V12 metabolische Netzwerke

- Metabolisches Netzwerk: Gesamtheit aller metabolischen Pfade einer Zelle / eines Organismus
- Methoden der Systembiologie ermöglichen die integrierte, simultane Betrachtung von kompletten metabolischen Netzwerken.

- **wichtige Fragen:**

  (1) wie soll man die Gesamtheit der metabolischen Flüsse im Netzwerk beschreiben?

  (2) wie können mathematische Techniken die gezielte Manipulation von Mikroorganismen im Hinblick auf die Produktion bestimmter Substanzen unterstützen?

- **Inhalt V12:**

  (1) stöchiometrische Matrix

  (2) Flux Balance Analysis - Methode

  (3) Elementary Flux Mode – Analyse (anschaulich + Anwendung)

Ecocyc : Datenbank mit Reaktionen von *E. coli*

Datentyp	Anzahl
Gene	4499
Genprodukte, die in Mini-Review behandelt werden	3706
Genprodukte mit exp. validierten GO-Termen	2462
Enzyme	1485
Metabolische Reaktionen	1577
Substanzen	2363
Transporter	264
Transportreaktionen	348
Transportierte Substrate	254
Transkriptionsfaktoren	188
Regulatorische Interaktionen	5827
Transkriptionsinitiation	3207
Transkription – Elongation	20
Regulation der Translation	114
Enzym-Modulierung	2468
andere	18

Keseler et al. *Nucl. Acids Res. (2013) 41: D605-D612*

Beispiel: Stoffwechsel von *E. coli*

Table 1. Most Frequently Used Metabolites in *E. coli* Central Metabolism

Occurrence	Name of metabolite
205	H_2O
152	ATP
101	ADP
100	phosphate
89	pyrophosphate
66	NAD
60	NADH
54	CO_2
53	H^+
49	AMP
48	NH_3
48	NADP
45	NADPH
44	Coenzyme A
43	L-glutamate
41	pyruvate
29	acetyl-CoA
26	O_2
24	2-oxoglutarate
23	*S*-adenosyl-L-methionine
18	*S*-adenosyl-homocysteine
16	L-aspartate
16	L-glutamine
15	H_2O_2

Im Mittel enthält jede Reaktion 4 Substrate.

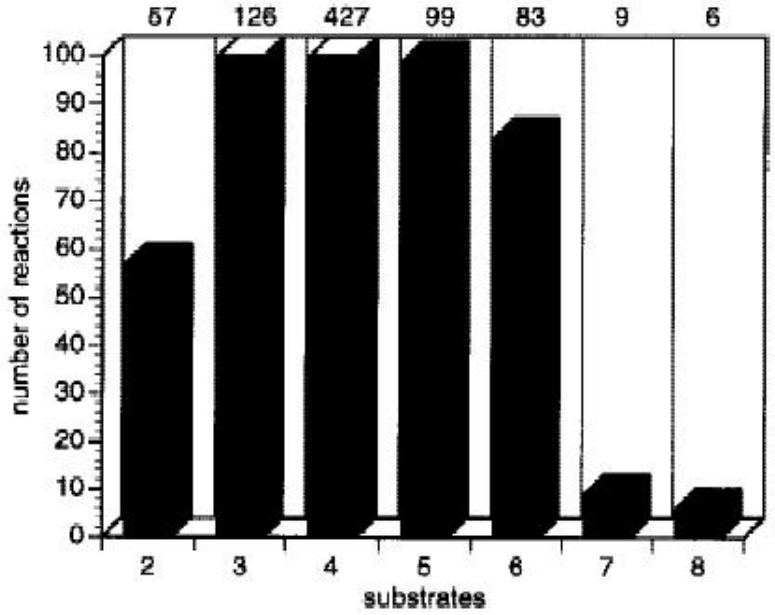


Figure 4 Diagram showing the number of reactions containing varying numbers of substrates (reactants plus products).

Ouzonis, Karp, Genome Research 10, 568 (2000)

Beispiel: Stoffwechsel von *E. coli*

Eine ältere Version von EcoCyc enthielt 131 Stoffwechsel-Pfade.

Die Länge der Pfade variiert von 1 bis 16. Der Mittelwert ist 5.4.

Von den 607 Enzymen sind 100 multifunktional. Purin-Nukleosid-Phosphorylase und Nukleosid-Diphosphatkinase katalysieren 7 bzw. 9 Reaktionen.

483 Reaktionen gehören zu einem Pfad, 99 Reaktionen gehören zu mehreren Pfaden.

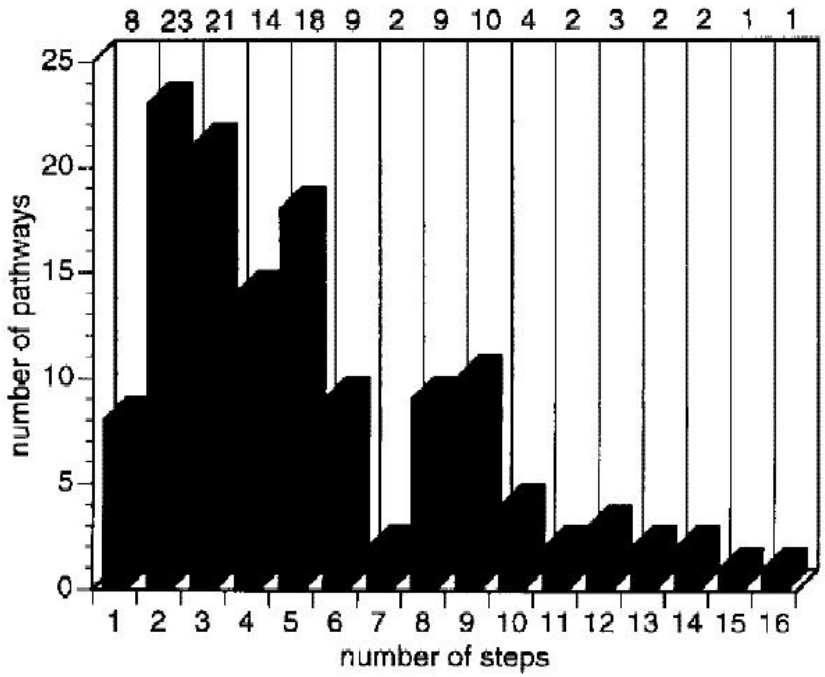


Figure 5 Length distribution of EcoCyc pathways; two pathways are not included because the number of steps is not known.

Ouzonis, Karp, Genome Research 10, 568 (2000)

Fazit

Stoffwechsel-Netzwerke von einfachen Organismen sind mittlerweile fast vollständig bekannt.

Ist die Beschreibung mit einzelnen Stoffwechsel-Wegen adäquat?

- Reaktionen, Enzyme und Substrate gehören oft zu mehreren Pfaden.

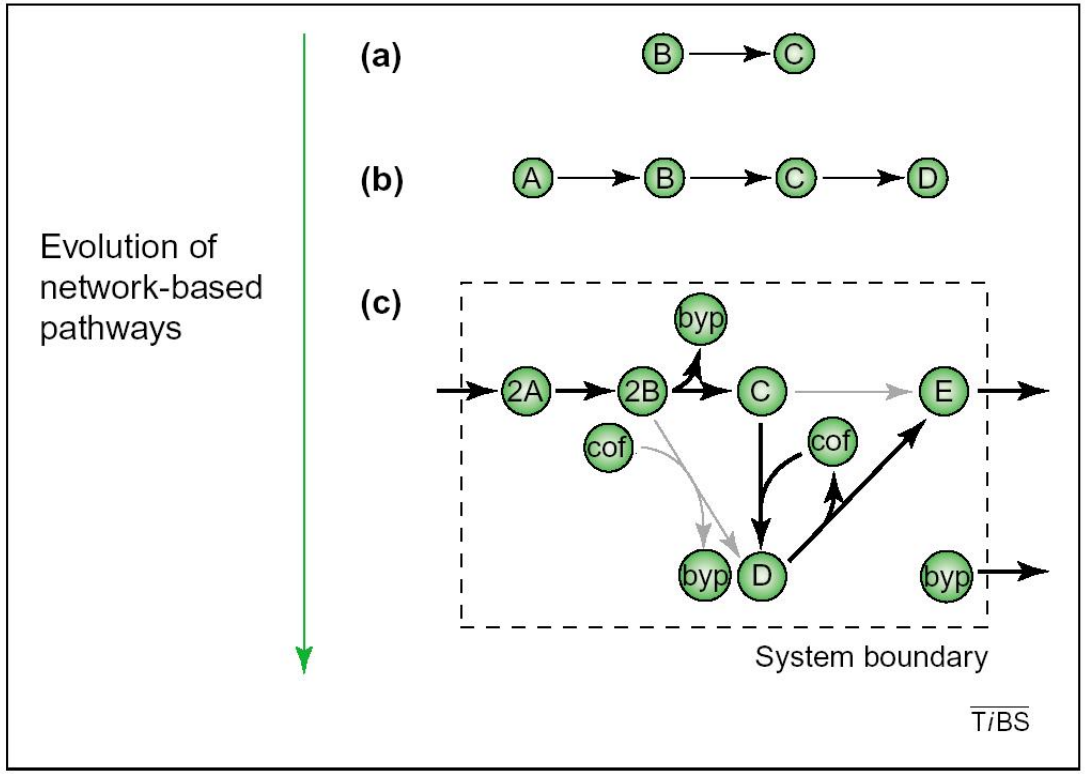
- Die Einteilung in einzelne Stoffwechsel-Pfade ist nicht immer eindeutig

Metabolische Pfade in der post-genomischen Ära

(a) klassische Biochemie bestimmt Stöchiometrien einzelner **Reaktionen**

(b) Katalogisierung vieler Reaktionen, Gruppierung nach gemeinsamen Metaboliten führt zu traditionellen **Pfaden** wie Glykolyse, Pentose-Phosphat-Pfad

(c) Durch komplette Information können nun die kompletten metabolischen Pfade zu einem **Netzwerk** kombiniert werden.



Metabolische Pfade in der post-genomischen Ära

Traditionelle metabolische Pfade dienen als konzeptioneller Rahmen für Forschung und Lehre.

Man kann dadurch Metabolismen verschiedener Organismen vergleichen.

Jedoch sind sie nicht für quantitative, systemische Bewertungen biologischer Reaktionsnetzwerke geeignet, da sie nur Teile der Netzwerke darstellen.

Sie wurden oft in Zelltypen entdeckt, in denen sie wichtige metabolische Funktionen übernehmen (z.G. Glykolyse in Hefe).

Man kann diese Pfade jedoch nicht einfach auf andere Zelltypen mit anderen Enzym- und Metabolit-Konzentrationen übertragen.

Flux Balance Analyse

Jede chemische Reaktion erfüllt die **Erhaltung der Massen**.

Mit dieser grundlegenden Bedingung können metabolische Systeme untersucht werden.

Man muss lediglich die Stöchiometrie aller metabolischen Pfade und die metabolischen Anforderungen kennen. Für jeden Metabolit gilt:

$$v_i = \frac{dX_i}{dt} = V_{synthesized} - V_{degraded} - (V_{used} - V_{transported})$$

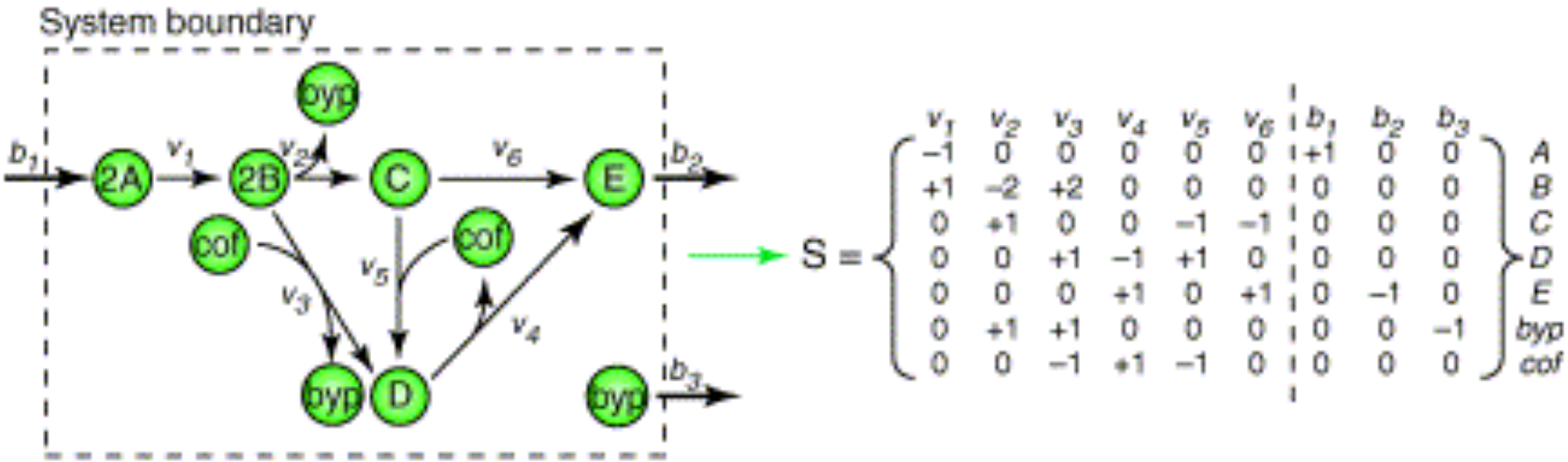
Im Gleichgewicht kann man die Massenerhaltung in einem metabolischen Netzwerk mathematisch durch eine Matrixgleichung darstellen:

$$S \cdot v = 0$$

Hierbei ist die Matrix S wiederum die $m \times n$ **stöchiometrische Matrix**,
m = Anzahl der Metabolite, n = Anzahl der Reaktionen im Netzwerk.

Der Vektor v beinhaltet alle Flüsse des metabolischen Netzwerks,
nämlich die internen Flüsse, die Transportflüsse, und das Zellwachstum.

Beschreibung vernetzter metabolischer Pfade



(a) aus genomischen, biochemischen, physiologischen Daten wird ein Reaktionsnetzwerk aufgestellt. Es gibt **interne Flüsse** innerhalb der Systemgrenzen und **externe Flüsse** zum Austausch mit der Umgebung.

(b) Dieses Netzwerk wird durch eine $m \times n$ **stöchiometrische Matrix** dargestellt, in der m Metabolite durch n Reaktionen miteinander verbunden werden..

Papin et al. TIBS 28, 250 (2003)

Lösungsraum der metabolischen Flüsse

Da die Zahl an Metaboliten gewöhnlich kleiner ist als die Zahl an Reaktionen ($m < n$) ist die Flussgleichung üblicherweise **unterbestimmt**.

Daher gibt es im Allgemeinen eine Vielzahl an erlaubten Flussverteilungen, die die Massenerhaltung erfüllen. Dieser Satz von Lösungen entspricht dem **Nullraum** der Matrix S.

Consider

$$\begin{pmatrix} 0 & 2 & 1 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Corresponds to

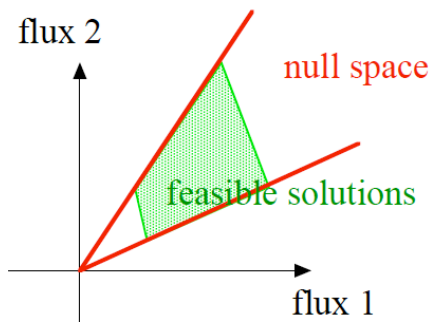
$$\begin{aligned} 2x_2 + x_3 &= 0 \\ 3x_1 - x_2 + x_3 &= 0 \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} 2x_2 &= -x_3 \\ 2x_1 &= -x_3 \end{aligned}$$

=> only one free parameter: x_3

null space: $\vec{x} = \begin{pmatrix} -a \\ -a \\ 2a \end{pmatrix}$

Add inequalities for external fluxes (here, e.g.: $x_3 \geq 0$)
=> **feasible** solutions for $a \geq 0$

Generally: null space is a cone,
constraints select part of it



E.coli in silico

Das am besten charakterisierte zelluläre System ist *E. coli*.

Im Jahr 2000 konstruierten Edwards & Palsson eine *in silico* Darstellung des *E.coli*- Metabolismus.

Dies erforderte eine Menge Handarbeit um die notwendigen Informationen
- aus der biochemischen Literatur,
- aus Genomannotationen und
- aus metabolischen Datenbanken wie EcoCyc und KEGG

zusammenzustellen.

Bernhard Palsson,
UC San Diego

Edwards & Palsson
PNAS 97, 5528 (2000)

Gene im *in silico* Modell für *E.coli*

Table 1. The genes Included In the *E. coli* metabolic genotype (21)

Zoom Out

Central metabolism (EMP, PPP, TCA cycle, electron transport)	*aceA, aceB, aceE, aceF, ackA, acnA, acnB, acs, adhE, agp, appB, appC, atpA, atpB, atpC, atpD, atpE, atpF, atpG, atpH, atpI, cydA, cydB, cydC, cydD, cyoA, cyoB, cyoC, cyoD, dld, eno, fba, fbp, fdhF, fdnG, fdnH, fdnI, fdoG, fdoH, fdoI, frdA, frdB, frdC, frdD, fumA, fumB, fumC, galM, gapA, gapC_1, gapC_2, glcB, glgA, glgC, glgP, glk, glpA, glpB, glpC, glpD, gltA, gnd, gpmA, gpmB, hyaA, hyaB, hyaC, hybA, hybC, hycB, hycE, hycF, hycG, icdA, lctD, ldhA, lpdA, malP, mdh, ndh, nuoA, nuoB, nuoE, nuoF, nuoG, nuoH, nuoI, nuoJ, nuoK, nuoL, nuoM, nuoN, pckA, pfkA, pfkB, pflA, pflB, pflC, pflD, pgi, pgk, pntA, pntB, ppc, ppsA, pta, purT, pykA, pykF, rpe, rpiA, rpiB, sdhA, sdhB, sdhC, sdhD, sfcA, sucA, sucB, sucC, sucD, talB, tktA, tktB, tpiA, trxB, zwf, pgl* (30), *maeB* (30)
Alternative carbon source	*adhC, adhE, agaY, agaZ, aldA, aldB, aldH, araA, araB, araD, bglX, cpsG, deoB, fruK, fucA, fucI, fucK, fucO, galE, galK, galT, galU, gatD, gatY, glk, glpK, gntK, gntV, gpsA, lacZ, manA, melA, mtlD, nagA, nagB, nanA, pfkB, pgi, pgm, rbsK, rhaA, rhaB, rhaD, srlD, treC, xylA, xylB*
Amino acid metabolism	*adi, aldH, alr, ansA, ansB, argA, argB, argC, argD, argE, argF, argG, argH, argI, aroA, aroB, aroC, aroD, aroE, aroF, aroG, aroH, aroK, aroL, asd, asnA, asnB, aspA, aspC, avtA, cadA, carA, carB, cysC, cysD, cysE, cysH, cysI, cysJ, cysK, cysM, cysN, dadA, dadX, dapA, dapB, dapD, dapE, dapF, dsdA, gabD, gabT, gadA, gadB, gdhA, glk, glnA, gltB, gltD, glyA, goaG, hisA, hisB, hisC, hisD, hisF, hisG, hisH, hisI, ilvA, ilvB, ilvC, ilvD, ilvE, ilvG_1, ilvG_2, ilvH, ilvI, ilvM, ilvN, kbl, ldcC, leuA, leuB, leuC, leuD, lysA, lysC, metA, metB, metC, metE, metH, metK, metL, pheA, proA, proB, proC, prsA, putA, sdaA, sdaB, serA, serB, serC, speA, speB, speC, speD, speE, speF, tdcB, tdh, thrA, thrB, thrC, tnaA, trpA, trpB, trpC, trpD, trpE, tyrA, tyrA, tyrB, ygjG, ygjH, alaB* (42), *dapC* (43), *pot* (44), *prr* (44), *sad* (45), *methylthioadenosine nucleosidase* (46), *5-methylthioribose kinase* (46), *5-methylthioribose-1-phosphate isomerase* (46), *adenosyl homocysteinase* (47), *L-cysteine desulfhydrase* (44), *glutaminase A* (44), *glutaminase B* (44)
Purine & pyrimidine metabolism	*add, adk, amn, apt, cdd, cmk, codA, dcd, deoA, deoD, dgt, dut, gmk, gpt, gsk, guaA, guaB, guaC, hpt, mutT, ndk, nrdA, nrdB, nrdD, nrdE, nrdF, purA, purB, purC, purD, purE, purF, purH, purK, purL, purM, purN, purT, pyrB, pyrC, pyrD, pyrE, pyrF, pyrG, pyrH, pyrI, tdk, thyA, tmk, udk, udp, upp, ushA, xapA, yicP, CMP glycosylase* (48)
Vitamin & cofactor metabolism	*acpS, bioA, bioB, bioD, bioF, coaA, cyoE, cysG, entA, entB, entC, entD, entE, entF, epd, folA, folC, folD, folE, folK, folP, gcvH, gcvP, gcvT, gltX, glyA, gor, gshA, gshB, hemA, hemB, hemC, hemD, hemE, hemF, hemH, hemK, hemL, hemM, hemX, hemY, ilvC, lig, lpdA, menA, menB, menC, menD, menE, menF, menG, metF, mutT, nadA, nadB, nadC, nadE, ntpA, pabA, pabB, pabC, panB, panC, panD, pdxA, pdxB, pdxH, pdxJ, pdxK, pncB, purU, ribA, ribB, ribD, ribE, ribH, serC, thiC, thiE, thiF, thiG, thiH, thrC, ubiA, ubiB, ubiC, ubiG, ubiH, ubiX, yaaC, ygiG, nadD* (49), *nadF* (49), *nadG* (49), *panE* (50), *pncA* (49), *pncC* (49), *thiB* (51), *thiD* (51), *thiK* (51), *thiL* (51), *thiM* (51), *thiN* (51), *ubiE* (52), *ubiF* (52), *arabinose-5-phosphate isomerase* (22), *phosphopentothenate-cysteine ligase* (50), *phosphopentothenate-cysteine decarboxylase* (50), *phospho-pentotheine adenylyltransferase* (50), *dephosphoCoA kinase* (50), *NMN glycohydrolase* (49)
Lipid metabolism	*accA, accB, accD, atoB, cdh, cdsA, cls, dgkA, fabD, fabH, fadB, gpsA, ispA, ispB, pgpB, pgsA, psd, pssA, pgpA* (53)
Cell wall metabolism	*ddlA, ddlB, galF, galU, glmS, glmU, htrB, kdsA, kdsB, kdtA, lpxA, lpxB, lpxC, lpxD, mraY, msbB, murA, murB, murC, murD, murE, murF, murG, murI, rfaC, rfaD, rfaF, rfaG, rfaI, rfaJ, rfaL, ushA, glmM* (54), *lpoA* (55), *rfaE* (55), *tetraacyldisaccharide 4' kinase* (55), *3-deoxy-D-manno-octulosonic-acid 8-phosphate phosphatase* (55)
Transport processes	*araE, araF, araG, araH, argT, aroP, artI, artJ, artM, artP, artQ, brnQ, cadB, chaA, chaB, chaC, cmtA, cmtB, codB, crr, cycA, cysA, cysP, cysT, cysU, cysW, cysZ, dctA, dcuA, dcuB, dppA, dppB, dppC, dppD, dppF, fadL, focA, fruA, fruB, fucP, gabP, galP, gatA, gatB, gatC, glnH, glnP, glnQ, glpF, glpT, gltJ, gltK, gltL, gltP, gltS, gntT, gpt, hisJ, hisM, hisP, hisQ, hpt, kdpA, kdpB, kdpC, kgtP, lacY, lamB, livF, livG, livH, livJ, livK, livM, lktP, lysP, malE, malF, malG, malK, malX, manX, manY, manZ, melB, mglA, mglB, mglC, mtlA, mtr, nagE, nanT, nhaA, nhaB, nupC, nupG, oppA, oppB, oppC, oppD, oppF, panF, pheP, pitA, pitB, pnuC, potA, potB, potC, potD, potE, potF, potG, potH, potI, proP, proV, proW, proX, pstA, pstB, pstC, pstS, ptsA, ptsG, ptsI, ptsN, ptsP, purB, putP, rbsA, rbsB, rbsC, rbsD, rhaT, sapA, sapB, sapD, sbp, sdaC, srlA_1, srlA_2, srlB, tdcC, tnaB, treA, treB, trkA, trkG, trkH, tsx, tyrP, ugpA, ugpB, ugpC, ugpE, uraA, xapB, xylE, xylF, xylG, xylH, fruF* (56), *gntS* (57), *metD* (43), *pnuE* (49), *scr* (56)

Edwards & Palsson

PNAS 97, 5528 (2000)

E.coli in silico

Begrenze $\alpha_i = 0$ für irreversible interne Flüsse,

$\alpha_i = -\infty$ für reversible interne Flüsse (aufgrund der biochemischen Literatur)

Transportflüsse für PO_4^{2-}, NH_3, CO_2, SO_4^{2-}, K^+, Na^+ blieben unbeschränkt.

Für die anderen Metabolite wurden Obergrenzen verwendet $0 < v_i < v_i^{max}$
außer für die Metabolite, die das metabolische Netzwerk verlassen könne
(d.h. Acetat, Äthanol, Laktat, Succinat, Format, Pyruvat etc.)

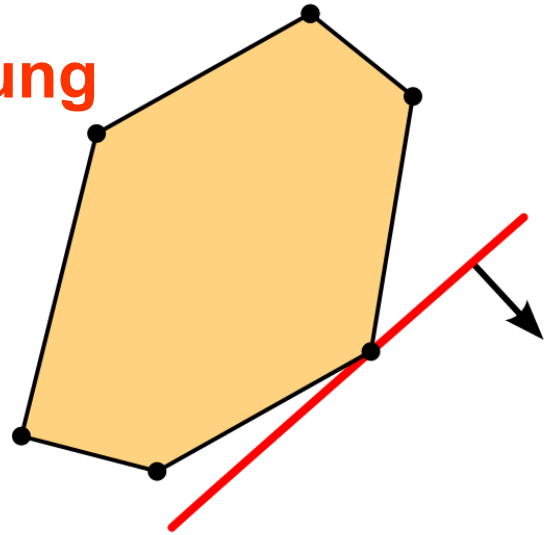
Aus dem Satz der möglichen Flussverteilungen wird mit **linearer Programmierung** eine bestimmte Lösung bestimmt, die eine bestimmte metabolische Zielfunktion Z maximiert (bzw. –Z minimiert).

$$Z = \sum c_i \cdot v_i = \langle \mathbf{c} \cdot \mathbf{v} \rangle$$

Als Zielfunktion Z wird hier die Summe aller Einzelflüsse definiert, was der Bildung maximaler Biomasse entspricht.

Edwards & Palsson, PNAS 97, 5528 (2000)

Lineare Programmierung



Lineare Programmierung ist eine Technik um optimale Werte für eine lineare objektive Funktion zu finden, wobei lineare Bedingungen für die Gleichheit und Ungleichheit für einzelne Variablen gelten.

Der Lösungsraum dieser Bedingungen ist ein konvexes Polytop. Jede Bedingung definiert eine begrenzende Fläche.

Die objektive Funktion ist eine lineare Funktion, die auf diesem Polyeder definiert ist.

Ein Algorithmus für lineare Programmierung findet einen Punkt in diesem Polyeder, wo diese Funktion den kleinsten (oder größten) Wert annimmt.

Dargestellt ist ein einfaches lineares Programm mit 2 Variablen (x und y) und 6 Ungleichungen.

Der Lösungsraum ist gelb gefärbt und bildet ein 2-dimensionales Polygon.

Die rote Linie repräsentiert die lineare Kostenfunktion.
Der Pfeil zeigt in die Richtung, in die wir optimieren.

www.wikipedia.org

Lineare Programmierung

Lineare Programme können in folgender kanonischer Form ausgedrückt werden:

$$\begin{aligned} \text{maximize} \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \\ \text{and} \quad & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Dabei ist **x** der zu bestimmende Vektor der Variablen, **c** und **b** sind Vektoren aus (bekannten) Koeffizienten, *A* ist eine (bekannte) Matrix mit Koeffizienten, und $(.)^T$ ist die transponierte Matrix.

$\mathbf{c}^T\mathbf{x}$ ist die zu optimierende objektive Funktion.

$\mathbf{c}^T = (c_1, c_2, .., c_k)$; $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \end{pmatrix}$; $\mathbf{c}^T\mathbf{x} = c_1 x_1 + c_2 x_2 + \dots + c_k x_k$

Die Ungleichungen $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq 0$ sind die Bedingungen, die den Lösungsraum definieren.

Wenn alle unbekannten Variablen ganzzahlige Werte annehmen müssen, nennt man das Problem ein **integer lineares Programmierungs** (ILP) problem.

15

www.wikipedia.org

E.coli in silico

Um die Korrektheit dieses Vorgehens zu testen, wurden Änderungen des metabolischen Netzwerkes simuliert, die sich durch hypothetische Gendeletionen ergeben.

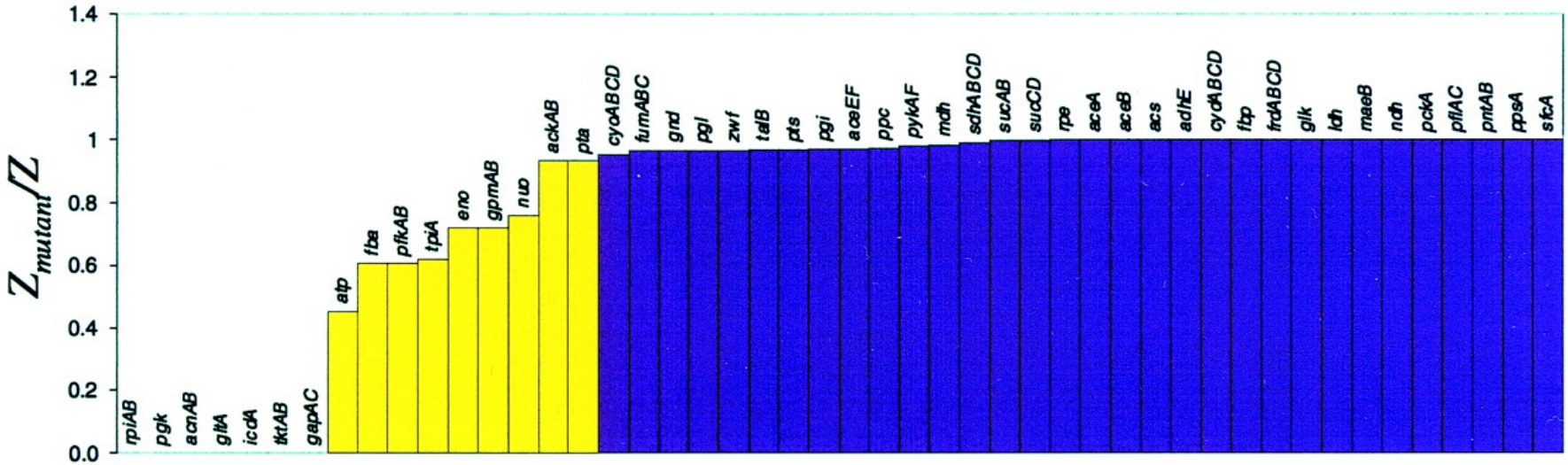
Dafür wird einfach der Fluss durch die entsprechende enzymatische Reaktion auf 0 gesetzt.

Dann wird der optimale Wert der Mutante (Z_{mutant}) mit dem des Wild-types verglichen (Z)

$$\frac{Z_{mutant}}{Z}$$

Edwards & Palsson
PNAS 97, 5528 (2000)

Gendeletionen im zentralen *E. coli* MG1655 Metabolismus



Maximale Flüsse für virtuelles Wachstum auf Glucose für alle möglichen Deletionen einzelner Gene in Pfaden des zentralen Metabolismus (Glykolyse, Pentose Phosphat Pfad (PPP), Zitratzyklus, Atmung).

Gelbe Balken: Gendeletionen, die die maximale Biomasse (Fluss) auf weniger von 95% des Wert für den *in silico* Wildtyp drücken.

Edwards & Palsson PNAS 97, 5528 (2000)

Interpretation der Ergebnisse

Die essentiellen Gene gehörten zur 3-Kohlenstoff-Stufe der Glykolyse, 3 Reaktionen des Zitratzyklus und mehrere Mitglieder des PPP.

Alle anderen Gene des zentralen Metabolismus konnten entfernt werden ohne nenneswerten Effekt auf das *in silico*-Wachstum von *E.coli*.

Edwards & Palsson PNAS 97, 5528 (2000)

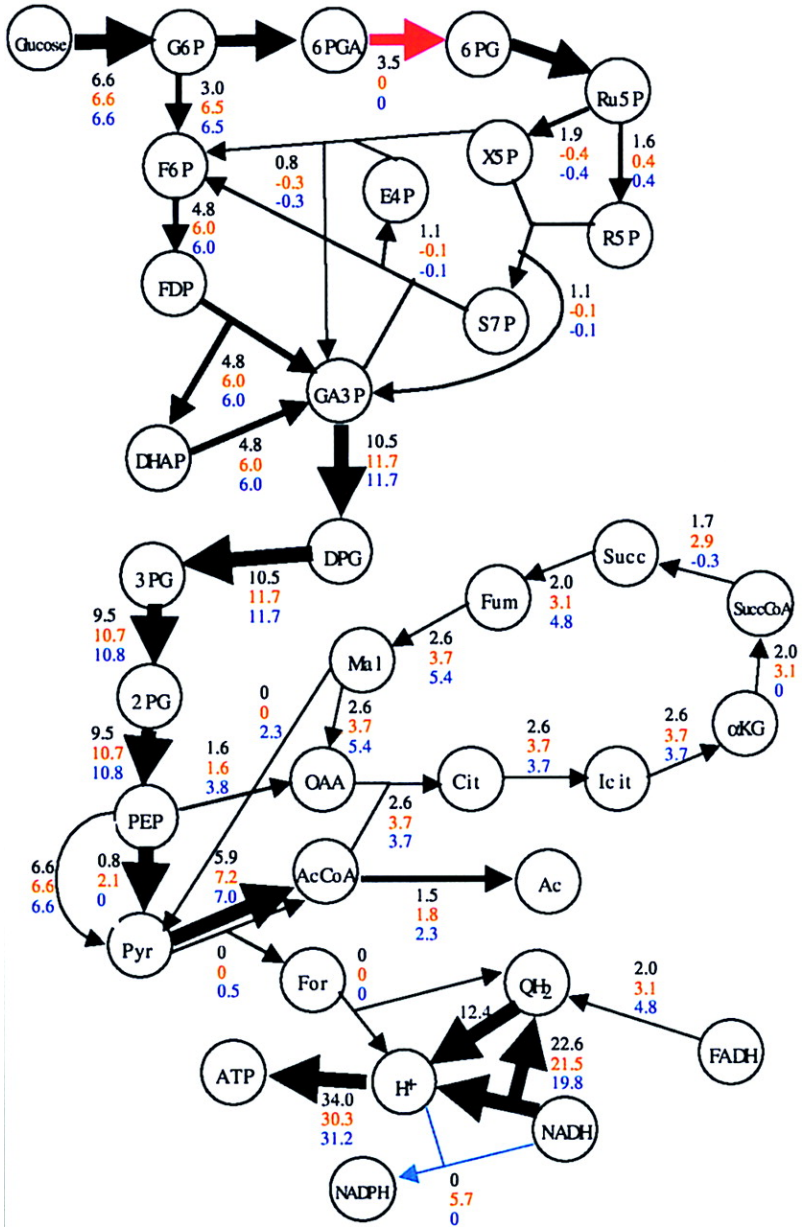
Umlenken der metabolischen Flüsse

(Schwarz) Flussverteilung für Wildtyp.

(Rot) *zwf*- Mutante. Die Biomasse-Erzeugung ist 99% des Wildtyps.

(Blau) *zwf*- *pnt*- Doppelmutante.
Biomasse-Erzeugung ist 92% des Wildtyps.

E.coli in silico umgeht die Deletion einer kritischen Reaktion (roter Pfeil) durch Erhöhung des Flusses durch die alternative G6P → P6P Reaktion.



Edwards & Palsson PNAS 97, 5528 (2000)

E.coli in silico

+ und – bedeuten Wachstum bzw. Nichtwachstum.
± heisst, dass Suppressormutanten beobachtet wurden, die diesem Mutanten-stamm Wachstum erlaubten.

glc: Glucose, gl: Glycerol,
succ: Succinat, ac: Acetat.

In 68 von 79 Fällen stimmte die Vorhersage mit dem Experiment überein.

Roten und gelbe Kreise kennzeichnen die vorhergesagten Mutationen, die Wachstum eliminieren bzw. reduzieren.

Edwards & Palsson
PNAS 97, 5528 (2000)

Table 2. Comparison of the predicted mutant growth characteristics from the gene deletion study to published experimental results with single mutants

Gene	glc	gl	succ	ac
aceA	+/+		+/+	–/–
aceB				–/–
*aceEF**	–/+			
ackA				+/+
acn	–/–			–/–
acs				+/+
cyd	+/+			
cyo	+/+			
eno†	–/+	–/+	–/–	–/–
fbaI	–/+			
fbp	+/+	–/–	–/–	–/–
frd	+/+		+/+	+/+
gap	–/–	–/–	–/–	–/–
glk	+/+			
gltA	–/–			–/–
gnd	+/+			
idh	–/–			–/–
mdh††	+/+	+/+	+/+	
ndh	+/+	+/+		
nuo	+/+	+/+		
pfk†	–/+			
pgi‡	+/+	+/–	+/–	
pgk	–/–	–/–	–/–	–/–
pgl	+/+			
pntAB	+/+	+/+	+/+	
ppc§	±/+	–/+	+/+	
pta				+/+
pts	+/+			
pyk	+/+			
rpi	–/–	–/–	–/–	–/–
sdhABCD	+/+		–/–	–/–
sucAB	+/+		–/+	–/+
tktAB	–/–			
*tpi***	–/+	–/–	–/–	–/–
unc	+/+		±/+	–/–
zwf	+/+	+/+	+/+	

EcoCyc FBA Modell

FBA Modell für die EcoCyc-Daten kann mit der Software MetaFlux berechnet werden.
Es gibt eine SBML-Datei davon http://biocyc.org/download.shtml.

Das EcoCyc FBA Modell enthält 1888 Reaktionen.

58 Metabolite produziere Biomass.
Auf einem minimalen Medium (Glukose, Ammoniak) haben 370 Reaktionen einen Fluss ungleich 0.

Für 383 Wachstumsbedingungen gibt das Modell in 72.6% der Fälle eine korrekte Vorhersage von Wachstum/Nicht-Wachstum.

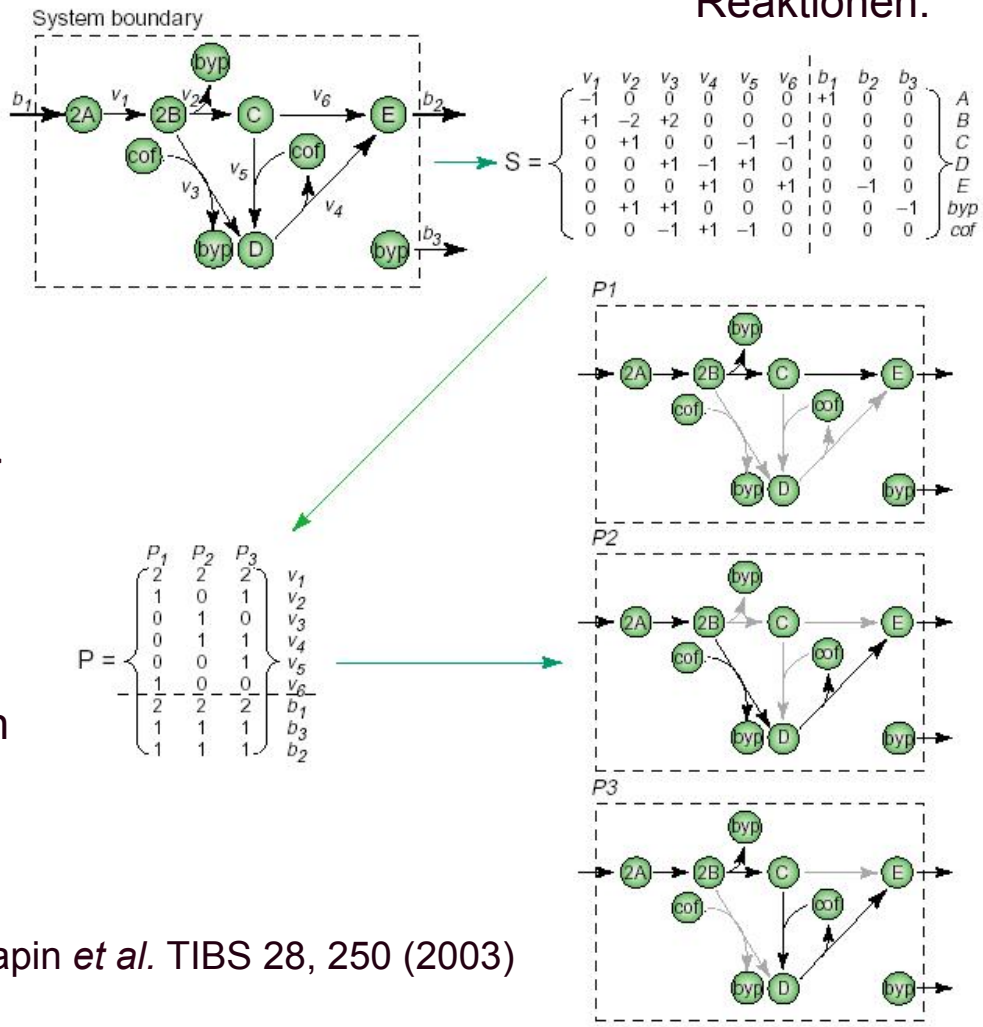
Für 4207 Deletionen einzelner Gene gibt das Modell in 91.2% der Fälle eine korrekte Vorhersage von Wachstum/Nicht-Wachstum.

Latendresse M, Krummenacker M, Trupp M, Karp PD
Bioinformatics 2012;28:388-396

Berechnung von Elementaren Flussmoden aus der stöchiometrischen Matrix

Darstellung des Reaktions-netzwerks mit stöchiometrischer Matrix S.

stöchiometrische Koeffizienten der einzelnen Reaktionen.

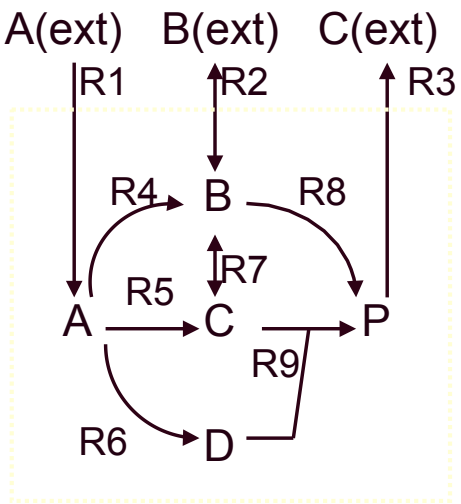


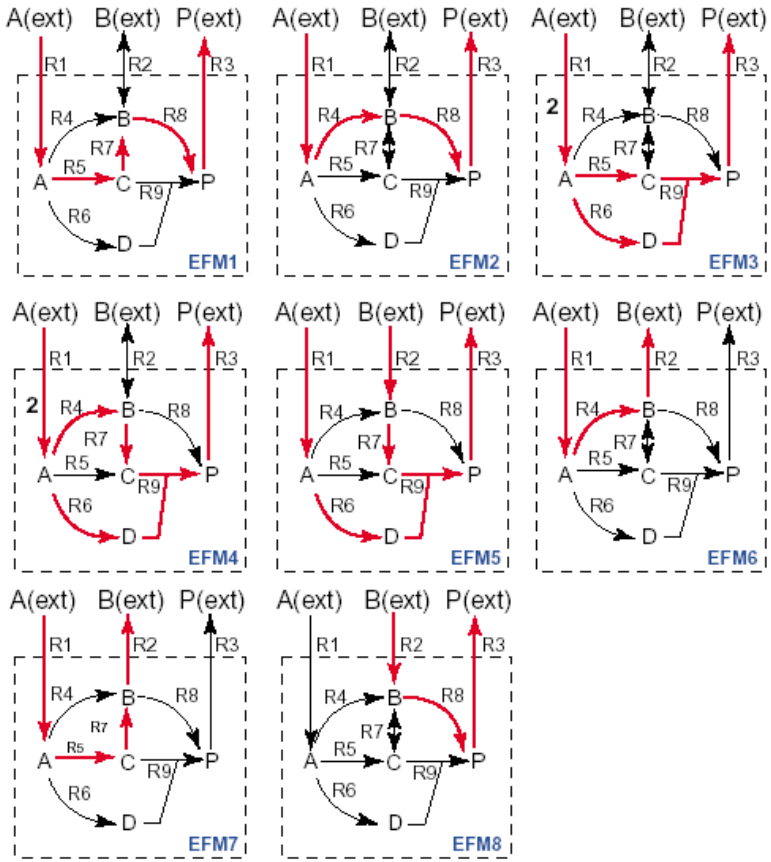
Analyse der Matrix S → Pathway-Darstellung P. Deren Zeilen enthalten den Reaktionen entsprechende Flüsse und die Spalten die sich ergebenden Pfade.

Elementare Flussmoden für dies System.

Papin *et al.* TIBS 28, 250 (2003)

Elementare Flussmoden in Modell-Netzwerk





Klamt & Stelling Trends Biotech 21, 64 (2003)

Stamm-Optimierung basierend auf EFM-Analyse

Metabolic Engineering 12 (2010) 112–122

Contents lists available at ScienceDirect

Metabolic Engineering

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

Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid

Pornkamol Unrean, Cong T. Trinh, Friedrich Srienc*

Department of Chemical Engineering and Materials Science, and BioTechnology Institute, University of Minnesota, 240 Gortner Laboratory, 1479 Gortner Ave, St. Paul, MN 55108, USA

Carotenoide (z.B. DPL und DPA) sind Licht-absorbierende Pigmente, schützen vor UV-Strahlung, regulieren die Fluidität von Membranen, und wirken antioxidierend.

Sie werden als Nahrungszusatzstoffe, Pharmazeutika und als Lebensmittelfarbstoffe verwendet.

Ziel des Projekts: erhöhe die Carotenoid-Synthese in *E.coli*

Unrean et al. Metabol Eng 12, 112-122 (2010)

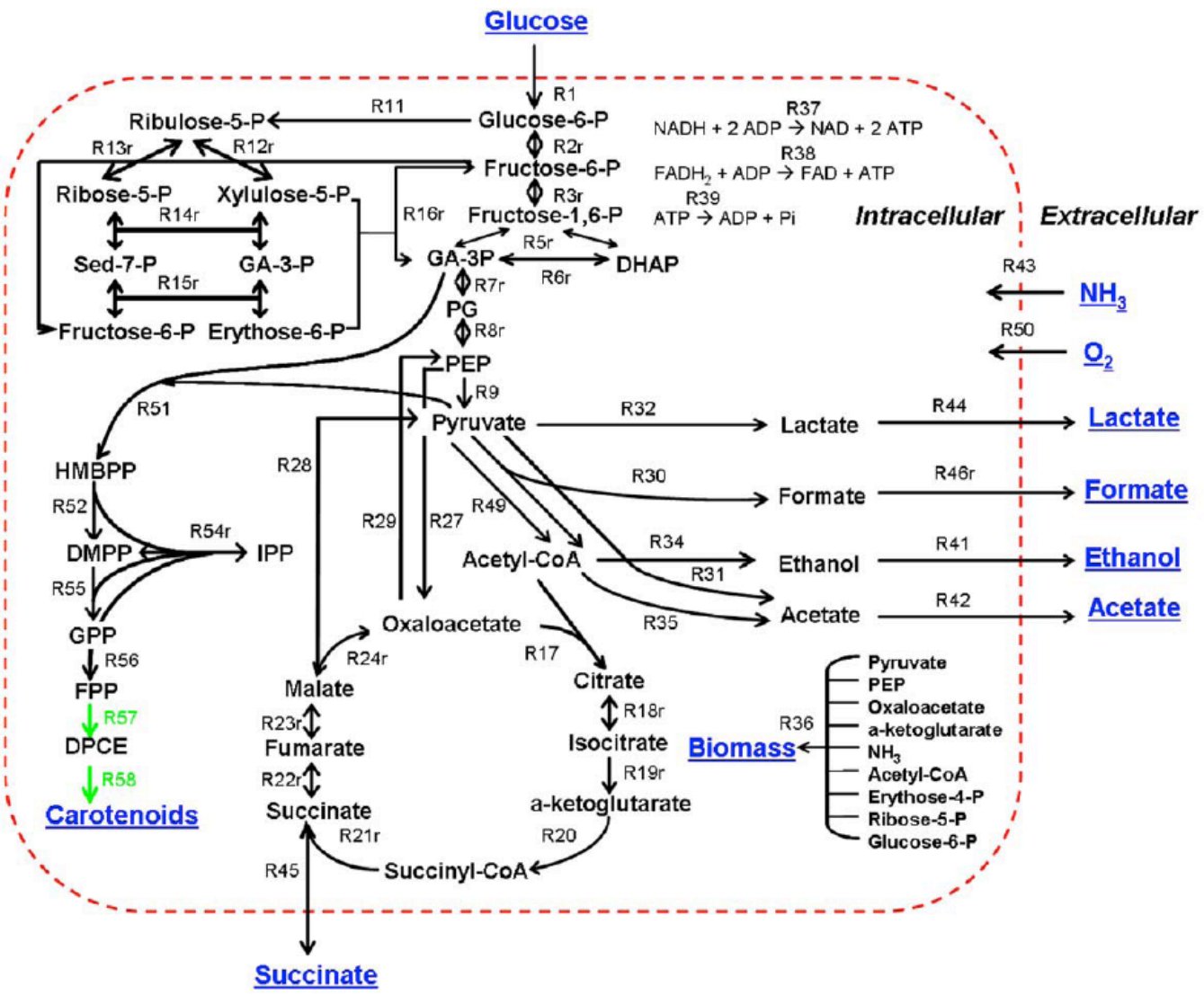
Metabolisches Netzwerk für rekombinantes *E.coli*

58 metabolische Reaktionen,
22 reversibel
36 irreversibel

57 Metabolite

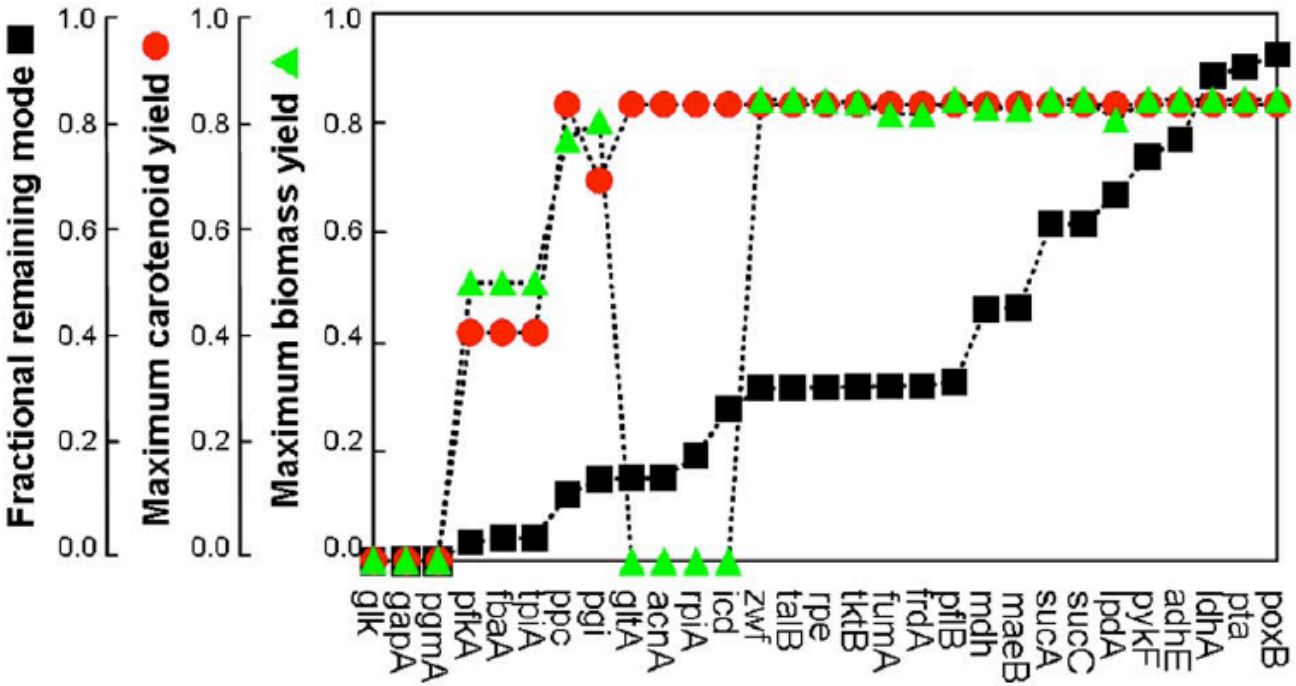
29532 EFMs

In 5923 EFMs ist die Produktion von Biomasse und DPA gekoppelt.



Unrean et al. Metabol Eng 12, 112-122 (2010)

Effekte einzelner Gendeletionen



Resultate für virtuelle Gen-Kockouts (Abzählen der EFMs und Berechnung der CRT- und Biomasse-Produktion gemäß Stöchiometrien).

Wähle Targetgene, deren Knockouts noch eine möglichst hohe Ausbeute an CRT sowie eine sinnvolle Biomasse-Produktion behalten, durch die jedoch möglichst viele EFMs eliminiert werden.

Unrean et al. Metabol Eng 12, 112-122 (2010)

Effekte einzelner Gendeletionen

Strain	Total modes	Aerobic modes	Anaerobic modes	Predicted CRT yield[a]
Wild-type	29,532	24,155	5377	0.0-426
Δ*ldhA*	15,662	13,405	2257	0.0-426
Δ*ldhA*Δ*frdA*	8573	7810	763	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*	7541	6861	680	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*	6171	5600	571	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*	4099	4099	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*	2573	2573	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*	375	375	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*Δ*maeB*	5	5	0	0.4-426

[a] Yield is in mg-diapolycopendioic acid/g–glucose.

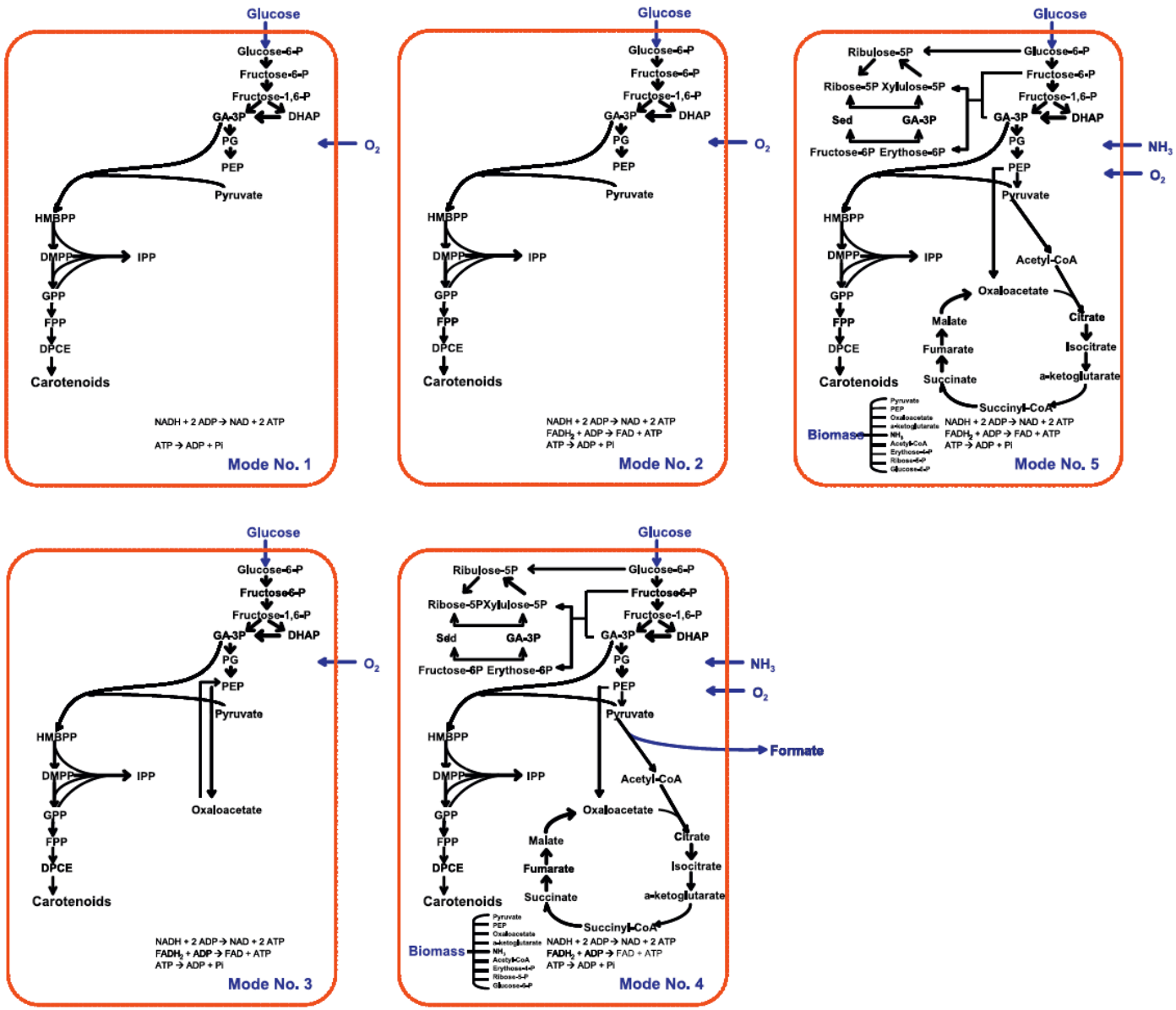
Deleted Reaction	Corresponding gene	Enzyme	Pathway
R9	*pykF*	Pyruvate kinase	Glycolysis
R11	*zwf*	Glucose-6-phosphate-1-dehydrogenase	Pentose phosphate
R22	*frdA*	Fumarate reductase	Fermentation
R28	*maeB*	Malate dehydrogenase	Anapleurotic
R31	*poxB*	Pyruvate oxidase	Fermentation
R32	*ldhA*	Lactate dehydrogenase	Fermentation
R34	*adhE*	Alcohol dehydrogenase	Fermentation
R35	*pta*	Phosphate acetyltransferase	Fermentation

Als optimale Lösung vorhergesagt:
der Knockout von 8 Genen sollte zur Überproduktion von DPL und DPA führen.

nur 5 EFMs verbleiben für das restliche Netzwerk.

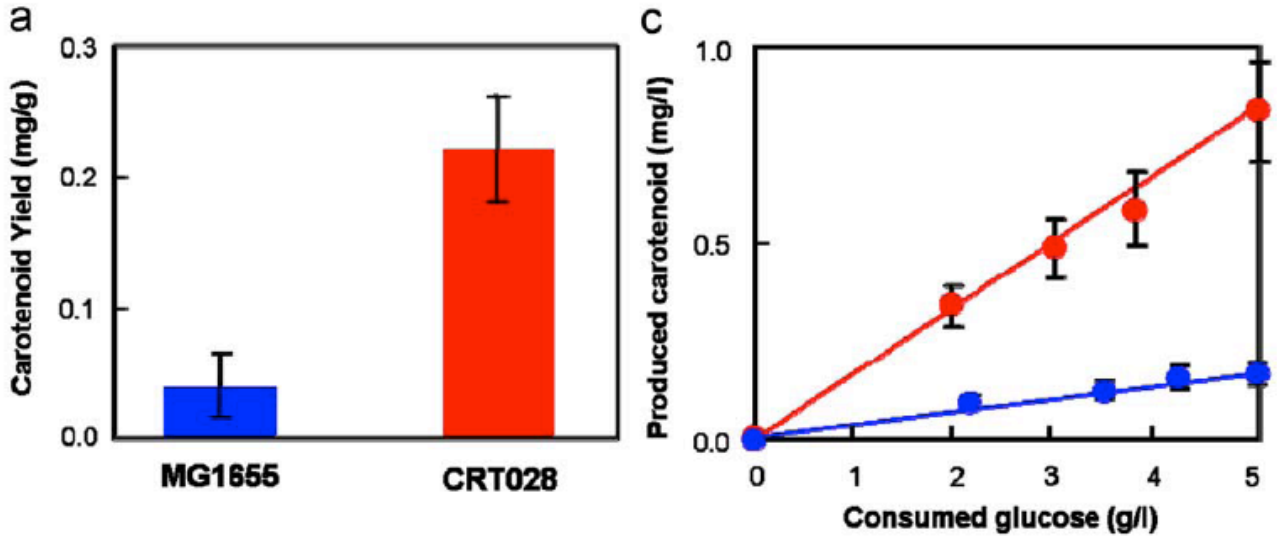
Unrean et al. Metabol Eng 12, 112-122 (2010)

Verbleibende EFMs



Unrean et al. Metabol Eng 12, 112-122 (2010)

Experimentelle Verifikation: erhöhte Carotenoid-Ausbeute



	MG1655/ pACMNOx	CRT028/ pACMNOx
Growth rate (/h)	0.17 ± 0.02	0.13 ± 0.01
Carotenoid production (mg/l)	0.19 ± 0.02	0.83 ± 0.20
Carotenoid yield (mg carotenoid/g glucose)	0.04 ± 0.00	0.17 ± 0.04
Specific production (mg carotenoid/g cell dry weight-h)	0.01 ± 0.00	0.10 ± 0.02

Mutante wächst langsamer, aber CRT-Produktion ist vierfach erhöht.

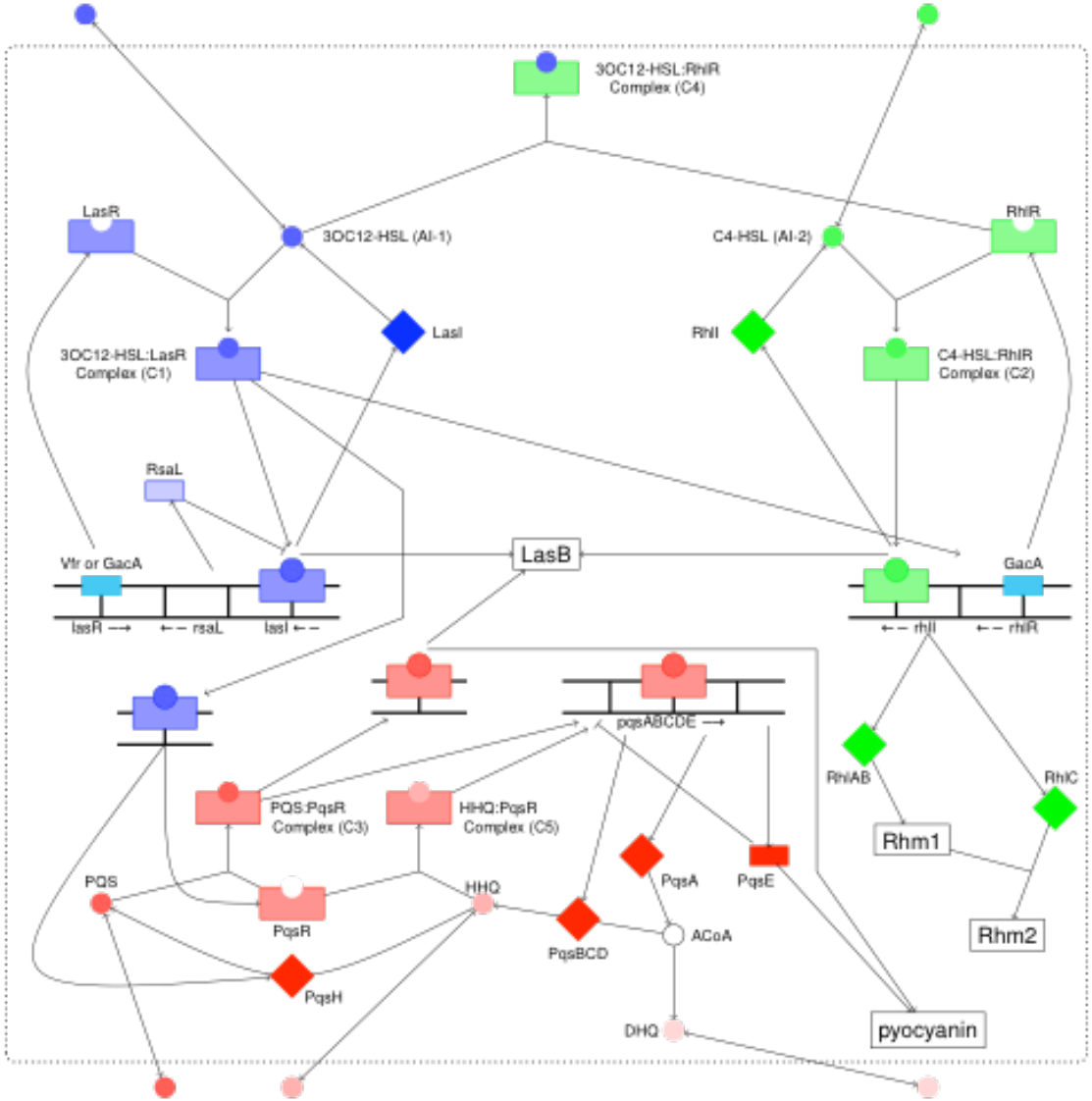
Unrean et al. Metabol Eng 12, 112-122 (2010)

Quorum Sensing in *P. aeruginosa*

In *P. aeruginosa*, besteht das QS-Netzwerk aus 3 Systemen (*las*, *rhl*, und *pqs*), die hierarchisch organisiert sind.

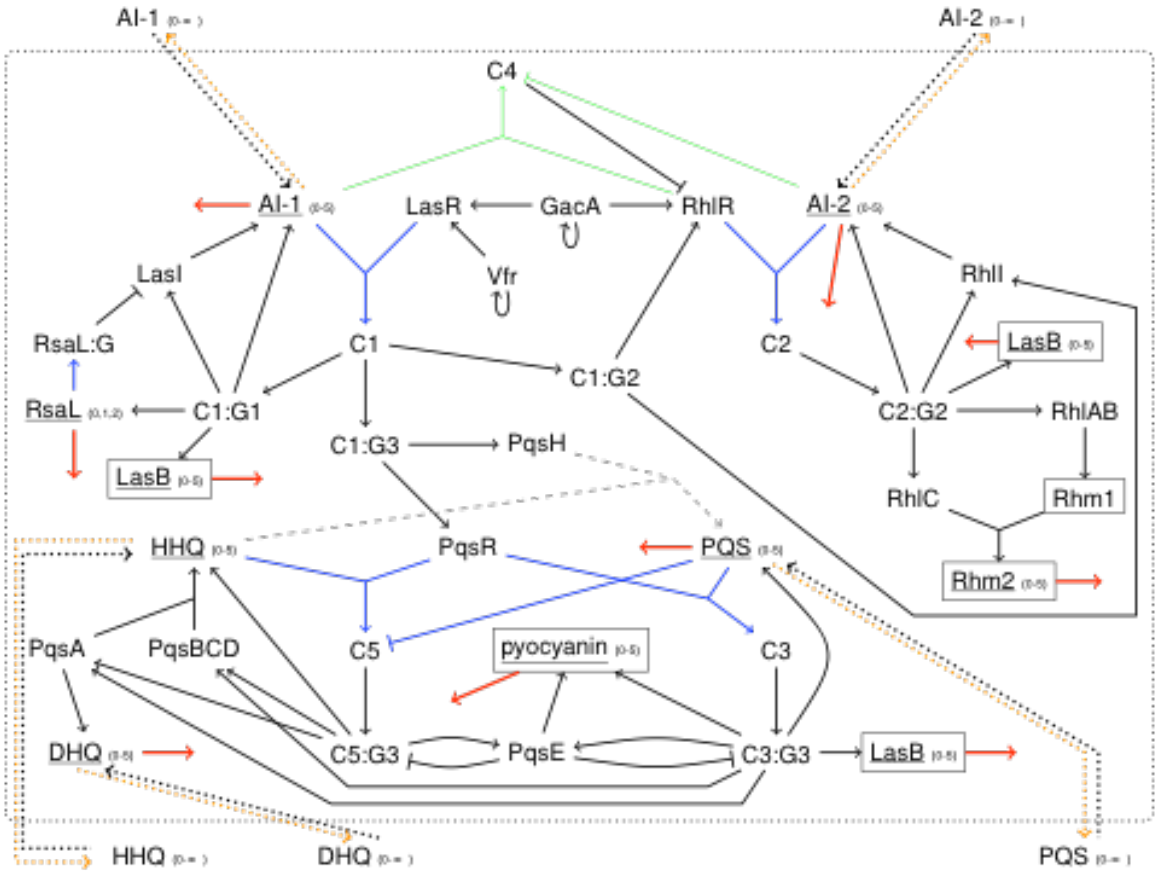
- Wenn man die QS-Maschinerie selektiv durch Inhibitoren der Signalübermittlung hemmt, kann dadurch evtl. die Entstehung von Resistenzmutationen vermeiden

Ziel: entwickle einfaches Computermodell, das die Einflüsse von inhibierenden Molekülen und Resistenzmutationen modelliert.



Schaadt, Steinbach, Helms
BMC Systems Biol. (2013)

QS-Netwerk als generalisierte Boole‘sche Topologie



schwarze Reaktion = Schwellenwert ist 1

blaue Reaktion = Zustand des unterstrichenen Knoten muss ≥ 2 sein;

orange Reaktion = Zustand des unterstrichenen Knoten muss ≥ 3 sein;

dünne **grüne** Reaktion = Zustand des unterstrichenen Knoten muss ≥ 4 sein;

Zahlen : mögliche Zustände des Knoten

Gepunktete Pfeile: Transportprozesse

dicke **rote** Reaktion : passiert nach einer bestimmten Anzahl an Zeitschritten (Abbau).

Gestrichelte **graue** Pfeile : Reaktionen, die zufällig mit einer bestimmten Wahrscheinlichkeit passieren.

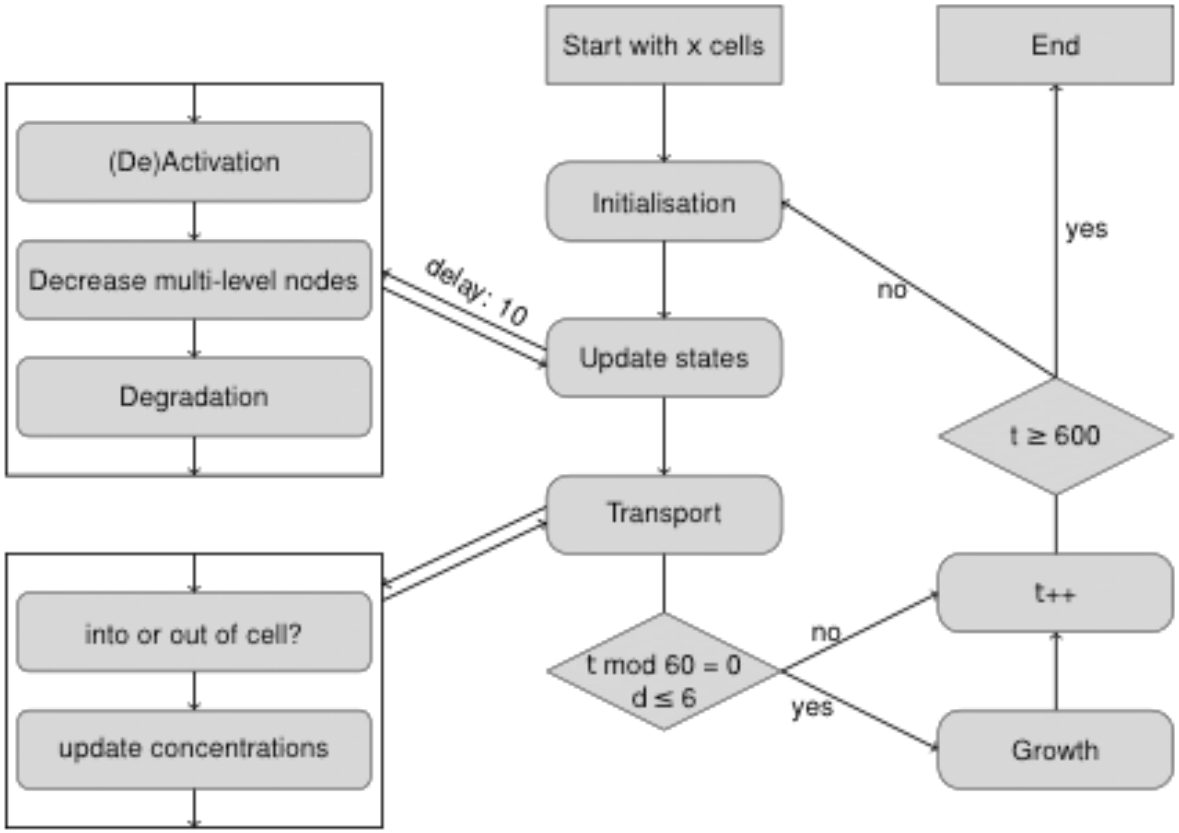
Mit C bezeichnete Knoten sind Komplexe eines Autoinducermoleküls und eines Rezeptors,
C:G ist ein an ein Operon gebundener Komplex.

Schaadt, Steinbach, Helms BMC Systems Biol. (2013)

Reaktionen im QS-System

Reaction	Type	Reference
GacA → LasR	activation, transcription + translation	Reimmann et al. (1997) Mol. Microbiol.
GacA → RhlR	activation, transcription + translation	Reimmann et al. (1997) Mol. Microbiol.
Vfr → LasR	activation, transcription + translation	Albus et al. (1997) J Bacteriol.
AI-1 + LasR → C1	association	Seed et al. (1995) J Bacteriol.
C1 → C1:G1	activation	Seed et al. (1995) J Bacteriol.
C1:G1 → LasI	transcription + translation	Seed et al. (1995) J Bacteriol.
C1:G1 → RsaL	transcription + translation	de Kievit et al. (1999) J Bacteriol.
RsaL → RsaL:G	activation	de Kievit et al. (1999) J Bacteriol.
RsaL:G −| LasI	inhibition	de Kievit et al. (1999) J Bacteriol.
LasI → AI-1	enzymatic reaction (formation)	Passador et al. (1993) Science
AI-1 + RhlR → C4	association	Pesci et al. (1997) J Bacteriol.
C1 → C1:G2	activation	Pesci et al. (1997) J Bacteriol.
C1:G2 → RhlR	transcription + translation	Pesci et al. (1997) J Bacteriol.
AI-2 −| C4	blocking	Pesci et al. (1997) J Bacteriol.
C1:G2 → RhlI	transcription + translation	- - - - -
AI-2 + RhlR → C2	association	Ochsner and Reiser (1995) PNAS
C2 → C2:G2	activation	Ochsner and Reiser (1995) PNAS
C2:G2 → RhlI	transcription + translation	Ochsner and Reiser (1995) PNAS
RhlI → AI-2	enzymatic reaction (formation)	Ochsner and Reiser (1995) PNAS
C1 → C1:G3	activation	Rampioni et al. (2010) Environ. Microbiol.
C1:G3 → PqsR	transcription + translation	Rampioni et al. (2010) Environ. Microbiol.
C1:G3 → PqsH	transcription + translation	Rampioni et al. (2010) Environ. Microbiol.
HHQ + PqsH → PQS	enzymatic reaction (formation)	Deziel et al. (2004) PNAS
PQS + PqsR → C3	association	Deziel et al. (2004) PNAS
C3 → C3:G3	activation	Deziel et al. (2004) PNAS
C3:G3 → PqsABCDE	transcription + translation	Deziel et al. (2004) PNAS
PqsA + PqsBCD → HHQ	enzymatic reaction (formation)	Deziel et al. (2004) PNAS
HHQ + PqsR → C5	association	Xiao et al. (2006) Mol. Microbiol.
C5 → C5:G3	activation	Xiao et al. (2006) Mol. Microbiol.
C5:G3 → PqsABCDE	transcription + translation	Xiao et al. (2006) Mol. Microbiol.
PqsA → DHQ	enzymatic reaction (formation)	

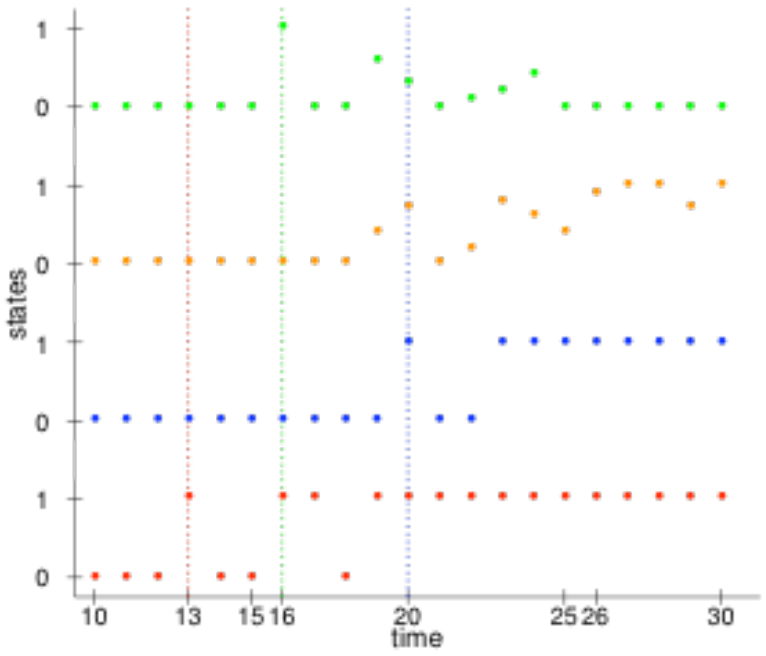
Update-Workflow für Zustand des Netzwerks



Beispieltrajektorie

time step	HHQ	PQS	C3	C5	C3:G3	C5:G3	PqsA	PqsBCD	PqsE
10	0	0	0	0	1	1	0	0	0
11	1	1	0	0	0	0	1	1	1
12	2	1	0	0	0	0	0	0	0
13	2	1	0	0	0	0	0	0	0
14	2	1	0	0	0	0	0	0	0
15	2	1	0	0	0	0	0	0	0
16	0	2	0	1	0	0	0	0	0
17	0	2	0	0	0	1	0	0	0
18	1	2	0	0	0	0	1	1	1
19	2	1	1	0	0	0	0	0	0
20	0	1	0	1	1	0	0	0	0
21	0	2	0	0	0	1	1	1	1
22	2	1	1	0	0	0	1	1	1
23	2	1	0	1	0	0	0	0	0
24	0	2	0	1	0	1	0	0	0
25	1	1	1	0	0	1	1	1	1
26	2	2	0	0	0	0	1	1	1
27	1	2	1	0	0	0	0	0	0
28	0	2	1	0	1	0	0	0	0
29	0	2	1	0	1	0	1	1	1
30	1	2	1	0	0	0	1	1	1
31	1	2	1	0	0	0	0	0	0
32	0	2	1	0	1	0	0	0	0

Simulationsstart



Für minimal notwendige Startbedingungen findet man:

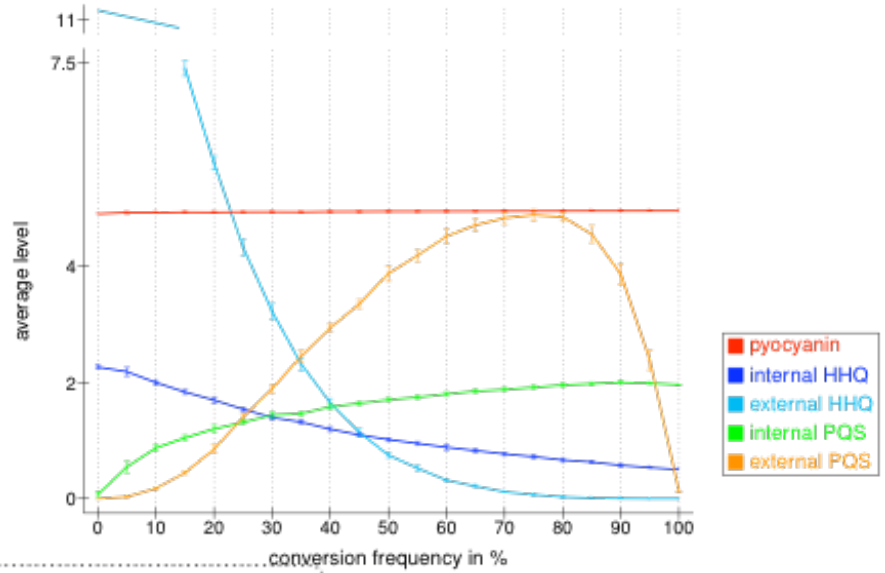
Rot: erster Komplex aus AI–1 und LasR

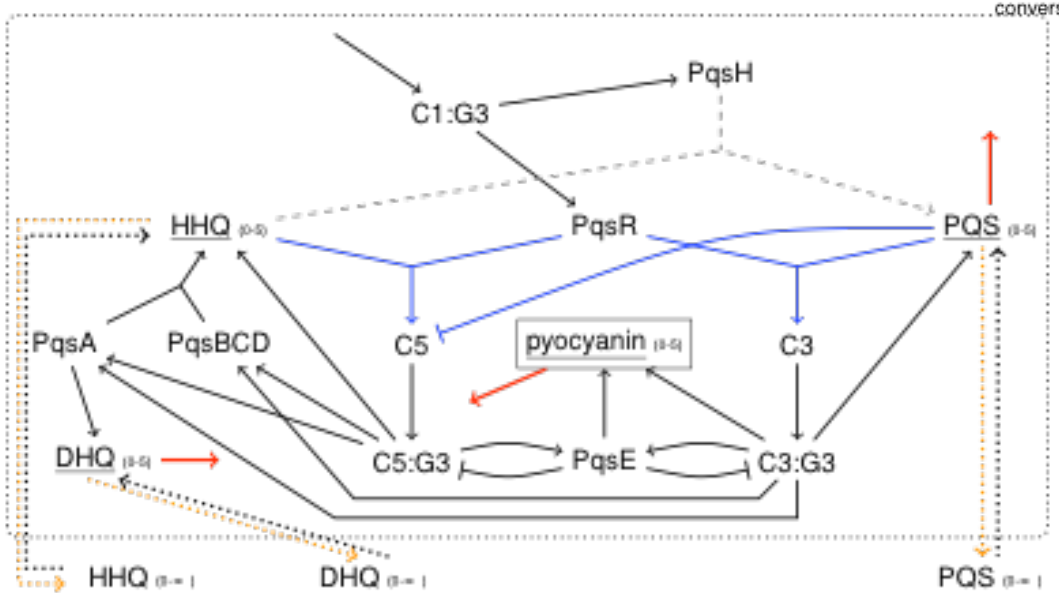
Blau: zweiter Komplex aus AI–2 und RhlR.

Oranger: Komplex C3 des *pqs*-Systems zwischen PQS und PqsR

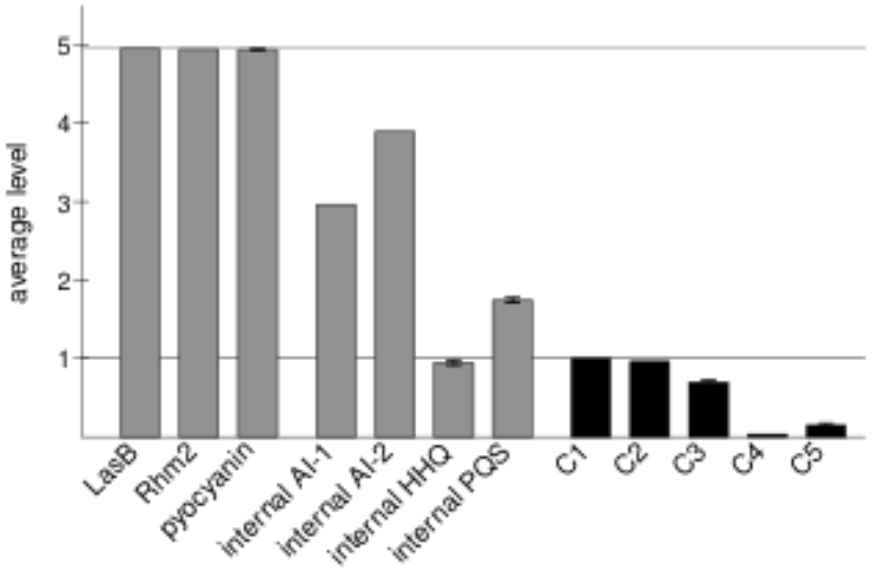
Grün: Komplex C5 aus HHQ und PqsR.

Auswirkung der PQS-Produktionsrate



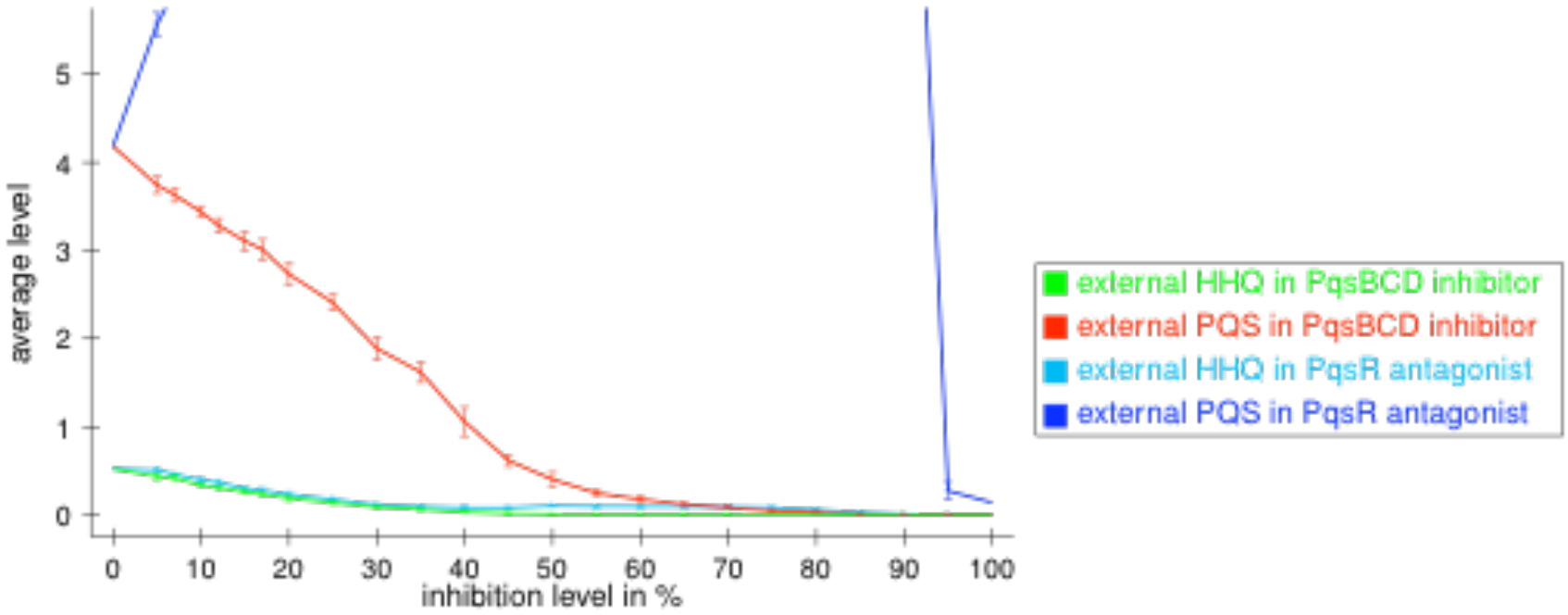


Verhalten des Wildtyps

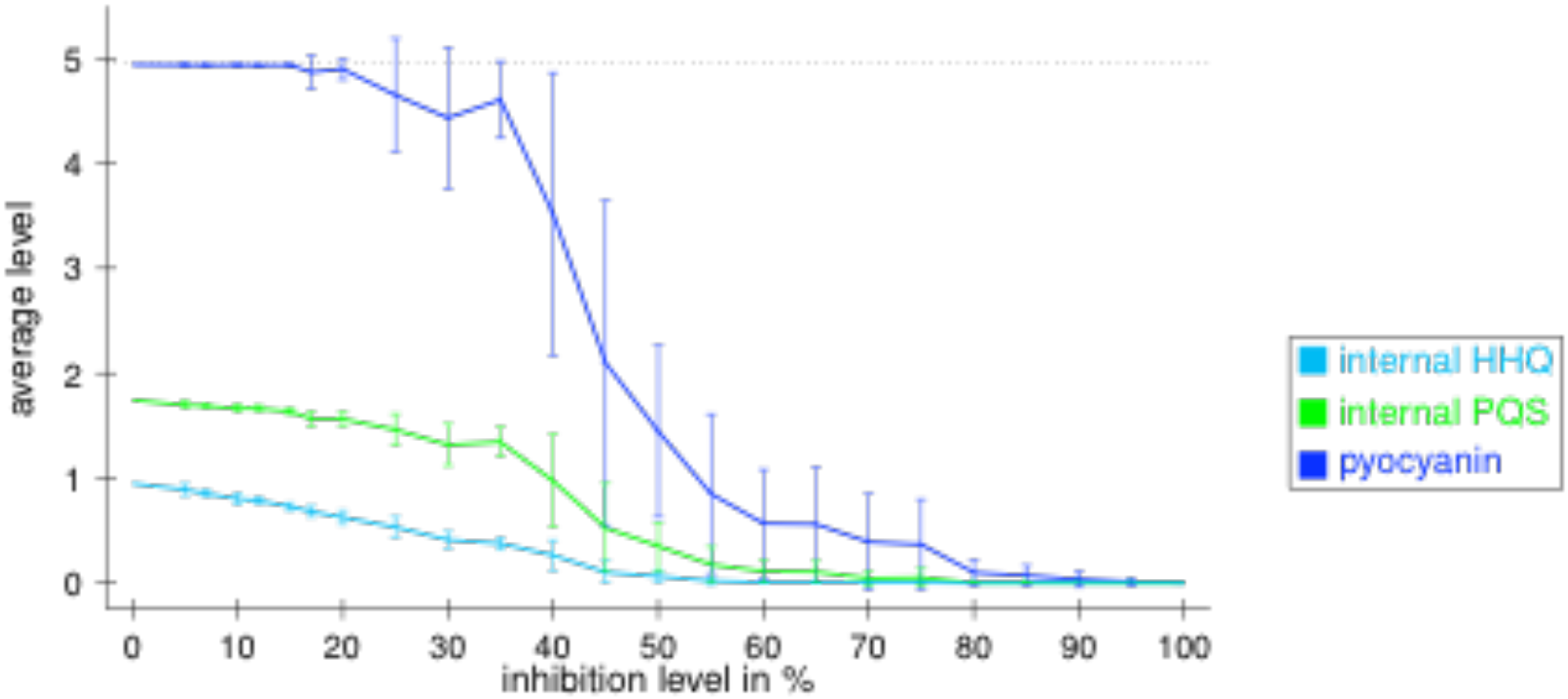


Linien: maximal erreichbare Werte

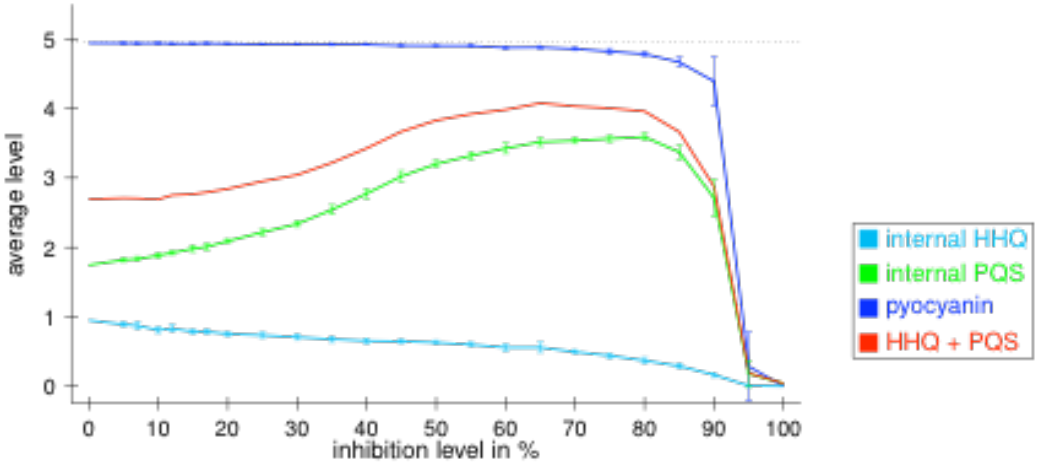
Simulierte PQS- und Pyocyanin-Levels für Wildtyp und Knock–out Mutanten



Externe HHQ und PQS-Level für Wildtyp und verschiedene Inhibitionslevel der zugefügten Inhibitoren



Inhibierung des PqsBCD-Enzyms mit verschiedener Stärke



Simuliere (künstlich) modifizierte Netzwerke mit zusätzlichen Reaktionen

Table 2 Behavior of updated networks

Network	Used reactions	Results			
		PqsA⁻–PqsBCD⁻	PqsE⁻	PqsR antagonists	PqsBCD inhibitors
N1	R1, R2	- - - - -	Pyocyanin	Pyocyanin	- - - - -
N2	R18	- - - - -	- - - - -	Pyocyanin	- - - - -
N3	R19	- - - - -	- - - - -	Pyocyanin	PQS
N4	R20, R2	- - - - -	Pyocyanin	Pyocyanin	- - - - -
N5	R1, R2, R4	Pyocyanin	Pyocyanin	Pyocyanin	- - - - -
N6	R18, R4	Pyocyanin	Pyocyanin	Pyocyanin	- - - - -
N7	R19, R4	Pyocyanin	Pyocyanin	Pyocyanin	PQS
N8	R20, R2, R4	Pyocyanin	Pyocyanin	Pyocyanin	- - - - -
N9	R21	- - - - -	- - - - -	- - - - -	- - - - -
N10	R22	- - - - -	- - - - -	- - - - -	PQS
N11	R23, R2	- - - - -	Pyocyanin	Pyocyanin	- - - - -
N12	R24, R2	- - - - -	Pyocyanin	Pyocyanin	- - - - -

Entries denote deviations (too high levels) from the expected behavior.

Zusammenfassung

- Regelbasierte Simulationen reproduzierten das aus der Literatur bekannte Verhalten der externen Levels der Autoinducer.
- Die Zugabe von PqsBCD-Inhibitoren reduziert die externen Levels von HHQ und PQS deutlich. Die Stäre der Abnahme hängt davon ab, wie hoch der Inhibitionslevel ist.
- Der Pyocyanin-Pfad scheint jedoch unvollständig zu sein.

- Simulationen mit modifizierten Netzwerken deuten auf zusätzliche regulatorische Interaktionen hin. Dann kann man die experimentelle Befunde besser nachstellen.