©Коллектив авторов

НЕЙРОСЕТЕВОЕ МОДЕЛИРОВАНИЕ МУЛЬТИТАРГЕТНОЙ RAGE-ИНГИБИРУЮЩЕЙ АКТИВНОСТИ

П.М. Васильев*, А.А. Спасов, Л.Р. Яналиева, А.Н. Кочетков, В.В. Ворфоломеева, В.Г. Клочков, Д.Т. Аппазова

Волгоградский государственный медицинский университет,
400131, Волгоград, пл. Павших борцов, 1; *эл. почта: pvassiliev@mail.ru

На основе методологии искусственных нейронных сетей построены модели, описывающие зависимость уровня RAGE-ингибирующей активности от аффинности соединений к белкам-мишеням сигнального пути RAGE-NF-κB. Для этого была сформирована валидированная база данных по структурам и уровням активности 183 известных соединений, испытанных, по литературным данным, на RAGE-ингибирующую активность. Проведён анализ сигнальных путей AGE-RAGE, найдены 14 основных узлов сигнального пути RAGE-NF-κB, для которых выявлены 34 релевантных белка-мишени. Сформирована база данных по 66 валидным 3D-моделям 22 белков-мишеней сигнальной цепочки RAGE-NF-κB. Выполнен ансамблевый молекулярный докинг 183 известных ингибиторов RAGE в сайты 66 валидных 3D-моделей 22 релевантных RAGE белков-мишеней и определены минимальные величины оценочной функции докинга для каждого соединения в отношении каждой мишени. По методике искусственных многослойных перцептронных нейронных сетей построены классификационные модели для прогноза уровня RAGE-ингибирующей активности по расчётной аффинности соединений к значимым белкам-мишеням сигнальной цепи RAGE-NF-κB. Проведена оценка прогностической способности этих моделей; максимальная точность по данным ROC анализа составила 90% для высокого уровня активности. Выполнен анализ чувствительности разработанных нейросетевых мультитаргетных моделей, определены наиболее значимые мишени цепочки передачи сигнала RAGE-NF-κB. Найдено, что для высокого уровня RAGE-ингибирующей активности наиболее значимыми биомишенями являются не AGE-рецепторы, а восемь сигнальных киназ пути RAGE-NF-κB и транскрипционный фактор NF-κB1. Предположено, что известные соединения с высокой RAGE-ингибирующей активностью по сути являются преимущественно ингибиторами сигнальных киназ.

Ключевые слова: сигнальный путь RAGE-NF-κB; RAGE-ингибиторы; мультитаргетная аффинность; молекулярный докинг; искусственные нейронные сети

DOI: 10.18097/PBMC20196502091

ВВЕДЕНИЕ

При сахарном диабете значительно усиливается неферментативное гликирование белков вследствие хронической гипергликемии, обусловленной данным заболеванием. В результате этого процесса образуются конечные продукты гликирования (advanced glycation end products, AGE), которые являются лигандами соответствующего специфического рецептора (RAGE) [1]. Сигнальная цепочка RAGE–NF-κB является одним из провоспалительных сигнальных путей, определяющих активацию RAGE [2]. Конечная точка приёма сигнала – ядерный фактор каппа-би (NF-κB) – является транскрипционным фактором многих генов, обуславливающих различные клеточные повреждения; его активация может приводить к усилению апоптоза, воспалительным процессам, аутоиммунным и опухолевым заболеваниям. NF-κB является транскрипционным фактором гена AGER, экспрессия которого приводит к синтезу новых AGE рецепторов (RAGE), а это, за счёт положительной обратной связи, приводит к существенному усилению воспалительных процессов, осуществляемых при их активации [3]. Показано, что активация RAGE является одним из значимых факторов при развитии осложнений сахарного диабета и болезни Альцгеймера [1, 3-6]. Следовательно, ингибиторы RAGE могут рассматриваться как новые перспективные фармакологические соединения для лечения данных патологий. При направленном поиске таких веществ весьма важно учитывать не только прямое ингибирование RAGE, но и пострецепторное действие новых соединений. Таким образом, построение *in silico* системной модели RAGE-ингибирующей активности, учитывающей влияние химических соединений на ключевые белки-мишени, образующие сигнальный путь RAGE–NF-κB, может способствовать поиску системных полифункциональных мультитаргетных ингибиторов рецепторов конечных продуктов гликирования – принципиально новых лекарственных веществ для лечения патологий при сахарном диабете и болезни Альцгеймера.

Целью проведенного исследования была разработка нейросетевой мультитаргетной модели для прогноза уровня RAGE-ингибирующей активности по расчётной аффинности соединений к значимым белкам-мишеням сигнальной цепи RAGE–NF-κB.

МЕТОДИКА

Формирование валидированной базы данных по структуре и уровню активности известных RAGE-ингибиторов

Поиск информации по структуре и активности соединений, испытанных на RAGE-ингибирующую активность для человека, был выполнен в базах

данных ChEMBL [7], BindingDB [8] и PubChem [9]. Найдено 364 единичных записи. Осуществлены экспертный анализ и проверка полученной информации, исправлены ошибки и проведена унификация представления химических структур, устранены дублирующие данные. Создана валидированная база данных из 183 соединений, для которых известны количественные значения RAGE-ингибирующей активности. С помощью программы Statistica 8.0 [10] методом k-средних выполнен кластерный анализ количественных значений RAGE-ингибирующей активности. Так как в разных методиках для RAGE-ингибирующей активности регистрируются разные показатели активности (например, IC_{50}, % ингибирования в нескольких фиксированных концентрациях и др.), кластерный анализ выполняли отдельно по каждому такому показателю, затем результаты объединяли. Таким образом, были сформированы четыре класса активности с градациями: High (38 соединений), Moderate (61 соединение), Low (39 соединений) и Inactive (45 соединений). Метки уровня RAGE-ингибирующей активности соединений были занесены в базу данных. На созданную базу данных получено свидетельство о государственной регистрации [11].

Формирование валидированной базы данных по 3D-моделям белков-мишеней, релевантных сигнальному пути RAGE–NF-κB

I. Ранее на основе базы данных KEGG была найдена общая схема сигнальных путей AGE–RAGE человека [2], выделена сигнальная цепь RAGE–NF-κB и определены 14 основных узлов данного сигналинга. Из баз данных UniProtKB [12] и PubMed [13] получена информация о возможных изоформах, подтипах и типах белков-мишеней, которые могут участвовать в передаче сигнала по цепи RAGE–NF-κB. Для 14 основных узлов этого сигнального пути найдены 78 монофункциональных белков, субъединиц и изоформ [14].

II. По каждому отдельному белку был выполнен контекстный поиск в PubMed [13], определяющий соотносимость его как биомишени с осложнениями сахарного диабета или с болезнью Альцгеймера. Всего из 78 белков выделено 34 белка-мишени, релевантных сигнальной цепи RAGE–NF-κB [14].

III. В банке данных экспериментальных моделей PDBe [15] и в базе данных теоретических моделей ModBase [16] после проверки на адекватность и первичной валидации с помощью программы VMD 1.9.2 [17] были найдены 826 3D-моделей 34 белков-мишеней *Homo sapiens*, релевантных сигнальной цепи RAGE–NF-κB.

IV. Для каждого из 34 видов релевантных биомишеней было определено местоположение специфического сайта связывания и ключевые аминокислотные остатки, ответственные за фиксацию лигандов в этом сайте. Определение набора ключевых аминокислотных остатков, участвующих в связывании ингибиторов, выполняли последовательно по следующей схеме:

1. Из доступных в PubMed [13] публикаций по экспериментальному изучению активности изоформ с различными точечными мутациями в специфическом сайте связывания ингибиторов определяли первичный набор ключевых аминокислотных остатков.

2. Полученные данные (в том числе, при их отсутствии) дополняли результатами сравнительного анализа экспериментальных рентгеноструктурных 3D-моделей комплексов конкретного белка с его известными ингибиторами. В каждой 3D-модели с использованием программ LigPlot+ 1.4.5 [18] и LigandScout 4.2.1 [19] выявляли набор аминокислотных остатков, обеспечивающих фиксацию лиганда в сайте. С помощью оригинальной программы BestSites 18.3.19 по всем 3D-моделям этого белка проводили сопоставление полученных данных с определением ключевых аминокислотных остатков.

3. Если в результате выполнения двух предыдущих шагов для данного белка-мишени не были найдены ключевые связывающие аминокислотные остатки, то в PubMed [13] проводили поиск публикаций по молекулярному моделированию связывания лигандов с этим белком. При наличии в таких публикациях данных о ключевых остатках, их принимали наравне с экспериментальными результатами для дальнейшего использования.

4. Если в итоге для данного белка-мишени всё-таки не были найдены ключевые аминокислотные остатки, такую биомишень исключали из дальнейшего рассмотрения.

Всего из 34 белков-мишеней *Homo sapiens*, релевантных сигнальному пути RAGE–NF-κB, были отобраны 22 биомишени, для которых была получена достоверная информация о локализации специфических сайтов связывания и их ключевых связывающих аминокислотах.

V. Валидацию 3D-моделей белков-мишеней, релевантных сигнальному пути RAGE–NF-κB, выполняли в несколько этапов.

1. Для каждого белка-мишени в базах данных IUPHAR/BPS [20] и ChEMBL [7] была найдена информация о его высокоактивных референсных ингибиторах. Отбиралось от 5 до 21 таких структур (в зависимости от полноты информации).

2. С помощью оригинальной системы Microcosm BioS 18.1.9 для каждого референсного ингибитора было найдено по одному структурно наиболее сходному с ним соединению, неактивному в отношении заданного белка-мишени. В состав системы Microcosm BioS 18.1.9 входит оригинальная QSAR-база, содержащая информацию по структуре 625888 известных химических соединений, испытанных, по литературным данным, на 11509 различных видов биологической активности.

3. Для всех найденных референсных ингибиторов и неактивных соединений были построены оптимальные конформации. Сначала отдельно для каждой из структур методами молекулярной механики с помощью программы MarvinSketch 17.1.23 [21] были построены по 10 конформеров с наименьшей энергией.

Все построенные конформеры были оптимизированы с помощью программы MOPAC2016 [22] с использованием полуэмпирического квантово-химического метода PM7. Среди оптимизированных конформеров было отобрано для каждого соединения по одному конформеру с наименьшей общей энергией.

4. Для каждой 3D-модели выполняли ансамблевый молекулярный докинг в ингибирующий сайт референсных и неактивных соединений, соответствующих данному белку. Докинг выполняли с помощью программы AutoDock Vina 1.1.2 [23] с пятикратным повтором для каждого лиганда. Отдельно по всем референтным и по всем неактивным соединениям вычисляли среднее значение минимальных энергий докинга. Метрикой валидности конкретной 3D-модели служила разность $\Delta E_{Валидации} = \Delta E_{Референсов} - \Delta E_{Неактивных}$. Для заданного белка-мишени отбирали три наиболее валидных модели с минимальными значениями $\Delta E_{Валидации}$.

В итоге была сформирована база данных по 66 валидным 3D-моделям 22 белков-мишеней, релевантных сигнальной цепочке RAGE–NF-κB.

Молекулярный докинг известных RAGE-ингибиторов в сайты белков-мишеней, релевантных сигнальному пути RAGE–NF-κB

Оптимальные 3D-конформации 183 известных RAGE-ингибиторов были рассчитаны последовательно методами молекулярной механики с помощью программы MarvinSketch 17.1.23 [21] и полуэмпирическим квантово-химическим методом PM7 с помощью программы MOPAC2016 [22], по схеме, описанной выше в предыдущем разделе в п. V.3.

Для каждого из 22 релевантных белков-мишеней в каждую из трёх соответствующих ему валидных 3D-моделей выполняли ансамблевый молекулярный докинг в ингибирующий сайт оптимизированных 3D-моделей всех 183 известных RAGE-ингибиторов. Докинг выполняли с помощью программы AutoDock Vina 1.1.2 [23] с пятикратным повтором для каждого лиганда. Для каждого соединения и каждой биомишени по полученным 15 значениям определяли минимальную энергию докинга. В результате по массиву из 60390 значений энергий докинга была сформирована матрица аффинности 183 известных ингибиторов RAGE в отношении 22 белков-мишеней сигнальной цепочки RAGE–NF-κB, включающая 4026 минимальных значений ΔE.

Построение нейросетевых мультитаргетных моделей зависимости RAGE-ингибирующей активности соединений от их расчётной аффинности к белкам-мишеням сигнального пути RAGE–NF-κB

Классификационные модели для прогноза уровня RAGE-ингибирующей активности были построены с применением искусственных многослойных перцептронных нейронных сетей с помощью модуля Statistica Neural Networks [24] программы Statistica 8.0 [10] по следующей схеме.

По результатам кластерного анализа уровня активности 183 известных RAGE ингибиторов для целей бинарной классификации были сформированы следующие три пары классов соединений, соответствующие бинарным переменным: LevH = High – not High; LevHM = High or Moderate – Low or Inactive; LevA = Active – Inactive.

Затем была сформирована обучающая выборка по 183 известным ингибиторам RAGE, включающая метки уровня активности соединений и значения ΔE их аффинности к 22 белкам-мишеням.

После выбора в Statistica классификационной модели и загрузки обучающей выборки в качестве выходной переменной выбирали одну из колонок с метками уровня активности (например, LevH), а в качестве входных переменных использовались показатели аффинности ΔE к 22 белкам-мишеням.

В соответствии с теоремой Колмогорова [25], зависимость любой сложности может быть аппроксимирована с помощью двухслойной искусственной нейронной сети. Поэтому в настоящем исследовании использовалась архитектура сети в виде двухслойного перцептрона MLP k-m-2 с узким горлом (bottleneck). Здесь k – число входных нейронов, в данном случае 22; m – число скрытых нейронов, устанавливается программой от 3 до 21, поскольку 2<m<k. В программе Statistica при построении перцептронных сетей для скрытого и выходного слоев используются пять наиболее распространенных активационных функций (Identity, Logictic, Tanh, Exponential, Sine) [24], попарный перебор которых также осуществляется программой.

Обучение нейросетей в программе Statistica проводится с использованием алгоритма обратного распространения ошибок. С целью достижения наилучшего результата обучения число сетей для обучения было увеличено с 20 до 200, а число наилучших автоматически отбираемых сетей – с 5 до 50. После окончания обучения для заданного уровня активности из 50 лучших отобранных программой нейросетей по совокупности характеристик точности обучения и тестирования вручную отбирали пять лучших.

Для 15 полученных нейросетей с помощью модуля Statistica Neural Networks [24] программы Statistica 8.0 [10] был проведён ROCK-анализ, по результатам которого для последующего исследования чувствительности было отобрано по одной наилучшей нейросети для каждого уровня RAGE-ингибирующей активности.

Анализ чувствительности построенных нейросетевых мультитаргетных моделей и определение наиболее значимых биомишеней сигнального пути RAGE–NF-κB

Был проведён анализ чувствительности двух лучших нейросетевых моделей для градаций RAGE-ингибирующей активности High и High or Moderate; анализ выполнен с помощью модуля Statistica Neural Networks [24] программы Statistica 8.0 [10].

Известно, что ингибирование протеинкиназ может быть сопряжено с высокой токсичностью соединений [26]. Поэтому для 38 соединений с высокой RAGE-ингибирующей активностью

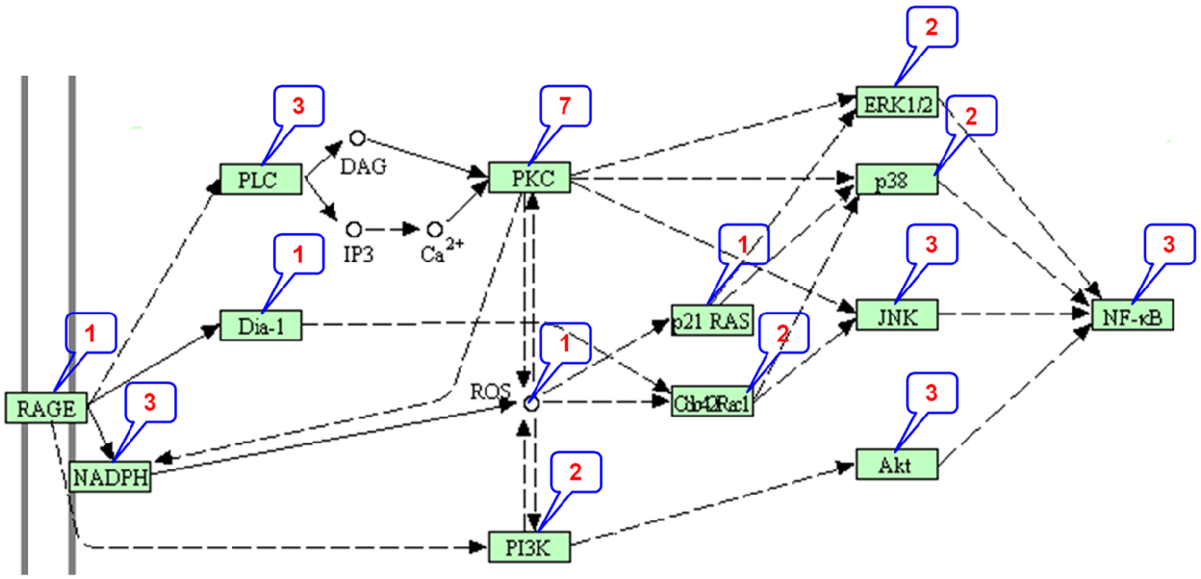

Рисунок 1. Сигнальный путь RAGE-NF-κB: цифрами указано число релевантных белков-мишеней для данного узла сигнальной цепи.

был выполнен консенсусный прогноз значений класса острой токсичности с использованием интернет-ресурсов GUSAR [27] и ProTox-II [28], а также оригинальной системы Microcosm ADMET 17.6.6, прогноз в которой осуществляется на основе структурного сходства прогнозируемого соединения с 40589 известными соединениями, изученными, по литературным данным, на острую токсичность в отношении мышей и крыс.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Сигнальный путь RAGE–NF-κB с 14 основными узлами приведен на рисунке 1.

Для этих 14 узлов найдено 78 белков-мишеней: RAGE – 1; NADPH (здесь и далее использованы обозначения белков в KEGG, в данном случае имеется в виду NADPH-оксидаза) – 13; PLC – 13; Dia-1 – 1; PKC – 15; ROS – 1; PI3K – 14; p21 RAS – 1; CDC42/RAC1 – 2; ERK – 2; p38 – 4; JNK– 3; AKT – 3; NF-κB – 5. В результате контекстного поиска среди этих 78 белков-мишеней выявлено 34 биомишени, релевантных осложнениям диабета и болезни Альцгеймера: NF-κB1 (p50); NF-κB2 (p52); RelA (p65, NF-κB3); ERK1 (MAPK3); ERK2 (MAPK1); p38 MAPK13; p38 MAPK14; JNK1 (MAPK8); JNK2 (MAPK9); JNK3 (MAPK10); AKT1; AKT2; AKT3; p21 RAS (RASA1); CDC42; RAC1; PRKCA; PRKCB; PRKCG; PRKCD; PRKCE; PRKCH; PRKCQ; ROS1; PIK3CA; PIK3CB; PLCB1; PLCG1; PLCD1; Dia-1 DIAPH1; NOX1 (NOH-1, MOX1); NOX2 (gp91phox); NOX4 (RENOX); RAGE [14]. Число релевантных белков-мишеней для каждого узла сигнального пути RAGE–NF-κB указано на рисунке 1.

Среди этих 34 биомишеней выявлено 22 белка, для которых была найдена достоверная информация о расположении специфического сайта связывания и о его ключевых аминокислотных остатках. При подготовке к валидации для этих 22 биомишеней было найдено 527 экспериментальных и 100 теоретических адекватных 3D-моделей, 220 референсных лигандов и 220 структурно сходных с ними неактивных соединений (табл. 1).

После валидации для этих 22 белков-мишеней было отобрано 66 3D-моделей. Обобщённые результаты ансамблевого молекулярного докинга 183 RAGE-ингибиторов в эти 66 3D-моделей приведены в таблице 2.

Видно, что средняя энергия докинга соединений с различным уровнем RAGE-ингибирующей активности достаточно сильно различается только для четырёх мишеней: MAPK14, AKT2, PRKCA, PRKCQ; для остальных мишеней средние значения ΔE по классам активности сопоставимы.

В таблице 3 приведены характеристики трёх лучших нейросетевых мультитаргетных моделей для различных уровней RAGE-ингибирующей активности.

Все три модели статистически высоко достоверны, $p<<0{,}001$ по биномиальному критерию [29].

На рисунке 2 приведена структура наиболее активного соединения RAGE-0093 из сформированной базы данных [11], которое имеет RAGE-ингибирующую активность IC_{50}=16,5 мкМ и относится к классу высокоактивных веществ.

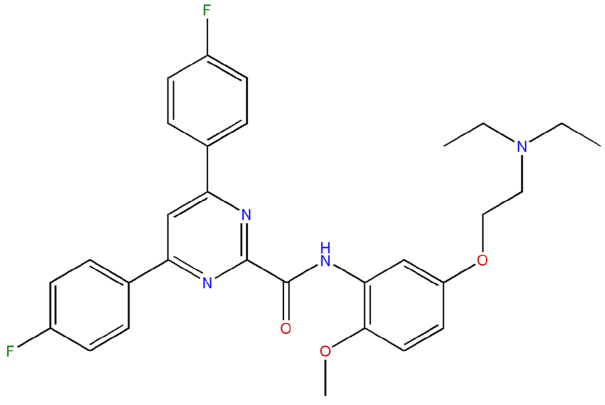

Рисунок 2. Структурная формула соединения RAGE-0093.

Таблица 1. Число 3D-моделей релевантных белков-мишеней сигнального пути RAGE-NF-κB

Шифр[1]	Название[2]	Число 3D-моделей[3]		Число соединений [4]
		PDBe	ModBase	
NF-κB1	Nuclear factor NF-kappa-B p105 subunit	3	—	14
ERK1	Extracellular signal-regulated kinase 1	2	10	16
ERK2	Extracellular signal-regulated kinase 2	71	—	22
MAPK13	Mitogen-activated protein kinase 13	7	—	22
MAPK14	Mitogen-activated protein kinase 14	222	—	42
JNK1	c-Jun N-terminal kinase 1	28	—	20
JNK2	c-Jun N-terminal kinase 2	2	5	20
JNK3	c-Jun N-terminal kinase 3	47	—	26
AKT1	RAC-alpha serine/threonine-protein kinase	15	—	22
AKT2	RAC-beta serine/threonine-protein kinase	16	—	22
CDC42	Cell division control protein 42 homolog	28	—	12
RAC1	Ras-related C3 botulinum toxin substrate 1	30	—	12
PRKCA	Protein kinase C alpha	2	4	20
PRKCB	Protein kinase C beta	1	18	26
PRKCG	Protein kinase C gamma	0	17	18
PRKCD	Protein kinase C delta	0	17	14
PRKCE	Protein kinase C epsilon	0	6	14
PRKCH	Protein kinase C eta	1	10	16
PRKCQ	Protein kinase C theta	5	—	10
ROS1	Proto-oncogene tyrosine-protein kinase ROS	2	13	12
PIK3CA	Phosphatidylinositol 4,5-bisphosphate 3-kinase C alpha	35	—	42
RAGE	Advanced glycosylation end product-specific receptor	10	—	18
Всего		527	100	440

Примечание: 1 – общепринятые сокращённые обозначения белков по данным UniProtKB [12]; 2 – номенклатурное название, рекомендуемое UniProtKB [12]; 3 – число моделей: экспериментальных из PDBe [15] и теоретических из ModBase [16]; 4 – число референсов для валидации моделей равно числу структурно сходных с ними неактивных соединений.

Таблица 2. Результаты ансамблевого молекулярного докинга известных RAGE-ингибиторов в сайты валидных 3D-моделей белков-мишеней, релевантных сигнальному пути RAGE-NF-κB

Шифр	ΔE, ккал/моль[1]				
	High[2]	Moderate[3]	Low[4]	Inactive[5]	δ[6]
NF-κB1	-7,59	-7,44	-7,31	-7,42	0,14
ERK1	-9,01	-8,87	-8,83	-8,87	0,09
ERK2	-9,31	-9,29	-9,17	-9,24	0,08
MAPK13	-9,35	-9,31	-9,14	-9,05	0,18
MAPK14	-8,44	-8,09	-8,08	-8,35	**0,22**
JNK1	-9,18	-9,12	-8,96	-9,00	0,13
JNK2	-9,48	-9,43	-9,28	-9,29	0,12
JNK3	-9,50	-9,42	-9,24	-9,35	0,14
AKT1	-9,95	-9,84	-9,70	-9,61	0,19
AKT2	-10,48	-10,5	-10,19	-9,95	**0,32**
CDC42	-8,38	-8,23	-8,21	-8,39	0,12
RAC1	-8,01	-7,92	-7,89	-8,00	0,07
PRKCA	-9,97	-9,85	-9,67	-9,53	**0,25**
PRKCB	-8,97	-8,82	-8,72	-8,73	0,14
PRKCG	-9,36	-9,29	-9,33	-9,34	0,03
PRKCD	-9,86	-9,77	-9,65	-9,56	0,18
PRKCE	-9,71	-9,68	-9,52	-9,58	0,11
PRKCH	-10,10	-9,92	-9,91	-9,88	0,11
PRKCQ	-10,02	-9,85	-9,55	-9,69	**0,26**
ROS1	-9,03	-8,93	-8,81	-8,74	0,17
PIK3CA	-8,92	-8,83	-8,74	-8,84	0,09
RAGE	-7,18	-7,09	-6,93	-6,97	0,15

Примечание: 1 – средняя энергия докинга; 2 – для класса соединений с градацией активности High; 3 – для класса соединений с градацией активности Moderate; 4 – для класса соединений с градацией активности Low; 5 – для класса соединений с градацией активности Inactive; 6 – средняя разность абсолютных значений ΔE при попарном сравнении классов.

При прогнозе с использованием трёх нейросетевых моделей (табл. 3) получаем следующие расчетные оценки уровней активности соединения RAGE-0093: High (0,528), High or Moderate (0,728), Active (0,749); в скобках указаны значения функции принадлежности к данному классу активности. Все три прогнозных оценки являются правильными и соответствуют высокому уровню экспериментальной активности указанного соединения.

На рисунке 3 приведены результаты анализа чувствительности двух лучших нейросетевых мультитаргетных моделей для градаций RAGE-ингибирующей активности High и High or Moderate.

В обоих случаях чувствительность наиболее важного входного нейрона белка RAGE оказалась весьма низкой, примерно 1,5 относительных единицы (минимальное значение 1,0). В то же время, чувствительность общих для двух градаций активности входных нейронов киназных белков-мишеней оказалась значительно выше: ERK2 – 1,78 и 2,01; MAPK14 – 2,62 и 1,80; JNK3 – 1,74 и 1,84; PRKCG – 1,82 и 2,28; PRKCD – 3,02 и 1,89; PRKCQ – 3,41 и 2,63; для градаций High и High or Moderate, соответственно.

Для уровня активности High существенную роль играют ещё три белка-мишени; чувствительность их входных нейронов: NF-κB1 – 1,90; JNK1 – 2,49; PRKCA – 1,83.

Наиболее существенно на высокий уровень RAGE-ингибирующей активности влияет ингибирование четырёх киназ: PRKCQ, PRKCD, MAPK14 и JNK1.

Таким образом, высокий уровень RAGE-ингибирующей активности определяет ингибирование не собственно AGE-рецептора, а ингибирование восьми его сигнальных киназ ERK2, MAPK14, JNK1, JNK3, PRKCA, PRKCG, PRKCD, PRKCQ и транскрипционного фактора NF-κB1.

Таблица 3. Характеристики нейросетевых мультитаргетных моделей зависимости RAGE-ингибирующей активности соединений от их расчётной аффинности к белкам-мишеням сигнального пути RAGE-NF-κB

Уровень активности	Архитектура сети	Точность модели, %		
		Train[1]	Test[2]	AUC[3]
High	MLP 22-10-2 Exp Log	90,5	83,3	89,7
High or Moderate	MLP 22-6-2 Exp Log	86,4	77,8	85,8
Active	MLP 22-15-2 Log Tanh	85,0	80,6	81,0

Примечание: 1 – при обучении; 2 – в независимом тестировании, тестовая выборка формировалась при каждом обучении нейросети с помощью генератора случайных чисел в объёме 20% от исходной выборки; 3 – по данным ROCK анализа.

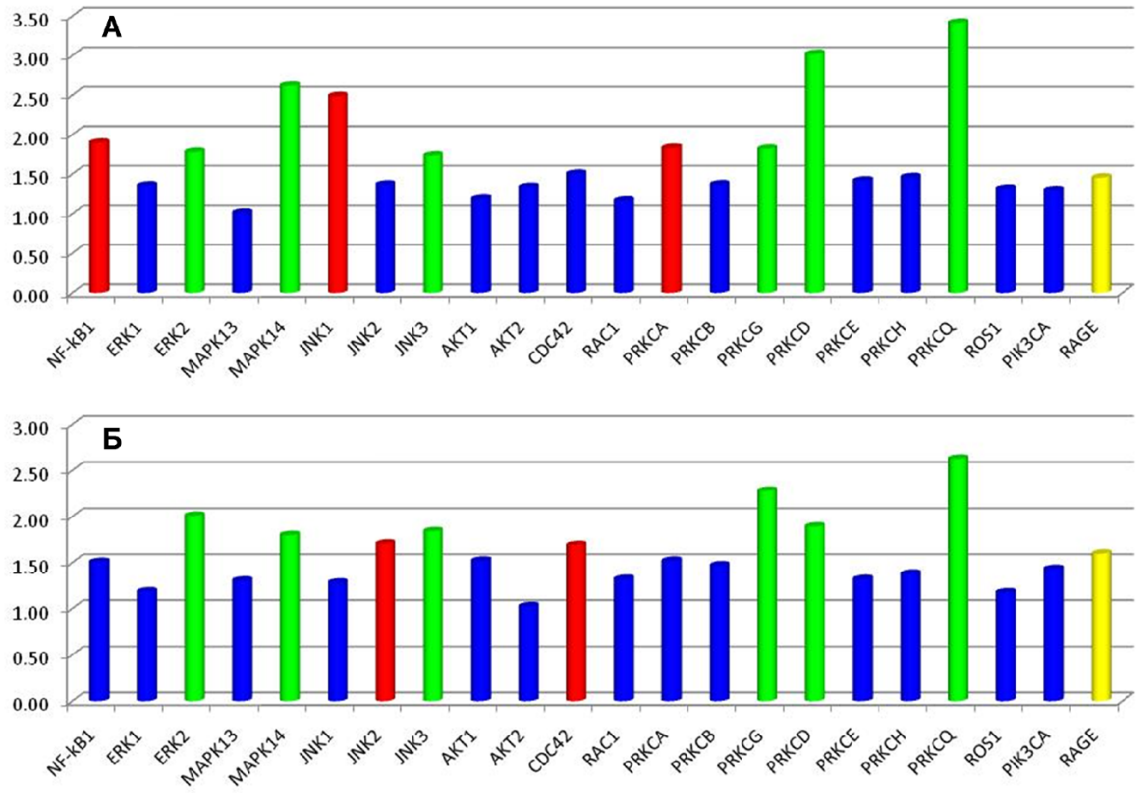

Рисунок 3. Анализ чувствительности нейросетевых мультитаргетных моделей RAGE-ингибирующей активности: А – для градации активности High; Б – для градации активности High or Moderate; ось Y – относительная чувствительность входных нейронов в условных безразмерных единицах.

Выше было упомянуто, что ингибирование киназ может обуславливать высокую токсичность соединений [26]. Однако консенсусный прогноз значений класса острой токсичности для 38 соединений с высокой RAGE-ингибирующей активностью показал, что 82% этих веществ относятся к 4 классу токсичности (мало токсичны), а остальные 18% – к 5 классу токсичности (практически не токсичны). Можно предположить, что ингибиторы AGE рецептора, активные в отношении киназ сигнального пути RAGE–NF-κB, не будут проявлять сколько-нибудь высокой острой токсичности, что характеризует их как перспективных лекарственных кандидатов.

Необходимо отметить, что чувствительность всех 22 входных нейронов для градации RAGE-ингибирующей активности “active” оказалась примерно одинаковой и колебалась в диапазоне от 1,00 до 1,35.

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

Методом искусственных нейронных сетей построены мультитаргетные модели зависимости уровня RAGE-ингибирующей активности соединений от их расчётной аффинности к белкам-мишеням сигнального пути RAGE–NF-κB, показана их высокая статистическая достоверность.

Анализ чувствительности полученных нейросетевых мультитаргетных моделей RAGE-ингибирующей активности позволил выявить наиболее значимые биомишени цепочки передачи сигнала RAGE–NF-κB.

На основании проведённых расчётов установлено, что высокий уровень RAGE-ингибирующей активности соединений определяется не только ингибированием AGE-рецептора, но преимущественно ингибированием восьми его сигнальных киназ ERK2, MAPK14, JNK1, JNK3, PRKCA, PRKCG, PRKCD, PRKCQ и транскрипционного фактора NF-κB1. Таким образом, решающий вклад в ингибирование RAGE вносит пострецепторное действие.

Полученные нейросетевые модели будут использованы в направленном поиске системных полифункциональных мультитаргетных ингибиторов рецепторов конечных продуктов гликирования – принципиально новых лекарственных веществ для лечения патий при сахарном диабете и болезни Альцгеймера.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 18-015-00499).

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит каких-либо исследований с участием людей или с использованием животных в качестве объектов.

ЛИТЕРАТУРА

1. *Ансари Н.А., Рашид З.* (2010) Биомед. химия, **56**, 168-178.
2. KEGG: AGE-RAGE signaling pathway in diabetic complications – *Homo sapiens* (2017), https://www.genome.jp/kegg-bin/show_pathway?map=hsa04933 &show_description=show.
3. *Tobon-Velasco J.C., Cuevas E., Torres-Ramos M.A.* (2014) CNS Neurol. Disord. Drug Targets, **13**, 1615-1626.
4. *Yan S.F., Ramasamy R., Schmidt A.M.* (2008) Nat. Clin. Pract. Endocrinol. Metab., **4**, 285-293.
5. *Singh V.P., Bali A., Singh N., Jaggi A.S.* (2014) Korean. J. Physiol. Pharmacol., **18**, 1-14.
6. *Matrone C., Djelloul M., Taglialatela G., Perrone L.* (2015) Histol. Histopathol., **30**, 125-39.
7. ChEMBL (2018), https://www.ebi.ac.uk/chembl/
8. BindingDB (2018), https://www.bindingdb.org/bind/index.jsp
9. PubChem (2018), https://pubchem.ncbi.nlm.nih.gov/
10. Statistica (2007), http://www.statsoft.com/Products/STATISTICA-Features
11. *Васильев П.М., Яналиева Л.Р., Спасов А.А., Кочетков А.Н., Ворфоломеева В.В., Клочков В.Г.* (2019) Свидетельство о государственной регистрации базы данных №2019620160. Ингибиторы рецепторов конечных продуктов гликирования (24.01.2019), Офиц. бюлл. Программы для ЭВМ. БД. ТИМС, **2**, RU 2019620160.
12. UniProtKB (2018), https://www.uniprot.org/uniprot/
13. PubMed (2018), https://www.ncbi.nlm.nih.gov/pubmed/
14. *Васильев П.М., Яналиева Л.Р., Кочетков А.Н., Ворфоломеева В.В., Клочков В.Г.* (2018) Вестник ВолгГМУ, **3**, 133-138
15. PDBe (2018), http://www.ebi.ac.uk/pdbe/
16. ModBase (2018), https://modbase.compbio.ucsf.edu/modbase-cgi/
17. VMD (2016), https://www.ks.uiuc.edu/Research/vmd/
18. *Laskowski R.A., Swindells M.B.* (2011) J. Chem. Inf. Model., **51**, 2778-2786.
19. Inte:ligand (2018), http://www.inteligand.com/
20. IUPHAR/BPS (2018), http://www.guidetopharmacology.org/
21. ChemAxon: Marvin (2017), https://chemaxon.com/products/marvin
22. MOPAC (2016), http://openmopac.net/
23. *Trott O., Olson A.J.* (2010) J. Comp. Chem., **31**, 455-461.
24. Нейронные сети. Statistica Neural Networks: Методология и технология современного анализа данных (2008), Горячая линия - Телеком, М.
25. *Колмогоров А.Н.* (1958) Докл. АН СССР, **114**, 953-956.
26. *Yang B., Papoian T.* (2018) J. Appl. Toxicol., **38**, 790-800.
27. GUSAR AcuteRodentstoxicity (2019), http://www.way2drug.com/gusar/acutoxpredict.html
28. ProTox-II (2019), http://tox.charite.de/protox_II/index.php?site=compound_input
29. *Глотов Н.В., Животовский Л.А., Хованов Н.В., Хромов-Борисов Н.Н.* (1982) Биометрия, Изд-во Ленингр. ун-та, Л.

Поступила в редакцию: 20. 01. 2019.
После доработки: 19. 02. 2019.
Принята к печати: 13. 03. 2019.

NEURAL NETWORK MODELING OF MULTITARGET RAGE INHIBITORY ACTIVITY

P.M. Vassiliev*, A.A. Spasov, L.R. Yanaliyeva, A.N. Kochetkov, V.V. Vorfolomeyeva, V.G. Klochkov, D.T. Appazova

Volgograd State Medical University,
1, Pavshikh bortsov sq., Volgograd, 400131 Russia; *e-mail: pvassiliev@mail.ru

Based on the methodology of artificial neural networks, models describing the dependence of the level of RAGE inhibitory activity on the affinity of compounds for target proteins of the RAGE-NF-κB signal pathway have been costructed. A validated database of the structures and activity levels of 183 known compounds, which were tested for RAGE inhibitory activity was formed. The analysis of the AGE-RAGE signaling pathways was carried out, 14 key RAGE-NF-κB signal pathway nodes were found, for which 34 relevant target proteins were identified. A database of 66 valid 3D models of 22 target proteins of the RAGE-NF-κB signal chain was compiled. Ensemble molecular docking of 3D models of 183 known RAGE inhibitors into sites of 66 valid 3D models of 22 relevant RAGE target proteins was performed and minimum docking energies for each compound were determined for each target. According to the method of artificial multilayer perceptron neural networks, classification models were constructed to predict level of RAGE inhibitory activity based on the calculated affinity of compounds for significant target proteins of the RAGE-NF-κB signaling chain. The prognostic ability of these models of RAGE-inhibitory activity was evaluated, the maximum accuracy according to ROC-analysis was 90% for a high level of activity. The sensitivity analysis of the developed multitarget models were carried out, the most significant targets of the RAGE-NF-κB signal transmission chain were determined. It was found that for high level of RAGE inhibitory activity, the most significant biotargets are not AGE receptors, but eight signaling kinases of the RAGE-NF-κB pathway and transcription factor NF-κB1. Thus, it is suggested that known compounds with high RAGE-inhibitory activity are preferential inhibitors of signal kinases.

Key words: RAGE-NF-κB signaling pathway; RAGE inhibitors; multitarget affinity; molecular docking; artificial neural networks