

Математическое моделирование метаболических путей

Общие подходы

Последние годы ознаменовались значительным прогрессом в нашем понимании процессов клеточного метаболизма, генной регуляции, передачи сигнала в клетке.

Информация о последовательности генов собрана в различные базы данных и позволяет реконструировать весь массив метаболических путей для отдельного организма. Так, например, построены метаболические карты для бактерии *E.coli*, которые в свою очередь собраны в базы данных, представляющие собой описание совокупности всех метаболических путей и регуляций между ними. Для объединения и осознания накопленного экспериментального материала возникла необходимость в специальных математических моделях. В результате появилась новая область в математическом моделировании – метаболическое моделирование.

Такие модели успешно применяются для анализа фенотипа организма при различных делециях генов, для изучения генной экспрессии, приводящей к определенной структуре метаболической сети, и для определения регуляторных генов.

Метаболическое моделирование является перспективным подходом для предсказания *in silico* функционирования клетки на основе взаимосвязи и взаимодействия всех клеточных компонент.

Существуют два больших класса моделей, используемых для описания и анализа метаболических сетей: стехиометрические и динамические. Стехиометрические модели представляют собой как бы детальный мгновенный снимок метаболических путей, находящихся в стационарном режиме, установившемся в данных условиях. Динамические модели описывают изменения метаболитов во времени используются для возможности предсказания изменений в метаболизме в ответ на действия факторов внешней среды или генетических модификаций.

Хорошие результаты получаются при сочетании этих двух подходов. Существует большое количество моделей, построенных на основе анализа метаболических потоков (Flux Balance Analysis - FBA). FBA не требует сложных математических выкладок и легко реализуется в виде программ для проведения вычислительных экспериментов. Основы этого метода изложены в следующем разделе.

Несмотря на широкое применение метод анализа метаболических потоков, строго говоря, ограничен ситуацией стационарного состояния. Более того, FBA позволяет определять лишь «чистые» потоки от одного метаболита к другому и не способен оценить вклад потока, направленного в обратную сторону.

Для экспериментального измерения метаболических потоков используются методы ЯМР, газовой хроматографии (ГХ), масс-спектрометрии (МС). Эти методы дают дополнительную информацию для FBA о соотношении метаболических потоков в клетке. Их комбинированное применение в сочетании с использованием меченого углерода ^{13}C позволяет рассчитать отношения потоков в ключевых точках метаболических путей и определить степень обратимости отдельных стадий в метаболической сети. Однако и в этом случае FBA применим только для случаев стационарного состояния.

Для того, чтобы предсказать поведение системы во времени (изменение концентраций метаболитов и ферментов во времени) стехиометрическая модель должна быть дополнена описанием кинетических свойств системы.

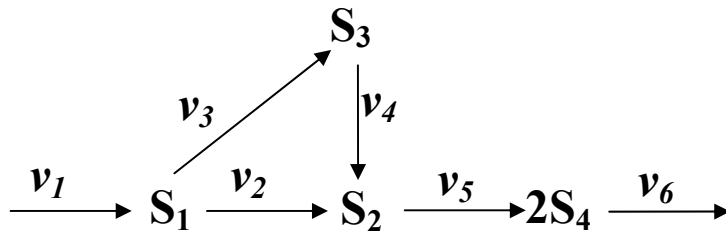
Ключевым отличием между рассмотренной выше традиционной системой химических реакций и метаболической сетью является наличие регуляторных связей и генетический контроль активности метаболитов. В биологических системах все уровни метаболического функционирования (транскрипция, трансляция и каталитическая активность) тесно взаимосвязаны, обеспечивая, тем самым, возможность адаптации к изменяющимся внешним условиям.

В тех случаях, когда существует подробная информация о механизмах генетической и пост-трансляционной регуляции на уровне закона действующих масс могут быть построены общие кинетические модели, описывающие регуляцию ферментативной активности.

Анализ баланса метаболических потоков

(задача линейного программирования ЛП)

Для математического описания и анализа метаболических потоков используется подход, основанный на предположении о балансе метаболитов и применении метода линейного программирования, называемый методом анализа стационарных метаболических потоков (в англоязычной литературе Flux Balance Analysis, FBA). Под метаболическим потоком понимается скорость превращения вещества в метаболическом пути. Проиллюстрируем этот метод на простом примере. Рассмотрим схему метаболических реакций, $v_1 \dots v_6$, в результате которых исходный метаболит S_1 превращается в метаболит S_4 , являющимся конечным продуктом.



Здесь S_i – метаболиты, v_i – скорости реакций.

Будем считать, что скорости, направленные на образование метаболита, имеют положительный знак, а скорости реакций, в которых метаболит расходуется, имеют отрицательный знак. Запишем в общем виде систему дифференциальных уравнений, описывающую изменения во времени всех метаболитов:

$$\begin{aligned}
 \frac{dS_1}{dt} &= v_1 - v_2 - v_3 \\
 \frac{dS_2}{dt} &= v_2 + v_4 - v_5 \\
 \frac{dS_3}{dt} &= v_3 - v_4 \\
 \frac{dS_4}{dt} &= 2v_5 - v_6
 \end{aligned}
 \tag{1}$$

Метод анализа стационарных потоков (FBA) рассматривает стационарную ситуацию, когда в системе установилось равновесие. Для каждого метаболита существует баланс метаболических потоков, то есть, все потоки, направленные на образование конкретного метаболита уравновешены потоками, в результате которых концентрация этого метаболита уменьшается. Сумма всех потоков, участвующих в изменении концентрации данного метаболита, таким образом, равна нулю. Для системы дифференциальных уравнений это означает, что сумма всех скоростей реакций, в которых участвует данный метаболит, равна нулю, тогда все правые части уравнений приравниваются нулю, и система дифференциальных уравнений вырождается в систему алгебраических уравнений. Такой подход позволяет оценить распределение метаболических потоков (или стационарных скоростей) и их относительный вклад в образование данного метаболита. Запишем уравнения баланса для всех метаболитов таким образом, чтобы отразить степень участия каждого метаболита в каждой реакции.

$$\text{Для первого метаболита, } S_1: 1 \cdot v_1 - 1 \cdot v_2 - 1 \cdot v_3 + 0 \cdot v_4 + 0 \cdot v_5 + 0 \cdot v_6 = 0.$$

$$\text{Для второго метаболита, } S_2: 0 \cdot v_1 + 1 \cdot v_2 + 0 \cdot v_3 + 1 \cdot v_4 - 1 \cdot v_5 + 0 \cdot v_6 = 0. \tag{2}$$

$$\text{Для третьего метаболита, } S_3: 0 \cdot v_1 + 0 \cdot v_2 + 1 \cdot v_3 - 1 \cdot v_4 + 0 \cdot v_5 + 0 \cdot v_6 = 0.$$

$$\text{Для четвертого метаболита, } S_4: 0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 + 0 \cdot v_4 + 2 \cdot v_5 - 1 \cdot v_6 = 0.$$

Здесь множитель 0 обозначает, что метаболит не участвует в данной реакции, множители 1 и 2 обозначают, что в реакциях образования (знак «+») или реакциях распада (знак «-») участвует одна или соответственно две молекулы метаболита.

Соответствующие стехиометрическая матрица и вектор скоростей реакций выглядят как

$$\mathbf{N} = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix}.$$

Тогда систему уравнений можно представить в матричном виде:

$$\begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = 0 \quad (3)$$

(По правилу умножения матрицы на вектор элемент строки матрицы умножается на элемент вектора- столбца). Каждая колонка стехиометрической матрицы представляет реакцию, а элементы ряда представляют вклад определенного метаболита с учетом стехиометрии в каждую из реакций. Умножая матрицу стехиометрических коэффициентов на вектор скоростей, получим систему алгебраических уравнений (2).

Представляя вектор метаболитов как $\mathbf{S} = (S_1, S_2, S_3, S_4)$ исходную систему дифференциальных уравнений (1) можно записать как

$$\frac{d\mathbf{S}}{dt} = \mathbf{N}\mathbf{v} \quad (4)$$

$$\text{и в стационарных условиях: } \mathbf{N}\mathbf{v} = 0 \quad (5)$$

Такой вид записи используется в большинстве вычислительных программ, разработанных для расчета распределения метаболических потоков в стационарном состоянии. Таким образом, сформулирована задача для поиска стационарных метаболических потоков (стационарных скоростей реакций), $v_1 \dots v_6$.

Полученная для рассматриваемой схемы система уравнений не может быть решена однозначно, поскольку число переменных, равное числу стационарных скоростей (6) в ней больше, чем число уравнений, равное числу метаболитов (4). Однако можно

поставить задачу оптимизации (задачу линейного программирования), то есть выделить некое пространство, в котором мы сможем найти максимум или минимум функционала, определяемого из нашей системы. Такое пространство выделяется естественным образом, поскольку величины потоков всегда ограничены положительными значениями. Для оптимизации выбирается целевая функция, обычно это один или сумма нескольких потоков, суммарная величина которых по биологическому смыслу задачи должна быть максимальной (минимальной). Пусть в рассматриваемой схеме нас интересует максимальный выход конечного продукта S_4 , определяемый скоростью v_6 . Тогда, задача линейного программирования для рассматриваемой схемы будет выглядеть следующим образом:

Найти максимум целевой функции z ,

$$z = v_6 \rightarrow \max$$

при следующих условиях:

$$\begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = 0$$

$$0 \leq v_1 \dots v_6 \leq \infty.$$

Поставим задачу оптимизации для произвольного числа метаболитов и реакций.

Первым шагом является построение стехиометрической матрицы на основании схемы метаболических потоков. Предполагается, что концентрации всех интермедиатов находятся в псевдо-стационарном равновесии, а изменения, связанные с ростовыми процессами считаются незначительными. Стехиометрическая модель дает набор алгебраических уравнений, отражающих существование концентрационного баланса для всех рассматриваемых метаболитов

$$n_{11} \cdot v_{11} + n_{12} \cdot v_{12} + \dots + n_{1r} \cdot v_{1r} = 0$$

$$n_{21} \cdot v_{21} + n_{22} \cdot v_{22} + \dots + n_{2r} \cdot v_{2r} = 0$$

.....

$$n_{m1} \cdot v_{m1} + n_{m2} \cdot v_{m2} + \dots + n_{mr} \cdot v_{mr} = 0$$

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

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

Полученную систему уравнений можно представить в виде:

$$\mathbf{N}\mathbf{v} = 0.$$

Матрица \mathbf{N} содержит стехиометрические коэффициенты, вектор \mathbf{v} представляет величины потоков метаболических реакций. Размерность \mathbf{N} в общем виде можно записать $m \times r$, где m - это число метаболитов, участвующих в реакциях, а r - число реакций.

Поскольку число реакций обычно больше, чем число внутриклеточных метаболитов ($n > m$), система уравнений, полученная на основе стехиометрической матрицы является недоопределенной, и поэтому существует бесконечное число решений, т.е. бесконечное число возможных распределений потоков. В численных экспериментах определенное распределение потоков может быть найдено, если сформулировать соответствующую задачу линейного программирования или задачу оптимизации, которую часто называют анализом стационарных потоков FBA. Чтобы сформулировать задачу линейного программирования, необходимо задать целевую функцию z , которая, например, может описывать максимальный выход продукта с учетом всех реакций v_i , в которых он образуется.

Например, найти максимум функции

$$z = \sum_i v_i \rightarrow \max \quad (6)$$

при следующих условиях:

$$\mathbf{N}\mathbf{v} = 0 \quad (7)$$

$$\alpha_i \leq v_i \leq \beta_i \quad (8)$$

Линейные неравенства $\alpha_i \leq v_i \leq \beta_i$ используются для получения дополнительных ограничений на отдельные потоки. Ограничения устанавливают верхнюю и нижнюю границы изменения величины потока, для каждой реакции. α_i - нижняя граница потока v_i и β_i - его верхняя граница. Если нет никакой информации о величине потока, то в случае необратимых потоков значение α_i приравнивается нулю. Для всех других случаев α_i и β_i не ограничены слева и могут принимать как положительные, так и отрицательные значения. Далее с помощью метода ЛП исследуются свойства системы. Реакция, выбранная в качестве целевой функции, оптимизируется (ищется максимум или минимум). Полученное решение удовлетворяет всем ограничениям (7) и (8). Результатом

является оптимальное распределение потоков, такое, что через выбранную целевую функцию будет осуществляться максимальный (минимальный) поток данного продукта.

С помощью метода анализа стационарных потоков можно теоретически оценить различные режимы работы метаболической сети. Данный метод базируется на естественных физико-химических ограничениях метаболической сети и требует только информации о стехиометрии метаболических путей.

В качестве примера рассмотрим окислительно-восстановительный статус клетки, а именно, состояние, связанное с присутствием NAD(P)H.

Первым шагом для исследования состояния NAD(P)H в клетке является построение схемы центрального метаболизма растительной клетки, а именно, гликолиза, и пентозофосфатного пути.

Для анализа состояния NAD(P)H в клетке исследуем два возможных распределения метаболических потоков. В первом случае метаболический поток через пентозофосфатный путь составляет 10% от потока через гликолитический путь. Во втором случае исследуем возможность реализации такого распределения потоков, при котором выход NAD(P)H является максимальным.

Каждой i -той скорости реакции сопоставляется переменная v_i (рис. III.15). При этом часть последовательных стадий от GAP до P_{ru} может быть редуцирована в одну, поскольку результат оптимизации в данном случае не зависит от числа последовательных стадий.

Редуцированный метаболический путь

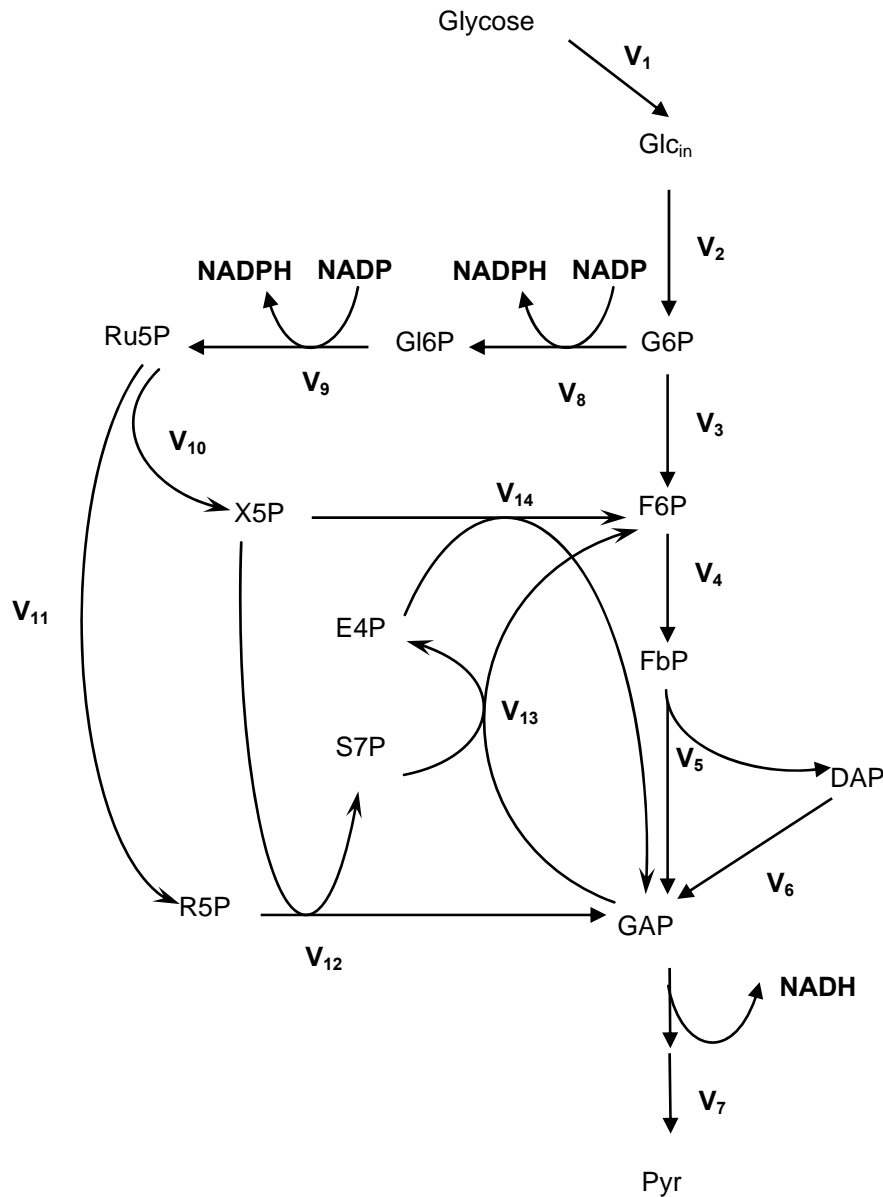


Рис. III.15. Редуцированный метаболический путь для решения задачи оптимизации. Каждой i -той реакции в соответствие поставлен поток v_i .

Далее исходя из основного принципа метода стационарных потоков, предполагается, что все потоки находятся в квази-стационарном состоянии. Тогда, учитывая, что величина любого потока определяется неотрицательным числом, можно записать систему уравнений и неравенств.

Далее определяем целевую функцию для окончательной формулировки задачи оптимизации. Как можно видеть из рисунка III.15, основными стадиями образования NAD(P)H являются реакции v_7 , v_8 и v_9 . Поэтому для анализа максимально возможного выхода NAD(P)H выбираем следующую целевую функцию: $z = v_7 + v_8 + v_9$.

В качестве исходного распределения потоков рассмотрим случай, когда 90% метаболического потока идет через гликолиз и только 10% - через пентозофосфатный путь. Все потоки метаболического пути нормируем на величину входящего потока для получения относительных единиц измерения. Результат исходного распределения представлен на рисунке (III.16а). Далее будем максимизировать целевую функция z . Результат распределения потоков при максимальном выходе NAD(P)H представлен на рисунке (III.16б).

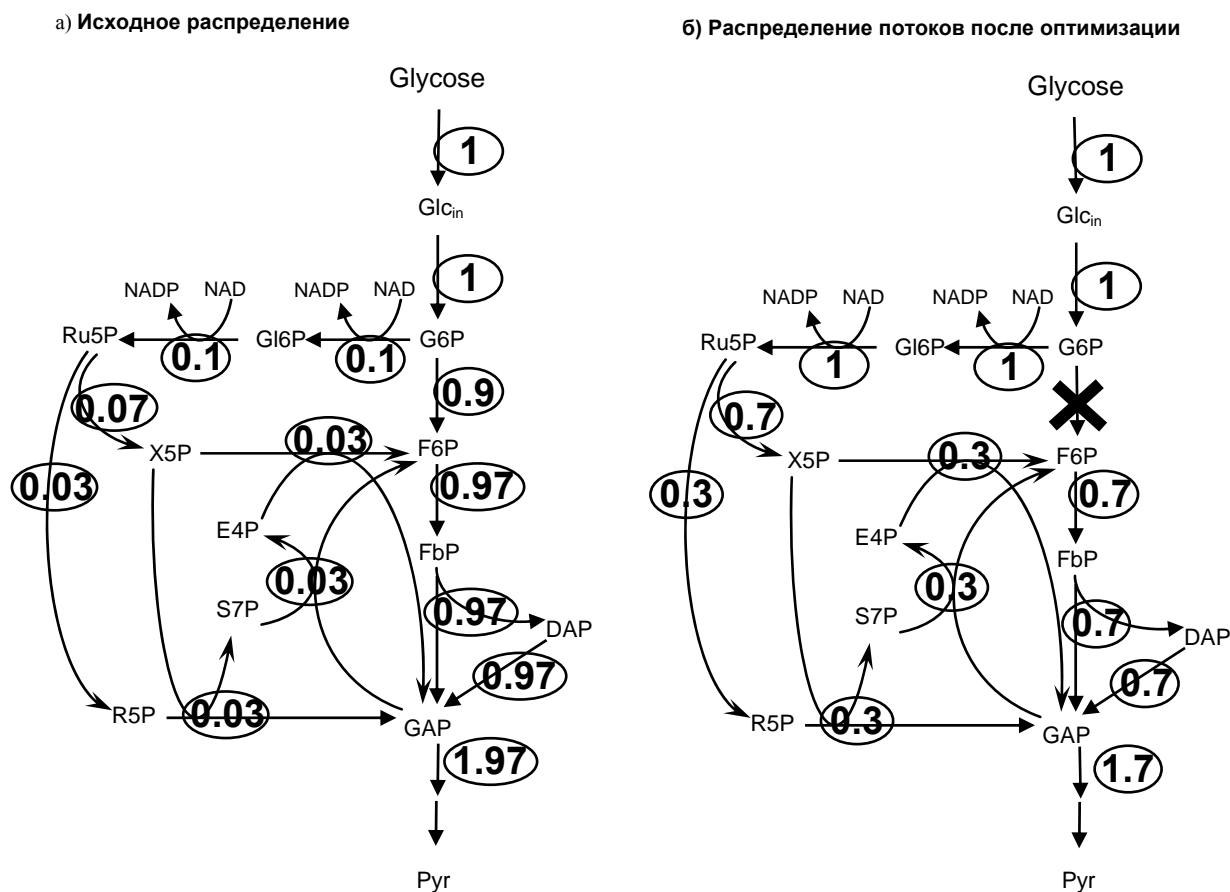


Рис. III.16 а) исходное распределение потоков. б) распределение для случая максимального выхода NAD(P)H. Относительные значения потоков приведены в кружках.

Как можно видеть, в случае максимального выхода NAD(P)H весь метаболический поток идет через пентозофосфатный шунт, при этом гликолитическая реакция, катализируемая фосфоглюкоизомеразой, перестает «работать», что соответствует выключению *pgi* гена.

С помощью анализа метаболических потоков в центральном метаболическом пути мы попытались оценить состояние NAD(P)H в различных ситуациях. Можно предположить, что в случае, когда требуется увеличение уровня NAD(P)H, происходит инактивация *pgi* гена, что ведет блокированию соответствующе реакции в гликолизе и перераспределению метаболических потоков, такому, что максимальный поток идет через пентозофосфатный шунт.

Список литературы

1. <http://www.brenda.uni-koeln.de>
2. <http://biocyc.org>
3. Aiba, S. & Matsuoka, M. (1979) Identification of metabolic model: citrate production from glucose by *Candida lipolytica*. *Biotechnol. Bioeng.* 21, 1373–1386.
4. Papoutsakis E.T. (1984) Equations and calculations for fermentations of butyric acid bacteria. *Biotechnol. Bioeng.* 26, 174-187.
5. Papoutsakis E.T., Meyer C.L. (1985) Fermentation equations for propionic-acid bacteria and production of assorted oxychemicals from various sugars. *Biotechnol. Bioeng.* 27, 67-80.
6. Papoutsakis E.T., Meyer C.L. (1985) Equations and calculations for products yields and preferred pathways for butanediol and mixed-acid fermentations. *Biotechnol. Bioeng.* 27, 50-66.
7. Cortassa S., Aon J.C., Aon M.A., (1995) Fluxes of carbon, phosphorylation, and redox intermediates during growth of *Saccharomyces cerevisiae* on different carbon sources. *Biotechnol. Bioeng.* 47, 193-208
8. Vallino J.J. Stephanopoulos G. (1987) Intelligent sensors in biotechnology applications for the monitoring of fermentations and cellular metabolism. *Annals NY Acad Sci*, 506, 4215-4230.
9. Vallino J.J. Stephanopoulos G. (1994) Carbon flux distribution at the pyruvate branch point in *Corynebacterium glutamicum* during lysine overproduction *Biotechnol Prog*, 10, 327-334.
10. Park S.M., Sinskey A.J., Stephanopoulos G. (1997) Metabolic and physiological studies of *Corynebacterium glutamicum* mutants. *Biotechnol. Bioeng.* 55, 864-879.
11. Nyberg G.B., Balcarcel R.R., Follstad B.D., Stephanopoulos G., Wang D.I.C. (1999) Metabolism of peptide amino acid by Chinese Hamster ovary cells grown in a complex medium. *Biotechnol. Bioeng.* 62, 324-335.
12. Jorgensen H., Nielsen J., Villadsen H. (1995) Metabolic flux distributions in *Penicillium chrysogenum* during fed-batch cultivation. *Biotechnol. Bioeng.* 46, 117-131.
13. Sonntag K., Eggling L., de Gaaf A.A., Sahm H. (1993) Flux partitioning in the split pathway of lysine biosynthesis in *Corynebacterium glutamicum*: quantification by ¹³C- and ¹H-NMR spectroscopy. *Eur J Biochem*, 213, 1325-1331.
14. Marx A., de Graaf a.A., Wiechert W., Eggling L., Sahm H. (1996) Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 49, 111-129.
15. Zupke C., Stephanopoulos G. (1995) Intracellular flux analysis in hybridomas using mass balances and *in vitro* ¹³C-NMR. *Biotechnol. Bioeng.* 45, 292-303.

16. Schmidt K., Marx A. de Graaf A.A., Wiechert W., Sahn H., Nielsen J., Villadsen J. (1998) ^{13}C tracer experiments and metabolite balancing for metabolic flux analysis: comparing two methods. *Biotechnol. Bioeng.* 58, 254-257.
16. Domach M.M., Leung S.K., Cahn R.E., Cocks G.C., Shuler M.L. (1984) Computer model for glucose-limited growth of a single copy of *Escherichia coli* B/r-A. *Biotechnol. Bioeng.* 26, 203-216.
17. Peretti S.W., Bailey J.E. A mechanistically detailed model of cellular metabolism of glucose-limited growth of *Escherichia coli* B/r-A. *Biotechnol. Bioeng.* 28, 1672-1689.
18. Peretti S.W., Bailey J.E., Lee J.J. (1989) Transcription from plasmid genes, macromolecular stability and cell specific productivity in *Escherichia coli* carrying copy number mutant plasmids. *Biotechnol. Bioeng.* 34, 902-908.
19. Rizzi M., Baltes M., Theobald U., Reuss M. (1997) *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II mathematical model. *Biotechnol. Bioeng.* 55, 592-608.
20. Daae E.B., Gruys K.J., Dunnill P. (1999) Metabolic modeling as an experimental planning tool for the production of polyhydroxyalkanoate co-polymers in plants. *Metab Eng.*
21. Pissara P.N., Nielsen J., Bazin M.J. (1996) Pathway kinetics and metabolic control analysis of a high yielding strain of *Penicillium chrysogenum* during fed-batch cultivation. *Biotechnol. Bioeng.* 51, 168-176.
22. Hatzimanikatis V., Floudas C., Bailey J.E. (1996) Analysis and design of metabolic reaction networks via mixed-integer linear optimization. *AIChE J.* 42, 1277-1292.
23. Stephanopoulos, G.N., Aristidou, A.A. & Nielsen, J. (1998) *Metabolic Engineering – Principles and Methodologies*. Academic Press, San Diego, USA.
24. Nielsen, J. & Villadsen, J. (1994) *Bioreaction Engineering Principles*. Plenum Press, New York.
25. Bertsimas, D. & Tsitsiklis, J.N. (1997) *Introduction to Linear Optimization*. Athena Scientific, Belmont.
26. Varma, A. & Palsson, B.O. (1994) Metabolic flux balancing: basic concepts, scientific and practical use. *Bio/Technology* 12, 994–998.
27. Chvatal, V. (1983) *Linear programming*. W. H. Freeman and Company, New York, N.Y.
28. Edwards, J. S., R. U. Ibarra, and B. O. Palsson. (2001) *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* 19:125-130.
29. Edwards, J. S., and B. O. Palsson. (2000) The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* 97:5528-5533.