

УДК 544.165

©Коллектив авторов

ПРЕДСКАЗАНИЕ СПЕКТРА ДЕЙСТВИЯ ЛИГАНДОВ РЕЦЕПТОРОВ СТЕРОИДНЫХ ГОРМОНОВ

И.В. Федюшкина, И.В. Ромеро Рейес, А.А. Лагунин, В.С. Скворцов*

Федеральное государственное бюджетное учреждение
“Научно-исследовательский институт биомедицинской химии
имени В.Н. Ореховича” Российской академии медицинских наук, 119121
Москва, Погодинская ул., д. 10; тел.: 499-245-0768; эл. почта: ivf@ibmh.msk.ru

С использованием двух основных методов расчёта (PASS и SIMCA) созданы модели, предсказывающие тип физиологического ответа стероидных соединений при связывании с ядерными рецепторами стероидных гормонов. В работе рассмотрено 10 моделей агонистов и антагонистов эстрогеновых, прогестероновых, андрогенных, глюкокортикоидных и минералокортикоидных рецепторов. Для SIMCA были рассмотрены два варианта, когда в качестве дескрипторов использовались молекулярные и конституционные свойства в первом и индексы сходства во втором. Показано, что с использованием дискриминантного анализа для любого произвольного стероида на основании его структурной формулы можно предсказать спектр его действия со средней точностью 80-85%.

Ключевые слова: рецепторы стероидных гормонов, прогностическая модель, дискриминантный анализ, SIMCA.

ВВЕДЕНИЕ. Ядерные рецепторы стероидных гормонов входят в большое суперсемейство транскрипционных факторов, обеспечивающих функционирование многоклеточного организма на всех уровнях, включая эмбриональное развитие, рост и дифференцировку клеток, регуляцию апоптоза и обмен веществ [1]. К настоящему времени описано пять типов ядерных рецепторов стероидных гормонов: 1) эстрогеновые рецепторы (ЭР); 2) прогестероновые рецепторы (ПР); 3) андрогенные рецепторы (АР); 4) глюкокортикоидные рецепторы (ГР); 5) минералокортикоидные рецепторы (МР) [1]. В пределах одного типа, существуют также подтипы и варианты, но в данной работе они не рассматривались. Ряд стероидов могут взаимодействовать с различными типами этих рецепторов (перекрёстное действие). Так, минералокортикоидные рецепторы могут связывать кортизол, а рецепторы прогестерона способны связывать помимо гестагенов также глюкокортикоиды, минералокортикоиды и в меньшей степени андрогены [2]. Некоторые стероиды, не являющиеся гормонами, могут взаимодействовать со всеми 5 типами, обуславливая развитие определенных эффектов [3, 4].

* - адресат для переписки

ПРЕДСКАЗАНИЕ СПЕКТРА ДЕЙСТВИЯ СТЕРОИДНЫХ ЛИГАНДОВ

Таким образом, при создании лекарственного прототипа возникает проблема селективности лиганда, когда желательно заранее определить спектр возможных взаимодействий конкретного вещества. Кроме того, конечный эффект действия любого лиганда зависит не только от аффинности рецептора, но и от того какие конкретно изменения будут инициированы. В случае ядерных рецепторов это означает, что любой лиганд может оказаться как агонистом природного гормона, так и антагонистом. Разумеется, что любой тип физиологического ответа должен быть подтверждён экспериментально. Однако, предсказание соответствующего эффекта может существенно упростить процедуру планирования эксперимента и обосновать выбор кандидата. Данная работа посвящена созданию моделей предсказания типа физиологического ответа стероидных соединений при связывании с ядерными рецепторами стероидных гормонов и оценке качества технической процедуры предсказания. Методы анализа структура/активность (SAR), использованные в работе, основаны на концепции структурного сходства (как прямого – сходства структурных формул, так и опосредованного – подобия расчётных дескрипторов) и широко используются достаточно давно [5].

МЕТОДИКА. Для построения прогностических моделей из литературных источников были выбраны стероиды с различными гормональными активностями на ядерные рецепторы.

Рабочие выборки соединений стероидной природы с установленным экспериментально типом физиологического ответа были сформированы на основе литературных источников [6-18]. Всего было сформировано 10 выборок. В каждую из них входили вещества, для которых было показано наличие конкретного физиологического ответа (например, “агонист ПР” или другая выборка “антагонист ПР”) и соединения, для которых было показано его отсутствие. Для ПР в качестве агонистов использовались, применяемые в клинической практике, природные и синтетические прогестагены [6], в качестве его антагонистов – синтетические производные мифепристона (они же являются антагонистами и для ГР) [7]. Агонисты ГР – синтетические кортикостероиды – аналоги флутиказона [8]. Агонисты ЭР – натуральные и синтетические эстрогены [6, 9], антагонисты – синтетические производные 17 β -эстрадиола [10]. Агонисты АР – набор природных и синтетических андрогенов [11], антагонисты – синтетические стероиды, включающие дополнительные гетероциклы [12]. Агонисты МР - синтетические производные альдостерона [13], Антагонисты – производные 17-спиронолактона [14]. Учитывая, что для прогестагенов показана перекрестная гормональная активность [15], они были добавлены в выборки для других рецепторов согласно описанным эффектам. В качестве соединений “без эффекта” для АР были использованы агонисты МР [16] и прогестины [17]. Для остальных были взяты агонисты АР [11], так как они не имеют прямого перекрестного взаимодействия с другими рецепторами [18]. Известно, что слабые андрогены применяют при лечении гинекологических заболеваний. Однако, это не связано с прямым воздействием этих препаратов на ЭР, они образуют комплексы с глобулинами, связывающими половые гормоны (ГСПС) и кортикостероиды (ГСК), кроме того, свободный тестостерон является предшественником для эстрогенов [6]. Выборки доступны в качестве дополнительных материалов на [www-сайте журнала \(http://pbmc.ibmc.msk.ru\)](http://pbmc.ibmc.msk.ru).

Для дискриминантного анализа в работе были использованы два основных метода: PASS (Prediction of Activity Spectra for Substance) [19] и метод формального независимого моделирования аналогий классов (SIMCA) [20].

PASS – компьютерная программа для прогнозирования спектра биологической активности химических соединений на основе их структурной формулы с использованием топологических дескрипторов многоуровневых атомных окрестностей [21]. Спектр биологической активности отражает результат взаимодействия химического соединения с различными биологическими объектами. В работе использовали версию программы PASS 2012, которая в числе 6400 видов предсказываемой биологической активности включает и агонистические и антагонистические эффекты при взаимодействии с исследуемыми рецепторами. Для тестируемого соединения, вычисляются вероятности принадлежности к классу соединений, проявляющих/не проявляющих конкретный вид биологической активности. Результат прогноза PASS представляется в виде упорядоченного списка названий соответствующих активностей и вероятностей P_a “быть активным” и P_i “быть неактивным”. В исследовании были рассмотрены 2 случая: соединения считались активными, если P_a было больше P_i (критерий определения активности в PASS по умолчанию), или разность $P_a - P_i$ была больше 0,5. Обучающие выборки PASS формируются авторами программы и в случае рецепторов стероидных гормонов содержат соединения различной химической природы, включая нестероидные соединения. Программа доступна в виде сетевого сервиса по адресу <http://www.pharmaexpert.ru/PASSOnline/>.

Метод SIMCA решает задачу классификации с обучением на основе метода главных компонент. Особенность этого метода заключается в построении главных компонент отдельно для каждого класса. В качестве распознаваемых классов для моделей были установлены два: наличие эффекта (1) и его отсутствие (0). В работе использовали параллельно две группы дескрипторов, рассчитанных на основании структурных формул. Первая группа была рассчитана с помощью web-приложения E-Dragon [22, 23] (табл. 1).

Таблица 1. Молекулярные дескрипторы для прогностических моделей.

Конституционные дескрипторы	Дескрипторы водородных связей	Молекулярные свойства
MW – молекулярный вес	nHDon – количество атомов-доноров для водородных связей (азот и кислород)	U _i – индекс ненасыщенности
nAT – количество атомов	nHAcc – количество атомов-акцепторов для водородных связей (азот, фтор, кислород)	H _u – фактор гидрофильности
nBT – количество связей		ALOGP – коэффициент распределения октанол-вода (logP)
RBN – количество вращающихся связей		TPSA(NO) – топологическая полярная площадь поверхности с использованием полярных вкладов атомов кислорода и азота
nDB – количество двойных связей		
nAB – количество ароматических связей		
nH – количество атомов водорода		
nC – количество атомов углерода		
nN – количество атомов азота		
nO – количество атомов кислорода		
nX – количество атомов галогенов		
S _v – суммарное количество Ван-дер-Ваальсовых объемов атомов		
S _e – суммарное количество электроотрицательных атомов		
S _p – суммарное количество поляризованных атомов		
S _s – суммарное количество электротопологических состояний атомов (индекс Кир-Холла)		

ПРЕДСКАЗАНИЕ СПЕКТРА ДЕЙСТВИЯ СТЕРОИДНЫХ ЛИГАНДОВ

Вторая группа в качестве дескрипторов использовала коэффициенты (индикаторы) молекулярного сходства с набором из 15 стероидных молекул, выделенных как при иерархической кластеризации набора из 154 стероидов, как центральный представитель кластера (список представлен в дополнительных материалах). В качестве меры подобия использовали коэффициент Танимото, все расчёты для данной группы проведены с использованием пакета SYBYL-X [24].

Оценка качества предсказания с использованием PASS была проведена по всей выборке. Метод SIMCA оценивали при помощи метода кросс-валидации. Сначала для всех лигандов была рассчитана мера молекулярного подобия с природным агонистом рецептора – коэффициент Танимото [20] (пример распределения представлен на рисунке 1). Затем случайным образом общая выборка делилась на тестовую (20%) и обучающую (80%) выборки. При этом наборы выбирались так, чтобы точки равномерно покрывали весь диапазон изменения коэффициента Танимото. Данную процедуру повторяли 1000 раз. На каждой итерации вычисляли ошибку предсказания (рис. 2), равную отношению количества верных ответов на тестовом множестве к их общему числу, ошибки первого и второго рода. Средние значения этих ошибок приведены в таблице 2.

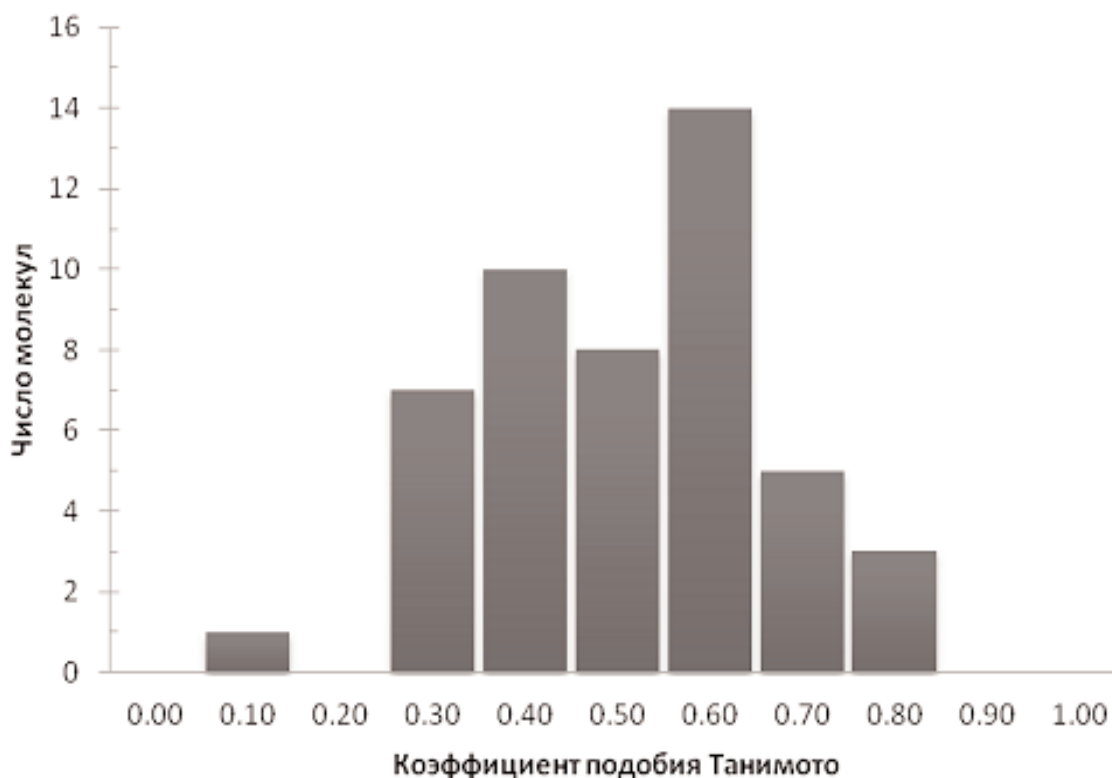


Рисунок 1.

Распределение выборки для антагонистов МР по молекулярному подобию с природным агонистом (альдостероном).

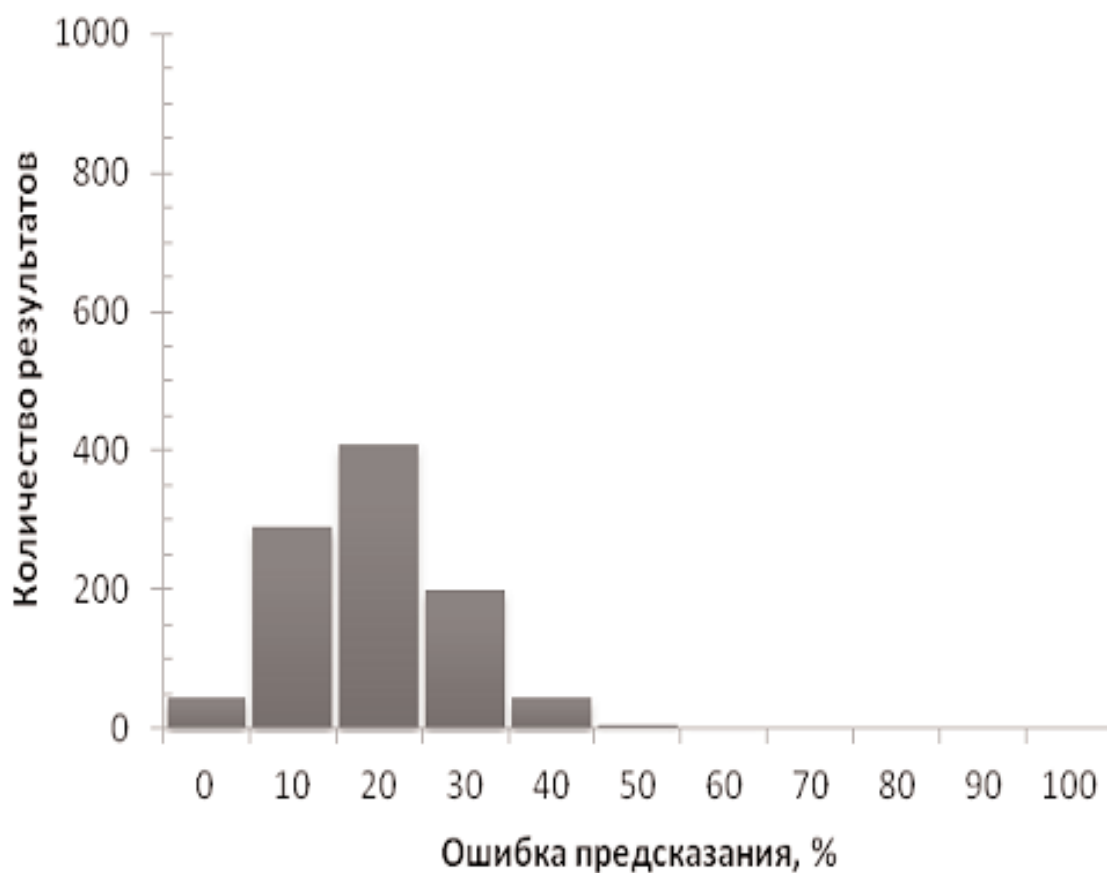


Рисунок 2.

Распределение ошибки предсказания модели (%) для 1000 итераций независимого обучения на примере антагонистов МР (дескрипторы Dragon).

Таблица 2. Результаты предсказания биологического эффекта программой PASS.

Эффект	N*	Pa>Pi							Pa-Pi>0,5				
		П	Н	ИП	ИН	ОП	ОН	СО (%)	ИП	ИН	ОП	ОН	СО (%)
АГ АР	219	24	31	24	0	31	0	56,36	23	15	16	1	30,91
АН АР	1280	21	34	21	0	34	0	61,82	18	13	21	3	43,64
АГ ЭР	2227	10	35	10	0	35	0	77,78	10	10	25	0	55,56
АН ЭР	734	22	23	22	0	23	0	51,11	2	15	8	20	62,22
АГ ГР	900	25	27	25	0	27	0	75,00	15	27	0	10	19,23
АН ГР	307	10	30	10	0	30	0	51,92	2	28	2	8	25
АГ МР	16	10	26	10	2	24	0	66,67	4	26	0	6	16,67
АН МР	441	14	34	14	0	34	0	70,83	12	34	0	2	4,17
АГ ПР	173	21	25	21	0	25	0	54,35	17	24	1	4	10,87
АН ПР	1088	10	24	10	0	24	0	70,59	10	23	1	0	2,94

Примечание. Здесь и в таблице 3 * - см. сокращения в тексте.

ПРЕДСКАЗАНИЕ СПЕКТРА ДЕЙСТВИЯ СТЕРОИДНЫХ ЛИГАНДОВ

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ. В таблице 2 представлены результаты работы программы PASS. Агонисты (АГ) и антагонисты (АН) сгруппированы по рецепторам (АР, ЭР, ГР, МР и ПР); вариант 1 – решение положительное если P_a больше P_i ; вариант 2 разность $P_a - P_i$ была больше 0,5. П и Н – позитивный (наличие эффекта) и негативный (отсутствие эффекта); ИП и ИН – истинно позитивный и истинно негативный, когда результаты предсказания совпадают реальностью; ОП и ОН – ошибочно позитивный и ошибочно негативный, когда результаты предсказания противоречат экспериментальным данным; СО – средняя ошибка. В данном случае “тестовой” является вся выборка. Несмотря на то, что с высокой долей вероятности большинство из тестируемых активных соединений были представлены в обучающей выборке PASS особенности алгоритма [25], особенно в части “неактивных” соединений позволяют допустить такой вариант. Для сравнения приведены размеры (N) обучающих выборок в PASS. Данные величины характеризуют количество активных соединений (включая нестероидные), неактивными считаются любые соединения, для которых заданная активность не упомянута в базе данных. Последнее представляется основной причиной того, что PASS не показал удовлетворительного результата. Легко видеть, что первый вариант ($P_a > P_i$) можно отбросить сразу. Практически все соединения в выборке отбираются как имеющие заданный эффект. Второй вариант отбора значительно лучше, однако и в этом случае для АР и ЭР средняя ошибка находится на уровне случайного выбора.

Результаты, показанные методом SIMCA на двух наборах дескрипторов (табл. 3), достаточно сходные. В общем случае модели, построенные на основе индикаторов сходства, имеют чуть лучшие параметры, однако в основном эти отличия не значимые (исключая АР). Следует помнить, что если несколько раз провести по 1000 итераций обучения, то средняя ошибка будет колебаться в пределах 1-2%. Различия проявляются, главным образом, при анализе ошибок первого (СО I) и второго (СО II) рода. Если модели, основанные на дескрипторах E-Dragon, в среднем ошибаются равнонаправленно, то модели на базе индикаторов сходства более склонны давать ложноположительные, а не ложноотрицательные предсказания.

Таблица 3. Результаты тестирования моделей, построенных методом SIMCA.

Эффект	П*	Н	Дескрипторы Dragon			Дескрипторы инд. сходства		
			СО (%)	СО I (%)	СО II (%)	СО (%)	СО I (%)	СО II (%)
АГ АР	24	31	29,94	17,54	12,40	22,44	14,45	7,98
АН АР	21	34	25,97	10,39	15,58	17,25	9,78	7,47
АГ ЭР	10	35	19,28	19,27	0,01	20,02	13,05	6,97
АН ЭР	22	23	33,55	13,56	19,99	29,49	22,4	7,08
АГ ГР	25	27	24,33	4,36	19,97	22,92	18,12	4,8
АН ГР	10	30	4,47	0	4,47	0	0	0
АГ МР	10	26	0,00	0	0,00	1,02	1	0,02
АН МР	14	34	19,48	15,44	4,02	18,82	12,12	6,7
АГ ПР	21	25	26,73	13,97	12,76	17,79	15,18	2,6
АН ПР	10	24	1,82	1,80	0,02	2	2	0

В качестве иллюстрации применения данного подхода можно привести спектр предсказанного действия для стероида растительного происхождения гуггулстерона [3], который используется в медицине, как антиатерогенное и противовоспалительное средство, а также применяется для спортивного питания в составе комплексных жиросжигателей. В таблице 4 показаны результаты предсказаний в сравнении с экспериментально подтверждёнными данными. Совпадающие величины выделены жирным шрифтом. Как видно из таблицы основные эффекты предсказываются правильно, но различные методы дают некоторые несоответствия. Отдельно стоящий случай – АГ ЭР. В работе Burriss T.P. et al. [3] указано, что гуггулстерон детектирован как слабый агонист ЭР, однако приведённые цифры говорят, что результаты показаны на грани чувствительности экспериментального метода. Последняя колонка в таблице даёт пример использования консенсуса различных методов для предсказания. В данном примере результат улучшается. Однако, по нашему мнению, трёх вариантов для использования консенсуса недостаточно.

Таблица 4. Пример предсказания эффектов, проявляющихся при взаимодействии гуггулстерона с рецепторами стероидных гормонов.

Эффект	Литературные данные	Предсказание PASS Pa-Pi>0.5	Предсказание SIMCA Dragon	Предсказание SIMCA инд. сходства	Предсказание Консенсус
АГ АР	Нет	Нет	Нет	Да	Нет
АН АР	Да	Да	Да	Нет	Да
АГ ЭР	<i>низкая активность</i>	<i>Да</i>	<i>Нет</i>	<i>Нет</i>	<i>Нет</i>
АН ЭР	Нет	Нет	Да	Нет	Нет
АГ ГР	Нет	Нет	Да	Нет	Нет
АН ГР	Да	Нет	Нет	Нет	Нет
АГ МР	Нет	Нет	Нет	Нет	Нет
АН МР	Да	Да	Да	Да	Да
АГ ПР	Да	Нет	Да	Да	Да
АН ПР	Нет	Нет	Нет	Нет	Нет
Совпадений:		7	6	6	8

Примечание: жирным шрифтом выделены совпадения с экспериментальными данными.

ЗАКЛЮЧЕНИЕ. В результате всего вышесказанного можно отметить, что с использованием дискриминантного анализа для любого произвольного стероида можно предсказать спектр его действия на рецепторы стероидных гормонов. Особо важно, что для их работы требуется только знание структурной формулы соединения. В самом худшем случае (АН ЭР) предсказание будет выполнено с точностью 67-70% при средней точности 80-85%. Результат можно считать вполне приемлемым для предварительного отбора первоочередных вариантов из множества кандидатов при планировании экспериментов и дальнейших исследований.

Авторы выражают благодарность заведующему лабораторией структурной биоинформатики ФГБУ ИБМХ РАМН д.б.н. А.В. Веселовскому за консультацию и помощь.

ЛИТЕРАТУРА

1. *Novac N., Heinzl T.* (2004) *Curr. Drug Targets Inflamm. Allergy*, **3**, 335-346.
2. *Lazar M.A.* (2002) in *Williams Textbook of Endocrinology* (Melmed S., Polonsky K.S., Larsen P.R., Henry M., eds.), Saunders, Philadelphia, USA, pp. 38-46.
3. *Burris T.P., Montrose C., Houck K.A., Osborne H.E., Bocchinfuso W.P., Yaden B.C., Cheng C.C., Zink R.W., Barr R.J., Hepler C.D., Krishnan V., Bullock H.A., Burris L.L., Galvin R.J., Bramlett K., Stayrook K.R.* (2005) *Mol. Pharmacol.*, **67**, 948-954.
4. *Wilkinson J.M., Hayes S., Thompson D., Whitney P., Bi K.* (2008) *J. Biomol. Screen*, **13**, 755-765.
5. *Nowell I.* (1986) *Endeavour*, **10**, 216-217.
6. *Бертрам Г.К.* (2008) *Базисная и клиническая фармакология*, Издательство "Диалект", Санкт-Петербург.
7. *Nickisch K., Elger W., Cessac J., Kesavaram N., Das B., Garfield R., Shi S.Q., Amelkina O., Meister R.* (2013) *Steroids*, **78**, 255-267.
8. *Millan D.S., Ballard S.A., Chunn S., Dybowski J.A., Fulton C.K., Glossop P.A., Guillabert E., Hewson C.A., Jones R.M., Lamb D.J., Napier C.M., Payne-Cook T.A., Renery E.R., Selby M.D., Tutt M.F., Yeadon M.* (2011) *Bioorg. Med. Chem. Lett.*, **21**, 5826-5830.
9. *Зайдиева Я.З.* (2001) *Заместительная гормонотерапия. Фармакология и клиническое применение*, Москва.
10. *Jiang X.R., Wang P., Smith C.L., Zhu B.T.* (2013) *J. Med. Chem.*, **56**, 2779-2790.
11. *Fang H., Tong W., Branham W.S., Moland C.L., Dial S.L., Hong H., Xie Q., Perkins R., Owens W., Sheehan D.M.* (2003) *Chem. Res. Toxicol.*, **16**, 1338-1358.
12. *Singh S.M., Gauthier S., Labrie F.* (2000) *Curr. Med. Chem.*, **7**, 211-247.
13. *Kirk D.N., Schroder M.H., Latif S.A., Souness G.W., Morris D.J.* (1993) *Steroids*, **58**, 59-63.
14. *Kandemirli F., Tokay N., Shvets N.M., Dimoglo A.S.* (2002) *Farmacol.*, **57**, 601-607.
15. *Kuhl H.* (2005) *Climacteric: the journal of the International Menopause Society*, **8 Suppl 1**, 3-63.
16. *Bledsoe R.K., Madauss K.P., Holt J.A., Apolito C.J., Lambert M.H., Pearce K.H., Stanley T.B., Stewart E.L., Trump R.P., Willson T.M., Williams S.P.* (2005) *J. Biol. Chem.*, **280**, 31283-31293.
17. *Kuhl H.* (2005) *Climacteric*, **8 Suppl 1**, 3-63.
18. *Дегтярь В.Т., Кушлинский, Н.Е.* (2005) *Сибирский онкологический журнал*, **3**, 58-67.
19. *Филимонов Д.А., Пороиков В.В.* (2006) *Российский химический журнал*, **50**, 66-75.
20. *Wold S., Sjostrom M.* (1976) in *Chemometrics: Theory and Application* (Kowalski B., ed.) American Chemical Society, Washington, pp. 243-282.
21. *Filimonov D., Poroikov V., Borodina Y.V., Glorizova T.* (1999) *J. Chem. Inf. Comput. Sci.*, **39**, 666-670.
22. *Tetko I.V., Gasteiger J., Todeschini R., Mauri A., Livingstone D., Ertl P., Palyulin V.A., Radchenko E.V., Zefirov N.S., Makarenko A.S., Tanchuk V.Y., Prokopenko V.V.* (2005) *J. Comput. Aided. Mol. Des.*, **19**, 453-463.

-
23. E-Dragon, (2005) VCCLAB. Available from: <http://www.vcclab.org>.
 24. SYBYL-X, Tripos, St. Louis, MO, USA
 25. *Poroikov V.V., Filimonov D.A., Borodina Y.V., Lagunin A.A., Kos A.* (2000) *J. Chem. Inf. Comput. Sci.*, **40**, 1349-1355.

Поступила: 09. 07. 2013.

MODE OF ACTION PREDICTION OF LIGANDS OF STEROID HORMONE RECEPTORS

I.V. Fedyushkina, I.V. Romero Reyes, A.A. Lagunin, V.S. Skvortsov

Orekhovich Institute of Biomedical Chemistry of Russian Academy of Medical Sciences,
Pogodinskaya ul., 10, Moscow, 119121 Russia; tel.: 499-245-0768; e-mail: ivf@ibmh.msk.ru

The several predictive models based on two well-known methods PASS and SIMCA were created. These models predict a type of physiological response of steroid compounds binding to nuclear receptors of steroid hormones. We considered 10 variants: the agonists and the antagonists of estrogen, progesterone, androgen, glucocorticoid and mineralocorticoid receptors respectively. Two different sets of descriptors were used during SIMCA (the Dragon descriptors and indices of similarity). The results of discriminant analysis are good enough with average accuracy of 80-85%.

Key words: steroid hormone receptors, predictive model, discriminant analysis, SIMCA.