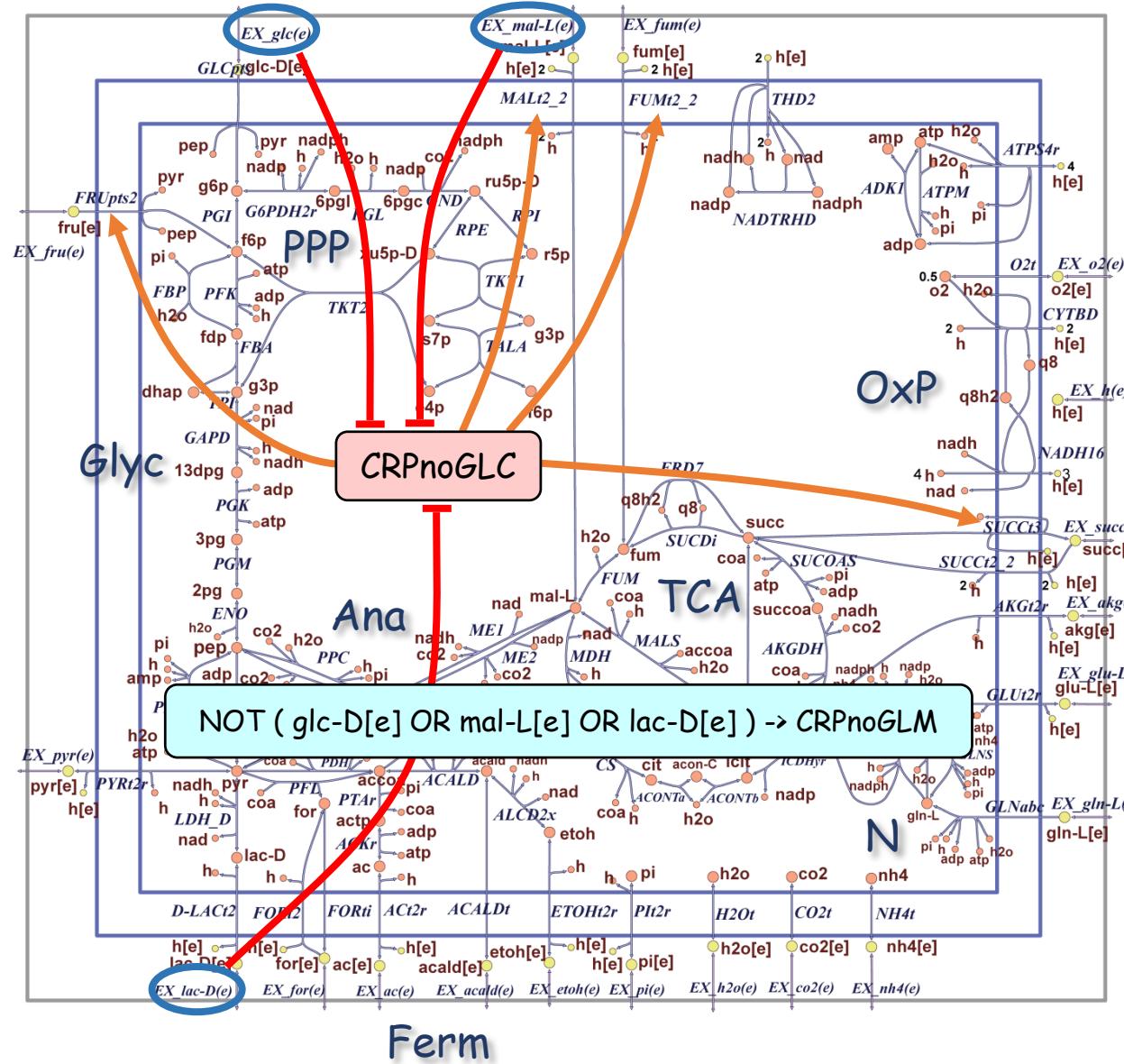
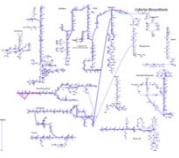
- The activity of the cAMP receptor protein, Crp is modeled when no glucose is present in the media using CRPnoGLC.
- Upregulates the reductive pathway in the TCA cycle (fumB)
- Upregulates the formate and acetate fermentation pathways
- Upregulates the conversion from glutamate to glutamine
- Downregulates the conversion from glutamine to glutamate

Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Crp Impact in *E.coli* K-12 (EcoCyc)

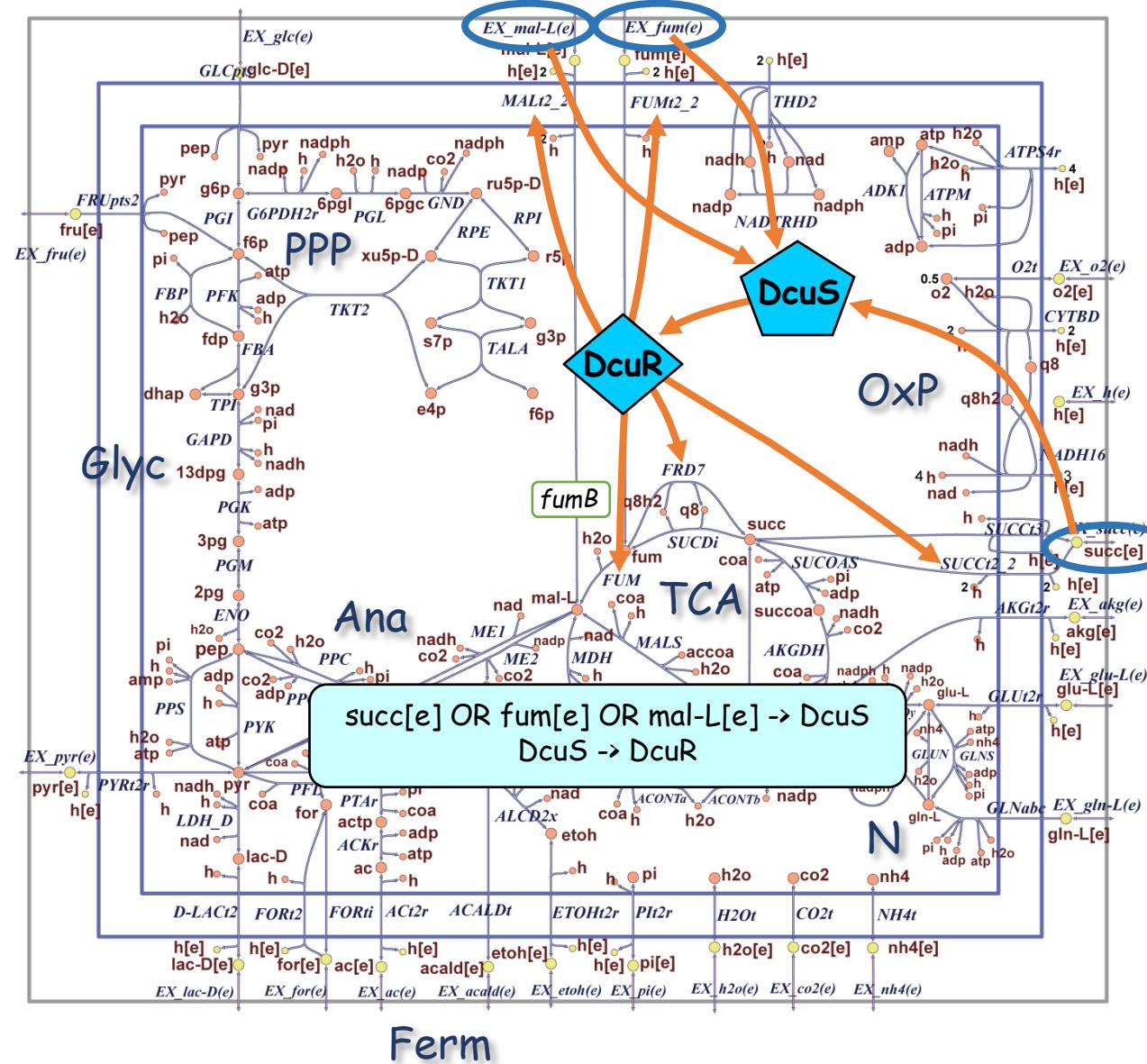




## CRPnoGLM

- GLM stands for glucose, lactate or malate
- The activity of the cAMP receptor protein, Crp is modeled when **no** glucose, malate or lactate are present in the media using CRPnoGLM.
- Upregulates the transport pathways for fructose, malate, fumarate, and succinate

Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



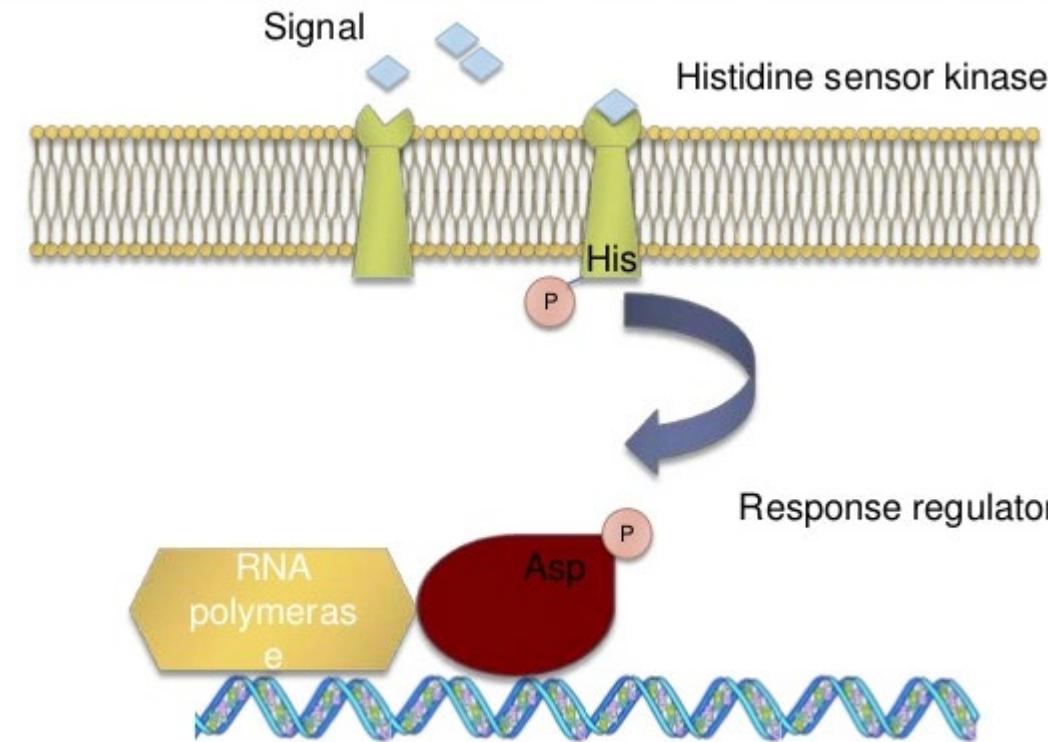
# DcuR & DcuS

- DcuR is a transcriptional regulator
  - DcuS is a sensory histidine kinase
  - Activated when malate, fumarate, or succinate are present in the media.
  - DcuS and DcuR form a two component histidine kinase system.
  - Upregulates the reductive pathway in the TCA cycle (fumB)
  - Upregulates the transport pathways for malate, fumarate, and succinate ( $C_4$ -dicarboxylate compounds)

## Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



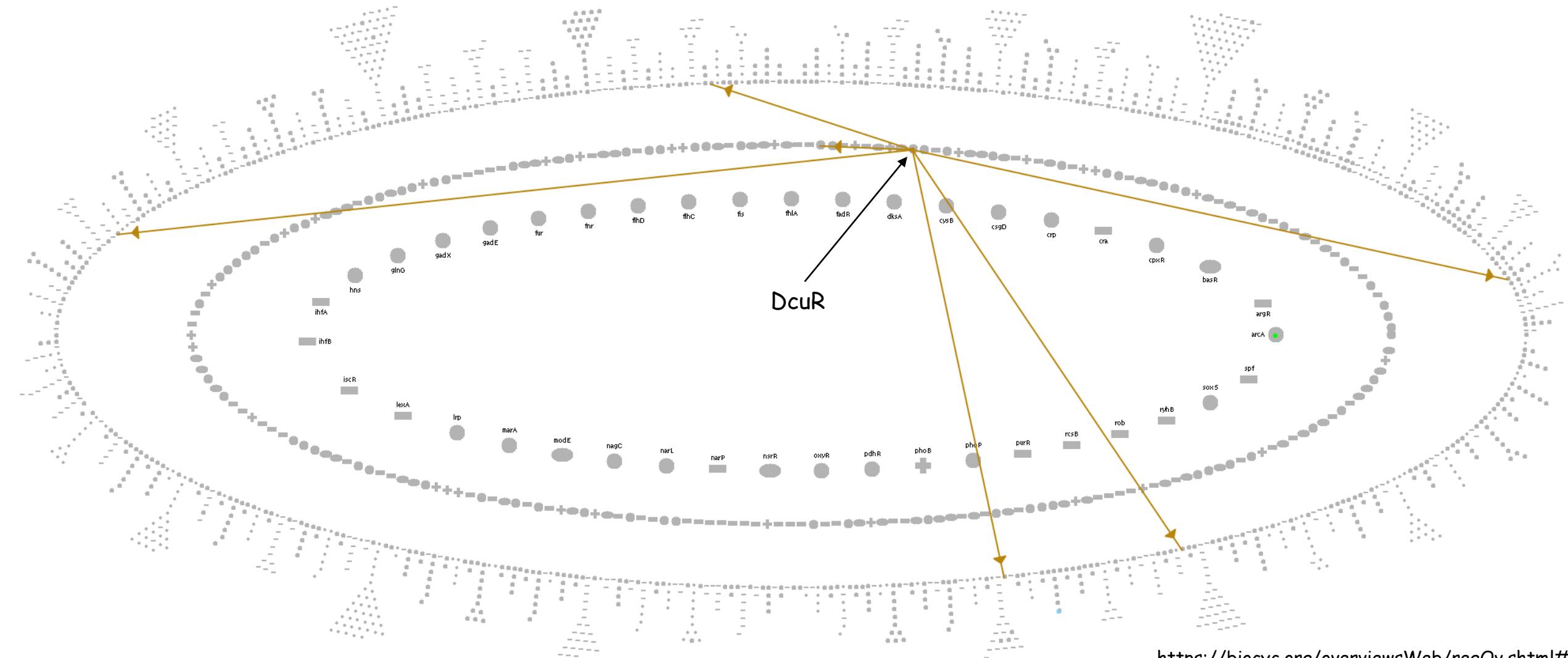
# Two-Component Regulatory System



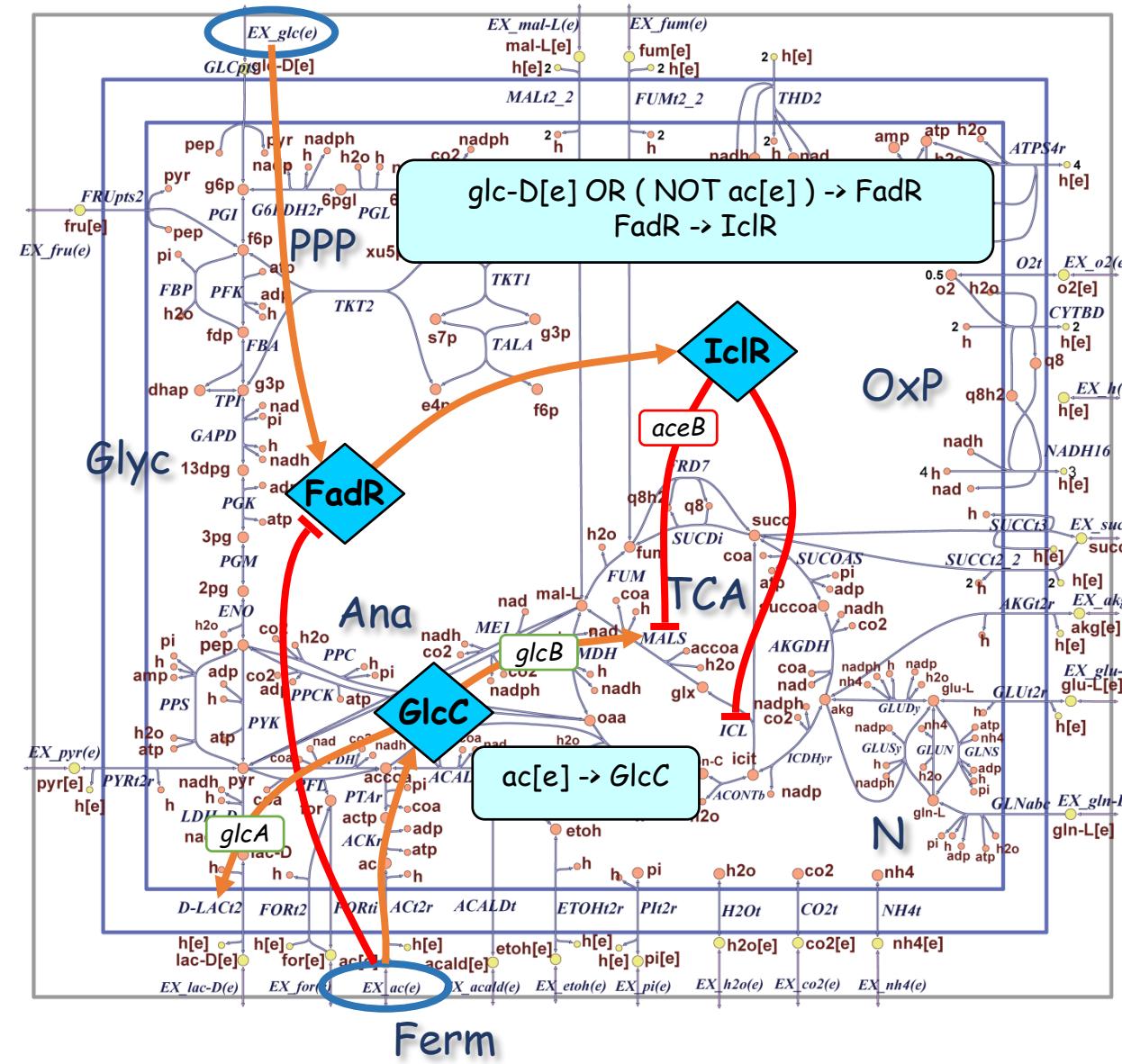
<https://image.slidesharecdn.com/bio305regulation2012-120105105328-phpapp02/95/bio305-lecture-on-gene-regulation-in-bacterial-pathogens-17-728.jpg?cb=1325761450>



# DcuR Impact in *E.coli* K-12 (EcoCyc)



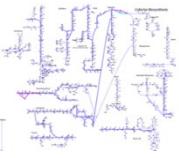
<https://biocyc.org/overviewsWeb/regOv.shtml#>



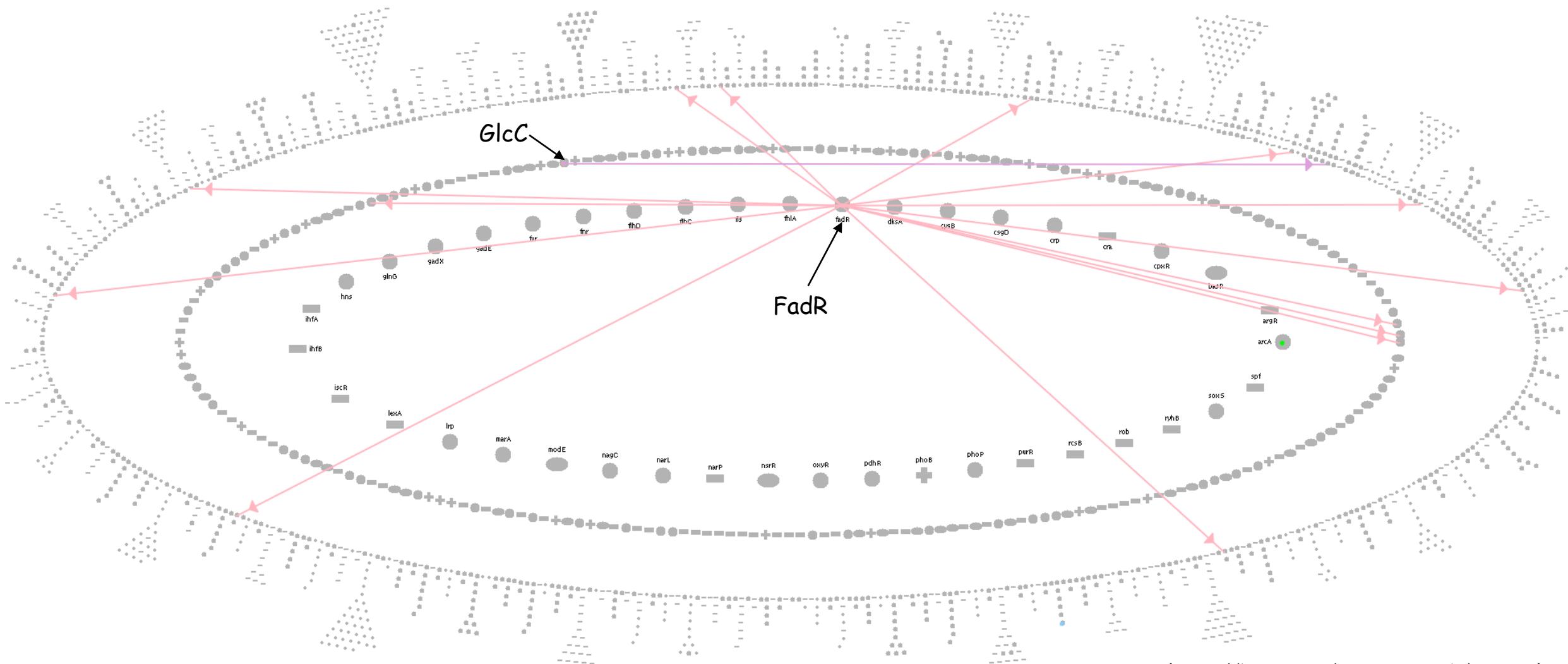
# FadR, IclR & GlcC

- FadR, IclR & GlcC are transcriptional regulators
  - FadR and IclR are activated when either glucose is present in the media or acetate is not.
  - FadR and IclR form a two component histidine kinase system.
  - Down regulates the glyoxylate cycle.
  - GlcC is activated when acetate is present in the media
  - Upregulates the transport pathway for D-lactate

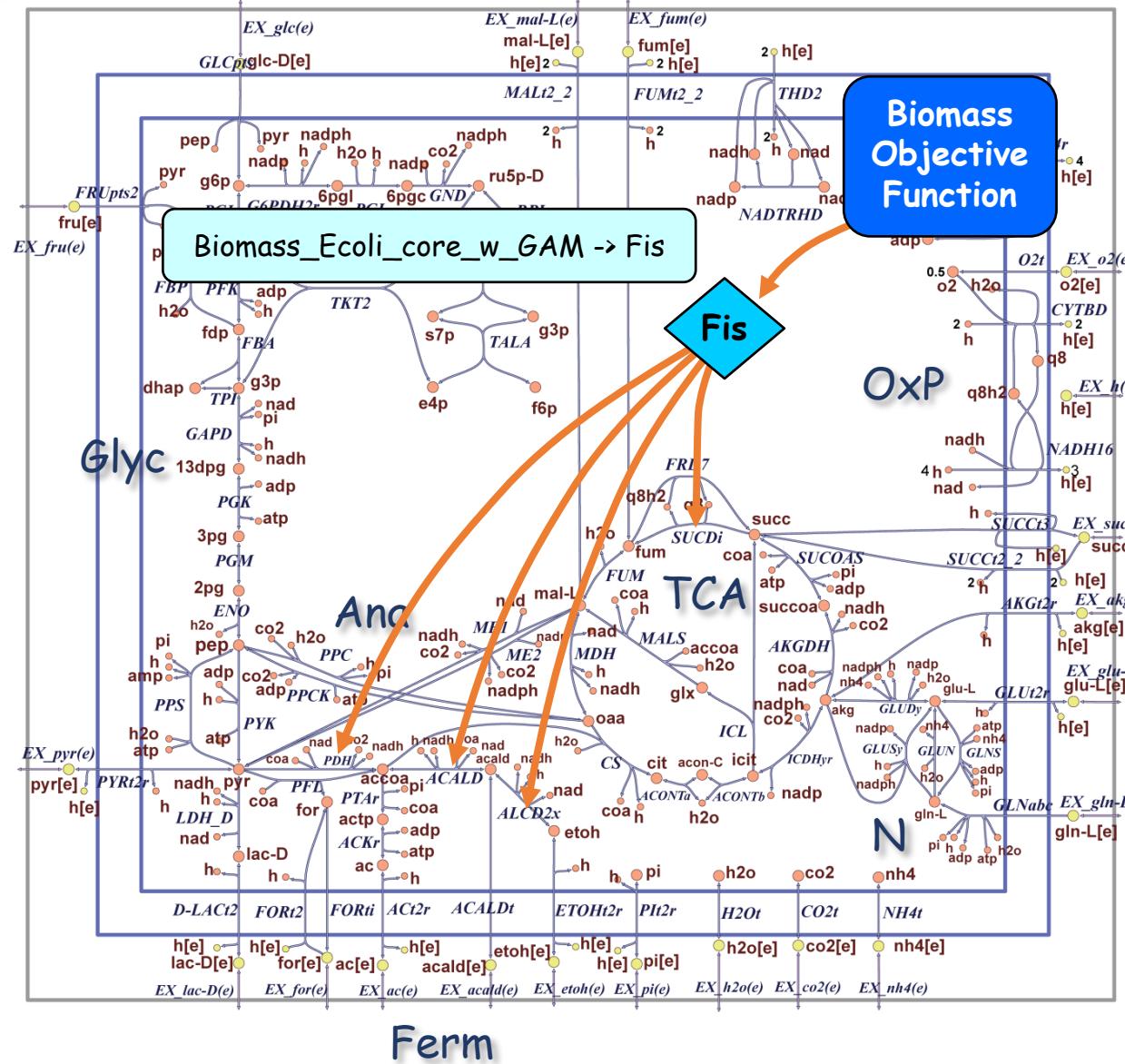
Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# FadR and GlcC Impact in *E.coli* K-12 (EcoCyc)



<https://biocyc.org/overviewsWeb/regOv.shtml#>



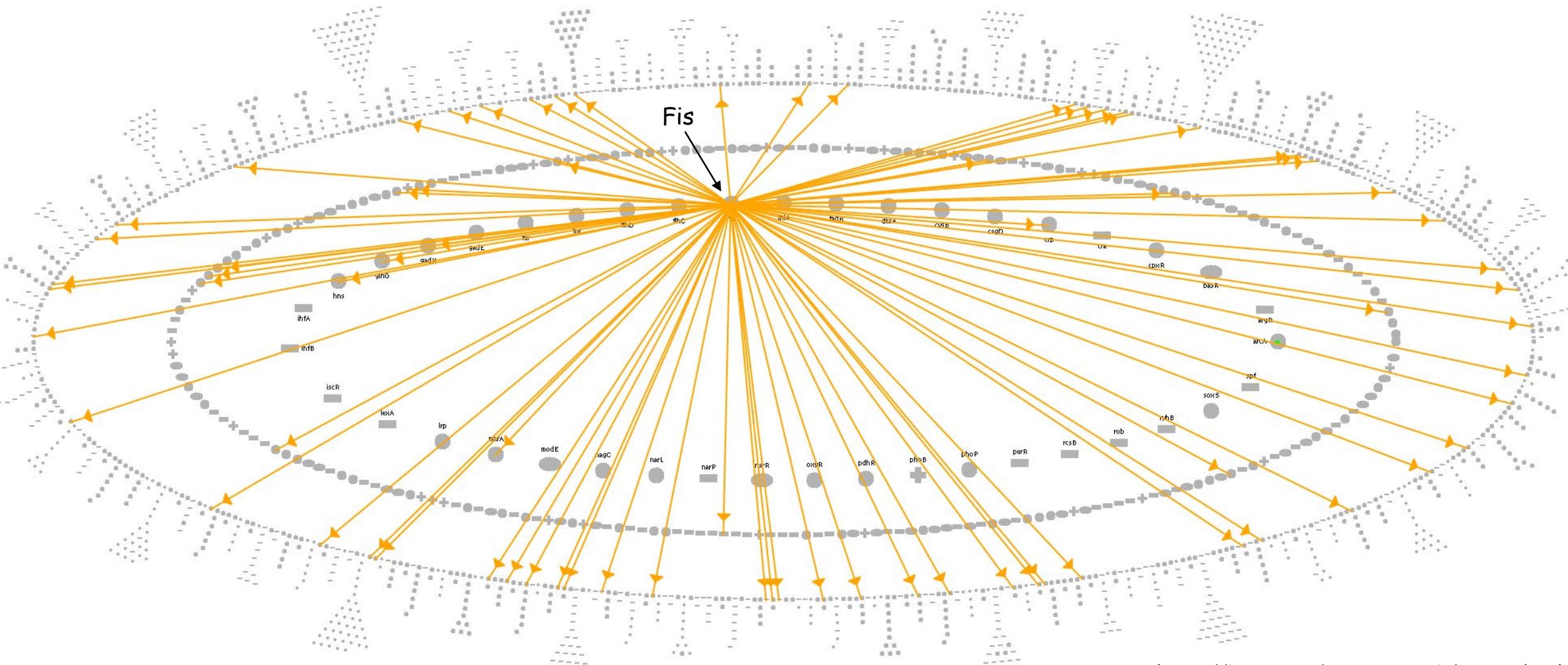
Fis

- Fis is a transcriptional regulator
  - Directly or indirectly regulates 21% of the genes in *E.coli*
  - Fis is activated when the cell is in exponential growth phase.
  - Up regulates the ethanol pathway
  - Up regulates SUCDi in the TCA cycle
  - When modeling the balanced steady state growth typical of the exponential growth phase, the state of Fis is always set to be true.

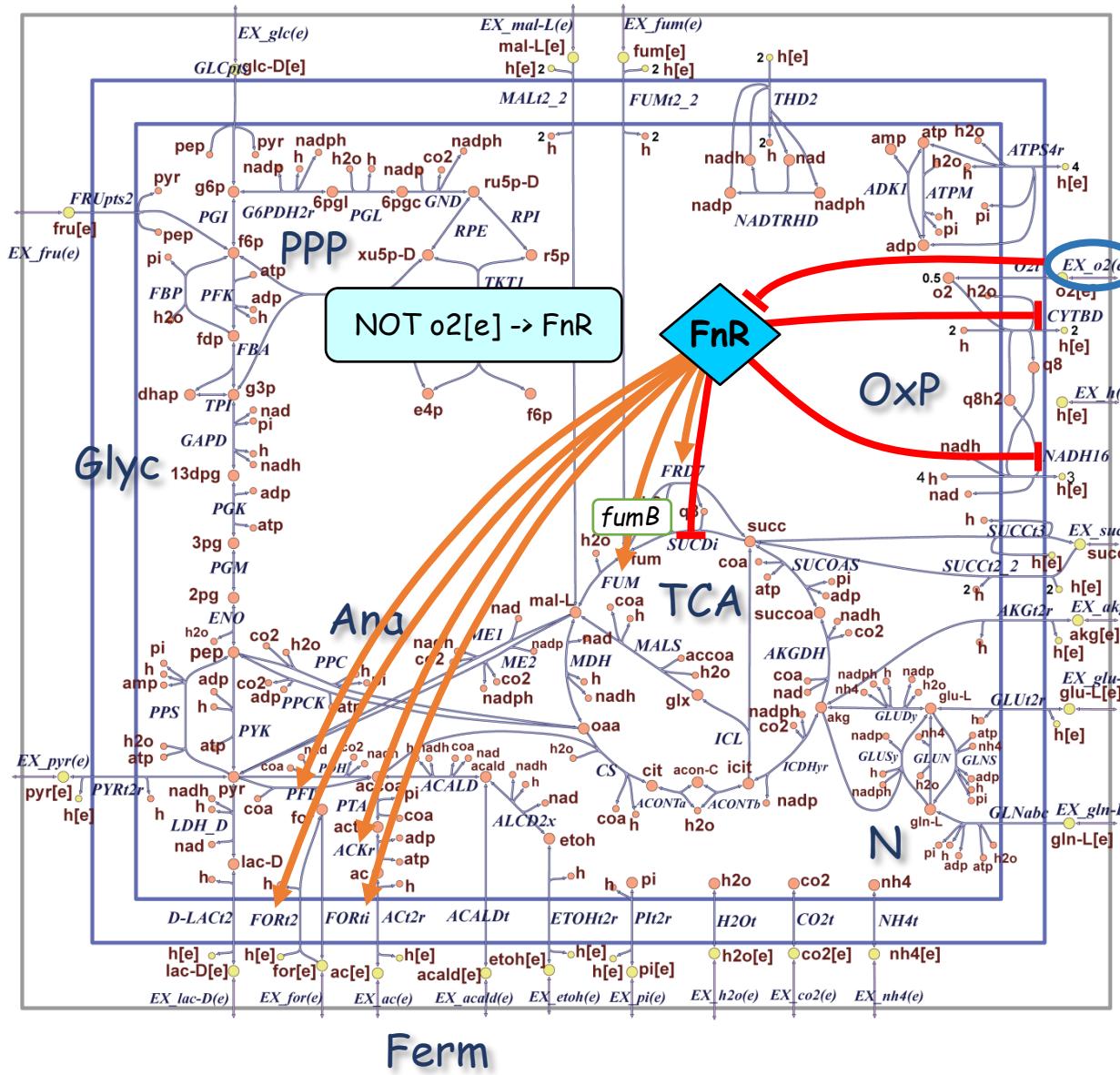
## Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Fis Impact in E.coli K-12 (EcoCyc)



<https://biocyc.org/overviewsWeb/regOv.shtml#>



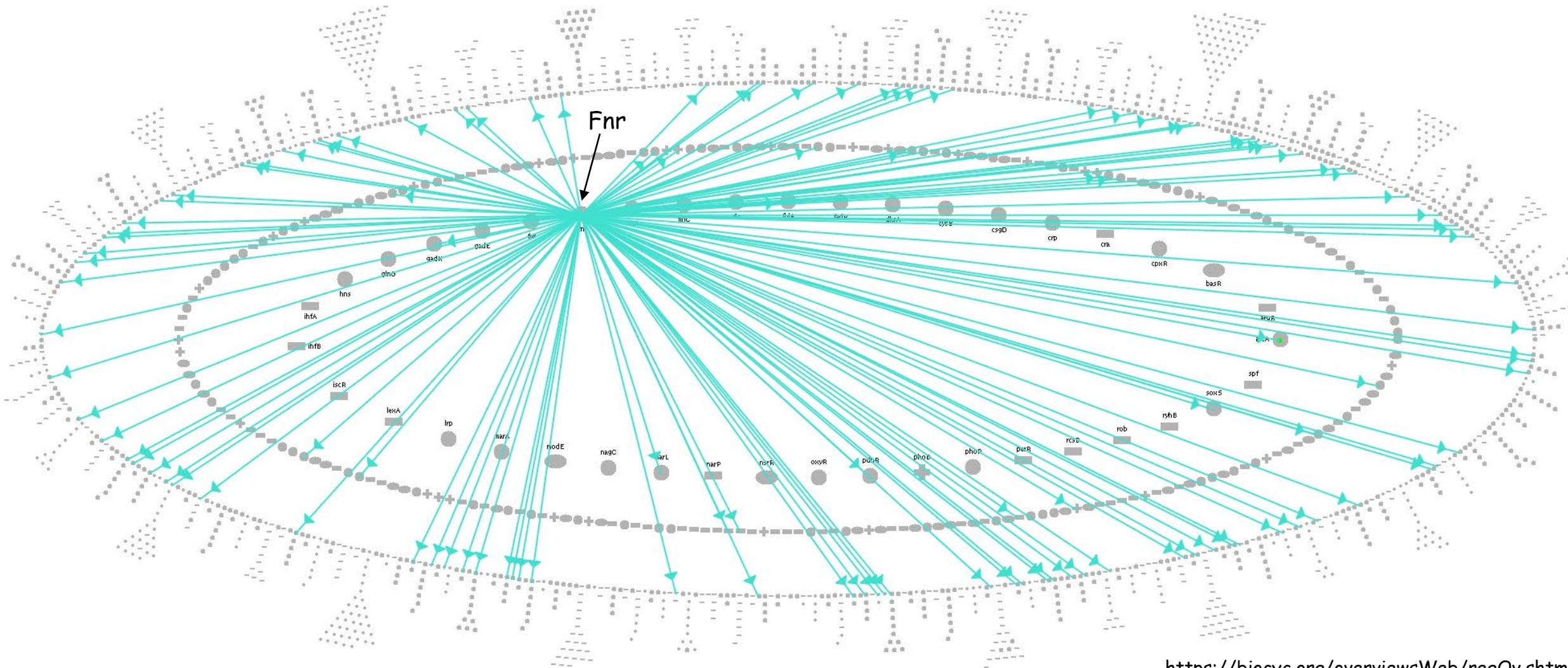
## FnR

- FnR is a transcriptional regulator
- FnR is activated in anaerobic conditions
- Downregulates the energy producing portion of the TCA cycle
- Upregulates the reductive pathway in the TCA cycle (fumB)
- Upregulates the fermentation pathway for formate and acetate
- Downregulates oxidative phosphorylation

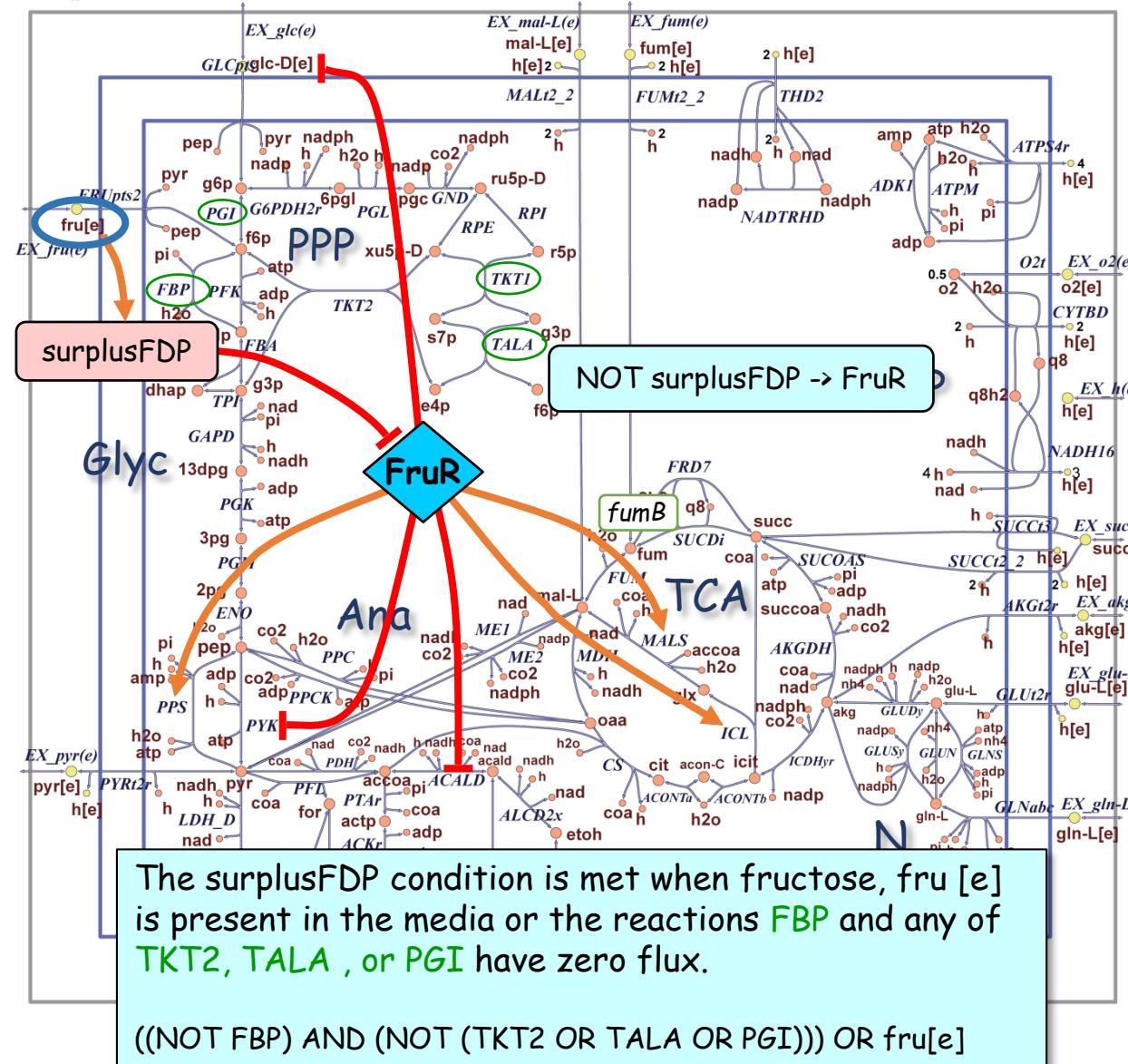
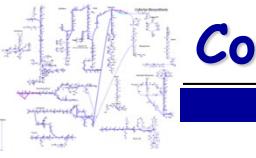
Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Fnr Impact in E.coli K-12 (EcoCyc)



<https://biocyc.org/overviewsWeb/regOv.shtml#>



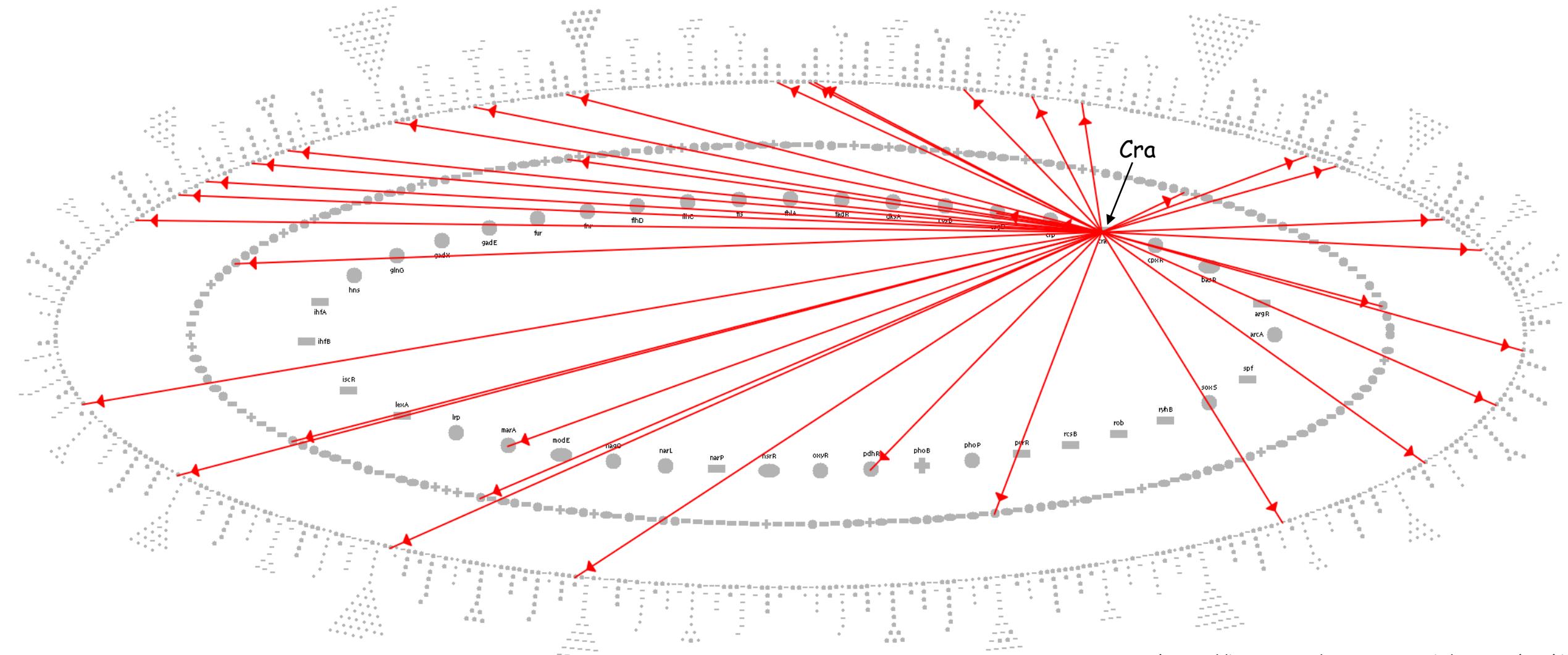
## FruR (Cra) & surplusFDP

- FruR is a transcriptional regulator
- FruR has been replaced by cra ("Catabolite repressor activator")
- FruR is activated by low fructose levels ( $\text{surplusFDP} = \text{false}$ )
- FruR reverses the flow of carbon to replenish glycolytic intermediates
- Upregulates the glyoxylate cycle
- Upregulates the Gluconeogenesis pathway
- Downregulates ethanol and acetaldehyde fermentation pathway
- Downregulates the uptake of glucose

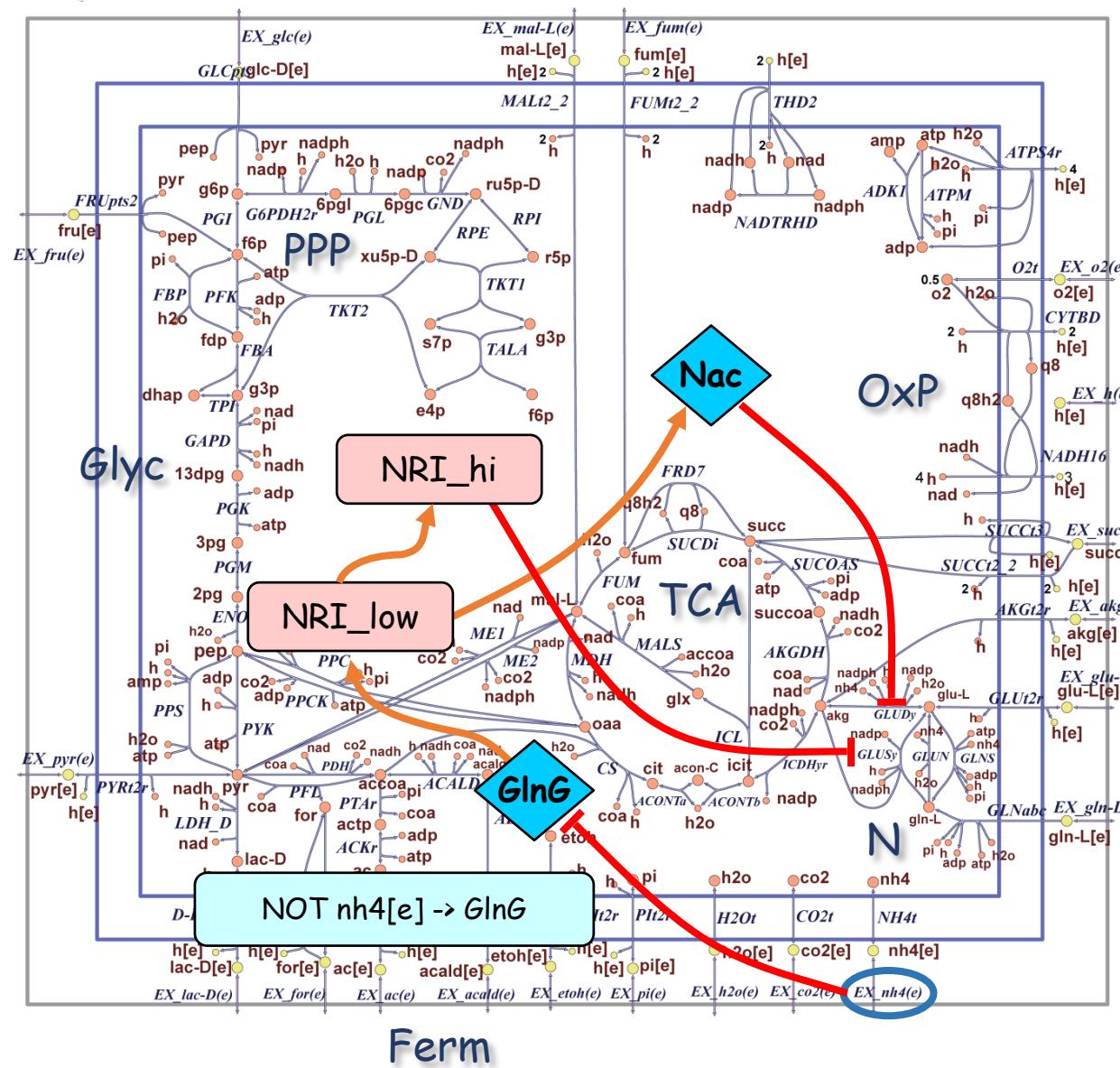
Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Fru (Cra) Impact in E.coli K-12 (EcoCyc)



<https://biocyc.org/overviewsWeb/regOv.shtml#>



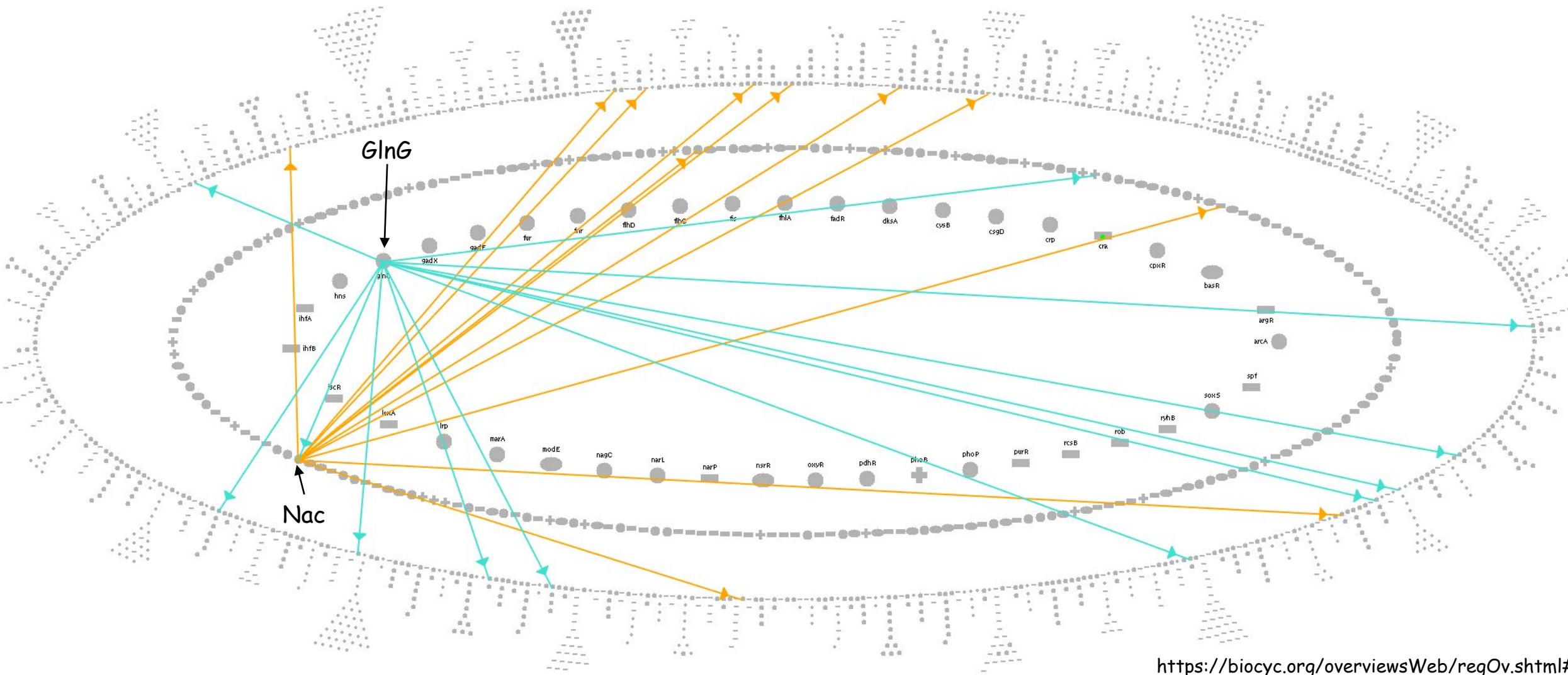
## GlnG, Nac & NRI

- GlnG, Nac & NRI are transcriptional regulators
- GlnG is activated by a low extracellular ammonium ( $\text{nh4[e]}$ ) concentration and then activates the low-level (fast) nitrogen response, NRI\_low.
- NRI\_low activates Nac which down regulates the production of L-glutamate from 2-Oxoglutarate.
- NRI\_low also activates NRI\_hi (high-level, slow response) which down regulates the production of L-glutamate from L-glutamine and 2-Oxoglutarate.
- The total response is to reduce the production of L-glutamate.

Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# GlnG and Nac Impact in E.coli K-12 (EcoCyc)



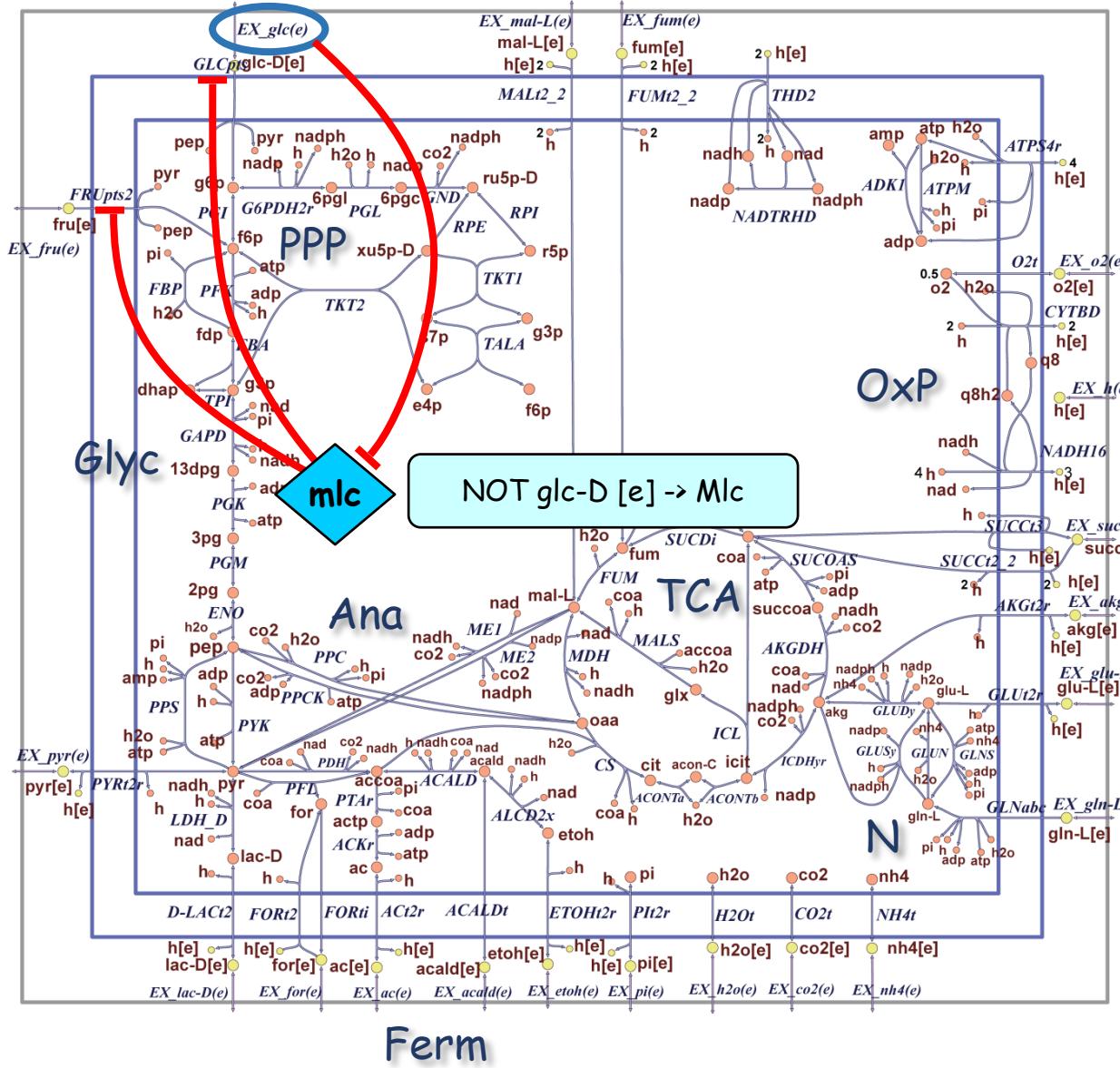
<https://biocyc.org/overviewsWeb/regOv.shtml#>



# Constraint-based Metabolic Reconstructions & Analysis

2021 H. Scott Hinton

56



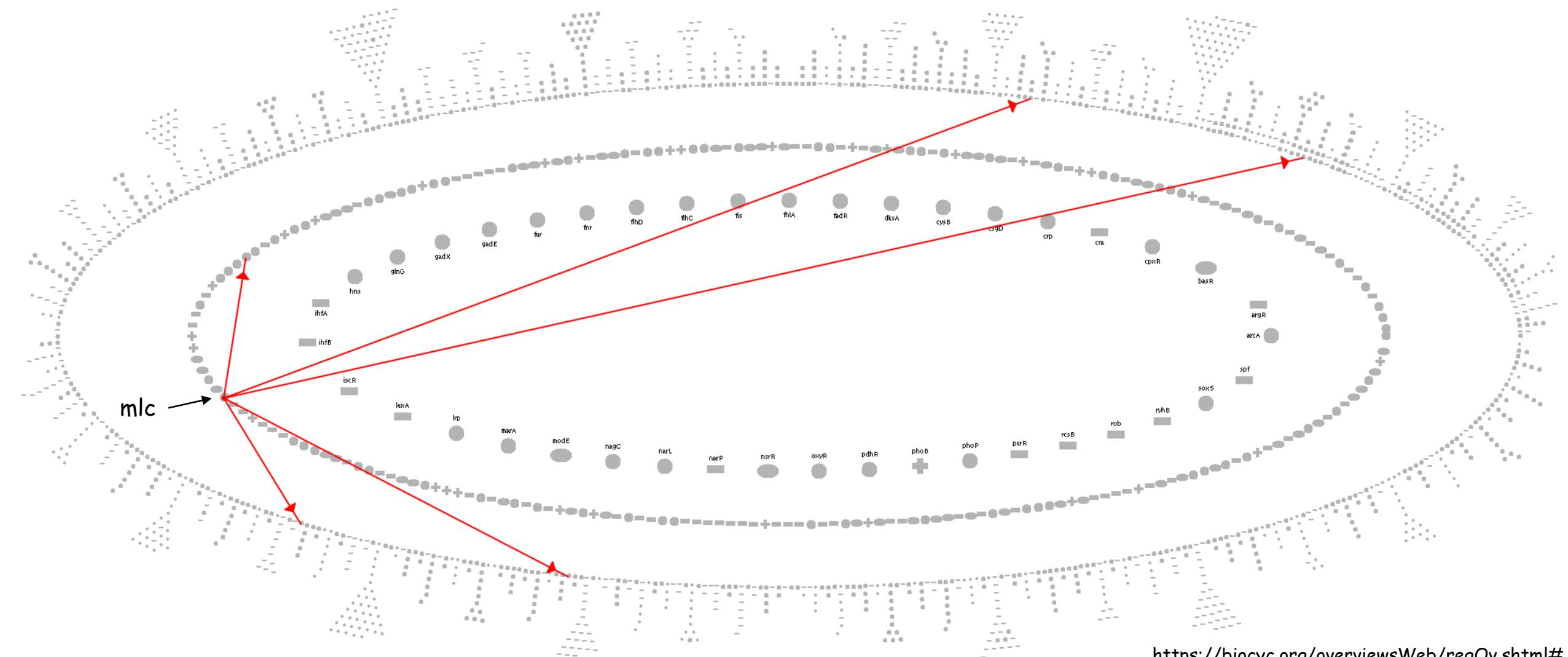
m<sub>lc</sub>

- *m<sub>lc</sub>* is a transcriptional regulator
- *m<sub>lc</sub>* is activated when no glucose is present in the media.
- When no glucose is present in the media, the transporters for both glucose and fructose are down regulated.

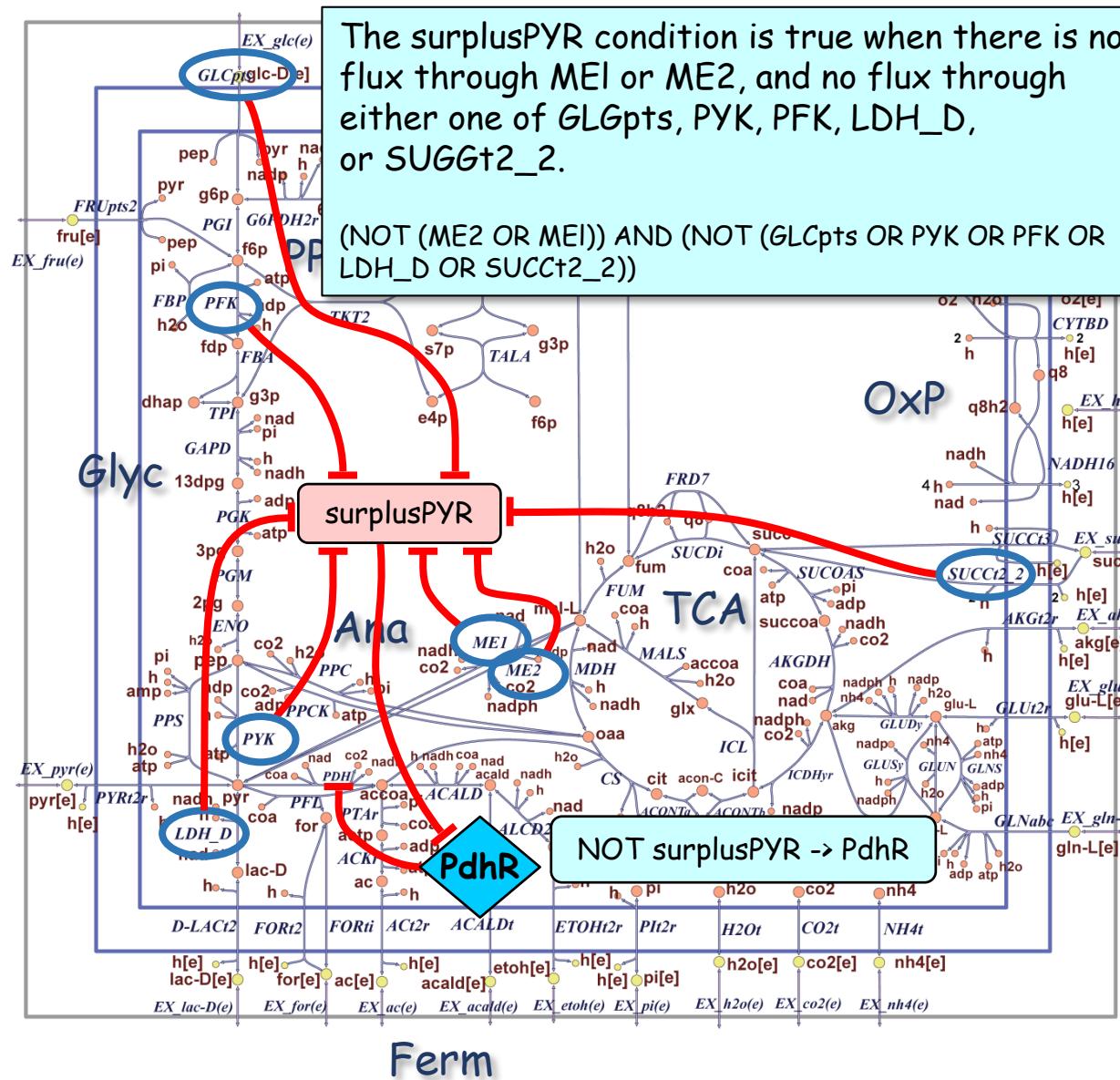
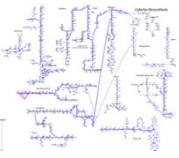
Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# mlc Impact in E.coli K-12 (EcoCyc)



<https://biocyc.org/overviewsWeb/regOv.shtml#>



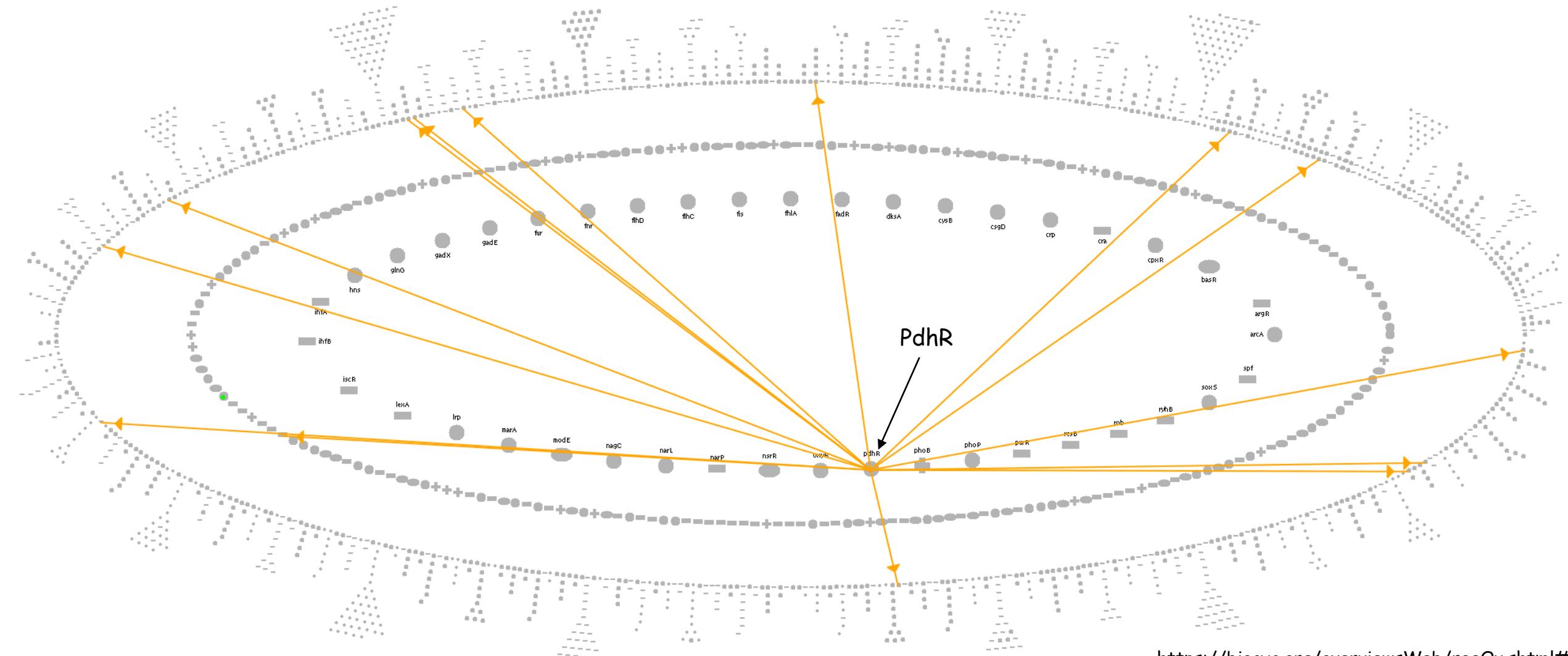
## PdhR & surplusPYR

- PdhR is a transcriptional regulator
- The dual transcriptional regulator PdhR down regulates pyruvate dehydrogenase, PDH, when the pyruvate concentration in the cell is low.
- High pyruvate concentration is represented by the variable surplusPYR, which is true when there is no flux through ME1 or ME2, and no flux through either one of GLGpts, PYK, PFK, LDH\_D, or SUGGT2\_2.

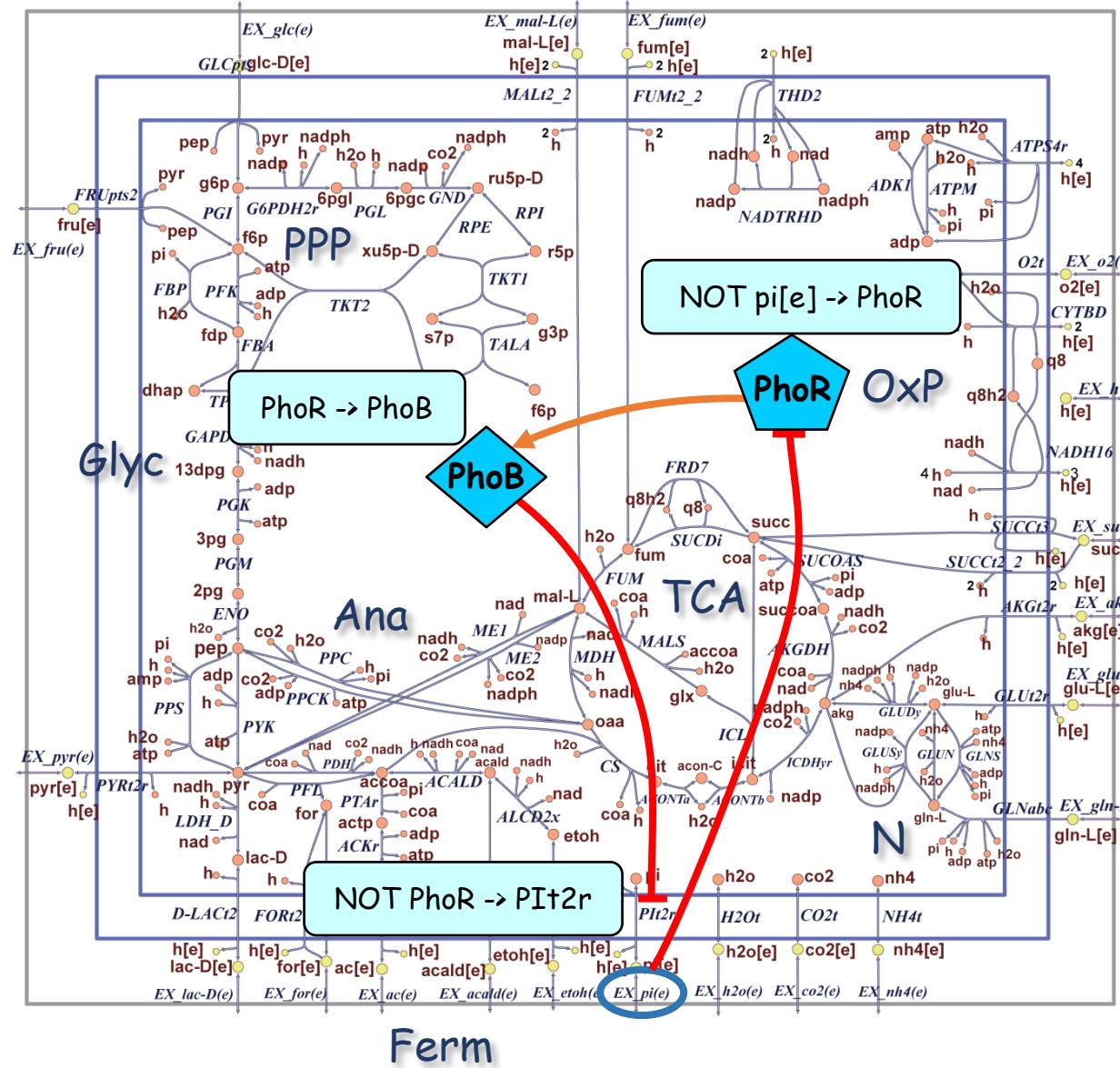
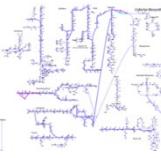
Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# PdhR Impact in E.coli K-12 (EcoCyc)



<https://biocyc.org/overviewsWeb/regOv.shtml#>



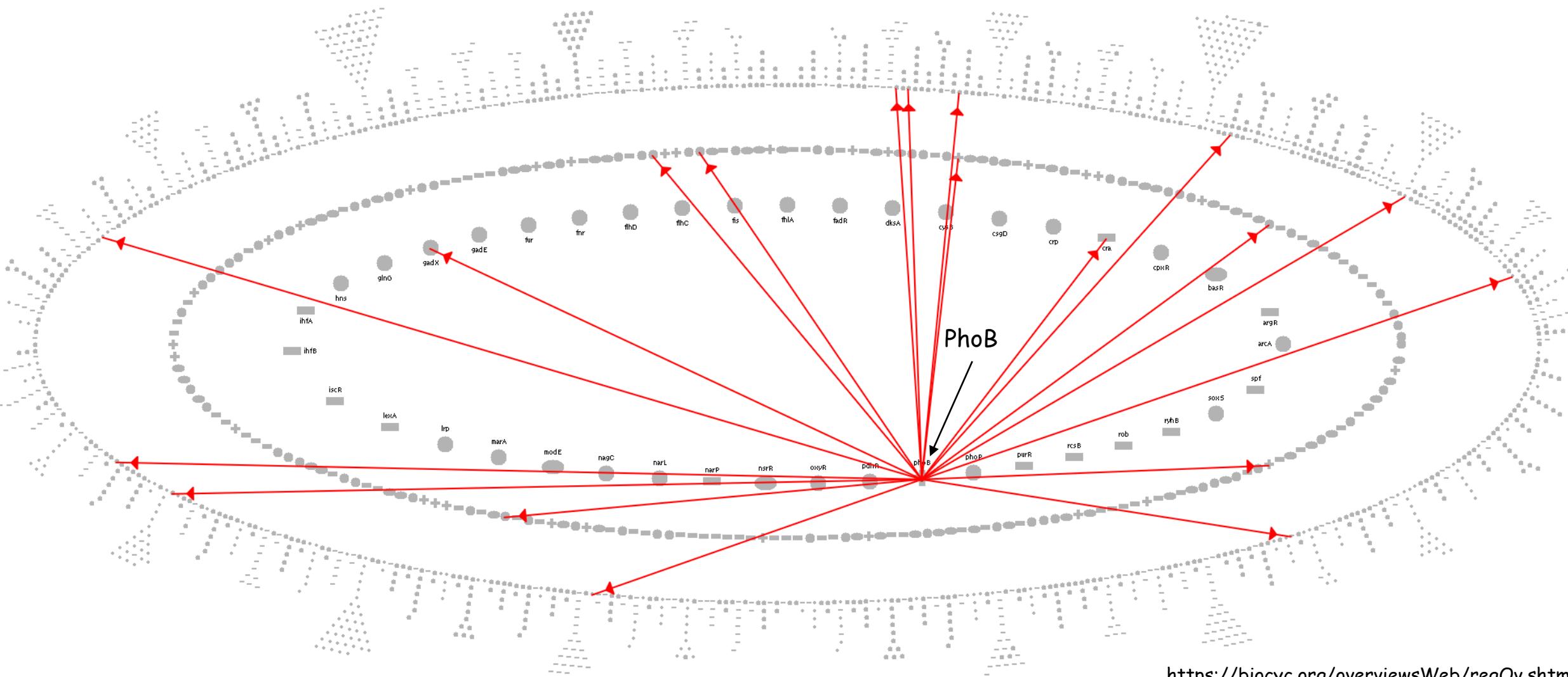
# PhoB & PhoR

- PhoB is a transcriptional regulator
- PhoR is a sensory histidine kinase
- Phosphorus uptake is regulated by the two-component system phoR/ phoB.
- phoR codes for a sensor kinase that is phosphorylated when extracellular inorganic phosphate is not present.
- The phosphorylated enzyme is activated, and it phosphorylates the transcriptional regulator PhoB.
- Phosphorylated PhoB then represses the phosphate transporter, PIt2r.
- The overall effect of phosphorus regulation is to down regulate the phosphate transport reaction, PIt2r, when no extracellular inorganic phosphate is present.

Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# PhoB Impact in E.coli K-12 (EcoCyc)



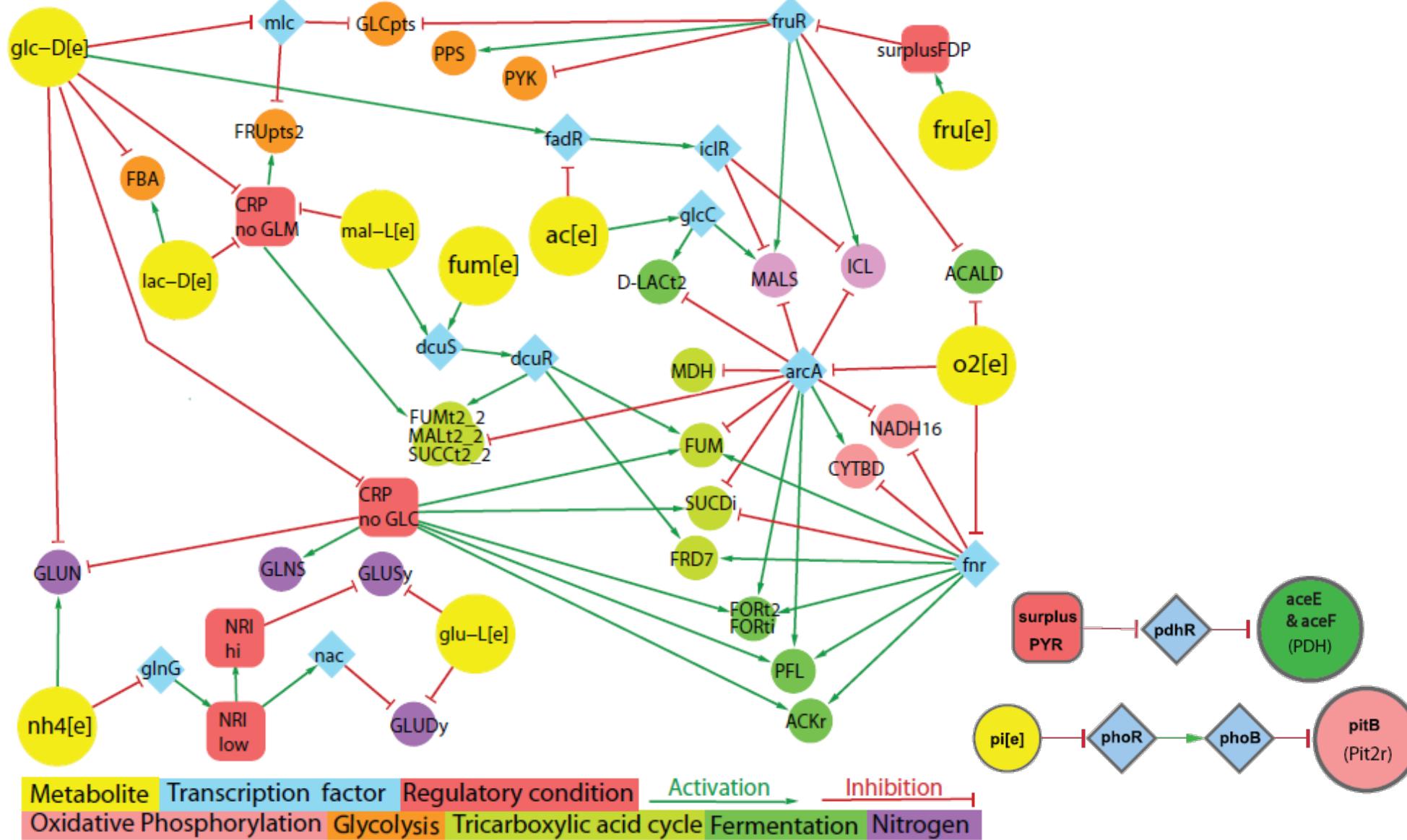
<https://biocyc.org/overviewsWeb/regOv.shtml#>



# Constraint-based Metabolic Reconstructions & Analysis

2021 H. Scott Hinton

62





# Lesson Outline

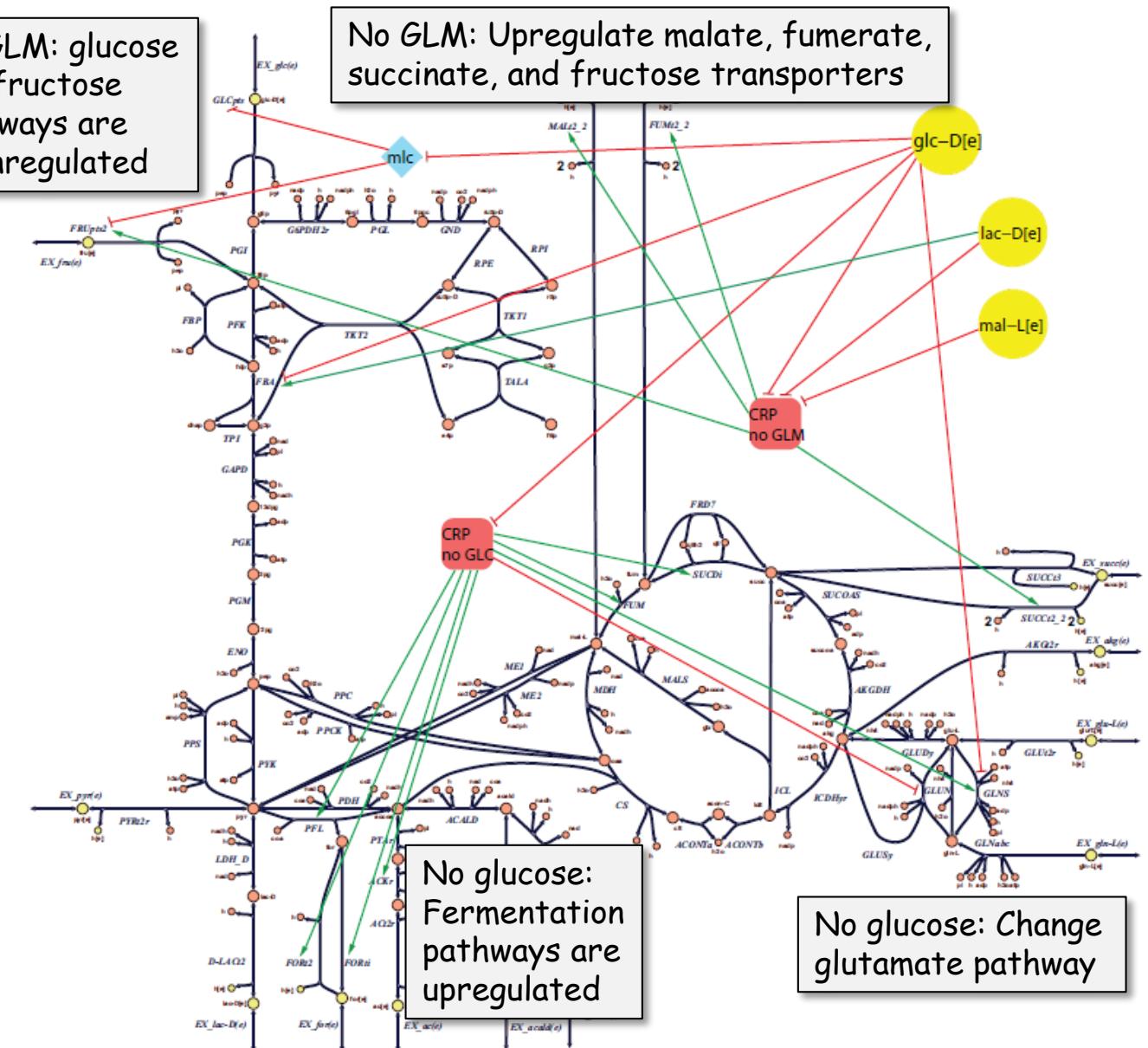
- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



# Catabolite Repression

- In media containing glucose and other sugar substrates (lactate or malate), *E. coli* preferentially catabolizes glucose until it is depleted, and then switches to the other substrates.
- The activity of the cAMP receptor protein, Crp is modeled using CRPnoGLM and CRPnoGLC.
- The CRPnoGLM regulatory condition is true when either glucose (glc-D[e]) , malate (mal-L[e]) or lactate (lac-D[e]) are not present.
- The CRPnoGLC regulatory condition is true when glucose (glc-D[e]) is absent.
- The transcription factor, mlc, is also activated when glucose is not present.

No GLM: glucose and fructose pathways are downregulated





# Aerobic Glucose & Fructose

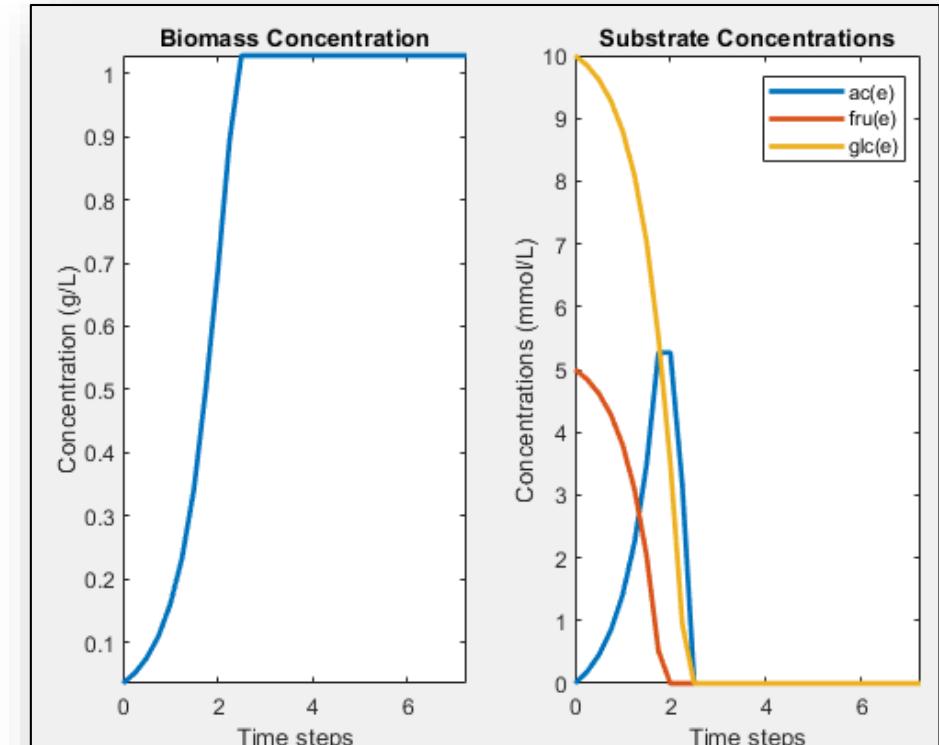
```
% GlucoseFructose_Aerobic.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -10, '1');
model = changeRxnBounds(model, 'EX_fru(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -30, '1');

substrateRxns = {'EX_fru(e)', 'EX_glc(e)'};
initConcentrations = [5 10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoh(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', 'EX_gln_L(e)', ...
    'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Low-Aerobic Glucose & Fructose

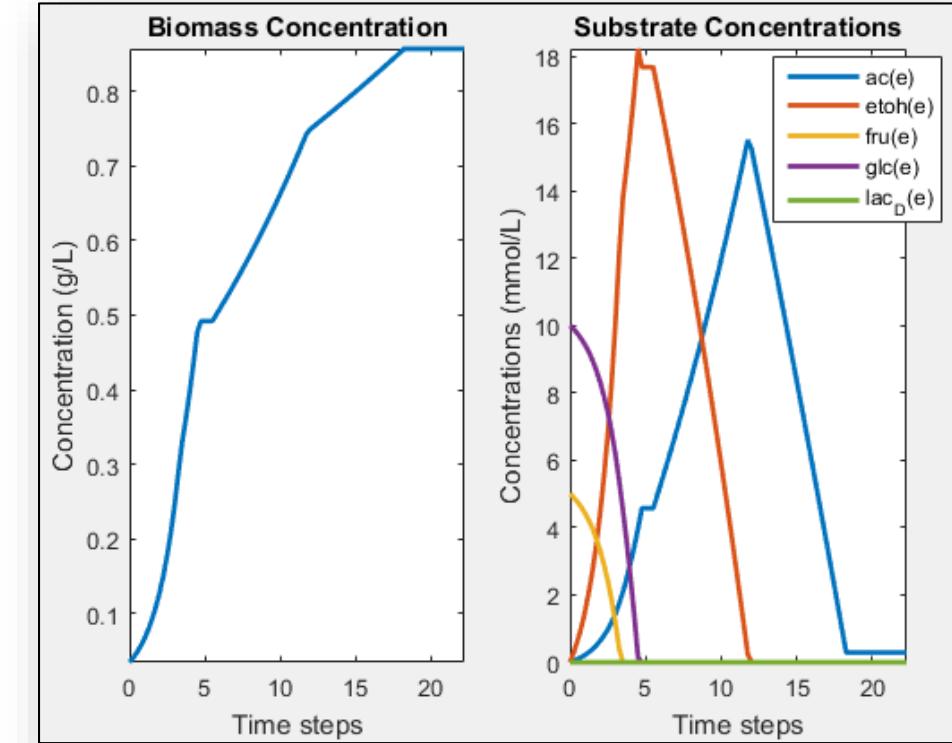
```
% GlucoseFructose_Aerobic.m
clear;

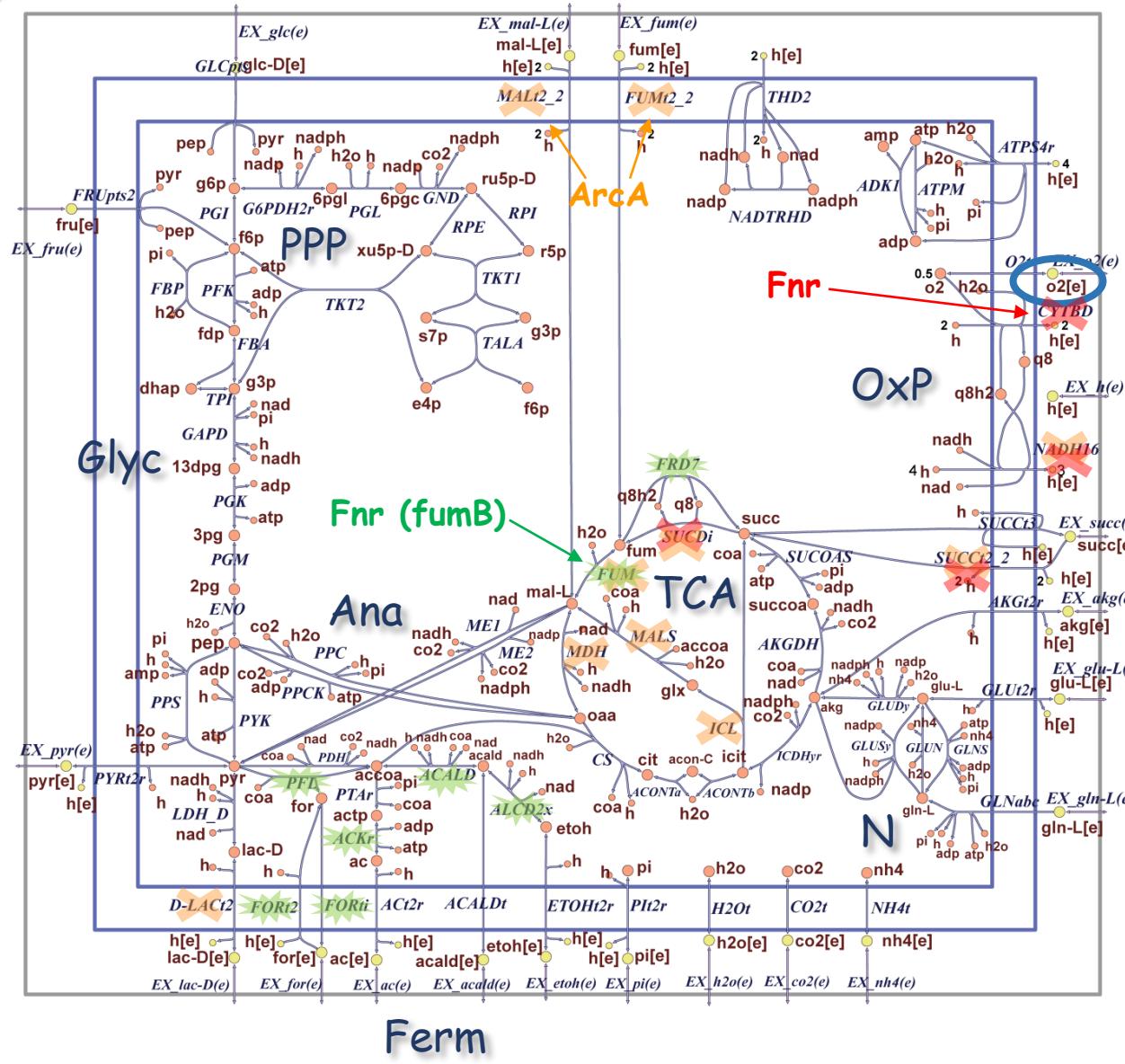
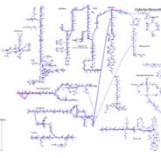
model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_fru(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_o2(e)', -5, 'l');

substrateRxns = {'EX_fru(e)', 'EX_glc(e)'};
initConcentrations = [5 10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etooh(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', 'EX_gln_L(e)', ...
    'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





## Anoxic Growth "Low Oxygen"

- Upregulates the fermentation pathway for formate, acetate, and ethanol
- Upregulates the reductive pathway in the TCA cycle
- Downregulates the transporters for malate, fumarase, lactate, and succinate.
- Downregulates the glyoxylate cycle
- Downregulates the energy producing portion of the TCA cycle
- Downregulates oxidative phosphorylation

Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Aerobic Glucose & Fumerate

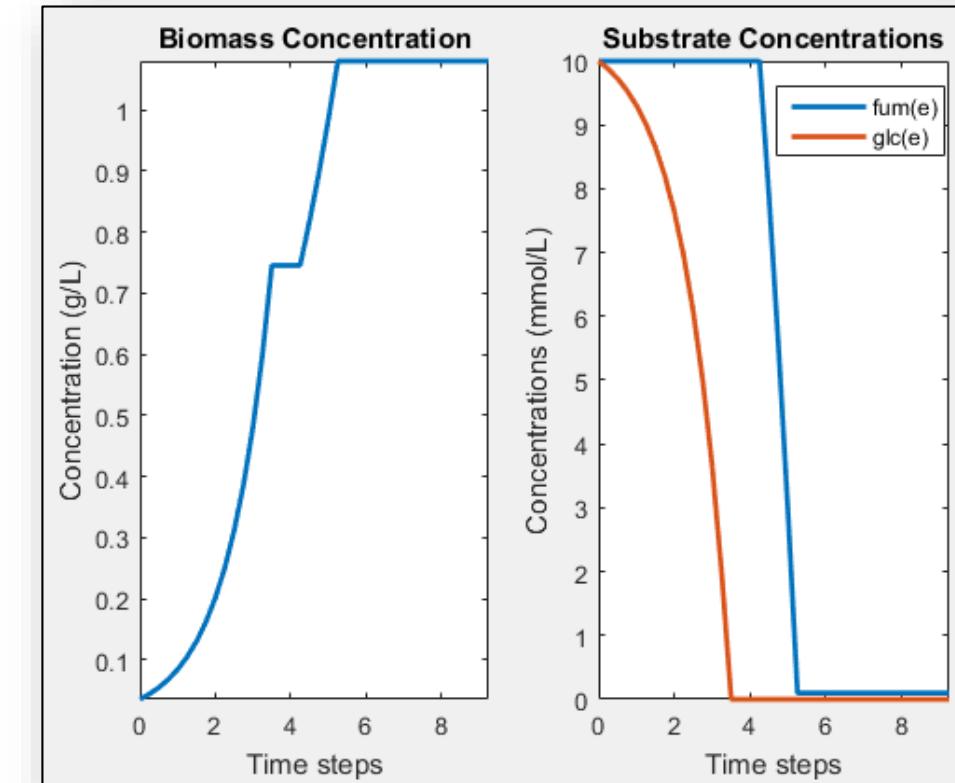
```
% GlucoseFumarateCatabolite_Repression.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_fum(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_o2(e)', -30, 'l');

substrateRxns = {'EX_glc(e)', 'EX_fum(e)'};
initConcentrations = [10 10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoh(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_fum(e)', ...
    'EX_glc(e)', 'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Low-Aerobic Glucose & Fumerate

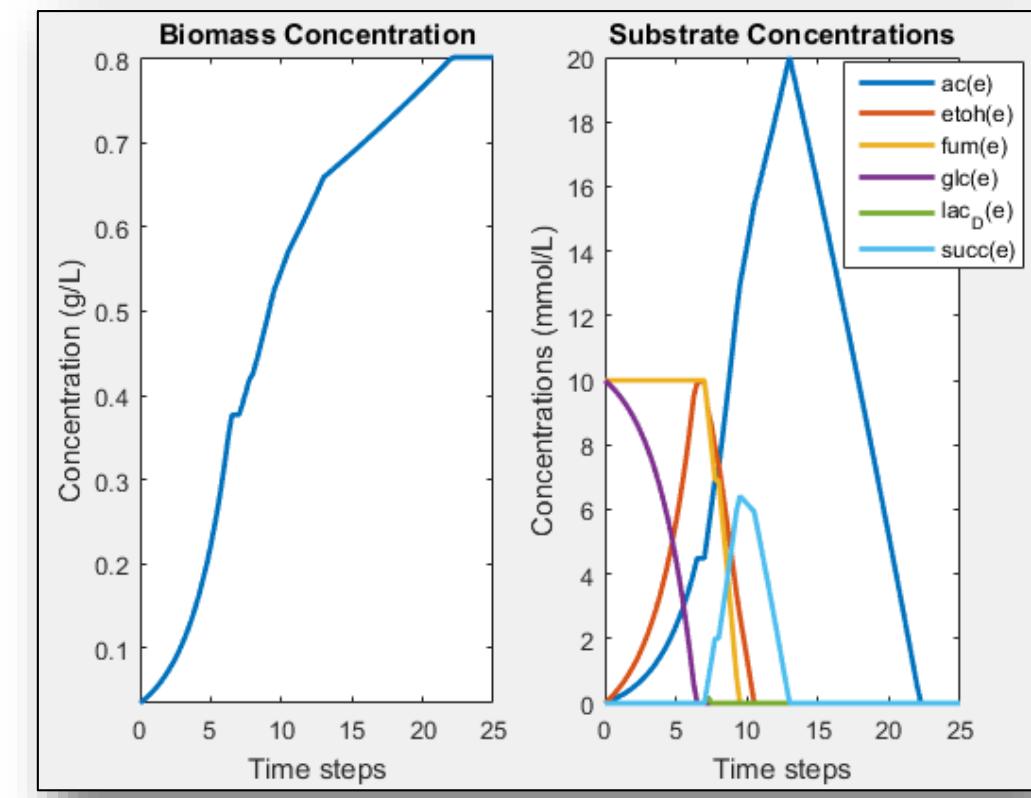
```
% GlucoseFumerateCatabolite_Repression.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -10, '1');
model = changeRxnBounds(model, 'EX_fum(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -5, '1');

substrateRxns = {'EX_glc(e)', 'EX_fum(e)'};
initConcentrations = [10 10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoH(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_fum(e)', ...
    'EX_glc(e)', 'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Aerobic Glucose, Fructose & Lactate

```
% GlucoseFructoseLactate_Aerobic.m
clear;

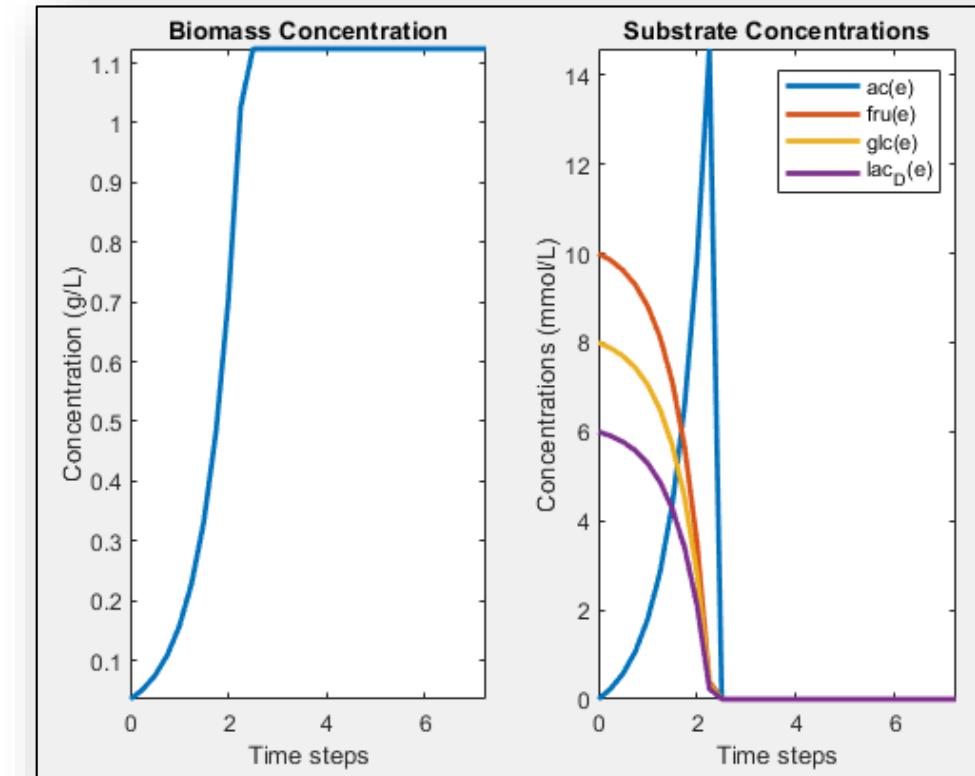
model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_fru(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_glc(e)', -8, 'l');
model = changeRxnBounds(model, 'EX_lac_D(e)', -6, 'l');

modelReg = changeRxnBounds(model, 'EX_o2(e)', -30, 'l');

substrateRxns = {'EX_fru(e)', 'EX_glc(e)', 'EX_lac_D(e)'};
initConcentrations = [10 8 6];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etho(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', 'EX_gln_L(e)', ...
    'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Low-Aerobic Glucose, Fructose & Lactate

```
% GlucoseFructoseLactate_Aerobic.m
clear;

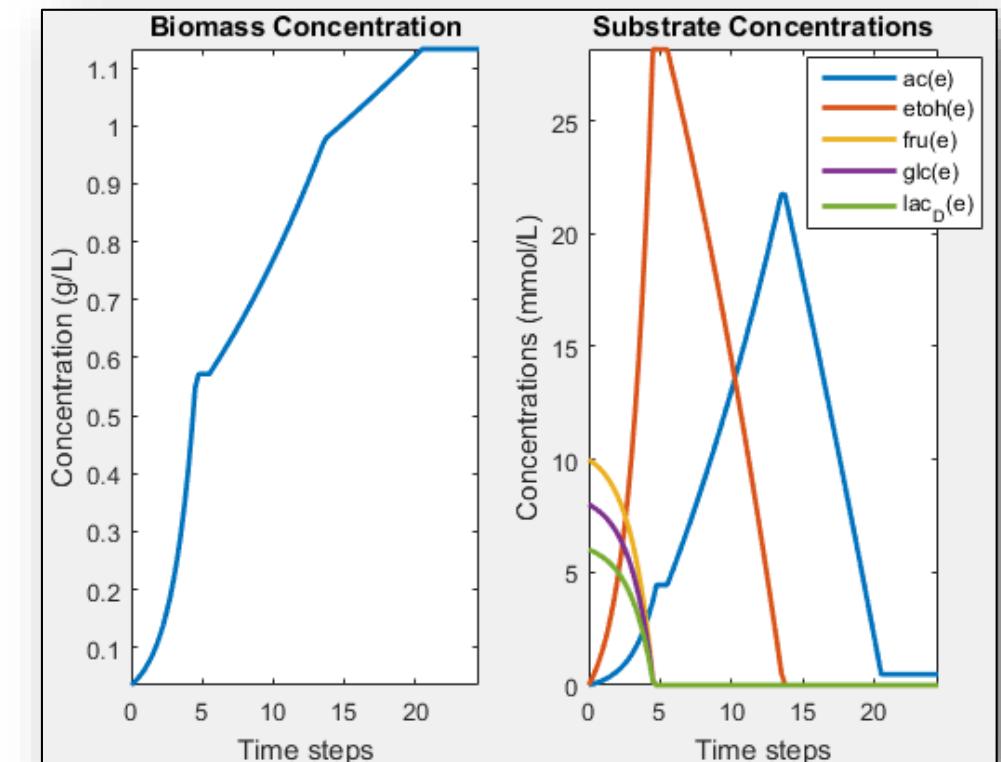
model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_fru(e)', -10, '1');
model = changeRxnBounds(model, 'EX_glc(e)', -8, '1');
model = changeRxnBounds(model, 'EX_lac_D(e)', -6, '1');

modelReg = changeRxnBounds(model, 'EX_o2(e)', -5, '1');

substrateRxns = {'EX_fru(e)', 'EX_glc(e)', 'EX_lac_D(e)'};
initConcentrations = [10 8 6];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoh(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', 'EX_gln_L(e)', ...
    'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

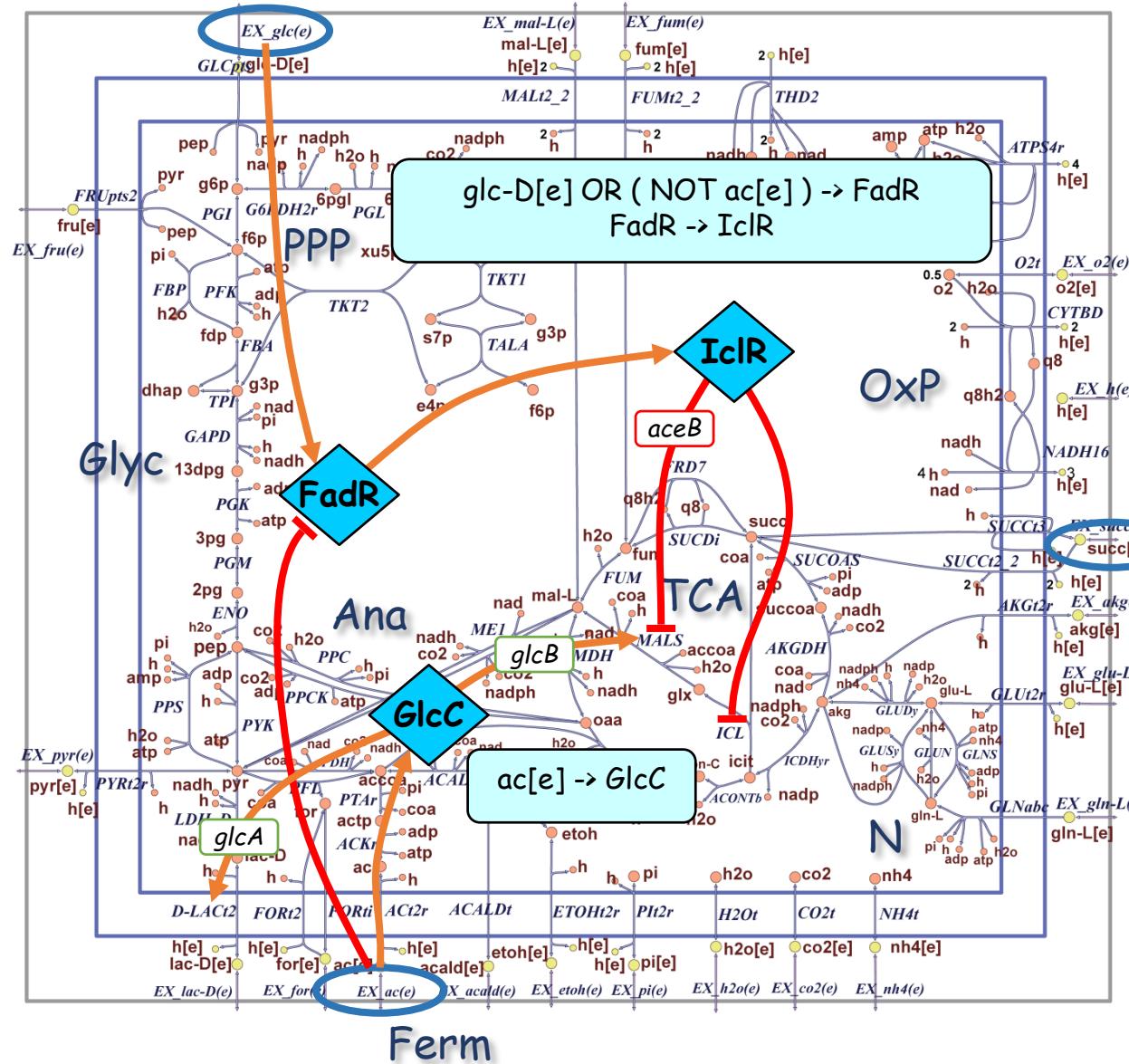
% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



## Growth on Acetate

- FadR and IclR are activated when either glucose is present in the media or acetate is not.
- FadR and IclR form a two component histidine kinase system.
- Down regulates the glyoxylate cycle.
- GlcC is activated when acetate is present in the media
- GlcC Upregulates the transport pathway for D-lactate

Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Aerobic Acetate & Lactate

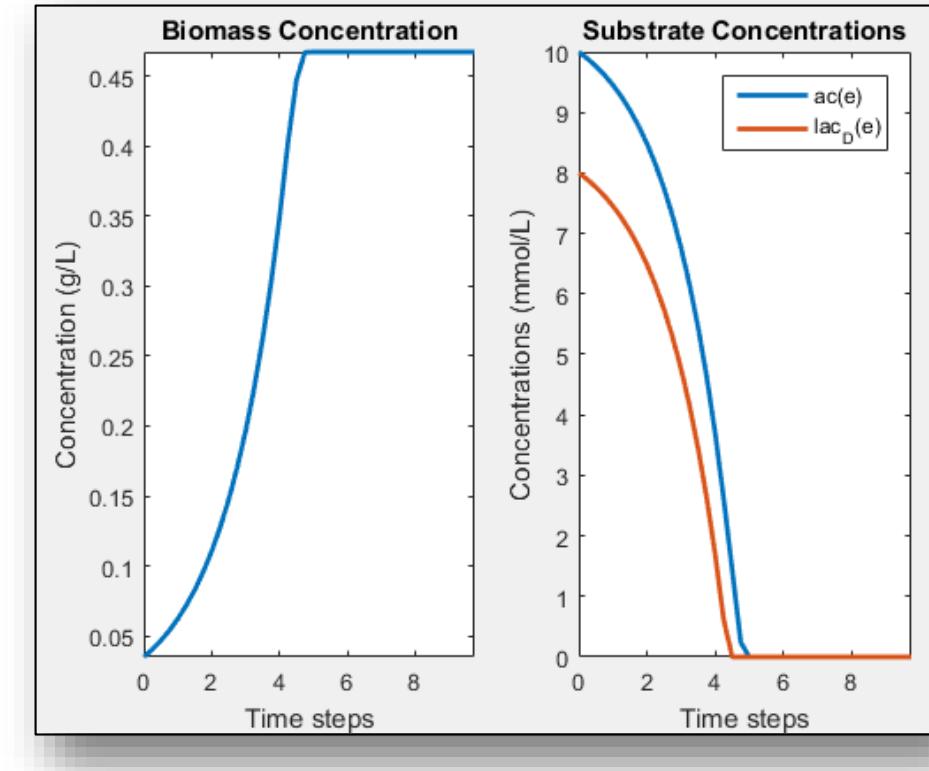
```
% AcetateLactate_Aerobic.m
clear;

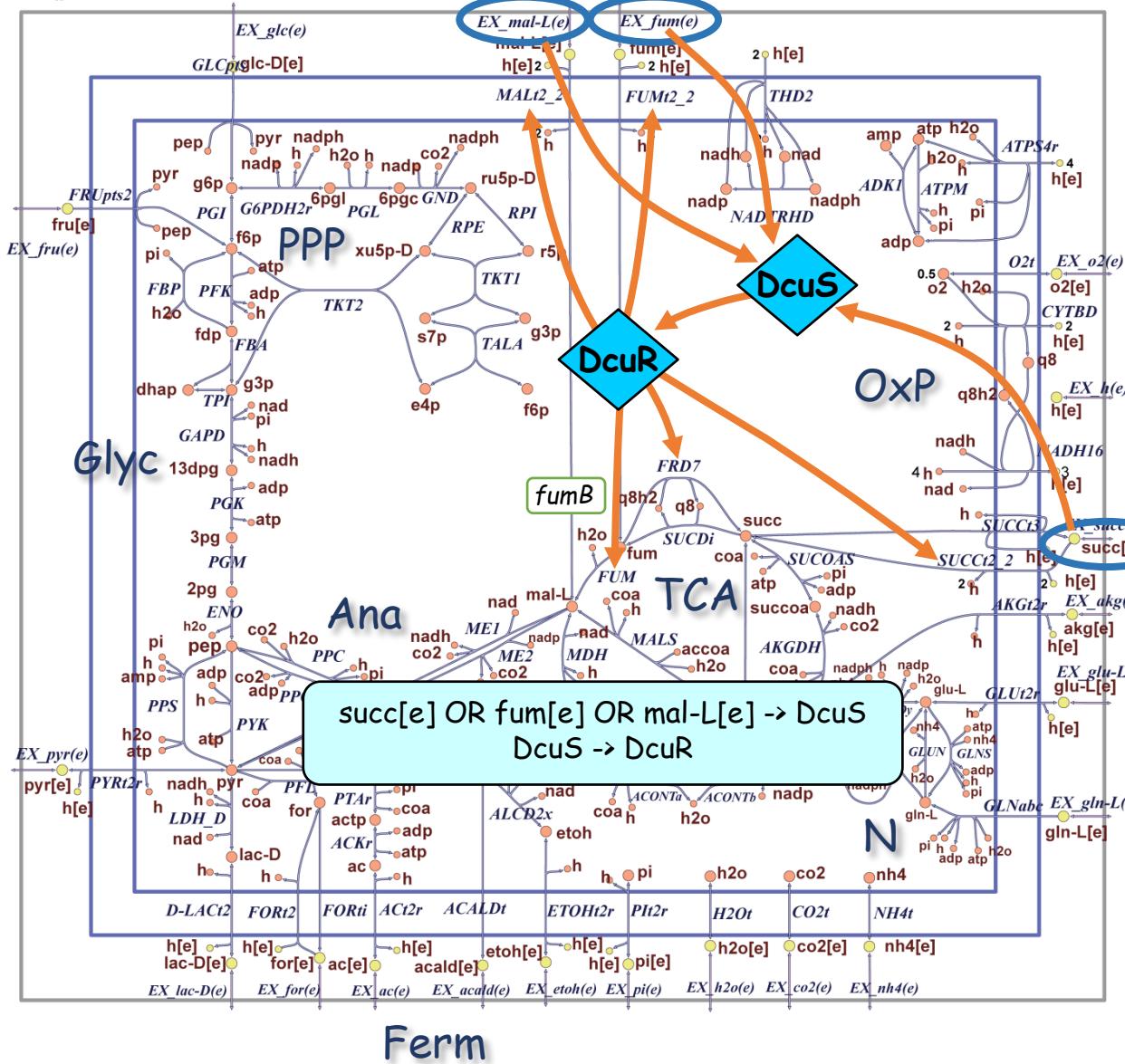
model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -0, '1');
model = changeRxnBounds(model, 'EX_ac(e)', -10, '1');
model = changeRxnBounds(model, 'EX_lac_D(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -30, '1');

substrateRxns = {'EX_ac(e)', 'EX_lac_D(e)'};
initConcentrations = [10 8];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etho(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', 'EX_gln_L(e)', ...
    'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





## Growth on C4-Dicarboxylate Compounds

- Activated when malate, fumarate, or succinate (C4-dicarboxylate compounds) are present in the media.
- DcuS and DcuR form a two component histidine kinase system.
- The presence of malate, fumarate, and succinate upregulates the reductive pathway in the TCA cycle (fumB)
- Upregulates the transport pathways for malate, fumarate, and succinate

Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



# Aerobic Fumerate

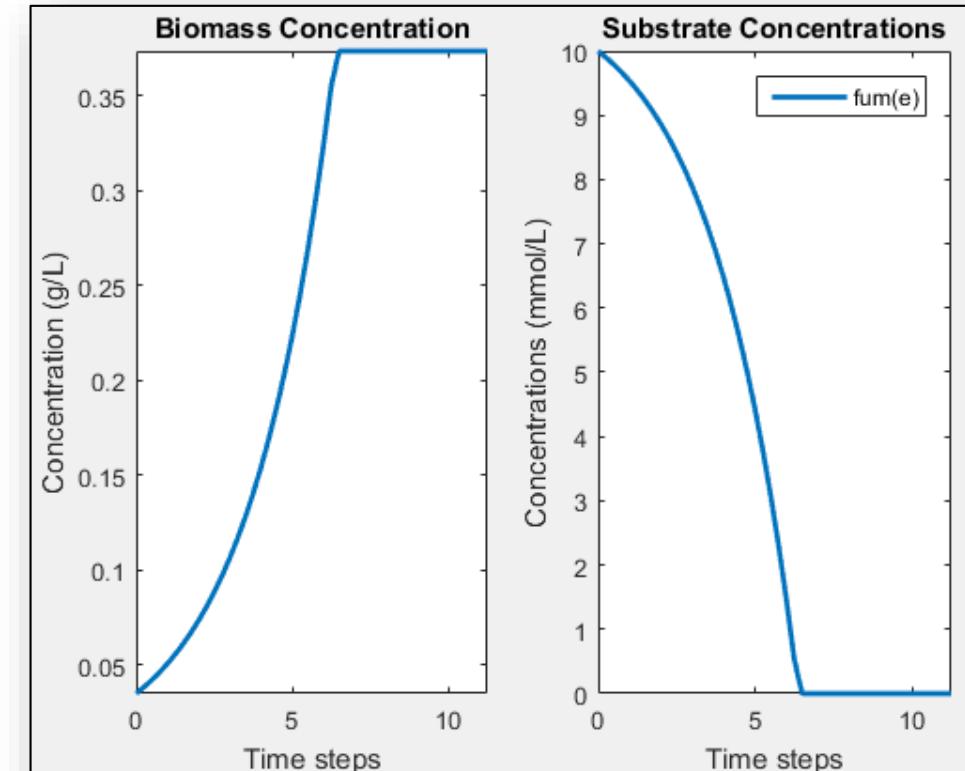
```
% Fumerate_Aerobic.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -0, '1');
model = changeRxnBounds(model, 'EX_fum(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -30, '1');

substrateRxns = {'EX_fum(e)'};
initConcentrations = [10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoh(e)', 'EX_for(e)', 'EX_fum(e)', 'EX_fru(e)', ...
    'EX_glc(e)', 'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxnNames, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Aerobic Succinate

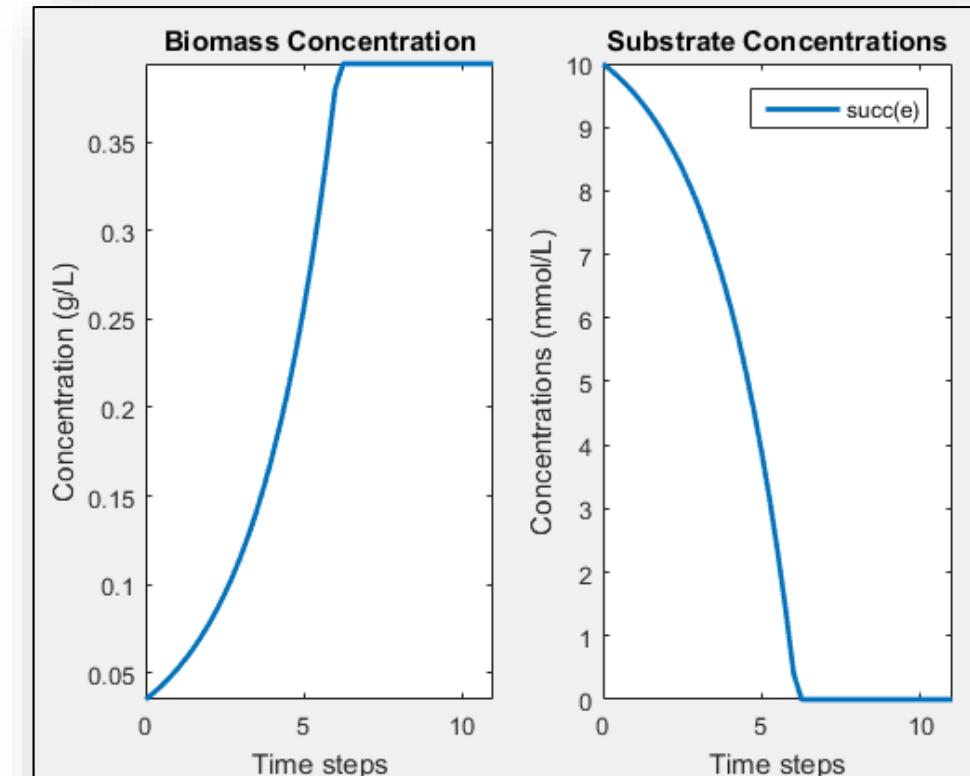
```
% Succinate_Aerobic.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -0, 'l');
model = changeRxnBounds(model, 'EX_succ(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_o2(e)', -30, 'l');

substrateRxns = {'EX_succ(e)'};
initConcentrations = [10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoh(e)', 'EX_for(e)', 'EX_fum(e)', 'EX_fru(e'), ...
    'EX_glc(e)', 'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxns, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Aerobic Fumerate & Succinate

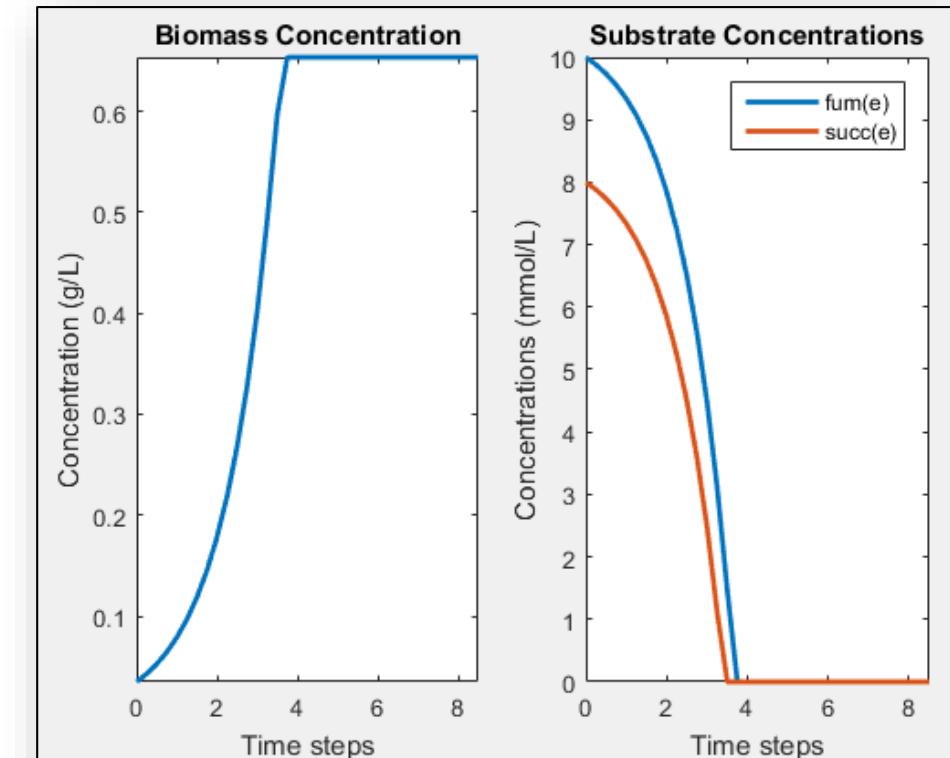
```
% FumerateSuccinate_Aerobic.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -0, '1');
model = changeRxnBounds(model, 'EX_fum(e)', -10, '1');
model = changeRxnBounds(model, 'EX_succ(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -30, '1');

substrateRxns = {'EX_fum(e)', 'EX_succ(e)'};
initConcentrations = [10 8];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etooh(e)', 'EX_for(e)', 'EX_fum(e)', 'EX_fru(e)', ...
    'EX_glc(e)', 'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxnNames, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Anoxic Pyruvate

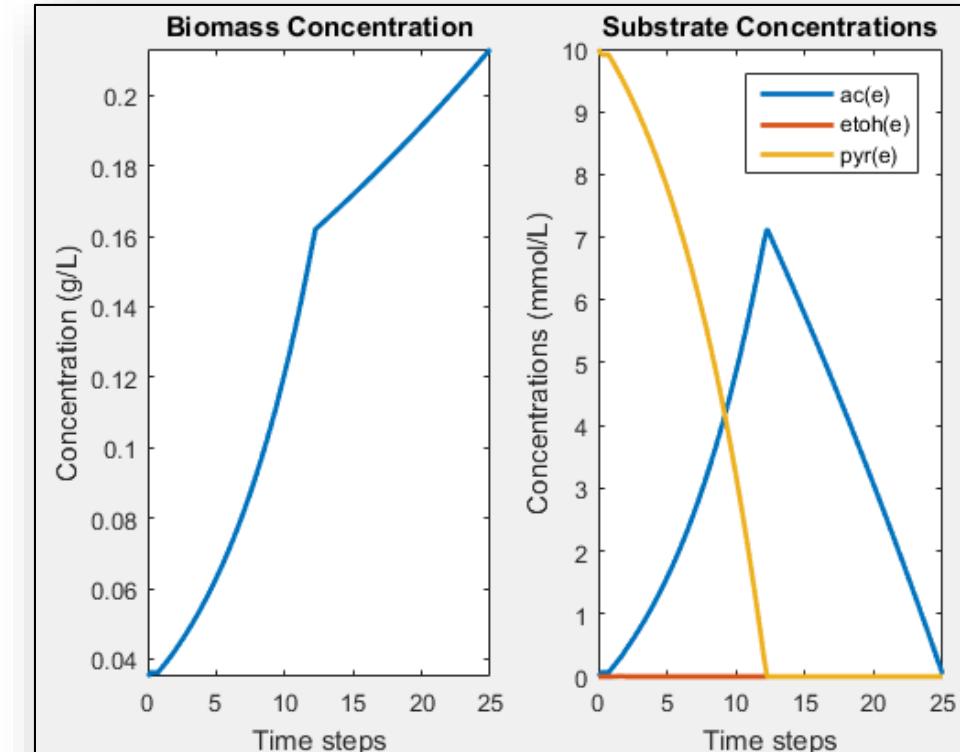
```
% Pyruvate_Anoxic.m
clear;

model = readCbModel('modelReg.mat');
model = changeRxnBounds(model, 'EX_glc(e)', -0, '1');
model = changeRxnBounds(model, 'EX_pyr(e)', -10, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -5, '1');

substrateRxns = {'EX_pyr(e)'};
initConcentrations = [10];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etoH(e)', 'EX_for(e)', 'EX_fum(e)', 'EX_fru(e)', 'EX_glc(e)', ...
    'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_pyr(e)', 'EX_succ(e)'};

[concentrationMatrix, excRxnNames, timeVec, biomassVec, drGenes, constrainedRxnNames, states] = ...
    dynamicRFBA(model, substrateRxns, initConcentrations, initBiomass, timeStep, nSteps, plotRxns);

% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# dynamicRFBA Limitations

- The dynamicRFBA tool cannot simulate the fed batch mode. There is no way to account for substrates that enter the medium via fed batch mode, so it can only show what becomes of the initial concentrations.
- The dynamicRFBA was created to optimize the biomass reaction, so there is currently no way to maximize reactions for protein production, or to maximize both the growth rate and protein production at the same time.
- The predicted growth rate can reach values higher than possible because the calculated growth rate is constantly in the exponential phase.
- The dynamicRFBA tool more useful for qualitative rather than quantitative study.



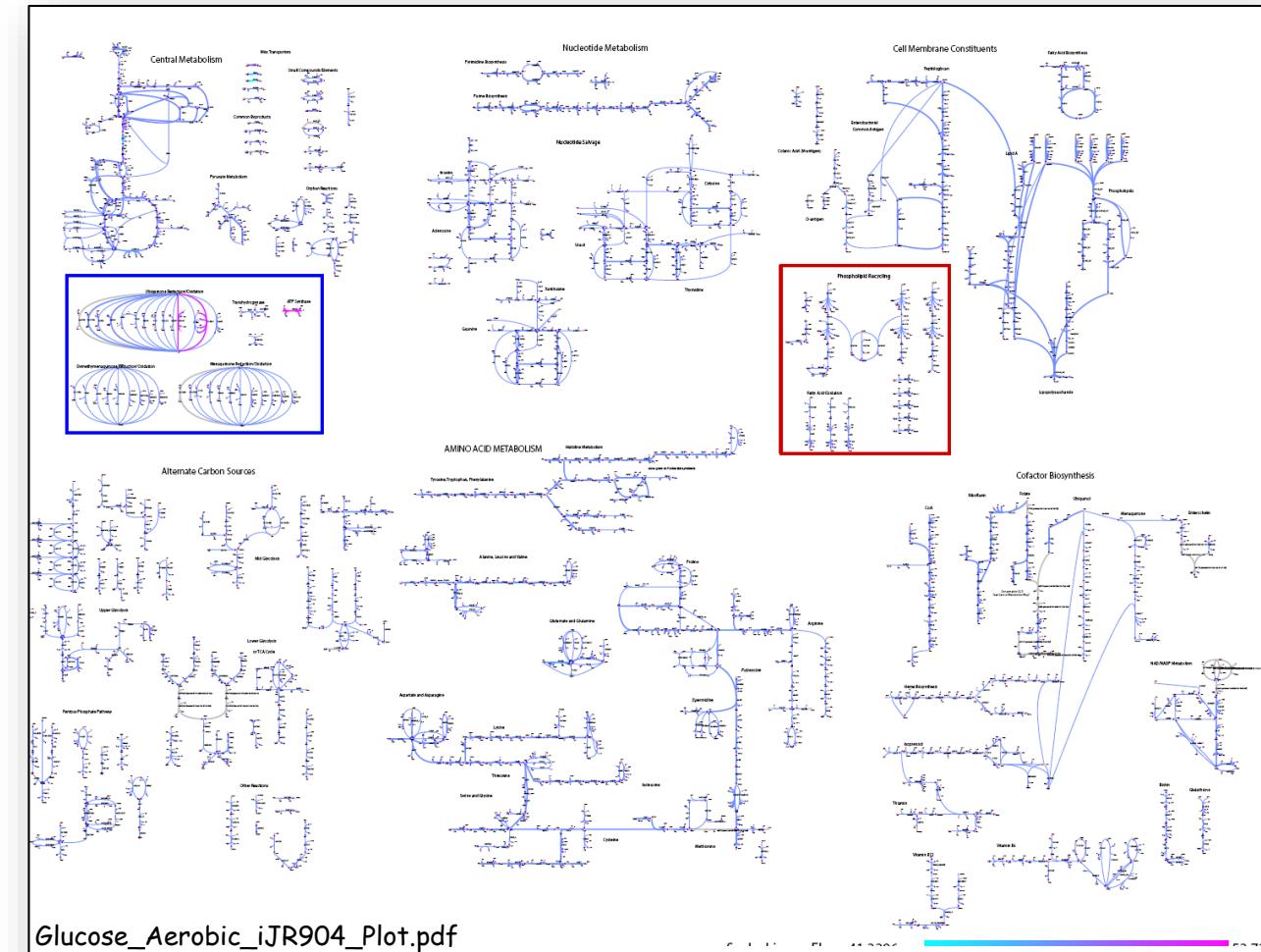
# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- • iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



# iMC1010 Regulatory Model

- Developed by Markus W. Covert
- Built on the iJR904 Model
- 1010 total genes
- 106 regulatory genes (includes 27 special regulatory variables)
- 1037 total genes & regulatory variables
- 96 external metabolites monitored
- 22 internal fluxes monitored
- Regulates the expression of 479 genes
- Validated the model against 13,750 growth phenotypes
- iMC1010<sup>v1</sup> manually curated regulation rules
- iMC1010<sup>v2</sup> modified regulation rules to align with high-throughput data



Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# iMC1010<sup>v1</sup> Model Specification

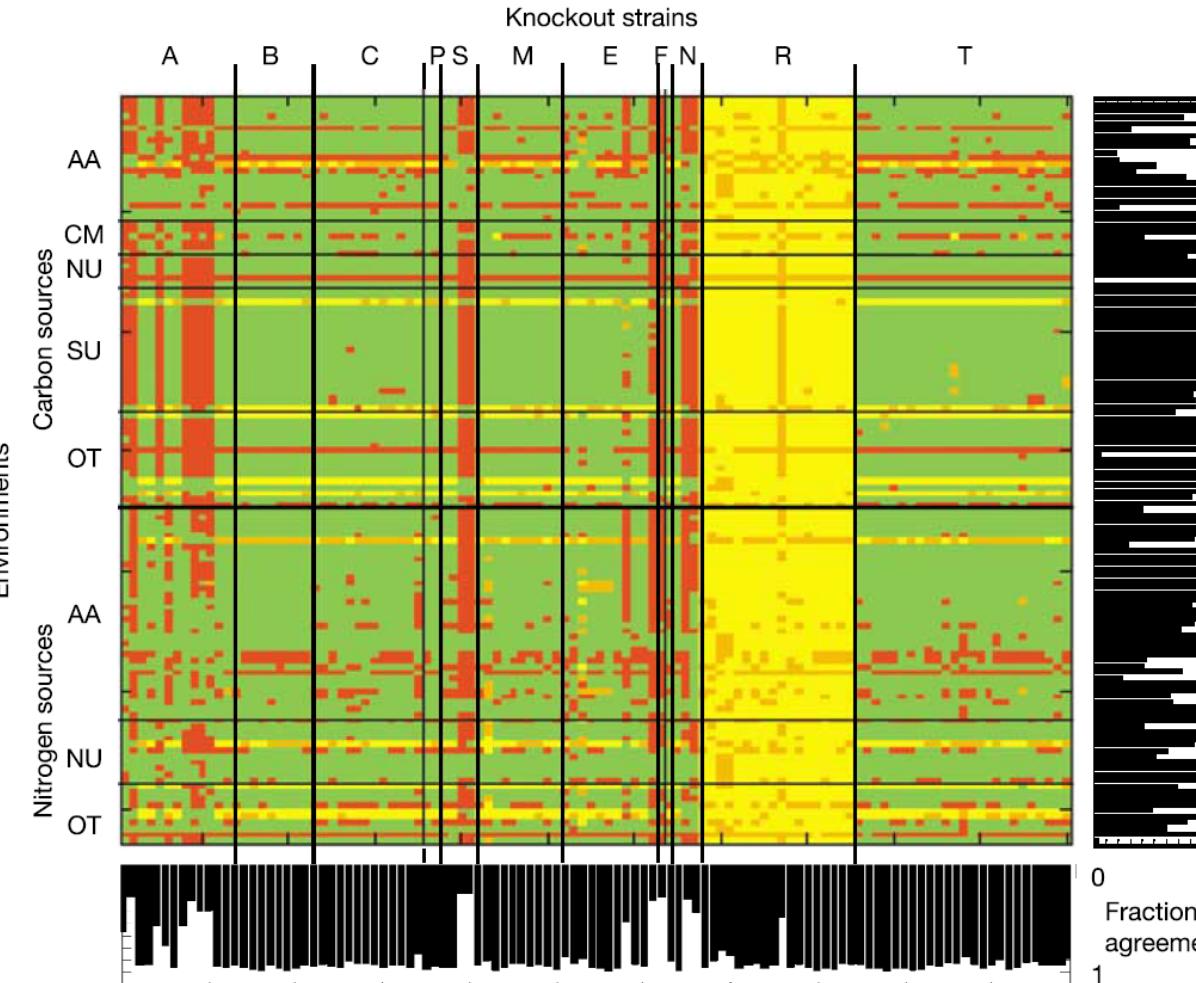
iMC1010v1.xls

bNum	Gene	Rule	Reference	Additional Comments on the Regulatory Rules			
b0002	thrA	(NOT (thr-L(e) > 0 OR ile-L(e) > 0))	NH 32	Note: Crp			
b0003	thrB	(NOT (thr-L(e) > 0 OR ile-L(e) > 0))	NH 32	Crp has complex regulation based on the level of cAMP in the cell. To describe this using Boolean logic, we divided the responses into categories based on the data of PMID: 5337847 and assuming			
b0004	thrC	(NOT (thr-L(e) > 0 OR ile-L(e) > 0))	NH 32	that repression by a "higher" level substrate was complete until the substrate was exhausted. Additionally, because most Crp testing involved only glucose, we have a more general Crp statement			
b0007	yaJ			which depends on glucose only. The resulting statements are shown below:			
b0008	talB						
b0019	nhaA	((NhaR) OR (RpoS))	PMID: 11133959	CRP GLCN	(glcn(e) > 0)		
b0020	nhaR	(na1(e) > 0)	PMID: 11133959	CRP noARAB	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0)))		
b0025	ribF			CRP noGL	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0) OR (mal-L(e) > 0) OR (glyc(e) > 0)))		
b0029	lytB			CRP noGLC	(NOT((glcn(e) > 0) OR (glc-D(e) > 0)))		
b0031	dapB	(NOT lys-L(e) > 0)	NH 32	CRP noGLCN	(NOT((glcn(e) > 0)))		
b0032	carA	(NOT ArgR)	NH 25; PMID: 9457878	CRP noGLT	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0) OR (mal-L(e) > 0) OR (glyc(e) > 0) OR (sbt-D(e) > 0)))		
b0033	carB	(NOT ArgR)	NH 25; PMID: 9457878	CRP noLAC	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0) OR (mal-L(e) > 0) OR (glyc(e) > 0) OR (sbt-D(e) > 0) OR (lac-D(e) > 0)))		
b0034	caiF	(Fnr AND Crp AND NOT NarL)	PMID: 10564497, 8631699	CRP noMAL	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0) OR (mal-L(e) > 0)))		
b0036	caiD	(Crp AND CaiF)	PMID: 10564497	CRP noMAN	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0) OR (mal-L(e) > 0) OR (glyc(e) > 0) OR (sbt-D(e) > 0) OR (lac-D(e) > 0) OR (man(e) > 0)))		
b0038	caiB	(Crp AND CaiF)	PMID: 10564497	CRP noRIB	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0)))		
b0040	caiT	(Crp AND CaiF)	PMID: 10564497	CRP noSUCC	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0) OR (rib-D(e) > 0) OR (mal-L(e) > 0) OR (glyc(e) > 0) OR (sbt-D(e) > 0) OR (lac-D(e) > 0) OR (man(e) > 0) OR (succ(e) > 0)))		
b0048	folA			CRP noXYL	(NOT((glcn(e) > 0) OR (glc-D(e) > 0) OR (arab-L(e) > 0) OR (xyl-D(e) > 0)))		
b0049	apaH						
b0052	pdxA	(RpoE)	PMID: 11844765	Note: "Surplus" internal metabolites.			
b0061	araD	(AraC OR (AraC AND Crp))	NH 20	The method we describe does not currently allow for the calculation of internal metabolite concentrations. For most cases where internal metabolite concentrations are involved in activation/repression,			
b0062	araA	(AraC OR (AraC AND Crp))	NH 20	we simply use concentration of related external metabolites as an approximation (e.g., for induction of the lac operon, we consider external lactose, rather than internal allolactose, the inducer).			
b0063	araB	(AraC OR (AraC AND Crp))	NH 20	However, in the case of important central metabolites, we use the values of connected fluxes to approximate concentration qualitatively, as shown below:			
b0064	araC	(arab-L(e) > 0)	NH 20				
b0066	sfuC			Surplus FDP	(( NOT ( FBP > 0) AND NOT (TKT2 > 0 OR TALA > 0 OR PGI > 0)) OR fru(e) > 0)		
b0067	sfuB			Surplus PYR	(( NOT ((ME2 > 0 OR ME1 > 0) AND NOT (GLCpts > 0 OR PYK > 0 OR PFK > 0 OR LDH_D < 0 OR LDH_D2 > 0 OR SUCCT2_2 > 0 OR SUCCT2_3 > 0)))		

Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# iMC1010<sup>v1</sup> Growth Phenotype Study

**b**

a	Exp/Met/Reg	Percentage
Reg and Met models predict correctly		
+/-/+	6,222	45.3%
-/-/-	2,094	15.2%
Reg model only predicts correctly		
-/+/-	657	4.8%
+/-/+	0	0%
+/n/+	1,350	9.8%
-/n/-	505	3.7%
Met model only predicts correctly		
+/n/-	242	1.8%
-/n/+	153	1.1%
-/-/+	0	0%
+/-/-	257	1.9%
Neither model predicts correctly		
-/+/+	702	5.1%
+/-/-	1,568	11.4%
Met	8,968	65.2%
Reg	10,828	78.7%
Total	13,750	100%

Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# iMC1010<sup>v1</sup> Growth on Different Carbon Sources

iMC1010v1\_carbon\_source\_growth.xls

Genes knocked out

	Medium	b3172	b3359	b3572	b3671	b3771	b3829	b3870	b3957	b3958	b3959	b4013	b4054	b4117
OD600 growth on D-Alanine	ala-D(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on D-Serine	ser-D(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Alanine	ala-L(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Asparagine	asn-L(e)-nh4(e)	-/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Aspartic Acid	asp-L(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Glutamic Acid	glu-L(e)-nh4(e)	-/-	-/-	-/+	+/+	-/-	+/+	-/+	-/-	-/-	-/-	-/-	-/+	-/+
OD600 growth on L-Glutamine	gln-L(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Proline	pro-L(e)-nh4(e)	+/-	+/-	+/+	-/+	+/-	-/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Serine	ser-L(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on L-Threonine	thr-L(e)-nh4(e)	+/-	+/-	+/+	+/+	+/-	+/+	+/+	+/-	+/-	+/-	+/-	+/+	+/+
OD600 growth on g-Amino Butyric Acid	4abut(e)-nh4(e)	-/-	-/-	-/+	-/+	-/-	+/+	-/+	+/-	+/-	+/-	+/-	-/+	-/+
OD600 growth on Glycine	gly(e)-nh4(e)	-/-	-/-	-/+	+/-	-/-	+/-	-/+	+/-	+/-	+/-	+/-	+/-	+/-
OD600 growth on L-Arginine	arg-L(e)-nh4(e)	-/+	-/+	-/+	-/+	+/-	-/+	-/+	+/-	-/+	-/+	-/+	+/-	-/+
OD600 growth on L-Histidine	his-L(e)-nh4(e)	-/-	-/-	+/-	-/-	-/-	+/-	+/-	-/-	-/-	-/-	-/-	-/-	-/-
OD600 growth on L-Isoleucine	ile-L(e)-nh4(e)	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-
OD600 growth on L-Leucine	leu-L(e)-nh4(e)	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	+/-	+/-	-/-	-/-
OD600 growth on L-Lysine	lys-L(e)-nh4(e)	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	+/-	-/-	-/-	-/-
OD600 growth on L-Methionine	met-L(e)-nh4(e)	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-
OD600 growth on L-Ornithine	orn(e)-nh4(e)	-/-	-/+	-/+	-/+	+/-	-/+	-/+	+/-	-/+	-/+	-/+	-/+	-/+

Red - Experimental different from simulated results

Green - All three results are the same

Yellow - Experimental and regulated have the same results

Orange - Experimental and non-regulated have the same results

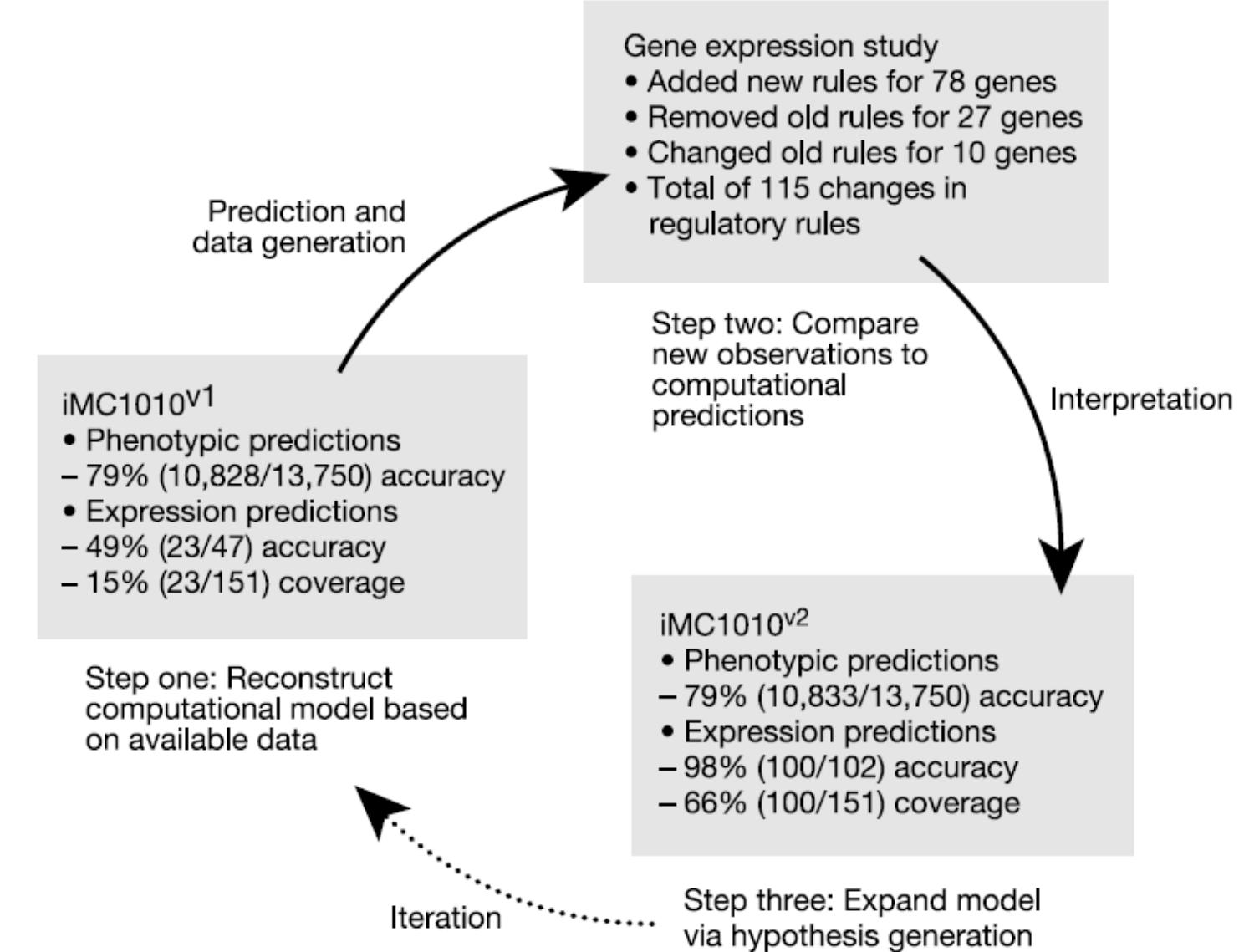
Experimental/Simulated with no regulation/ Simulation with regulation

+ = predicted or observed growth, - = no growth, n = regulatory gene not in no regulation simulation

Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# Iterative Process to Create iMC1010<sup>v2</sup> From iMC1010<sup>v1</sup>



Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# iMC1010<sup>v2</sup> Model Specification

iMC1010v2.xls

Fxn	Bnum	Gene	L2R	Old	New	Old Rule	New Rule	Comments
A	b0242	proB	-0.55	5	5	none	(ON)	Essential for growth on Arginine
A	b0828	ybiK	-0.56	5	5	none	(ON)	Essential for WT growth
A	b1261	trpB	0.28	5	5	(NOT TrpR)	((NOT TrpR))	Essential for WT growth
A	b1761	gdhA	-0.96	5	5	(NOT ((Nac) OR (glu-L(e) > 0)))	(NOT ((Nac) OR (glu-L(e) > 0)))	Essential for WT growth
A	b2021	hisC	-0.54	5	5	none	(ON)	Essential for WT growth
A	b2478	dapA	-0.24	5	5	none	none	Very small shift: ANOVA: (ArcA and Fnr) or OxyR
						(NOT(leu-L(e) > 0 OR ile-L(e) > 0 OR val-L(e) > 0) AND Lrp)	((NOT(leu-L(e) > 0 OR ile-L(e) > 0 OR val-L(e) > 0) AND Lrp) AND NOT (OxyR))	
A	b3767	ilvG_1	-0.36	5	1			
A	b3769	ilvM	-0.53	5	1	(NOT(leu-L(e) > 0 OR ile-L(e) > 0 OR val-L(e) > 0) AND Lrp)	((NOT(leu-L(e) > 0 OR ile-L(e) > 0 OR val-L(e) > 0) AND Lrp) AND (Fnr))	
A	b3770	ilvE	-0.27	5	5	none	(ON)	Essential for WT growth
A	b3771	ilvD	-0.33	5	5	none	(ON)	Essential for WT growth
A	b3957	argE	0.39	5	5	(NOT ArgR)	(NOT ArgR)	Essential for WT growth
B	b0068	sfuA	0.33	5	1	none	(o2(e) > 0)	No knockout exhibited abolished shift
B	b0133	panC	-0.48	5	5	none	(ON)	Essential for WT growth
B	b0595	entB	1.20	5	3	(NOT (Fur))	(NOT (Fur))	Fur transcription is directly opposite to activity
B	b0776	bioF	0.48	5	1	(NOT (BirA))	(NOT (BirA) AND (o2(e) > 0))	No knockout exhibited abolished shift
B	b0778	bioD	0.43	5	1	(NOT (BirA))	(NOT (BirA) AND (o2(e) > 0))	No knockout exhibited abolished shift
B	b1210	hemA	-0.31	5	5	none	(ON)	Essential for WT growth
B	b1991	cobT	-0.27	5	1	(cbi(e)>0)	((cbi(e)>0) OR (Fnr))	
B	b1993	cobU	-0.17	5	1	(cbi(e)>0)	((cbi(e)>0) OR (Fnr))	
B	b2153	folE	0.85	5	5	none	(ON)	Essential for WT growth
B	b3041	ribB	-0.96	5	5	none	(ON)	Essential for WT growth
B	b3368	cysG	-0.63	1	1	(Fnr OR NarL)	(AppY OR SoxS OR NarL)	Correct
B	b3805	hemC	-0.19	5	1	none	(NOT (o2(e) > 0))	No knockout exhibited abolished shift
B	b3929	meng	-0.50	5	5	none	none	Complex rule -- ANOVA: Fnr and not ArcA
B	b3990	thiH	0.69	5	1	none	(o2(e) > 0)	No knockout exhibited abolished shift

Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# Characterization of the Regulatory Network Related to the Aerobic-Anaerobic Shift

## a Functional groups: A,B,C,E

Locus	Gene	L2R	v1	v2
b0242	<i>proB</i>	-0.55		
b0828	<i>ybiK</i>	-0.56		
b1261	<i>trpB</i>	0.28		
b1761	<i>gdhA</i>	-0.96		
b2021	<i>hisC</i>	-0.54		
b2478	<i>dapA</i>	-0.24		
b3767	<i>ilvG_1</i>	-0.36		
b3769	<i>ilvM</i>	-0.53		
b3770	<i>ilvE</i>	-0.27		
b3771	<i>ilvD</i>	-0.33		
b3957	<i>argE</i>	0.39		
b0068	<i>sfuA</i>	0.33		
b0133	<i>panC</i>	-0.48		
b0595	<i>entB</i>	1.20		
b0776	<i>bioF</i>	0.48		
b0778	<i>bioD</i>	0.43		
b1210	<i>hemA</i>	-0.31		
b1991	<i>cobT</i>	-0.27		
b1993	<i>cobU</i>	-0.17		
b2153	<i>folE</i>	0.85		
b3041	<i>ribB</i>	-0.96		
b3368	<i>cysG</i>	-0.63		
b3805	<i>hemC</i>	-0.19		
b3929	<i>menG</i>	-0.50		

## E,F,I,N,P,R

Locus	Gene	L2R	v1	v2
b2277	<i>nuoM</i>	0.54		
b2278	<i>nuoL</i>	0.34		
b2279	<i>nuoK</i>	0.34		
b2280	<i>nuoJ</i>	0.34		
b2281	<i>nuoI</i>	0.43		
b2282	<i>nuoH</i>	0.60		
b2283	<i>nuoG</i>	0.46		
b2284	<i>nuoF</i>	0.32		
b2285	<i>nuoE</i>	0.38		
b2287	<i>nuoB</i>	0.28		
b2288	<i>nuoA</i>	0.23		
b2296	<i>ackA</i>	-1.49		
b2723	<i>hycC</i>	-3.08		
b2779	<i>eno</i>	-0.42		
b2925	<i>fbaA</i>	-0.45		
b2926	<i>pgk</i>	-0.59		
b3236	<i>mdh</i>	2.23		
b3425	<i>glpE</i>	0.83		
b3892	<i>fdol</i>	-0.04		
b3893	<i>fdoH</i>	0.61		
b3894	<i>fdoG</i>	0.90		
b3916	<i>pfkA</i>	-1.06		
b3919	<i>tpiA</i>	-0.56		
b3952	<i>pflC</i>	0.12		

## R,T,U

Locus	Gene	L2R	v1	v2
b0993	<i>torS</i>	-0.97		
b1187	<i>fadR</i>	0.54		
b1221	<i>narL</i>	0.56		
b1323	<i>tyrR</i>	-0.62		
b1334	<i>fnr</i>	0.63		
b1531	<i>marA</i>	0.90		
b1827	<i>kdgR</i>	-0.47		
b2087	<i>gatR_1</i>	0.29		
b2573	<i>rpoE</i>	-0.62		
b2707	<i>srlR</i>	-0.36		
b2731	<i>fhIA</i>	0.07		
b3357	<i>crp</i>	-0.16		
b3423	<i>glpR</i>	0.13		
b3806	<i>cyaA</i>	-0.54		
b3961	<i>oxyR</i>	0.07		
b4062	<i>soxS</i>	-1.13		
b4124	<i>dcuR</i>	0.09		
b4125	<i>dcuS</i>	-0.17		
b4401	<i>arcA</i>	-0.69		
b0314	<i>betT</i>	0.51		
b0336	<i>codB</i>	0.43		
b0401	<i>brnQ</i>	-0.65		
b0653	<i>gltK</i>	0.73		
b0854	<i>potF</i>	0.83		

- ▲ = Change in model prediction of gene expression
- △ = Change in experimental data of gene expression
- = No change in model prediction of gene expression
- = No change in experimental data of gene expression
- ? = No gene in model or no experimental data

## b

	v1	v2
Pred total	75	128
Exp total	437	437
Exp in model	151	151
▲△ or ▼▽	23	100
■▬	606	629
▲▽ or ▼△	1	2
▲▬ or ▼▬	23	0
▬△ or ▬▽	127	49

No comparison possible

▲ ? or ▼ ?	28	26
▬ ?	197	199
? ▽ or ? △	286	286
?▬	2,067	2,067
? ?	989	989

log2 ratio (L2R) of gene expression  
(aerobic to anaerobic)

Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." Nature 429(6987): 92-96.



# Anoxic Glucose Example

```
% Glucose_Low_Aerobic_iMC1010v2.m
clear;

model = readCbModel('iMC1010v2.mat'); % Reconstructed model found in lesson Matlab files

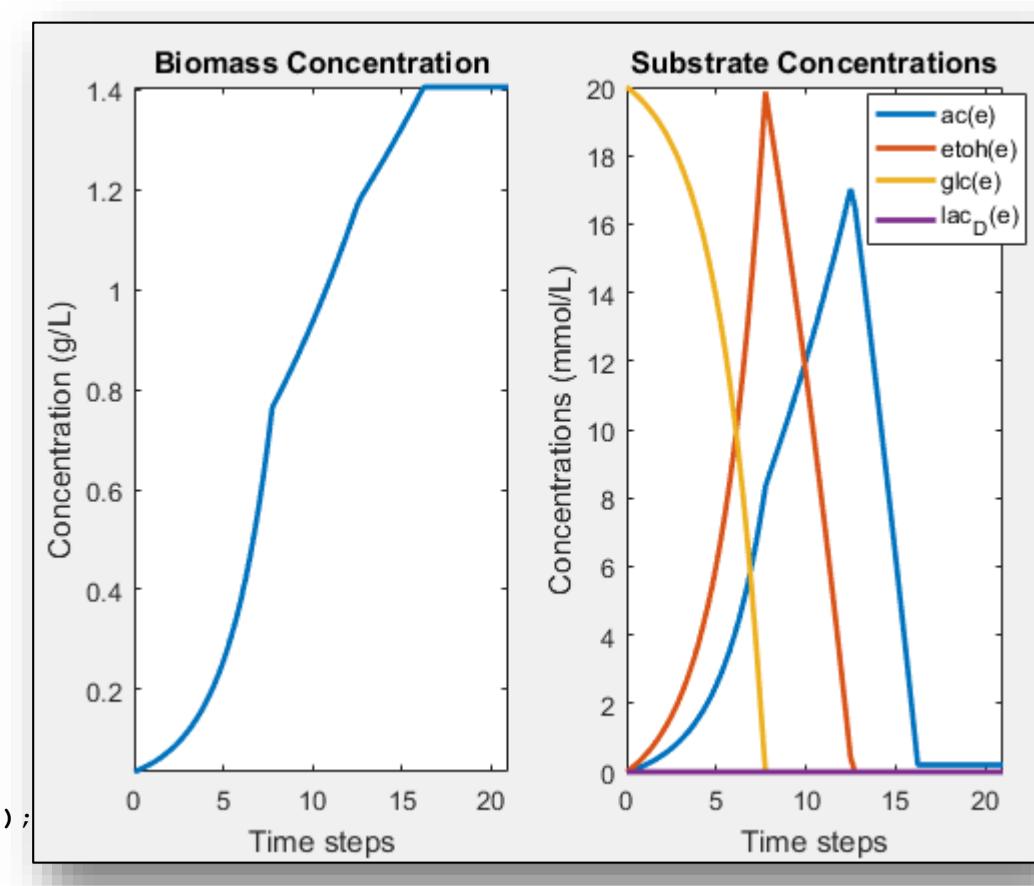
model = changeRxnBounds(model, 'EX_glc(e)', -10, 'l');
model = changeRxnBounds(model, 'EX_o2(e)', -5, 'l');

[FBAAsols,DRgenes,constrainedRxns,cycleStart,states]= optimizeRegModel(model);

substrateRxns = {'EX_glc(e)'};
initConcentrations = [20];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_etooh(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_glc(e)', ...
    'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_lac_D(e)', 'EX_mal_L(e)', 'EX_succ(e)'};

[concentrationMatrix,excRxnNames,timeVec,biomassVec,drGenes,constrainedRxns,states] = ...
    dynamicRFBA(model,substrateRxns,initConcentrations,initBiomass,timeStep,nSteps,plotRxns);

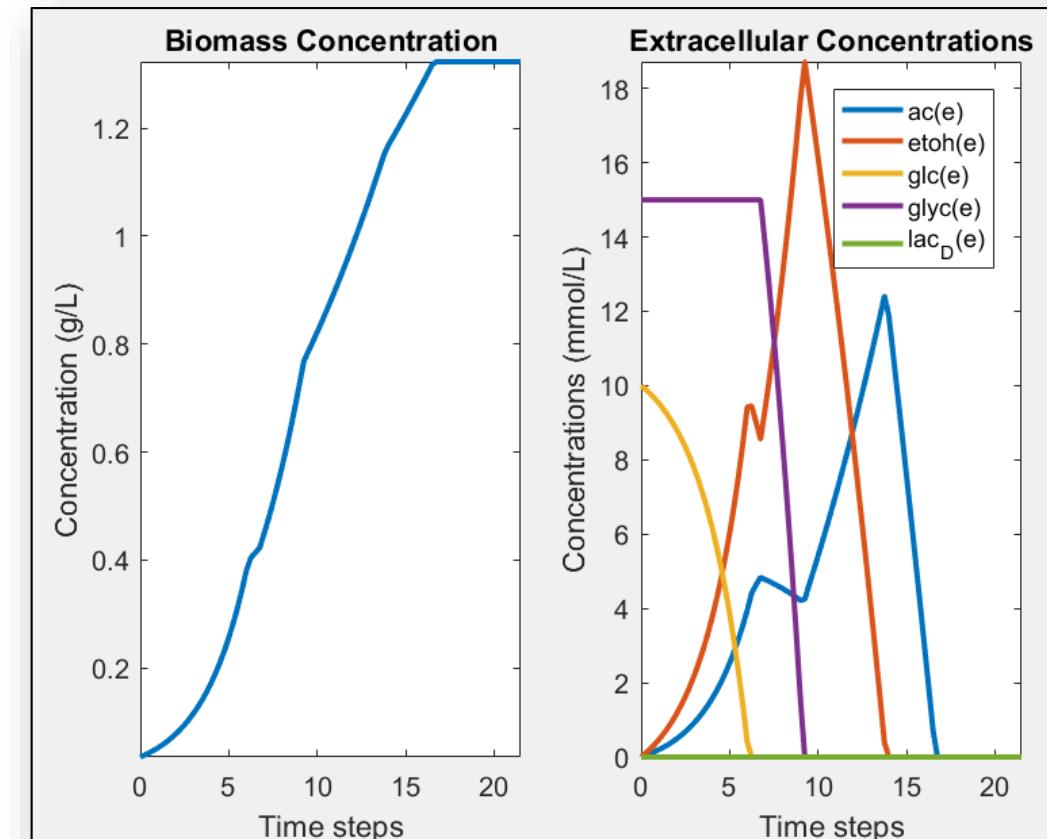
% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





# Catabolite Repression - Glucose & Glycerol

```
% Catabolite_Repression_iMC1010v2.m
clear;
model = readCbModel('iMC1010v2.mat'); % Reconstructed model found in lesson Matlab files
% Begin analysis
model = changeRxnBounds(model, 'EX_glc(e)', -10, '1');
model = changeRxnBounds(model, 'EX_lcts(e)', -0, '1');
model = changeRxnBounds(model, 'EX_glyc(e)', -10, '1');
model = changeRxnBounds(model, 'EX_succ(e)', -0, '1');
model = changeRxnBounds(model, 'EX_fum(e)', -0, '1');
model = changeRxnBounds(model, 'EX_o2(e)', -5, '1');
[FBAsol,DRgenes,constrainedRxns,cycleStart,states]= optimizeRegModel(model);
printFluxVector(model, FBAsol{1,1}.x, true)
substrateRxns = {'EX_fum(e)', 'EX_glc(e)', 'EX_glyc(e)', 'EX_lcts(e)', 'EX_succ(e)'};
initConcentrations = [0 10 15 0 0];
initBiomass = .035;
timeStep = .25;
nSteps = 100;
plotRxns = {'EX_ac(e)', 'EX_co2(e)', 'EX_etch(e)', 'EX_for(e)', 'EX_fru(e)', 'EX_fum(e)', ...
'EX_glc(e)', 'EX_gln_L(e)', 'EX_glu_L(e)', 'EX_glyc(e)', 'EX_lac_D(e)', 'EX_lcts(e)', ...
'EX_mal_L(e)', 'EX_pro_L(e)', 'EX_pyr(e)', 'EX_succ(e)'};
[concentrationMatrix,excRxnNames,timeVec,biomassVec,drGenes,constrainedRxns,states] = ...
dynamicRFBA(model,substrateRxns,initConcentrations,initBiomass,timeStep,nSteps,plotRxns);
cMatrix = full(concentrationMatrix);
% Plot labels
subplot(1,2,1); title('Biomass Concentration'); xlabel('Time steps'); ylabel('Concentration (g/L)');
subplot(1,2,2); title('Substrate Concentrations'); xlabel('Time steps'); ylabel('Concentrations (mmol/L)');
```





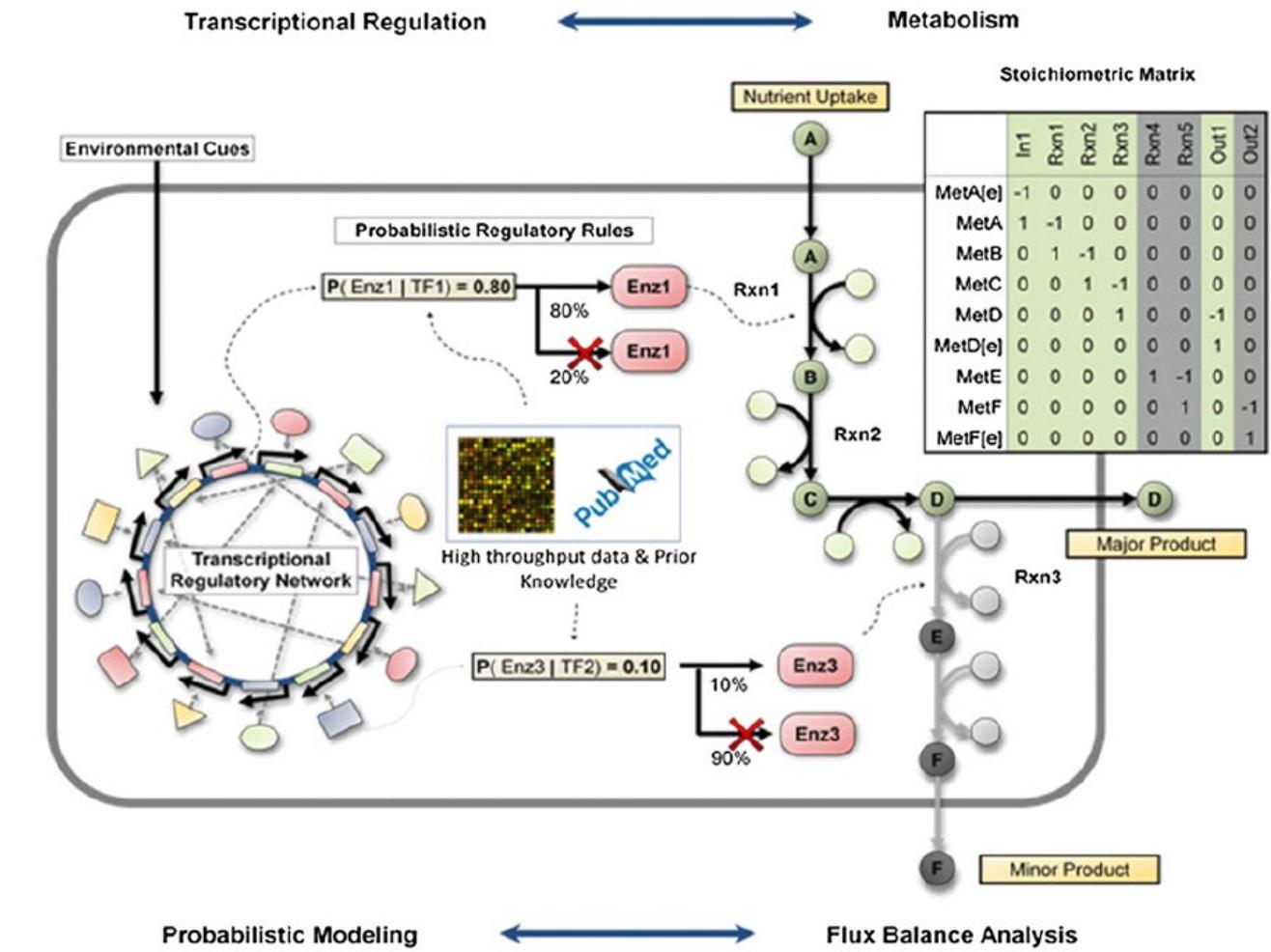
# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- ➡ • Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



# Probabilistic Regulation of Metabolism (PROM)

- PROM is an algorithm that can build and model integrated metabolic-regulatory networks in an automated fashion from high-throughput data.
- The algorithm makes use of high-throughput data, as much of the data generated currently for most new organisms are high-throughput in nature.
- PROM enables direct integration of the transcriptional and metabolic networks for modeling and overcomes the need for manually writing the Boolean rules by automatically quantifying the interactions from high-throughput data, thereby greatly increasing the capacity to generate genome-scale integrated models.
- The PROM algorithm uses conditional probabilities for modeling transcriptional regulation, similar to the probabilistic Boolean networks.
- PROM's accuracy in comparison with RFBA is highly significant, given that PROM computationally quantified the interactions using high throughput data whereas the Boolean rules for RFBA were constructed through detailed manual curation of literature.



Chandrasekaran, S (2010). "Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in Escherichia coli and Mycobacterium tuberculosis." Proc Natl Acad Sci U S A 107(41): 17845-17850.



# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- • Other Regulatory-based Model Approaches



# Other Regulatory-based Model Approaches

- SR-FBA
  - This method works by iteratively predicting a regulatory and metabolic steady state for short successive time intervals. For each time interval, a regulatory state that is consistent with the metabolic steady state of the previous interval (and with the availability of nutrients in the changing growth media) is computed. Then, FBA is used to find a steady-state flux distribution that is consistent with the regulatory state of the current time interval.
  - Shlomi, T., Y. Eisenberg, et al. (2007). "A genome-scale computational study of the interplay between transcriptional regulation and metabolism." *Molecular Systems Biology* 3: 101.
- Feuer
  - Combination of iAF1260 and iMC1010v2: The computation of a regulatory model combined with metabolic model was outlined by Covert et al.
  - Based on dynamicRFBA
  - Feuer, R., K. Gottlieb, et al. (2012). "Model-based analysis of an adaptive evolution experiment with Escherichia coli in a pyruvate limited continuous culture with glycerol." *EURASIP J Bioinform Syst Biol* 2012(1): 14.
- Tiger
  - TIGER converts a series of generalized, Boolean or multilevel rules into a set of mixed integer inequalities. The package also includes implementations of existing algorithms to integrate high-throughput expression data with genome-scale models of metabolism and transcriptional regulation.
  - Jensen, P. A., K. A. Lutz, et al. (2011). "TIGER: Toolbox for integrating genome-scale metabolic models, expression data, and transcriptional regulatory networks." *BMC systems biology* 5: 147.



# Lesson Outline

- Transcriptional Regulatory Networks
- Dynamic Regulatory Flux Balance Analysis (dynamicRFBA)
- Regulated *E.coli* Core Model
  - ✓ Regulatory Genes/Transcription Factors
  - ✓ Catabolite Repression Examples
  - ✓ Growth on Non-glucose Carbon Sources
- iMC1010 Regulatory Model
- Probabilistic Regulation of Metabolism (PROM)
- Other Regulatory-based Model Approaches



# Transcriptional Regulatory Networks Review Questions

1. What is a transcriptional regulatory network?
2. What is a Boolean regulatory network?
3. What is a probabilistic transcriptional regulatory network?
4. What is the relationship between a regulatory network and a metabolic network?
5. What are the key inputs required for dynamic regulatory FBA operation?
6. What is a regulated model? How is it different from a normal FBA model?
7. What is a regulatory rule?
8. What role do the external metabolite inputs play in dRFBA?
9. What role do the internal reaction inputs play in dRFBA?
10. What is the difference between a regulatory gene and a gene that is controlled by regulation?
11. What is catabolite repression?
12. What are some of the differences between the output of dynamicFBA and dynamicRFBA?
13. What is the difference between the iMC1010v1 and iMC1010v2 models?
14. What is the difference between dRFBA and PROM regulatory algorithms?
15. What are the strengths of dynamic regulatory FBA?
16. What are the limitations of dynamic regulatory FBA?



# References

## Dynamic Regulatory Flux Balance Analysis

- Feuer, R., K. Gottlieb, et al. (2012). "Model-based analysis of an adaptive evolution experiment with *Escherichia coli* in a pyruvate limited continuous culture with glycerol." *EURASIP J Bioinform Syst Biol* 2012(1): 14.
- Jensen, P. A., K. A. Lutz, et al. (2011). "TIGER: Toolbox for integrating genome-scale metabolic models, expression data, and transcriptional regulatory networks." *BMC systems biology* 5: 147.
- Barua, D., J. Kim, et al. (2010). "An automated phenotype-driven approach (GeneForce) for refining metabolic and regulatory models." *PLoS computational biology* 6(10): e1000970.
- Gianchandani, E. P., A. R. Joyce, et al. (2009). "Functional states of the genome-scale *Escherichia coli* transcriptional regulatory system." *PLoS computational biology* 5(6): e1000403.
- Gianchandani, E. P., J. A. Papin, et al. (2006). "Matrix formalism to describe functional states of transcriptional regulatory systems." *PLoS computational biology* 2(8): e101.
- Barrett, C. L., C. D. Herring, et al. (2005). "The global transcriptional regulatory network for metabolism in *Escherichia coli* exhibits few dominant functional states." *Proceedings of the National Academy of Sciences of the United States of America* 102(52): 19103-19108.
- Herrgard, M. J., M. W. Covert, et al. (2004). "Reconstruction of microbial transcriptional regulatory networks." *Current opinion in biotechnology* 15(1): 70-77.  
Covert, M. W., E. M. Knight, et al. (2004). "Integrating high-throughput and computational data elucidates bacterial networks." *Nature* 429(6987): 92-96.
- Herrgard, M. J., M. W. Covert, et al. (2003). "Reconciling gene expression data with known genome-scale regulatory network structures." *Genome research* 13(11): 2423-2434.
- Covert, M. W., I. Famili, et al. (2003). "Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?" *Biotechnology and bioengineering* 84(7): 763-772.
- Covert, M. W. and B. O. Palsson (2002). "Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*." *The Journal of biological chemistry* 277(31): 28058-28064.
- Covert, M. W., C. H. Schilling, et al. (2001). "Regulation of gene expression in flux balance models of metabolism." *Journal of theoretical biology* 213(1): 73-88.



# References (2)

## Dynamic Analysis

- [B. Ø. Palsson, "Systems Biology: Properties of Reconstructed Networks," Cambridge University Press, 2006.](#)
- [Lee, J. M., E. P. Ganchandani, et al. \(2008\). "Dynamic analysis of integrated signaling, metabolic, and regulatory networks." PLoS computational biology 4\(5\): e1000086.](#)

## Dynamic FBA Growth Simulations

- [Becker, S. A., A. M. Feist, et al. \(2007\). "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox." Nature protocols 2\(3\): 727-738. - See section on Dynamic FBA growth simulations \(batch growth simulations\)- pp. 731,734](#)
- [Varma, A. and B. O. Palsson \(1994\). "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110." Applied and Environmental Microbiology 60\(10\): 3724-3731.](#)

## Integrated Dynamic Regulatory Flux Balance Analysis

- [Imam, S., S. Schauble, et al. \(2015\). "Data-driven integration of genome-scale regulatory and metabolic network models." Frontiers in microbiology 6: 409.](#)
- [Carrera, J., R. Estrela, et al. \(2014\). "An integrative, multi-scale, genome-wide model reveals the phenotypic landscape of Escherichia coli." Molecular Systems Biology 10: 735.](#)
- [Covert, M. W., N. Xiao, et al. \(2008\). "Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli." Bioinformatics 24\(18\): 2044-2050. \(iFBA\)](#)
- [Shlomi, T., Y. Eisenberg, et al. \(2007\). "A genome-scale computational study of the interplay between transcriptional regulation and metabolism." Molecular Systems Biology 3: 101.](#)



# References (3)

## Probabilistic Regulation

- [Chandrasekaran, S. and N. D. Price \(2013\). "Metabolic constraint-based refinement of transcriptional regulatory networks." PLoS computational biology 9\(12\): e1003370.](#)
- [Simeonidis, E., S. Chandrasekaran, et al. \(2013\). "A guide to integrating transcriptional regulatory and metabolic networks using PROM \(probabilistic regulation of metabolism\)." Methods in molecular biology 985: 103-112](#)
- [Chandrasekaran, S \(2010\). "Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in Escherichia coli and Mycobacterium tuberculosis." Proc Natl Acad Sci U S A 107\(41\): 17845-17850.](#)
- [Shmulevich, I., E. R. Dougherty, et al. \(2002\). "Probabilistic Boolean Networks: a rule-based uncertainty model for gene regulatory networks." Bioinformatics 18\(2\): 261-274.](#)
- [Shmulevich, I., E. R. Dougherty, et al. \(2002\). "Gene perturbation and intervention in probabilistic Boolean networks." Bioinformatics 18\(10\): 1319-1331.](#)

## General Transcriptional Regulation

- [Cho, B. K., S. Federowicz, et al. \(2012\). "Deciphering the transcriptional regulatory logic of amino acid metabolism." Nature chemical biology 8\(1\): 65-71.](#)
- [Thiele, I., R. M. Fleming, et al. \(2010\). "Functional characterization of alternate optimal solutions of Escherichia coli's transcriptional and translational machinery." Biophysical journal 98\(10\): 2072-2081.](#)
- [Thiele, I., N. Jamshidi, et al. \(2009\). "Genome-scale reconstruction of Escherichia coli's transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization." PLoS computational biology 5\(3\): e1000312.](#)
- [Shlomi, T., O. Berkman, et al. \(2005\). "Regulatory on/off minimization of metabolic flux changes after genetic perturbations." Proceedings of the National Academy of Sciences of the United States of America 102\(21\): 7695-7700.](#)
- [Gutierrez-Rios, R. M., D. A. Rosenblueth, et al. \(2003\). "Regulatory network of Escherichia coli: consistency between literature knowledge and microarray profiles." Genome research 13\(11\): 2435-2443.](#)
- [Shen-Orr, S. S., R. Milo, et al. \(2002\). "Network motifs in the transcriptional regulation network of Escherichia coli." Nature genetics 31\(1\): 64-68.](#)