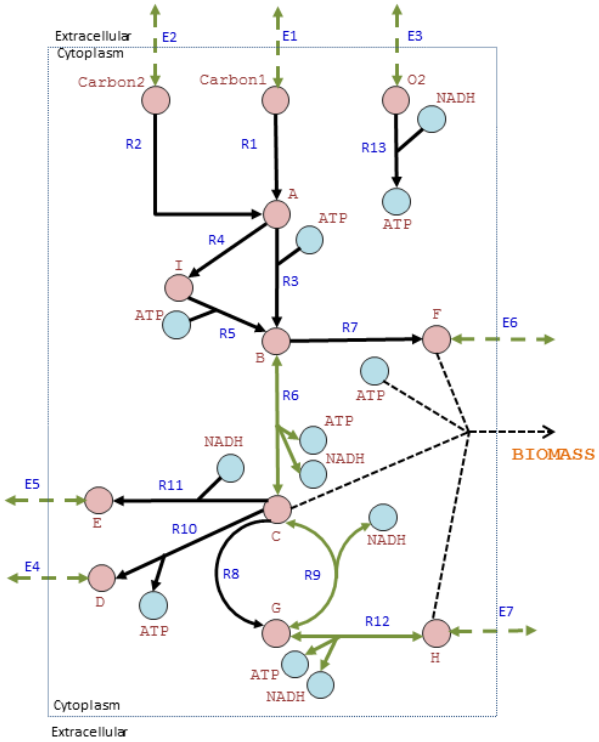
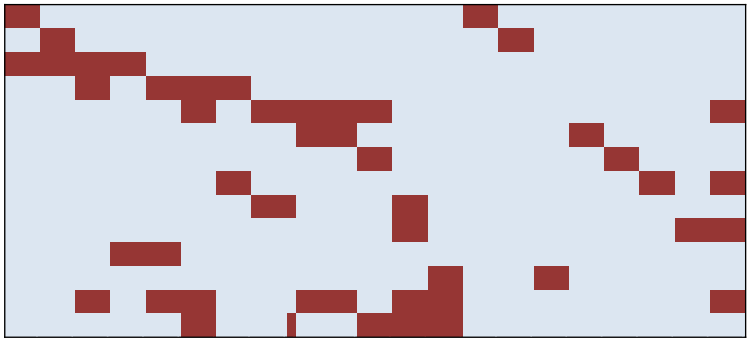
Introduction to Flux Balance Analysis (FBA)





Areejit Samal

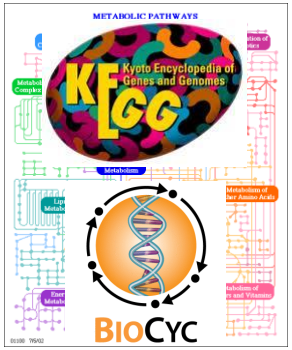
The Institute of Mathematical Sciences (IMSc)

Chennai

Genome-scale metabolic model reconstruction pipeline

Genome sequence with annotation

Biochemical Knowledge

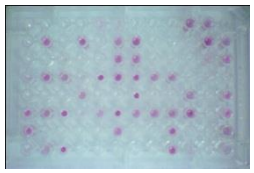


BRENDA

swissprot

TransportDB

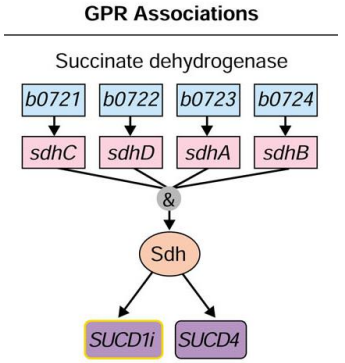
Physiology



Genome-scale reconstructed metabolic network

List of metabolic reactions that can occur in an organism along with gene-protein -reaction (GPR association)

Abbreviatio	OfficialName	Equation (note [c] and [e] at the beginning refer to th	Subsystem	ProteinClass	Description
ALATA_L	L-alanine transaminase	[c]akg + ala-L <==> glu-L + pyr	Alanine and aspartate metaboli	EC-2.6.1.2	
ALAR	alanine racemase	[c]ala-L <==> ala-D	Alanine and aspartate metaboli	EC-5.1.1.1	
ASNN	L-asparaginase	[c]asn-L + h2o --> asp-L + nh4	Alanine and aspartate metaboli	EC-3.5.1.1	
ASNS2	asparagine synthetase	[c]asp-L + atp + nh4 --> amp + asn-L + h + ppi	Alanine and aspartate metaboli	EC-6.3.1.1	
ASNS1	asparagine synthase (glutamine-hydrolysing)	[c]asp-L + atp + gln-L + h2o --> amp + asn-L + glu-L + h	Alanine and aspartate metaboli	EC-6.3.5.4	
ASPT	L-aspartase	[c]asp-L --> fum + nh4	Alanine and aspartate metaboli	EC-4.3.1.1	
ASPTA	aspartate transaminase	[c]akg + asp-L <==> glu-L + oaa	Alanine and aspartate metaboli	EC-2.6.1.1	
VPAMT	Valine-pyruvate aminotransferase	[c]3mob + ala-L --> pyr + val-L	Alanine and aspartate metaboli	EC-2.6.1.66	
DAAD	D-Amino acid dehydrogenase	[c]ala-D + fad + h2o --> fadh2 + nh4 + pyr	Alanine and aspartate metaboli	EC-1.4.99.1	
ALARi	alanine racemase (irreversible)	[c]ala-L --> ala-D	Alanine and aspartate metaboli	EC-5.1.1.1	
FFSD	beta-fructofuranosidase	[c]h2o + suc6p --> fru + g6p	Alternate Carbon Metabolism	EC-3.2.1.26	
ASPISO	arabinose-5-phosphate isomerase	[c]ru5p-D <==> ara5p	Alternate Carbon Metabolism	EC-5.3.1.13	
MME	methylmalonyl-CoA epimerase	[c]mmcoa-R <==> mmcoa-S	Alternate Carbon Metabolism	EC-5.1.99.1	
MICITD	2-methylisocitrate dehydratase	[c]2mcacn + h2o --> micit	Alternate Carbon Metabolism	EC-4.2.1.99	
ALCD19	alcohol dehydrogenase (glycerol)	[c]glyald + h + nadh <==> glyc + nad	Alternate Carbon Metabolism	EC-1.1.1.1	
LCADi	lactaldehyde dehydrogenase	[c]h2o + lald-L + nad --> (2) h + lac-L + nadh	Alternate Carbon Metabolism	EC-1.2.1.22	
TGBPA	Tagatose-bisphosphate aldolase	[c]tagdp-D <==> dhap + g3p	Alternate Carbon Metabolism	EC-4.1.2.40	
LCAD	lactaldehyde dehydrogenase	[c]h2o + lald-L + nad <==> (2) h + lac-L + nadh	Alternate Carbon Metabolism	EC-1.2.1.22	
ALDD2x	aldehyde dehydrogenase (acetaldehyde, NAD)	[c]acald + h2o + nad --> ac + (2) h + nadh	Alternate Carbon Metabolism	EC-1.2.1.3	
ARAI	L-arabinose isomerase	[c]arab-L <==> rbl-L	Alternate Carbon Metabolism	EC-5.3.1.4	
RBK_L1	L-ribulokinase (L-ribulose)	[c]atp + rbl-L --> adp + h + ru5p-L	Alternate Carbon Metabolism	EC-2.7.1.16	
RBP4E	L-ribulose-phosphate 4-epimerase	[c]ru5p-L <==> xu5p-D	Alternate Carbon Metabolism	EC-5.1.3.4	
ACACCT	acetyl-CoA:acetoacetyl-CoA transferase	[c]acac + accoa --> aacoa + ac	Alternate Carbon Metabolism		
BUTCT	Acetyl-CoA:butyrate-CoA transferase	[c]accoa + but --> ac + btcoa	Alternate Carbon Metabolism	EC-2.8.3.8	
AB6PGH	Arbutin 6-phosphate glucohydrolase	[c]arbt6p + h2o --> g6p + hqn	Alternate Carbon Metabolism	EC-3.2.1.86	
PMANM	phosphomannomutase	[c]man1p <==> man6p	Alternate Carbon Metabolism	EC-5.4.2.8	
PPM2	phosphopentomutase 2 (deoxyribose)	[c]2dr1p <==> 2dr5p	Alternate Carbon Metabolism	EC-5.4.2.7	



Genome-scale reconstructed metabolic model

Exchange and transport reactions defining possible nutrients

ACALDt	acetaldehyde reversible transport	acald[e] <==> acald[c]	Transport, Extracellular
GUAt	Guanine transport	gua[e] <==> gua[c]	Transport, Extracellular
HYXNt	Hypoxanthine transport	hxan[e] <==> hxan[c]	Transport, Extracellular
XANt	xanthine reversible transport	xan[e] <==> xan[c]	Transport, Extracellular
NACUP	Nicotinic acid uptake	nac[e] --> nac[c]	Transport, Extracellular
ASNabc	L-asparagine transport via ABC system	asn-L[e] + atp[c] + h2o[c] --> adp[c] + asn-L[c] + h[c] + p	Transport, Extracellular
ASNt2r	L-asparagine reversible transport via proton sympo	asn-L[e] + h[e] <==> asn-L[c] + h[c]	Transport, Extracellular
DAPabc	M-diaminopimelic acid ABC transport	26dap-M[e] + atp[c] + h2o[c] --> 26dap-M[c] + adp[c] + h	Transport, Extracellular
CYSabc	L-cysteine transport via ABC system	atp[c] + cys-L[e] + h2o[c] --> adp[c] + cys-L[c] + h[c] + p	Transport, Extracellular
ACt2r	acetate reversible transport via proton symport	ac[e] + h[e] <==> ac[c] + h[c]	Transport, Extracellular
ETOHt2r	ethanol reversible transport via proton symport	etoh[e] + h[e] <==> etoh[c] + h[c]	Transport, Extracellular
PYRt2r	pyruvate reversible transport via proton symport	h[e] + pyr[e] <==> h[c] + pyr[c]	Transport, Extracellular
O2t	o2 transport (diffusion)	o2[e] <==> o2[c]	Transport, Extracellular
CO2t	CO2 transporter via diffusion	co2[e] <==> co2[c]	Transport, Extracellular
H2Ot	H2O transport via diffusion	h2o[e] <==> h2o[c]	Transport, Extracellular
DHAt	Dihydroxyacetone transport via facilitated diffusior	dha[e] <==> dha[c]	Transport, Extracellular
NH3t	ammonia reversible transport	nh4[e] <==> nh4[c]	Transport, Extracellular
ARBt2r	L-arabinose transport via proton symport	arab-L[e] + h[e] <==> arab-L[c] + h[c]	Transport, Extracellular

Cellular Objective: Biomass Reaction

(0.05) 5mthf + (5.0E-5) accoa + (0.488) ala-L + (0.0010) amp + (0.281) arg-L + (0.229) asn-L + (0.229) asp-L + (45.7318) atp + (1.29E-4) clpn_EC + (6.0E-6) coa + (0.126) ctp + (0.087) cys-L + (0.0247) datp + (0.0254) dctp + (0.0254) dgtp + (0.0247) dttp + (1.0E-5) fad + (0.25) gln-L + (0.25) glu-L + (0.582) gly + (0.154) glycogen + (0.203) gtp + (45.5608) h2o + (0.09) his-L + (0.276) ile-L + (0.428) leu-L + (0.0084) lps_EC + (0.326) lys-L + (0.146) met-L + (0.00215) nad + (5.0E-5) nadh + (1.3E-4) nadp + (4.0E-4) nadph + (0.001935) pe_EC + (0.0276) peptido_EC + (4.64E-4) pg_EC + (0.176) phe-L + (0.21) pro-L + (5.2E-5) ps_EC + (0.035) ptrc + (0.205) ser-L + (0.0070) spmd + (3.0E-6) succoa + (0.241) thr-L + (0.054) trp-L + (0.131) tyr-L + (0.0030) udpg + (0.136) utp + (0.402) val-L --> (45.5608) adp + (45.56035) h + (45.5628) pi + (0.7302) ppi

Constraint-based modeling approach

- Currently, lack of kinetic information for most enzymes and reactions in a genome-scale metabolic network prohibits kinetic modeling of metabolism.

- In the absence of adequate kinetic data, the alternate *constraint-based modeling approach* and *Flux Balance Analysis* (FBA) has been widely used to analyze metabolic networks.

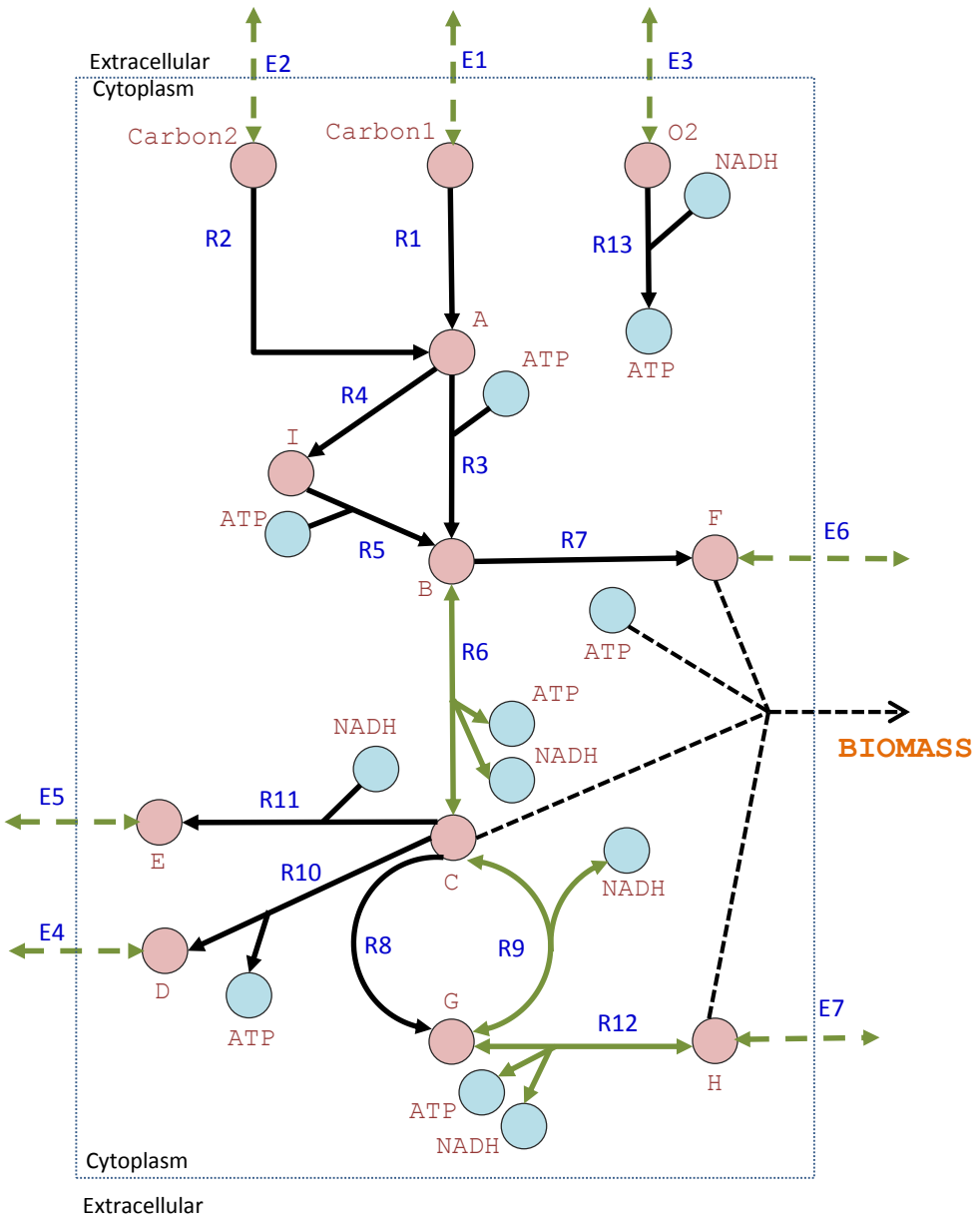
- FBA is a mathematical approach to calculate the flows in a metabolic network, and in certain situations, the method can be used to predict the growth rate of an organism.

- FBA and related approaches can used to investigate the effect of environmental and genetic perturbations on cellular metabolism, and, thus, the method gives quantitative insights into metabolic genotype-phenotype relationships.

- In this lecture, I will describe in detail the conceptual framework associated with FBA.

Genome-scale metabolic models for real organisms contain from hundreds to few thousand reactions. Instead for simplicity, a toy metabolic model will be used in this lecture to describe the FBA framework.

Toy model mimicking core metabolism



S. No.	Abbreviation	Equation
1	R1	Carbon1 -> A
2	R2	Carbon2 -> A
3	R3	A + ATP -> B
4	R4	A -> I
5	R5	I + ATP -> B
6	R6	B -> C + (2) ATP + (2) NADH
7	R7	B -> F
8	R8	C -> G
9	R9	G <-> (0.8) C + (2) NADH
10	R10	C -> (3) D + (2) ATP
11	R11	C + (4) NADH -> (3) E
12	R12	G + ATP + (2) NADH <-> H
13	R13	O2 + NADH -> ATP
14	E1	Carbon1 <->
15	E2	Carbon2 <->
16	E3	O2 <->
17	E4	D <->
18	E5	E <->
19	E6	F <->
20	E7	H <->
21	BIOMASS	C + F + H + (10) ATP -> Biomass

Internal Reactions

Exchange Flux

Network contains: m=14 metabolites and n=21 reactions (13 internal reactions, 7 exchange fluxes and 1 Biomass reaction).

Metabolite Cofactor

Irreversible Reversible

RP Vivek Ananth & A Samal, Biosystems (2016)

Mathematical representation of the network

List of Reactions

S. No.	Abbreviation	Equation
1	R1	Carbon1 -> A
2	R2	Carbon2 -> A
3	R3	A + ATP -> B
4	R4	A -> I
5	R5	I + ATP -> B
6	R6	B -> C + (2) ATP + (2) NADH
7	R7	B -> F
8	R8	C -> G
9	R9	G <-> (0.8) C + (2) NADH
10	R10	C -> (3) D + (2) ATP
11	R11	C + (4) NADH -> (3) E
12	R12	G + ATP + (2) NADH <-> H
13	R13	O2 + NADH -> ATP
14	E1	Carbon1 <->
15	E2	Carbon2 <->
16	E3	O2 <->
17	E4	D <->
18	E5	E <->
19	E6	F <->
20	E7	H <->
21	BIOMASS	C + F + H + (10) ATP -> Biomass

Stoichiometric Matrix S

Internal Reactions | Exchange Flux

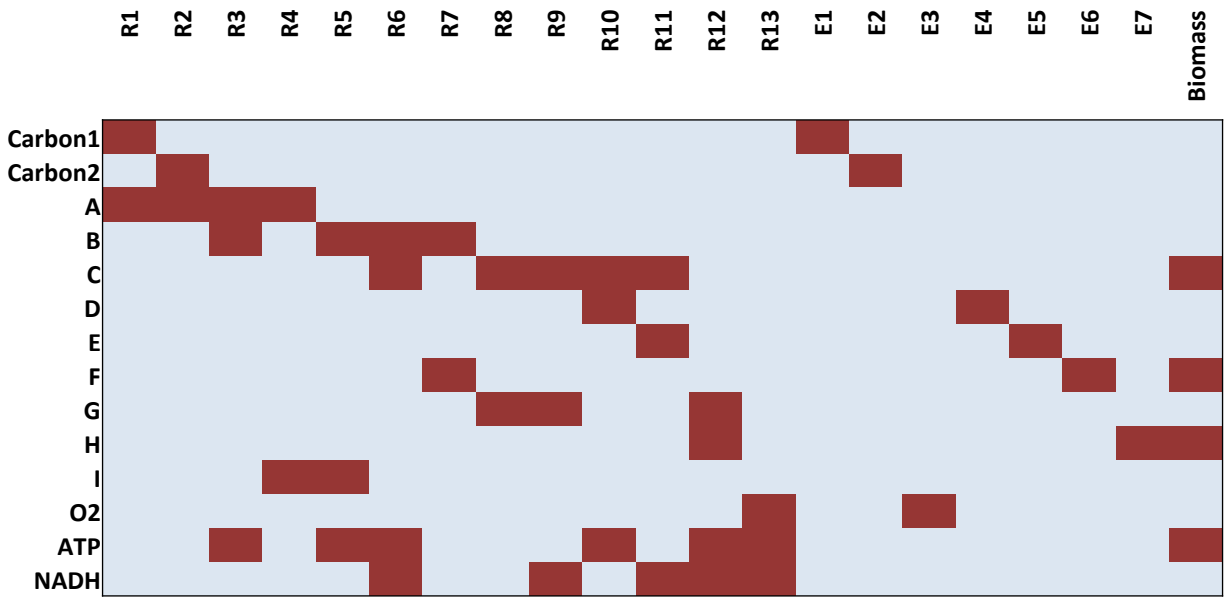
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	E1	E2	E3	E4	E5	E6	E7	Biomass
Carbon1	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
Carbon2	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
A	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	1	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	1	0	-1	0.8	-1	-1	0	0	0	0	0	0	0	0	0	-1
D	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0
F	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	0	-1
G	0	0	0	0	0	0	0	1	-1	0	0	-1	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	-1
I	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-1	0	0	0	0	0
ATP	0	0	-1	0	-1	2	0	0	0	2	0	-2	1	0	0	0	0	0	0	0	-10
NADH	0	0	0	0	0	2	0	0	2	0	-4	-1	-1	0	0	0	0	0	0	0	0

Each row corresponds to a unique metabolite in the network and each column corresponds to a specific reaction.

Negative coefficients represent metabolite is consumed in the reaction and Positive coefficients represent metabolite is produced in the reaction.

The matrix has dimensions m x n where m is the number of metabolites and n is the number of reactions.

Stoichiometric matrix is typically sparse



Nonzero entries of the stoichiometric matrix are colored brown in this visualization.

Stoichiometric constraint

The rate of change of concentration of metabolites is given by:

$$\frac{dX}{dt} = S \bullet v$$

where $\boldsymbol{X}$ is the vector of metabolite concentrations of length equal to number of metabolites, $\boldsymbol{S}$ is the stoichiometric matrix, $\boldsymbol{v}$ is the vector of reaction fluxes of length equal to number of reactions.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	E1	E2	E3	E4	E5	E6	E7	Biomass
Carbon1	**-1**	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
Carbon2	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
A	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	1	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	1	0	-1	0.8	-1	-1	0	0	0	0	0	0	0	0	0	-1
D	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0
F	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	0	-1
G	0	0	0	0	0	0	0	1	-1	0	0	-1	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	-1
I	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-1	0	0	0	0	0
ATP	0	0	-1	0	-1	2	0	0	0	2	0	-2	1	0	0	0	0	0	0	0	-10
NADH	0	0	0	0	0	2	0	0	2	0	-4	-1	-1	0	0	0	0	0	0	0	0

$\boldsymbol{S}$

v
V_{R1}
V_{R2}
V_{R3}
V_{R4}
V_{R5}
V_{R6}
V_{R7}
V_{R8}
V_{R9}
V_{R10}
V_{R11}
V_{R12}
V_{R13}
V_{E1}
V_{E2}
V_{E3}
V_{E4}
V_{E5}
V_{E6}
V_{E7}
$V_{BIOMASS}$

$\boldsymbol{v}$

In any metabolic steady state, the total amount of any metabolite being produced is equal to the total amount consumed.
At steady state, we have:

$$\frac{dX}{dt} = S \bullet v = 0$$

This is the stoichiometric mass-balance constraint.

Stoichiometric constraint leads to a system of linear equations

The mass-balance constraint coming from the stoichiometric matrix gives us a system of linear equations relating various reaction fluxes.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	E1	E2	E3	E4	E5	E6	E7	Biomass
Carbon1	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
Carbon2	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
A	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	1	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	1	0	-1	0.8	-1	-1	0	0	0	0	0	0	0	0	0	-1
D	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0
F	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	0	-1
G	0	0	0	0	0	0	0	1	-1	0	0	-1	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	-1
I	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-1	0	0	0	0	0
ATP	0	0	-1	0	-1	2	0	0	0	2	0	-2	1	0	0	0	0	0	0	0	-10
NADH	0	0	0	0	0	2	0	0	2	0	-4	-1	-1	0	0	0	0	0	0	0	0

S

$\cdot$

v

$= 0$

Each metabolite or row in the stoichiometric matrix gives one linear equation relating fluxes.
The third row corresponding to metabolite A gives the linear equation:

$$V_{R1} + V_{R2} - V_{R3} - V_{R4} = 0$$

Balances: Stoichiometric constraint

The mass-balance constraint coming from stoichiometric matrix gives one linear equation relating various fluxes for each metabolite in the network. Thus, we get 14 linear equations (corresponding to m=14 metabolites) in the network which are:

$$-V_{R1} - V_{E1} = 0$$
$$-V_{R2} - V_{E2} = 0$$
$$V_{R1} + V_{R2} - V_{R3} - V_{R4} = 0$$
$$V_{R3} + V_{R5} - V_{R6} - V_{R7} = 0$$
$$V_{R6} - V_{R8} + 0.8\,V_{R9} - V_{R10} - V_{R11} - V_{BIOMASS} = 0$$
$$3\,V_{R10} - V_{E4} = 0$$
$$3\,V_{R11} - V_{E5} = 0$$
$$V_{R7} - V_{E6} - V_{BIOMASS} = 0$$
$$V_{R8} - V_{R9} - V_{R12} = 0$$
$$V_{R13} - V_{E7} - V_{BIOMASS} = 0$$
$$V_{R4} - V_{R5} = 0$$
$$-V_{R13} - V_{E3} = 0$$
$$-V_{R3} - V_{R5} + 2\,V_{R6} + 2\,V_{R10} - 2\,V_{R12} + V_{R13} - 10\,V_{BIOMASS} = 0$$
$$2\,V_{R6} + 2\,V_{R9} - 4\,V_{R11} - V_{R12} - V_{R13} = 0$$

These 14 linear equations relate 21 reaction fluxes.

Thus, the number of equations m (rows in $\boldsymbol{S}$ or metabolites) is less than the number of unknowns n (columns in $\boldsymbol{S}$ or reactions) which is true for any genome-scale metabolic network.

This system is underdetermined and the set of equations gives a range of possible solutions.

Any flux vector $\boldsymbol{v}$ satisfying the above equations is said to be in the null space of $\boldsymbol{S}$.

Further constraints: Bounds

Thermodynamic constraints

Thermodynamic constraints may render certain reactions to be (practically) irreversible. Since irreversible reactions proceed only in forward direction, their flux is bounded from one side. Thus, for an irreversible reaction j: $V_j \geq 0$.

Enzyme capacity constraints

Based on enzyme capacity measurements, specific upper bounds can be set on the flux of reactions in the network. Thus, for a reaction j, the enzyme capacity constraint gives: $V_j \leq V_{max,j}$

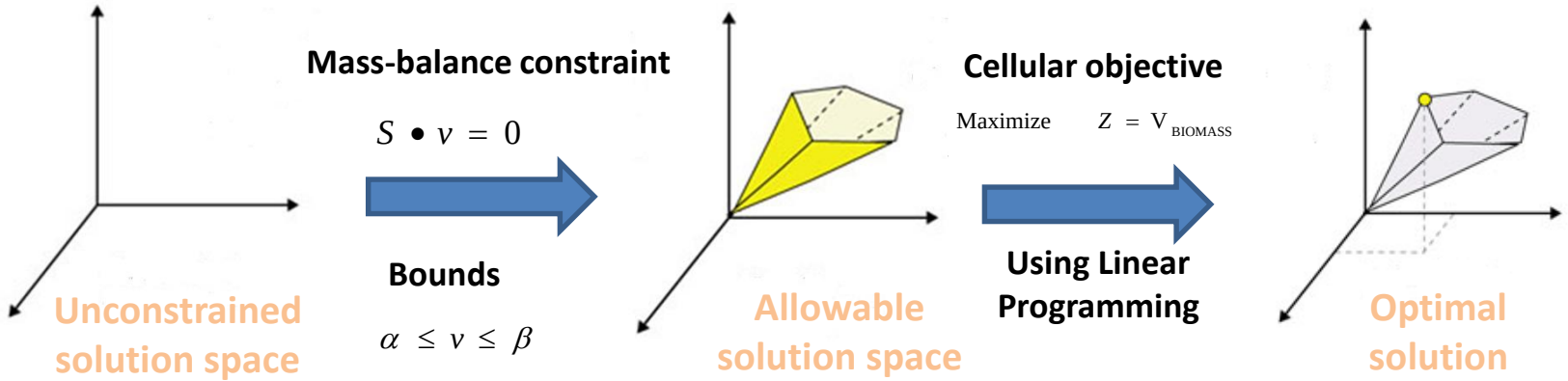
Environmental constraints

Based on nutrient metabolites available in the external environment, bounds are set on the exchange fluxes.
If a nutrient metabolite is not available in the environment, lower bound (LB) of the exchange flux is set to 0, the upper bound (UB) of the exchange flux is left unconstrained for excretion.

Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	-1000
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	-1000	1000
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-10	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	-1000	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

These bound constraints limit the size of the solution space.

Flux balance analysis (FBA) uses linear programming to determine a particular solution



To determine a particular flux distribution ***v*** in the allowable solution space, linear programming is used along with a cellular objective (reflecting evolutionary pressure). In principle, the objective function can be any linear combination of reaction fluxes ***v***.

Usually an 'artificial biomass reaction' consuming biomass precursors in experimentally measured proportions to simulate biomass production is included in the model. The biomass reaction is scaled such that the flux through it equals the exponential growth rate μ of the organism.

The Linear Programming (LP) formulation of the FBA problem is:

Solve

$$S.v = 0$$

$$\alpha \leq v \leq \beta$$

to obtain v that

maximizes $Z = V_{BIOMASS}$

Nature Biotechnology **12**, 994 - 998 (1994)
doi:10.1038/nbt1094-994

Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use

Amit Varma[2] & Bernhard O. Palsson[1,*]

Watson MR, Metabolic maps for the Apple II. 12, 1093-1094 (1984)

Biochem. J. (1986) **238**, 781–786 (Printed in Great Britain)

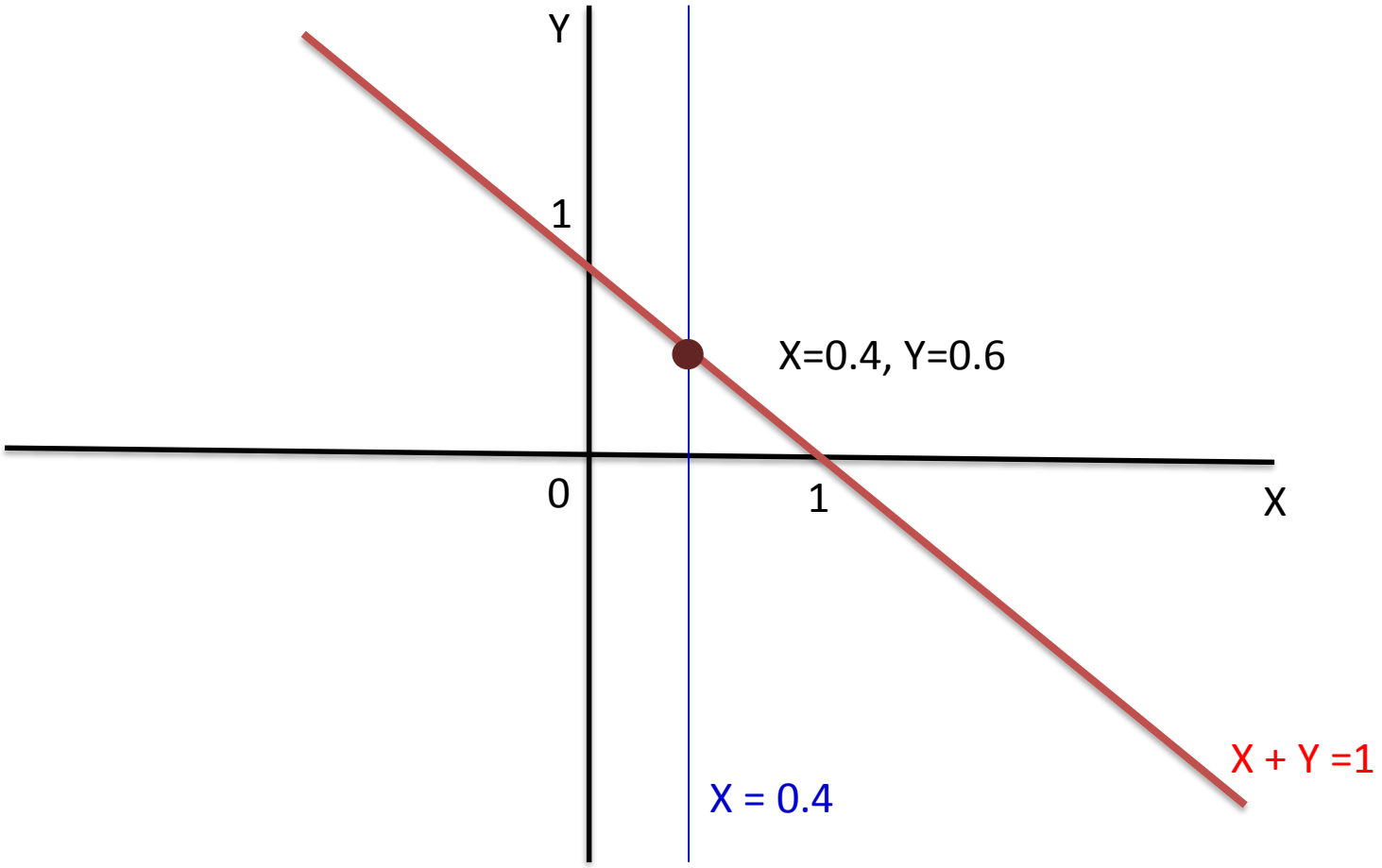
Fat synthesis in adipose tissue

An examination of stoichiometric constraints

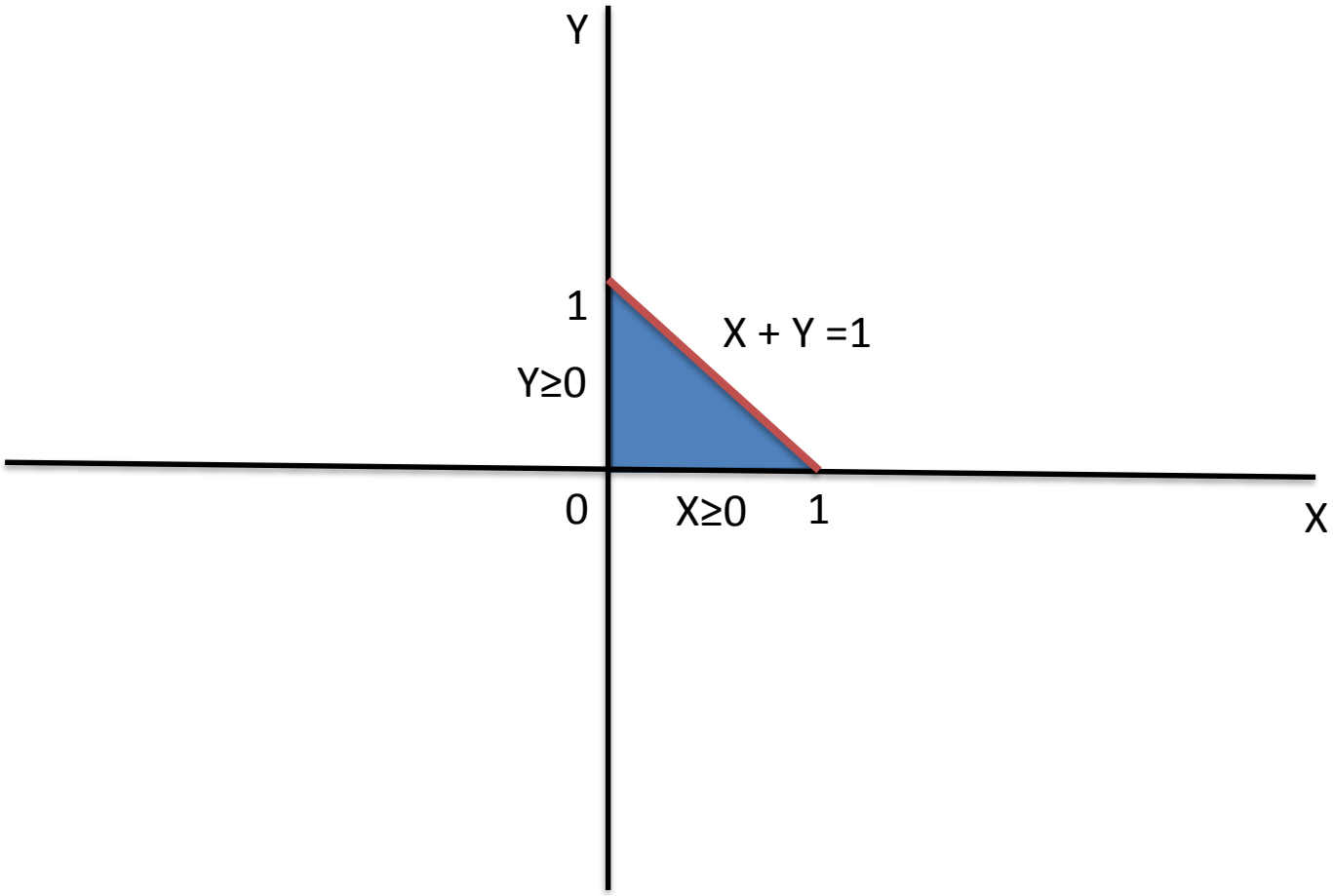
David A. FELL and J. Rankin SMALL

Department of Biology, Oxford Polytechnic, Headington, Oxford OX3 OBP, U.K.

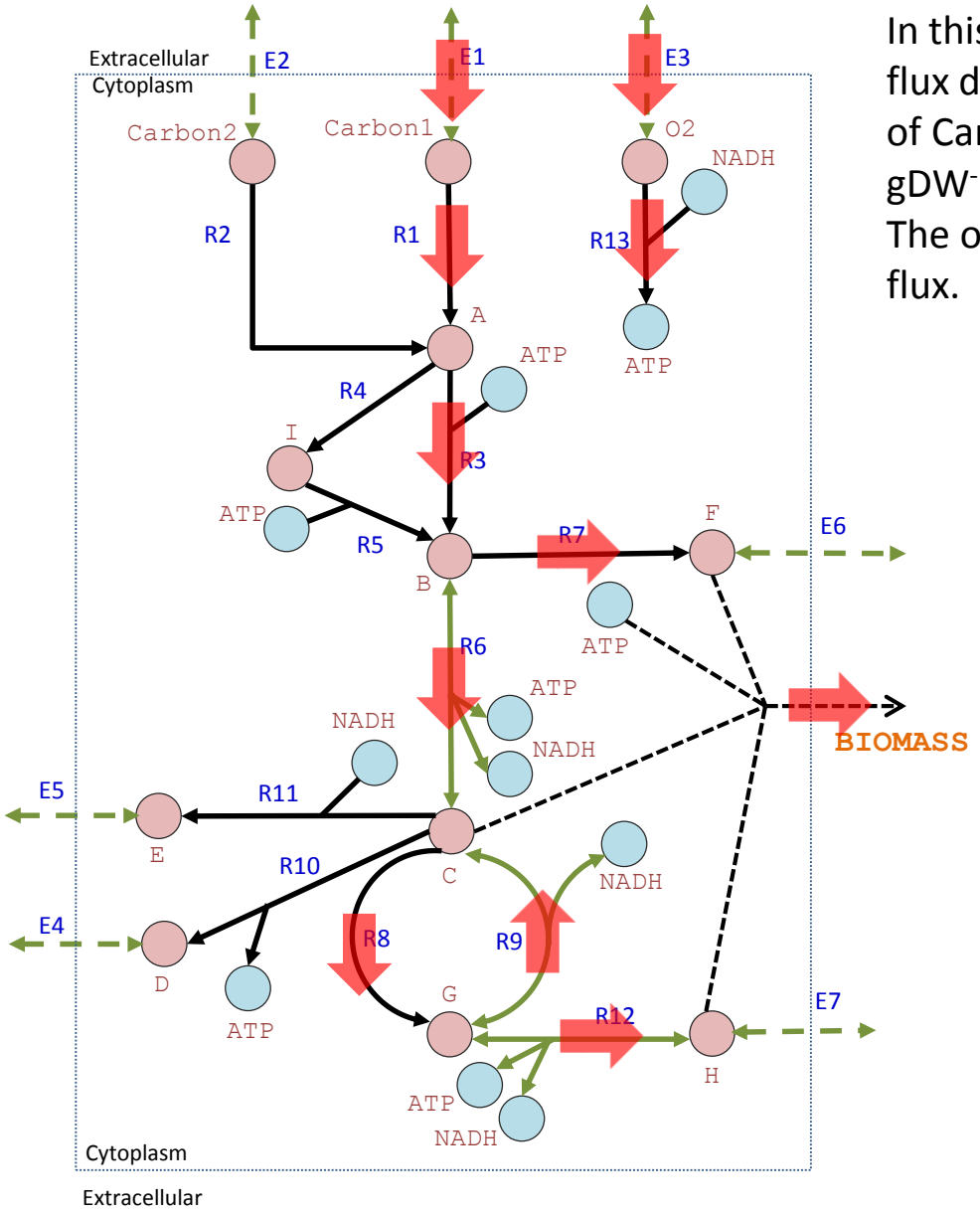
Simple Example: Balances



Simple Example: Balances and Bounds



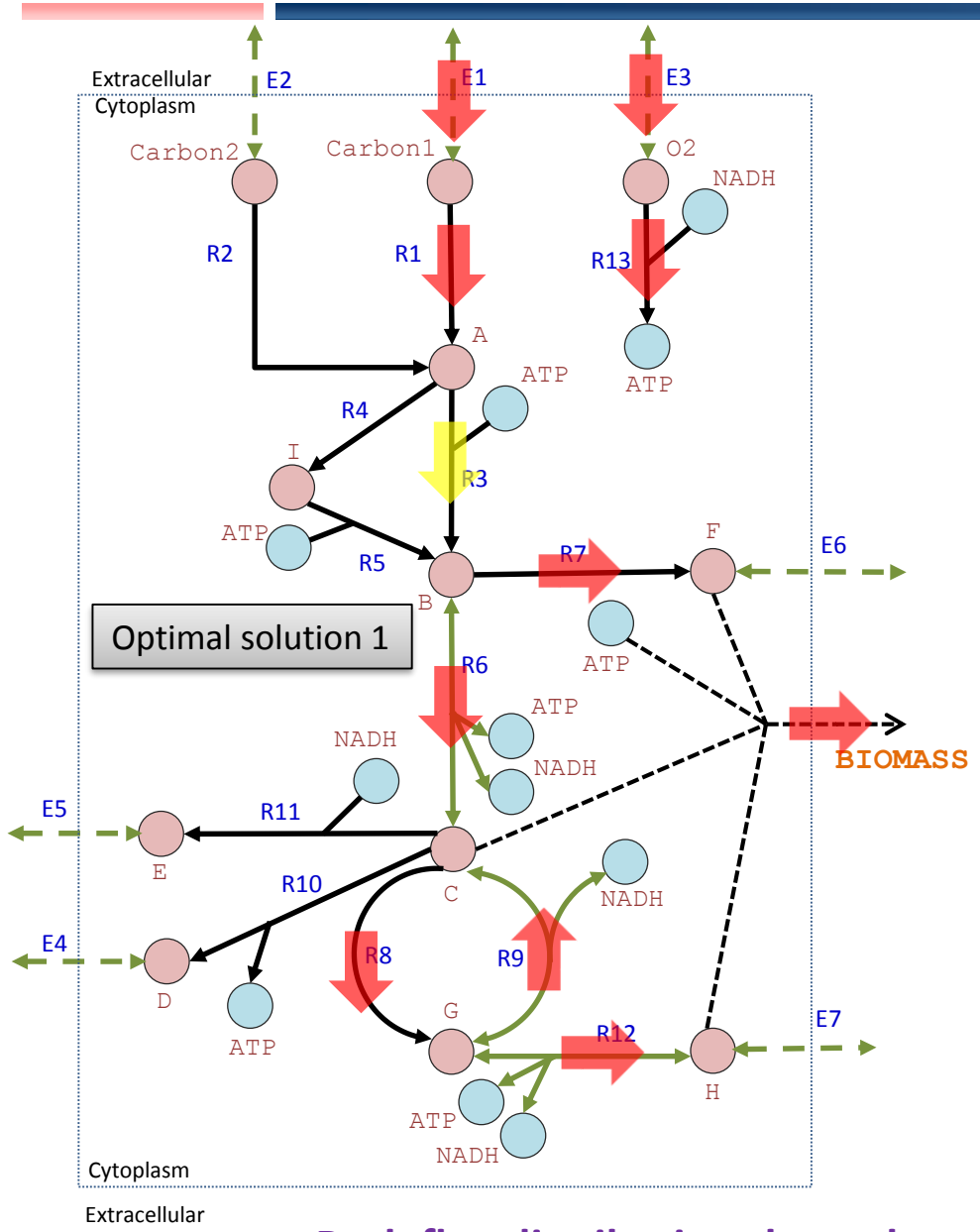
Example 1: Growth in carbon-limited aerobic condition

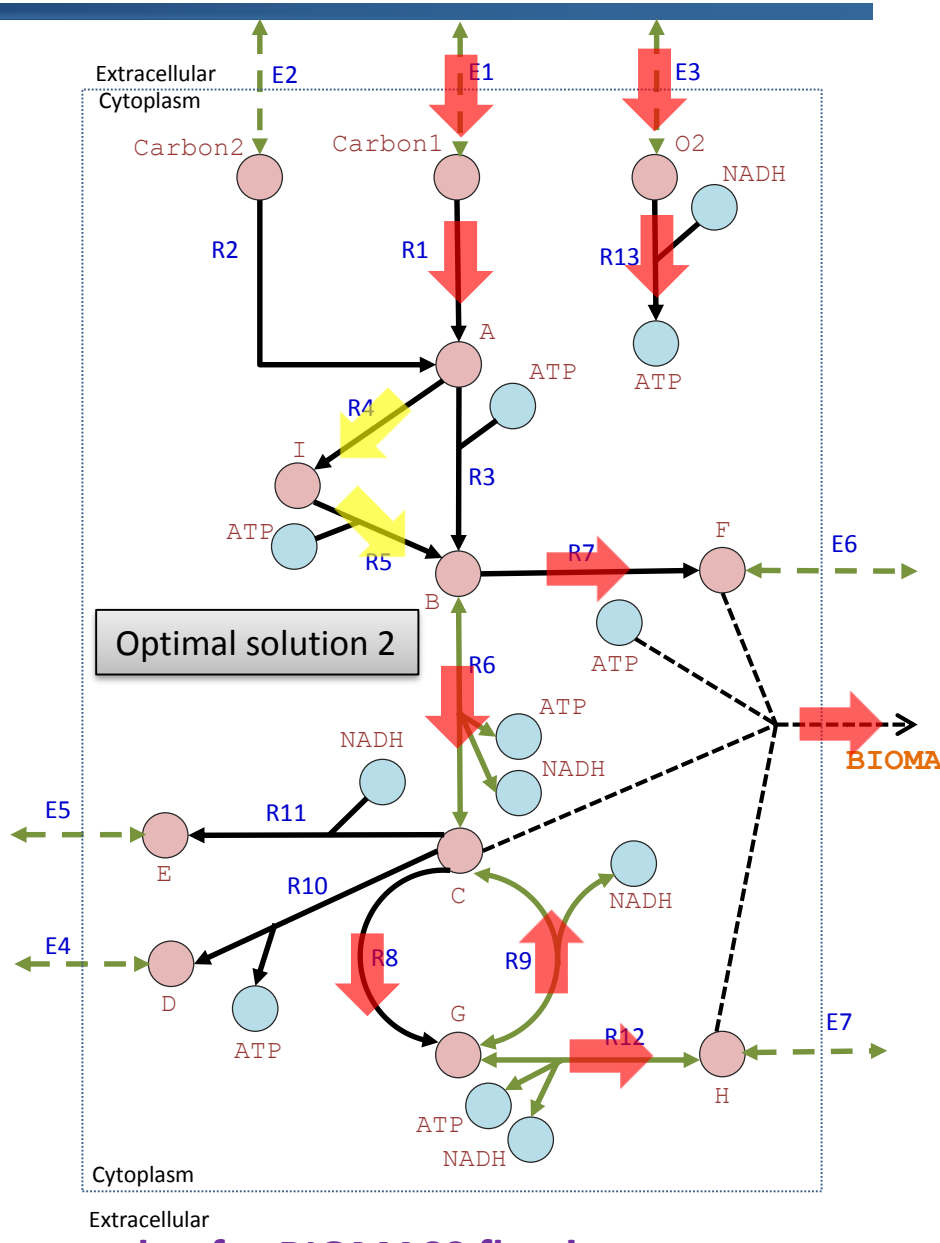


In this example, FBA is used to determine a optimal flux distribution in an environment where the uptake of Carbon1 is constrained to be less than 10 mmol gDW^{-1} h^{-1} while Oxygen (O2) uptake is unconstrained. The objective function is maximization of BIOMASS flux.

Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	1000
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	-1000	1000
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-10	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	-1000	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

Alternate optimal solutions exist due to redundancy in metabolic networks





Both flux distributions have the same value for BIOMASS flux !

Determining alternate optimal flux distributions

Mixed Integer Linear Programming

ELSEVIER
Computers and Chemical Engineering 24 (2000) 711–716
Computers & Chemical Engineering
www.elsevier.com/locate/compchemeng

Recursive MILP model for finding all the alternate optima in LP models for metabolic networks

Sangbum Lee, Chan Phalakornkule, Michael M. Domach, Ignacio E. Grossmann *

A **mixed integer linear programming (MILP)** based algorithm which can be used to sample alternate optimal flux distributions which have the same value of the objective function (i.e., Biomass flux). However, the number of alternate optima in most genome-scale metabolic networks are typically very large and it is computational infeasible to enumerate all optimal solutions.

Flux Variability Analysis

Although, it is difficult to sample all alternate optimal flux distributions in genome-scale metabolic networks**, Flux Variability Analysis (FVA)** allows one to compute the set of reactions (in the network) whose fluxes vary across different alternate optima.

Metab Eng. 2003 Oct;5(4):264-76.

The effects of alternate optimal solutions in constraint-based genome-scale metabolic models.

Mahadevan R, Schilling CH.
Genomatica, Inc., Bioprocessing Division, 5405 Morehouse Drive, Suite 210, San Diego, CA 92121, USA. rmahadevan@genomatica.com

The LP formulation of the FVA problem is as follows:

Step 1: Use FBA to determine the maximum biomass flux for the given environment ($=V_{optimal,BIOMASS}$).

Step 2: For each reaction j, solve two LP problems: Maximize (and Minimize) V_j such that

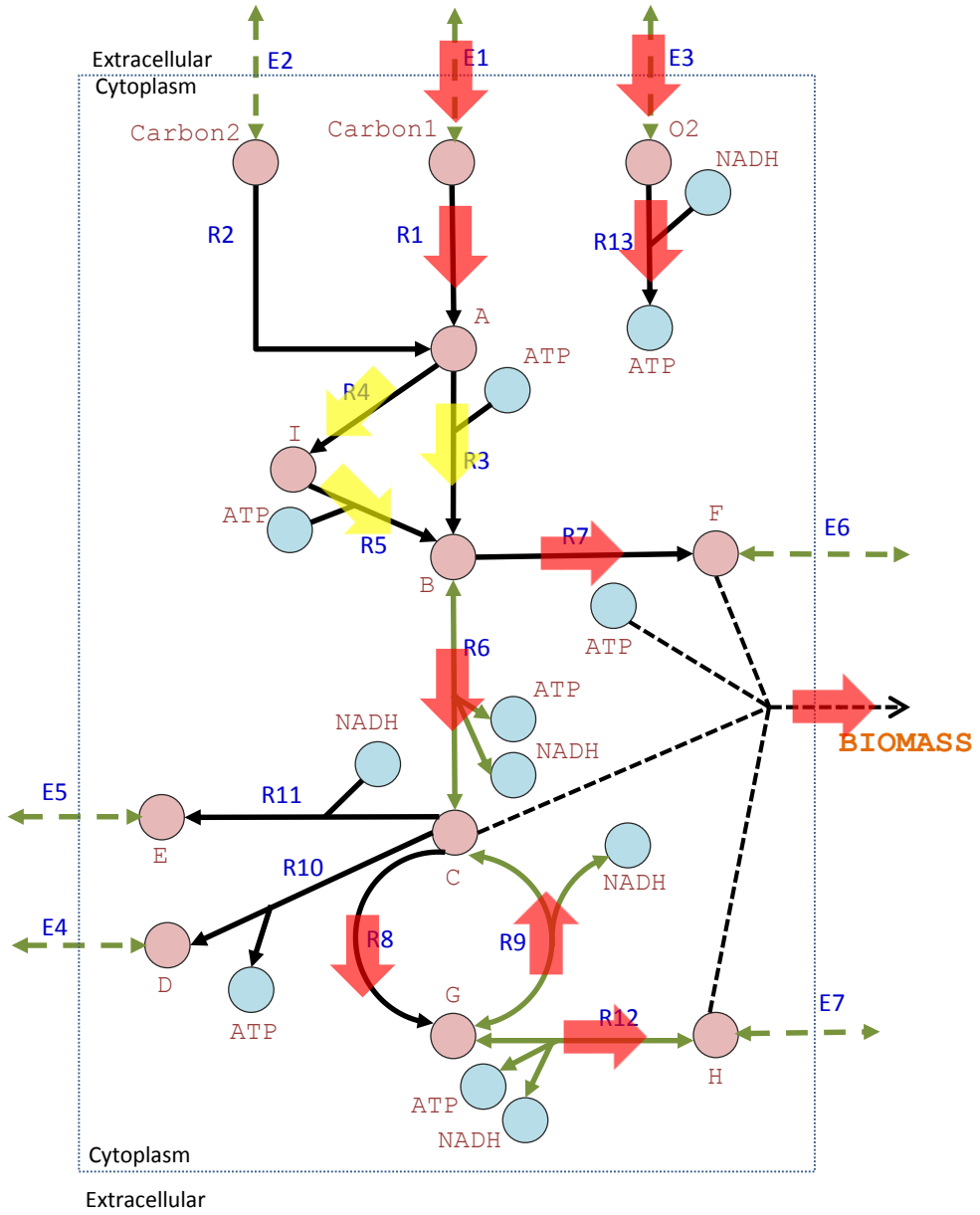
$$S.v = 0$$

$$\alpha \leq v \leq \beta$$

$$V_{BIOMASS} = V_{optimal,BIOMASS}$$

The difference between the maximum and minimum flux for each reaction j determines its variability across alternate optima.

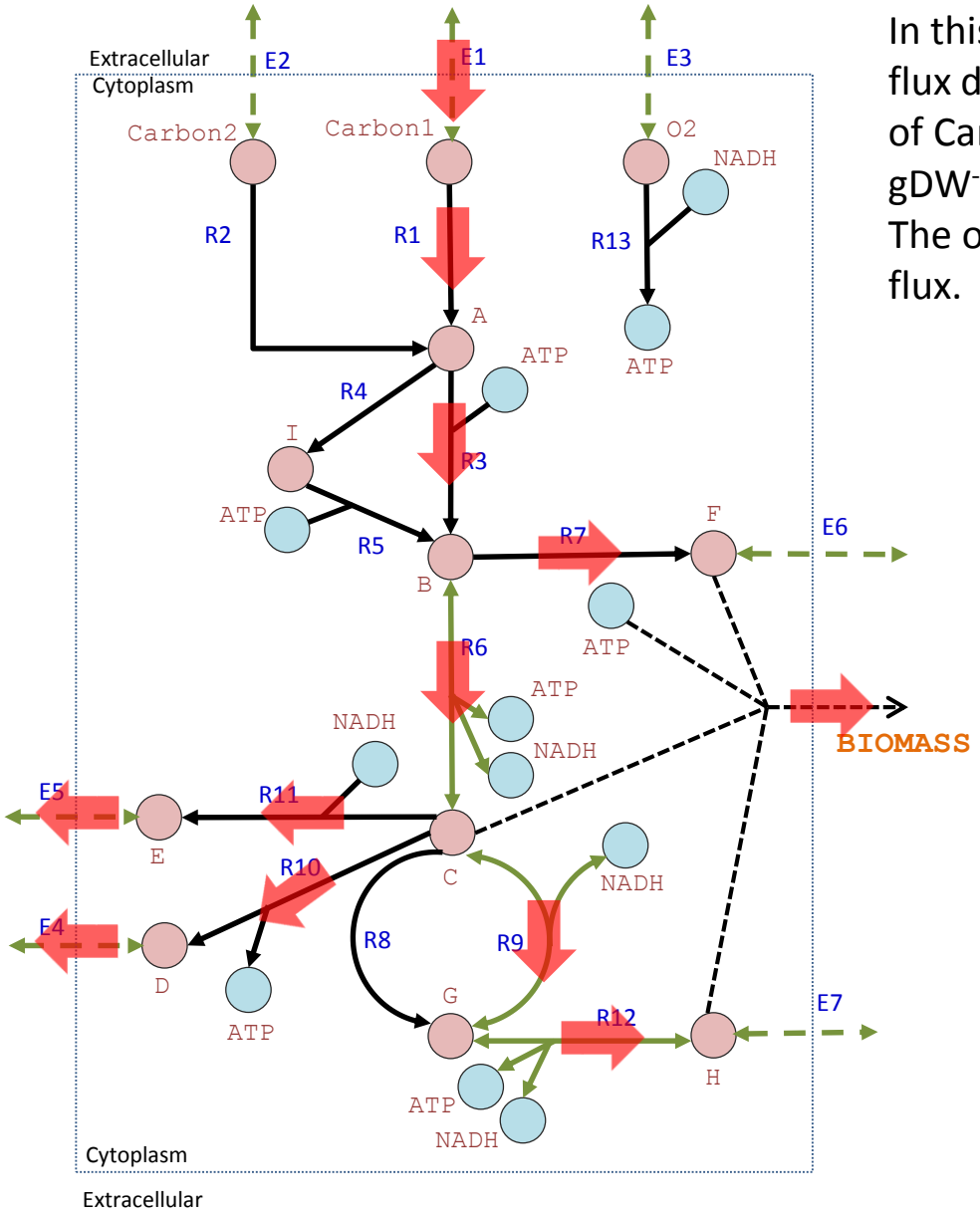
Flux Variability in carbon-limited aerobic condition



In this example, Flux Variability Analysis (FVA) is used to determine the variability of each internal reaction flux across different optimal flux distributions in an environment where the uptake of Carbon1 is constrained to be less than 10 mmol gDW^{-1} h^{-1} while Oxygen (O2) uptake is unconstrained.

The reactions shown with yellow arrows have variable flux across different alternate optima.

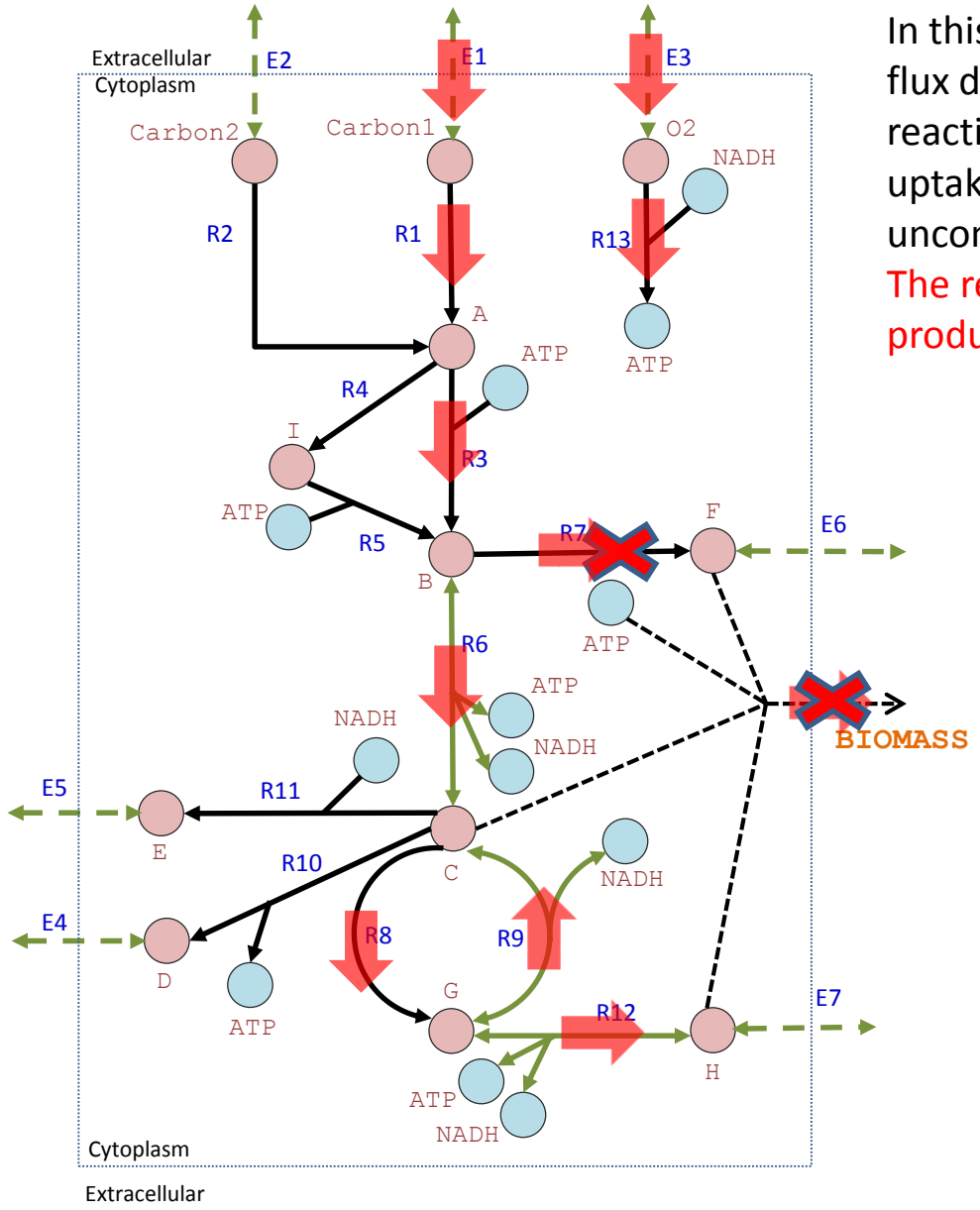
Example 2: Growth in carbon-limited anaerobic condition



In this example, we use FBA to determine the optimal flux distribution in an environment where the uptake of Carbon1 is constrained to be less than 10 mmol gDW^{-1} h^{-1} while Oxygen (O2) is not available.
The objective function is maximization of BIOMASS flux.

Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	1000
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	-1000	1000
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-10	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	0	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

Example 3: Genetic perturbation under carbon-limited aerobic condition

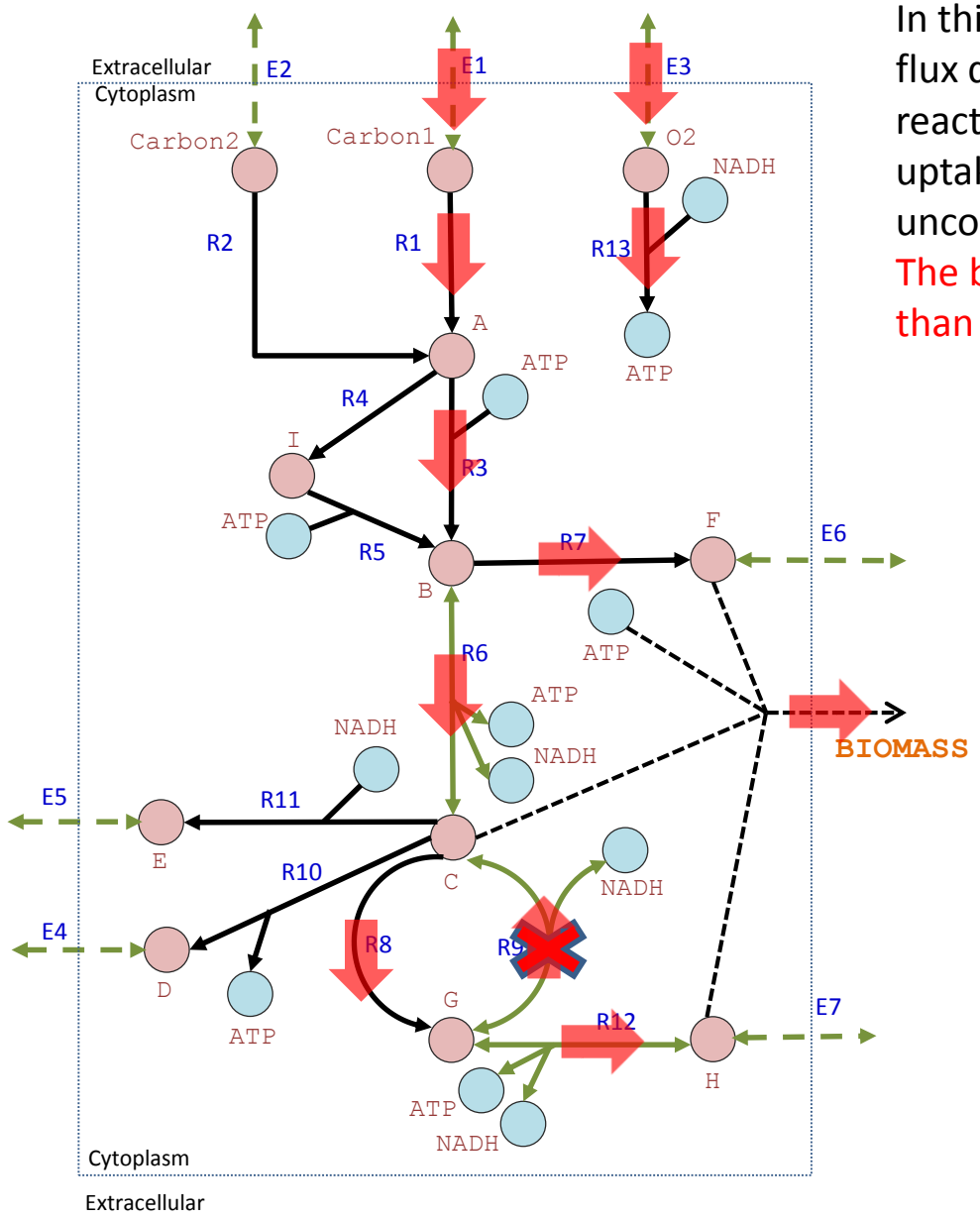


In this example, FBA is used to determine the optimal flux distribution after the enzyme associated with reaction R7 is knocked out in an environment where uptake of Carbon1 is constrained while Oxygen (O2) is unconstrained.

The reaction R7 is found to be essential for biomass production.

Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	0
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	-1000	1000
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-10	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	-1000	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

Example 4: Genetic perturbation under carbon-limited aerobic conditions

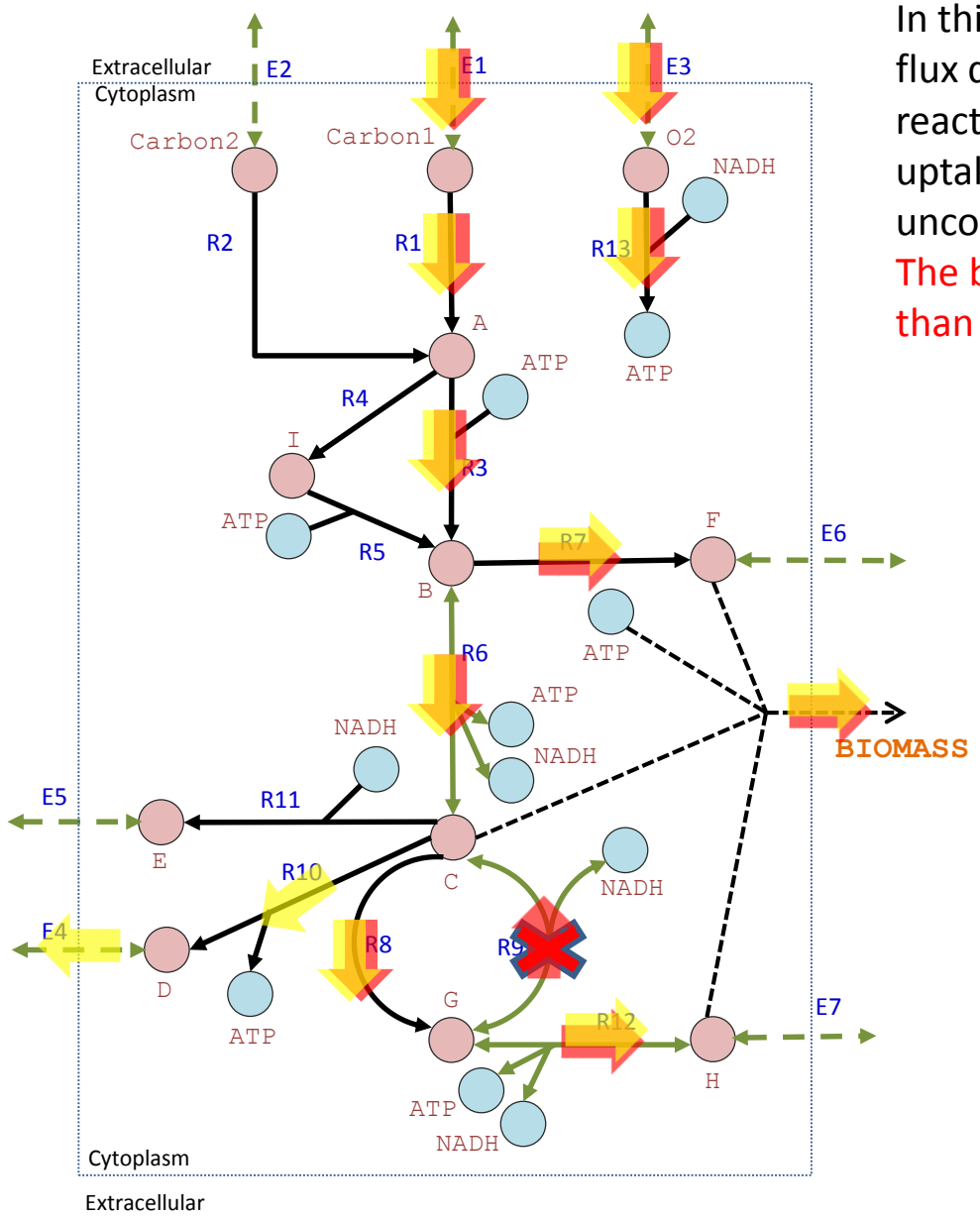


In this example, FBA is used o determine the optimal flux distribution after the enzyme associated with reaction R9 is knocked out in an environment where uptake of Carbon1 is constrained while Oxygen (O2) is unconstrained.

The biomass flux of the perturbed network is less than that in the wild type network.

Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	-1000
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	0	0
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-10	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	-1000	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

Example 4: Genetic perturbation under carbon-limited aerobic conditions



In this example, FBA is used o determine the optimal flux distribution after the enzyme associated with reaction R9 is knocked out in an environment where uptake of Carbon1 is constrained while Oxygen (O2) is unconstrained.

The biomass flux of the perturbed network is less than that in the wild type network.

Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	-1000
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	0	0
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-10	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	-1000	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

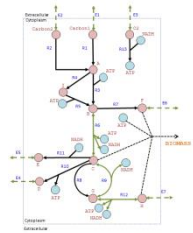
Flux Balance Analysis (FBA) framework

Genome-scale metabolic model

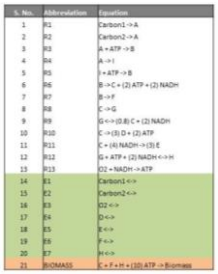
Stoichiometric matrix

Allowable Solution Space

Optimal Solution

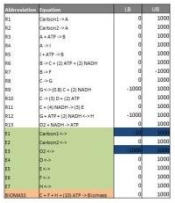


	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	E1	E2	E3	E4	E5	E6	E7	Biomass
Carbon1	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
Carbon2	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
A	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	1	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	1	0	-1	0.8	-1	-1	0	0	0	0	0	0	0	0	0	-1
D	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	-1	0	0	0
F	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	0	-1
G	0	0	0	0	0	0	0	1	-1	0	0	-1	0	0	0	0	0	0	0	-1	0
H	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	-1	-1
I	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-1	0	0	0	0	0
ATP	0	0	-1	0	-1	2	0	0	0	2	0	-2	1	0	0	0	0	0	0	0	-10
NADH	0	0	0	0	0	2	0	0	2	0	-4	-1	-1	0	0	0	0	0	0	0	0



S. No.	Abbreviation	Equation
1	R1	Carbon1 -> A
2	R2	Carbon2 -> A
3	R3	A + ATP -> B
4	R4	A -> I
5	R5	I + ATP -> B
6	R6	B -> C + (2) ATP + (2) NADH
7	R7	B -> F
8	R8	C -> G
9	R9	G <-> (0.8) C + (2) NADH
10	R10	C -> (3) D + (2) ATP
11	R11	C + (4) NADH -> (3) E
12	R12	G + ATP + (2) NADH <-> H
13	R13	O2 + NADH -> ATP
14	E1	Carbon1 <->
15	E2	Carbon2 <->
16	E3	O2 <->
17	E4	D <->
18	E5	E <->
19	E6	F <->
20	E7	H <->
21	BIOMASS	C + F + H + (10) ATP -> Biomass

Other constraints



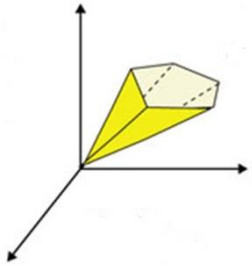
Abbreviation	Equation	LB	UB
R1	Carbon1 -> A	0	1000
R2	Carbon2 -> A	0	1000
R3	A + ATP -> B	0	1000
R4	A -> I	0	1000
R5	I + ATP -> B	0	1000
R6	B -> C + (2) ATP + (2) NADH	0	1000
R7	B -> F	0	-1000
R8	C -> G	0	1000
R9	G <-> (0.8) C + (2) NADH	-1000	1000
R10	C -> (3) D + (2) ATP	0	1000
R11	C + (4) NADH -> (3) E	0	1000
R12	G + ATP + (2) NADH <-> H	-1000	1000
R13	O2 + NADH -> ATP	0	1000
E1	Carbon1 <->	-1000	1000
E2	Carbon2 <->	0	1000
E3	O2 <->	-1000	1000
E4	D <->	0	1000
E5	E <->	0	1000
E6	F <->	0	1000
E7	H <->	0	1000
BIOMASS	C + F + H + (10) ATP -> Biomass	0	1000

Balance

$$S \bullet v = 0$$

$$\alpha \leq v \leq \beta$$

Bounds

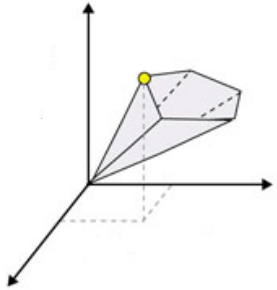


Cellular objective

Maximize

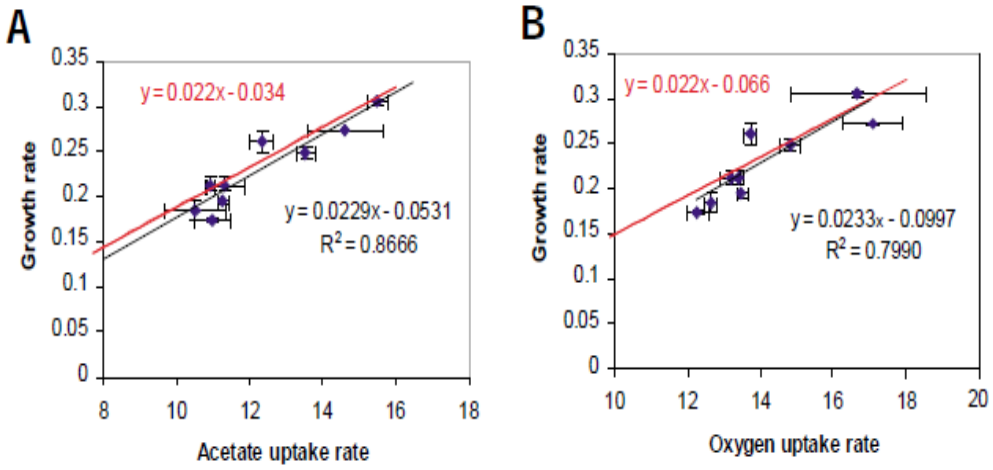
$$Z = V_{BIOMASS}$$

Using Linear Programming



Flux balance analysis (FBA) predictions match with experiments

- Growth rate measurements of microbes such as *E. coli*, especially, in minimal media
- Prediction of essential metabolic genes
- Prediction of synthetic-lethal gene pairs



Nature Biotechnology **19**, 125–130 (1 February 2001) | doi:10.1038/84

In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data

Jeremy S. Edwards , Rafael U. Ibarra & Bernhard O. Palsson

Limitations of Flux Balance Analysis (FBA)

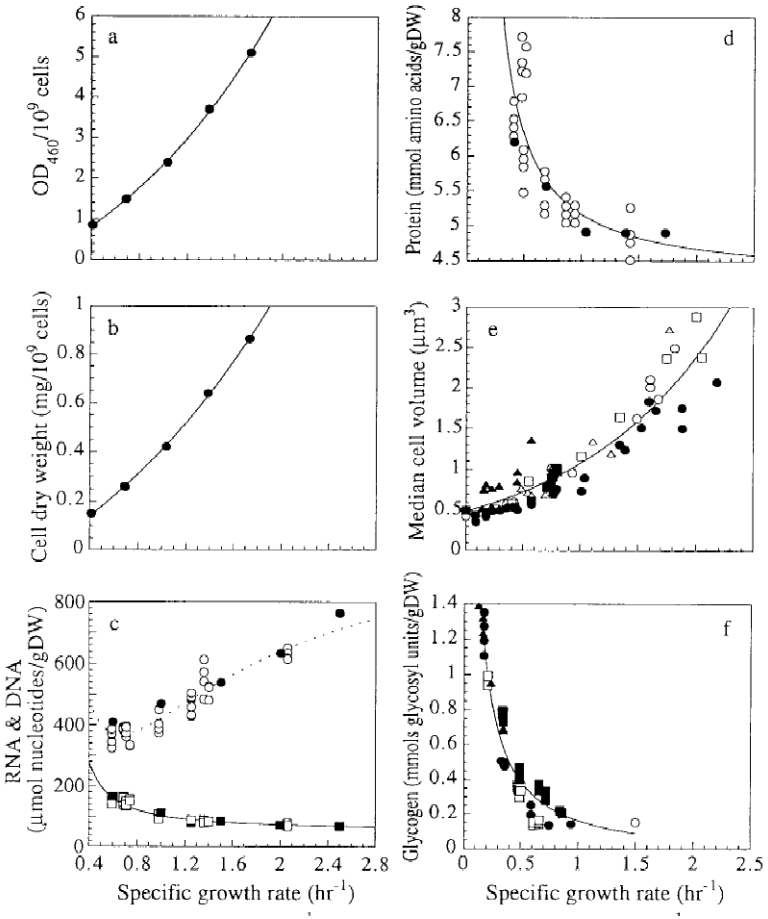
- FBA uses steady-state assumption and cannot predict metabolite concentrations.

  Specialist enzymes were found to carry on average higher flux and their flux was highly variable across different environmental shifts. Specialist enzymes were more regulated in terms of small molecule allosteric inhibition and post-translation modifications.
  In the FBA framework, we cannot model allosteric regulation.

- The simplest FBA models do not account for regulation, and thus, the FBA predictions are not always accurate.

Assumption of fixed biomass composition

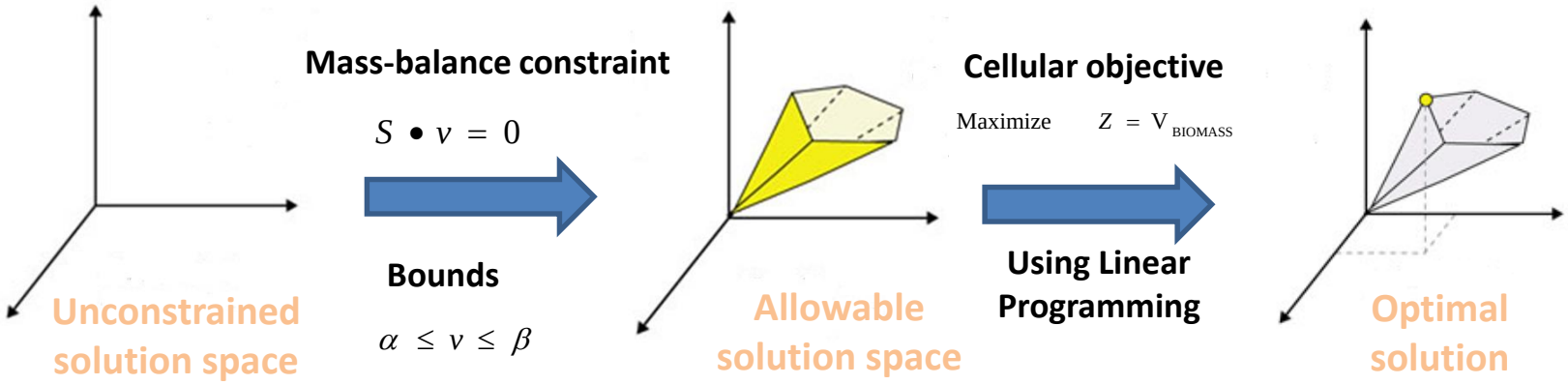
- Experiments have shown conclusively that macromolecular composition of *E. coli* changes with growth rate.
- RNA content increases with growth rate while Protein and DNA content decreases.
- Biomass composition and energy requirements is a function of specific growth rate.
- Most of the current genome-scale metabolic reconstructions impose a fixed biomass composition which is independent of growth rate.



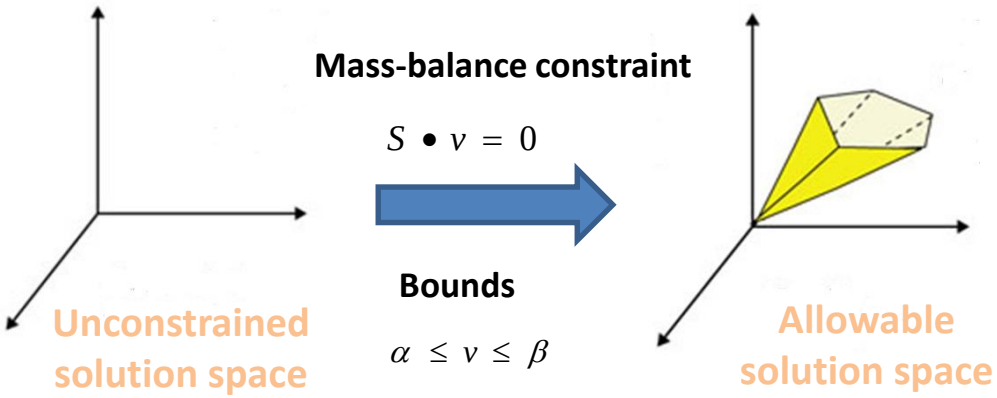
References:
Bremer and Dennis (1987);
Neidhardt (1987)

Other constraint-based metabolic modeling methods

Flux Balance Analysis (FBA) – Optimal metabolic states in the solution space



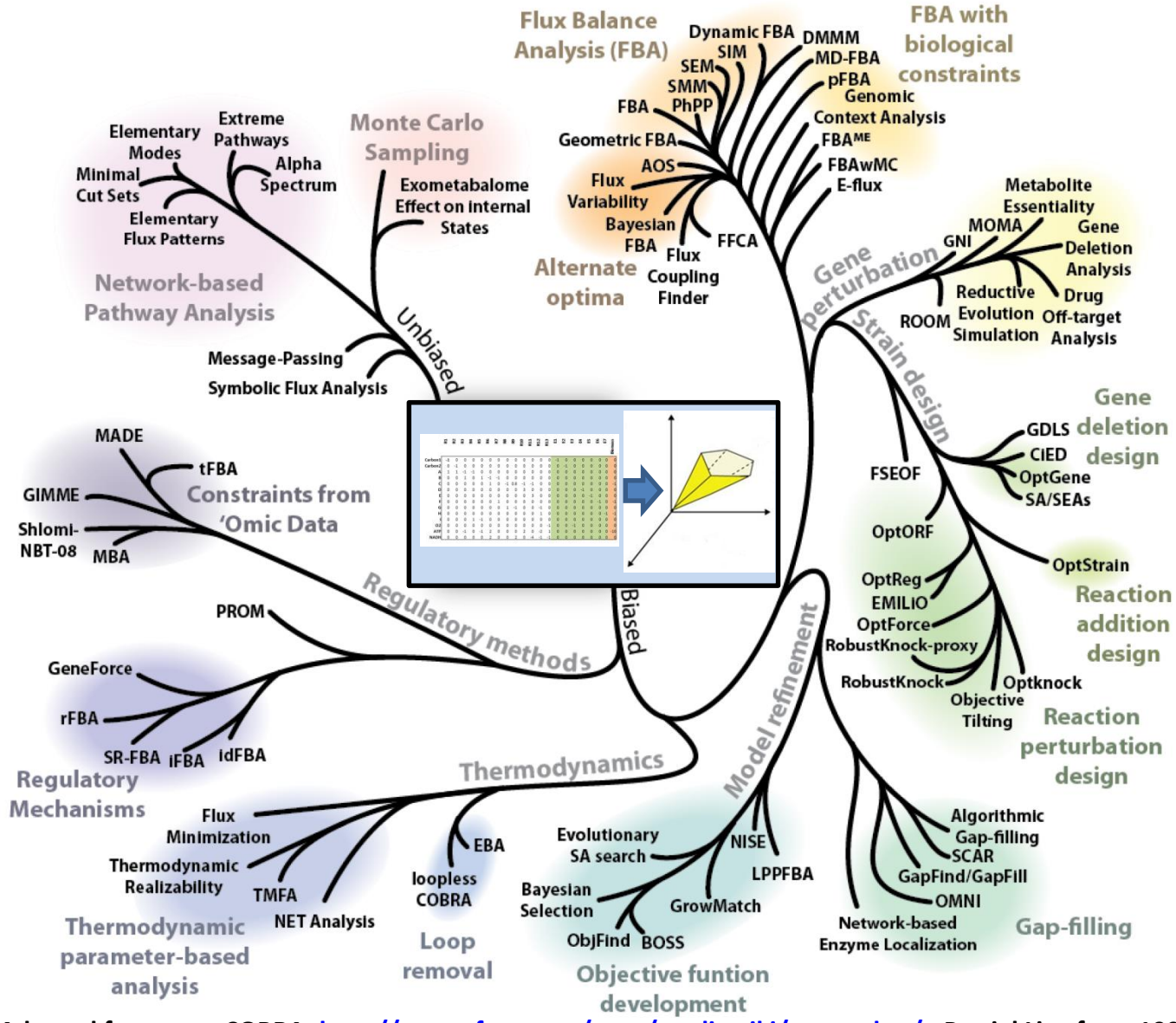
Global characterization of solution spaces



- Network-based pathway definitions
  - Elementary Flux Modes
  - Extreme Pathways
- Uniform (Markov Chain Monte Carlo) sampling of solution space

Any feasible flux vector in the allowable solution space can be written as a linear combination of basis vectors.

The phylogeny of constraint-based modeling methods



- Finding optimal or best states
- Determining Flux dependencies using Linear Programming
- Characterizing the allowable solution space.
- Studying altered solution spaces
- Application of additional constraints (Thermodynamic, regulatory, etc.)

Adapted from openCOBRA: http://sourceforge.net/apps/mediawiki/opencobra/ - Partial List from 100+ COBRA methods proposed till now

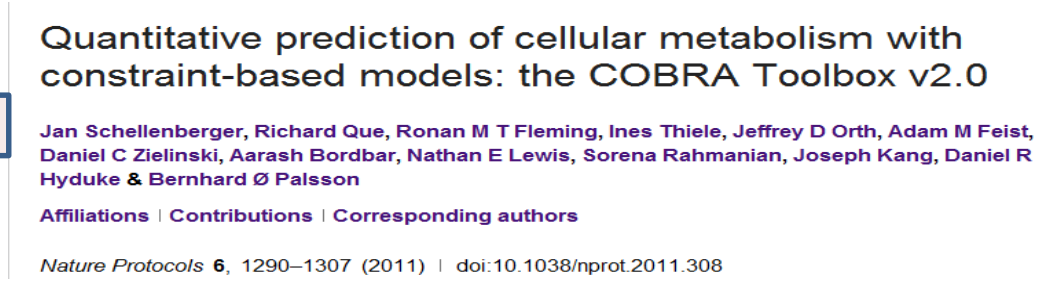
Software tools for metabolic pathway analysis

Software Tools

Software Package	Authors
a-c-o-r-n	Sroka, et al.
anNET	Zamboni, et al.
BioMet Toolbox	Cvijovic, et al.
CellNetAnalyzer	Klamt, et al.
Cobra toolbox	Schellenberger, et al.
CycSim	Le Fevre, et al.
FASIMU	Hoppe, et al.
FBA-SimVis	Grafahrend-Belau, et al.
iMAT	Zur, et al.
Metabolica	Heino, et al.
MetaFluxNet	Lee, et al.
Metatool	Kamp and Schuster
Omix	Droste, et al.
OptFlux	Rocha, et al.
ScurmPy	Poolman, et al.
SNA	Urbanczik
SurreyFBA	Gevorgyan, et al.
The Systems Biology Research Tool	Wright and Wagner
TIGER	Jensen, et al.
WEbcoli	Jung, et al.
Yana	Schwarz, et al.
YanaSquare	Schwarz, et al.

openCOBRA Project

http://opencobra.sourceforge.net/openCOBRA/Welcome.html
http://sourceforge.net/apps/mediawiki/opencobra/index.php?title=Software_Packages



Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0

Jan Schellenberger, Richard Que, Ronan M T Fleming, Ines Thiele, Jeffrey D Orth, Adam M Feist, Daniel C Zielinski, Aarash Bordbar, Nathan E Lewis, Sorena Rahmanian, Joseph Kang, Daniel R Hyduke & Bernhard Ø Palsson

Affiliations | Contributions | Corresponding authors

Nature Protocols **6**, 1290–1307 (2011) | doi:10.1038/nprot.2011.308

Linear Programming Solvers

Solver	Licence
LP_Solve	LGPL
glpk	GPL
CBC	GPL
CLP	GPL
AMPL	Proprietary
CPLEX	Proprietary
GAMS	Proprietary

Free for academic users !!

Suggested References

What is flux balance analysis?

Jeffrey D Orth, Ines Thiele & Bernhard Ø Palsson

Affiliations | Corresponding author

Nature Biotechnology **28**, 245–248 (2010) | doi:10.1038/nbt.1614

Note:
Supplementary material gives code to perform various analysis using COBRA toolbox.

Nature Reviews Microbiology **2**, 886-897 (November 2004) | doi:10.1038/nrmicro1023

Genome-scale models of microbial cells: evaluating the consequences of constraints

Nathan D. Price[1], Jennifer L. Reed[1] & Bernhard Ø. Palsson[1] About the authors

BioSystems 147 (2016) 1–10

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Review article

Advances in the integration of transcriptional regulatory information into genome-scale metabolic models

R.P. Vivek-Ananth, Areejit Samal*

The Institute of Mathematical Sciences (IMSc), Chennai, India