How does a bacterium achieve optimal flux partitioning?

Lei-Han Tang

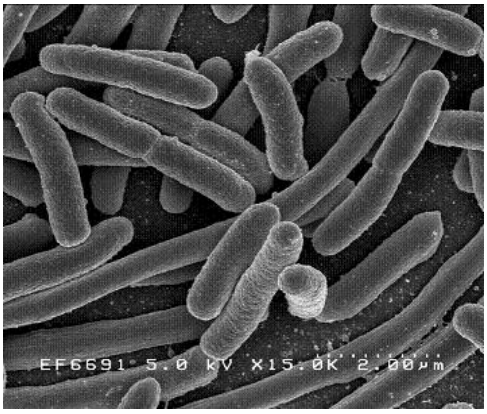
Department of Physics, Hong Kong Baptist University

Fast growth = efficient usage of resources ⊕ speed

addressed by FBA

regulation at branch points of metabolic flow

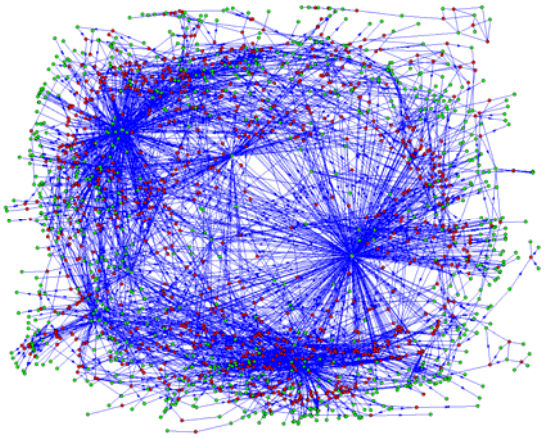
Further coarse-graining (examine global gene expression) to identify nonessential activities and bottlenecks



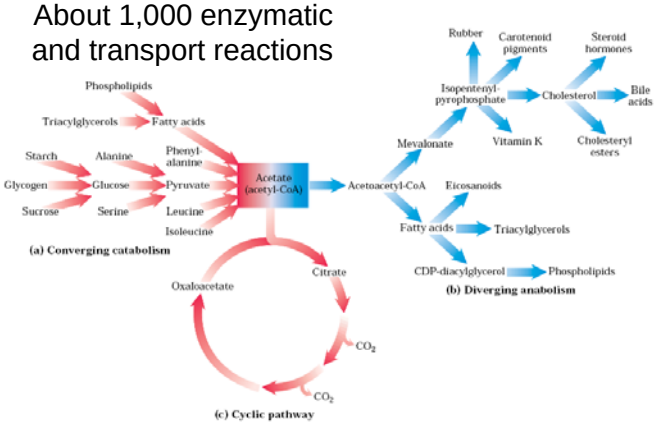
Makes a run when condition is good

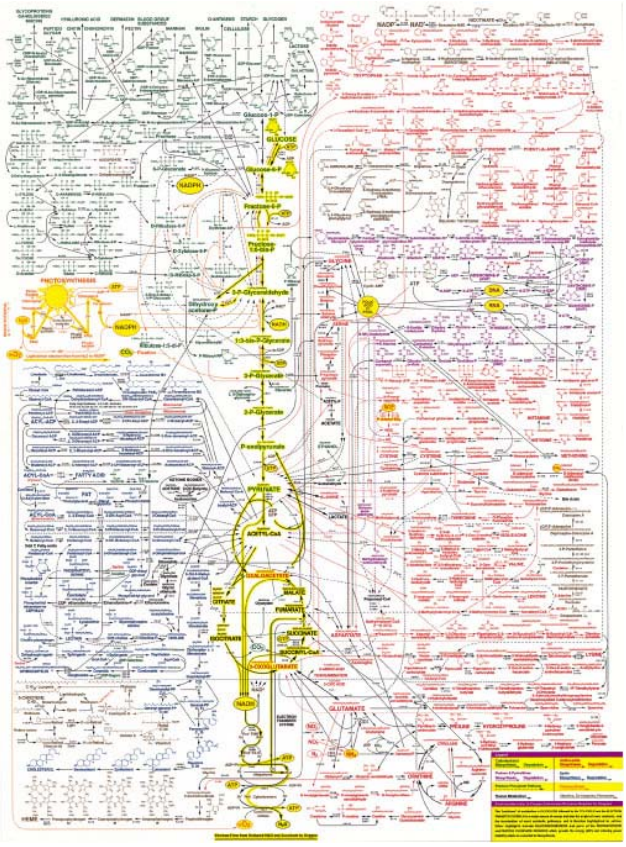
⇒ Divides in approximately 30 min.

What's underlying cell growth?



About 1,000 enzymatic and transport reactions





The Grand Synthesis

Genomic data

EcoCyc

KEGG Kyoto Encyclopedia of Genes and Genomes

4,200 genes

Gene expression

EcoliHub

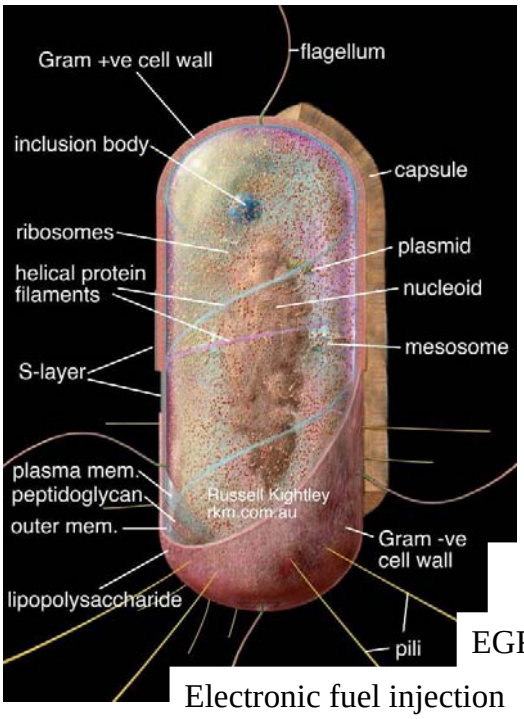
ASAP

Over 270
transcription factors

Compound concentrations

BRENDA The Comprehensive Enzyme Information System

1,300 interactions
by metabolites



John Doyle:

Organized complexity

Temperature control

Air bags

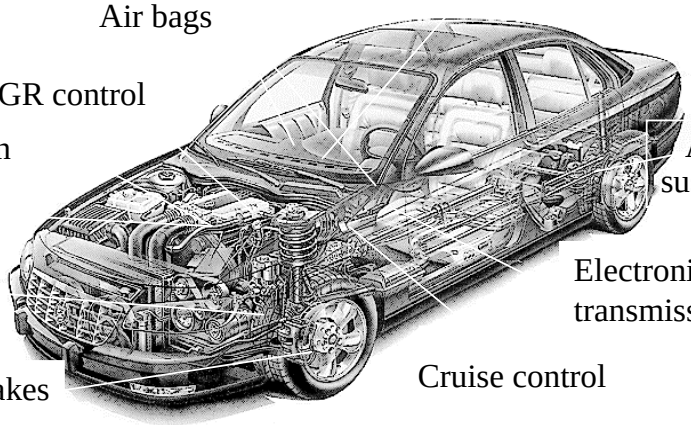
EGR control

Electronic fuel injection

Electronic ignition

Electric power steering (PAS)

Anti-lock brakes



Active suspension

Electronic transmission

Cruise control

Optimal performance: Flux Balance Analysis (FBA)

N. D. Price, J. L. Reed, B. O. Palsson, *Nature Reviews Microbiology* **2**, 886-897 (2004)

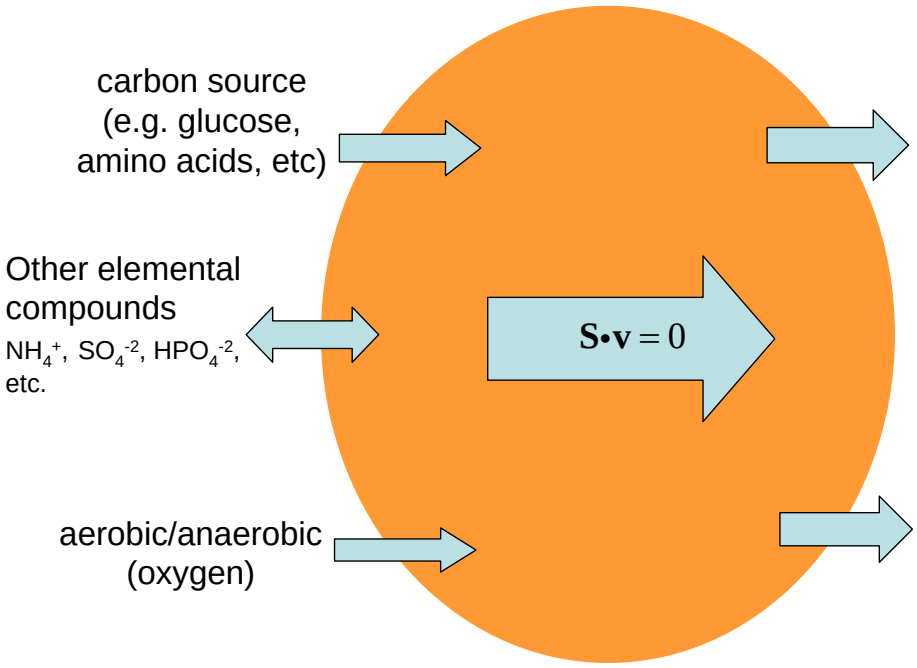


Table 9. Cellular components of *S. cerevisiae*

ALA	0.459	CMP	0.05
ARG	0.161	dAMP	0.0036
ASN	0.102	dCMP	0-0024
ASP	0.297	dGMP	0.0024
CYS	0.007	DTMP	0.0036
GLU	0.302	TAGLY	0.007
GLN	0.105	ERGOST	0.0007
GLY	0.290	ZYMST	0.015
HIS	0.066	PA	0.0006
ILE	0.193	PINS	0.005
LEU	0.296	PS	0.002
LYS	0.286	PE	0.005
MET	0.051	PC	0.006
PHE	0.134	GLYCOGEN	0.519
PRO	0.165	TRE	0.023
SER	0.185	Mannan	0.809
THR	0.191	13GLUCAN	1.136
TRP	0.028	SLF	0.02
TYR	0.102	ATP	23.9166
VAL	0.265	ADP	23.9166
AMP	0.051	PI	23.9456
GMP	0.051	Biomass	1
UMP	0.067		

biomass

Use linear programming to determine flux pattern for optimal biomass production

$\Rightarrow$ Upper limit on growth speed

All is good, except cell does not know linear programming!

And at times, it does not follow the optimal growth predicted by the FBA

What is the basic scheme of metabolic regulation?

- Identify decision points to reach optimal flux pattern
- Regulatory strategies (3-levels)
  - Transcriptional regulation to eliminate futile cycles (as well as setting the enzyme copy number to a suitable level), minutes
  - Small molecule regulation to achieve fast equilibrium, ms to second.
  - Homeostasis of global commodities
- Incorporate proteomic, metabolomic and transcriptomic data to iteratively improve the model parameters

Network simplification to highlight carbon flow

An automated program to reveal carbon flow network

Free-standing

compound	frequency
h[c]	497
h2o[c]	293
pi[c]	149
ppi[c]	75
nh4[c]	38
o2[c]	16

compound	frequency	formula
co2[c]	45	CO2
pyr[c]	16	C3H3O3
ac[c]	8	C2H3O2
for[c]	8	CHO2
succ[c]	7	C4H4O4
fum[c]	3	C4H2O4
hco3[c]	2	CHO3

Adenosine deaminase: adn + ~~h~~ + ~~h2o~~ --> ins + ~~nh4~~

Hexokinase: ~~atp~~ + glc-D --> ~~adp~~ + g6p + ~~h~~

Carriers

Cmpd1	Cmpd2	cargo	frequency
nadh[c]	nad[c]	H	71
nadph[c]	nadp[c]	H	49
q8h2[c]	q8[c]	H2	17
mqn8[c]	mql8[c]	H2	16
2dmmql8[c]	2dmmq8[c]	H2	10
trdrd[c]	trdox[c]	H2	10
fadh2[c]	fad[c]	H2	8
fum[c]	succ[c]	H2	5

Cmpd1	Cmpd2	cargo	frequnecy
gln-L[c]	glu-L[c]	H2NO	13
asn-L[c]	asp-L[c]	H2NO	3
glu-L[c]	akg[c]	H4NO	20
atp[c]	adp[c]	O3P	136
atp[c]	amp[c]	O6P2	25
pep[c]	pyr[c]	HO3P	16
gtp[c]	gdp[c]	O3P	6
adp[c]	amp[c]	O3P	3

Currency compounds

Nitrogenous: NH4, NO, NO2, NO3, etc.

Phosphates: PO4, diphosphate, etc.

Sulfate/sulfite: SO3, SO4, etc.

Metal ions: Fe, Na, K, etc.

Water, hydrogen, oxygen

Carriers

ATP/ADP/AMP, GTP/GDP,CTP/CMP

nad/nadh, nadp/nadph, fad/fadh

q8/q8h2, mql8/mqn8, 2dmmq8/2dmmql8

akg/glu/gln

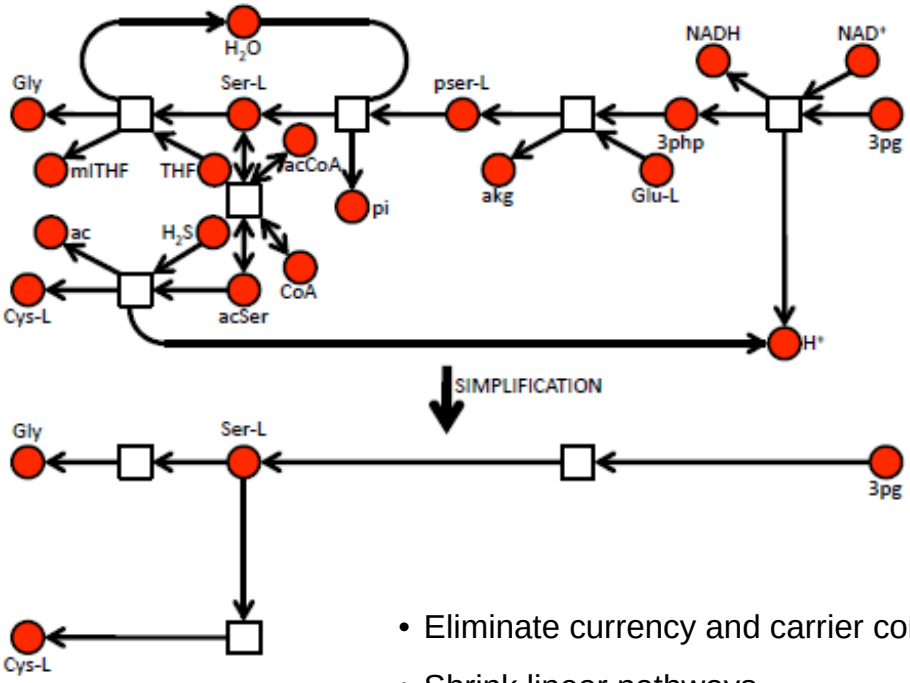
acCoA/CoA, sucCoA/CoA, pep/pyr

Coenzymes/cofactors

ACP, THF, udcpp, etc.

High degree nodes are currency and carrier compounds!

An example of the network simplification scheme

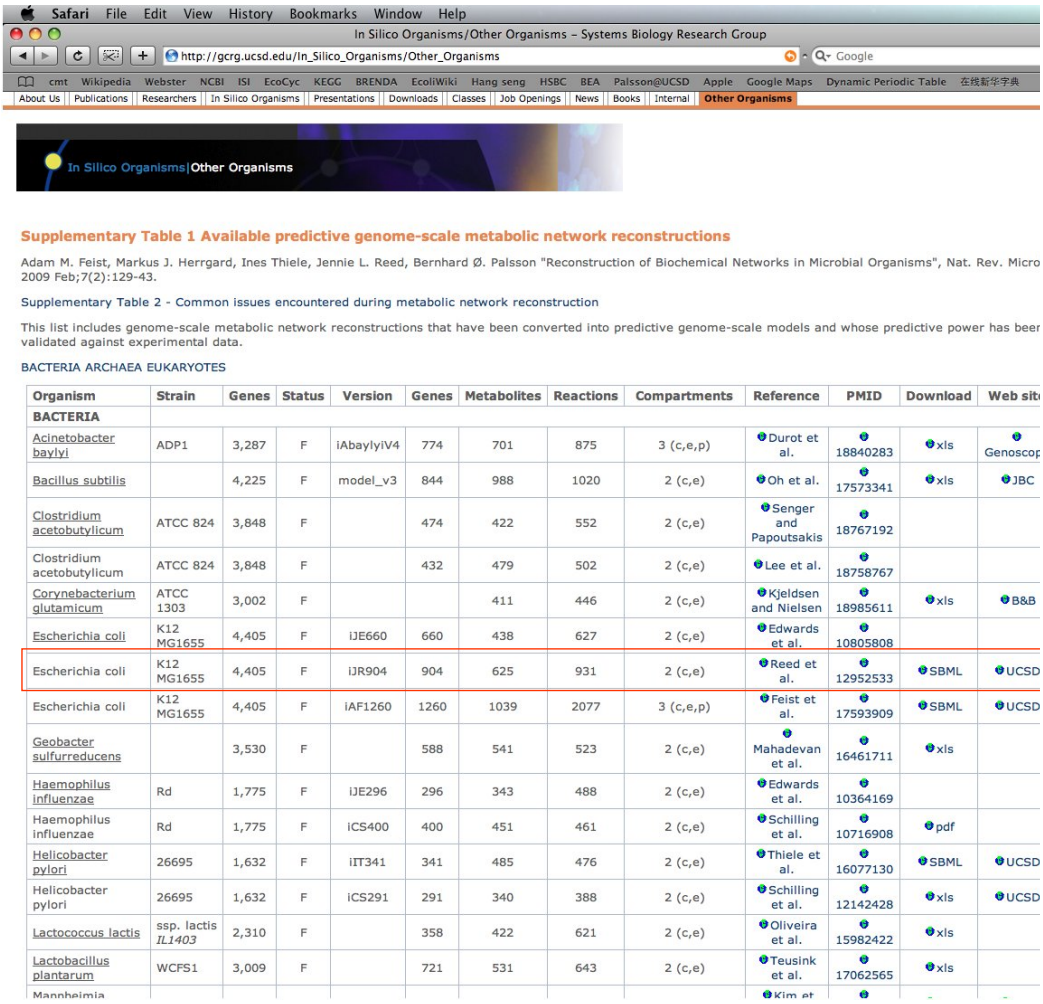


- Eliminate currency and carrier compounds
- Shrink linear pathways

in silico organisms

constructed by Palsson's group at UCSD

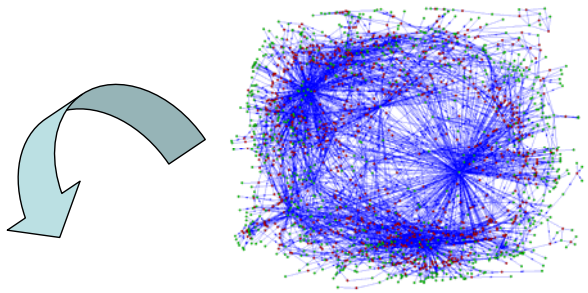
- Collection of organism-specific reactions leading to biomass production
- Linear programming to find the optimal state
- Allow for simulation of different growth conditions (nutrient uptake, O2 availability, etc.)
- **Outcome**: growth rate and flux pattern

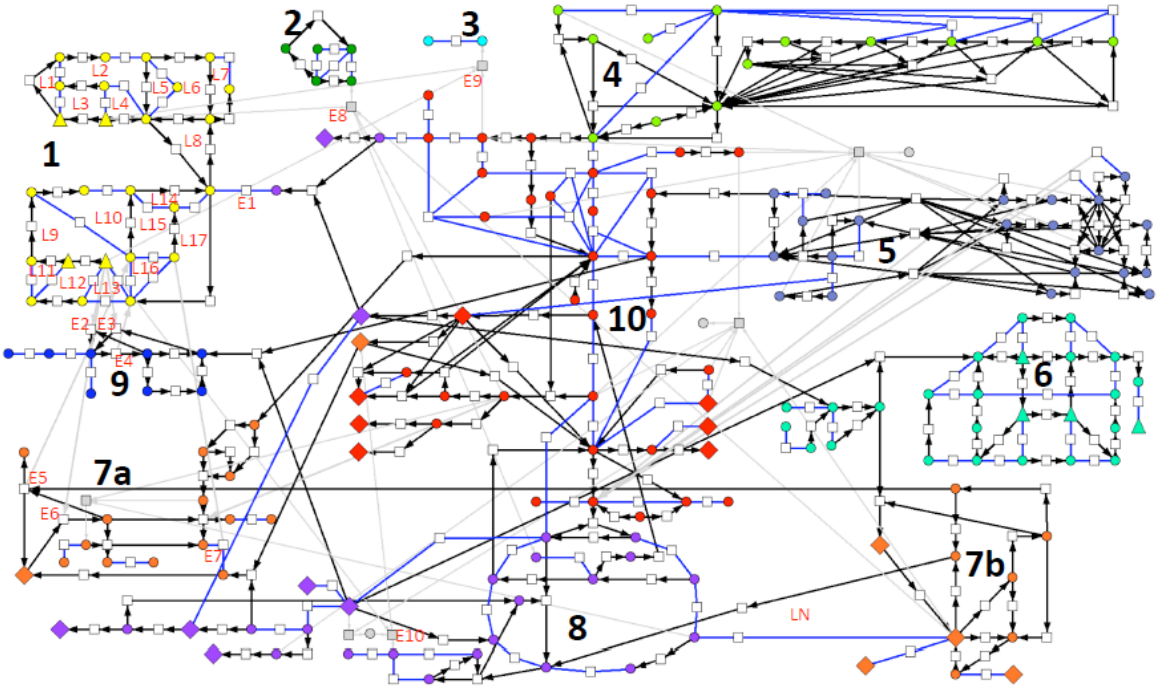


Organism	Strain	Genes	Status	Version	Genes	Metabolites	Reactions	Compartments	Reference	PMID	Download	Web sit
BACTERIA												
Acinetobacter baylyi	ADP1	3,287	F	iAbaylyiV4	774	701	875	3 (c,e,p)	Durot et al.	18840283	xls	Genoscop
Bacillus subtilis		4,225	F	model_v3	844	988	1020	2 (c,e)	Oh et al.	17573341	xls	JBC
Clostridium acetobutylicum	ATCC 824	3,848	F		474	422	552	2 (c,e)	Senger and Papoutsakis	18767192		
Clostridium acetobutylicum	ATCC 824	3,848	F		432	479	502	2 (c,e)	Lee et al.	18758767		
Corynebacterium glutamicum	ATCC 1303	3,002	F			411	446	2 (c,e)	Kjeldsen and Nielsen	18985611	xls	B&B
Escherichia coli	K12 MG1655	4,405	F	iJE660	660	438	627	2 (c,e)	Edwards et al.	10805808		
Escherichia coli	K12 MG1655	4,405	F	iJR904	904	625	931	2 (c,e)	Reed et al.	12952533	SBML	UCSD
Escherichia coli	K12 MG1655	4,405	F	iAF1260	1260	1039	2077	3 (c,e,p)	Feist et al.	17593909	SBML	UCSD
Geobacter sulfurreducens		3,530	F		588	541	523	2 (c,e)	Mahadevan et al.	16461711	xls	
Haemophilus influenzae	Rd	1,775	F	iJE296	296	343	488	2 (c,e)	Edwards et al.	10364169		
Haemophilus influenzae	Rd	1,775	F	iCS400	400	451	461	2 (c,e)	Schilling et al.	10716908	pdf	
Helicobacter pylori	26695	1,632	F	iIT341	341	485	476	2 (c,e)	Thiele et al.	16077130	SBML	UCSD
Helicobacter pylori	26695	1,632	F	iCS291	291	340	388	2 (c,e)	Schilling et al.	12142428	xls	UCSD
Lactococcus lactis	ssp. lactis *IL1403*	2,310	F		358	422	621	2 (c,e)	Oliveira et al.	15982422	xls	
Lactobacillus plantarum	WCFS1	3,009	F		721	531	643	2 (c,e)	Teusink et al.	17062565	xls	
Mannheimia									Kim et			

Simplified network based on iJR904

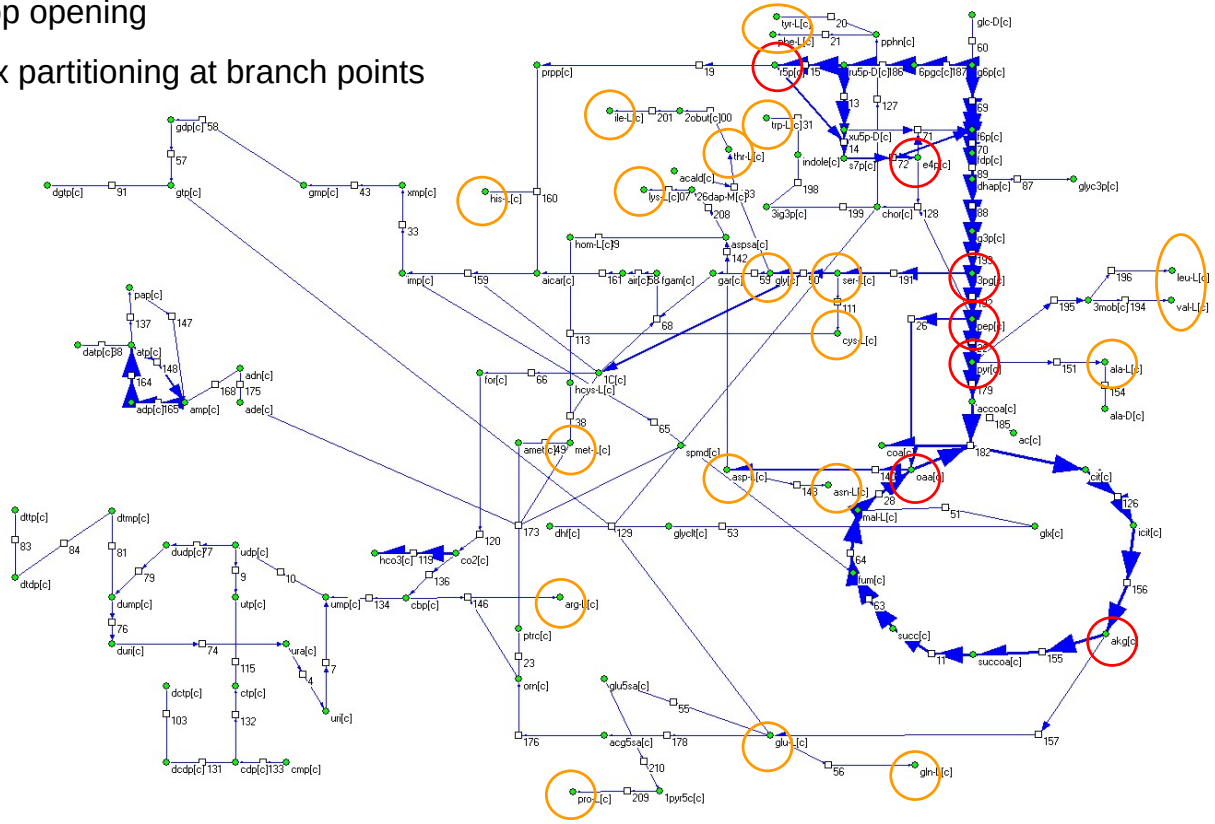
- tree like
- community structure
- loops





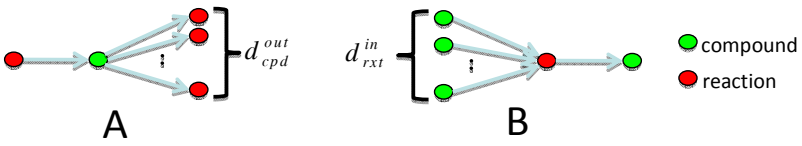
Optimal flow identified by FBA

- Loop opening
- Flux partitioning at branch points



DEGREES OF FREEDOM

Due to flux balance, the actual degree of freedom (312, including 124 transport reactions in iJR904) is much less than the number of reaction flux variables (1075).



Rules to count DoF on a network of compounds and irreversible reactions*:

1. For compounds (figure A):

$$DoF_{cpd} = \sum_{cpd}\left(d^{out}_{cpd} - 1\right)$$

Branching increases degrees of freedom

2. For reactions (figure B):

$$DoF_{rxn} = -\sum_{rxn}\left(d^{in}_{rxn} - 1\right)$$

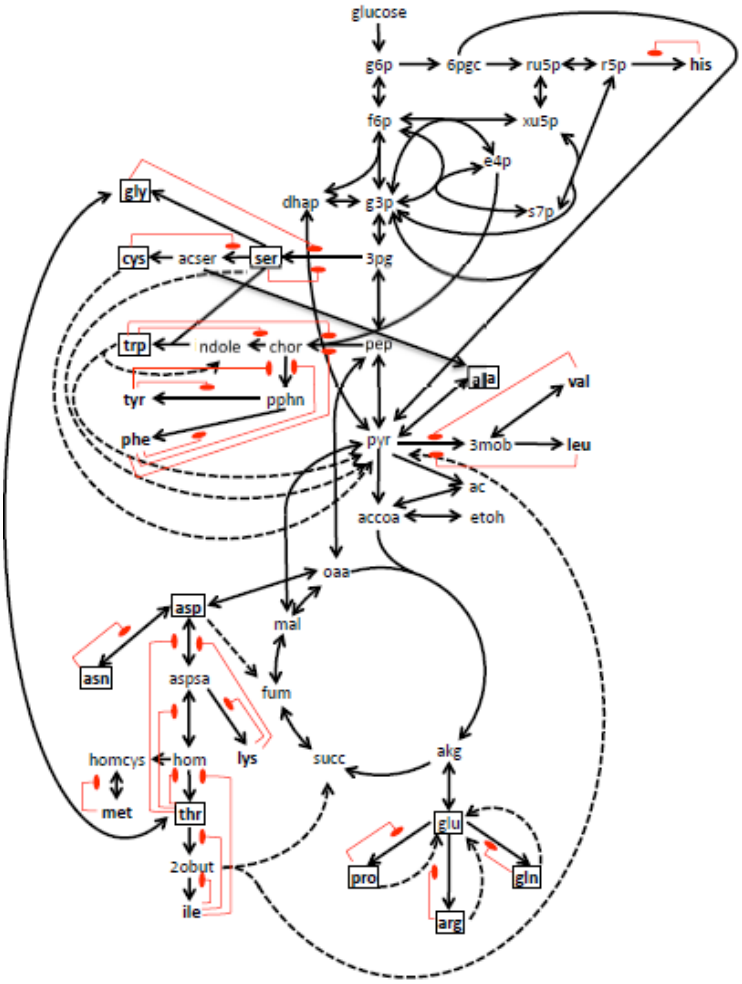
Synthesis decreases degrees of freedom

3. Thus

$$DoF_{total} = DoF_{cpd} + DoF_{rxn}$$

How does E. coli do it?

Amino acid biosynthesis subnetwork



Carbon flow network for amino acid metabolism

Amino acids can be both synthesized from sugar or used as carbon source

transcriptional regulation to eliminate futile cycles in carbon flow

Allen TE, et al. (2003) Genome-scale analysis of the uses of the Escherichia coli genome: model-driven analysis of heterogeneous data sets. J Bacteriol 185(21):6392-6399

Fold = $\log_2$ (LB/glucose)

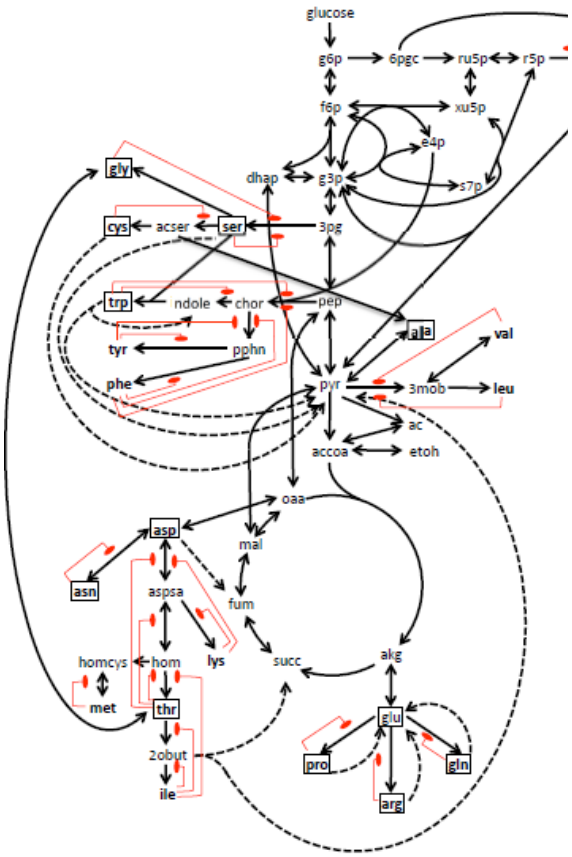
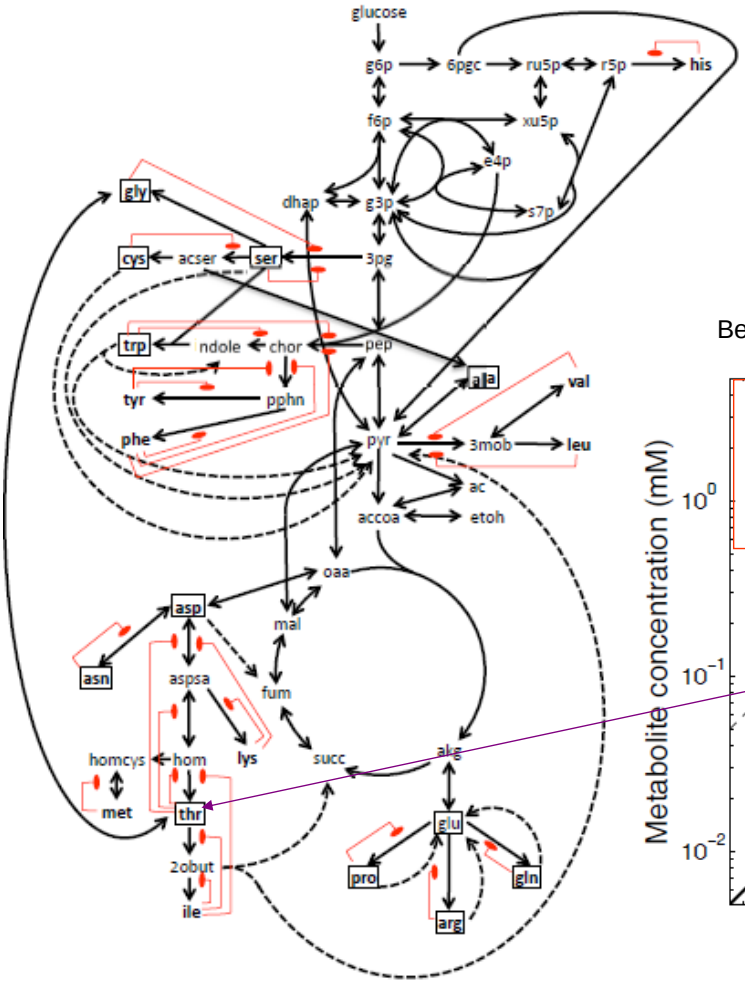


Table 4. The TUs and genes for the loop backward pathways of amino acids.

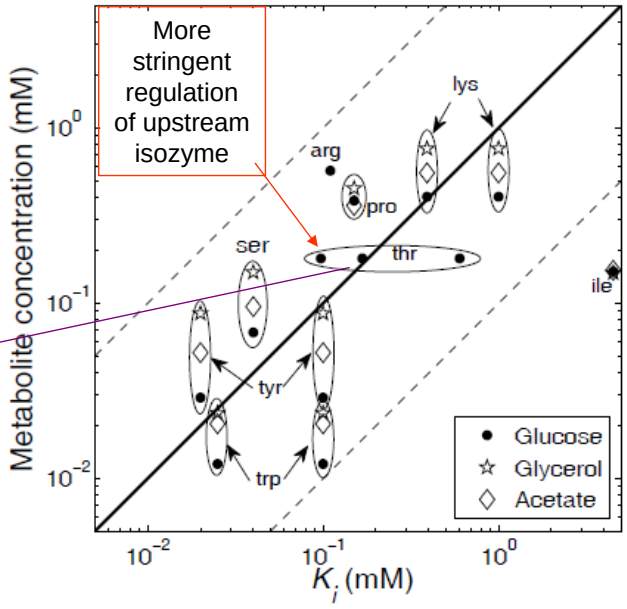
AA	TU	Gene	Reaction	Fold
Trp	tnaCAB	tnaB	h[e] + trp[e] ⟺ h + trp	**2.230**
		tnaA	h2o + trp ⟺ indole + nh4 + pyr	**6.296**
Ser	sdaCB	sdaC	ser[e] + h[e] → ser + h	**2.614**
		sdaB	ser → nh4 + pyr	**1.014**
	sdaA	sdaA		**2.108**
Asp	aspA-dcuA	dcuA	asp[e] + succ → asp + succ[e]	-0.281
	aspA	aspA	asp → fum + nh4	**2.505**
Gln	yneHG	yneH	gln + h2o → glu + nh4	**2.343**
Pro	putA*	putA	fad + pro → 1pyr5c + fadh2 + h 1pyr5c + (2) h2o + nad → glu + h + nadh	**3.448**
Cys	malXY	malY	cys + h2o → h2s + pyr + nh4	0.815
	metC	metC		-0.513
Arg	astCADBE	astA	arg + succoa → coa + h + sucarg	-0.368
		astB	(2) h + (2) h2o + sucarg → co2 + (2) nh4 + sucorn	-0.524
		astC	akg + sucorn → glu + sucgsa	-0.332
		astD	h2o + nad + sucgsa → (2) h + nadh + sucglu	-0.442
		astE	h2o + sucglu → glu + succ	0.829
Thr	tdcABCDEFG	tdcC	thr[e] + h[e] → thr + h	-0.029
		tdcB	thr → 2obut + nh4	0.222
		tdcE	2obut + coa → for + ppcoa	-0.795
	prpBCDE	prpC	h2o + oaa + ppcoa → 2mcit + coa + h	-0.225
		prpD	2mcit → 2mcacn + h2o	-0.121
		prpB	micit ⟺ pyr + succ	-0.678
	acnB	acnB	2mcacn + h2o → micit	-0.620

* There is no data of gene putA in measurement of LB/glucose minimal media comparison. We compare change between L-proline minimal media and glucose minimal media using data from same literature.
The column Fold is ln-2 fold change of gene expression between growth in LB/L-proline minimal media and growth in glucose minimal media [40]. Significant changes of gene expression are highlighted in bold.



End-product inhibition via allosteric regulation ensures appropriate flux ratio at branching points of the carbon flow

Bennett et al. (2009) Nat. Chem. Biol. **5**, 593-9



Regulation of global commodities

Table 3. Evidences for homeostasis of currencies and carriers.

Compounds	Catergory	References
NAD^+/NADH, $NADP^+$/NADPH	redox	[20–22]
SO_3^{2-}, SO_4^{2-}, HS^-, sulfonate	sulfur	[23–25]
NH_4^+, glutamine, glutamate	nitrogen	[26, 27]
Fe^{2+},Fe^{3+}	ions	[28–32]
Pyruvate	small carbon unit	[33, 34]
ATP/ADP/AMP	energy	[35–37]
PO_4^{3-}, pyrophosphate, metaphosphate	phosphorus	[38]

- Regulatory strategies (3-levels)

— Transcriptional regulation to eliminate futile cycles (as well as setting the enzyme copy number to a suitable level), minutes

excess enzymes are kept within 10% of respective population

— Small molecule regulation to achieve fast equilibrium, ms to second.

metabolite concentration measurements show that they are used not only during transition periods to a new mode of growth, but even in steady-state.

— Homeostasis of global commodities

growth rate dependent?

Beyond FBA: growth physiology

Is 30 minutes the physical limit for doubling of any cell?

Where does the bottleneck lie?

Why doesn't the bug use multiple carbon source to speed up growth?

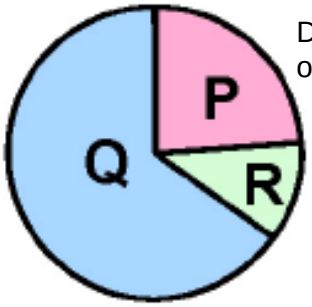
What determines the order of preference for the sugars?

How should one attempt to answer such questions?

Interdependence of cell growth and gene expression: origins and consequences
M Scott, C W Gunderson, E M Mateescu, Z Zhang & Terence Hwa, *Science* **330**:1099-102 (2010)

Protein synthesis = major consumer/expenditure of metabolism

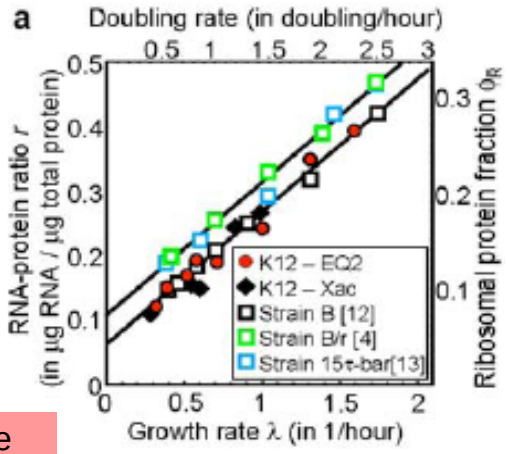
Partitioning of protein content in a bacterial cell



Dispensable overhead

Growth machinery

Non-dispensable overhead

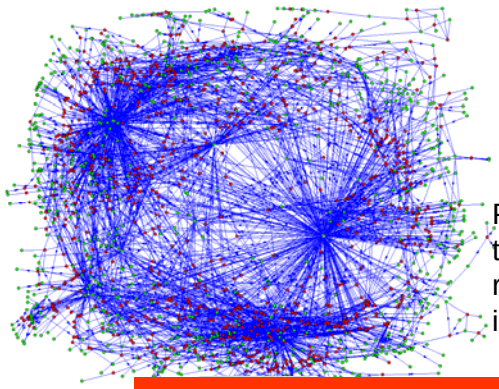


Ribosome makes ribosome
⇒ doubling time 25 min!

Summary

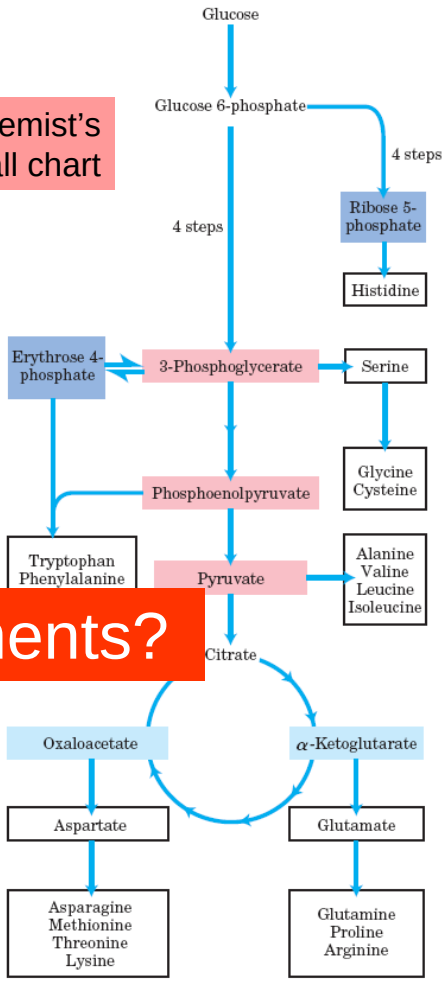
Automated network simplification schemes with biochemical input

Unedited network



Plus tens of thousands of regulatory interactions

Biochemist's wall chart



Evolutionary improvements?

Happily swimming and dividing bacteria

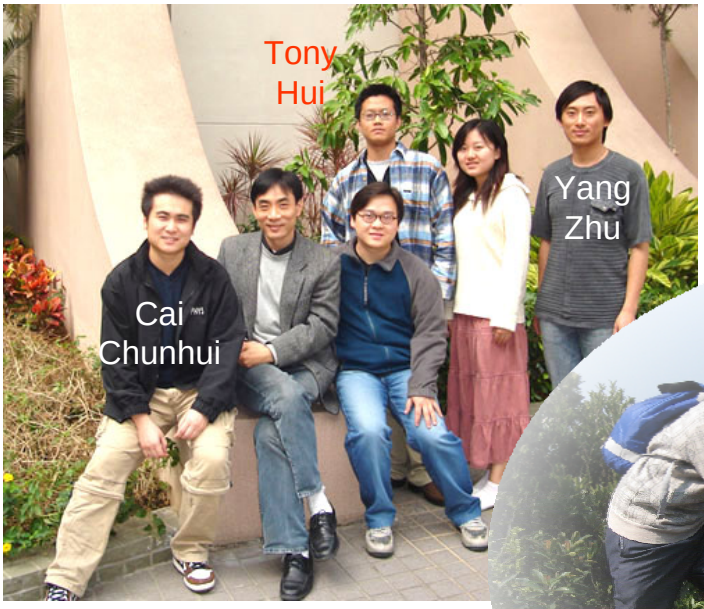
Network topology analysis to identify key variables

Interpretation of regulatory interactions

The Team

Prof. Terry Hwa, UCSD

Dr. Pan-Jun Kim, KAIST
and UIUC



Supported by the RGC
of the HKSAR