НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
сентябрь–октябрь 2026 Том 26 № 5 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
September–October 2026 Vol. 26 No 5 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

COMPUTER SCIENCE

doi: 10.17586/2226-1494-2026-26-5-945-953
УДК 004.94

Метод метаболических потенциалов для учета ациклических ограничений при моделировании стационарных метаболических потоков с помощью целочисленного линейного программирования

Иван Андреевич Сафонов[1], Алексей Александрович Сергушичев[2]✉

[1] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
[2] Университет Вашингтона в Сент-Луисе, Сент-Луис, 63110, США

[1] isafonov27@gmail.com, https://orcid.org/0009-0006-8483-7415
[2] asergushichev@wustl.edu✉, https://orcid.org/0000-0003-1159-7220

Аннотация

Введение. При моделировании стационарных метаболических потоков в метаболических сетях широко применяется метод баланса потоков, который, однако, в классической формулировке допускает нефизичные решения с термодинамически невозможными циклами. Для учета соответствия законам термодинамики добавляются ациклические ограничения, что приводит задачу к классу NP-трудных. Это делает актуальной разработку эффективных формулировок в виде задачи целочисленного линейного программирования. **Метод.** Предложен новый эффективный метод метаболических потенциалов для задачи метода баланса потоков с ациклическими ограничениями, который основан на новой формулировке в виде задачи целочисленного линейного программирования. Проведенный анализ метода Fast-SNP, который ускорил моделирование на порядки, позволил выделить и оценить вклад двух оптимизационных идей: использования разреженных ограничений и сокращения числа целочисленных переменных. Представленный метод метаболических потенциалов применяет эти идеи и строит эквивалентную формулировку ациклических ограничений для задачи целочисленного линейного программирования. **Основные результаты.** Выполнено сравнение эффективности метода метаболических потенциалов на метаболических моделях из базы BiGG. Предложенный метод демонстрирует сопоставимую с методом Fast-SNP эффективность на метаболических моделях малого и среднего размеров и улучшение времени примерно на 30 % на большой метаболической модели клеток человека RECON1. Показано, что сокращение числа целочисленных переменных дает наибольший вклад в оптимизацию эффективности. **Обсуждение.** Разработанный метод метаболических потенциалов сохраняет преимущества оптимизационных идей метода Fast-SNP и имеет такую же или лучшую эффективность, поэтому может быть использован в качестве замены метода Fast-SNP и других стандартных методов в задаче метода баланса потоков с ациклическими ограничениями.

Ключевые слова

метаболическое моделирование, метод баланса потоков, стационарные метаболические потоки, ациклические ограничения, целочисленное линейное программирование

Ссылка для цитирования: Сафонов И.А., Сергушичев А.А. Метод метаболических потенциалов для учета ациклических ограничений при моделировании стационарных метаболических потоков с помощью целочисленного линейного программирования // Научно-технический вестник информационных технологий, механики и оптики. 2026. Т. 26, № 5. С. 945–953. doi: 10.17586/2226-1494-2026-26-5-945-953

© Сафонов И.А., Сергушичев А.А., 2026

Metabolic potentials method for adding loopless constraints in steady-state metabolic flux modeling using mixed-integer linear programming

Ivan A. Safonov[1], Alexey A. Sergushichev[2]✉

[1] ITMO University, Saint Petersburg, 197101, Russian Federation
[2] Washington University School of Medicine in St. Louis, St. Louis, 63110, USA

[1] isafonov27@gmail.com, https://orcid.org/0009-0006-8483-7415
[2] asergushichev@wustl.edu✉, https://orcid.org/0000-0003-1159-7220

Abstract
In modeling of steady-state metabolic fluxes in metabolic networks, Flux Balance Analysis (FBA) is widely used. However, the standard FBA formulation can admit biochemically infeasible solutions due to thermodynamically infeasible (futile) cycles. Enforcing thermodynamic feasibility with loopless constraints yields an NP-hard problem, which motivates the development of efficient Mixed-Integer Linear Programming (MILP) formulations of loopless FBA. In this work, we propose a new efficient metabolic potentials method for FBA with loopless constraints. This method is based on a novel MILP formulation of loopless constraints. Our analysis of the prior Fast-SNP method, which achieved an order-of-magnitude speedup, allowed us to identify and quantify the contributions of two key optimization ideas: using sparse constraints and reducing the number of integer variables. The proposed metabolic potentials method adopts these ideas and introduces an equivalent MILP formulation of loopless constraints. We compare the performance of the metabolic potentials method using metabolic models from the BiGG database. The proposed method shows performance comparable to Fast-SNP on small and medium scale metabolic models and achieves an approximately 30 % reduction in runtime on a large genome-scale human cell metabolic model RECON1. We also demonstrate that reducing the number of integer variables is the dominant factor in the overall speedup. The proposed metabolic potentials method retains the benefits of the optimization ideas introduced in Fast-SNP and achieves the same or better efficiency. Therefore, it can be used as a drop-in replacement for Fast-SNP and other standard methods for FBA with loopless constraints.

Keywords
metabolic modeling, Flux Balance Analysis, steady-state flux solutions, loopless constraints, Mixed-Integer Linear Programming

For citation: Safonov I.A., Sergushichev A.A. Metabolic potentials method for adding loopless constraints in steady-state metabolic flux modeling using mixed-integer linear programming. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2026, vol. 26, no. 5, pp. 945–953 (in Russian). doi: 10.17586/2226-1494-2026-26-5-945-953

Введение

Метаболизм биохимических веществ в клетках организмов служит фундаментом для многих биологических процессов [1, 2]. Вычислительное моделирование является одним из важных инструментов изучения метаболизма и используется как в биотехнологических приложениях [3], так и при исследовании заболеваний человека [4]. Метод баланса потоков [5] (Flux Balance Analysis, FBA), рассматриваемый в настоящей работе, является одним из популярных методов такого вычислительного моделирования [6]. Этот метод базируется на законе сохранения масс и направлен на предсказание метаболических потоков в стационарном приближении — при отсутствии изменения скоростей реакций. Популярность метода обеспечивается его простотой, вычислительной эффективностью и возможностью использования как составного блока в решении более сложных задач метаболического моделирования.

Важное ограничение классического метода баланса потоков — отсутствие учета законов термодинамики, что делает допустимыми нефизичные потоки, содержащие циклы [7]. Эта проблема может быть решена через добавление в задачу дополнительных математических ограничений — ациклических ограничений [8]. Однако это делает задачу NP-трудной [9], что существенно увеличивает время работы и требует разработки эффективных подходов к ее решению. Так, в работе [10] предложен метод Fast-SNP, позволяющий ускорить решение задачи метода баланса потоков с ациклическими ограничениями на порядки, а в [9, 11] представлены альтернативные методы, дающие ускорение. Отметим, что для больших метаболических моделей, например, для клеток человека, время работы методов может быть очень большим. В результате актуальным является развитие вычислительного аппарата для более эффективного решения этой задачи.

В настоящей работе разработан новый метод метаболических потенциалов для добавления ациклических ограничений в методе баланса потоков. Установлено, какие оптимизации вносят основной вклад в эффективность метода Fast-SNP, и рассмотрено их применение в новом виде. Экспериментальное исследование продемонстрировало улучшение времени работы разработанного метода по сравнению с методом Fast-SNP на метаболической модели клеток человека RECON1.

Задача метода баланса потоков

Введем формальную постановку задач метаболического моделирования. В данных задачах метаболическая модель характеризуется множествами метаболитов (химических веществ) $\mathcal{M}$ и реакций $\mathcal{R}$. Реакции связаны с метаболитами через стехиометрическую матрицу $\mathbf{S} \in \mathbb{R}^{\mathcal{M}\times\mathcal{R}}$. Каждая строка матрицы соответствует метаболиту, каждый столбец — реакции. Отрицательные значения $S_{m,r} < 0$ соответствуют метаболитам m, которые потребляются реакцией r, а положительные значения $S_{m,r} > 0$ — метаболитам m, которые произво-

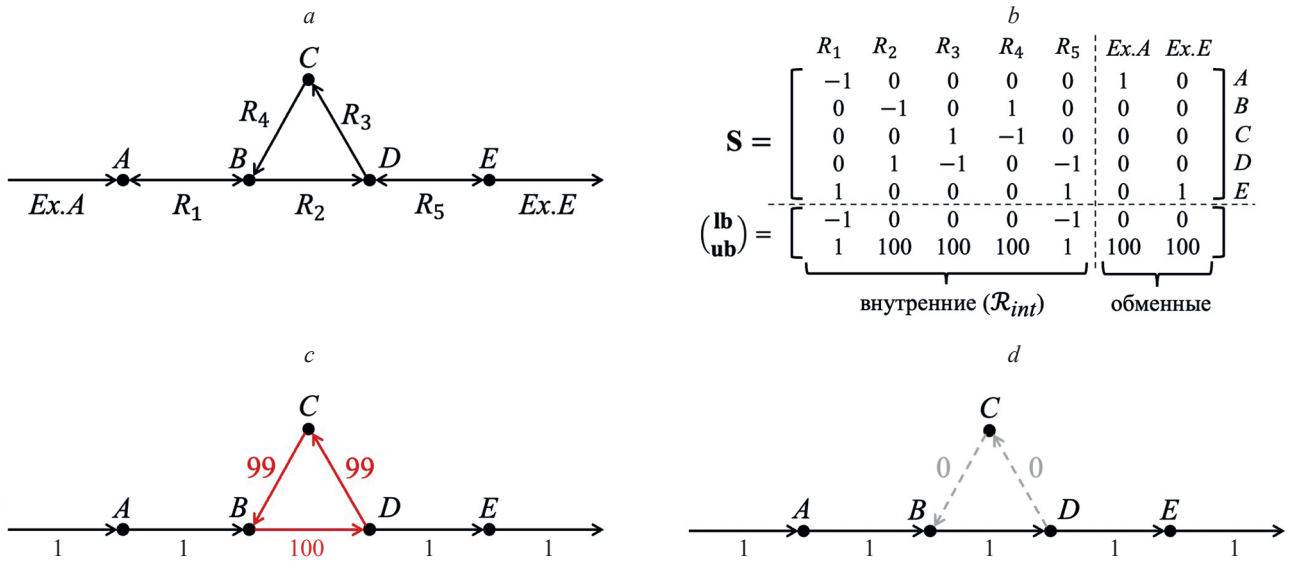


Рис. 1. Пример метаболической модели (5 метаболитов *A*, *B*, *C*, *D*, *E*; 7 реакций R_1, R_2, R_3, R_4, R_5, *Ex.A*, *Ex.E*): структура реакций и их возможные направления (*a*); стехиометрическая матрица **S**, верхние и нижние границы потоков **lb** и **ub**, множество внутренних реакций $\mathcal{R}_{int}$ (*b*); стационарный поток, максимизирующий поток через реакцию R_2 без (*c*) и с учетом (*d*) ациклических ограничений. Числа около реакций на (*c*), (*d*) показывают потоки через них. Красным цветом на (*c*) обозначены реакции, образующие цикл

Fig. 1. Metabolic model example (five metabolites *A*, *B*, *C*, *D* and *E*; seven reactions R_1, R_2, R_3, R_4, R_5, *Ex.A*, and *Ex.E*): structure of reactions and their possible directions (*a*); stoichiometric matrix **S**, flux bounds **lb**, **ub**, and set of internal reactions $\mathcal{R}_{int}$ (*b*); steady-state flux distribution maximizing the flux through reaction R_2, without loopless constraints (*c*) and with loopless constraints (*d*). Numbers near reactions in (*c*), (*d*) show fluxes through them. Reactions making a cycle are colored red in (*c*)

дятся реакцией *r*. Метаболиты *m*′, не участвующие в реакции *r*, имеют нулевые коэффициенты $S_{m',r} = 0$. На рис. 1, *a*, *b* изображен пример метаболической модели и ее стехиометрической матрицы.

В настоящей работе рассмотрена задача метода анализа баланса потоков. При этом объектом моделирования является вектор потоков $\mathbf{v} \in \mathbb{R}^{\mathcal{R}}$ или вектор мгновенных скоростей реакций. Основное предположение задачи FBA — *стационарность* — состояние, в котором концентрации метаболитов и скорости реакций не изменяются со временем. Условие стационарности может быть записано в виде $\mathbf{Sv} = \mathbf{0}$. Поток **v**, удовлетворяющий этому ограничению, называется стационарным.

Задача FBA состоит в поиске максимума целевой функции $\mathbf{c}^{\mathsf{T}}\mathbf{v}$ по всем стационарным потокам **v** при заданном векторе стоимостей реакций $\mathbf{c} \in \mathbb{R}^{\mathcal{R}}$. Дополнительно требуется соответствие потока **v** нижним и верхним границам $\mathbf{lb}, \mathbf{ub} \in \mathbb{R}^{\mathcal{R}}$ — параметрам метаболической модели. Таким образом, задача FBA может быть записана так:

$$\begin{aligned} &\mathbf{c}^{\mathsf{T}}\mathbf{v} \to \max, && \text{при условии:} \\ &\mathbf{Sv} = \mathbf{0}; && \\ &lb_i \le v_i \le ub_i, && \forall i \in \mathcal{R}. \end{aligned} \tag{1}$$

На рис. 1, *c* изображен пример стационарного потока, максимизирующего поток через реакцию R_2 (равносильно использованию вектора **c** в (1), в котором у реакции R_2 значение 1, а все остальные значения 0).

Оптимизационная задача (1) может быть напрямую решена с помощью линейного программирования (Linear Programming, LP), поскольку все условия и целевая функция являются линейными. С использованием современных решателей LP задача FBA эффективно решается даже для больших моделей, с десятками тысяч метаболитов и реакций.

Задача FBA — базовая задача для некоторых более специализированных задач метаболического моделирования. Например, задача вариации потока по реакциям [8] (Flux Variability Analysis, FVA) состоит в поиске минимальной и максимальной достижимых границ потока для каждой реакции по всем стационарным потокам. Она может быть решена с помощью многократного решения задачи FBA с целевыми функциями, равными оптимизируемому потоку через реакцию или минус оптимизируемому потоку через реакцию.

Ациклические ограничения

Основной проблемой задачи FBA (1) является то, что она допускает нефизичные решения, не удовлетворяющие законам термодинамики [7]. Такие решения могут содержать циклы — сбалансированные подпотоки, которые не потребляют и не производят метаболиты из внешней среды. Вдоль таких циклов можно пропускать сколь угодно большой поток, который не нарушает ограничение стационарности. В контексте задачи FBA это приводит к тому, что оптимум целевой функции может достигаться на таком нефизичном потоке (рис. 1, *c*). Таким образом, встает вопрос введения дополнительных ограничений, обеспечивающих соответствие законам термодинамики.

Определим циклы формально. Для этого разобьем множество реакций на обменные (с участием метабо-

литов из внешней среды) и внутренние (все остальные). Множество внутренних реакций обозначим $\mathcal{R}_{int}$. Циклом называется вектор потоков $\mathbf{q} \in \mathbb{R}^{\mathcal{R}}$, для которого выполнен баланс масс $\mathbf{Sq} = \mathbf{0}$, и все потоки через обменные реакции $r_e \in \mathcal{R}\backslash\mathcal{R}_{int}$ нулевые: $q_{r_e} = 0$. Цикл $\mathbf{q}$ принадлежит решению $\mathbf{v}$, если знаки всех ненулевых потоков в $\mathbf{q}$ совпадают со знаками соответствующих потоков в $\mathbf{v}$. Другими словами, из решения $\mathbf{v}$ можно вычесть цикл $\varepsilon\mathbf{q}$ для некоторого $\varepsilon > 0$, так что все знаки потоков в $\mathbf{v}$ и $\mathbf{v} - \varepsilon\mathbf{q}$ будут одинаковыми.

Одним из вариантов модификации задачи FBA для учета соответствия законам термодинамики является задача loopless Flux Balance Analysis (ll-FBA) [8]. В данной задаче кроме условий (1) вводятся дополнительные ограничения, называемые ациклическими и запрещающие все стационарные потоки, содержащие циклы. Ациклические ограничения также могут быть добавлены в другие задачи метаболического анализа. На рис. 1, *d* приведен пример стационарного потока, максимизирующего поток через реакцию R_2 с учетом ациклических ограничений — с их учетом максимальный возможный поток уменьшился со 100 до 1.

Один из основных подходов для добавления ациклических ограничений, который используется в популярных пакетах COBRA Toolbox [12] и COBRApy [13], был предложен в работе [8]. В нем ацикличность решения $\mathbf{v}$ гарантируется с помощью вспомогательного вектора $\mathbf{G} \in \mathbb{R}^{\mathcal{R}_{int}}$. Решение $\mathbf{v}$ не содержит циклов тогда и только тогда, когда существует вектор $\mathbf{G}$, такой что $G_i \neq 0$ для всех $i \in \mathcal{R}_{int}$, $\mathbf{N}_{int}\mathbf{G} = \mathbf{0}$ и $\text{sign}(v_i) \neq \text{sign}(G_i)$ для всех $v_i \neq 0$. Здесь $\mathbf{N}_{int}$ — базис линейного пространства циклов, который, например, может быть получен с помощью сингулярного разложения (Singular Value Decomposition, SVD) матрицы $\mathbf{S}_{int}$ — матрицы $\mathbf{S}$, суженной на подмножество внутренних реакций. Равносильность этого подхода ациклическим ограничениям доказана в работе [14]. Таким образом, задача ll-FBA может быть записана так:

$$
\begin{aligned}
&\mathbf{c}^{\mathsf{T}}\mathbf{v} \to \max, && \text{при условии:}\\
&\mathbf{Sv} = \mathbf{0}; &&\\
&lb_i \le v_i \le ub_i, && \forall i \in \mathcal{R};\\
&a_i \in \{0, 1\}, && \forall i \in \mathcal{R}_{int}; \quad (2)\\
&-M(1 - a_i) \le v_i \le Ma_i, && \forall i \in \mathcal{R}_{int};\\
&-Ma_i + 1(1 - a_i) \le G_i \le M(1 - a_i) - 1a_i, && \forall i \in \mathcal{R}_{int};\\
&\mathbf{N}_{int}\mathbf{G} = \mathbf{0}, &&
\end{aligned}
$$

где $M = 10^4$ — большая константа.

Оптимизационная задача (2) может быть напрямую решена с помощью целочисленного линейного программирования (Mixed Integer Linear Programming, MILP), поскольку все условия и целевая функция являются линейными, а также присутствует ограничение на целочисленность переменных a_i, которые соответствуют знакам $\text{sign}(v_i)$.

В отличие от задачи FBA задача ll-FBA является NP-трудной задачей [9]. Однако в некоторых случаях эта задача может быть решена с помощью специализированных MILP-решателей, используя (2). Время работы таких MILP-решателей зависит от многих факторов, таких как размер и структура ограничений MILP-задачи. Задачу ll-FBA можно представлять в виде MILP-задачи различными способами, будем называть такие способы формулировками. Соответственно, можно оптимизировать ациклические ограничения в формулировках за счет использования других, эквивалентных условий, чтобы добиться улучшения времени решения задачи ll-FBA.

Анализ метода Fast-SNP

Выполним анализ метода Fast-SNP, предложенного в работе [10]. Метод использует оптимизированную формулировку, позволившую достичь ускорения на один-два порядка в задаче ll-FBA относительно формулировки (2).

Во-первых, в [10] отмечено, что стандартный способ вычисления матрицы $\mathbf{N}_{int}$ — базиса пространства циклов — через SVD-разложение дает плотную матрицу, и предположено, что использование разреженной матрицы может уменьшить сложность формулировки ациклических ограничений для MILP-решателя. В результате был разработан эвристический алгоритм, позволяющий построить разреженную матрицу $\mathbf{N}_{int}$, которая может быть использована в (2) вместо базиса, построенного с помощью SVD. Будем называть такую формулировку, использующуюся в методе Fast-SNP, формулировкой Fast-SNP.

Отметим, что на эффективность формулировки Fast-SNP влияет не только разреженность получаемой матрицы $\mathbf{N}_{int}$, но и то, что при ее построении учитываются знаки ограничений lb_i и ub_i. Для этого при построении $\mathbf{N}_{int}$ рассматривается линейное пространство циклов, которое включает в себя только циклы с допустимыми знаками, размерность которого часто значительно ниже размерности пространства всех циклов. В [10] было предложено создавать целочисленные переменные a_i только для реакций, для которых соответствующий столбец $\mathbf{N}_{int}$ содержит хотя бы один ненулевой элемент. Действительно, если все элементы столбца нулевые, то в (2) ограничение $\mathbf{N}_{int}\mathbf{G} = \mathbf{0}$ никак не ограничивает соответствующий элемент вектора $\mathbf{G}$. За счет этого формулировка Fast-SNP позволяет значительно уменьшить число используемых целочисленных переменных, что важно для производительности MILP-решателей.

Использование метаболических потенциалов

На основе идеи про использование разреженной матрицы для построения ациклических ограничений в настоящей работе предложена альтернативная формулировка, в которой разреженность обеспечивается разреженностью стехиометрической матрицы $\mathbf{S}$.

Заметим, что условие $\mathbf{N}_{int}\mathbf{G} = \mathbf{0}$ из (2) равносильно тому, что существует вектор $\mathbf{p} \in \mathbb{R}^{\mathcal{M}}$ метаболических потенциалов такой, что $\mathbf{G} = \mathbf{S}_{int}^{\mathsf{T}}\mathbf{p}$. Действительно, из $\mathbf{N}_{int}\mathbf{G} = \mathbf{0}$ следует, что $\mathbf{G}$ лежит в пространстве, ортогональном $\mathbf{N}_{int}$ — пространстве $\mathbf{S}_{int}$ (линейное пространство строк соответствующей матрицы). В результате следует, что $\mathbf{G}$ представим в виде линейной комбинации элементов $\mathbf{S}_{int}$, т. е. соответствует $\mathbf{S}_{int}^{\mathsf{T}}\mathbf{p}$. Аналогично

обратное: если $\mathbf{G} = \mathbf{S}_{int}^{\top}\mathbf{p}$, то $\mathbf{G}$ лежит в пространстве $\mathbf{S}_{int}$, ортогональном $\mathbf{N}_{int}$, а значит $\mathbf{N}_{int}\mathbf{G} = \mathbf{0}$.

Таким образом, можно в ограничениях (2) вместо вектора $\mathbf{G}$ подставить вектор $\mathbf{S}_{int}^{\top}\mathbf{p}$ и убрать явное ограничение $\mathbf{N}_{int}\mathbf{G} = \mathbf{0}$, которое теперь выполняется автоматически. Подобная формулировка была впервые предложена в работе [7], но не предлагалась как альтернативная формулировка для решения задачи ll-FBA с использованием MILP-решателей. Следовательно, задача ll-FBA может быть записана в альтернативном виде:

$$
\begin{aligned}
&\mathbf{c}^{\top}\mathbf{v} \to \max, && \text{при условии:}\\
&\mathbf{Sv} = \mathbf{0}; &&\\
&lb_i \le v_i \le ub_i, && \forall i \in \mathcal{R}; \quad (3)\\
&a_i \in \{0, 1\}, && \forall i \in \mathcal{R}_{int};\\
&-M(1 - a_i) \le v_i \le Ma_i, && \forall i \in \mathcal{R}_{int};\\
&-Ma_i + 1(1 - a_i) \le \mathbf{p}^{\top}\mathbf{S}_{int,i} \le M(1 - a_i) - 1a_i, && \forall i \in \mathcal{R}_{int}.
\end{aligned}
$$

Отличие формулировки (3) от классической (2) заключается в том, что она использует вектор метаболических потенциалов $\mathbf{p}$, а также $\mathbf{S}_{int}$ вместо $\mathbf{N}_{int}$.

В метаболических моделях на практике стехиометрическая матрица является разреженной — обычно большинство реакций превращают несколько метаболитов в несколько других метаболитов. В результате автоматически получим формулировку для задачи ll-FBA, в которой участвует разреженная матрица, без запуска каких-либо предварительных шагов.

Оптимизация с помощью циклических реакций

В описанной выше формулировке (3) с помощью метаболических потенциалов была применена идея использования разреженной матрицы для построения ациклических ограничений. Однако в ограничениях используется полная матрица $\mathbf{S}_{int}$ и целочисленные переменные a_i создаются для всех внутренних реакций. Предложим способ сократить число целочисленных переменных, как в формулировке Fast-SNP.

Для начала введем понятие циклической реакции, которое поможет выразить множество реакций, участвующих в ациклических ограничениях, построенных в формулировке Fast-SNP. Реакция i называется *циклической*, если существует хотя бы один цикл $\mathbf{q}$, такой, что $q_i \neq 0$. При этом будем рассматривать только циклы, знаки потоков которых допускаются возможными направлениями реакций, которые получаются из соответствующих ограничений lb_i и ub_i. Например, в метаболической модели, изображенной на рис. 1, *а*, реакции R_2, R_3, R_4 являются циклическими, а R_1, R_5 не являются.

Заметим, что в ациклических ограничениях, построенных в формулировке Fast-SNP, целочисленные переменные a_i будут созданы в точности только для циклических реакций. Если для реакции i создана целочисленная переменная a_i, т. е. существует хотя бы один ненулевой элемент в соответствующем столбце матрицы $\mathbf{N}_{int}$, тогда строка с этим элементом будет являться допустимым циклом, содержащим реакцию i, следовательно, i будет циклической реакцией. Поскольку линейное пространство строк матрицы $\mathbf{N}_{int}$ содержит в себе все допустимые циклы, для всех циклических реакций будет существовать хотя бы одна строка матрицы $\mathbf{N}_{int}$ с ненулевым элементом, следовательно, для всех циклических реакций будут созданы целочисленные переменные a_i.

Заметим, что в формулировках (2) и (3) ациклические ограничения достаточно добавлять только для циклических реакций, а не для всех внутренних, потому что нециклические реакции не могут участвовать в циклах по определению. Найти циклические реакции можно с помощью алгоритма из метода Fast-SNP, но в настоящей работе выбран упрощенный метод поиска, основанный на определении циклической реакции.

Проверим по очереди все внутренние реакции на цикличность. Для этого для каждой реакции $r \in \mathcal{R}_{int}$ можно максимизировать и минимизировать q_r, при условии, что $\mathbf{q}$ является циклом. Если получится найти ненулевое значение целевой функции, то реакция является циклической, иначе нет. Таким образом, для каждой реакции r требуется решить две следующие задачи линейного программирования:

$$
\begin{aligned}
&q_r \to \max/\min, && \text{при условии:}\\
&\mathbf{Sq} = \mathbf{0}; &&\\
&q_i = 0, && \forall i \in \mathcal{R} \setminus \mathcal{R}_{int};\\
&q_i \le 0, && \forall ub_i \le 0,\ i \in \mathcal{R}_{int};\\
&q_i \ge 0, && \forall lb_i \ge 0,\ i \in \mathcal{R}_{int}.
\end{aligned}
$$

Используя найденное множество циклических реакций $\mathcal{R}_{cyc}$ совместно с метаболическими потенциалами (3), задача ll-FBA может быть записана в виде:

$$
\begin{aligned}
&\mathbf{c}^{\top}\mathbf{v} \to \max, && \text{при условии:}\\
&\mathbf{Sv} = \mathbf{0}; &&\\
&lb_i \le v_i \le ub_i, && \forall i \in \mathcal{R}; \quad (4)\\
&a_i \in \{0, 1\}, && \forall i \in \mathcal{R}_{cyc};\\
&-M(1 - a_i) \le v_i \le Ma_i, && \forall i \in \mathcal{R}_{cyc};\\
&-Ma_i + 1(1 - a_i) \le \mathbf{p}^{\top}\mathbf{S}_{cyc,i} \le M(1 - a_i) - 1a_i, && \forall i \in \mathcal{R}_{cyc}.
\end{aligned}
$$

В формулировке (4) вектор потенциалов $\mathbf{p}$ будет иметь размерность $|\mathcal{M}_{cyc}|$, где $\mathcal{M}_{cyc}$ — множество метаболитов, которые участвуют хотя бы в одной циклической реакции.

В настоящей работе предложена новая формулировка (4). Она, как и формулировка Fast-SNP, использует разреженную матрицу в ограничениях, а также аналогичное уменьшенное число целочисленных переменных. Новый метод решения задачи ll-FBA с помощью формулировки (4) назовем методом метаболических потенциалов.

Экспериментальное исследование

Выполним анализ времени работы MILP-решателя с использованием формулировки (4), сравнив ее с другими формулировками. Проведем анализ независимого вклада двух идей оптимизации в эффективность добавляемых ациклических ограничений: разреженных

ограничений и сокращения числа целочисленных переменных до числа циклических реакций.

В сравнении будут участвовать три группы формулировок задачи ll-FBA. Первая группа основана на базовой формулировке (2) [8], в которой матрица $\mathbf{N}_{int}$ строится с помощью SVD (группа SVD). Вторая группа основана на формулировке Fast-SNP [10], в которой матрица $\mathbf{N}_{int}$ строится с помощью алгоритма из метода Fast-SNP (группа Fast-SNP). Третья группа состоит из предложенных в настоящей работе формулировок (3) и (4), в которых ациклические ограничения строятся на основе стехиометрической матрицы (группа Potentials). В случае, если в MILP-задаче используется оптимизация числа целочисленных переменных до числа циклических реакций, к названию формулировки будем добавлять (Cyclic), иначе — (Full). В результате выполним сравнение 5 вариантов формулировок: SVD (Full) (2); Potentials (Full) (3); SVD (Cyclic) (2 с Cyclic-оптимизацией); Fast-SNP (Cyclic) [10] не рассматривается, потому что метод Fast-SNP всегда строит матрицу с ненулевыми столбцами только для циклических реакций. Проведена оценка, что специальное добавление целочисленных переменных a_i для всех реакций, в том числе для тех, у которых соответствующий столбец матрицы $\mathbf{N}_{int}$ нулевой, не влияет на время работы MILP-решателя.

Для сравнения формулировок применим задачу loopless Flux Variability Analysis (ll-FVA) [8] — ациклический вариант задачи FVA, в которой ациклические ограничения используются для определения минимальной и максимальной достижимых границ потока для каждой реакции. Эта задача позволяет протестировать формулировки задачи ll-FBA на нескольких целевых функциях. Также она часто применяется как первый шаг метаболического анализа модели для уточнения достижимых границ потоков.

Для оценки эффективности формулировки рассчитаем среднее время поиска ответа по всем циклическим реакциям метаболической модели. При анализе не будет учтено время поиска ответа для нециклических реакций, потому что ответ для них может быть найден без добавления ациклических ограничений гораздо эффективнее с помощью LP-решателей и формулировки (1). Также не будет учитываться время построения оптимизационной задачи и ациклических ограничений, так как для их построения используются только LP-решатели, и оно занимает небольшую часть общего времени решения задачи.

В результате для сравнения использованы модели из базы метаболических моделей BiGG [15] (табл. 1). Все эксперименты проведены на процессоре Apple M1 Pro (16 ГБ, 8 ядер) с применением LP- и MILP-решателей IBM ILOG CPLEX версии 22.1.1. Все алгоритмы и эксперименты были реализованы на языке Python с использованием библиотеки COBRApy [13] для работы с метаболическими моделями и библиотеки optlang [16] для взаимодействия с LP- и MILP-решателями.

Сначала проведем анализ разреженности ациклических ограничений, которые строятся в рассматриваемых формулировках. В табл. 2 приведено число ненулевых коэффициентов в матрицах $\mathbf{N}_{int}$ (для формулировок SVD и Fast-SNP) и $\mathbf{S}_{int}$ (для Potentials), на которых базируются ациклические ограничения. Видно, что число ненулевых коэффициентов для SVD значительно превосходит Fast-SNP и Potentials. Для SVD и Potentials использование циклических реакций (режим Cyclic) сильно сокращает размерность задачи, что позволяет сократить и число ненулевых коэффициентов. В Fast-SNP число ненулевых коэффициентов обычно в три-четыре раза меньше, чем в Potentials, однако в обоих случаях матрицы являются разреженными. В случае использования циклических реакций число ненулевых коэффициентов составляет примерно 80–100 % от матрицы для SVD и 1–5 % от матрицы для Fast-SNP и Potentials. Таким образом, показано, что предложенная в настоящей работе формулировка Potentials позволяет сделать разреженные ограничения и число ненулевых коэффициентов сравнимыми с формулировкой Fast-SNP.

Проанализируем вклад двух идей оптимизации формулировки в эффективность: использование разреженной матрицы вместо плотной матрицы и сокращение числа целочисленных переменных до числа циклических реакций. На рис. 2 показано сравнение эффективности на различных моделях и формулировках SVD и Potentials. Видно, что идея про сокращение числа целочисленных переменных до числа циклических реакций дает наибольший вклад в оптимизацию и по-

Таблица 1. Размеры метаболических моделей

Table 1. Metabolic model sizes

Модель	Число			
	реакций	метаболитов	внутренних реакций	циклических реакций
e_coli_core	95	72	75	2
iIT341	554	485	477	22
iAF692	690	628	620	30
iNJ661	1025	825	938	53
iAF1260	2382	1668	2078	68
iYL1228	2262	1658	1971	59
STM_v1_0	2545	1802	2200	52
iJO1366	2583	1805	2253	76
RECON1	3741	2766	3311	958

Таблица 2. Число ненулевых коэффициентов в используемых формулировками матрицах $\mathbf{N}_{int}$ (для SVD, Fast-SNP) и $\mathbf{S}_{int}$ (для Potentials)

Table 2. The number of nonzero coefficients in matrices used by MILP-formulations: $\mathbf{N}_{int}$ (for SVD, Fast-SNP) and $\mathbf{S}_{int}$ (for Potentials)

Модель	Режим	SVD	Fast-SNP	Potentials
e_coli_core	Full	650	—	340
	Cyclic	2	2	8
iIT341	Full	5434	—	2237
	Cyclic	132	24	98
iAF692	Full	12 560	—	2997
	Cyclic	255	38	136
iNJ661	Full	78 147	—	4706
	Cyclic	1094	64	241
iAF1260	Full	611 279	—	8927
	Cyclic	1510	84	262
iYL1228	Full	482 604	—	8474
	Cyclic	1188	77	227
STM_v1_0	Full	648 355	—	9461
	Cyclic	965	59	189
iJO1366	Full	717 116	—	9853
	Cyclic	2114	96	296
RECON1	Full	1 452 449	—	13 869
	Cyclic	438 596	4126	3946

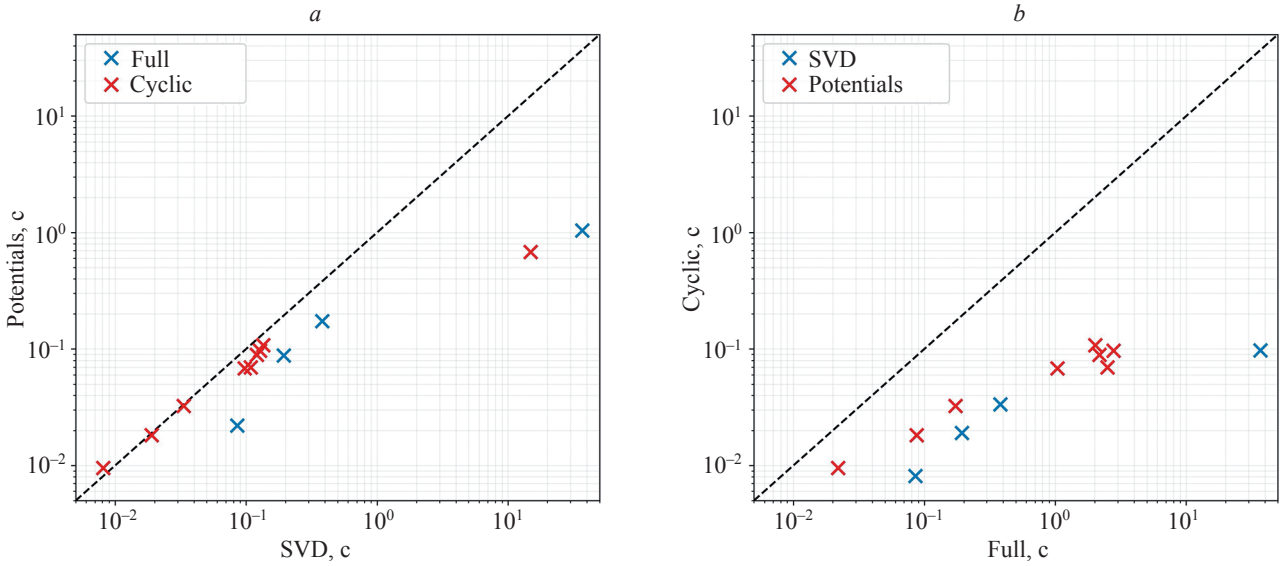


Рис. 2. Сравнение вклада двух идей оптимизации формулировки в эффективность.

Использование плотной матрицы (SVD) против разреженной матрицы (Potentials) для различных метаболических моделей и режимов (Full и Cyclic) (*a*). Режим Full против Cyclic для различных метаболических моделей и методов (SVD и Potentials) (*b*).

Координаты каждой точки — среднее по 10 независимым запускам время поиска ответа на одну циклическую реакцию, усредненное по всем циклическим реакциям метаболической модели

Fig. 2. Comparison of the contribution of two optimization ideas to the MILP-formulation efficiency.

Dense matrix (SVD) versus sparse matrix (Potentials) for different metabolic models and regimes (Full and Cyclic) (*a*). Full versus Cyclic for different metabolic models and methods (SVD and Potentials) (*b*).

The coordinates of each point are the mean time over ten independent runs to find the answer for a single cyclic reaction, averaged for all cyclic reactions of the metabolic model

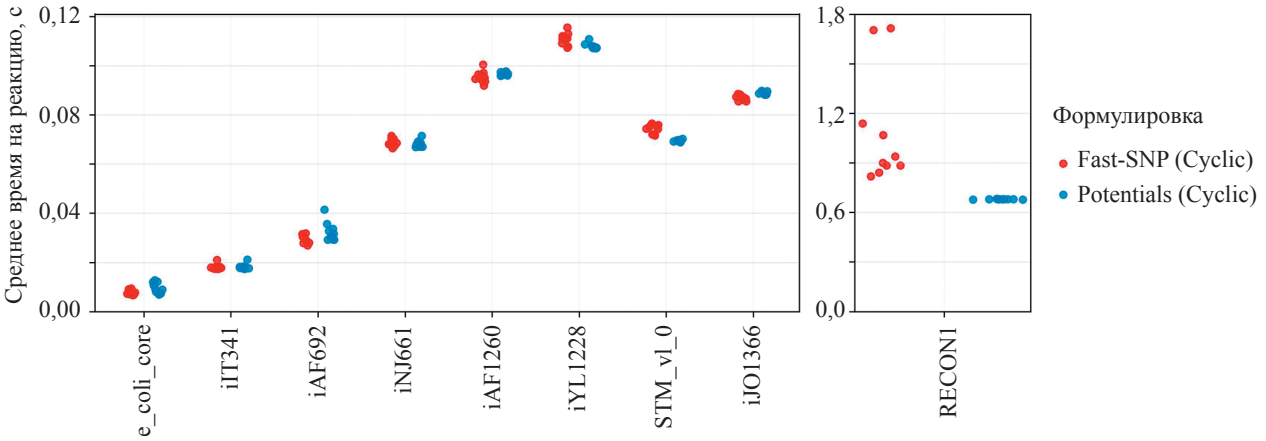


Рис. 3. Сравнение эффективности формулировок Fast-SNP (Cyclic) и Potentials (Cyclic). Для каждой формулировки и метаболической модели вычисления для всех циклических реакций выполнялись независимо 10 раз

Fig. 3. Performance comparison of Fast-SNP (Cyclic) and Potentials (Cyclic). For each MILP-formulation and metabolic model, computations for all cyclic reactions were run independently ten times

зволяет ускорить поиск ответа для реакции на один-два порядка. Идея про использование разреженной матрицы также дает улучшение, однако она обеспечивает ускорение в 2–10 раз. Таким образом, формулировка Potentials (Cyclic), которая использует обе идеи, имеет наибольшее преимущество в эффективности за счет сокращения числа циклических реакций.

Рассмотрим эффективность решения ll-FVA с использованием двух наиболее эффективных формулировок — Fast-SNP (Cyclic) и Potentials (Cyclic), предложенных в настоящей работе. На рис. 3 показано среднее время поиска ответа для реакции при использовании этих двух формулировок. Заметно, что предложенная в настоящей работе формулировка Potentials (Cyclic) показывает такую же эффективность, как и Fast-SNP (Cyclic) на метаболических моделях малого и среднего размера, а для большой метаболической модели клеток человека RECON1 работает в среднем на 30 % эффективнее. Также виден больший разброс времени работы при использовании формулировки Fast-SNP (Cyclic) по сравнению с Potentials (Cyclic), для которой разброс незначительный.

Заключение

Представлен новый метод метаболических потенциалов для задачи ll-FBA, основанный на метаболических потенциалах. Разработанная формулировка (4), как и формулировка Fast-SNP, использует две оптимизации ациклических ограничений — использование разреженной матрицы и сокращение числа целочисленных переменных. Отметим, что предложенная формулировка не строит матрицу базиса пространства циклов $\mathbf{N}_{int}$, а напрямую использует матрицу $\mathbf{S}_{cyc}$ — стехиометрическую матрицу, суженную на множество циклических реакций модели. Результаты экспериментов показали, что новый метод метаболических потенциалов на различных метаболических моделях имеет такую же или лучшую эффективность по сравнению с методом Fast-SNP. Например, на самой большой рассмотренной метаболической модели клеток человека RECON1 метод ищет ответ в среднем на 30 % эффективнее. Показано, что обе оптимизации формулировки являются полезными, но сокращение числа целочисленных переменных имеет наибольший вклад. Таким образом, представленные в настоящей работе результаты имеют как теоретическую, так и практическую значимость. С точки зрения теории был предложен новый метод метаболических потенциалов, который строит новую альтернативную формулировку ациклических ограничений, а также был проведен анализ независимого вклада идей оптимизации формулировок ациклических ограничений. С точки зрения практики было показано, что метод метаболических потенциалов может применяться в качестве основного метода для метода баланса потоков с ациклическими ограничениями.

Литература

1. Judge A., Dodd M.S. Metabolism // Essays in Biochemistry. 2020. V. 64. N 4. P. 607–647. doi: 10.1042/EBC20190041
2. Chandel N.S. Navigating Metabolism. Cold Spring Harbor Laboratory Press, 2014. 264 p.
3. Copeland W.B., Bartley B.A., Chandran D., Galdzicki M., Kim K.H., Sleight S.C., et al. Computational tools for metabolic engineering // Metabolic Engineering. 2012. V. 14. N 3. P. 270–280. doi: 10.1016/j.ymben.2012.03.001
4. Mardinoglu A., Nielsen J. Systems medicine and metabolic modelling // Journal of Internal Medicine. 2012. V. 271. N 2. P. 142–154. doi: 10.1111/j.1365-2796.2011.02493.x

References

1. Judge A., Dodd M.S. Metabolism. *Essays in Biochemistry*, 2020, vol. 64, no. 4, pp. 607–647. doi: 10.1042/EBC20190041
2. Chandel N.S. *Navigating Metabolism*. Cold Spring Harbor Laboratory Press, 2014, 264 p.
3. Copeland W.B., Bartley B.A., Chandran D., Galdzicki M., Kim K.H., Sleight S.C., et al. Computational tools for metabolic engineering. *Metabolic Engineering*, 2012, vol. 14, no. 3, pp. 270–280. doi: 10.1016/j.ymben.2012.03.001
4. Mardinoglu A., Nielsen J. Systems medicine and metabolic modelling. *Journal of Internal Medicine*, 2012, vol. 271, no. 2, pp. 142–154. doi: 10.1111/j.1365-2796.2011.02493.x

5. Orth J.D., Thiele I., Palsson B.Ø. What is flux balance analysis? // Nature Biotechnology. 2010. V. 28. N 3. P. 245–248. doi: 10.1038/nbt.1614
6. Gianchandani E.P., Chavali A.K., Papin J.A. The application of flux balance analysis in systems biology // WIREs Systems Biology and Medicine. 2010. V. 2. N 3. P. 372–382. doi: 10.1002/wsbm.60
7. Beard D.A., Babson E., Curtis E., Qian H. Thermodynamic constraints for biochemical networks // Journal of Theoretical Biology. 2004. V. 228. N 3. P. 327–333. doi: 10.1016/j.jtbi.2004.01.008
8. Schellenberger J., Lewis N.E., Palsson B.Ø. Elimination of thermodynamically infeasible loops in steady-state metabolic models. // Biophysical Journal. 2011. V. 100. N 3. P. 544–553. doi: 10.1016/j.bpj.2010.12.3707
9. Müller A.C., Bockmayr A. Fast thermodynamically constrained flux variability analysis // Bioinformatics. 2013. V. 29. N 7. P. 903–909. doi: 10.1093/bioinformatics/btt059
10. Saa P.A., Nielsen L.K. Fast-SNP: a fast matrix pre-processing algorithm for efficient loopless flux optimization of metabolic models // Bioinformatics. 2016. V. 32. N 24. P. 3807–3814. doi: 10.1093/bioinformatics/btw555
11. Chan S.H.J., Wang L., Dash S., Maranas C.D. Accelerating flux balance calculations in genome-scale metabolic models by localizing the application of loopless constraints // Bioinformatics. 2018. V. 34. N 24. P. 4248–4255. doi: 10.1093/bioinformatics/bty446
12. Heirendt L., Arreckx S., Pfau T., Mendoza S.N., Richelle A., Heinken A., et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0 // Nature Protocols. 2019. V. 14. N 3. P. 639–702. doi: 10.1038/s41596-018-0098-2
13. Ebrahim A., Lerman J.A., Palsson B.O., Hyduke D.R. COBRApy: COnstraints-Based Reconstruction and Analysis for Python // BMC Systems Biology. 2013. V. 7. P. 74. doi: 10.1186/1752-0509-7-74
14. Noor E., Lewis N.E., Milo R. A proof for loop-law constraints in stoichiometric metabolic networks // BMC Systems Biology. 2012. V. 6. P. 140. doi: 10.1186/1752-0509-6-140
15. Schellenberger J., Park J.O., Conrad T.M., Palsson B.O. BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions // BMC Bioinformatics. 2010. V. 11. P. 213. doi: 10.1186/1471-2105-11-213
16. Jensen K., Cardoso J.G.R., Sonnenschein N. Optlang: An algebraic modeling language for mathematical optimization // The Journal of Open Source Software. 2017. V. 2. N 9. P. 139. doi: 10.21105/joss.00139

5. Orth J.D., Thiele I., Palsson B.Ø. What is flux balance analysis? *Nature Biotechnology*, 2010, vol. 28, no. 3, pp. 245–248. doi: 10.1038/nbt.1614
6. Gianchandani E.P., Chavali A.K., Papin J.A. The application of flux balance analysis in systems biology. *WIREs Systems Biology and Medicine*, 2010, vol. 2, no. 3, pp. 372–382. doi: 10.1002/wsbm.60
7. Beard D.A., Babson E., Curtis E., Qian H. Thermodynamic constraints for biochemical networks. *Journal of Theoretical Biology*, 2004, vol. 228, no. 3, pp. 327–333. doi: 10.1016/j.jtbi.2004.01.008
8. Schellenberger J., Lewis N.E., Palsson B.Ø. Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophysical Journal*, 2011, vol. 100, no. 3, pp. 544–553. doi: 10.1016/j.bpj.2010.12.3707
9. Müller A.C., Bockmayr A. Fast thermodynamically constrained flux variability analysis. *Bioinformatics*, 2013, vol. 29, no. 7, pp. 903–909. doi: 10.1093/bioinformatics/btt059
10. Saa P.A., Nielsen L.K. Fast-SNP: a fast matrix pre-processing algorithm for efficient loopless flux optimization of metabolic models. *Bioinformatics*, 2016, vol. 32, no. 24, pp. 3807–3814. doi: 10.1093/bioinformatics/btw555
11. Chan S.H.J., Wang L., Dash S., Maranas C.D. Accelerating flux balance calculations in genome-scale metabolic models by localizing the application of loopless constraints. *Bioinformatics*, 2018, vol. 34, no. 24, pp. 4248–4255. doi: 10.1093/bioinformatics/bty446
12. Heirendt L., Arreckx S., Pfau T., Mendoza S.N., Richelle A., Heinken A., et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nature Protocols*, 2019, vol. 14, no. 3, pp. 639–702. doi: 10.1038/s41596-018-0098-2
13. Ebrahim A., Lerman J.A., Palsson B.O., Hyduke D.R. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, 2013, vol. 7, pp. 74. doi: 10.1186/1752-0509-7-74
14. Noor E., Lewis N.E., Milo R. A proof for loop-law constraints in stoichiometric metabolic networks. *BMC Systems Biology*, 2012, vol. 6, pp. 140. doi: 10.1186/1752-0509-6-140
15. Schellenberger J., Park J.O., Conrad T.M., Palsson B.O. BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinformatics*, 2010, vol. 11, pp. 213. doi: 10.1186/1471-2105-11-213
16. Jensen K., Cardoso J.G.R., Sonnenschein N. Optlang: An algebraic modeling language for mathematical optimization. *The Journal of Open Source Software*, 2017, vol. 2, no. 9, pp. 139. doi: 10.21105/joss.00139

Авторы

Сафонов Иван Андреевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0006-8483-7415, isafonov27@gmail.com
Сергушичев Алексей Александрович — кандидат технических наук, доцент, Университет Вашингтона в Сент-Луисе, Сент-Луис, 63110, США, SC 55772694000, https://orcid.org/0000-0003-1159-7220, asergushichev@wustl.edu

Authors

Ivan A. Safonov — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0006-8483-7415, isafonov27@gmail.com
Alexey A. Sergushichev — PhD, Assistant Professor, Washington University School of Medicine in St. Louis, St. Louis, 63110, USA, SC 55772694000, https://orcid.org/0000-0003-1159-7220, asergushichev@wustl.edu

Статья поступила в редакцию 20.01.2026
Одобрена после рецензирования 20.08.2026
Принята к печати 21.09.2026

Received 20.01.2026
Approved after reviewing 20.08.2026
Accepted 21.09.2026

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»