

СРАВНЕНИЕ МЕТОДОВ МОДЕЛИРОВАНИЯ ТРЕТИЧНОЙ СТРУКТУРЫ ИЗОФОРМЫ УРОКИНАЗЫ

Байроченко Д.С. (2 курс, лечебный факультет), Ринейская О.Н. (к.м.н., доцент, кафедра биоорганической химии)

Белорусский государственный медицинский университет, г. Минск

Аннотация. В настоящей работе проведено сравнение качества 3D-моделей изоформы урокиназы P00749-2. Поиск первичной последовательности изоформы осуществлялся в ресурсе UniProt. Моделирование третичных структур было проведено с использованием программ MODELLER, Phyre2, Robetta и SWISS-MODEL. Для оценки качества 3D-структур были использованы сервисы SAVES v6.0 и Structure Assessment. В результате выполненной работы было установлено, что сервис SWISS-MODEL является оптимальным для моделирования изоформы P00749-2 урокиназы.

Ключевые слова: гомологичное моделирование, SWISS-MODEL, Robetta, Phyre2, MODELLER, урокиназа.

Введение. Установление трехмерных структур белков крайне важно в понимании патогенеза ряда заболеваний, делает возможным проведение таргетной терапии и разработку высокоаффинных лекарственных препаратов. Сегодня важная роль в данном направлении отводится методам *in silico*, получившим широкое распространение в современной научной биологической и медицинской практике.

Цель исследования. Сравнение качества третичных структур белковых молекул, полученных путем моделирования *in silico*, с использованием ряда онлайн-сервисов и программ в контексте определения наиболее подходящего сервиса для установления третичной структуры изоформы P00749-2 урокиназы.

Материал и методы. Данные о первичной структуре изоформ урокиназы были взяты из ресурса UniProt [2]. Были найдены два основных варианта фермента, обозначаемые как P00749-1 и P00749-2. Вариант P00749-1 имеет длину 431 аминокислотных остатка и массу около 48,5 кДа, является каноничным, для него ранее была предсказана *in silico* трехмерная структура, которую можно найти в базе данных AlphaFold. Вариант P00749-2 образуется в результате альтернативного сплайсинга, имеет длину 414 аминокислотных остатка и массу около 46,9 кДа. Именно для изоформы P00749-2 проводилось моделирование 3D-структуры.

Далее в каждой из четырех вышеназванных программ было проведено *in silico* моделирование данной изоформы белка. Программы SWISS-MODEL [3], Robetta [1] и MODELLER [4] конструируют несколько моделей на базе одной первичной последовательности. В связи с чем была проведена первичная оценка глобального качества 3D-структур с целью отбора по одному наиболее успешному образцу от каждой программы для последующего сравнительно анализа. Первичная оценка производилась посредством инструментов, предлагаемых каждой из программ.

С целью выявления наиболее подходящей программы для трехмерного моделирования изоформы урокиназы были использованы следующие сервисы:

1. SAVES v6.0. Программа использует следующие инструменты: ERRAT – анализирует белковые структуры на предмет наличия несвязанных остатков в пределах 3,5 Å между различными атомами, рассчитывая общий коэффициент качества. Для высококачественной модели значение должно превышать 50; VERIFY3D – задействует структурную базу данных и сопоставляет трехмерную структуру с аминокислотной последовательностью, проводя сравнение качества предоставленной структуры с белковыми структурами с высоким разрешением из базы данных); PROCHECK – анализирует общую геометрию и оценивает стереохимическое качество прогнозируемой модели, строит карту Рамачандрана с разрешением 2,0 Å.

2. Инструмент Structure Assessment. Рассчитывает показатель QMEANDisCo, который позволяет дать абсолютную оценку качества модели. Этот параметр определяет локальное качество для каждого остатка путем оценки соответствия попарных расстояний остаток-остаток с учетом пространственных ограничений. Программа определяет потенциалы взаимодействия, зависящие от расстояния, на основе атомов C β и всех атомов, соответственно; значения эффекта сольватации; угла торсии для трех последовательных аминокислот (оценивается качество вторичных структур); QMEAN (описывает основные геометрические аспекты белковых структур, используя вышеперечисленные параметры). Названные показатели формируются на основании сравнения анализируемой модели и экспериментально определенных структур аналогичного размера; структуры получившие значения близкие к нулю считаются нативными; значения меньше «-4» обычно считаются неудовлетворительными. Строит карту Рамачандрана с разрешением 2,5 Å.

Результаты исследования. При построении карты Рамачандрана с использованием инструмента Structure Assessment лучшие результаты соответствовали модели SWISS-MODEL. По значениям QMEANDisCo и QMEAN данная модель также занимает лидирующую позицию. Полученные данные представлены в ниже (табл. 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
Потенциал взаимодействия (C β атомы)	0,69	-0,38	-2,94	-5,68
Потенциал взаимодействия (все атомы)	-1,87	-2,73	-5,72	-6,89
Эффект сольватации	-1,31	-0,99	-3,83	-7,96
Угол торсии	0,33	-0,04	-3,37	-3,77
QMEAN	-0,02	-0,49	-4,94	-7,00
QMEANDisCo	0,77	0,70	0,64	0,48

Далее с целью получения дополнительных данных о качестве структуры сконструированных белковых молекул, а также уточнения значений в картах Рамачандрана для каждой из четырех моделей был использован сервис SAVES v6.0, результаты представлены в сводной таблице (табл. 2).

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

Показатели	SWISS-MODEL	Robetta	Phyre2	MODELLER
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
ERRAT	89,26	90,5	42,47	38,48
VERIFY3D, %	70,63	73,43	76,09	59,66

Образец SWISS-MODEL обладает лучшими показателями по всем параметрам карты Рамачандрана, за исключением ERRAT и VERIFY3D. Также данная модель не имеет остатков в запрещенном регионе карты, что является весомым аргументом в пользу высокого качества структуры. Ниже представлено изображение трехмерной структуры модели SWISS-MODEL (рис. 1).

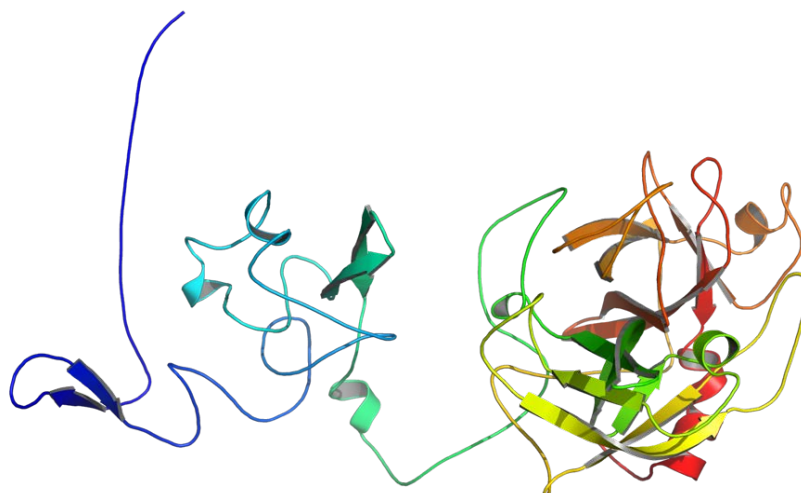


Рисунок 3 – Модель, полученная с помощью SWISS-MODEL

Заключение. Был построен ряд трехмерных моделей с использованием четырех сервисов. На основании 15-ти проанализированных показателей,

полученных с использованием Structure Assessment и SAVES v6.0, наилучшими результатами обладает модель SWISS-MODEL, лидируя по 10 параметрам.

Для конструирования третичной структуры изоформы P00749-2 урокиназы было рекомендовано использовать сервис SWISS-MODEL.

Список литературы:

1. Baek, M. Accurate prediction of protein structures and interactions using a three-track neural network / M. Baek [et al] // Science. – 2021. – Vol. 373, № 6557. – P. 871–876. – Mode of access: <https://www.science.org/doi/10.1126/science.abj8754>. – Date of access: 11.08.2023.

2. UniProt: the Universal Protein Knowledgebase in 2023 / UniProt Consortium // Nucleic Acids Res. – 2023. – Vol. 51, № D1. – P. D523-D531. – Mode of access: <https://academic.oup.com/nar/article/51/D1/D523/6835362?login=false>. – Date of access: 03.08.2023.

3. Waterhouse, A. SWISS-MODEL: homology modelling of protein structures and complexes / A. Waterhouse [et al] // Nucleic Acids Res. – 2018. – Vol. 46, № W1. – P. W296–W303. – Mode of access: <https://academic.oup.com/nar/article/46/W1/W296/5000024>. – Date of access: 15.08.2023.

4. Webb, B. Protein Structure Modeling with MODELLER / B. Webb, A. Sali // Methods Mol Biol. – 2021. – Vol. 2199. – P. 239-255. – Mode of access: <https://escholarship.org/content/qt2023k96v/qt2023k96v.pdf?t=rdxpn>. – Date of access: 08.08.2023.

**МИНИСТЕРСТВО ЗДРАВООХРАНЕНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ «ВИТЕБСКИЙ
ГОСУДАРСТВЕННЫЙ ОРДЕНА ДРУЖБЫ НАРОДОВ
МЕДИЦИНСКИЙ УНИВЕРСИТЕТ»**

**«СТУДЕНЧЕСКАЯ МЕДИЦИНСКАЯ НАУКА
XXI ВЕКА»**

*XXIII Международная научно-практическая конференция
студентов и молодых ученых*

26-27 октября 2023 г.

Витебск, 2023