

V10 metabolische Netzwerke / Protein-Interaktionsnetzwerke

- 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 V10:**
 - (1) stöchiometrische Matrix
 - (2) Flux Balance Analysis - Methode
 - (3) Elementary Flux Mode – Analyse (anschaulich + Anwendung)
 - (4) Aufbau und Visualisierung von Protein-Protein-Interaktionsnetzwerken

Analyse von Stoffwechselwegen: Beispiel *E. coli*

verwende Daten aus Datenbank **EcoCyc**

(siehe auch **MetaCyc**: Stoffwechsel von > 150 Organismen)

EcoCyc enthält 905 Reaktionen für *E.coli*

davon gehören 161 nicht zum Stoffwechsel kleiner Moleküle, z.B. DNA Replikation, von den verbleibenden 744 wurden 569 mindestens einem Pfad zugeordnet

Dagegen gibt es 607 Enzyme.

Es gibt also keine 1:1 Zuordnung zwischen Enzymen und Reaktionen, denn

- (1) manche Enzyme katalysieren mehrere Reaktionen,
und manche Reaktionen werden von mehreren Enzymen katalysiert
- (2) nicht zu allen Reaktionen sind die Enzyme bekannt, die sie katalysieren.

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

Beispiel: Stoffwechsel von *E. coli*

Die 744 Reaktionen enthalten
791 verschiedene Substrate

Im Mittel enthält jede Reaktion
4 Substrate.

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

Occurrence	Name of metabolite
205	H ₂ O
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 ₂ O ₂

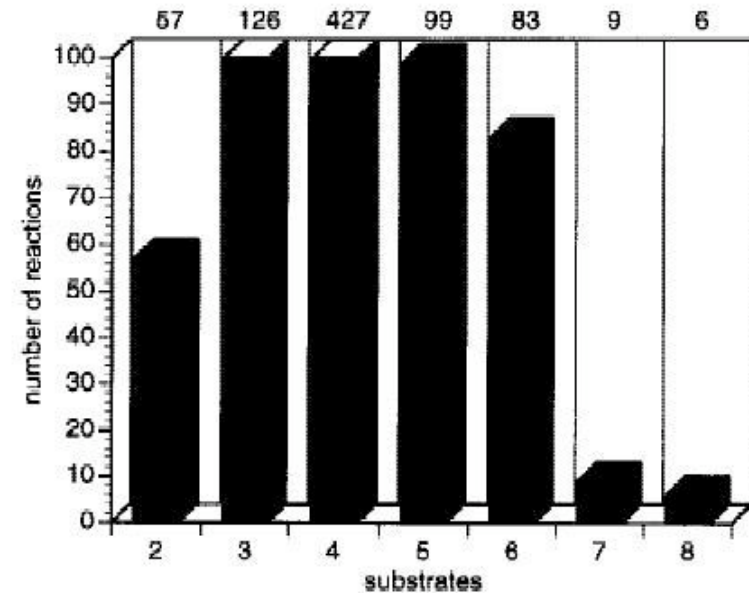


Figure 4 Diagram showing the number of reactions containing varying numbers of substrates (reactants plus products).

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

Beispiel: Stoffwechsel von *E. coli*

EcoCyc enthält 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.

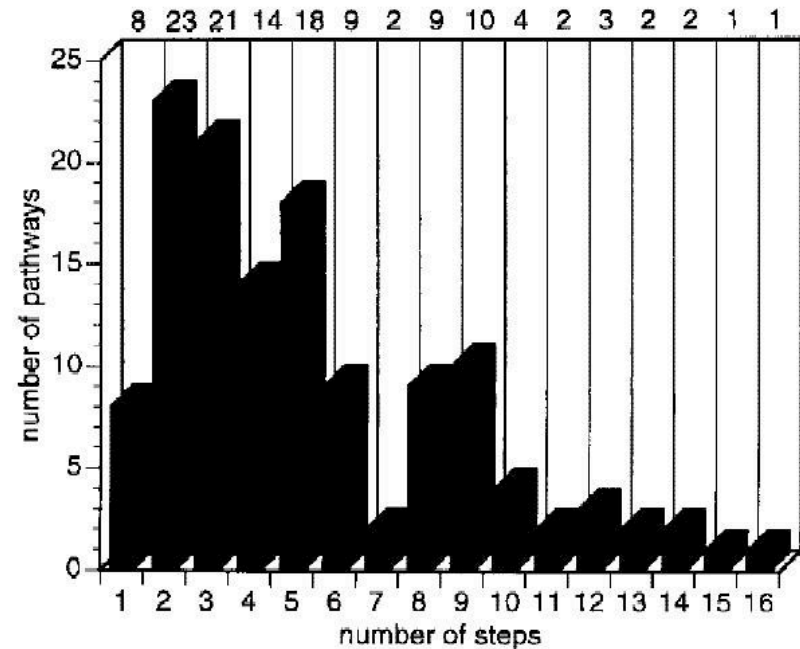


Figure 5 Length distribution of EcoCyc pathways; two pathways are not included because the number of steps is not known.

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?

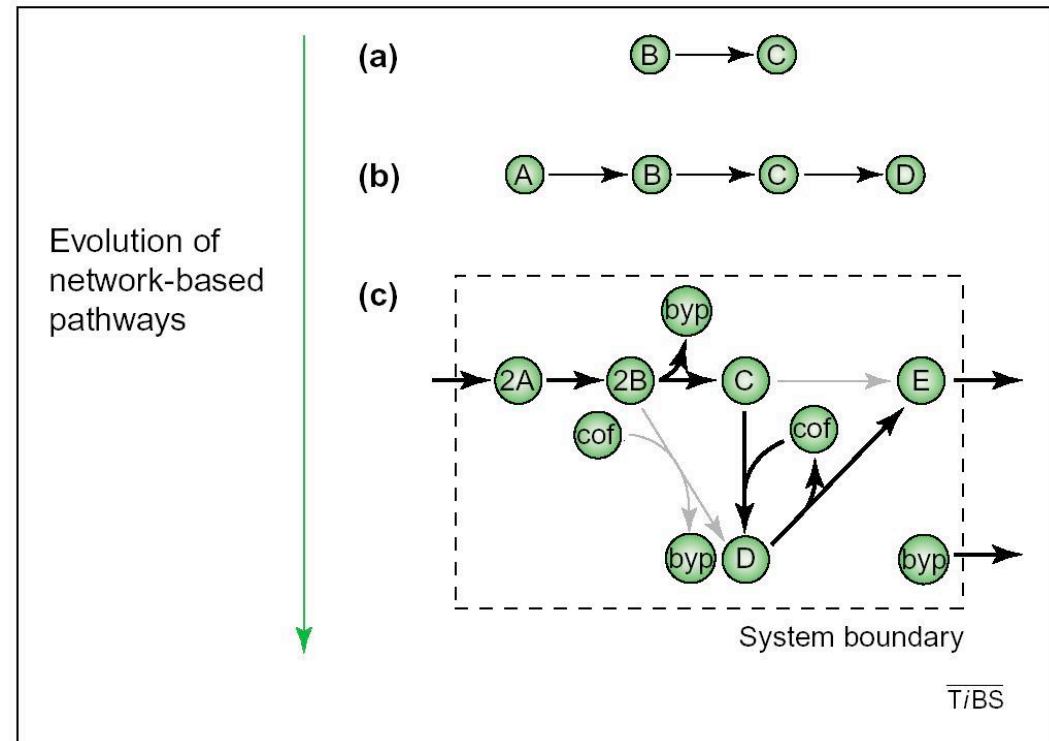
- Reaktionen, Enzyme und Substrate gehören oft zu mehreren Pfaden.
- Die Einteilung in einzelne Stoffwechsel-Pfade ist nicht immer eindeutig

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.

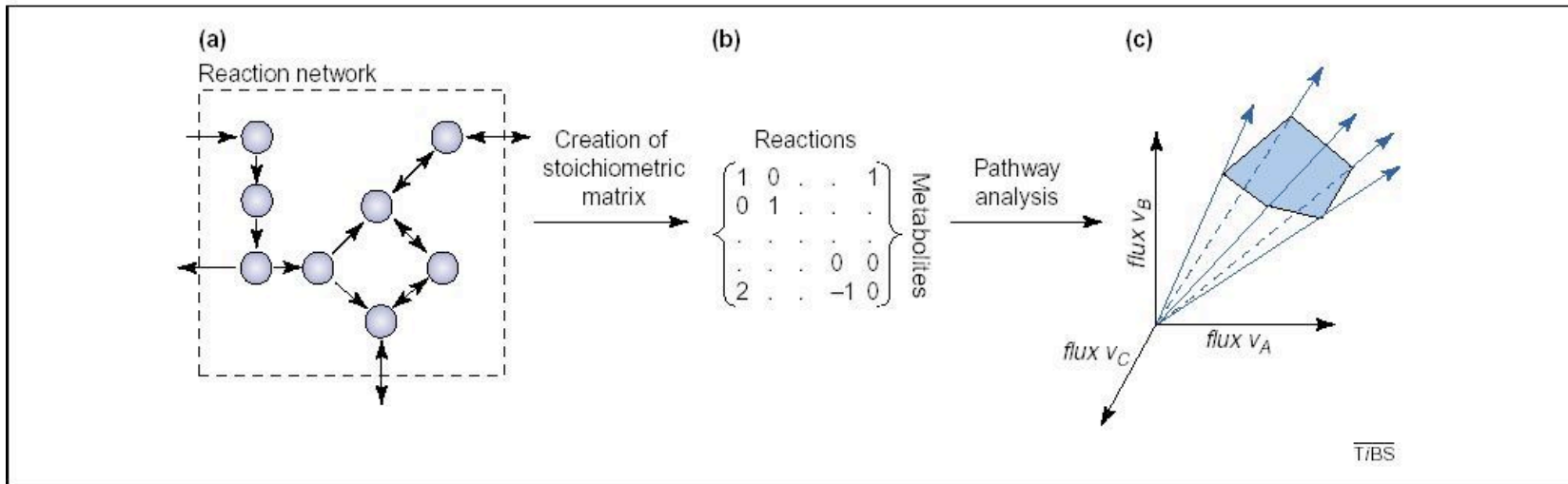
Man kann dadurch Metabolismen verschiedener Organismen vergleichen.

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.

Beschreibung vernetzter metabolischer Pfade



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

(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.

(c) Mögliche Zustände der Zelle aufgrund dieser Matrix werden mit Techniken wie „**elementary modes**“ oder „**extreme pathways**“ identifiziert. Die möglichen Zustände liegen innerhalb eines Konus im durch die verschiedenen Flüsse aufgespannten Koordinatensystem.

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:

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

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.

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 .

Um den „wahren“ biologischen Fluss zu finden (\rightarrow e.g. Heinzle/UdS) benötigt man zusätzliche Informationen, z.B. aus Experimenten mit C^{13} -Markierung.

Alternativ kann man Grenzwerte $\alpha_i \leq v_i \leq \beta_i$

für den Minimal- und Maximalbetrag jedes einzelnen metabolischen Flusses vorschreiben.

Der Durchschnitt des Nullraums und des durch diese linearen Ungleichungen eingeschränkten Raums definiert dann ein Volumen im hochdimensionalen Flussraum, das alle möglichen Flussverteilungen enthält.

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*

Zoom Out

Table 1. The genes included in the *E. coli* metabolic genotype (21)

Central metabolism (EMP, PPP, TCA cycle, electron transport)	<i>aceA, aceB, aceE, aceF, ackA, ackB, acnB, acs, adhE, agp, appB, appC, atpA, atpB, atpC, atpD, atpE, atpF, atpG, atpH, atpI, cydA, cydB, cydC, cydD, cyoA, cyoB, cyoC, cyoD, dkl, eno, fba, fbp, fchF, fdnG, fdnH, fdnI, fdoG, fdoH, fdoI, frcA, frcB, frcC, frcD, fumA, fumB, fumC, galM, gbaA, gapC_1, gapC_2, glcB, glgA, glgC, glgP, glk, glpA, glpB, glpC, glpD, gltA, gnd, gpmA, gpmB, hyaA, hyaB, hyaC, hyaB, hycB, hycE, hycF, hycG, icdA, ictD, ldhA, lpcA, malP, meth, ndh, nuoA, nuoB, nuoE, nuoF, nuoG, nuoH, nuol, nuol, 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, sudA, sucB, sucC, sucD, talB, tktA, tktB, tpiA, trxB, zwf, pgi (30), maeB (30)</i>
Alternative carbon source	<i>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</i>
Amino acid metabolism	<i>adi, aktH, 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, cysK, cysM, cysN, dacA, dadX, dapA, dapB, dapD, dapE, dapF, dacB, gabT, gadA, gabB, gadC, gadH, 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, kbI, 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, tdk, thrA, thrB, thrC, trnA, trpA, trpB, trpC, trpD, trpE, tyrA, tyrB, yjiG, yjiH, alaB (42), dapC (43), pat (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)</i>
Purine & pyrimidine metabolism	<i>adk, adl, 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, trnK, udk, udp, upp, ushA, xapA, yicP, CMP glycosylase (48)</i>
Vitamin & cofactor metabolism	<i>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, pdxI, 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-pentothenate adenylyltransferase (50), dephosphoCoA kinase (50), NMN glycohydrolase (49)</i>
Lipid metabolism	<i>accA, accB, accD, atoR, coh, cdsA, ck, dgkA, fabD, fabH, fabB, gpaA, ispA, ispB, pggB, pgsA, pxd, pssA, pppA (53)</i>
Cell wall metabolism	<i>dclA, dclB, galF, galU, glmS, glmU, htrB, kdsA, kdsB, kdtA, lpxA, lpxB, lpxC, lpxD, mraY, msbB, murA, murB, murC, murD, murE, murF, murG, murI, rfaC, rfaD, rfaE, 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)</i>
Transport processes	<i>araE, araF, araG, araH, argT, aroP, artI, artJ, artM, artP, artQ, brnQ, cadB, chaA, chaB, chaC, cmtA, cmtB, codB, crr, cycA, cysA, cysB, cysT, cysU, cysW, cysZ, dcta, dcaA, dcuB, dppA, dppB, dppC, dppD, dppF, fadL, fcaA, fruA, fruB, fucP, gabP, galP, galV, gatB, gatC, glnH, glnP, glnQ, glpF, glpT, gltI, gltK, gltL, gltP, gltS, gntT, gpt, hisI, hisM, hisP, hisQ, hpt, kdpA, kdpB, kdpC, kgtP, lacY, lamB, livF, livG, livH, livI, livK, livM, lkpP, lypP, malE, malF, malG, malK, malX, manX, manY, manZ, melB, mgIA, mgIB, mgIC, mtIA, mt, 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, ptaA, ptaB, ptaC, ptaD, ptaE, ptaF, ptaG, ptaH, ptaI, ptaJ, ptaK, ptaL, ptaM, ptaN, ptaO, ptaP, ptaQ, ptaR, ptaS, ptaT, ptaU, ptaV, ptaW, ptaX, ptaY, ptaZ, ptaAA, ptaAB, ptaAC, ptaAD, ptaAE, ptaAF, ptaAG, ptaAH, ptaAI, ptaAJ, ptaAK, ptaAL, ptaAM, ptaAN, ptaAO, ptaAP, ptaAQ, ptaAR, ptaAS, ptaAT, ptaAU, ptaAV, ptaAW, ptaAX, ptaAY, ptaAZ, ptaBA, ptaBB, ptaBC, ptaBD, ptaBE, ptaBF, ptaBG, ptaBH, ptaBI, ptaBJ, ptaBK, ptaBL, ptaBM, ptaBN, ptaBO, ptaBP, ptaBQ, ptaBR, ptaBS, ptaBT, ptaBU, ptaBV, ptaBW, ptaBX, ptaBY, ptaBZ, ptaCA, ptaCB, ptaCC, ptaCD, ptaCE, ptaCF, ptaCG, ptaCH, ptaCI, ptaCJ, ptaCK, ptaCL, ptaCM, ptaCN, ptaCO, ptaCP, ptaCQ, ptaCR, ptaCS, ptaCT, ptaCU, ptaCV, ptaCW, ptaCX, ptaCY, ptaCZ, ptaDA, ptaDB, ptaDC, ptaDD, ptaDE, ptaDF, ptaDG, ptaDH, ptaDI, ptaDJ, ptaDK, ptaDL, ptaDM, ptaDN, ptaDO, ptaDP, ptaDQ, ptaDR, ptaDS, ptaDT, ptaDU, ptaDV, ptaDW, ptaDX, ptaDY, ptaDZ, ptaEA, ptaEB, ptaEC, ptaED, ptaEE, ptaEF, ptaEG, ptaEH, ptaEI, ptaEJ, ptaEK, ptaEL, ptaEM, ptaEN, ptaEO, ptaEP, ptaEQ, ptaER, ptaES, ptaET, ptaEU, ptaEV, ptaEW, ptaEX, ptaEY, ptaEZ, ptaFA, ptaFB, ptaFC, ptaFD, ptaFE, ptaFF, ptaFG, ptaFH, ptaFI, ptaFJ, ptaFK, ptaFL, ptaFM, ptaFN, ptaFO, ptaFP, ptaFQ, ptaFR, ptaFS, ptaFT, ptaFU, ptaFV, ptaFW, ptaFX, ptaFY, ptaFZ, ptaGA, ptaGB, ptaGC, ptaGD, ptaGE, ptaGF, ptaGG, ptaGH, ptaGI, ptaGJ, ptaGK, ptaGL, ptaGM, ptaGN, ptaGO, ptaGP, ptaGQ, ptaGR, ptaGS, ptaGT, ptaGU, ptaGV, ptaGW, ptaGX, ptaGY, ptaGZ, ptaHA, ptaHB, ptaHC, ptaHD, ptaHE, ptaHF, ptaHG, ptaHH, ptaHI, ptaHJ, ptaHK, ptaHL, ptaHM, ptaHN, ptaHO, ptaHP, ptaHQ, ptaHR, ptaHS, ptaHT, ptaHU, ptaHV, ptaHW, ptaHX, ptaHY, ptaHZ, ptaIA, ptaIB, ptaIC, ptaID, ptaIE, ptaIF, ptaIG, ptaIH, ptaII, ptaIJ, ptaIK, ptaIL, ptaIM, ptaIN, ptaIO, ptaIP, ptaIQ, ptaIR, ptaIS, ptaIT, ptaIU, ptaIV, ptaIW, ptaIX, ptaIY, ptaIZ, ptaJA, ptaJB, ptaJC, ptaJD, ptaJE, ptaJF, ptaJG, ptaJH, ptaJI, ptaJJ, ptaJK, ptaJL, ptaJM, ptaJN, ptaJO, ptaJP, ptaJQ, ptaJR, ptaJS, ptaJT, ptaJU, ptaJV, ptaJW, ptaJX, ptaJY, ptaJZ, ptaKA, ptaKB, ptaKC, ptaKD, ptaKE, ptaKF, ptaKG, ptaKH, ptaKI, ptaKJ, ptaKK, ptaKL, ptaKM, ptaKN, ptaKO, ptaKP, ptaKQ, ptaKR, ptaKS, ptaKT, ptaKU, ptaKV, ptaKW, ptaKX, ptaKY, ptaKZ, ptaLA, ptaLB, ptaLC, ptaLD, ptaLE, ptaLF, ptaLG, ptaLH, ptaLI, ptaLJ, ptaLK, ptaLL, ptaLM, ptaLN, ptaLO, ptaLP, ptaLQ, ptaLR, ptaLS, ptaLT, ptaLU, ptaLV, ptaLW, ptaLX, ptaLY, ptaLZ, ptaMA, ptaMB, ptaMC, ptaMD, ptaME, ptaMF, ptaMG, ptaMH, ptaMI, ptaMJ, ptaMK, ptaML, ptaMN, ptaMO, ptaMP, ptaMQ, ptaMR, ptaMS, ptaMT, ptaMU, ptaMV, ptaMW, ptaMX, ptaMY, ptaMZ, ptaNA, ptaNB, ptaNC, ptaND, ptaNE, ptaNF, ptaNG, ptaNH, ptaNI, ptaNJ, ptaNK, ptaNL, ptaNM, ptaNN, ptaNO, ptaNP, ptaNQ, ptaNR, ptaNS, ptaNT, ptaNU, ptaNV, ptaNW, ptaNX, ptaNY, ptaNZ, ptaOA, ptaOB, ptaOC, ptaOD, ptaOE, ptaOF, ptaOG, ptaOH, ptaOI, ptaOJ, ptaOK, ptaOL, ptaOM, ptaON, ptaOO, ptaOP, ptaOQ, ptaOR, ptaOS, ptaOT, ptaOU, ptaOV, ptaOW, ptaOX, ptaOY, ptaOZ, ptaPA, ptaPB, ptaPC, ptaPD, ptaPE, ptaPF, ptaPG, ptaPH, ptaPI, ptaPJ, ptaPK, ptaPL, ptaPM, ptaPN, ptaPO, ptaPP, ptaPQ, ptaPR, ptaPS, ptaPT, ptaPU, ptaPV, ptaPW, ptaPX, ptaPY, ptaPZ, ptaQA, ptaQB, ptaQC, ptaQD, ptaQE, ptaQF, ptaQG, ptaQH, ptaQI, ptaQJ, ptaQK, ptaQL, ptaQM, ptaQN, ptaQO, ptaQP, ptaQQ, ptaQR, ptaQS, ptaQT, ptaQU, ptaQV, ptaQW, ptaQX, ptaQY, ptaQZ, ptaRA, ptaRB, ptaRC, ptaRD, ptaRE, ptaRF, ptaRG, ptaRH, ptaRI, ptaRJ, ptaRK, ptaRL, ptaRM, ptaRN, ptaRO, ptaRP, ptaRQ, ptaRR, ptaRS, ptaRT, ptaRU, ptaRV, ptaRW, ptaRX, ptaRY, ptaRZ, ptaSA, ptaSB, ptaSC, ptaSD, ptaSE, ptaSF, ptaSG, ptaSH, ptaSI, ptaSJ, ptaSK, ptaSL, ptaSM, ptaSN, ptaSO, ptaSP, ptaSQ, ptaSR, ptaSS, ptaST, ptaSU, ptaSV, ptaSW, ptaSX, ptaSY, ptaSZ, ptaTA, ptaTB, ptaTC, ptaTD, ptaTE, ptaTF, ptaTG, ptaTH, ptaTI, ptaTJ, ptaTK, ptaTL, ptaTM, ptaTN, ptaTO, ptaTP, ptaTQ, ptaTR, ptaTS, ptaTT, ptaTU, ptaTV, ptaTW, ptaTX, ptaTY, ptaTZ, ptaUA, ptaUB, ptaUC, ptaUD, ptaUE, ptaUF, ptaUG, ptaUH, ptaUI, ptaUJ, ptaUK, ptaUL, ptaUM, ptaUN, ptaUO, ptaUP, ptaUQ, ptaUR, ptaUS, ptaUT, ptaUU, ptaUV, ptaUW, ptaUX, ptaUY, ptaUZ, ptaVA, ptaVB, ptaVC, ptaVD, ptaVE, ptaVF, ptaVG, ptaVH, ptaVI, ptaVJ, ptaVK, ptaVL, ptaVM, ptaVN, ptaVO, ptaVP, ptaVQ, ptaVR, ptaVS, ptaVT, ptaVU, ptaVV, ptaVW, ptaVX, ptaVY, ptaVZ, ptaWA, ptaWB, ptaWC, ptaWD, ptaWE, ptaWF, ptaWG, ptaWH, ptaWI, ptaWJ, ptaWK, ptaWL, ptaWM, ptaWN, ptaWO, ptaWP, ptaWQ, ptaWR, ptaWS, ptaWT, ptaWU, ptaWV, ptaWW, ptaWX, ptaWY, ptaWZ, ptaXA, ptaXB, ptaXC, ptaXD, ptaXE, ptaXF, ptaXG, ptaXH, ptaXI, ptaXJ, ptaXK, ptaXL, ptaXM, ptaXN, ptaXO, ptaXP, ptaXQ, ptaXR, ptaXS, ptaXT, ptaXU, ptaXV, ptaXW, ptaXX, ptaXY, ptaXZ, ptaYA, ptaYB, ptaYC, ptaYD, ptaYE, ptaYF, ptaYG, ptaYH, ptaYI, ptaYJ, ptaYK, ptaYL, ptaYM, ptaYN, ptaYO, ptaYP, ptaYQ, ptaYR, ptaYS, ptaYT, ptaYU, ptaYV, ptaYW, ptaYX, ptaYY, ptaYZ, ptaZA, ptaZB, ptaZC, ptaZD, ptaZE, ptaZF, ptaZG, ptaZH, ptaZI, ptaZJ, ptaZK, ptaZL, ptaZM, ptaZN, ptaZO, ptaZP, ptaZQ, ptaZR, ptaZS, ptaZT, ptaZU, ptaZV, ptaZW, ptaZX, ptaZY, ptaZZ, fruF (56), gntS (57), metD (43), pnuE (49), ser (56)</i>

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 können
(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)

E.coli in silico

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

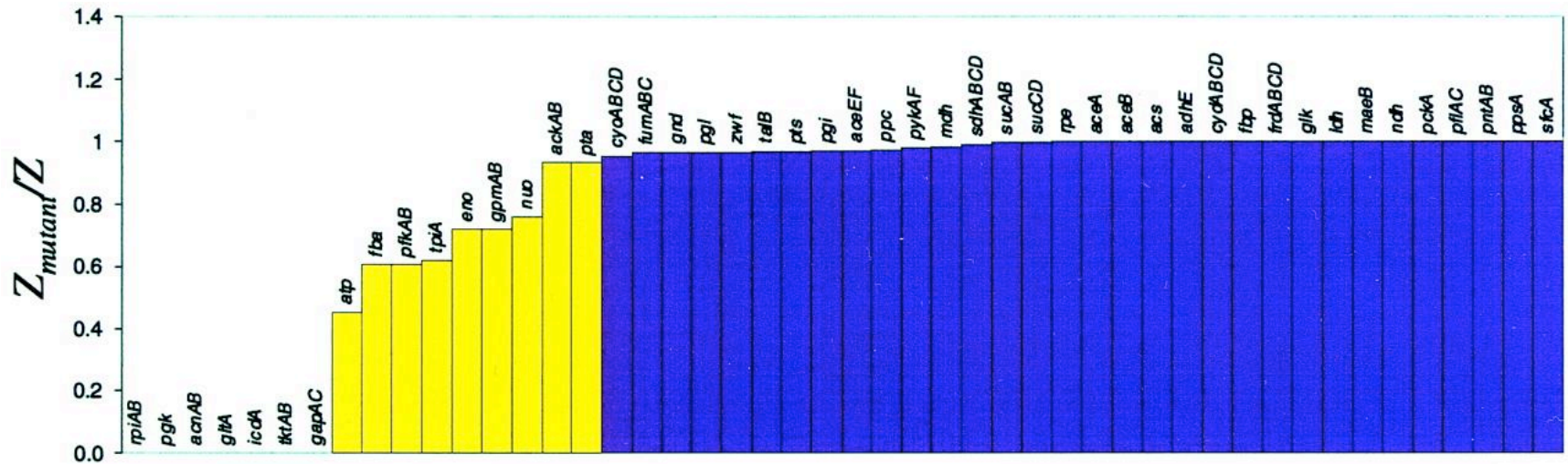
Dafür wird einfach der Fluss durch die entsprechende enzymatische Reaktion auf 0 gesetzt.

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 nennenswerten Effekt auf das *in silico*-Wachstum von *E.coli*.

Edwards & Palsson PNAS 97, 5528 (2000)

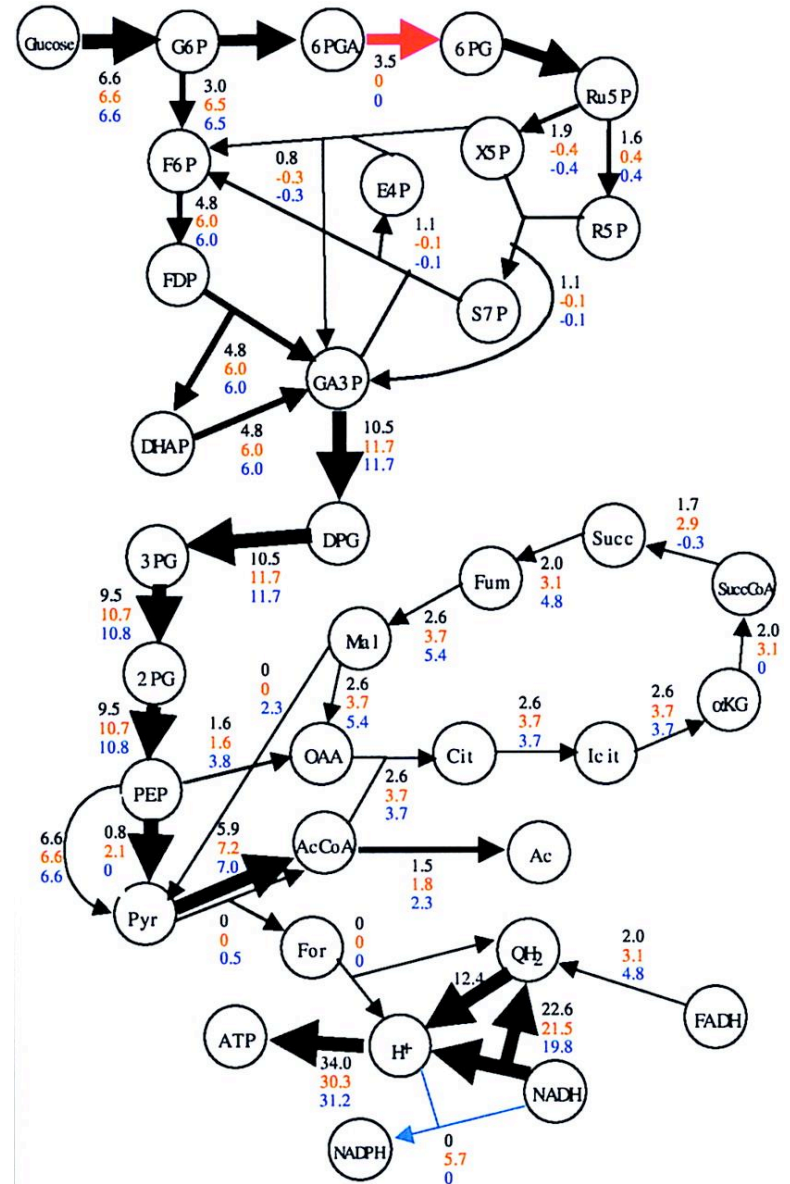
Umlenken der metabolischen Flüsse

(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 Mutantenstamm 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.

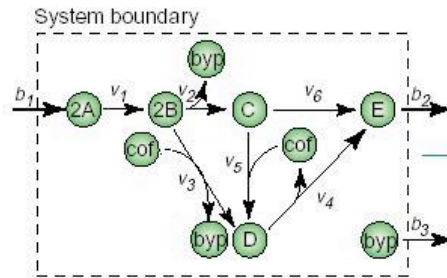
Edwards & Palsson
PNAS 97, 5528 (2000)

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
<i>aceA</i>	+/+		+/+	-/-
<i>aceB</i>				-/-
<i>aceEF*</i>	-/+			
<i>ackA</i>				+/+
● <i>acn</i>	-/-			-/-
<i>acs</i>				+/+
<i>cyd</i>	+/+			
<i>cyo</i>	+/+			
● <i>eno</i> [†]	-/+	-/+	-/-	-/-
● <i>fbaI</i>	-/+			
<i>fbp</i>	+/+	-/-	-/-	-/-
<i>frd</i>	+/+		+/+	+/+
● <i>gap</i>	-/-	-/-	-/-	-/-
<i>glk</i>	+/+			
● <i>gltA</i>	-/-			-/-
<i>gnd</i>	+/+			
<i>idh</i>	-/-			-/-
<i>mdh</i> ^{††}	+/+	+/+	+/+	
<i>ndh</i>	+/+	+/+		
● <i>nuo</i>	+/+	+/+		
● <i>pfk</i> [†]	-/+			
<i>pgi</i> [‡]	+/+	+/-	+/-	
● <i>pgk</i>	-/-	-/-	-/-	-/-
<i>pgl</i>	+/+			
<i>pntAB</i>	+/+	+/+	+/+	
<i>ppc</i> [§]	±/+	-/+	+/+	
<i>pta</i>				+/+
<i>pts</i>	+/+			
<i>pyk</i>	+/+			
● <i>rpi</i>	-/-	-/-	-/-	-/-
<i>sdhABCD</i>	+/+		-/-	-/-
<i>sucAB</i>	+/+		-/+	-/+
● <i>tktAB</i>	-/-			
● <i>tpi</i> ^{**}	-/+	-/-	-/-	-/-
<i>unc</i>	+/+		±/+	-/-
<i>zwf</i>	+/+	+/+	+/+	

Berechnung von Elementaren Flussmoden aus der stöchiometrischen Matrix

Darstellung des Reaktionsnetzwerks mit stöchiometrischer Matrix S.

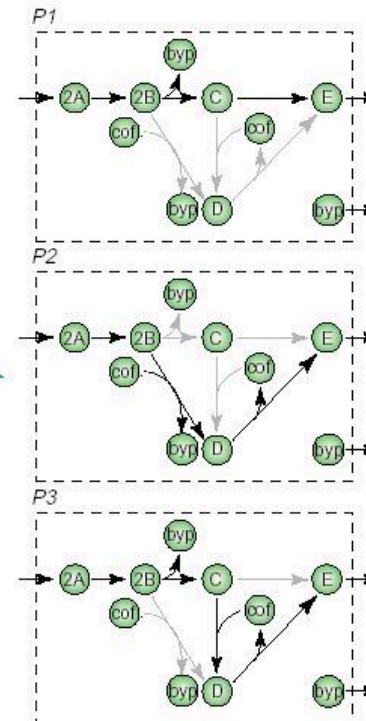


$$S = \begin{array}{cccccc|ccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & b_1 & b_2 & b_3 \\ \hline -1 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ +1 & -2 & +2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & +1 & -1 & +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 & 0 & +1 & 0 & -1 & 0 \\ 0 & +1 & +1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & +1 & -1 & 0 & 0 & 0 & 0 \end{array} \begin{array}{l} A \\ B \\ C \\ D \\ E \\ \text{byp} \\ \text{cof} \end{array}$$

Metabolite

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

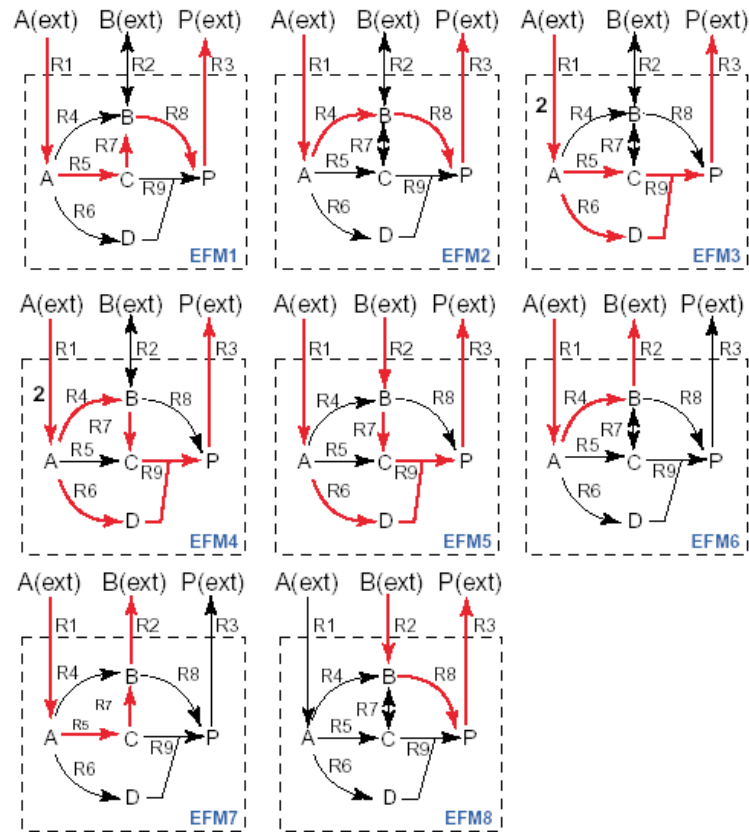
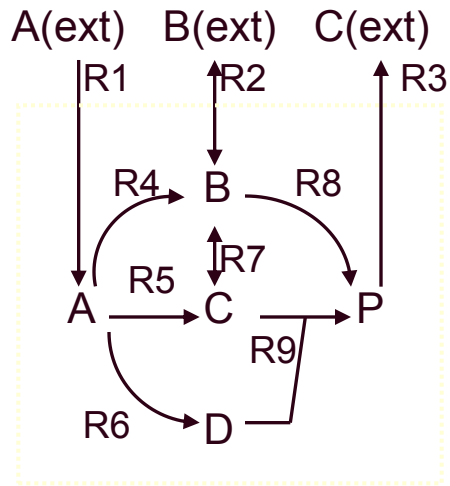
$$P = \begin{array}{ccc|cccccc} P_1 & P_2 & P_3 & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\ \hline 2 & 2 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 2 & 2 & 2 & b_1 & b_2 & b_3 & b_1 & b_2 & b_3 \\ 1 & 1 & 1 & & & & & & \\ 1 & 1 & 1 & & & & & & \end{array}$$



Elementare Flussmoden für dies System.

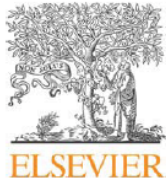
Papin *et al.* TIBS 28, 250 (2003)

Elementare Flussmoden in Modell-Netzwerk



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 Sreenc*

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

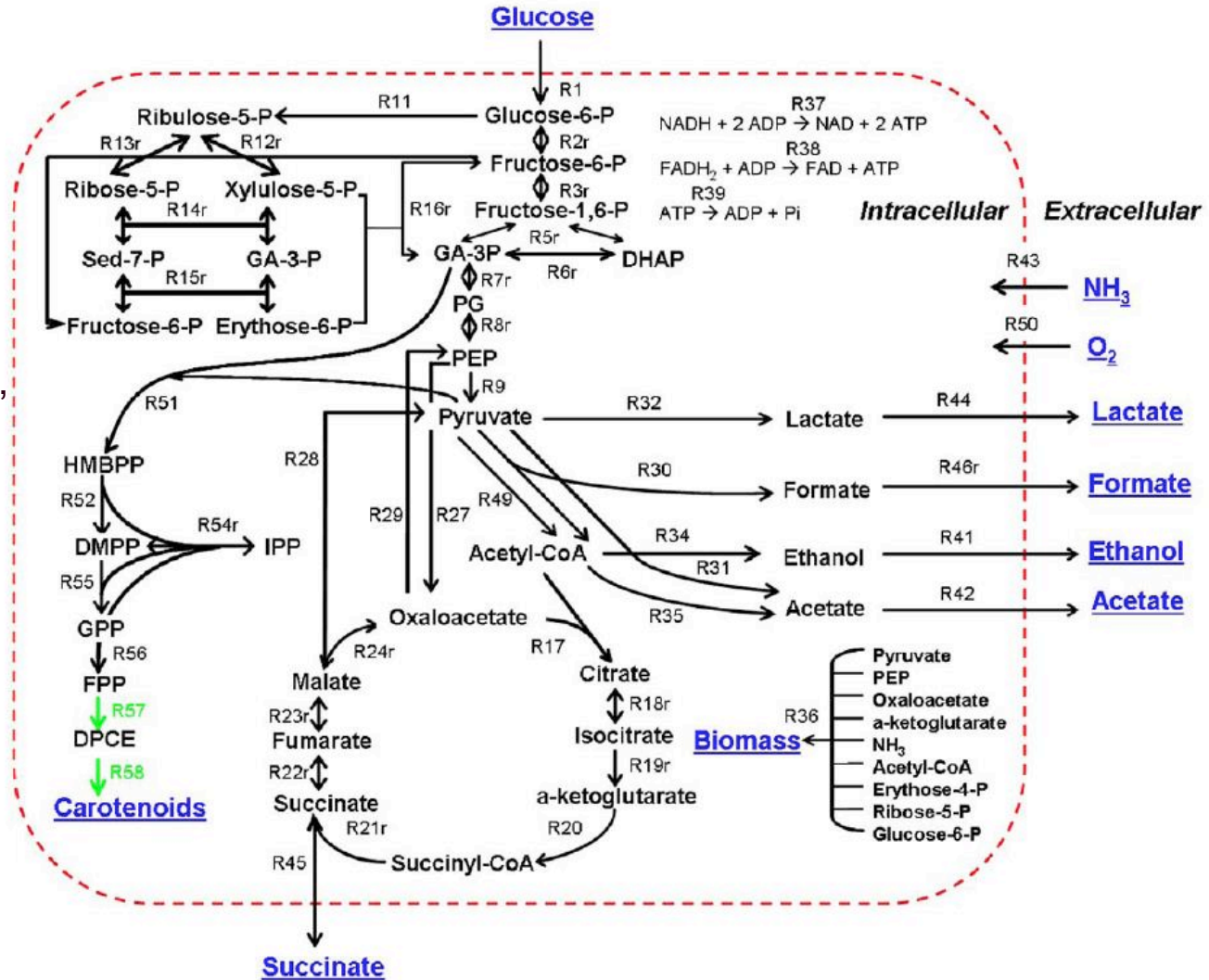
Carotenoide (z.B. DPL und DPA) sind Licht-absorbierende Pigmente, schützen vor UV-Strahlung, regulieren die Fluidität von Membranen, und wirken antioxidierend.

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

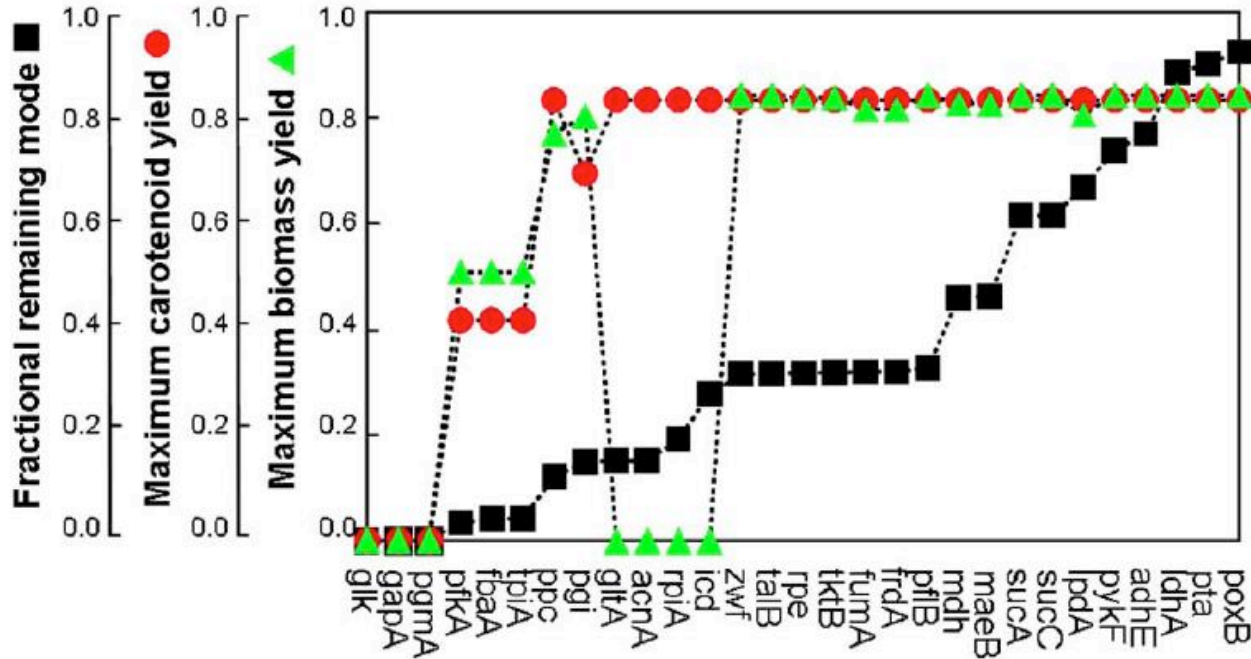
57 Metabolite

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-Knockouts (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 Gendelektionen

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

^a Yield is in mg-diapolycompedioic acid/g-glucose.

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

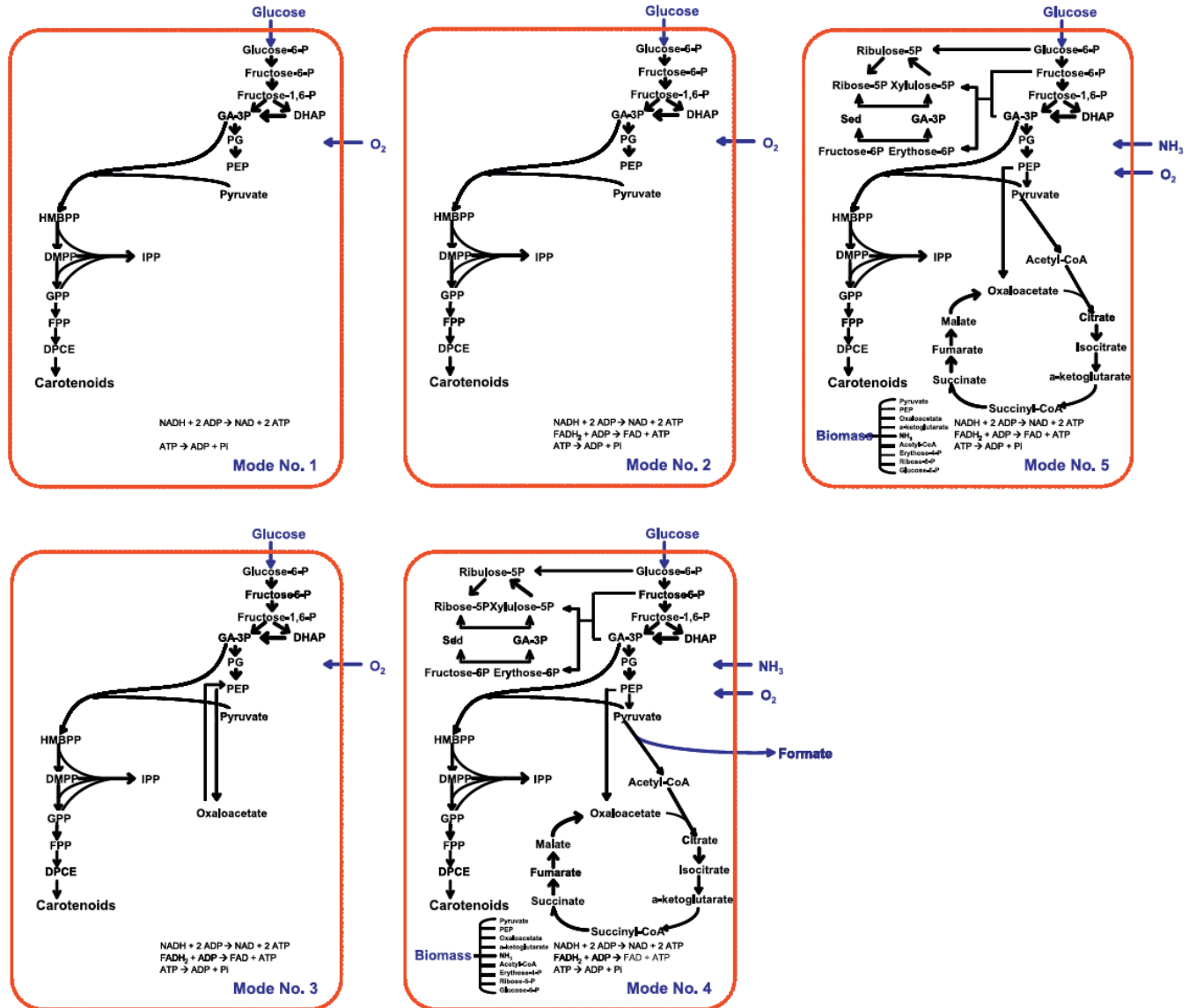
Als optimale Lösung vorhergesagt:

der Knockout von 8 Genen sollte zur Überproduktion von DPL und DPA führen.

nur 5 EFMs verbleiben für das restliche Netzwerk.

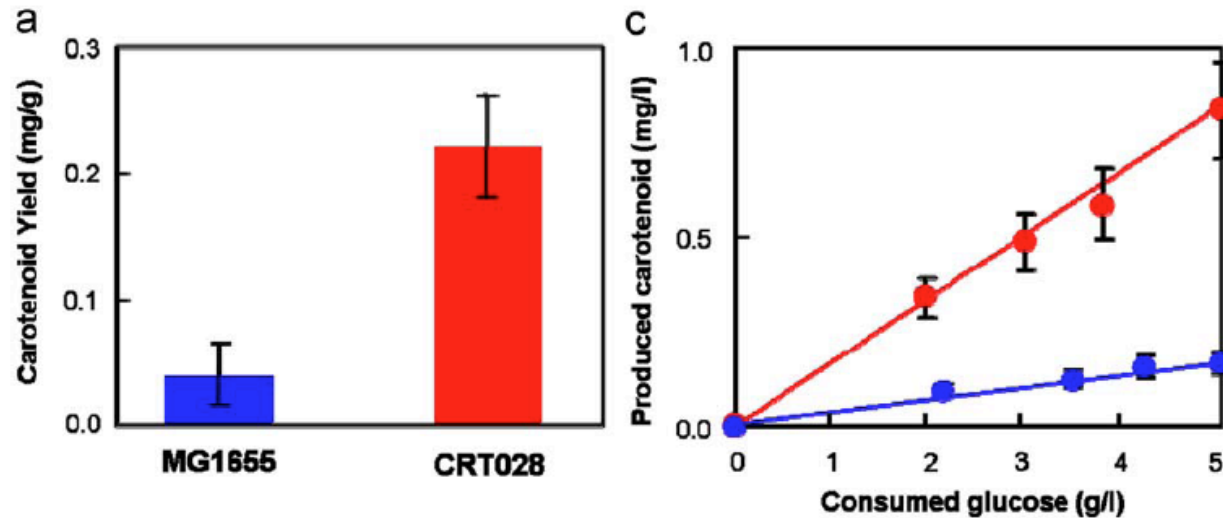
Unrean et al. Metabol Eng 12, 112-122 (2010)

Verbleibende EFM's



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

Experimentelle Verifikation: erhöhte Carotenoid-Ausbeute



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

	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

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

26

Protein-Interaktionsnetzwerk in *S. cerevisiae*

Abb. zeigt das Protein–Protein Interaktionsnetzwerk in *Saccharomyces cerevisiae*, basierend auf **yeast two-hybrid** Experimenten.

Knoten: einzelne Proteine

Verbindungen/Kanten: physikalische Interaktionen.

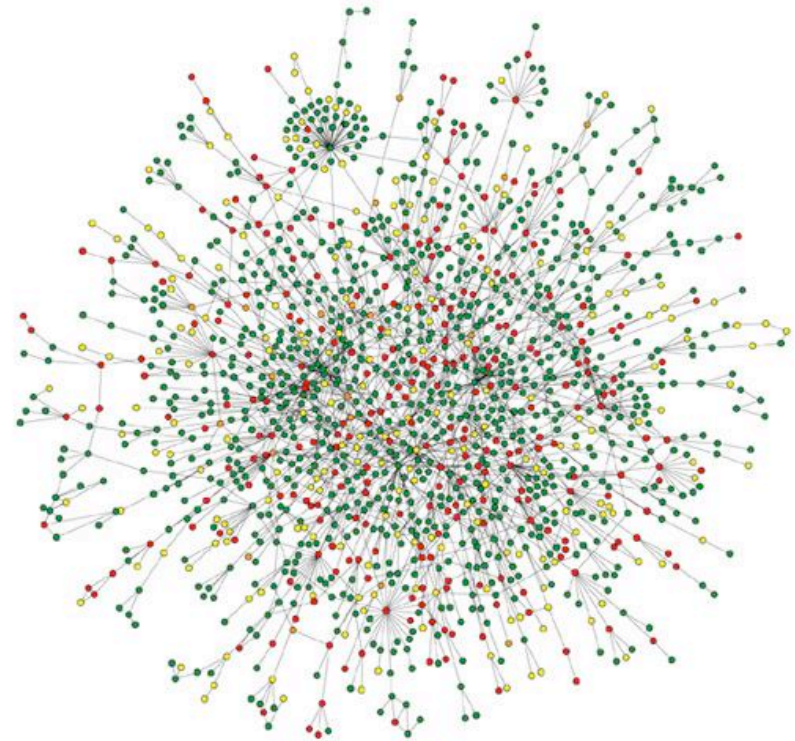
Das Netzwerk wird von einigen Knoten mit sehr vielen Verbindungen zusammengehalten. Diese nennt man **Hubs**.

Dieser größte Cluster enthält 78% aller Proteine.

Rot: Gendeletion ist tödlich

Grün: Gendeletion ist nicht tödlich

Gelb: Effekt der Gendeletion ist unbekannt.



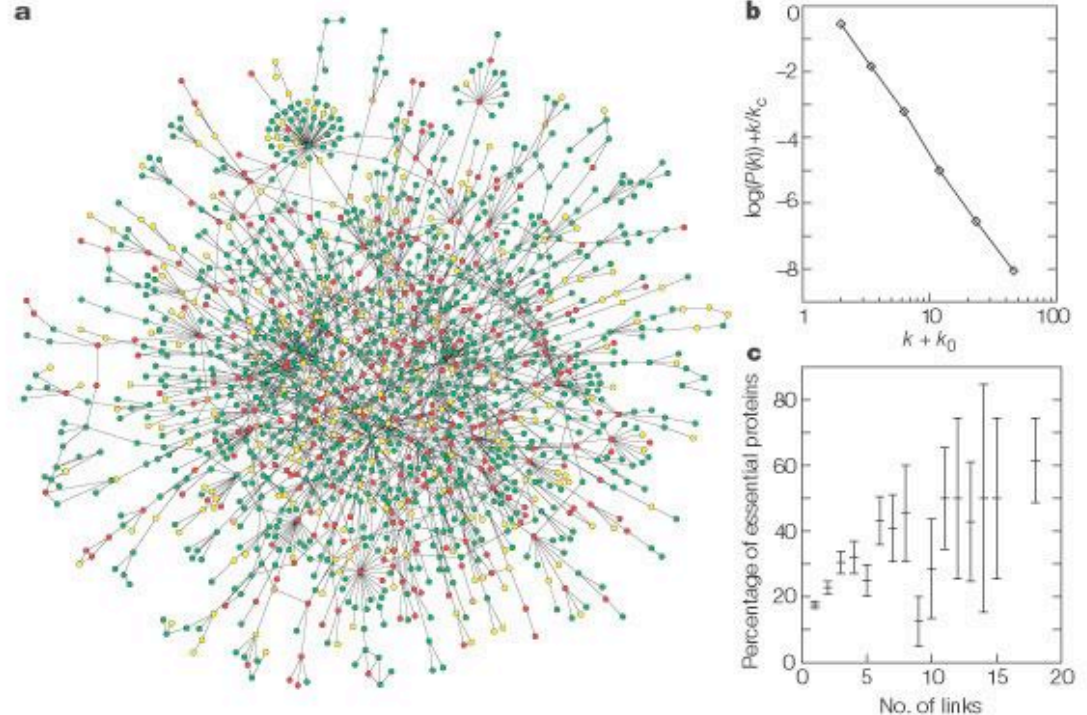
Nature Reviews | Genetics

Barabasi & Oltvai, Nature Rev Gen 5, 101 (2004)

Welche Proteindeletionen sind tödlich?

(b) Die Häufigkeit von Hub-Proteinen mit k Interaktionen nicht exponentiell ab (wie in einem Zufallsnetzwerk mit Poisson-Verteilung der Verknüpfungen erwartet), sondern gemäss einem Potenzgesetz (*engl. power law*).

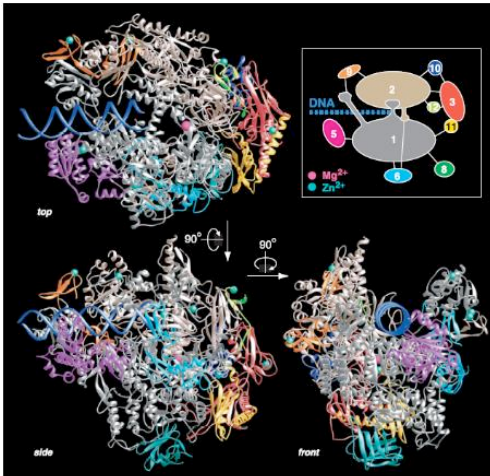
(c) Die Deletion von Genen, die für Hub-Proteine kodieren, die viele Links besitzen, ist eher tödlich als die Deletion von Proteinen, die mit wenigen anderen Proteinen interagieren.



H. Jeong, S. P. Mason, A.-L. Barabási and Z. N. Oltvai
Nature **411**, 41-42

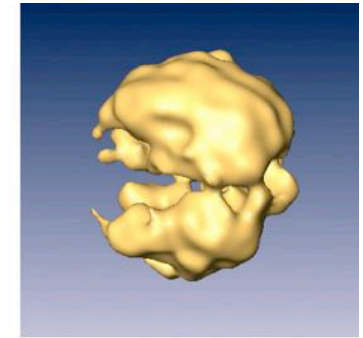
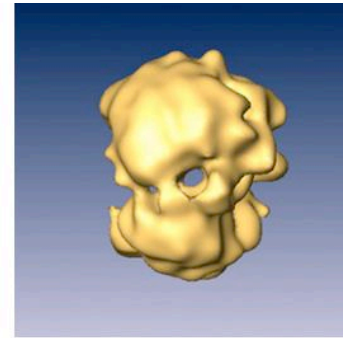
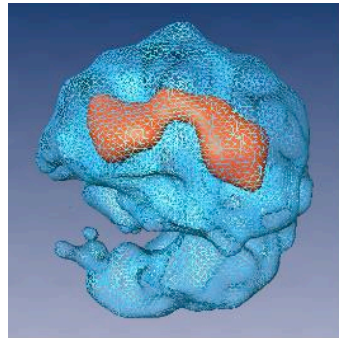
Beispiele für wichtige Proteinkomplexe

RNA Polymerase II



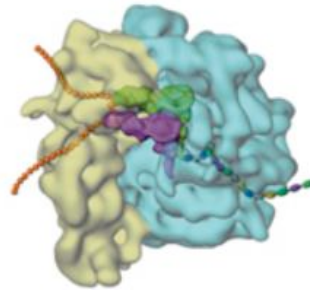
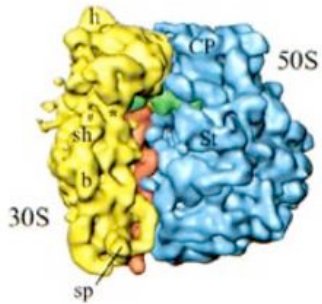
Cramer et al., Science 288, 640 (2000)

Spliceosome



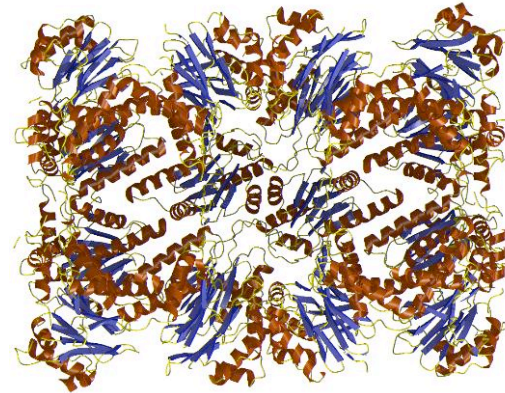
<http://www.weizmann.ac.il/>

Ribosom



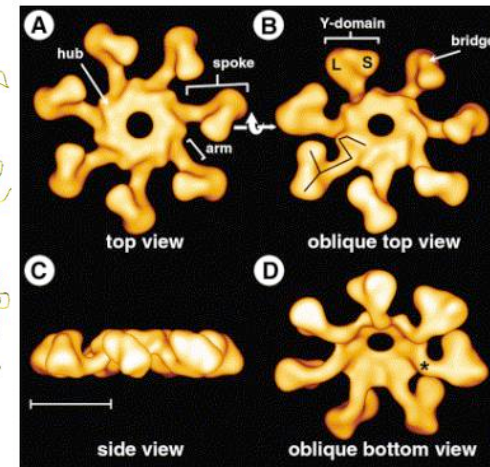
<http://www.millerandlevine.com/chapter/12/cryo-em.html>

Proteasome



<http://www.biochem.mpg.de>

Apoptosom



Acehan et al. Mol. Cell 9, 423 (2002)

Proteininteraktionsnetzwerke

The screenshot displays the Cytoscape Desktop interface (Session: humanANDmouseIntact.cys). The main window is divided into several panels:

- Control Panel:** Contains a table of network statistics and an "Import Network From Database" button.
- Network Table:**

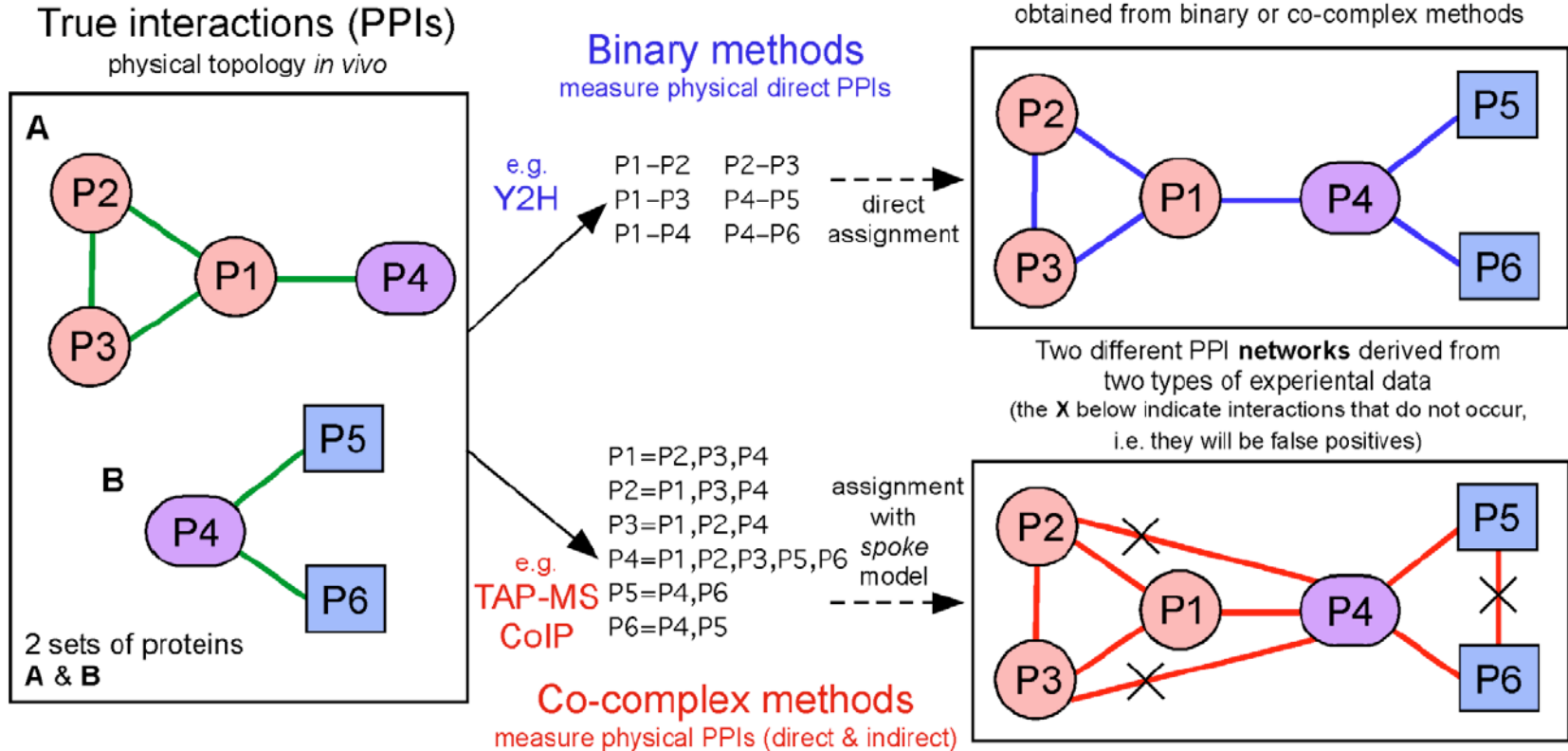
Network	Nodes	Edges
Human Interactions From IntAct	3007(1)	5825(134)
Mouse Interactions From IntAct	8578(52)	30491(817)
MCODE Result 1, Cluster 1	4(0)	69(0)
MCODE Result 1, Cluster 11	13(0)	161(0)
MCODE Result 1, Cluster 12	52(0)	406(0)
- Main View:** A large network graph with nodes and edges. A zoomed-in view of a cluster is shown in the top right, and a detailed view of a specific cluster (MCODE Result 1, Cluster 1) is shown in the bottom right, highlighting nodes like Q00987, P93099, and P04637.
- Data Panel:** Located at the bottom, it shows a table of node attributes from Ensembl:

ID	MI	ensembl
P62158	[0499]	[ENSG00000143933, ENSG00000160014, EN...
- Right Panel:** A sidebar with various tool options and a "Screenshots" label at the bottom.

At the bottom of the interface, there are instructions: "Welcome to Cytoscape 2.6.0", "Right-click + drag to ZOOM", and "Middle-click + drag to PAN".

Proteininteraktionsnetzwerke

Verschiedene experimentelle Techniken messen verschiedene Eigenschaften von Proteinkomplexen.



De Las Rivas, PLOS Comp Biol. 6, e1000807 (2010)

Datenbanken zu Proteininteraktionen

Table 1. Description of PPI databases and repositories.

Acronym	Database Full Name and URL	PPI Sources	Type of MI	Species	<i>n</i> Proteins (Dec. 2009)	<i>n</i> Interactions (Dec. 2009)
Primary Databases: PPI experimental data (curated from specific SSc & LSc published studies)						
BIND	Biomolecular Interaction Network Database, http://bond.unleashedinformatics.com/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[31,972]	[58,266]
BioGRID	Biological General Repository for Interaction Datasets, http://www.thebiogrid.org/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[28,717]	[108,691]
DIP	Database of Interacting Proteins, http://dip.doe-mbi.ucla.edu/dip/	Ssc & Lsc published studies (literature-curated)	Only PPIs	All	20,728	57,683
HPRD	Human Protein Reference Database, http://www.hprd.org/	Ssc & Lsc published studies (literature-curated)	Only PPIs	Human	27,081	38,806
IntAct	IntAct Molecular Interaction Database, http://www.ebi.ac.uk/intact/	Ssc & Lsc published studies (literature-curated)	PPIs & others	All	[60,504]	[202,826]
MINT	Molecular INTeraction database, http://mint.bio.uniroma2.it/mint/	Ssc & Lsc published studies (literature-curated)	Only PPIs	All	30,089	83,744
MIPS-MPact	MIPS protein interaction resource on yeast, http://mips.gsf.de/genre/proj/mpact/	Derived from CYGD	Only PPIs	Yeast	1,500	4,300
MIPS-MPPI	MIPS Mammalian Protein-Protein Interaction Database, http://mips.gsf.de/proj/ppi	Ssc published studies (literature-curated)	Only PPIs	Mammalian	982	937
Meta-Databases: PPI experimental data (integrated and unified from different public repositories)						
APID	Agile Protein Interaction DataAnalyzer, http://bioinfow.dep.usal.es/apid/	BIND, BioGRID, DIP, HPRD, IntAct, MINT	Only PPIs	All	56,460	322,579
MPIDB	The Microbial Protein Interaction Database, http://www.jcvi.org/mpidb/	BIND, DIP, IntAct, MINT, other sets (exp & lit.-curated)	Only PPIs	Microbial	7,810	24,295
PINA	Protein Interaction Network Analysis platform, http://csbi.ltdk.helsinki.fi/pina/	BioGRID, DIP, HPRD, IntAct, MINT, MPact	Only PPIs	All	[?]	188,823
Prediction Databases: PPI experimental and predicted data ("functional interactions", i.e., interactions <i>lato sensu</i> derived from different types of data)						
MiMI	Michigan Molecular Interactions, http://mimi.ncibi.org/MimiWeb/	BIND, BioGRID, DIP, HPRD, IntAct, & nonPPI data	PPIs & others	All	[45,452]	[391,386]
PIPs	Human PPI Prediction database, http://www.compbio.dundee.ac.uk/www-pips/	BIND, DIP, HPRD, OPHID, & nonPPI data	PPIs & others	Human	[?]	[37,606]
OPHID	Online Predicted Human Interaction Database, http://ophid.utoronto.ca/	BIND, BioGRID, HPRD, IntAct, MINT, MPact, & nonPPI data	PPIs & others	Human	[?]	[424,066]
STRING	Known and Predicted Protein-Protein Interactions, http://string.embl.de/	BIND, BioGRID, DIP, HPRD, IntAct, MINT, & nonPPI data	PPIs & others	All	[2,590,259]	[88,633,860]
UniHI	Unified Human Interactome, http://www.mdc-berlin.de/unihi/	BIND, BioGRID, DIP, HPRD, IntAct, MINT, & nonPPI data	PPIs & others	Human	[22,307]	[200,473]

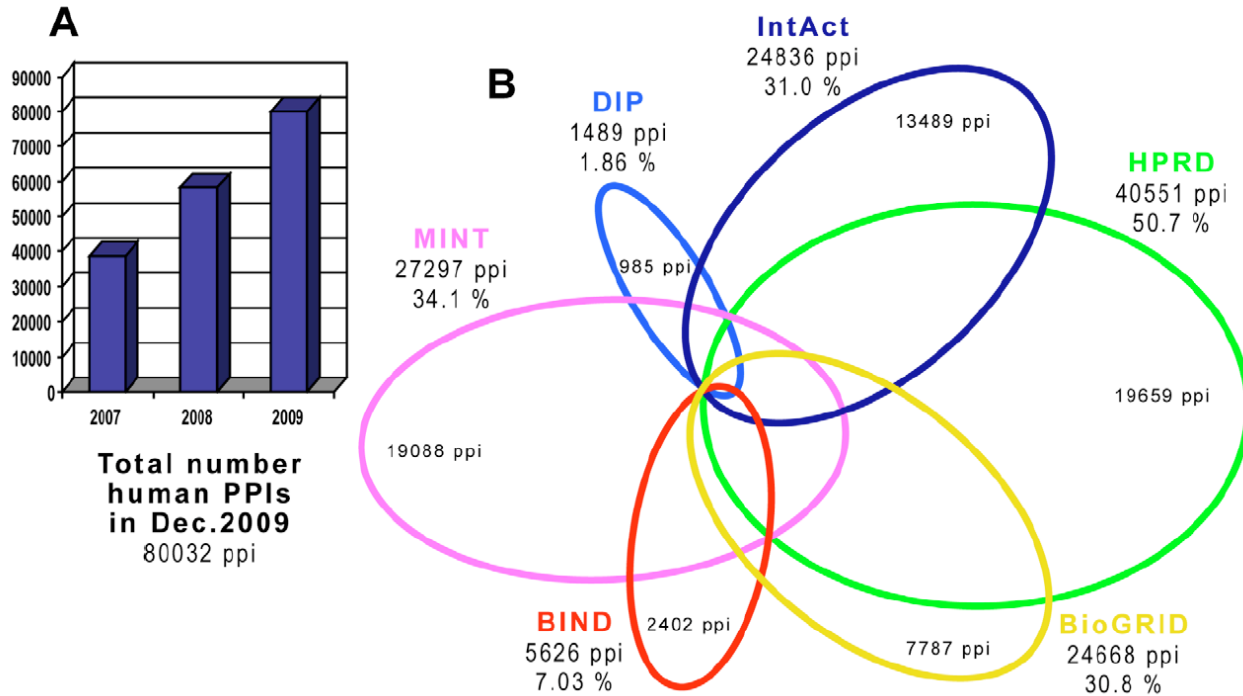
The table divided in three sections: **primary databases**, which include PPIs from large- and small-scale (Lsc & Ssc) experimental data that are usually obtained from curation of research articles (8 resources included: BIND, BioGRID, DIP, HPRD, IntAct, MINT, MIPS-MPACT, MIPS-MPPI); **meta-databases**, which include PPIs derived from integration and unification of several primary repositories (3 resources: APID, MPIDB, PINA); **prediction databases**, which include PPIs from experimental analyses together with predicted PPIs obtained from the analyses of heterogenous biological data (5 resources: MiMI, PIPs, OPHID, STRING, UniHI). The table shows the total number of proteins and interactions that were reported by each repository in December 2009 (as far as we could see in the respective Web site). The numbers are in brackets [] when the repository includes PPIs and other types of interactions (e.g., protein-ligand interactions or for the case of prediction databases nonPPI data). The question mark [?] indicates that the number of distinct proteins included in such repository could not be found in the Web. doi:10.1371/journal.pcbi.1000807.t001

De Las Rivas, PLOS Comp Biol. 6, e1000807 (2010)

Proteininteraktionsnetzwerke

Human Interactome

Coverage of human PPIs on major public repositories



Der Überlapp zwischen verschiedenen Datenbanken ist leider gering.

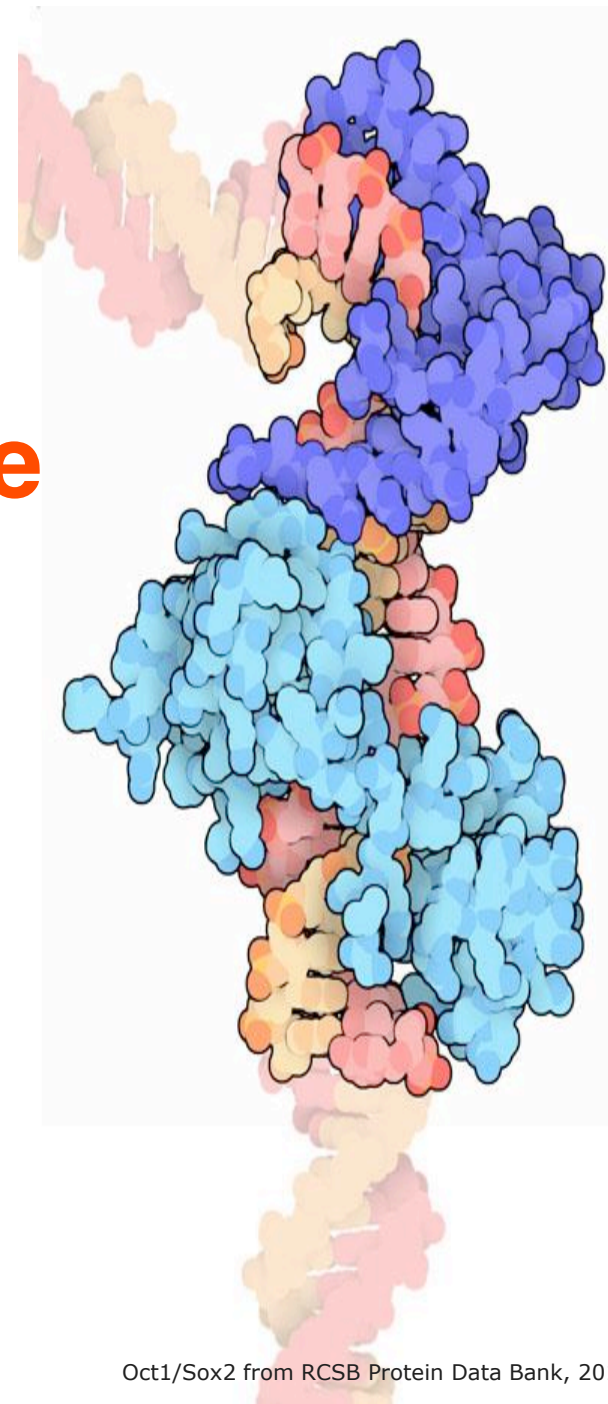
Experimentell nachgewiesen wurden ca. 80 000 Interaktionen von menschlichen Proteinen.

Man schätzt, dass es in *S. cerevisiae* (6000 Gene) ca. 20 000 – 30 000 Proteininteraktionen gibt.

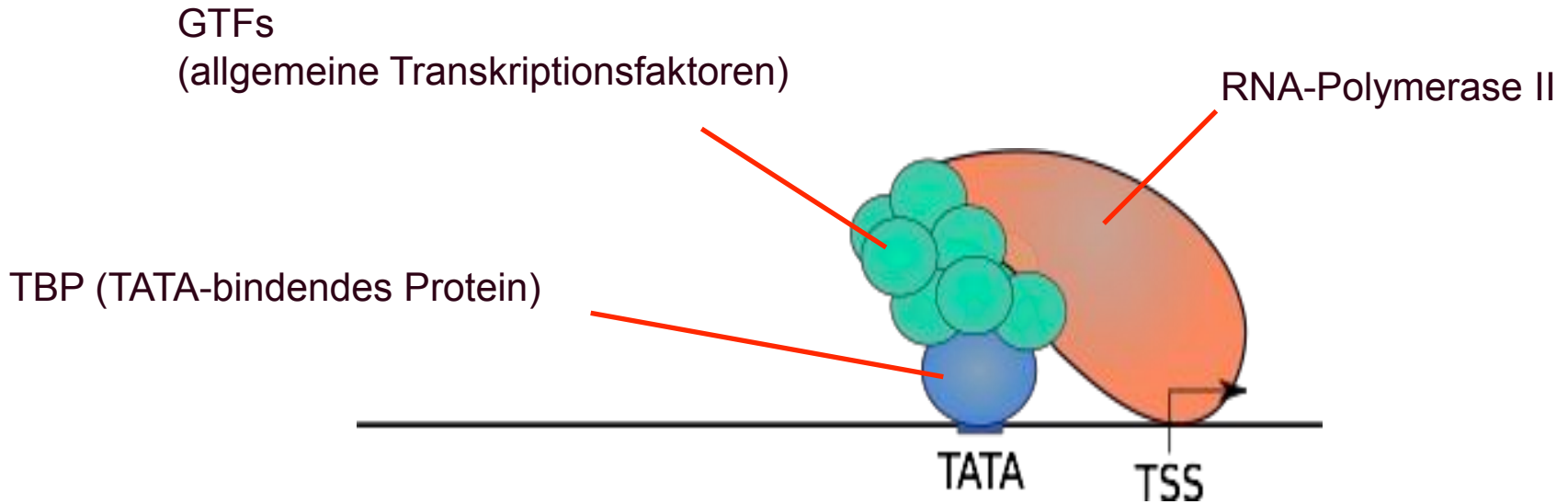
D.h. im Mensch wird eine erheblich höhere Anzahl (ca. 300.000) erwartet.

De Las Rivas, PLOS Comp Biol. 6, e1000807 (2010)

Transkriptionsfaktor- komplexe in Hefe und ihre Rolle

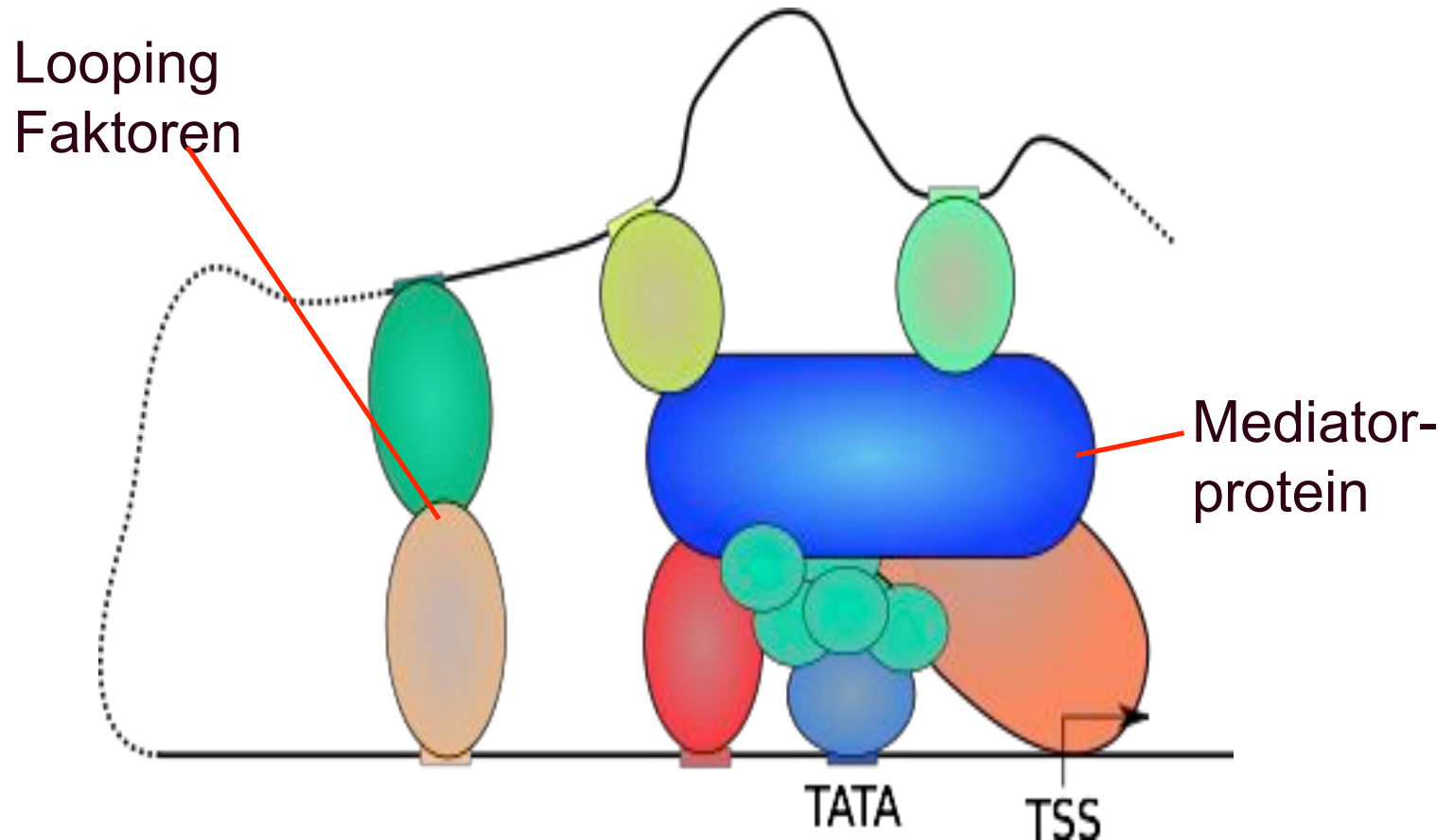


Transkription: Rolle von TFs



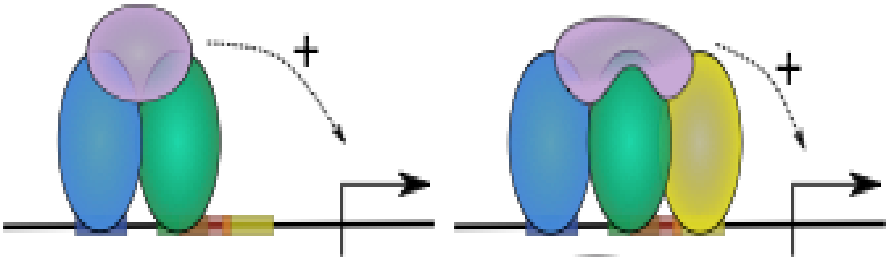
TBP bindet an die DNA, wechselwirkt mit den allgemeinen Transkriptionsfaktoren und rekrutiert RNA-Polymerase II

Kombinatorische Vielfalt vieler TFs, Bindung weiterer Proteine möglich

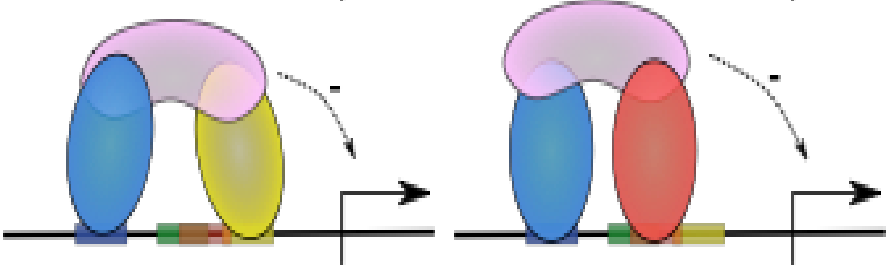


cis-regulatorische Module

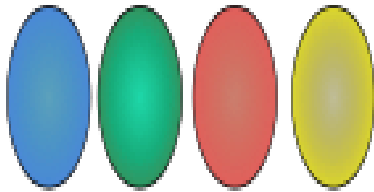
Koaktivatoren



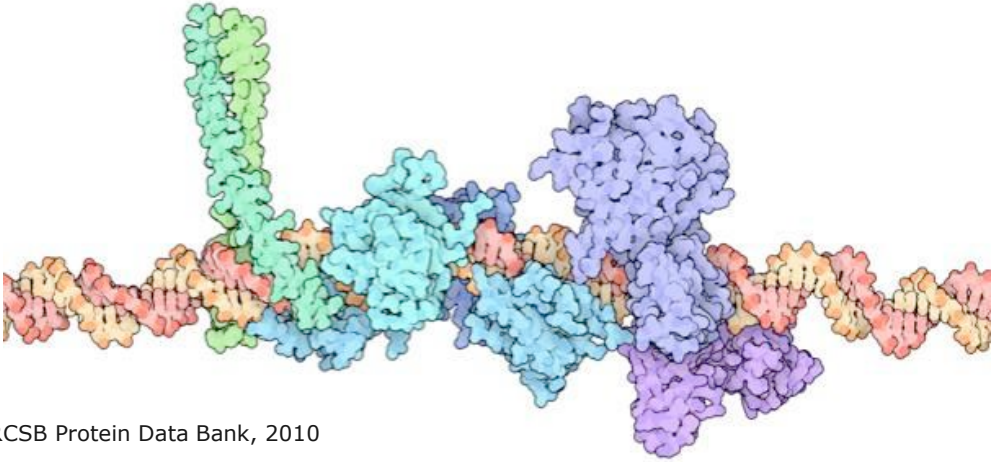
Korepressoren



TFs

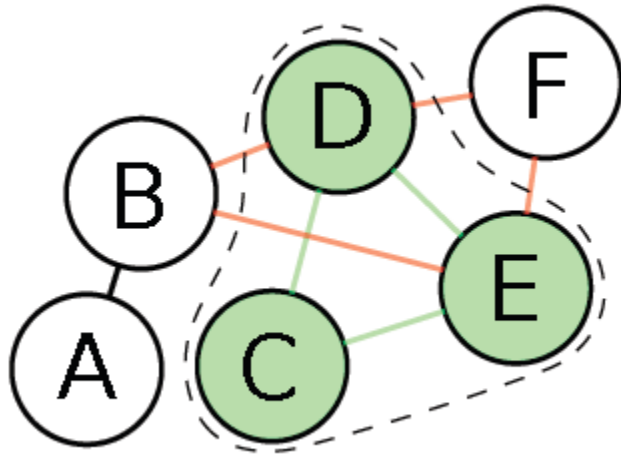


37



IFN-enhanceosome from RCSB Protein Data Bank, 2010

identifiziere Proteinkomplexe, die TFs beinhalten aus PPI-Netzwerk

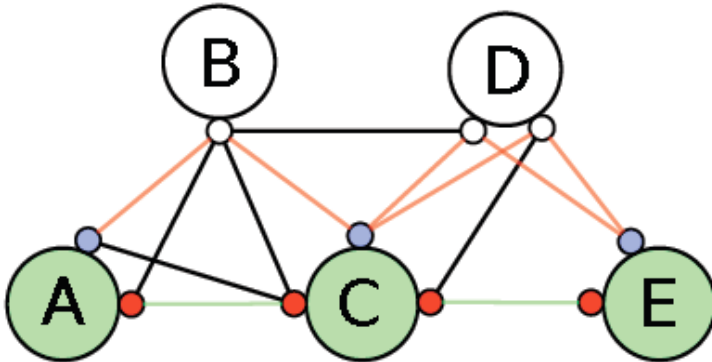


Vewende Idee der Methode ClusterOne:
Identifiziere Kandidaten für TF-
Komplexe im Protein-Interaktionsgraph
durch Optimierung der Cohesiveness:

$$f(V) = \frac{w^{in}(V)}{w^{in}(V) + w^{bound}(V)}$$

Domänen-Domänen Repräsentation des PPI-Netzwerks

Annahme: jede Domäne kann nur an einer Interaktion beteiligt sein.



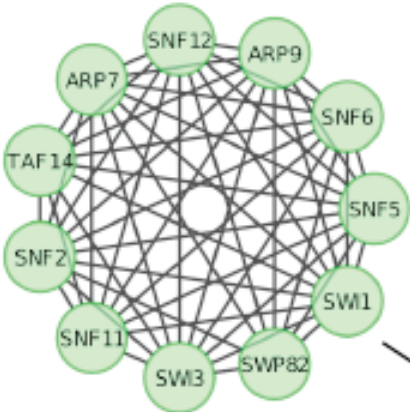
Grüne Proteine A, C, E bilden aktuellen Komplex.

Ihre roten Domänen stehen über die grünen Kanten in Kontakt.

B und D sind Kandidaten für einen größeren Komplex.

Sie könnten über die roten Kanten neue Interaktionen mit den nicht belegten (blauen) Domänen von A, C, E ausbilden.

Datenquellen



weighted protein-protein interaction data, such as

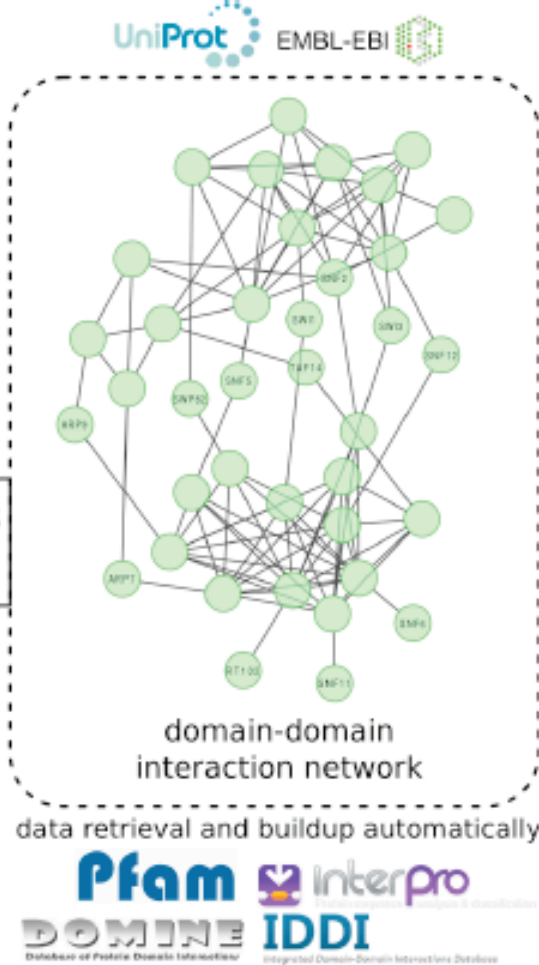


SIF format with UniProt IDs

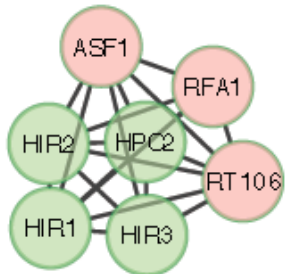
seed proteins,
threshold for pairs,
max. depth of search

domain-aware cohesiveness optimization algorithm

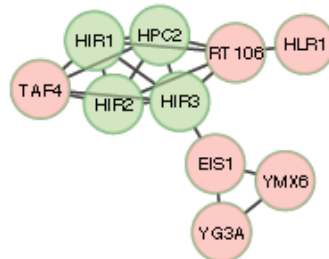
transcription factor complex candidates



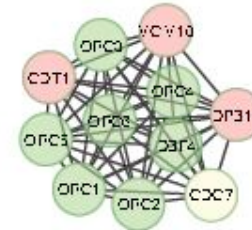
Beispiele für TF-Komplexe (DACA vs. ClusterONE)



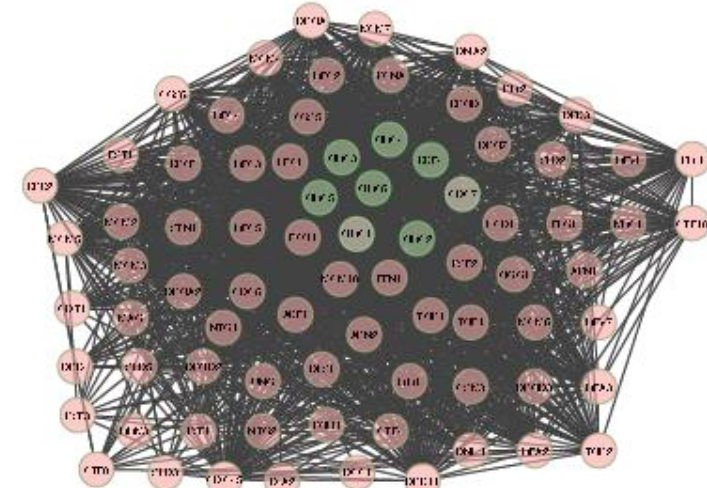
(a) HIR(SGD) / DACA



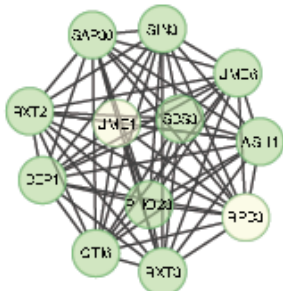
(b) HIR(SGD) / ClusterONE



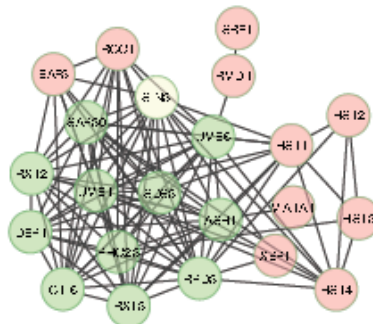
(e) ORC(MIPS) / DACA



(f) ORC(MIPS) / ClusterONE



(c) RPD3L(CYC2008) / DACA

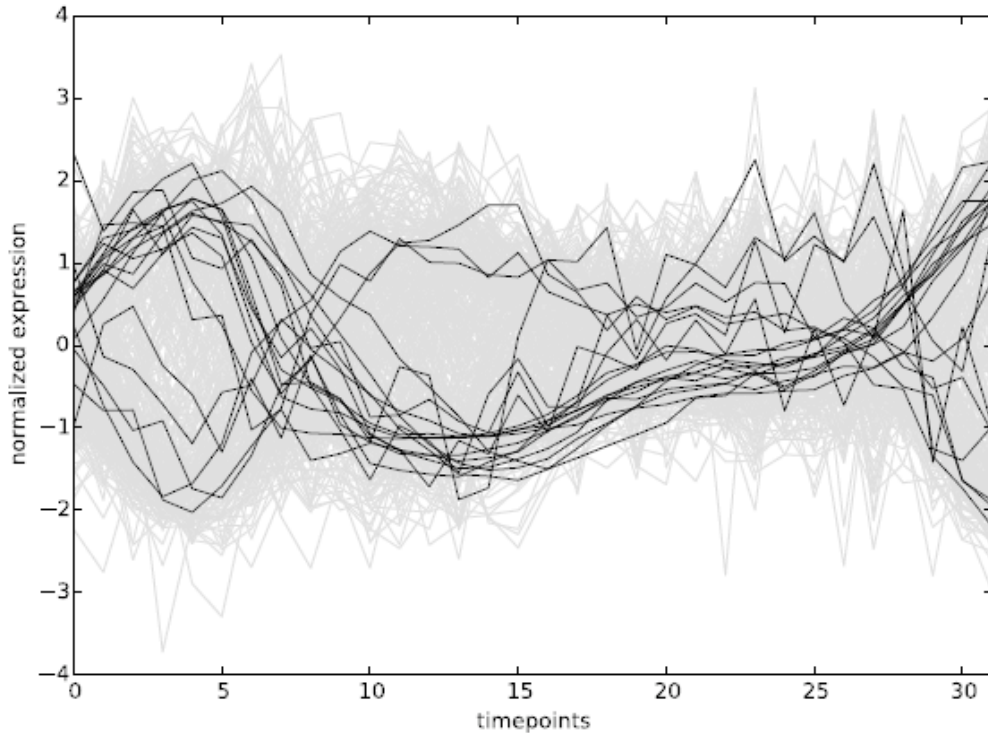


(d) RPD3L(CYC2008) / ClusterONE

Grüne Knoten: Proteine im Referenzkomplex, die korrekt vorhergesagt wurden.

Rote Knoten: weitere Proteine, die als Teil des Komplexes vorhergesagt werden, die aber experimentell nicht bestätigt sind.

Targetgene sind TF Komplexen sind ko-exprimiert!



Grau: Expression der Targetgene von MET4 **oder** MET32 während des Zellzyklus

Schwarz: Expression der Targetgene, die sowohl von MET4 **und** von MET32 reguliert werden.

-> es gibt 2 Gruppen von Targetgenen, die sich sehr ähnlich verhalten.

Funktionelle Rolle von TF Komplexen

TFs	P_{dECS}	bind. mode	targets	reg. influence	GO process enrichment ($P < 0.05$) in targets
MET4/MET32	0.0010	coloc.	19	+	methionine metabolic process
TBP/HAP5	0.0335	med.	47	+	/
GLN3/DAL80	0.0009	med.	28	/	allantoïn catabolic process
DIG1/STE12/SWI6	0.0369	all	15	/	fungal-type cell wall organization
FHL1/RAP1	0.0001	coloc.	116	+	rRNA transport
RPH1/GIS1	0.0001	med.	100	-	hexose catabolic process
CBF1/MET32	0.0002	coloc.	33	o	sulfate assimilation
DIG1/STE12	0.0003	med.	34	-	response to pheromone
GCN4/RAP1	0.033	med.	62	+	/
MSN4/MSN2	0.0021	med.	105	+	oligosaccharide biosynthetic process
DAL80/GZF3	0.0044	med.	20	-	purine nucleobase metabolic process
SWI6/SWI4	0.0039	med.	53	+	regulation of cyclin-dependent protein serine/threonine kinase activity
STB1/SWI6	0.0275	all	47	+	/
TBP/SWI6	0.0159	med.	14	+	/
GLN3/GZF3	0.0120	adj.	31	/	allantoïn catabolic process
MBP1/SWI6/SWI4	0.0307	med.	18	+	regulation of cyclin-dependent protein serine/threonine kinase activity
MBP1/SWI6	0.0124	adj.	25	/	cell cycle process