Моделирование метаболических путей

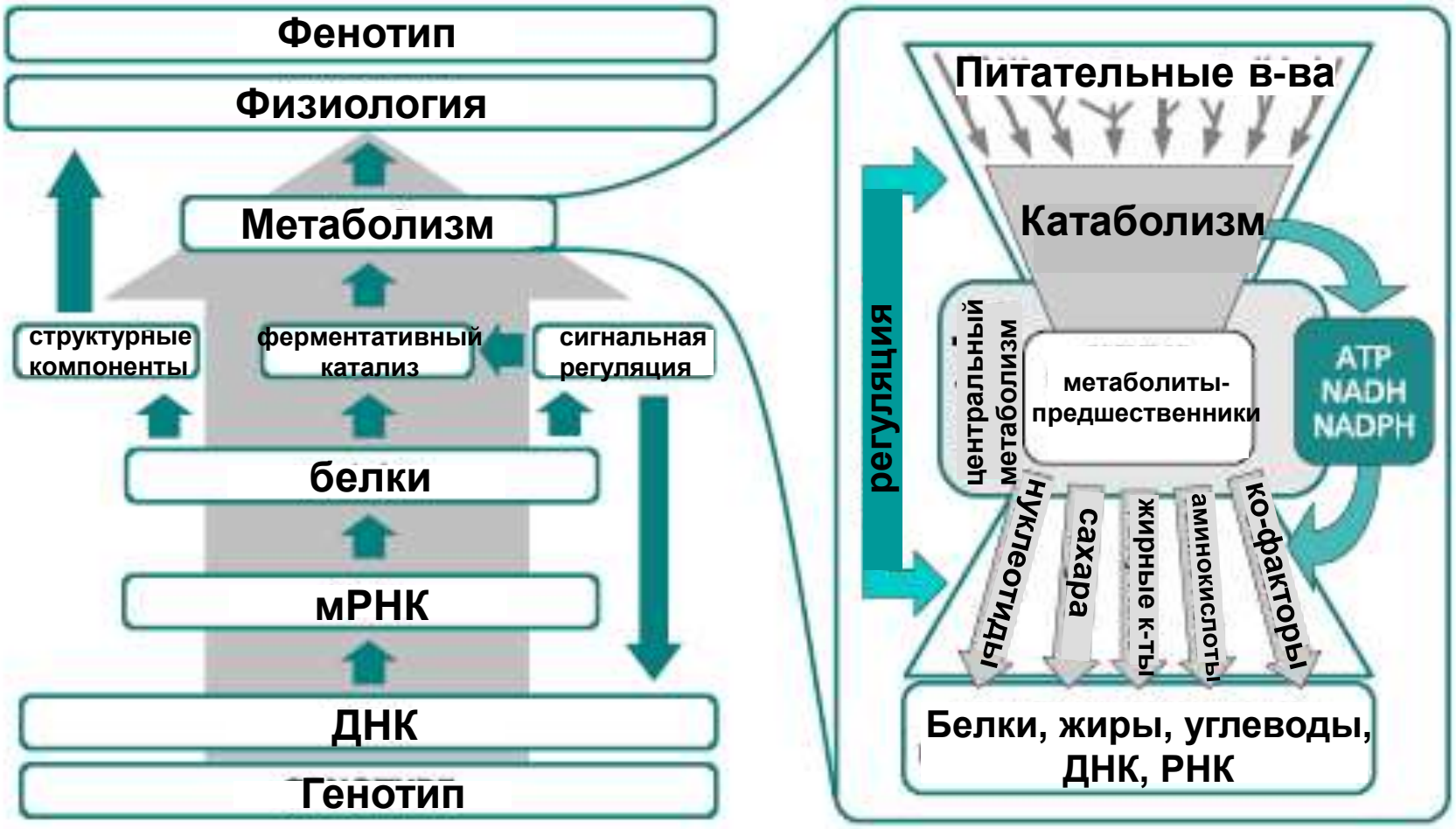
Плюснина Татьяна Юрьевна
Доц-т каф.биофизики
Биологического ф-та МГУ

Метаболи́зм (от греч. μεταβολή — «превращение, изменение»), — набор химических реакций, которые протекают в живом организме и способствуют поддержанию жизни.

Метаболический путь – совокупность метаболических реакций

Метаболиты – субстраты и продукты, участвующие в реакциях

Организация клеточного метаболизма



Цели и задачи метаболического моделирования

Научные задачи:

•исследование механизмов клеточной регуляции

•объяснение экспериментально установленных фактов

•предсказание новых характеристик внутриклеточных процессов

•систематизация накопленных экспериментальных данных

Биоинженерные задачи:

•получение необходимых веществ из бактериальных клеток и других организмов

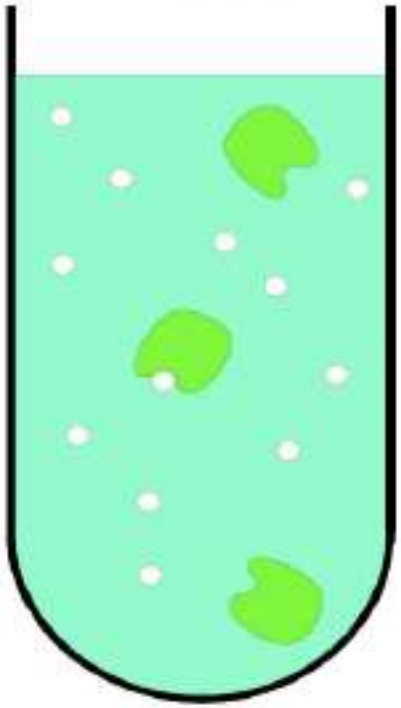
-разработка новых лекарств

-получение биотоплива

•разработка новых штаммов с наперед заданными свойствами

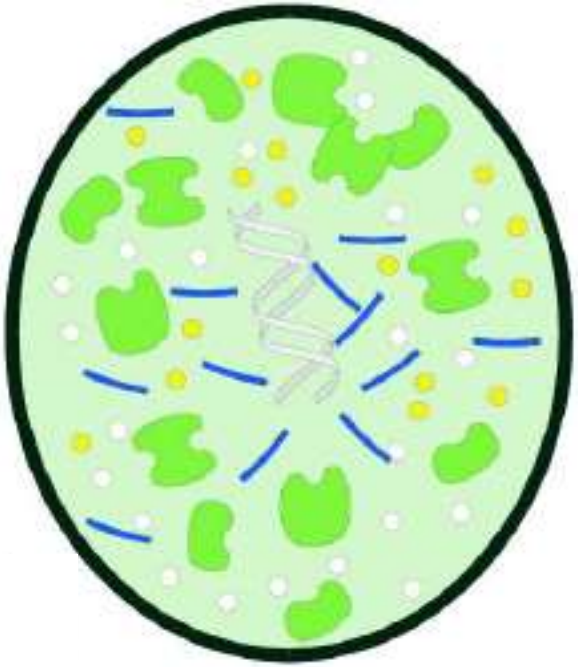
Уровни исследования метаболических реакций

In vitro



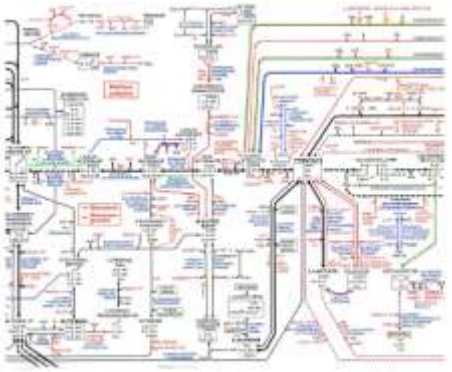
свойства отдельных метаболитов и их динамика

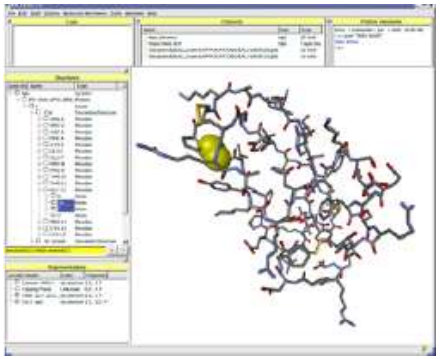
In vivo



описание характеристик метаболизма клетки в целом

In silico

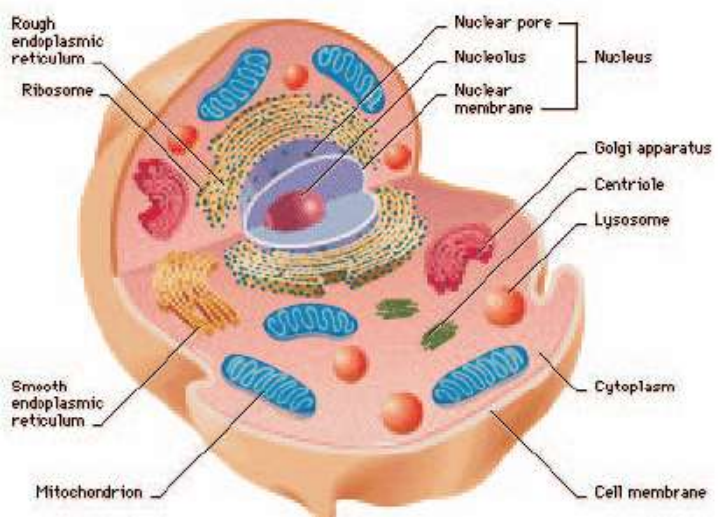




построение «электронной клетки»

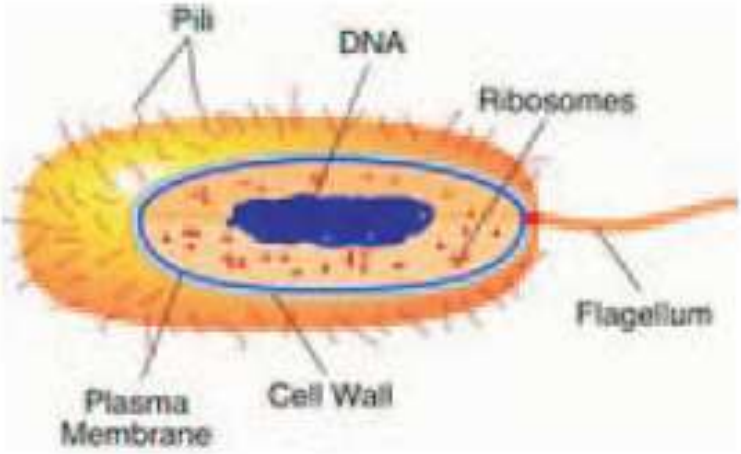
Разные уровни сложности организации клеток

Клетка эукариот



Клеточные процессы локализованы в разных компартментах

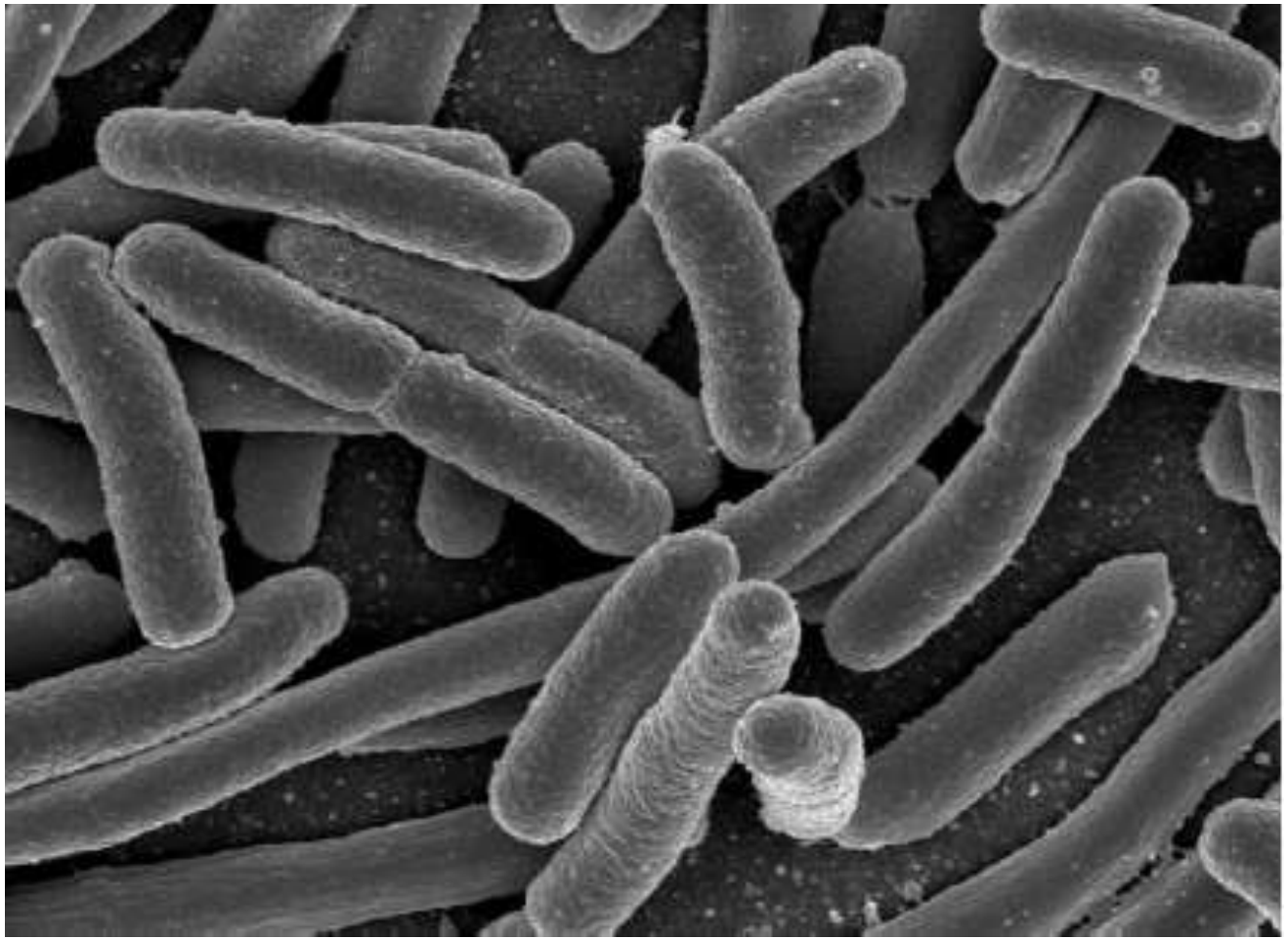
Клетка прокариот



Все процессы протекают в одном компартменте

Первые модели – модели клеток прокариот

Escherichia coli



http://www3.niaid.nih.gov/NR/rdonlyres/49477C30-0513-47BE-88FC-17974CB1F952/0/e_coli.jpg)

Escherichia coli – удобный объект для моделирования клеточного метаболизма

Метаболические пути *Escherichia coli*

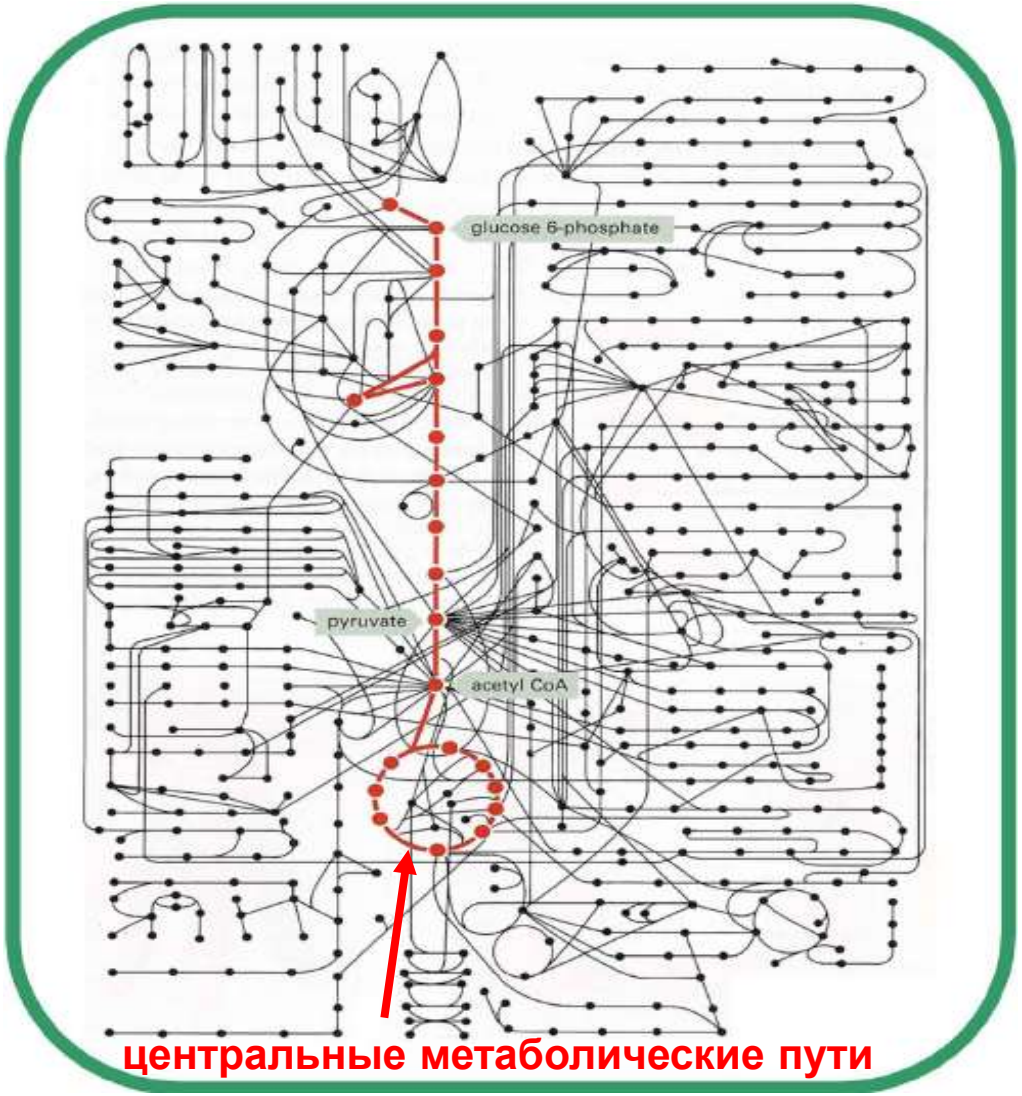


Figure from Ermir Qeli
Information Visualization Techniques for Metabolic Engineering
Dissertation zur Erlangung des akademischen Grades Doktor der Naturwissenschaften (Dr. rer. nat)

> 800 метаболических процессов
> 400 метаболитов
> 500 ферментов

Описание метаболизма включает:

•потребление необходимых для роста веществ из среды (мембранные транспортные процессы)

•центральные катаболические пути

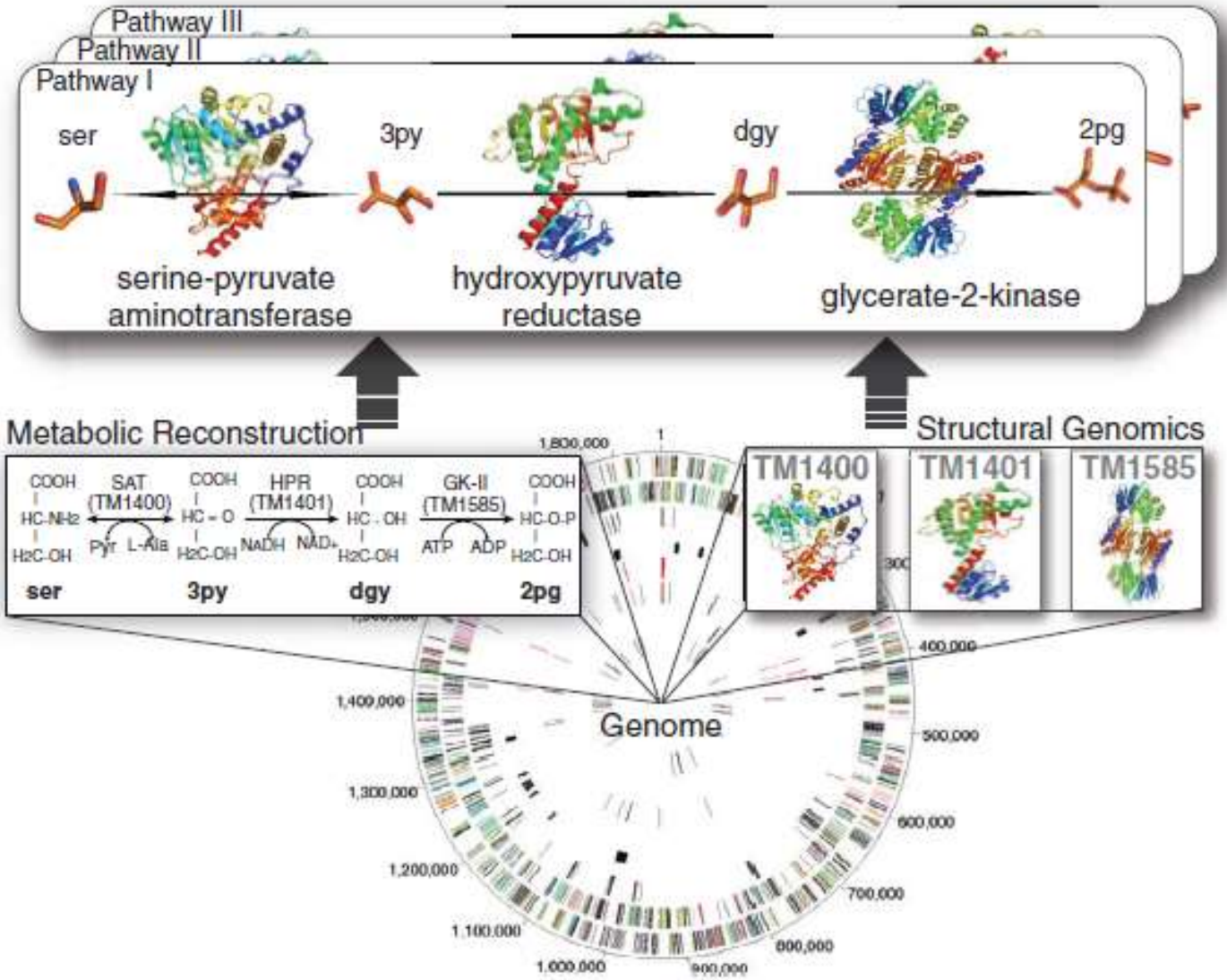
•биосинтетические пути (продукция всех компонент биомассы)

Построение метаболической модели:

реконструкция метаболических путей:
построение общей схемы метаболических путей

математическое описание:
вывод уравнений скорости и построение системы уравнений в соответствии с метаболической картой

Реконструкция метаболических путей

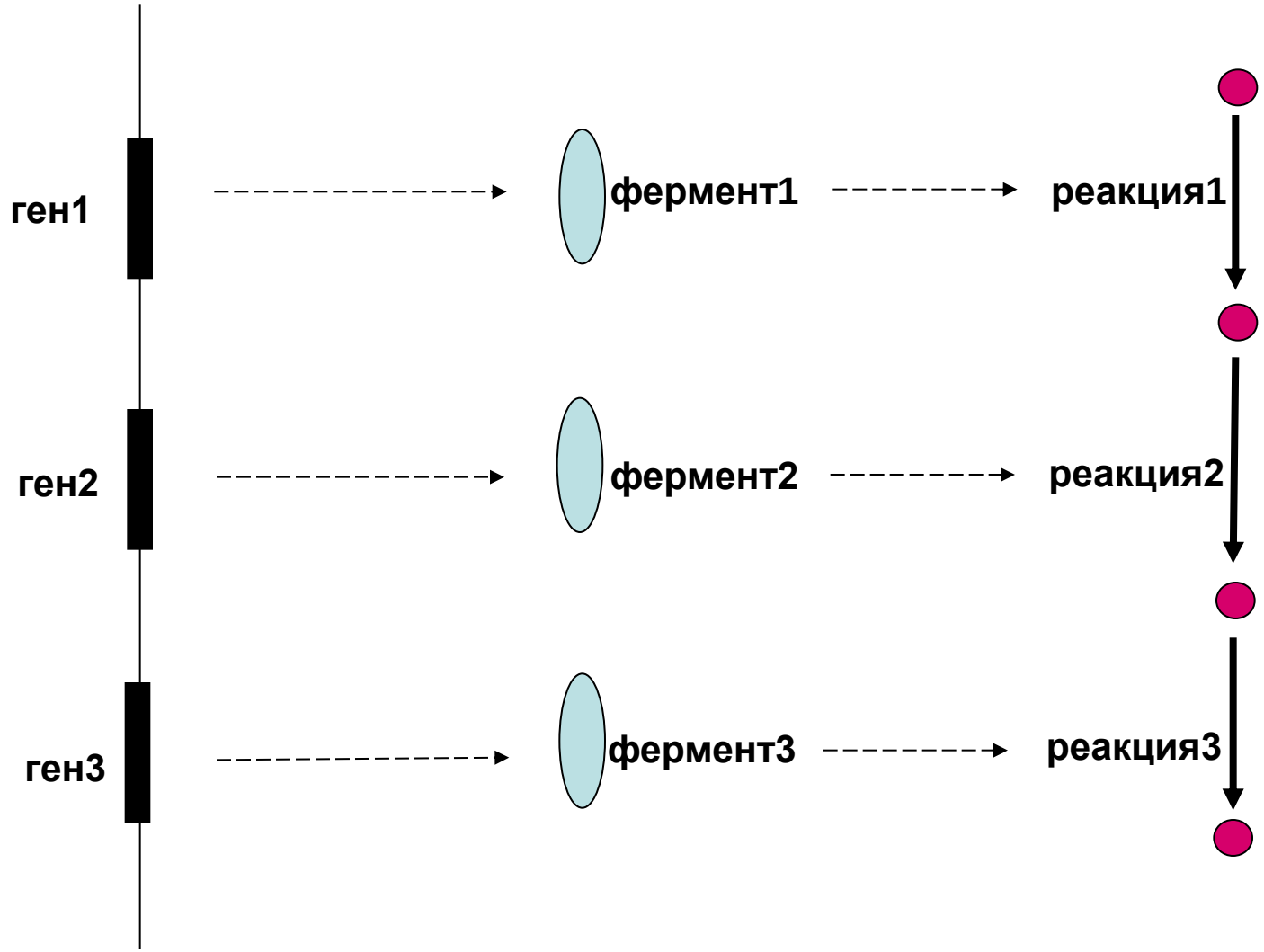


Science **325**, 1544 (2009);

Геном
(совокупность генов)

Протеом
(совокупность белков)

Метаболом
(совокупность метаболических реакций)

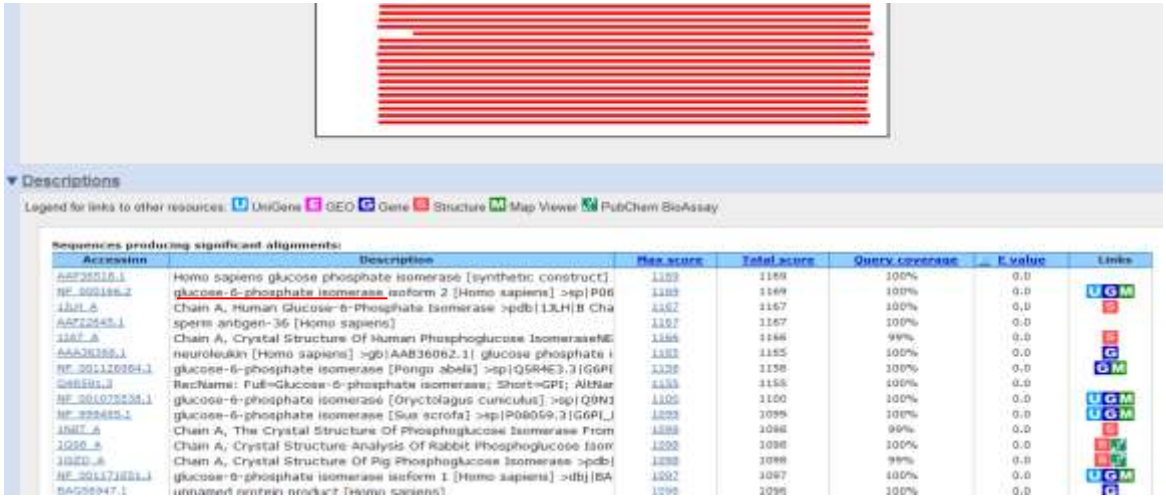


Пример аминокислотной последовательности

MAALTRDPQFQKLQQWYREHRSELNLRRLFDANKDRFNHFSLTLNTNHGHILVDYSKNLVTEDVMRMLVDLAKSRGVEAA
RERMFNGEKINYTEGRAVLHVALRNRSNTPILVDGKDVMPEVNKVLDKMKSFCQRVRSGDWKGYTGKTITDVINIGIGGSD
LGPLMVTEALKPYSSGGPRVWYVSNIDGTHIAKTLAQLNPESSLFIIASKTFTTQETITNAETAKEWFLQAAKDPSAVA
KHFVALSTNTTKVKEFGIDPQNMFEFWDWVGGRYSLWSAIGLSIALHVGFDNFEQLLSGAHWMDQHFRTTPLEKNAPVLL
ALLGIWYINCFGCETHAMLPYDQYLHRFAAYFQQGDMESNGKYITKSGTRVDHQTGPIVWGEPGTNGQHAFYQLIHQGT
KMIPCDFLIPVQTQHPIRKGLHHKILLANFLAQTEALMRGKSTEEARKELQAAGKSPEDLERLLPHKVFEGNRPTNSIVFT
KLTPFMLGALVAMYEHKIFVQGIIWDINSFDQWGVELGKQLAKKIEPELDGSAQVTSHDASTNGLINFIKQQREARVQ

↓

BLAST База данных по сравнению последовательностей



↓

фермент, соответствующий выбранной последовательности

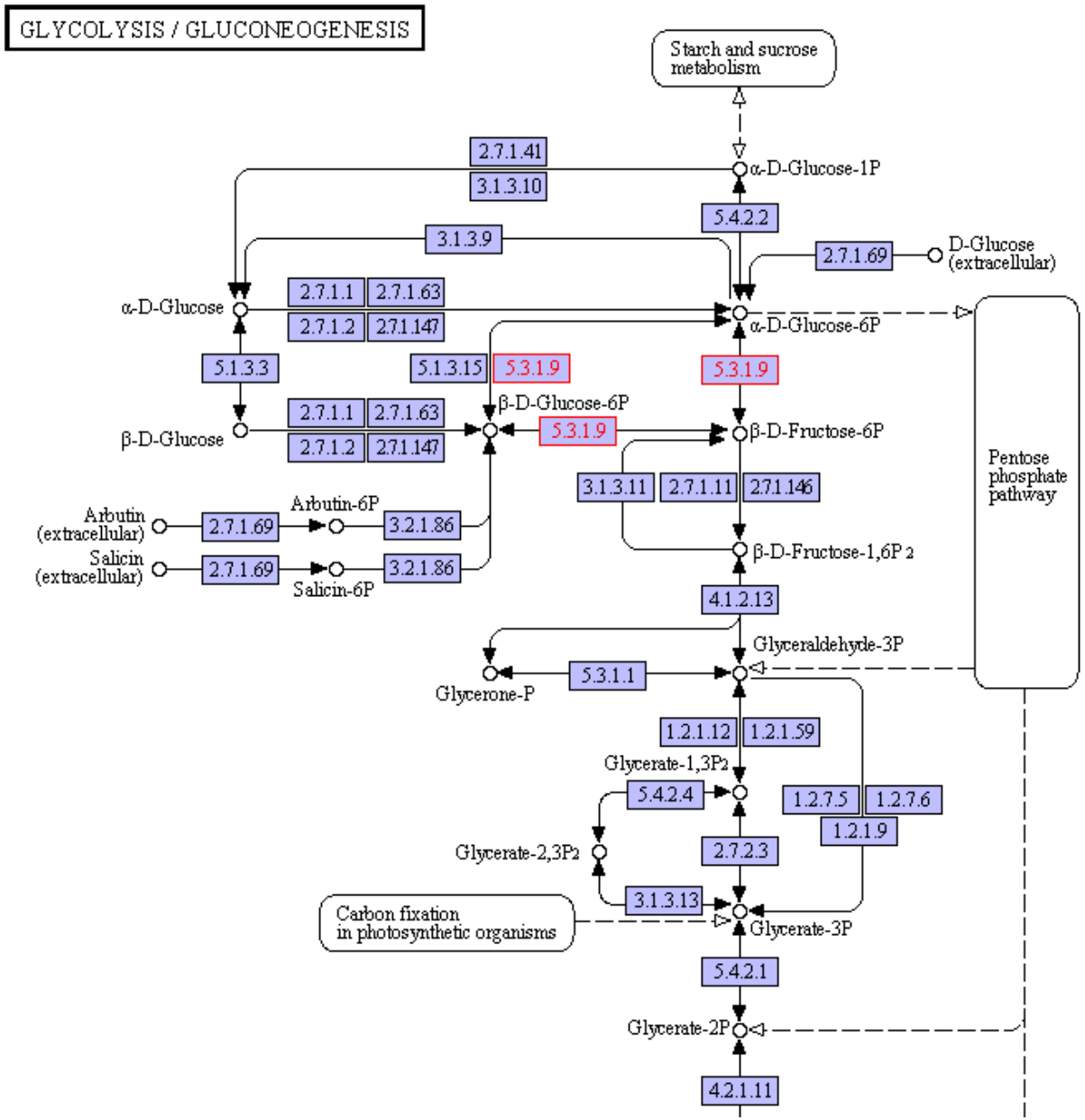
glucose-6-phosphate isomerase

↓

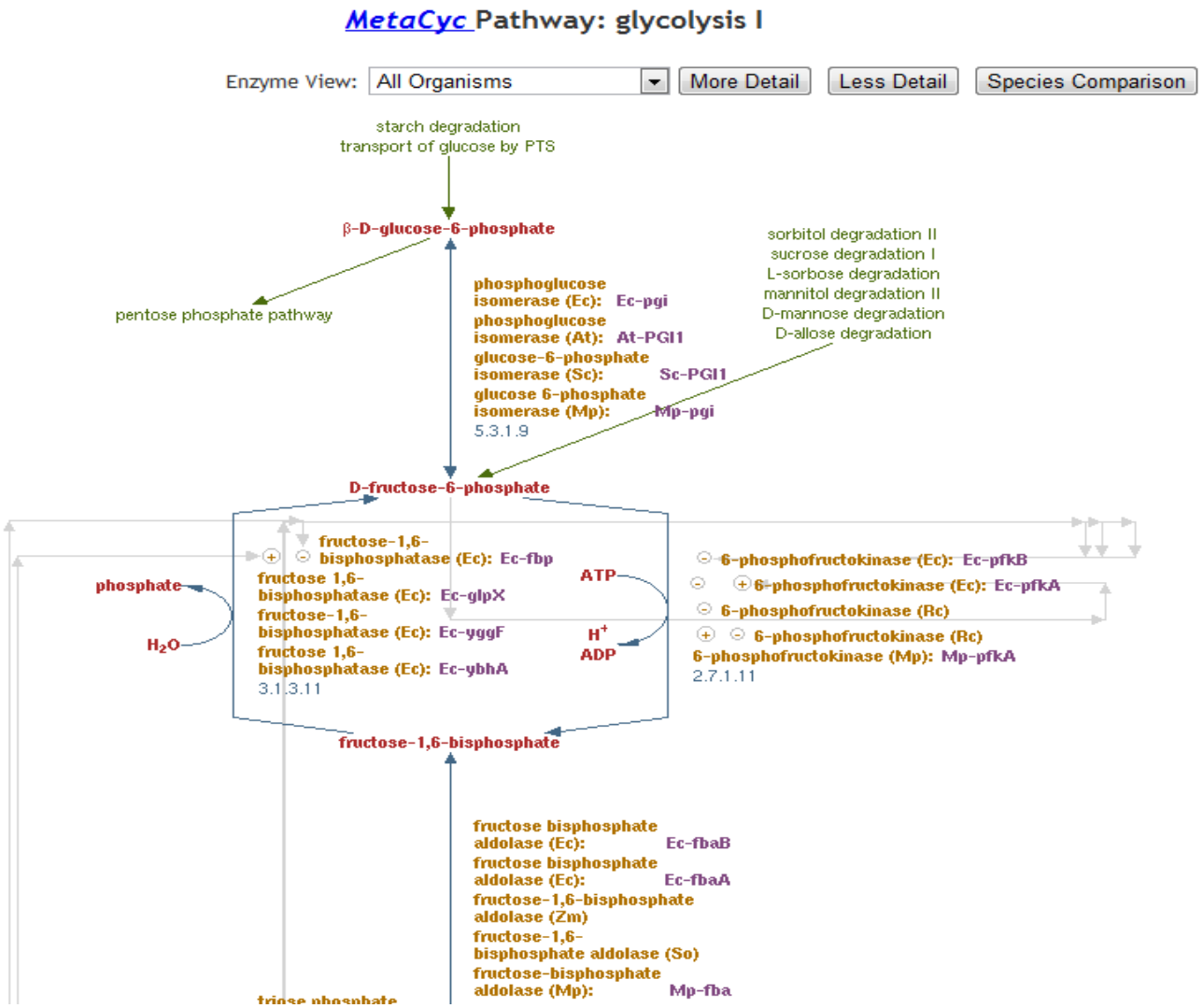
реакция, катализируемая ферментом → **определение всех реакций метаболического пути**

G6P → F6P

KEGG База данных по метаболическим путям



MetaCyc База данных по метаболическим путям



Представление скорости реакции

В соответствии с законом действующих масс

(скорость реакции пропорциональна вероятности столкновения реагентов, а вероятность в свою очередь пропорциональна концентрации реагентов с учетом молекулярности реакции)

Для реакции $S_1 + S_2 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} 2P$

$$v = k_1 S_1 S_2 - k_{-1} P^2 = v_+ - v_-$$

v - общая скорость реакции

v_+ - скорость прямой реакции

v_- - скорость обратной реакции

В общем виде:

$$v = v_+ - v_- = k_+ \prod_i S_i^{m_i} - k_- \prod_j P_j^{m_j}$$

m_i и m_j соответсвуют молекулярности S_i и P_j

Скорость реакции может быть выражена как через **концентрации** реагирующих веществ, так и в виде баланса **скоростей** прямых и обратных реакций.

Типы метаболических моделей

кинетические модели – системы ОДУ

- переменная модели – концентрация метаболита
- описание отдельных метаболических путей
- решение модели – динамическое поведение метаболитов во времени
- (~10-50 уравнений, ~100-500 параметров)

стехиометрические (потоковые) модели – системы линейных алгебраических уравнений

- переменная модели – метаболический поток
- описание метаболизма целой клетки
- решение модели – стационарное распределение метаболических потоков
- (~100-1000 уравнений)

Стехиометрические модели

Стехиометрические коэффициенты

Стехиометрические коэффициенты обозначают пропорции субстратов и продуктов, участвующих в реакции.

Пример: $S_1 + S_2 \overset{v}{\rightleftarrows} 2P$

Стехиометрические коэффициенты для S_1 S_2 и P: -1, -1 и 2.

Набор стехиометрических коэффициентов не единственен:

если считать, что для получения одного моля используется по половине моля каждого субстрата S_1 и S_2, можно записать: -1/2, -1/2 и 1;

или если изменить направление реакции, тогда можно записать: 1,1 и -2.

Стехиометрическая матрица

Стехиометрические коэффициенты n_{ij}, соответствующие метаболиту S_i и реакции v_j могут быть объединены в стехиометрическую матрицу:

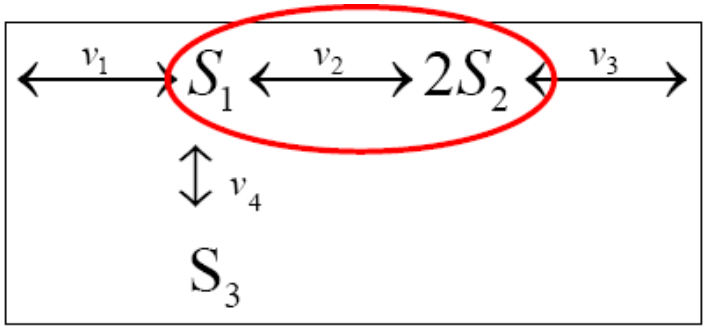
$$N = \begin{array}{c} \begin{matrix} v_1 & v_2 & \dots & v_r \end{matrix} \\ \begin{pmatrix} n_{11} & n_{12} & \dots & n_{1r} \\ \dots & \dots & \dots & \dots \\ n_{m1} & n_{m2} & \dots & n_{mr} \end{pmatrix} \begin{matrix} S_1 \\ \\ S_m \end{matrix} \end{array}$$

$$\boldsymbol{N} = \{n_{ij}\},$$
$$i = 1,\dots,m \; ; \; j = 1,\dots,r$$

каждый столбец соответствует реакции (r реакций)
каждая строка соответствует метаболиту (m метаболитов)

Стехиометрическая матрица

для реакции v_2:



$$N = \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array}$$

Строка 1: 1 молекула S_1 расходуется (-1)

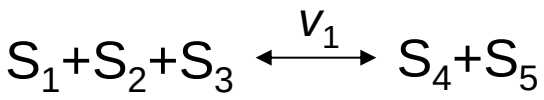
Строка 2: 2 молекулы S_2 образуются (2)

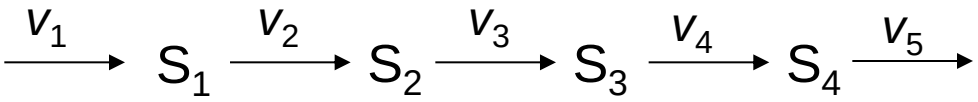
Строка 3: S_3 в этой реакции не участвует (0)

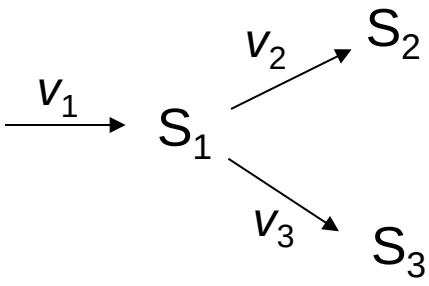
Реакции метаболической сети могут быть обратимы.

Чтобы определить знак коэффициентов, направление стрелок искусственно выбирается как положительное «слева направо» и «сверху вниз».

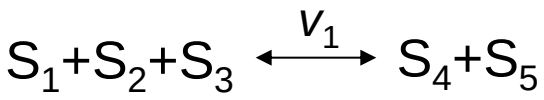
Построить стехиометрическую матрицу



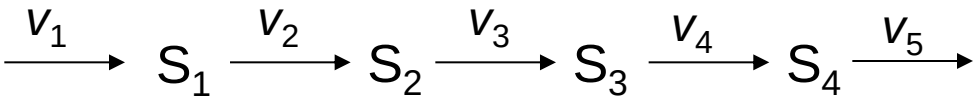




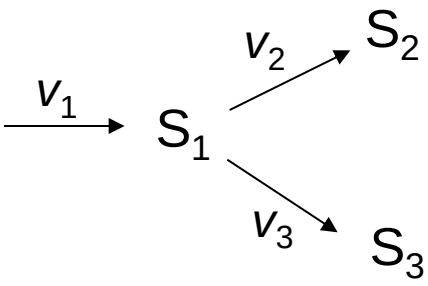
Построить стехиометрическую матрицу



$$N = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

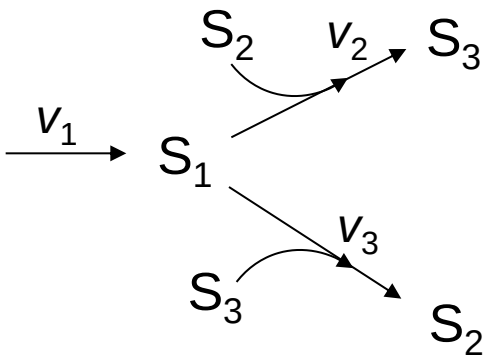


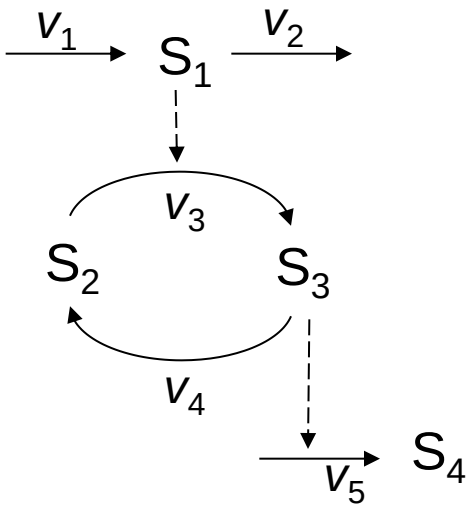
$$N = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$



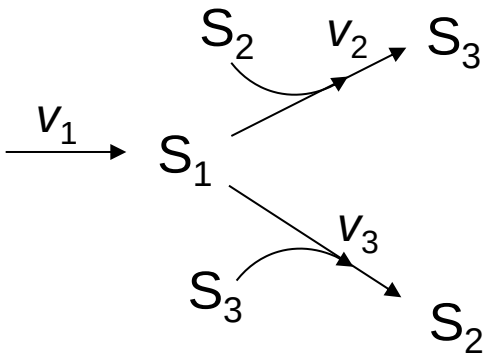
$$N = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Построить стехиометрическую матрицу

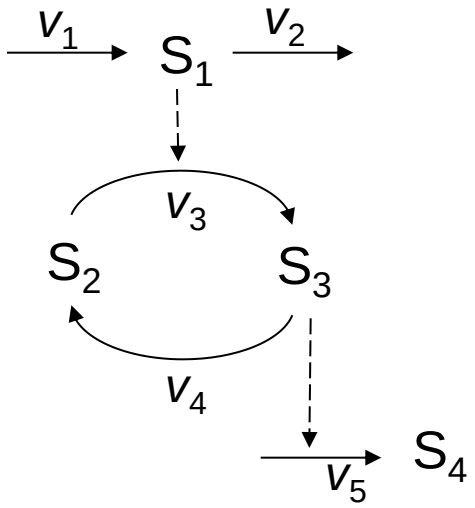




Построить стехиометрическую матрицу



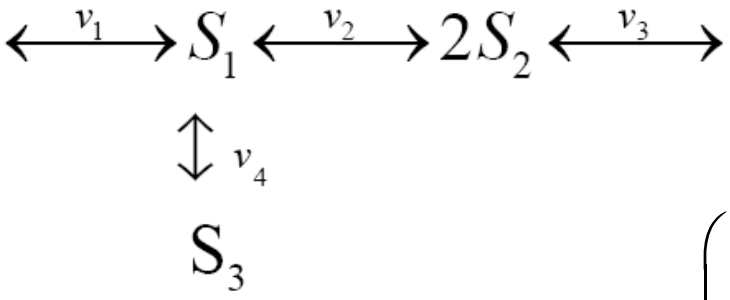
$$N = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$



$$N = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Построение уравнений баланса

Пример:



$$S = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

вектор метаболитов

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$$

вектор скоростей

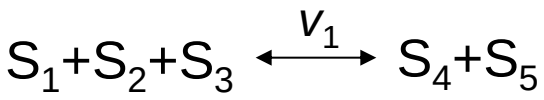
$$N = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

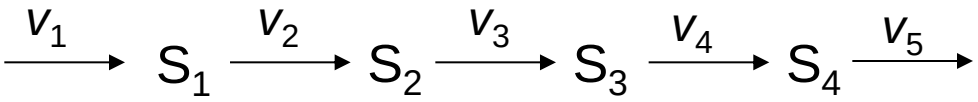
стехиометрическая матрица

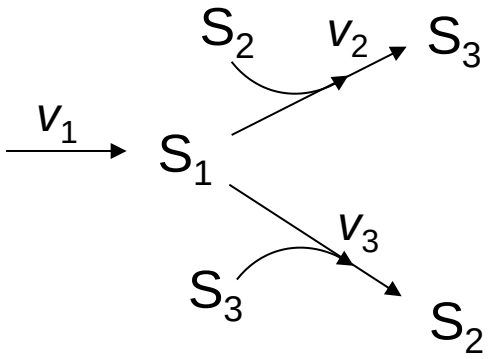
$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 1 \cdot v_1 - 1 \cdot v_2 + 0 \cdot v_3 - 1 \cdot v_4 \\ 0 \cdot v_1 + 2 \cdot v_2 - 1 \cdot v_3 + 0 \cdot v_4 \\ 0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 + 1 \cdot v_4 \end{pmatrix}$$

уравнения баланса

Записать уравнения баланса







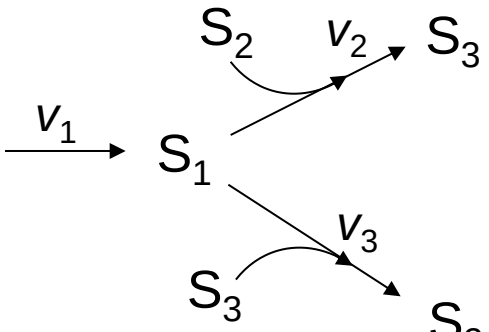
Записать уравнения баланса

$$S_1+S_2+S_3 \xleftrightarrow{v_1} S_4+S_5$$

$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \\ \frac{dS_4}{dt} \\ \frac{dS_5}{dt} \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \cdot (v_1) = \begin{pmatrix} -v_1 \\ -v_1 \\ -v_1 \\ v_1 \\ v_1 \end{pmatrix}$$

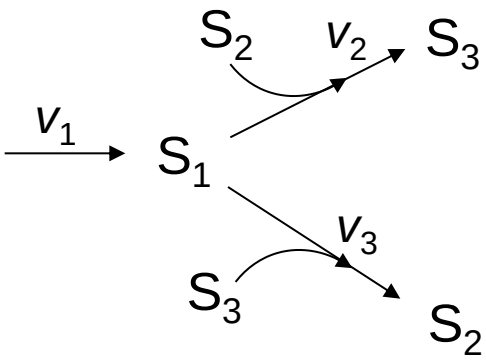
$$\xrightarrow{v_1} S_1 \xrightarrow{v_2} S_2 \xrightarrow{v_3} S_3 \xrightarrow{v_4} S_4 \xrightarrow{v_5}$$

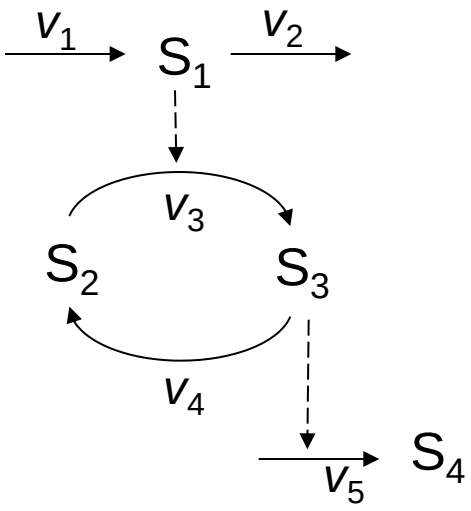
$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \\ \frac{dS_4}{dt} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} v_1 - v_2 \\ v_2 - v_3 \\ v_3 - v_4 \\ v_4 - v_5 \end{pmatrix}$$



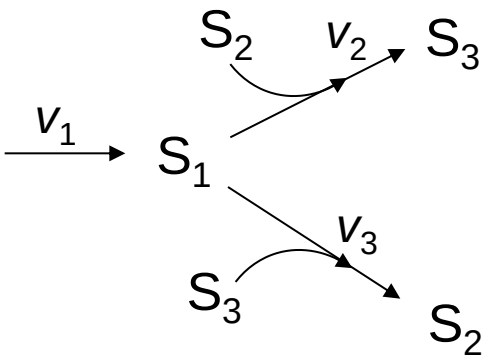
$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 - v_2 - v_3 \\ -v_2 + v_3 \\ v_2 - v_3 \end{pmatrix}$$

Записать уравнения баланса

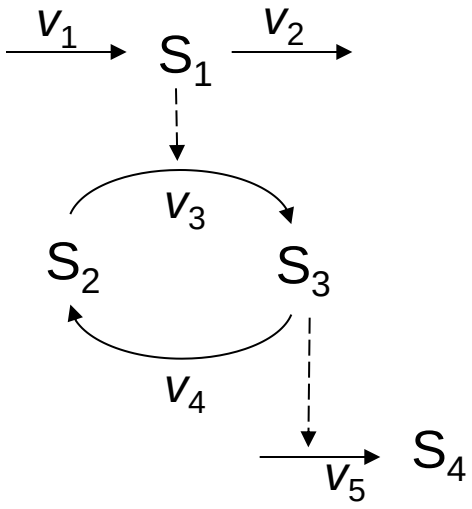




Записать уравнения баланса



$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 - v_2 - v_3 \\ -v_2 + v_3 \\ v_2 - v_3 \end{pmatrix}$$



$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \\ \frac{dS_4}{dt} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} v_1 - v_2 \\ -v_3 + v_4 \\ v_3 - v_4 \\ v_5 \end{pmatrix}$$

Построение стехиометрической модели

$S = (S_1, S_2, ..., S_n)^T$ вектор концентраций метаболитов

$v = (v_1, v_2, ..., v_n)^T$ вектор скоростей реакций

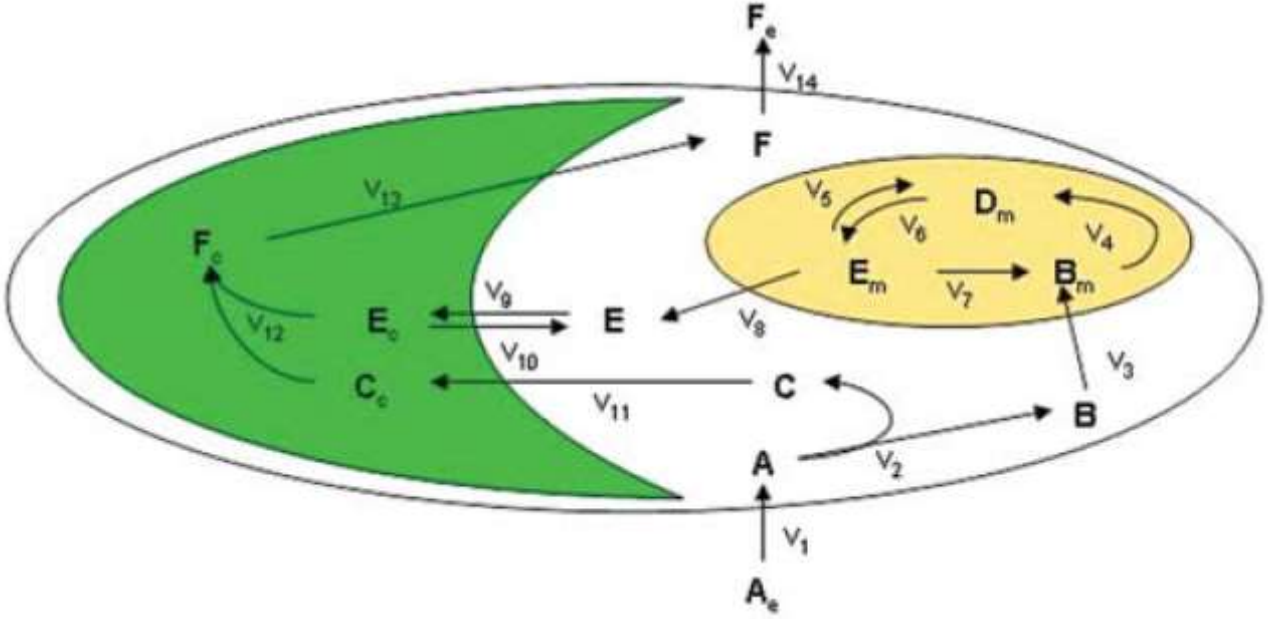
$\boldsymbol{N}$ стехиометрическая матрица

$\frac{dS}{dt} = Nv$ уравнения баланса

$\frac{dS}{dt} = 0$ стационарное состояние

$Nv = 0$ система алгебраических уравнений

Гипотетическая схема метаболических потоков



Метаболиты

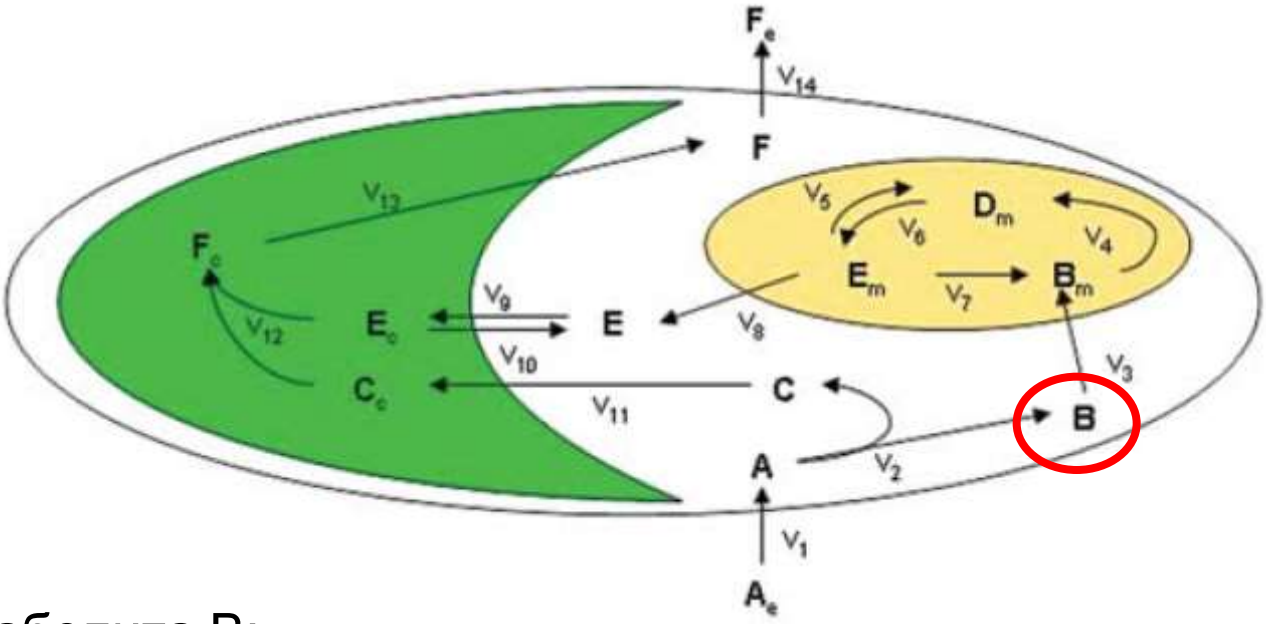
внешние:
A_e, F_e

внутренние:
A, B, C, E, F,– в цитоплазме
B_m, D_m, E_m – в митохондрии
C_c, E_c, F_c – в хлоропласте

Процессы

транспорт через внешнюю мембрану:
v_1, v_{14}
транспорт через внутренние мембраны:
v_3, v_8 – через мембрану митохондрии
v_9, v_{10}, v_{11}, v_{13} – через мембрану хлоропласта
реакции превращения субстрата в продукт
v_4, v_5, v_6, v_7, v_{12}

Гипотетическая схема метаболических потоков



для метаболита B:

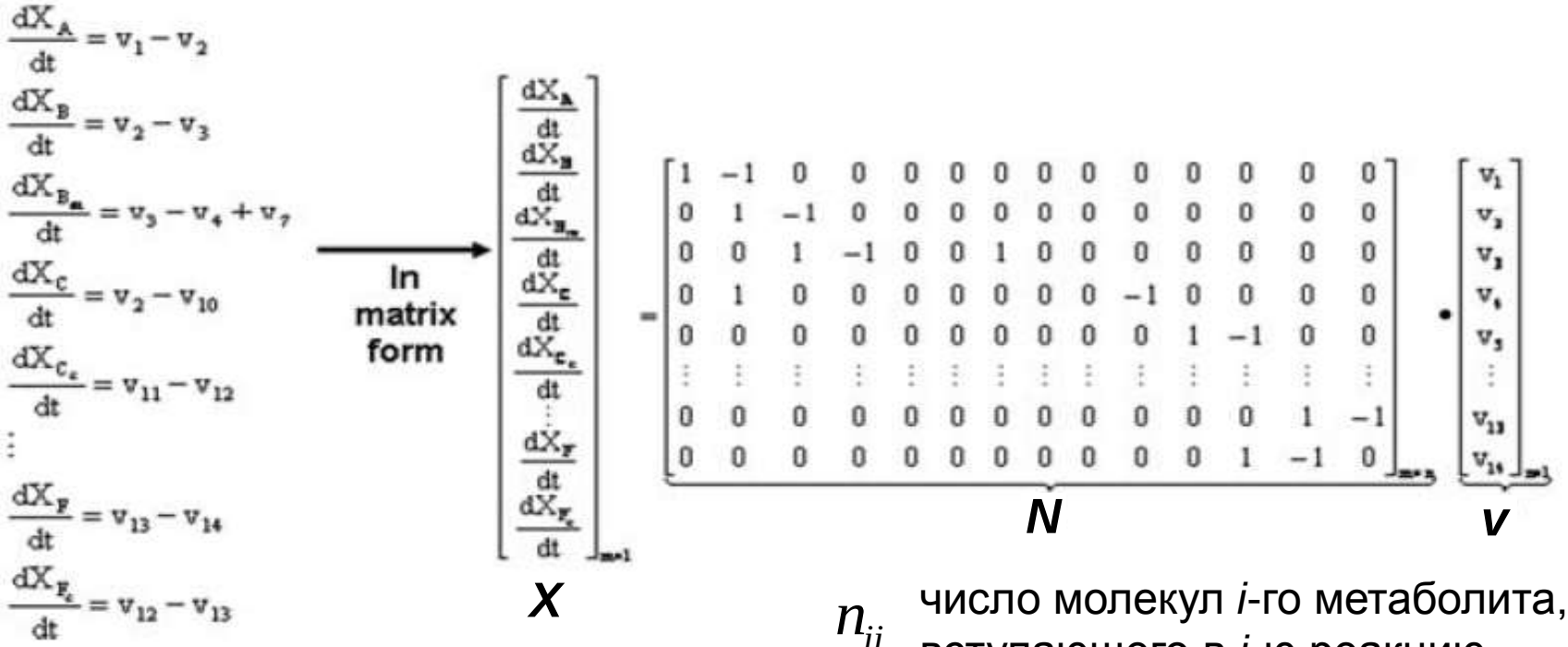
$$\frac{dX_B}{dt} = v_2 - v_3$$

(Общая скорость образования метаболита B) =
= (скорость образования B) – (скорость потребления B)

Метаболический поток – нормированная скорость реакции.

Размерность потока – моль/ед.биомассы/время.

В матричном виде систему дифференциальных уравнений можно записать как :



$$\frac{dX}{dt} = Nv \qquad Nv = 0$$

n_{ij} число молекул *i*-го метаболита, вступающего в *j*-ю реакцию

v_{ij} скорость *j*-й реакции, в которой участвует *i*-й метаболит

Размерность ***N*** в общем виде можно записать $m \times r$,
m – это число метаболитов, участвующих в реакциях,
r – число реакций.

Соотношение числа метаболитов *m* и числа потоков *r*

В условиях стационарного состояния: $$Nv = 0$$

m – количество уравнений
r – количество переменных

Возможны три варианта:

1. ***m = r*** полностью определенная система. Решение единственно.

2. ***m < r*** недоопределенная система.
Означает, что недостаточно уравнений баланса для определения всех потоков.
Система уравнений может иметь бесконечное число решений.
(Наиболее часто встречаемый случай)
Используется **метод линейного программирования.**

3. ***m > r*** переопределенная система.
В случае, когда доступно множество экспериментальных измерений.
Решается методом псевдо-инверсий Мура-Пенроуза (Moore–Penrose)

Постановка задачи линейного программирования для метаболической сети

(Метод анализа стационарных потоков, **Flux Balance Analysis, FBA**)

I. Определить пространство решений:

- Уравнения баланса массы для всех метаболитов.

$$N\nu = 0$$

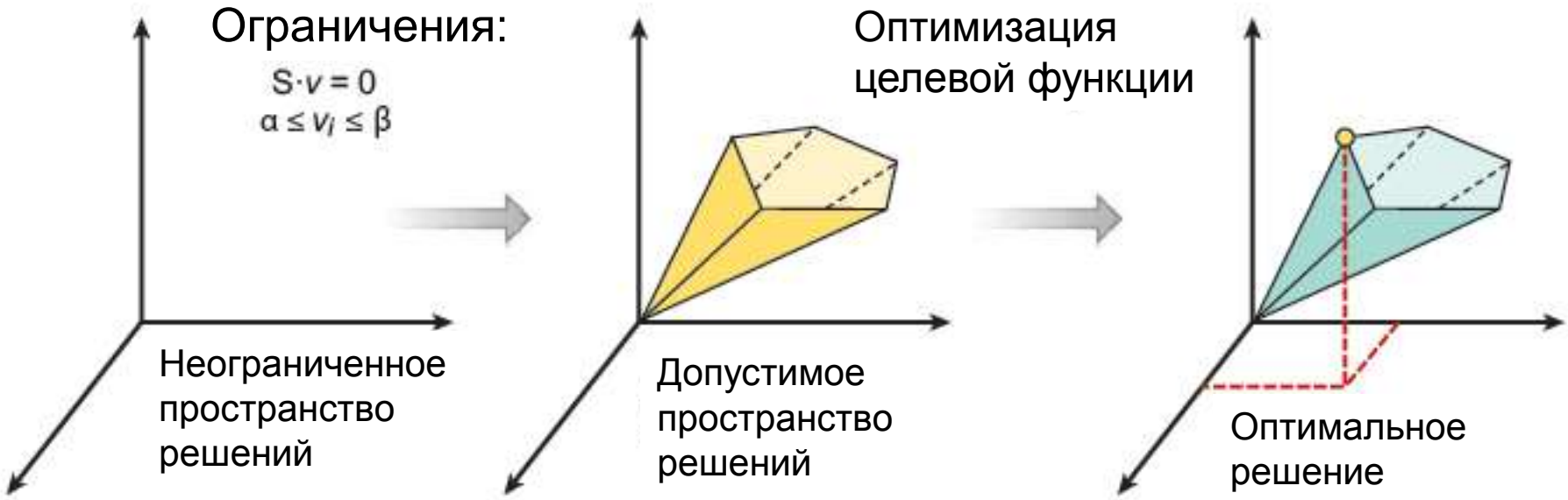
- Ограничения на потоки - максимальные и минимальные значения потоков (если известно).

$$\alpha_i \le \nu_i \le \beta_i$$

II. Определить целевую функцию.

$$z = \sum_j c_{ij} \nu_j$$

Поиск оптимального решения



Reed JL, Vo TD, Schilling CH, Palsson BO. 2003. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* 4:R54

Задача линейного программирования

Пример:

максимизировать $x_1 + 2x_2$ ← Целевая функция

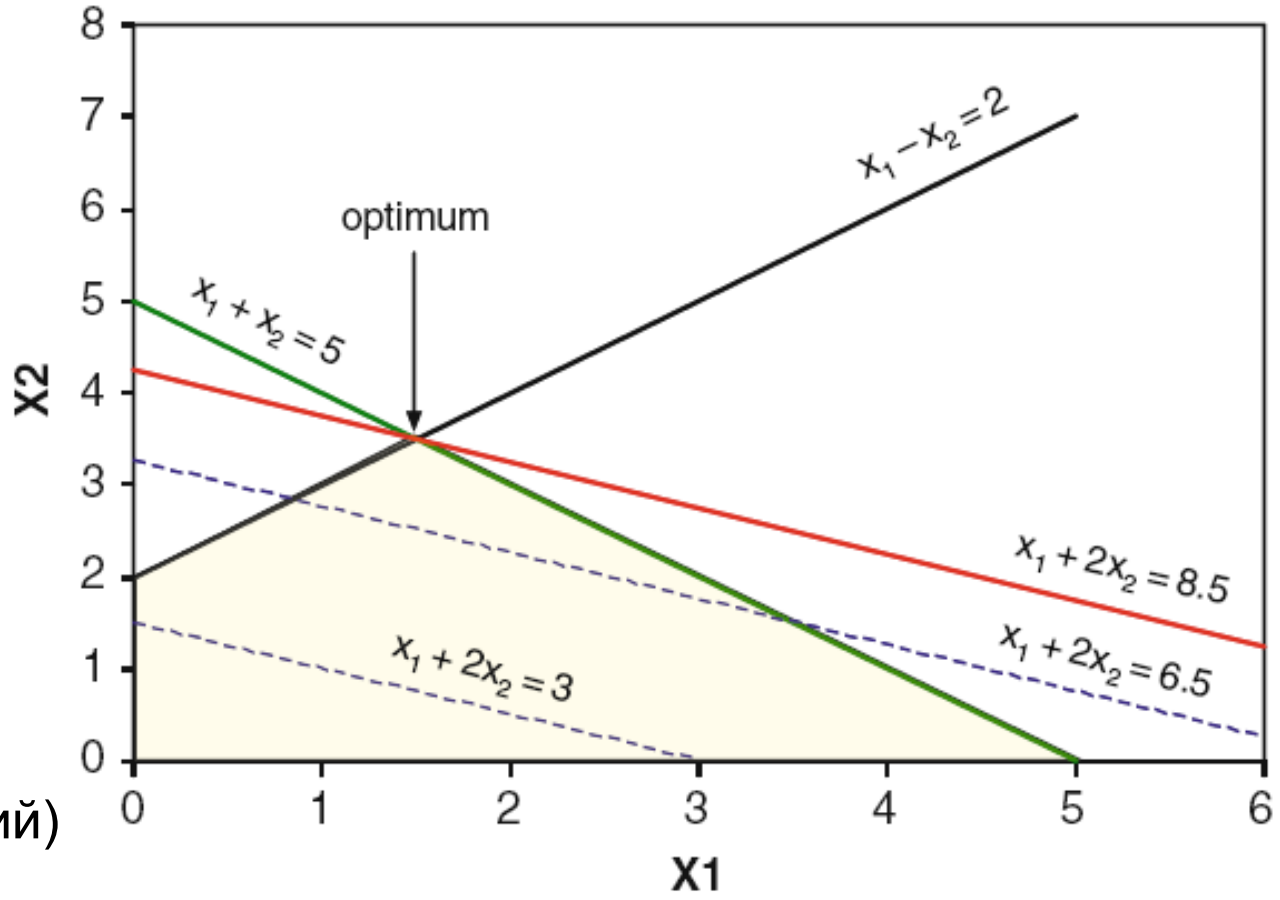
при условии:

$$x_1 - x_2 - 2 \leq 0$$

$$x_1 + x_2 - 5 \leq 0$$

$$x_1 \geq 0, x_2 \geq 0$$

Ограничения
(пространство решений)



Выбор целевой функции

Целевая функция определяется:

1. биологическими потребностями клетки,

2. биотехнологической задачей.

Примеры целевой функции:

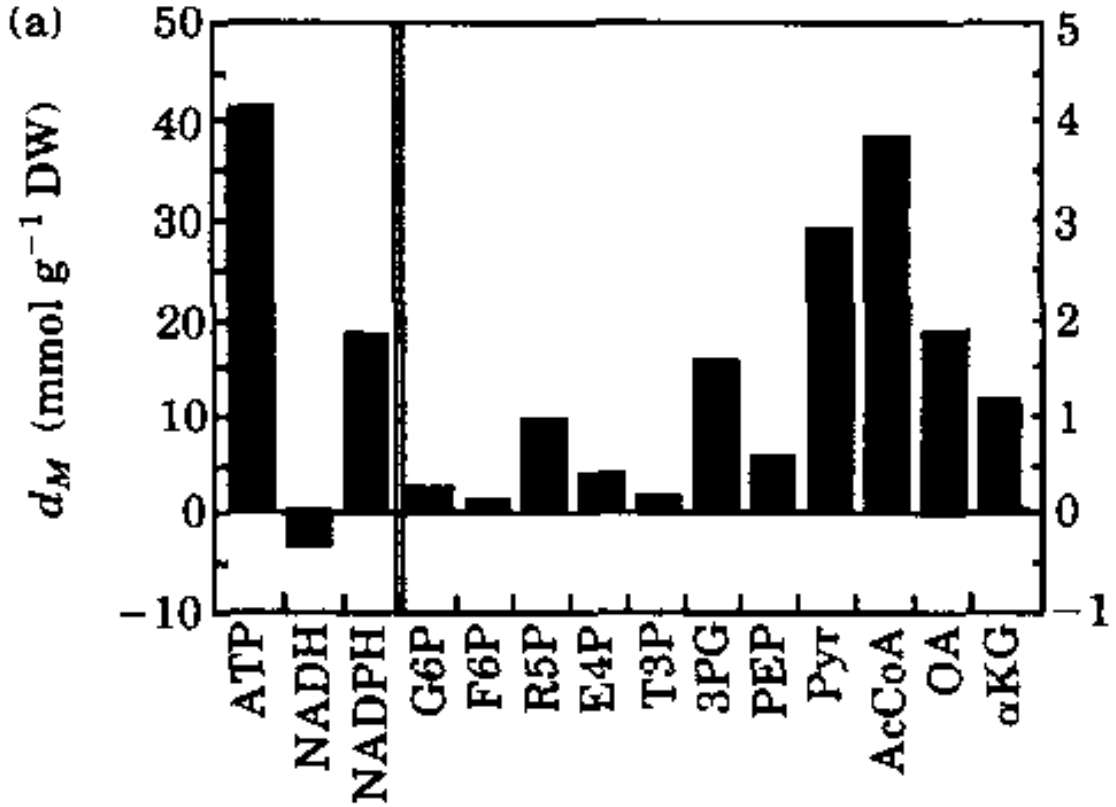
1. Максимизировать биомассу.

2. Минимизировать поток питательных веществ из среды.

3. Максимизировать продукцию определенного метаболита.

4. Максимизировать продукцию АТФ.

Расчет биомассы



Metabolic requirements for the production of 1 g of *E. coli* cells, from Ingraham *et al.* (1983).

Расчет биомассы:

a_1 белок + a_2 углеводы + a_3 липиды + a_4 ДНК + a_5РНК + b АТФ →

→ с биомассы+ b АДФ+ продукты

Анализ стационарных потоков

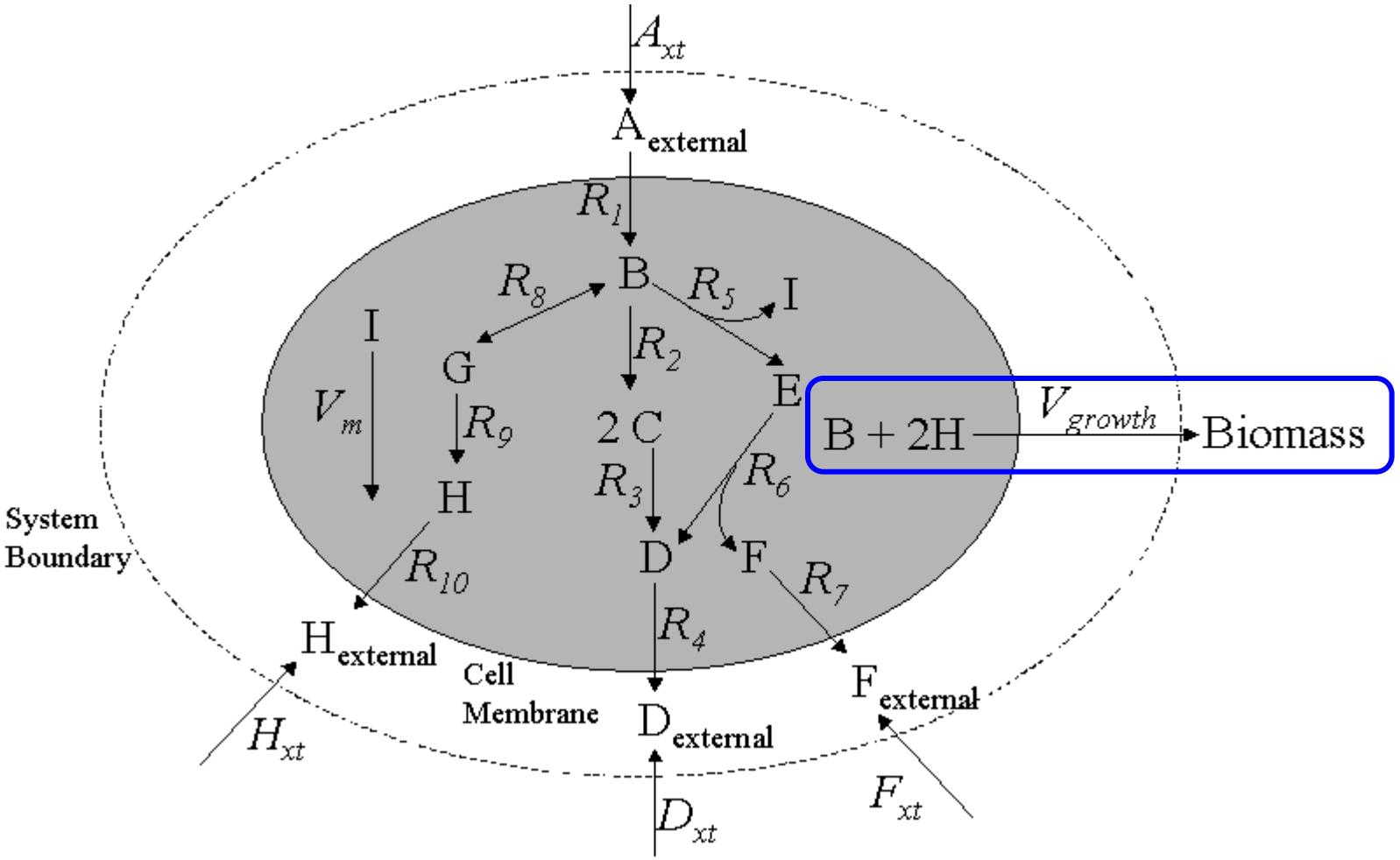
http://www.nature.com/nbt/web_extras/supp_info/nbt0201_125/info_frame.html

Gene	Enzyme	Flux
$Gene_1$	A Transporter	R_1
$Gene_2$	$Enzyme_2$	R_2
$Gene_3$	$Enzyme_3$	R_3
$Gene_4$	D Transporter	R_4
$Gene_5$	$Enzyme_5$	R_5
$Gene_6$	$Enzyme_6$	R_6
$Gene_7$	F Transporter	R_7
$Gene_8$	$Enzyme_8$	R_8
$Gene_9$	$Enzyme_9$	R_9
$Gene_{10}$	H Transporter	R_{10}
-	A Exchange	A_{xt}
-	D Exchange	D_{xt}
-	F Exchange	F_{xt}
-	H Exchange	H_{xt}

Выписываем все метаболические потоки

Анализ стационарных потоков

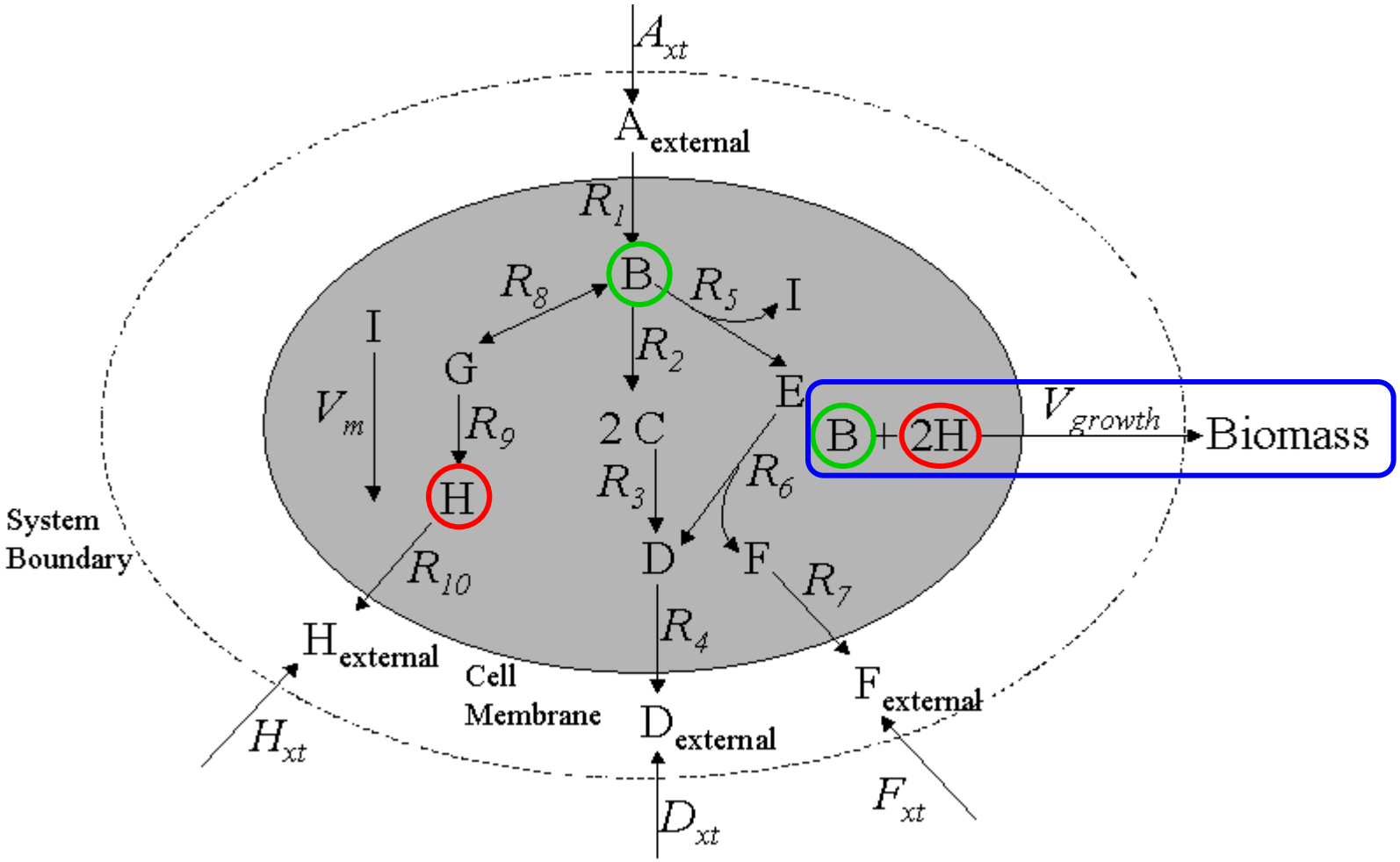
http://www.nature.com/nbt/web_extras/supp_info/nbt0201_125/info_frame.html



Выписываем все метаболиты и объединяем их в схему

Анализ стационарных потоков

http://www.nature.com/nbt/web_extras/supp_info/nbt0201_125/info_frame.html



Как записать поток на биомассу (V_{growth})?

Анализ стационарных потоков

http://www.nature.com/nbt/web_extras/supp_info/nbt0201_125/info_frame.html

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	V_m	V_{growth}	A_{xt}	D_{xt}	F_{xt}	H_{xt}			
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		R_1	0
B	1	−1	0	0	−1	0	0	−1	0	0	0	−1	0	0	0	0		R_2	0
C	0	2	−1	0	0	0	0	0	0	0	0	0	0	0	0	0		R_3	0
D	0	0	1	−1	0	1	0	0	0	0	0	0	0	0	0	0		R_4	0
E	0	0	0	0	1	−1	0	0	0	0	0	0	0	0	0	0		R_5	0
F	0	0	0	0	0	1	−1	0	0	0	0	0	0	0	0	0		R_6	0
G	0	0	0	0	0	0	0	1	−1	0	0	0	0	0	0	0		R_7	0
H	0	0	0	0	0	0	0	0	1	−1	0	−2	0	0	0	0		R_8	0
I	0	0	0	0	1	0	0	0	0	0	−1	0	0	0	0	0		R_9	0
$A_{external}$	−1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0		R_{10}	0
$D_{external}$	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0		V_m	0
$F_{external}$	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0		V_{growth}	0
$H_{external}$	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1		A_{xt}	0
																		D_{xt}	0
																		F_{xt}	0
																		H_{xt}	0

На формирование биомассы, V_{growth} ,
идет одна молекула B и две молекулы H

Анализ стационарных потоков

http://www.nature.com/nbt/web_extras/supp_info/nbt0201_125/info_frame.html

уравнения баланса

$$\mathbf{B}: R_1 - R_2 - R_5 - R_8 - V_{growth} = 0$$
$$\mathbf{C}: 2R_2 - R_3 = 0$$
$$\mathbf{D}: R_3 + R_6 - R_4 = 0$$
$$\mathbf{E}: R_5 - R_6 = 0$$
$$\mathbf{F}: R_6 - R_7 = 0$$
$$\mathbf{G}: R_8 - R_9 = 0$$
$$\mathbf{H}: R_9 - R_{10} - 2V_{growth} = 0$$
$$\mathbf{I}: R_5 - V_m = 0$$
$$\mathbf{A}_{external}: A_{xt} - R_1 = 0$$
$$\mathbf{D}_{external}: D_{xt} + R_4 = 0$$
$$\mathbf{F}_{external}: F_{xt} + R_7 = 0$$
$$\mathbf{H}_{external}: H_{xt} + R_{10} = 0$$

ограничения на потоки

$$0 \le R_1 \le \infty$$
$$0 \le R_2 \le \infty$$
$$0 \le R_3 \le \infty$$
$$0 \le R_4 \le \infty$$
$$0 \le R_5 \le \infty$$
$$0 \le R_6 \le \infty$$
$$0 \le R_7 \le \infty$$
$$-\infty \le R_8 \le \infty$$
$$0 \le R_9 \le \infty$$
$$0 \le R_{10} \le \infty$$
$$V_1 \le V_m \le V_1$$
$$0 \le V_{growth} \le \infty$$
$$Y_2 \le A_{up} \le Y_2$$
$$-\infty \le D_{up} \le 0$$
$$-\infty \le F_{up} \le 0$$
$$-\infty \le H_{up} \le 0$$

целевая функция

$$Z = V_{growth} \rightarrow \mathbf{max}$$

Ставим задачу линейного программирования

Применение анализа метаболических путей

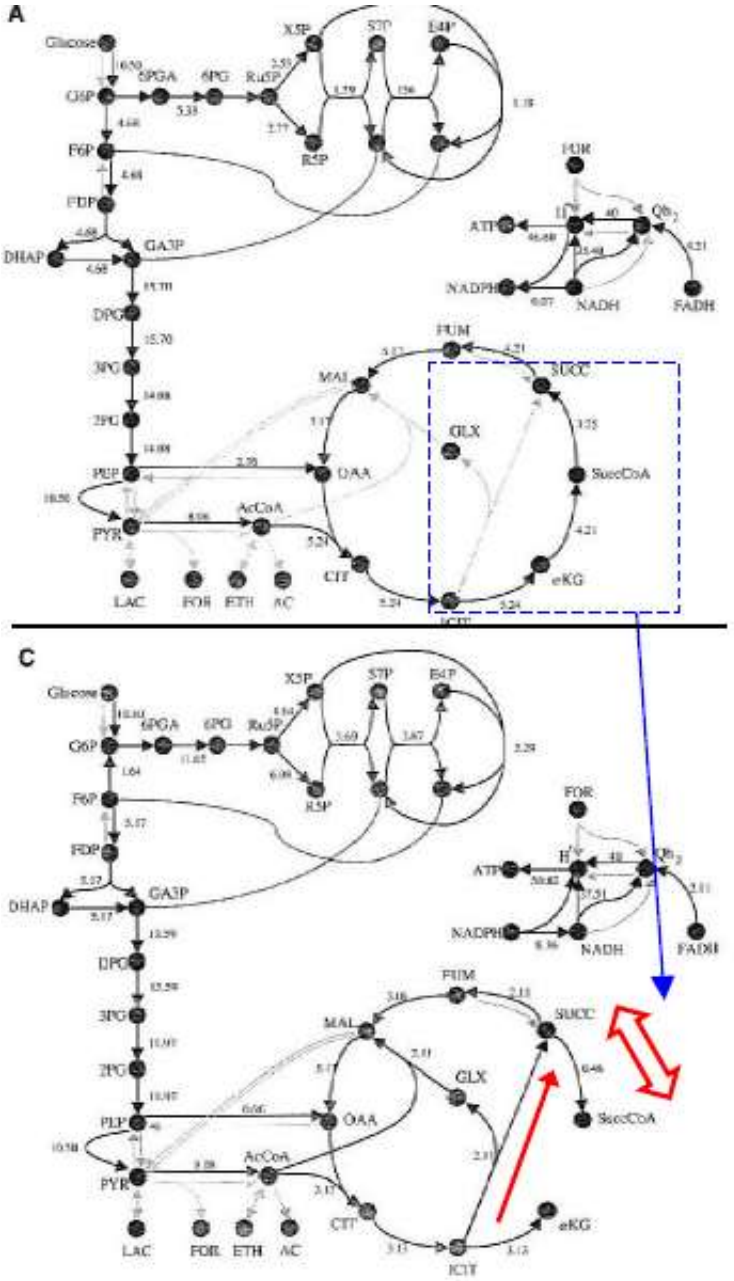
Метаболизм *E.coli*

Распределение метаболических потоков для полного набора генов.
Целевая функция – максимум биомассы при росте на глюкозе.

Способность E.coli адаптироваться к утрате функциональности гена может быть проанализирована путем удаления гена из основного набора

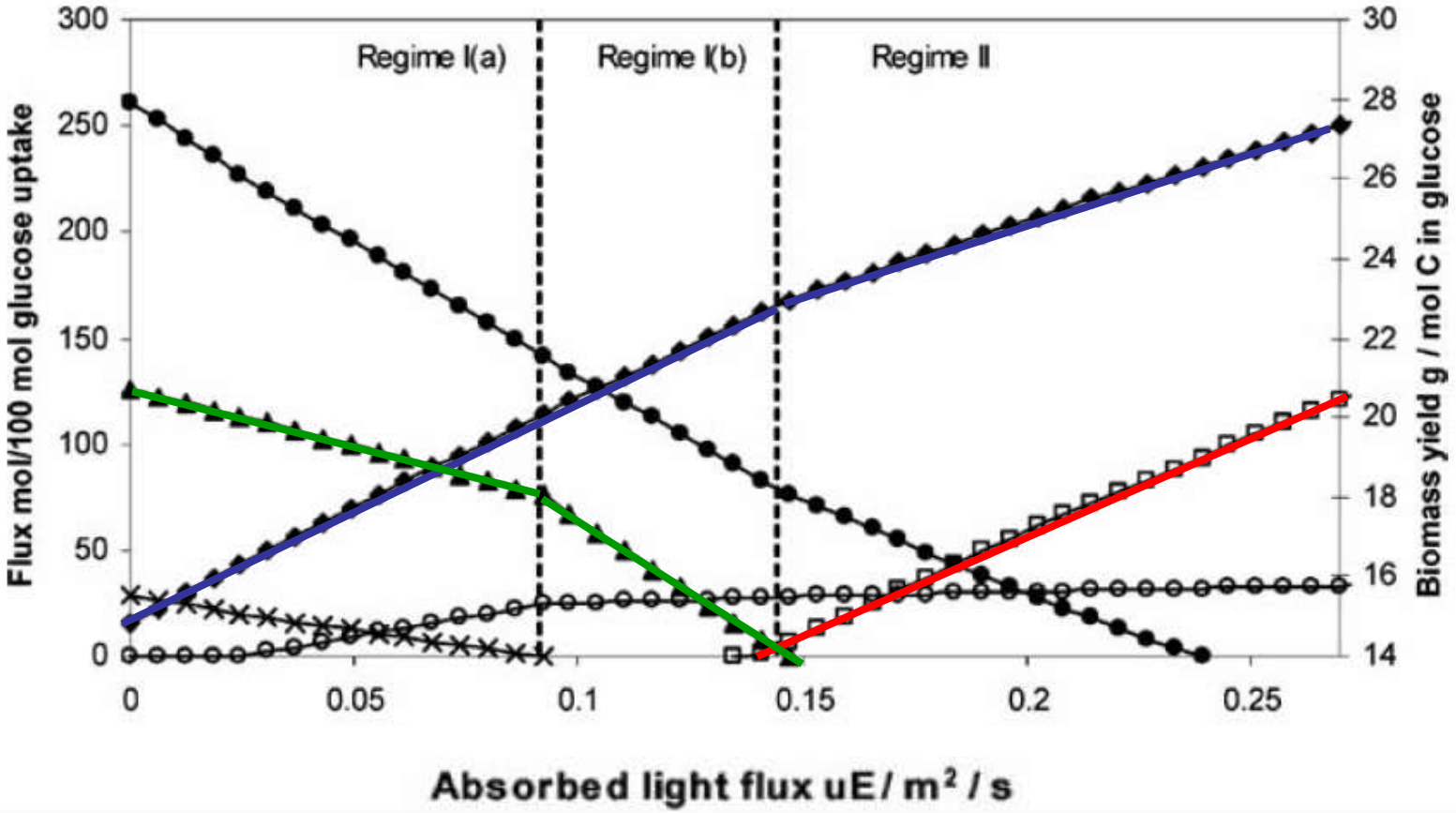
Распределение метаболических потоков, когда «выключен» ген sucA.
(ген sucA кодирует необходимый компонент реакции в ЦТК)

Поток перенаправляется через глиоксилатный шунт.
Реакция Succ—SucCoA обращается.



Применение анализа метаболических путей

Метаболизм синезеленых водорослей



N.R. Boyle et al J. Schwender (ed.), *Plant Metabolic Networks*, Springer Science+Business Media, LLC 2009