

ПРАКТИЧЕСКИЕ АСПЕКТЫ ВЫБОРА МЕТОДА МОДЕЛИРОВАНИЯ ТРЕТИЧНОЙ СТРУКТУРЫ ИЗОФОРМЫ P00749-2 УРОКИНАЗЫ

Д.С. Байроченко

Студент лечебного факультета

учреждения образования

«Белорусский государственный медицинский университет», г. Минск,

Беларусь

bairon0404@gmail.com;

***Аннотация.** Целью работы явилось сравнение качества третичных структур изоформы P00749-2 урокиназы (фермента, участвующего в процессах роста опухолей и распространении раковых клеток), полученных путем моделирования *in silico*, с использованием программ MODELLER, Phyre2, Robetta и SWISS-MODEL. После проведения сравнительного анализа на предмет качества полученных моделей сервис SWISS-MODEL был рекомендован как наиболее оптимальный для моделирования изоформ урокиназы.*

***Ключевые слова:** гомологичное моделирование; урокиназа; SWISS-MODEL; Robetta; Phyre2; MODELLER*

PRACTICAL ASPECTS OF CHOOSING A METHOD FOR MODELING THE TERTIARY STRUCTURE OF UROKINASE ISOFORM P00749-2

D.S. Bairachenka

Student of the Medical Faculty

of the Educational Institution

"Belarusian State Medical University", Minsk, Belarus

bairon0404@gmail.com;

***Annotation.** The aim of the study was to compare the quality of tertiary structures of the urokinase isoform P00749-2 (urokinase is an enzyme involved in the processes of tumor growth and the spread of cancer cells), which were obtained by modeling *in silico* in MODELLER, Phyre2, Robetta and SWISS-MODEL. After carrying out a comparative analysis for the quality of the received models, the SWISS-MODEL service was recommended as the most optimal for the modeling of urokinase isoforms.*

***Keywords:** homology modelling; urokinase; SWISS-MODEL; Robetta; Phyre2; MODELLER*

В настоящее время классические методы установления третичной структуры молекул в связи с большими затратами времени на их проведение не отвечают набранному темпу биомолекулярных исследований. Вследствие чего всё более важную роль в науке играют исследования *in silico*. Использование данных методов позволяет делать прогнозы и выдвигать теории с достаточно высокой степенью достоверности; экономить человеческие и денежные ресурсы при проведении исследований.

Методы предсказания третичной структуры белковых молекул могут использоваться не только для моделирования структуры отдельных белков, но и для конструирования их изоформ, что может быть полезно при разработке высокоаффинных ингибиторов для белков-мишеней [1], а также в контексте комплексного изучения *in silico* эффективности действия потенциальных лекарственных средств.

Урокиназный активатор плазминогена (урокиназа) – это внеклеточная сериновая протеаза, состоящая из 411 аминокислотных остатков, кодируемая геном PLAU (10q22.2). Данный фермент активирует плазминоген, способен подвергать гидролизу некоторые компоненты внеклеточного матрикса. Он также активирует представителей семейства матриксных металлопротеиназ, ответственных за ремоделирование и деградацию внеклеточного матрикса, регуляцию пролиферации, миграции и дифференциации клеток [2]. В опухолевых клетках экспрессия гена PLAU может усиливаться в несколько раз [3]. Это ведет к повышению активности урокиназы и усилению процессов разрушения внеклеточного матрикса, что, в свою очередь, может способствовать прогрессированию опухолей и метастазированию [4]. Высокий уровень урокиназы коррелирует с неблагоприятным прогнозом и высокой частотой рецидивов [5].

Урокиназа является потенциальной мишенью для проведения антиопухолевой терапии. При разработке *in silico* ингибиторов урокиназы необходимо провести анализ эффективности действия предложенных лигандов в отношении изоформ фермента. Решением подобной задачи служит предсказание структур изоформ урокиназы, необходимых для последующего анализа эффективности связывания ингибиторов, например, Нафамостата и ряда его производных.

Целью данной работы явилось сравнение качества третичных структур изоформы P00749-2 урокиназы, полученных путем моделирования *in silico*, с использованием программ MODELLER, Phyre2, Robetta и SWISS-MODEL.

Материалы и методы. Поиск аминокислотных последовательностей изоформ урокиназы осуществлялся в ресурсе UniProt [6], содержащем данные о последовательностях белков и аннотации к ним. Изоформа P00749-2 образуется

в результате альтернативного сплайсинга, имеет длину 414 аминокислотных остатка и массу около 46,9 кДа. Было произведено построение ряда моделей в следующих высокоспециализированных программах: SWISS-MODEL, Robetta, Phyre2, MODELLER. В связи с тем, что ряд программ конструируют несколько моделей на базе одной аминокислотной последовательности, была проведена первичная оценка качества трехмерных структур, предложенных каждой из программ, с целью выбора наиболее успешных образцов (по одной от каждого сервиса) для последующего сравнительно анализа. Первичная оценка производилась посредством предлагаемых каждой программой инструментов и параметров.

Для сравнительного анализа качества 3D-структур, полученных с помощью вышеуказанных программ, использовались сервисы:

1. Инструмент Structure Assessment, SWISS-MODEL. Рассчитывает показатель QMEANDisCo, получаемый путем оценки попарных расстояний остаток-остаток с учетом пространственных ограничений. Определяет значения C β и All Atom (потенциалы взаимодействия на основе атомов C β и всех атомов), Solvation (эффект сольватации), Torsion (торсионный угол для 3-х последовательных аминокислот), QMEAN (описывает геометрию структур, используя вышеперечисленные параметры). Строит карту Рамачандрана с разрешением 2,5 Å.

2. UCLA-DOE LAB – SAVES v6.0. Программа использует следующие инструменты: ERRAT анализирует белковые структуры на предмет наличия несвязанных остатков, рассчитывая общий коэффициент качества; VERIFY3D сопоставляет трехмерную структуру с аминокислотной последовательностью; PROCHECK строит карту Рамачандрана с разрешением 2,0 Å.

Результаты. При построении карты Рамачандрана с использованием инструмента Structure Assessment лучшие результаты (наибольшее количество остатков в разрешенных регионах и наименьшее в запрещенных) соответствовали модели, сконструированной в SWISS-MODEL. По значениям QMEANDisCo и QMEAN данная модель также занимает лидирующую позицию, уступая лишь по показателям C β , Solvation, Torsion модели Robetta. Полученные данные представлены в таблице 1.

Таблица 1 – Инструмент Structure Assessment, количественные характеристики лучших моделей

Показатели	SWISS-MODEL	Robetta	Phyre2	MODELLER
Рамачандран основной регион, %	95,85	92,96	89,32	87,38

Рамачандран запрещенный регион, %	0,49	1,46	4,13	5,10
Рамачандран другие регионы, %	3,66	5,58	6,55	7,52
QMEANDisCo	0,77	0,70	0,64	0,48
QMEAN	-0,02	-0,49	-4,94	-7,00
C β	0,69	-0,38	-2,94	-5,68
All Atom	-1,87	-2,73	-5,72	-6,89
Solvation	-1,31	-0,99	-3,83	-7,96
Torsion	0,33	-0,04	-3,37	-3,77

На основании результатов Structure Assessment лучшей можно считать модель, полученную с помощью SWISS-MODEL. Для этой модели был построен график (рисунок 1) оценки локального качества каждого аминокислотного остатка. Отклонение для каждого остатка от глобального значения QMEANDisCo в большую (синий) или меньшую (красный) сторону продемонстрированы вертикальными линиями. Принято считать, что остаткам со значением ниже 0,6 присуще низкое качество структуры.

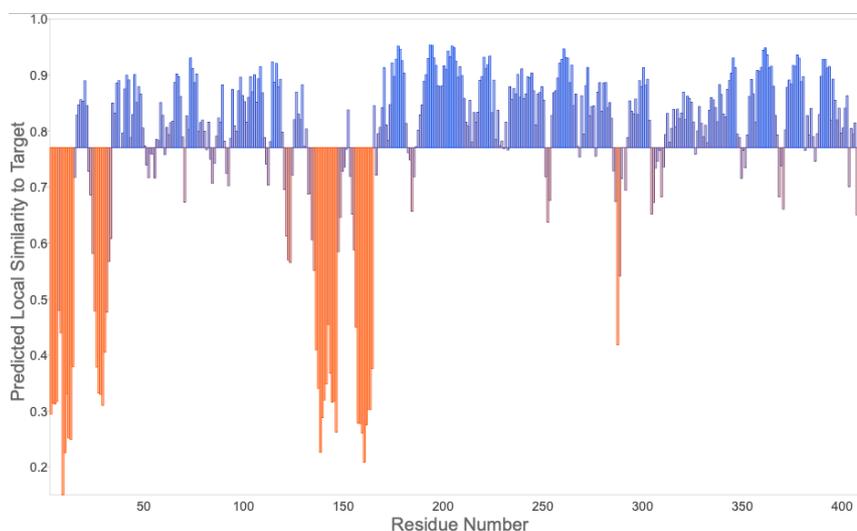


Рисунок 1 – Оценка локального качества модели SWISS-MODEL

На рисунке 2 представлен график, отражающий качество структуры SWISS-MODEL. Горизонтальная ось отражает длину белка. Вертикальная ось – значение QMEAN. Каждая точка – одна экспериментальная пространственная белковая структура, полученная из различных баз данных. Модель SWISS-MODEL обозначена красной звездочкой. Анализируемая модель относится к области с оценкой "QMEAN" в пределах одного стандартного отклонения ($|Z$ -

score| меньше 1) от среднего значения, что говорит о достаточно высоком качестве ее структуры.

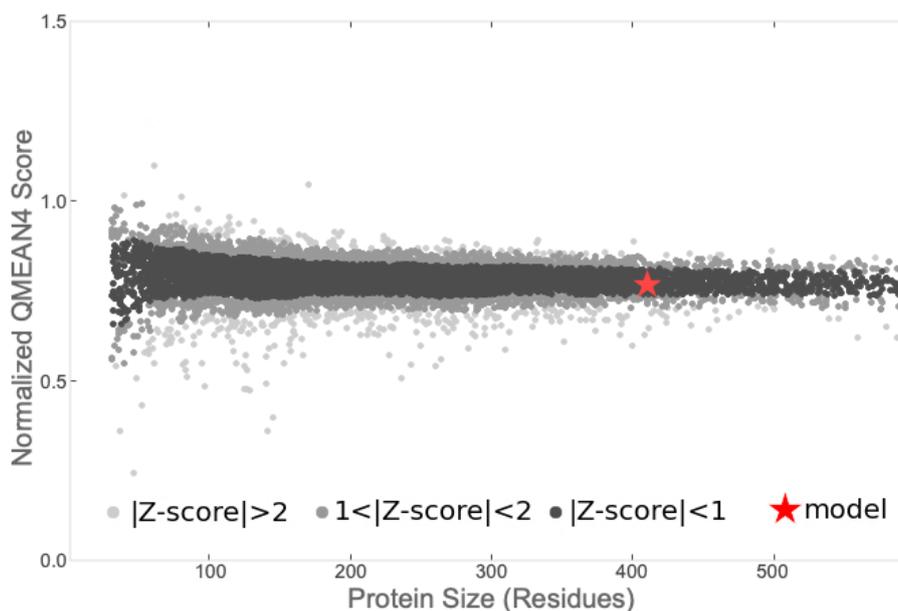


Рисунок 2 – Оценка глобального качества модели SWISS-MODEL

Далее был использован сервис SAVES v6.0 для уточнения значений в картах Рамачандрана и получения иных структурных данных для каждой из четырех моделей (таблица 2).

Таблица 2 – Инструмент SAVES v6.0, характеристики лучших моделей

Показатели	SWISS-MODEL	Robetta	Phyre2	MODELLER
ERRAT	89,26	90,5	42,47	38,48
VERIFY3D, %	70,63	73,43	76,09	59,66
Рамачандран основной, %	89,3	82,6	78,9	79,8
Рамачандран разрешенный, %	10,5	15,2	16,6	16,6
Рамачандран в целом разрешенный, %	0,3	1,7	2,8	2,2
Рамачандран запрещенный, %	0,0	0,6	1,7	1,4

По всем вышеперечисленным параметрам, кроме ERRAT и VERIFY3D, образец из SWISS-MODEL обладает лучшими значениями. Более того, данная

модель не имеет остатков, расположенных в запрещенном регионе карты Рамачандрана, что говорит в пользу высокой степени качества структуры.

Заключение. На основании проанализированных показателей, полученных с использованием Structure Assessment и SAVES v6.0, лучшими результатами обладает образец SWISS-MODEL.

Таким образом, для получения третичной структуры изоформы P00749-2 урокиназы целесообразно использовать сервис SWISS-MODEL.

Список литературы

1. Schmidt T., Bergner A., Schwede T. Modelling three-dimensional protein structures for applications in drug design. *Drug Discov Today*. 2014;19(7):890-897. doi:10.1016/j.drudis.2013.10.027.
2. Wang X., Khalil R. Matrix Metalloproteinases, Vascular Remodeling, and Vascular Disease. *Adv Pharmacol*. 2018;81:241-330. doi:10.1016/bs.apha.2017.08.002.
3. Mahmood N, Mihalcioiu C, Rabbani SA. Multifaceted Role of the Urokinase-Type Plasminogen Activator (uPA) and Its Receptor (uPAR): Diagnostic, Prognostic, and Therapeutic Applications. *Front Oncol*. 2018;8:24. Published 2018 Feb 12. doi:10.3389/fonc.2018.00024.
4. Madunić J. The Urokinase Plasminogen Activator System in Human Cancers: An Overview of Its Prognostic and Predictive Role. *Thromb Haemost*. 2018;118(12):2020-2036. doi:10.1055/s-0038-1675399.
5. Duffy MJ, McGowan PM, Harbeck N, Thomssen C, Schmitt M. uPA and PAI-1 as biomarkers in breast cancer: validated for clinical use in level-of-evidence-1 studies. *Breast Cancer Res*. 2014;16(4):428. Published 2014 Aug 22. doi:10.1186/s13058-014-0428-4.
6. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res*. 2023;51(D1):D523-D531. doi:10.1093/nar/gkac1052.