Supplementary file

List of Contents

1. Modified Control Effective Flux (mCEF) algorithm
2. Metabolic networks models

1 Modified Control Effective Flux (mCEF) algorithm

mCEF predicts the gene expression profile of genetically modified mutants where an enzyme coding gene is over-expressed (the relative enzyme activity ratio of a mutant versus wild type is > 1) or under-expressed (0 < enzyme activity ratio of mutant versus wild type < 1). Details of mCEF algorithm are described using a small metabolic network model.

The metabolic network model is shown in Figure S1. R6 is reversible reaction and the others are irreversible reactions. Ax, F and G are the external metabolites and the others are the internal metabolites.

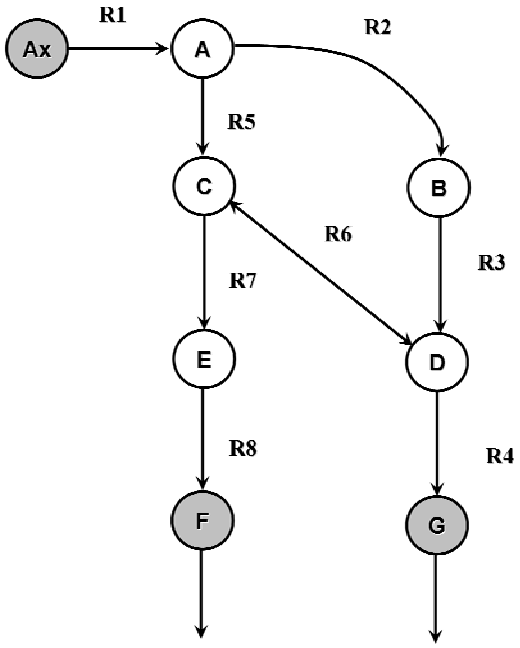

Figure S1. A small metabolic network map. The formation of F (R8: E→F) is the cellular function. Ax, F and G are the external metabolites.

The formation of F from E is regarded as a biological function. In a mutant the gene responsible for the reaction A --> C is 2-fold over-expressed, i.e., the relative enzyme activity parameter EAP of reaction 5 is set to 2.

There are four elementary modes (EMs) for the example network model. The fluxes could be decomposed onto EMs as follows:

$$\mathbf{v} = \mathbf{P} \bullet \boldsymbol{\lambda} = \begin{pmatrix} 1 & 6 & 6 & 3 \\ 0 & 3 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 1 & 0 & 0 & 3 \\ 0 & -2 & 0 & 1 \\ 1 & 6 & 0 & 0 \\ 1 & 6 & 0 & 0 \end{pmatrix} \bullet \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix}$$

where **v** is the flux; **P** is the EM matrix; λ is the EM coefficient (EMC) vector.

The elements in an EM are normalized by the element of the substrate uptake reaction in each EM.

Table S1 EMs normalized by the substrate uptake reaction

Reaction, *i*			Absolute element in the *j*-th EM $p_{i,j}$			
			1	2	3	4
1	R2	2 A → B	½	0	0	½
2	R3	3 B → 4 D	⅙	0	0	⅙
3	R4	2 D → G	0	⅓	0	⅓
4	R5	A → C	0	1	1	0
5	R6	3 C → 2 D	⅓	⅓	0	0
6	R7	C → E	1	0	1	0
7	R8	E → F	1	0	1	0

The efficiency of each EM for the mutant is calculated by:

$$\varepsilon^{m}_{j,CELLOBJ} = \frac{p_{CELLOBJ,j} \cdot EA_j}{\sum_i \left(\left| p_{i,j} \right| \cdot \eta_i \right)}$$

$$\eta_i = \begin{cases} EAP_i & (\text{if reaction } i \text{ is modified}) \\ 1 & (\text{if reaction } i \text{ is not modified}) \end{cases}$$

Table S2 Parameter η_i for each reaction

Reaction, *i*			η_i
1	R2	2 A → B	1
2	R3	3 B → 4 D	1
3	R4	2 D → G	1
4	R5	A → C	2
5	R6	3 C → 2 D	1
6	R7	C → E	1
7	R8	E → F	1

$$EA_j = \prod_{i=1}^{n} ge_{i,j}$$

$$ge_{i,j} = \begin{cases} EAP_i & (\text{if the } i\text{-th reaction is involved in the } j\text{-th EM}) \\ 1 & (\text{if the } i\text{-th reaction is not involved in the } j\text{-th EM}) \end{cases}$$

Table S3 Parameter EA_j for each EM

The *j*-th EM	EA_j
1	1
2	2
3	2
4	1

$$mCEF_i(mut) = \sum_{CELLOBJ} \frac{1}{p^{\max}_{CELLOBJ}} \frac{\sum_j \left(\varepsilon^m_{j,CELLOBJ} \cdot \left| p_{i,j} \right| \cdot \eta_i \right)}{\sum_j \varepsilon^m_{j,CELLOBJ}}$$

$$\varepsilon_{1,7} = \frac{1 \times 1}{\frac{1}{2} \times 1 + \frac{1}{6} \times 1 + 0 \times 1 + 0 \times 2 + \frac{1}{3} \times 1 + 1 \times 1 + 1 \times 1} = \frac{1}{3}$$

$$\varepsilon_{2,7} = \frac{0 \times 2}{0 \times 1 + 0 \times 1 + \frac{1}{3} \times 1 + 1 \times 2 + \frac{1}{3} \times 1 + 0 \times 1 + 0 \times 1} = 0$$

$$\varepsilon_{3,7} = \frac{1 \times 2}{0 \times 1 + 0 \times 1 + 0 \times 1 + 1 \times 2 + 0 \times 1 + 1 \times 1 + 1 \times 1} = \frac{1}{2}$$

$$\varepsilon_{4,7} = \frac{0 \times 1}{\frac{1}{2} \times 1 + \frac{1}{6} \times 1 + \frac{1}{3} \times 1 + 0 \times 2 + 0 \times 1 + 0 \times 1 + 0 \times 1} = 0$$

The mCEF for each reaction is calculated.

$$mCEF_1\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(\frac{1}{2} \times 1 \right) + 0 \cdot \left(0 \times 1 \right) + \frac{1}{2} \cdot \left(0 \times 1 \right) + 0 \cdot \left(\frac{1}{2} \times 1 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = \frac{1}{5}$$

$$mCEF_2\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(\frac{1}{6} \times 1 \right) + 0 \cdot \left(0 \times 1 \right) + \frac{1}{2} \cdot \left(0 \times 1 \right) + 0 \cdot \left(\frac{1}{6} \times 1 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = \frac{1}{15}$$

$$mCEF_3\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(0 \times 1 \right) + 0 \cdot \left(\frac{1}{3} \times 1 \right) + \frac{1}{2} \cdot \left(0 \times 1 \right) + 0 \cdot \left(\frac{1}{3} \times 1 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = 0$$

$$mCEF_4\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(0 \times 2 \right) + 0 \cdot \left(1 \times 2 \right) + \frac{1}{2} \cdot \left(1 \times 2 \right) + 0 \cdot \left(0 \times 2 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = \frac{6}{5}$$

$$mCEF_5\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(\frac{1}{3} \times 1 \right) + 0 \cdot \left(\frac{1}{3} \times 1 \right) + \frac{1}{2} \cdot \left(0 \times 1 \right) + 0 \cdot \left(0 \times 1 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = \frac{2}{15}$$

$$mCEF_6\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(1 \times 1 \right) + 0 \cdot \left(0 \times 1 \right) + \frac{1}{2} \cdot \left(1 \times 1 \right) + 0 \cdot \left(0 \times 1 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = 1$$

$$mCEF_7\left(mut\right) = \frac{1}{1} \cdot \frac{\frac{1}{3} \cdot \left(1 \times 1 \right) + 0 \cdot \left(0 \times 1 \right) + \frac{1}{2} \cdot \left(1 \times 1 \right) + 0 \cdot \left(0 \times 1 \right)}{\left(\frac{1}{3} + 0 + \frac{1}{2} + 0 \right)} = 1$$

The mCEFs of an over-expressing mutant are calculated below.

$$\Theta_i(w, mut) = \frac{mCEF_i(mut)}{mCEF_i(w)}$$

Table S4 Values of mCEFs of an over-expressing mutant

Reaction *i*			$mCEF_i(w)$	$mCEF_i(mut)$	$\Theta_i(w, mut)$
1	R2	2 A → B	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{4}{5}$
2	R3	3 B → 4 D	$\frac{1}{12}$	$\frac{1}{15}$	$\frac{4}{5}$
3	R4	2 D → 3 F	0	0	1
4	R5	A → C	$\frac{1}{2}$	$\frac{6}{5}$	$\frac{12}{5}$
5	R6	3 C → 2 D	$\frac{1}{6}$	$\frac{2}{15}$	$\frac{6}{5}$
6	R7	C → E	1	1	1
7	R8	E → F	1	1	1

2. Metabolic networks models

Table S5 Metabolic network model of *Escherichia coli*

Gene	Gene description	Reaction
ACEA	Isocitrate lyase	ICIT ==> SUCC + GLX
ACEB	Malase synthase A	ACCOA + GLX ==> MAL + COA
ACK	Acetate kinase	ACTP + ADP <==> ATP + AC
ACN	Aconitase	CIT <==> ICIT
ACXT	Acetate transport	AC ==> ACXT
ADK	Adenylate kinase	ATP + AMP <==> 2 ADP
ATP	F0F1-ATPase	PI + 4 HEXT + ADP <==> ATP
ATPDRAIN	ATP drain	ATP ==> PI + ADP
CO_2XT	CO_2 transport	CO_2 <==> CO_2XT
CYO	Cytochrome oxidase bo3	2 QH_2 + O_2 ==> 2 Q + 4 HEXT
EDA	6-phosphogluconate dehydratase	KetoPGluc ==> T3P1 + PYR
EDD	2-keto-3-deoxy-6-phosphogluconate aldolase	D6PGC ==> KetoPGluc
ENO	Enolase	P2G <==> PEP
FBA	Fructose-1,6-bisphosphatase aldolase	FDP <==> T3P1 + T3P2
FBP	Fructose-1,6-bisphosphatase	FDP ==> F6P + PI
FRD	Fumarate reductase	FADH + FUM ==> FAD + SUCC
FUM	Fumarase	FUM <==> MAL
GAP	Glyceraldehyde-3-phosphate dehydrogenase	PI + T3P1 + NAD <==> P13DG + NADH
GLCUP	uptake of glucose	GLCXT ==> GLC
GLT	Citrate synthase	ACCOA + OA ==> COA + CIT
GND	6-phosphoglycononate dehydrogenase	D6PGC + NADP ==> RL5P + CO_2 + NADPH
GPM	Phosphoglycerate mutase	P3G <==> P2G
GROWTH	Growth	0.1 F6P + 1.5 P3G + 3.7 ACCOA + 0.2 G6P + 41.3 ATP + 0.5 PEP + 1.1 AKG + 0.4 E4P + 18.2 NADPH + 1.8 OA + 0.9 R5P + 0.1 T3P1 + 3.5 NAD + 2.8 PYR ==> 41.3 PI + 3.7 COA + 41.3 ADP + 3.5 NADH + 18.2 NADP + BIOMASS
ICD	Isocitrate dehydrogenase	ICIT + NADP <==> AKG + CO_2 + NADPH
LPD	lipoamide dehydrogenase	COA + NAD + PYR ==> ACCOA + CO_2 + NADH
MAEB	Malic enzyme (NADP)	MAL + NADP ==> CO_2 + NADPH + PYR

Table S5 (*continued*)

Gene	Gene description	Reaction
MDH	Malase dehydrogenase	MAL + NAD <==> OA + NADH
NUO	NADH dehydrogenase	Q + NADH ==> QH2 + 2 HEXT + NAD
O_2XT	Oxygen transport	O_2 <==> O_2XT
PCK	Phosphoenolpyruvate carboxykinase	ATP + OA ==> PEP + CO_2 + ADP
PFK	Phosphofructokinase	F6P + ATP ==> ADP + FDP
PGI	Phosphoglucoisomerase	G6P <==> F6P
PGK	Phosphoglycerate kinase	ADP + P13DG <==> P3G + ATP
PGL	6-phosphoglyconolactonase	D6PGL ==> D6PGC
PIXT	Inorganic phosphate transport	PI <==> PIXT
PNTA	Pyridine nucleotide transhydrogenase A	NADPH + NAD ==> NADH + NADP
PNTB	Pyridine nucleotide transhydrogenase B	2 HEXT + NADH + NADP ==> NADPH + NAD
PPC	Phosphoenolpyruvate carboxylase	PEP + CO2 ==> PI + OA
PPS	Phosphoenolpyruvate synthase	ATP + PYR ==> PI + PEP + AMP
PTA	Phosphotransacetylase	ACCOA + PI <==> ACTP + COA
PTS	Glucose transport	PEP + GLC ==> G6P + PYR
PYK	Pyruvate kinase	PEP + ADP ==> ATP + PYR
RPE	Ribulose phosphate 3-epimerase	RL5P <==> X5P
RPI	Ribose-5-phosphate isomerase	RL5P <==> R5P
SDH	Succinate dehydrogenase	FAD + SUCC ==> FADH + FUM
SDH2	Succinate dehydrogenase complex	FADH + Q <==> FAD + QH2
SUCAB	2-Ketoglutarase dehydragenase	AKG + COA + NAD ==> CO_2 + SUCCOA + NADH
SUCCD	Succinyl-CoA synthetase	PI + ADP + SUCCOA <==> ATP + SUCC + COA
TALB	Transaldolase B	S7P + T3P1 <==> F6P + E4P
TKTA	Transketolase I	X5P + R5P <==> S7P + T3P1
TKTB	Transketolase II	X5P + E4P <==> F6P + T3P1
TPI	Triosphosphate isomerase	T3P1 <==> T3P2
ZWF	Glucose-6-phosphate 1-dehydrogenase	G6P + NADP <==> NADPH + D6PGL

Table S6 Metabolites in metabolic network model of *Escherichia coli*

Abbreviations	Full name	Abbreviations	Full name
			2-Keto-3-desoxy-6-phospho
AC	Acetate	KetoPGluc	gluconate
ACCOA	Acetyl coenzyme A	MAL	malate
			Nicotinamide adenine
ACTP	Acetyl phosphate	NAD	dinucleotide
			Nicotinamide adenine
ACXT	external acetate	NADH	dinucleotide - reduced
			Nicotinamide adenine
ADP	Adenosine diphosphate	NADP	dinucleotide phosphate
			Nicotinamide adenine dinucleotide
AKG	alpha-Ketoglutarate	NADPH	phosphate-reduced
	Adenosine		
AMP	5'-monophosphate	O_2	Oxygen
ATP	Adenosine triphosphate	O_2XT	External oxygen
BIOMASS	Biomass	OA	oxaloacetate
			3-Phospho-D-glyceroyl
CIT	Citrate	P13DG	phosphate
CO_2	Carbon dioxide	P2G	2-Phosphoglycerate
CO_2XT	external carbon dioxide	P3G	3-Phosphoglycerate
COA	Coenzyme A	PEP	Phosphoenolpyruvate
D6PGC	6-Phospho-D-gluconate	PI	Phosphate
	D-Glucono-1,5-lactone		
D6PGL	6-phosphate	PIXT	external phosphate
E4P	Erythrose-4-phosphate	PYR	Pyurvate
F6P	Fructose-6-phosphate	Q	Ubiquinone-8
FAD	Flavin adenine dinucleotide	QH2	Ubiquinol-8
	Flavin adenine		
FADH	dinucleotide-reduced	R5P	ribose-5-phosphate
	D-Fructose		
FDP	2,6-bisphosphate	RL5P	ribulose-5-phosphate
FUM	Fumarate	S7P	sedoheptulose-7-phosphate
G6P	Glucose-6-phosphate	SUCC	succinate
GLC	Glucose	SUCCOA	succinyl-coenzyme A
GLCXT	External glucose	T3P1	glyceraldehyde-3-phosphate
GLX	Glyoxylate	T3P2	dihydroxyacetate phosphate
HEXT	external hydrogen	X5P	xylulose-5-phosphate
ICIT	isocitrate		

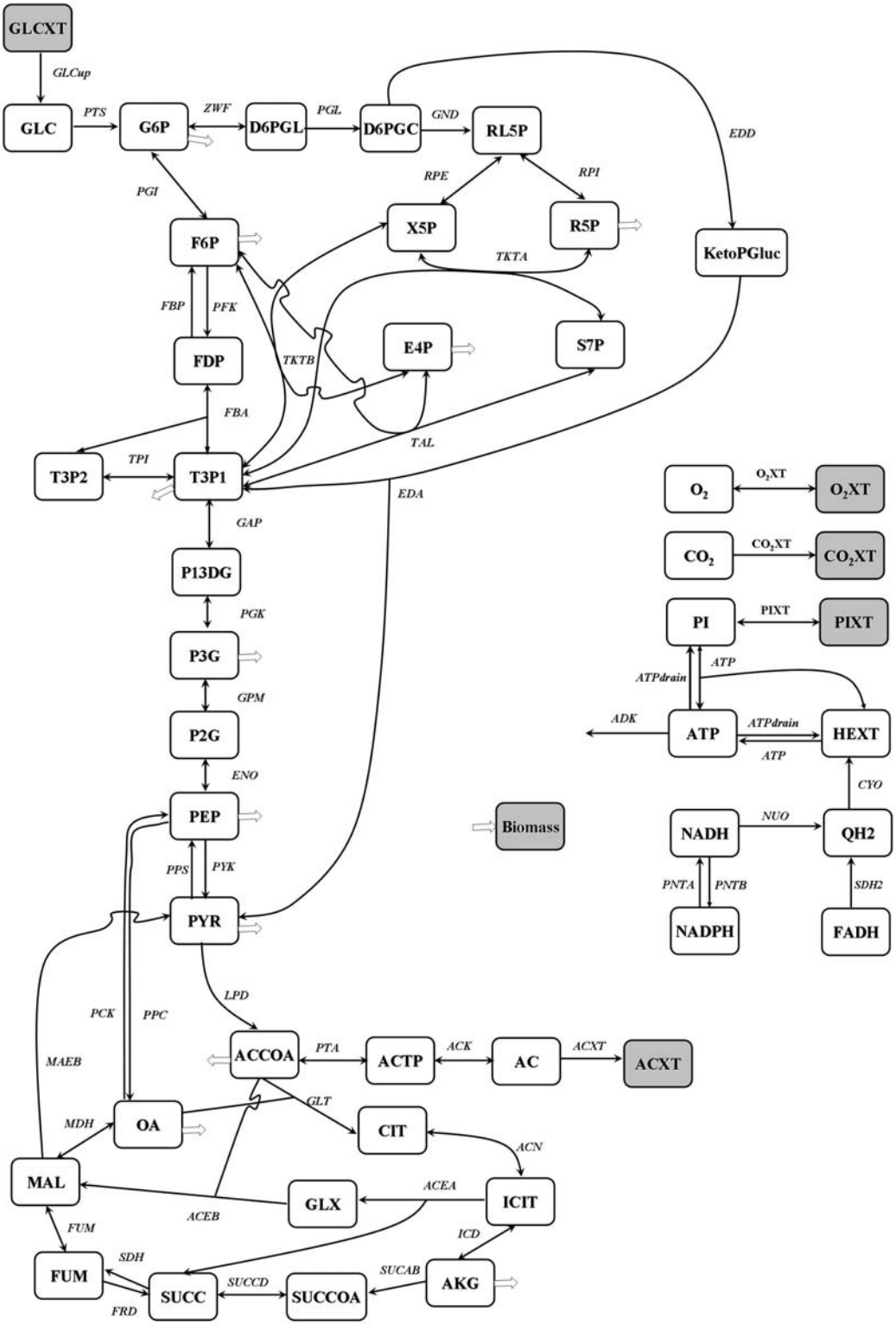

Figure S2 Central network model of *E. coli*. Grey ones are external metabolites.

Table S7 Metabolic network model for *Corynebacterium glutamicum*

Gene	Gene description	Reaction
asp	Asparagine transaminase	OAA + GLUT <==> ASP + AKG
glt, acn, icd	Citrate synthase, Aconitase, isocitrate dehydrogenase	OAA + ACCOA + H_2O + NADP <==> AKG + CO_2 + NADPH + COA
tkt II	Transketolase II	E4P + XYL5P <==> GAP + FRU6P
sucCD	SUCCCoA synthase	ADP + SUCCOA <==> ATP + COA + SUC
gap, pgk	Glyceraldehyde-3-phosphate dehydrogenase, Phosphoglycerate kinase	GAP + ADP + NAD <==> ATP + G3P + NADH
tal	Transaldolase	SED7P + GAP <==> E4P + FRU6P
glutDH	Glutamate dehydrogenase	AKG + NH_3 + NADPH <==> H_2O + NADP + GLUT
ots	trehalose-6-phosphate synthase	ATP + 2 GLC6P <==> TREHAL + ADP
pgi	Phosphoglucoisomerase	GLC6P <==> FRU6P
rpi	Ribose-5-phosphate isomerase	RIBU5P <==> RIB5P
rpe	Ribulose phosphate 3-epimerase	RIBU5P <==> XYL5P
gpm, eno	Phosphoglycerate mutase, Enolase	G3P <==> PEP + H_2O
O_2ext	Oxygen uptake	O_2 <==> O_2XT
tkt I	Transketolase I	RIB5P + XYL5P <==> SED7P + GAP
sucCD	SUCC-CoA synthase	FAD + H2O + NAD + SUC <==> OAA + FADH + NADH
Lysext	Lysine transport	LYSI ==> LYSIXT
pdh	Pyruvate dehydrogenase	COA + NAD + PYR ==> ACCOA + CO_2 + NADH
Treext	Trehalose transport	TREHAL ==> TREHALXT
sucAB	2-Ketoglutarase dehydragenase	AKG + COA + NAD ==> CO_2 + SUCCOA + NADH
pps, pyk	Phosphoenolpyruvate synthase, Pyruvate kinase	PEP + ADP ==> ATP + PYR
ilvBCDE	Acetolactate synthase I, Acetohydroxy acid isomeroreductase, Dihydroxy acid dehydratase, Branched chain amino acid aminotransferase	NADPH + GLUT + 2 PYR ==> VAL + AKG + CO_2 + H_2O + NADP
Growth	Biomass formation	332 ACCOA + 40 VAL + 3000 ATP + 80 ASP + 126 RIB5P + 33 LYSI + 150 G3P + 52 PEP + 13 GAP + 100 NADPH + 54 ALA + 25 GLUM + 7 FRU6P + 21 GLC6P + 446 GLUT + 30 PYR ==> 364 AKG + 143 CO_2 + 332 COA + 3000 ADP + 100 NADP + 1000 BIOMASS
ATP	ATP synthase	4 ADP + 2 NADH + O_2 ==> 4 ATP + 2 H_2O + 2 NAD

Table S7 (*continued*)

Gene	Gene description	Reaction
alaTr	Alanine transaminase	GLUT + PYR ==> AKG + ALA
ATPdrain	ATP drain	ATP ==> ADP
Lys	Diaminopimelate decarboxylase	MDAP ==> LYSI + CO_2
FADHDH	FADH dehydrogenase	2 FADH + 2 ADP + O_2 ==> 2 ATP + 2 FAD + 2 H2O
dapCEF	Succinyl diaminopimelate aminotransferase, Tetrahydrodipicolinate succinylase, Succinyl diaminopimelate desuccinylase	AKP + SUCCOA + H2O + GLUT ==> MDAP + AKG + COA + SUC
fbp	Fructose-1,6-bisphosphatase	H2O + FDP ==> FRU6P
pts	Phosphotransferase system	PEP + GLC ==> GLC6P + PYR
fba, tpi	Fructosebisphophaphate aldolase,Triosphosphate isomerase	FDP ==> 2 GAP
CO_2ext	CO_2 transport	CO_2 ==> CO_2XT
ppc	Phosphoenolpyruvate carboxylase	PEP + CO_2 ==> OAA
H_2Oext	Water transport	H_2O ==> H_2OXT
Thr, asd	Aspartate kinase, Aspartate semialdehyde	ATP + ASP + 2 NADPH + PYR ==> AKP + ADP + H_2O + 2 NADP
GLCup	Glucose uptake	GLCXT ==> GLC
Glum	Glutamine synthase	ATP + NH_3 + GLUT ==> ADP + GLUM
pfk	Phosphofructokinase	ATP + FRU6P ==> ADP + FDP
NH_3ext	Ammonia transport	NH3XT ==> NH_3
Zwf, pgl, gnd	Glucose-6-phosphate 1-dehydrogenase, 6-phosphoglyconolactonase, 6-phosphoglycononate dehydrogenase	H_2O + 2 NADP + GLC6P ==> CO_2 + 2 NADPH + RIBU5P

Table S8 Metabolites in metabolic network models for *Corynebacterium glutamicum*

Abbreviations	Full name	Abbreviations	Full name
OAA	Oxaloacetic acid	SUCCOA	Succinyl -coenzyme A
FADH	Flavin adenine dinucleotide (reduced)	H_2O	Water
ACCOA	Acetyl-coenzyme A	NADH	Nicotinamide adenine dinucleotide (reduced)
VAL	Valine	O_2	Oxygen
ATP	Adenosine triphosphate	ALA	Alanine
ASP	Aspartate	NAD	Nicotinamide adenine dinucleotide (oxidized)
SED7P	Sedoheptulose-7-phosphate	GLUM	Glutamine
RIB5P	Ribose-5-phosphate	GLC	Glucose
LYSI	Lysine	NADP	Nicotinamide adenine dinucleotide phosphate (oxidized)
G3P	3-phosphoglycerate	FRU6P	Fructose-6-phosphate
AKP	2-Amino-6-ketopimelate	XYL5P	Xylulose-5-phosphate
PEP	Phosphoenolpyruvate	GLC6P	Glucose-6-phosphate
MDAP	meso-diaminopimelate	FDP	Fructose 2,6-bisphosphate
GAP	Glyceraldehyde-3-phosphate	GLUT	Glutamate
AKG	alpha-ketoglutarate	SUC	Succinate
E4P	Erythrose-4-phosphate	PYR	Pyruvate
TREHAL	Trehalose	LYSIXT	External lysine
NH3	Ammonia	O_2XT	External oxygen
FAD	Flavin adenine dinucleotide (oxidized)	CO_2XT	External carbon dioxide
CO_2	Carbon dioxide	TREHALXT	External trehalose
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)	BIOMASS	Biomass
COA	Coenzyme A	H_2OXT	External Water
RIBU5P	Ribulose-5-phosphate	GLCXT	External glucose
ADP	Adenosine diphosphate	NH_3XT	External ammonia

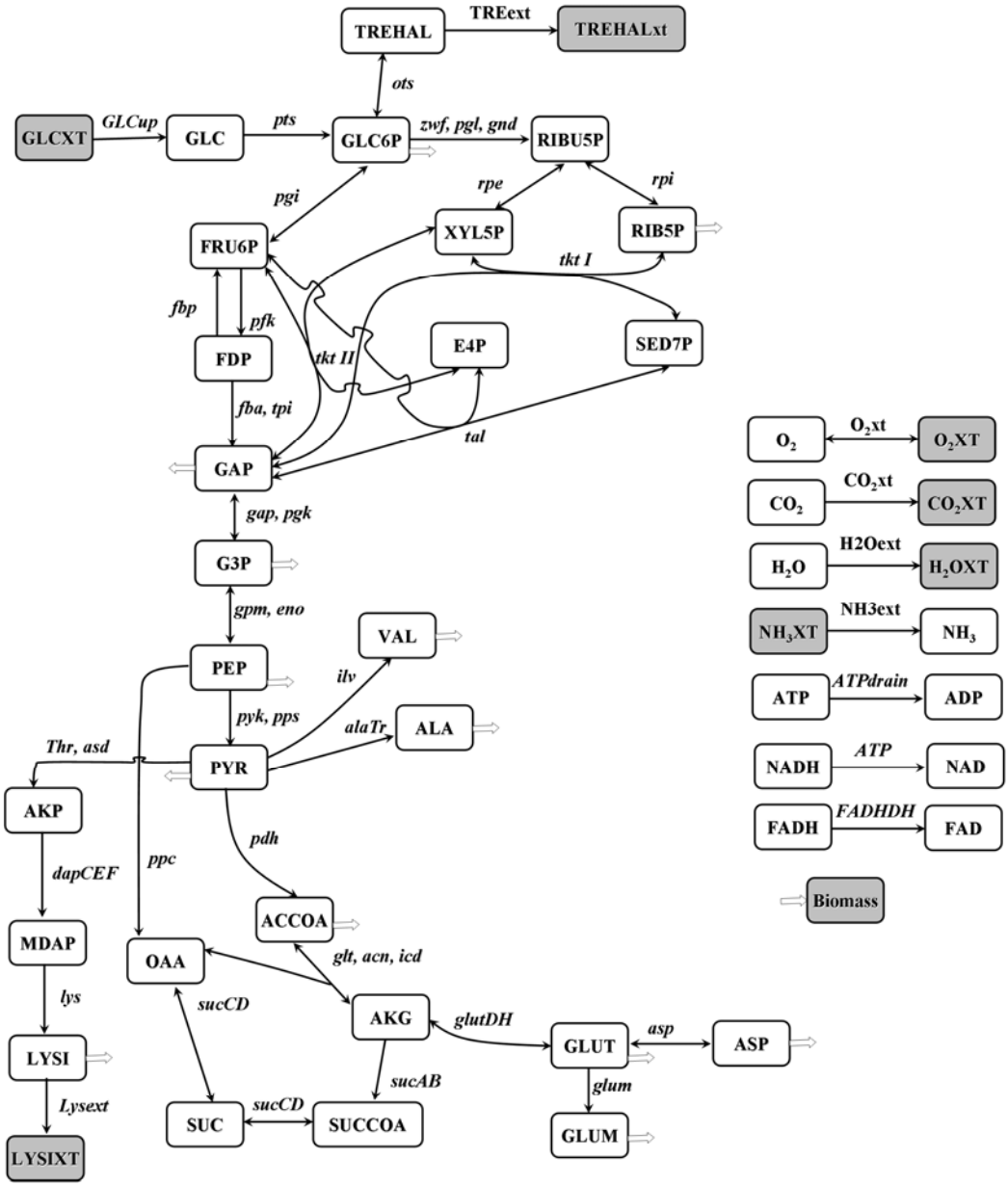

Figure S3 Central network model of *C. glutamicum*. Grey ones are external metabolites.

Table S9 Prediction errors for nine mutants of *E. coli* by FBA, MOMA and GMF

Mutant	FBA[a]	FBA[b]	MOMA	GMF
fbp	38.19	23.69	19.65	20.00
pps	35.31	20.66	10.78	9.97
rpe	32.65	34.84	11.64	9.20
tktA	38.44	40.82	15.17	14.12
tktB	21.01	23.25	4.78	4.68
fba	26.93	28.85	20.20	19.43
pgi	26.97	24.80	19.94	24.56
pgl	21.97	23.68	21.04	21.97
gnd	28.61	32.30	17.22	14.39

a: the objective function in FBA is the maximum biomass formation;

b: the objective function in FBA is the maximum ATP per flux unit (Schuetz *et al*, 2007);

The metabolic fluxes for mutants were cited from the reference (Ishii,N. *et al*., 2007)

FBA is performed by linprog in Matlab.

$$\min \sum_i c_i \mathbf{v}_i$$

$$subject\ to \quad \mathbf{S} \cdot \mathbf{v} = 0 \qquad v_{i,\min} \le v_i \le v_{i,\max} \ (i = 1,...,n)$$

In the metabolic network models of *E. coli* and *C. glutamicum*, the reaction for biomass formation was presented. The coefficients of the objective function, c_i, are -1 for this biomass formation reaction while they are zero for other reactions.