

## БИОИНФОРМАТИКА

УДК 544.165

©Коллектив авторов

### МОЛЕКУЛЯРНЫЙ ДОКИНГ И 3D-QSAR ПРОИЗВОДНЫХ 16 $\alpha$ ,17 $\alpha$ -ЦИКЛОАЛКАНОПРОГЕСТЕРОНА КАК ЛИГАНДОВ РЕЦЕПТОРА ПРОГЕСТЕРОНА

*И.В. Федюшкина<sup>1\*</sup>, В.С. Скворцов<sup>1</sup>, И.В. Ромеро Рейес<sup>1</sup>, И.С. Левина<sup>2</sup>*

<sup>1</sup>Федеральное государственное бюджетное учреждение  
“Институт биомедицинской химии” им. В.Н. Ореховича

Российской академии медицинских наук, 119121 Москва, ул. Погодинская, 10;  
тел.: +7-499-245-0768; факс: +7-499-245-0857; эл. почта: ivf@ibmh.msk.su

<sup>2</sup>Федеральное государственное бюджетное учреждение науки  
Институт органической химии им. Н.Д. Зелинского Российской академии наук,  
119991 Москва, Ленинский просп., д.47

По данным о 42 стероидах ряда прегна-D'-пентаранов был построен набор моделей, предсказывающих аффинность лигандов к рецептору прогестерона. Наилучший результат ( $Q^2=0,91$ ) был получен при использовании комбинации методов молекулярного докинга, симуляции молекулярной динамики и искусственных нейронных сетей. Хорошая предсказательная сила этой модели была подтверждена на группе из 8 синтезированных отдельно и проверенных *in vitro* пентаранов ( $R^2$  тестирования равно 0,77). Эту модель можно использовать для определения степени сродства лиганда к рецептору и точного ранжирования соединений по связыванию.

**Ключевые слова:** пентараны, рецептор прогестерона, лиганд-связывающий домен, аффинность, QSAR, вычислительные методы, COMFA, COMSIA.

**ВВЕДЕНИЕ.** Существуют реальные предпосылки для совершенствования стероидных аналогов прогестерона, которые находят широкое применение в контрацепции, заместительной терапии и как противоопухолевые препараты [1-3]. Во-первых, к настоящему времени получено и испытано довольно много соединений этого типа, что позволяет рассчитывать на выявление количественных корреляций между структурой и биологической активностью (QSAR). И такие работы на отдельных гомологичных рядах прогестиннов успешно проведены [4-6]. Во-вторых, в 1998 г. методом рентгеноструктурного анализа (РСА) была установлена пространственная структура комплекса прогестерона с лиганд-связывающим доменом (ЛС-доменом) рецептора прогестерона (РП) человека [7, 8], благодаря чему возникла возможность прогнозировать взаимодействие проектируемых аналогов с участком связывания вплоть до прямого моделирования комплексов белок – проектируемый лиганд. И, наконец, в-третьих, конформационная ригидность стероидных прогестинов упрощает сопоставление их молекул и позволяет рассчитывать на структурное сходство их комплексов с рецептором.

\* - адресат для переписки

Наглядным примером перспективности подходов, основанных на знании 3D-структуры рецептора, могут служить синтезированные в лаборатории химии стероидных соединений Института органической химии им. Н.Д. Зелинского РАН аналоги прогестерона, сочлененные в 16 $\alpha$ ,17 $\alpha$ -положениях с трёх-шестичленными карбоциклами (прегна-D'-пентараны) [9, 10]. Соединения пентаранового ряда имеют высокую селективную прогестагенную активность, один из них (мецигестон) предложен для лечения гормональной дисфункции у женщин. При связывании пентаранов с РП дополнительные циклы размещаются внутри полости лиганд-связывающего домена (ЛД), примыкающей к  $\alpha$ -области кольца D прогестерона, что способствует умножению лиганд-белковых контактов и росту энергии связывания. Очевидно, что оптимизация этого и других подобных дополнительных взаимодействий не только возможна на основе количественного анализа зависимостей между строением и активностью структурно-родственных соединений с применением QSAR технологий, но и представляется наиболее эффективной.

В настоящей работе представлены результаты такого анализа, совмещающего различные компьютерные методы, такие как 3D-QSAR и прямое конструирование и анализ структуры прогестин-белковых комплексов.

**МЕТОДИКА.** В настоящей работе анализ связи структура-активность применён к набору производных 16 $\alpha$ ,17 $\alpha$ -циклоалканопрогестерона (общая формула см. рисунок 1). Всего в качестве обучающей выборки в работе использованы данные о структуре 42 различных 16 $\alpha$ ,17 $\alpha$ -циклоалканопрогестеронов и данные по относительной конкурентной активности (ОКА). Синтез и биологическое тестирование этих пентаранов были выполнены ранее [9-12]. В таблице 1 представлены их структурные характеристики и величины ОКА. Значения относительной конкурентной активности в первом случае (ОКА 1) рассчитывали по соотношению  $K_d$  для прогестерона (I) и сравниваемого лиганда, измеренному по конкуренции с меченым тритием прогестероном. РП был получен из цитозоля матки крысы по описанной ранее методике [9, 10]. Значения для второго варианта (ОКА 2) были получены ранее относительно лиганд связывающего домена РП кролика. При создании моделей (уравнений) предсказания были использованы значения ОКА в виде десятичного логарифма. Сравнительный анализ последовательностей лиганд-связывающих доменов РП кролика, крысы и человека показал, что различий аминокислотных остатков очень немного, а при рассмотрении их положения на трёхмерной структуре белка легко увидеть, что все они локализуются в отдалении от места связывания прогестерона. Например, минимальное расстояние от прогестерона до аминокислотного остатка, различающегося в последовательностях ЛС-домена РП крысы и человека, составляет около 8 ангстрем (для остальных остатков значения лежат в промежутке от 10 до 30 ангстрем). Для аминокислотных последовательностей ЛС-домена РП крысы и кролика эти различия ещё меньше. Это даёт возможность предположить, что вклад аминокислотных остатков, различающихся у разных видов, в энергию связывания лиганда невелик. Тем более что для расчётов с использованием молекулярной механики 8-10 ангстрем это граница “горячей” зоны при использовании PME (Particle mesh Ewald) приближения [13], которое также было использовано в данной работе. Кроме того, сравнение величин ОКА пентаранов показывает высокую степень корреляций (рис. 2):  $R^2 = 0,69$  для пары крыса/человек и  $R^2 = 0,73$  для пары крыса/кролик [14]. Это позволяет при поиске корреляций использовать объединённые данные о связывании

изучаемых лигандов с РП крысы и кролика для прогнозирования взаимодействия проектируемых прогестинов с рецептором прогестерона человека. Данные были пересчитаны согласно уравнению, полученному при сравнении пересекающихся значений и представленному на рисунке 3. Во всех случаях, когда имелось значение ОКА 2 – использовалось оно. В результате диапазон значений  $\lg(\text{ОКА})$  в общей выборке из 42 лигандов составил от -5,0 до +0,35.

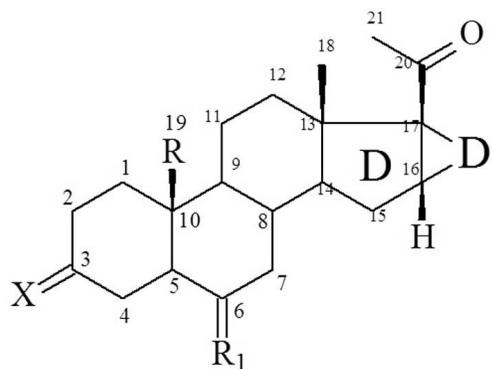


Рисунок 1.

Общая формула производных  $16\alpha,17\alpha$ -циклоалканопрогестерона (прегна-D'-пентанов).

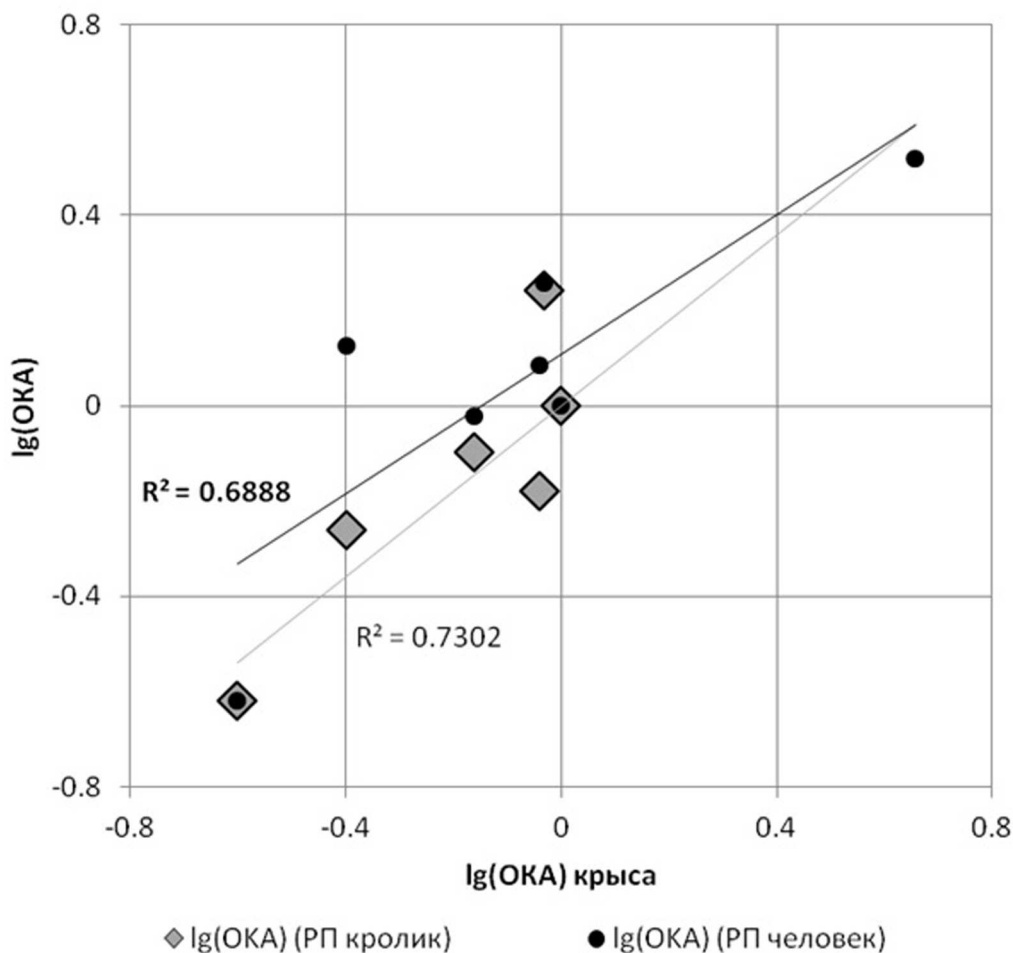


Рисунок 2.

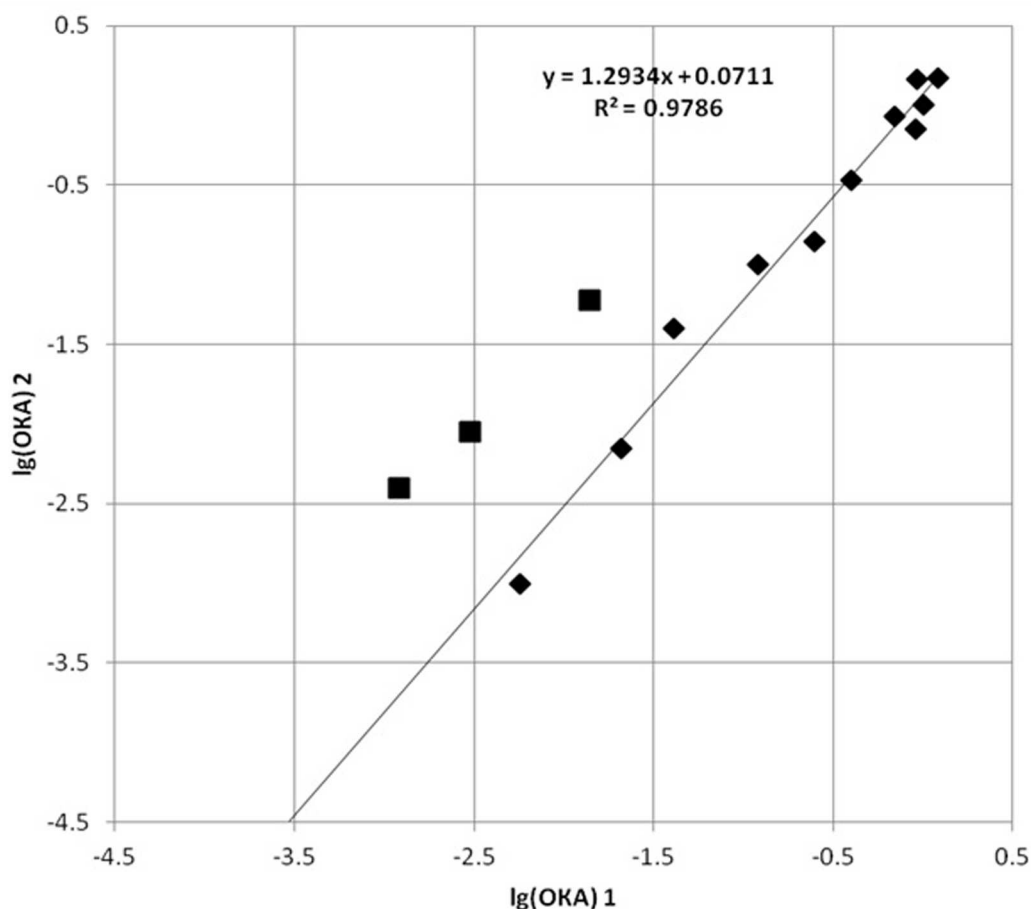
Сравнение значений относительной конкурентной активности набора лигандов для ЛС-домена прогестеронового рецептора крысы, кролика и человека [14].

Таблица 1. Структуры и значения affinности к РП прегна-D'-пентаранов обучающей выборки (см. рис.унок 1).

N	D' (16 $\alpha$ ,17 $\alpha$ )	X	Двойная связь	R	R <sub>1</sub>	ОКА изм.1	ОКА изм.2
I	-	O	4,5	Me	H <sub>2</sub>		1,00
II	-CH <sub>2</sub> -	O	4,5	Me	H <sub>2</sub>		0,93
III	-CH <sub>2</sub> -CH=CH-	O	4,5	Me	H <sub>2</sub>		1,21
IV	-(CH <sub>2</sub> ) <sub>3</sub> -	O	4,5	Me	H <sub>2</sub>		0,69
V	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	Me	H <sub>2</sub>		0,91
VI	-CH <sub>2</sub> -CH=CH-CH <sub>2</sub> -	O	4,5	Me	H <sub>2</sub>		0,40
VII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	-(5 $\delta$ H)	Me	H <sub>2</sub>		0,003
VIII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	-(5 $\delta$ H)	Me	H <sub>2</sub>		0,0012
IX	-CH <sub>2</sub> -CH=CH-CH <sub>2</sub> -	O	1,2	Me	H <sub>2</sub>		0,32
X	-CHMe- (CH <sub>2</sub> ) <sub>3</sub> -	O	4,5	Me	H <sub>2</sub>		0,021
XI	-CH <sub>2</sub> -CMe=CMe-CH <sub>2</sub> -	O	4,5	Me	H <sub>2</sub>		0,0057
XII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	Me	H, $\alpha$ Me		0,25
XIII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	1,2; 4,5	Me	H, $\alpha$ Me		0,076
XIV	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5; 6,7	Me	Me		0,12
XV	-CH <sub>2</sub> -	H, $\beta$ OH	5,6	Me	H <sub>2</sub>		0,027
XVI	-(CH <sub>2</sub> ) <sub>4</sub> -	H, $\beta$ OH	5,6	Me	H <sub>2</sub>		0,0015
XVII	-CH <sub>2</sub> -CH=CH-CH <sub>2</sub> -	H, $\beta$ OH	5,6	Me	H <sub>2</sub>		0,00018
XVIII	-CH <sub>2</sub> -CMe=CMe-CH <sub>2</sub> -	H, $\beta$ OH	5,6	Me	H <sub>2</sub>		0,0057
XIX	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CH <sub>2</sub> OH	H <sub>2</sub>		0,0138
XX	-(CH <sub>2</sub> ) <sub>4</sub> -	H, $\beta$ OH	5,6	CH <sub>2</sub> OH	H <sub>2</sub>		0,00021
XXI	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CHO	H <sub>2</sub>		0,122
XXII	-(CH <sub>2</sub> ) <sub>4</sub> -	H, $\beta$ OH	-	Me	O		0,003

Таблица 1. Структуры и значения аффинности к РП прегна-D'-пентаранов обучающей выборки (продолжение).

N	D' (16 $\alpha$ ,17 $\alpha$ )	X	Двойная связь	R	R <sub>1</sub>	ОКА изм.1	ОКА изм.2
XXIII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CH <sub>2</sub> -O-C(6)	H <sub>2</sub>		0,041
XXIV	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CH=NOMe	H <sub>2</sub>		1,21
XXV	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CH=NO(CH <sub>2</sub> ) <sub>3</sub> CO <sub>2</sub> Me	H <sub>2</sub>		0,0021
XXVI	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CH=CH <sub>2</sub>	H <sub>2</sub>		0,019
XXVII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CH <sub>2</sub> CH <sub>3</sub>	H <sub>2</sub>		0,027
XXVIII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	Me	(E) NO(CH <sub>2</sub> ) <sub>3</sub> CO <sub>2</sub> Me		0,00067
XXIX	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	Me	(Z) NO(CH <sub>2</sub> ) <sub>3</sub> CO <sub>2</sub> Me		0,0001
XXX	-CHMe-CH=CH-CH <sub>2</sub> -	O	4,5	Me	H <sub>2</sub>	0,004	
XXXI	-CH <sub>2</sub> -CH <sub>2</sub> -CO-CH <sub>2</sub> -	O	4,5	Me	H <sub>2</sub>	0,02	
XXXII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	1,2	Me	H <sub>2</sub>	0,63	
XXXIII	-(CH <sub>2</sub> ) <sub>4</sub> -	O	6,7	Me	H <sub>2</sub>	0,39	
XXXIV	-(CH <sub>2</sub> ) <sub>4</sub> -	O	1,2, - (5 $\beta$ H)	Me	H <sub>2</sub>	0,001	
XXXV	-(CH <sub>2</sub> ) <sub>4</sub> -	O	1,2 (-CH <sub>2</sub> -CH=CH-CH <sub>2</sub> -)	Me	H <sub>2</sub>	0,0001	
XXXVI	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	H	H <sub>2</sub>	1,52	
XXXVII	-CH <sub>2</sub> -CH=CH-CH <sub>2</sub> -	O	4,5	H	H <sub>2</sub>	0,86	
XXXVIII	-CH <sub>2</sub> -CH <sub>2</sub> -CH=CH-	O	4,5, -(18Me)	H	H <sub>2</sub>	2,17	
XXXIX	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	H	H <sub>2</sub>	0,78	
XL	-(CH <sub>2</sub> ) <sub>4</sub> -	-	4,5	H	H <sub>2</sub>	0,19	
XLI	-(CH <sub>2</sub> ) <sub>4</sub> -	-	-(5 $\alpha$ H)	H	H <sub>2</sub>	0,001	
XLII	-(CH <sub>2</sub> ) <sub>4</sub> -	-	-(5 $\alpha$ H)	H	H <sub>2</sub>	0,07	



**Рисунок 3.**

Сравнение значений относительной конкурентной активности ЛС-домена РП для крысы и кролика в обучающей выборке.

В работе использовали 2 группы основных методов: 3D-QSAR, включая CoMFA (сравнительный анализ молекулярных полей) или CoMSIA (сравнительный анализ молекулярных индексов), часто используемых при анализе QSAR [15], и процедуру, основанную на прямом моделировании комплексов лиганд/ЛС-домен.

Все молекулы прегна-D'-пентаранов были сконструированы на основе молекулы прогестерона, взятой в конфигурации, находящейся в кристаллической структуре. Учитывая, что в процессе работы всегда был этап оптимизации с использованием молекулярной механики, стартовую оптимизацию выполняли средствами программы Sybyl-X [15] с расстановкой частичных атомных зарядов по эмпирической схеме Gasteiger-Huckel, имеющейся в Sybyl-X.

Пространственное выравнивание молекул для 3D-QSAR анализа выполняли с использованием модуля Sybyl/GALAHAD [16]. Процедура проводилась с параметрами по умолчанию, как результат отбирались 20 лучших выравниваний, для каждого методом частичных наименьших квадратов на основании полей CoMFA и CoMSIA строились корреляционные уравнения предсказания аффинности. Лучшую модель рассматривали как конечный результат.

Моделирование комплексов осуществляли посредством докинга молекул потенциальных лигандов к участку связывания прогестерона ЛС-домена РП. Структура ЛС-домена РП была взята из файла 1A28 [7, 8] в Protein Data Bank. В работе использовали программу Dock 6.5 [17], допускающую конформационную подвижность лиганда при сохранении неизменной конформации белка. До 3-х вариантов виртуальных комплексов, отобранных по оценочной функции программы Dock, отбирали для дальнейшей оптимизации средствами программы Amber 9.0 [17] (поля сил AMBER99 и GAFF) по следующему алгоритму:

- Сольватация комплекса водой явным образом (модель TIP3P), слой воды до границ прямоугольного бокса не менее 4 ангстрем.
- Минимизация общей энергии системы до 1000 шагов.
- Разогрев системы от 0K до 300K - симуляция молекулярной динамики – 10 пс с шагом в 2 фс в периодических граничных условиях (NVT ансамбль) с заторможенной структурой белка.
- Выравнивание плотности системы при 300K - симуляция молекулярной динамики – 10 пс с шагом в 2 фс в периодических граничных условиях (NPT ансамбль) с заторможенной структурой белка.
- Уравновешивание системы при 300K – симуляция молекулярной динамики – 10 пс с шагом в 2 фс в периодических граничных условиях (NPT ансамбль).
- Симуляция молекулярной динамики (продуктивная динамика) при 300K, 10 пс с шагом в 2 фс в периодических граничных условиях (NPT ансамбль).

Продуктивная динамика служила для оптимизации структуры и для подготовки промежуточных данных для оценки методом ММ-PBSA виртуальной аффинности ЛС-домена к лигандам [18]. Усреднение производили по 10 наблюдениям, сохранённым через равные промежутки времени симуляции. Сольватационную составляющую вычисляли методами Poisson-Boltzmann (PB) и Generalized Born (GB). Отбор конечного варианта производили по минимальному значению изменения энергии комплекса, рассчитанному по методу ММ-PBSA.

Параллельно рассчитывали отдельные параметры для построения корреляционных уравнений: изменение величины электростатического взаимодействия (ELE); изменение величины ван-дер-ваальсовых взаимодействий (VDW); гидрофобный вклад в изменение свободной энергии, рассчитанной методом PB (PBSUR); вклад сольватации в изменение свободной энергии, рассчитанной методом PB (PBCAL); аналогичные двум последним параметрам величины, рассчитанные методом GB (GBSUR и GB). Также модулем NMODE [19] вычисляли значения трансляционного, ротационного и колебательного энтропийных вкладов. Для расчёта корреляционных уравнений для каждого лиганда были добавлены также его собственные характеристики: молекулярный вес, площадь поверхности, площадь полярной поверхности, полярный и общий объёмы молекулы.

Следует отметить, что для построения корреляционных моделей кроме классической линейной регрессии применялся также метод искусственных нейронных сетей (ИНС) [20]. С целью снижения ошибки переобучения ИНС начальная выборка была увеличена в 4 раза путём генерации новых точек в пределах погрешности (статистический бутстрап).

**РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.** На первом этапе работы было получено 20 пространственных выравниваний для 42 молекул, для которых были рассчитаны CoMFA-поля (стерическое, электростатическое, донорное и акцепторное) и CoMSIA-поля (четыре аналогичных и гидрофобное поле),

#### МОЛЕКУЛЯРНЫЙ ДОКИНГ ЛИГАНДОВ РЕЦЕПТОРА ПРОГЕСТЕРОНА

описывающие молекулы, выравненные в пространстве. Идеологически эти методы очень похожи, хотя имеются различия в способе расчёта каждого поля. Наилучший вариант выравнивания представлен на рисунке 4, соответствующие ему параметры качества моделей представлены в таблице 2. В таблице представлено 2 набора моделей, рассчитанных для полной выборки и с учётом так называемых “выбросов”. Наличие “выбросов” очень характерно для методов CoMFA и CoMSIA, так как оценка качества моделей ведётся по процедуре скользящего контроля. В этом случае наличие молекул, имеющих уникальные боковые радикалы или неудачно выравненных, может давать существенные флуктуации. Как правило, их отбрасывают. Данное свойство очень ограничивает применимость методов CoMFA и CoMSIA. Из данных, приведённых в таблице и представленных на рисунке 5, следует, что, несмотря на то, что формальный порог успешности модели ( $Q^2 > 0,6$ ) практически достигнут, предсказательная сила модели невелика, а средняя ошибка в процедуре скользящего контроля больше 1 (около 20% от всего диапазона значений).



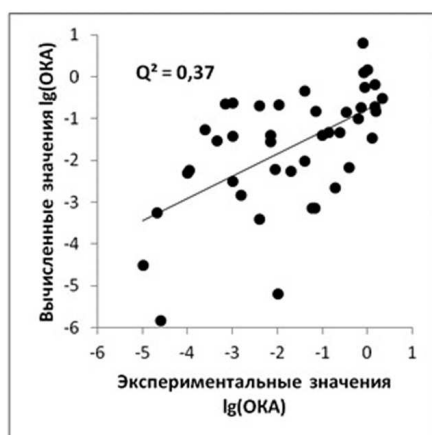
**Рисунок 4.**

Пространственное выравнивание прегна-D'-пентранов, использованное в 3D-QSAR.

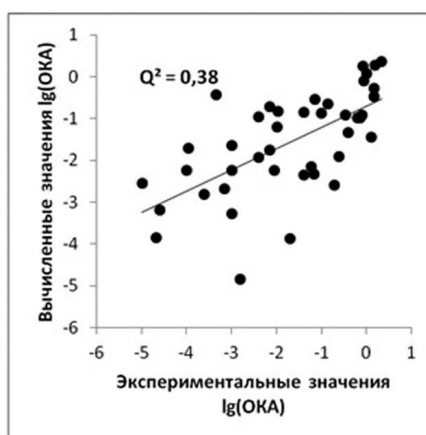


Таблица 2. Статистические параметры 3D-QSAR моделей.

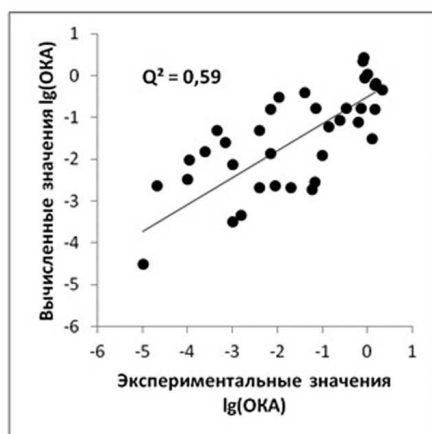
Использованный метод	Полная выборка		- 10% выбросов	
	CoMFA	CoMSIA	CoMFA	CoMSIA
Число соединений	42	42	37	37
$Q^2$ в процедуре скользящего контроля методом выкидывания по одному (loo)	0,37	0,38	0,59	0,57
Стандартная ошибка (SE) в loo	1,28	1,30	1,08	1,12
Оптимальное число компонент (n)	4	6	4	6
$R^2$ обучения	0,82	0,92	0,91	0,93
Стандартная ошибка (SE) обучения	0,68	0,47	0,52	0,41
Критерий Фишера	42,44	66,68	76,18	77,33
Относительный вклад стерического поля	0,24	0,11	0,26	0,13
Относительный вклад электростатического поля	0,24	0,19	0,24	0,22
Относительный вклад «гидрофобного» поля		0,18		0,19
Относительный вклад поля акцепторов водорода	0,28	0,31	0,34	0,29
Относительный вклад поля доноров водорода	0,24	0,21	0,17	0,19



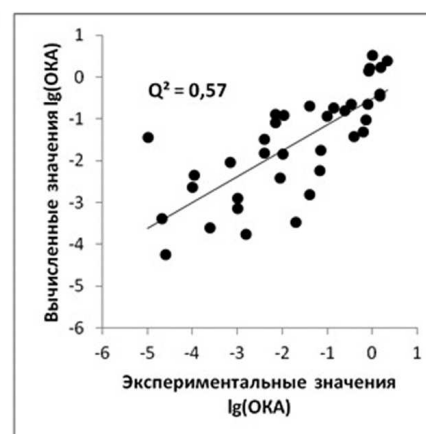
CoMFA (42 соединения)



CoMSIA (42 соединения)



CoMFA (37 соединений)



CoMSIA (37 соединений)

Рисунок 5.

Сравнение экспериментальных и вычисленных 3D-QSAR величин  $\lg(\text{OKA})$  в процедуре LVO (метод скользящего контроля).

## МОЛЕКУЛЯРНЫЙ ДОКИНГ ЛИГАНДОВ РЕЦЕПТОРА ПРОГЕСТЕРОНА

Следующим этапом для создания предсказательной модели была попытка построить линейную регрессионную модель, используя параметры комплексов лиганд/ЛС-домен (расчёты методом ММ-PBSA) и параметры, описывающие свойства лигандов, в качестве независимых переменных. Результат продемонстрирован на рисунке 6. Соединение XXIX было отброшено, так как процедура докирования не смогла найти решение, что вероятно связано с размерами заместителя  $R_1$ . Таким образом, в последующих рассуждениях выборка содержит 41 соединение.  $R^2$  в процедуре был равен 0,6, при процедуре скользящего контроля всего 0,41. Впрочем, визуально видно, что, несмотря на существенный разброс, имеет место наличие определённой зависимости.

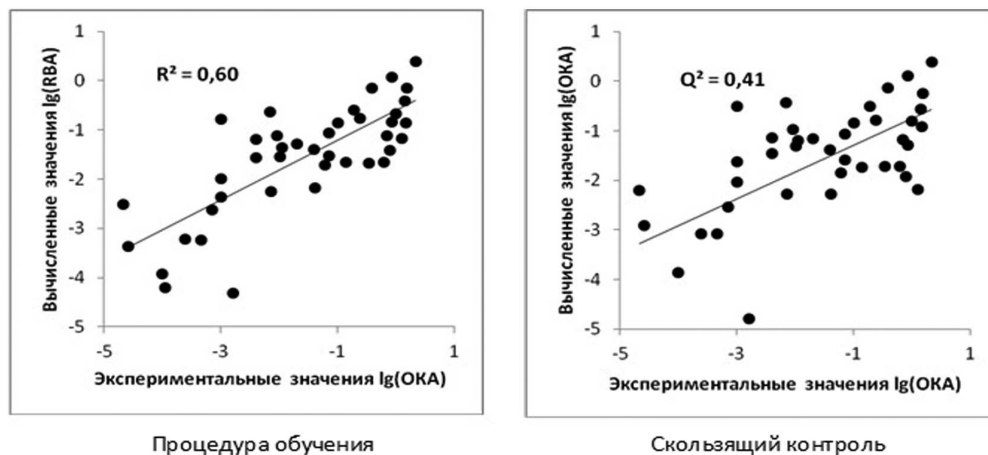


Рисунок 6.

Линейная модель предсказания  $\lg(RBA)$  лигандов РП (обучение и процедура скользящего контроля).

Третьим этапом являлось построение модели с использованием ИНС. Результаты представлены на рисунке 7. Левая часть рисунка характеризует процедуру обучения нейронной сети. Отсутствие существенного увеличения среднеквадратичной ошибки характеризует отсутствие переобучения сети, наиболее частой проблемы для ИНС. Рабочая модель выбрана из зоны, выделенной кругом. Характеристики модели представлены в таблице 3. Из всех построенных моделей ИНС даёт самые лучшие результаты ( $Q^2=0,91$ ), а средняя ошибка предсказания при скользящем контроле в 4 раза меньше чем в лучшем варианте моделей 3D-QSAR.

Таблица 3. Статистические параметры нейросетевой модели.

Этап	Параметер	Значение
	Число элементов в общей выборке*	164
Настройка сети	$R^2$ для обучающей выборки	0,98
	$R^2$ для контрольной выборки	0,97
	$R^2$ для тестовой выборки	0,96
	$R^2$ для общей выборки	0,98
Скользящий контроль	$Q^2$	0,91
	Среднеквадратичная ошибка	0,29
	Среднее отклонение	0,24
	$R^2$ для обучающей выборки	0,32

Примечание: \* - общая выборка = обучающая выборка (70%) + контрольная выборка (15%) + тестовая выборка (15%).

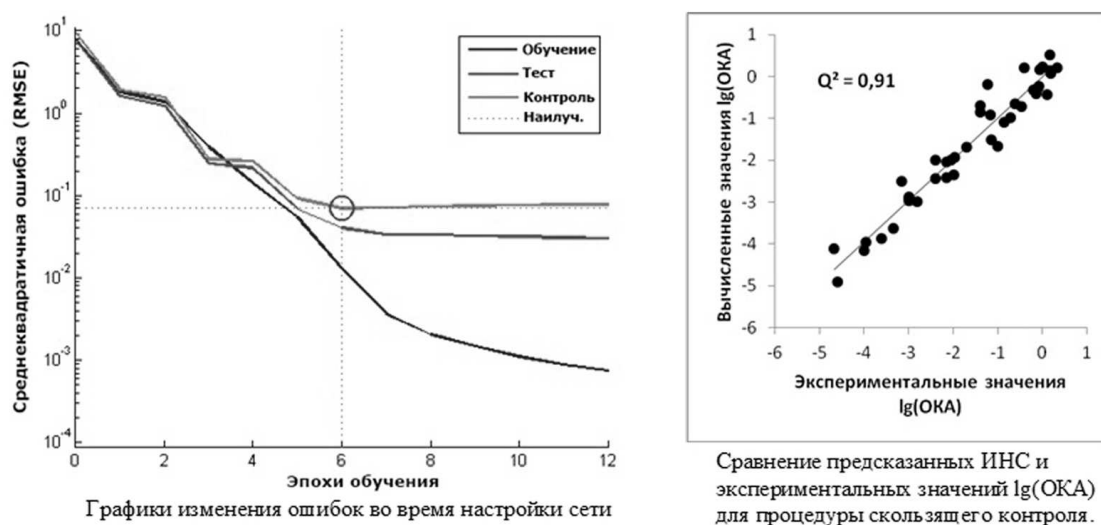


Рисунок 7.

Нейросетевая модель предсказания  $\lg(\text{RBA})$  лигандов РП  
(обучение и результаты предсказания в процедуре скользящего контроля).

Для оценки предсказательной способности полученных моделей нами была использована тестовая выборка синтезированных и изученных *in vitro* стероидов ряда пентаранов (табл. 4) [21-23]. Оказалось, что при сравнении с экспериментальными данными (рис. 8) рассчитанные ОКА в случае ИНС дают  $R^2$  тестирования равный 0,77, а для 3D-QSAR и линейной модели эта величина составляет всего 0,39 и 0,37. даже они способны отличить лиганды с высоким сродством от лигандов с низким сродством. Нейронная же сеть даёт очень хороший результат ( $R^2=0,77$ ).

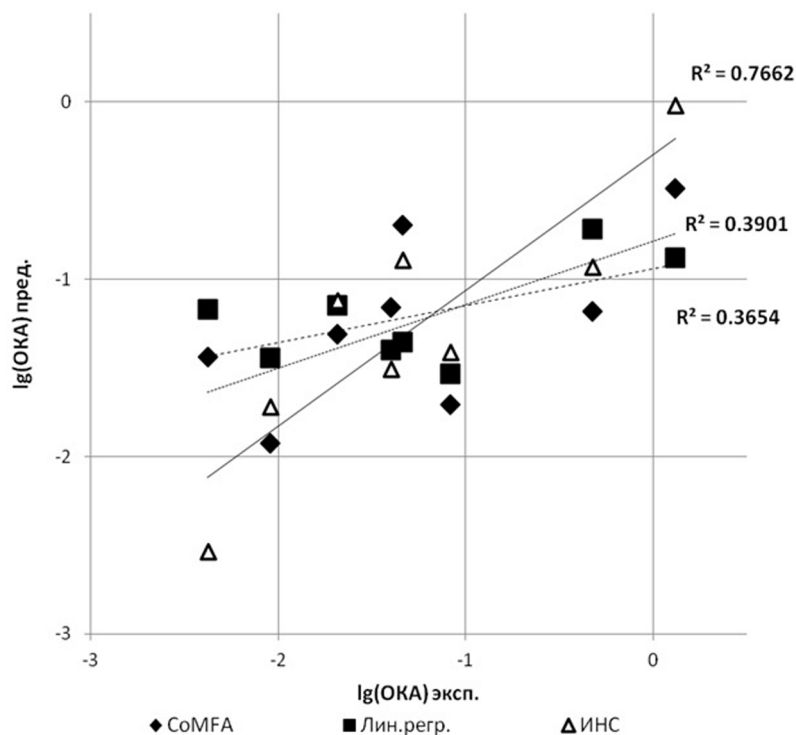




Рисунок 8.

Сравнение экспериментально измеренных и рассчитанных значений аффинности для тестового набора молекул прегна-D'-пентаранов.

Таблица 4. Тестовый набор структур (см. рисунок 1) и значения аффинности к РП прегна-D'-пентранов. Сравнение с предсказанными величинами (ОКА протгестерона принята за 100%).

N	D' (16 $\alpha$ , 17 $\alpha$ )	X	Двойная связь	R	R <sub>1</sub>	ОКА эксперт	ОКА CoMFA	ОКА Лин. модель	ОКА ИНС
<b>XLIII</b>		O	4,5	Me	H <sub>2</sub>	122	32,48	13,22	96,12
<b>XLIV</b>		O	4,5	Me	H <sub>2</sub>	3,7	20,28	4,43	12,81
<b>XLV</b>	-(CH <sub>2</sub> ) <sub>4</sub> -	(E) NOH	4,5	Me	H <sub>2</sub>	7,7	1,98	2,92	3,90
<b>XLVI</b>	-(CH <sub>2</sub> ) <sub>4</sub> -	(Z) NOH	4,5	Me	H <sub>2</sub>	3,7	6,92	4,00	3,12
<b>XLVII</b>	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	(Z) NOH	H <sub>2</sub>	1,93	4,90	7,09	7,55
<b>XLIX</b>	-(CH <sub>2</sub> ) <sub>4</sub> -	(E) NOMe	4,5	Me	H <sub>2</sub>	0,39	3,63	6,75	0,29
<b>CI</b>	-(CH <sub>2</sub> ) <sub>4</sub> -	O	4,5	CN	H <sub>2</sub>	43,8	6,58	19,20	11,80
<b>CX</b>	-(CH <sub>2</sub> ) <sub>4</sub> -	O	-(4N); 5,6	Me	H	0,84	1,18	3,58	1,91

**ЗАКЛЮЧЕНИЕ.** Таким образом, в итоге работы были получены 2 предсказательные модели. Первая, созданная с использованием методов 3D-QSAR, может быть использована для разделения производных 16 $\alpha$ ,17 $\alpha$ -циклоалканопрогестеронов на высоко- и низкоаффинные группы. Вторая, использующая комбинацию методов молекулярного докинга, симуляции молекулярной динамики и ИНС, может быть использована для определения степени сродства лиганда к рецептору прогестерона и точного ранжирования лигандов по связыванию в ряду соединений. Разработанные модели могут быть использованы для отбора новых стероидов с заданной аффинностью к РП при направленном конструировании и/или поиске новых лекарственных соединений.

## ЛИТЕРАТУРА

1. Schindler A.E., Campagnoli C., Druckmann R., Huber J., Pasqualini J.R., Schweppe K.W., Thijssen J.H. (2003) *Maturitas*, **46 Suppl 1**, S7-S16.
2. de Ziegler D., Fanchin R. (2000) *Steroids*, **65**, 671-679.
3. Lundgren S. (1992) *Acta Oncol.*, **31**, 709-722.
4. Bursi R., Groen M.B. (2000) *Eur. J. Med. Chem.*, **35**, 787-796.
5. So S.S., van Helden S.P., van Geerestein V.J., Karplus M. (2000) *J. Chem. Inf. Comput. Sci.*, **40**, 762-772.
6. Hillisch A., von Langen J., Menzenbach B., Driescher P., Kaufmann G., Schneider B., Elger W. (2003) *Steroids*, **68**, 869-878.
7. Williams S.P., Sigler P.B. (1998) *Nature*, **393**, 392-396.
8. Berman H., Henrick K., Nakamura H. (2003) *Nat. Struct. Biol.*, **10**, 980.
9. Камерницкий А.В., Левина И.С. (2005) *Биоорган. химия*, **31**, 227-238.
10. Камерницкий А.В., Левина И.С. (2005) *Биоорган. химия*, **31**, 115-129.
11. Smirnov A.N., Pokrovskaya E.V., Kogteva G.S., Shevchenko V.P., Levina I.S., Kulikova L.E., Kamernitzky A.V. (2000) *Steroids*, **65**, 163-170.
12. Левина И.С., Куликова Л.Е., Камерницкий А.В., Покровская Е.В., Смирнов А.Н. (2005) *Изв. АН. Сер. хим.*, **11**, 2579-2584.
13. Darden T., Perera L., Li L., Pedersen L. (1999) *Structure*, **7**, R55-R60.
14. Смирнов А.Н., Покровская Е.В., Шевченко В.П., Нагаев И.Ю., Мясоедов Н.Ф., Левина И.С., Куликова Л.Е., Камерницкий А.В. (2002) *Биоорган. химия*, **28**, 251-257.
15. SYBYL-X, Tripos, St. Louis, MO, USA
16. Richmond N.J., Abrams C.A., Wolohan P.R., Abrahamian E., Willett P., Clark R.D. (2006) *J. computer-aided molec. design*, **20**, 567-587.
17. Kuntz I.D., Blaney J.M., Oatley S.J., Langridge R., Ferrin T.E. (1982) *J. Mol. Biol.*, **161**, 269-288.
18. Kollman P.A., Massova I., Reyes C., Kuhn B., Huo S., Chong L., Lee M., Lee T., Duan Y., Wang W., Donini O., Cieplak P., Srinivasan J., Case D.A., Cheatham T.E. 3rd (2000) *Acc. Chem. Res.*, **33**, 889-897.
19. Case D.A., Darden T., Cheatham III T.E., Simmerling C., Wang J., Duke R.E., Luo R., Merz K.M., Pearlman D.A., Crowley M. (2006) University of California, San Francisco.
20. Romero Reyes I., Fedyushkina I., Skvortsov V., Filimonov D. (2013) *Intern. J. Math. Models and Methods appl. sci.*, **7**, 303-310.
21. Левина И.С.; Куликова Л.Е., Смирнов А.Н., Шимановский Н.Л., Семейкин А.В., Карева Е.Н., Федотчева Т.А., Болотова Е.Н. (2011) пат. 2426737, Бюл. № 23.

22. Левина И.С., Куликова Л.Е., Шулишов Е.В., Томилов Ю.В., Смирнов А.Н. (2013) Изв. Ак. Наук. Сер. хим., **6**, 1449-1453.
23. Levina I.S., Pokrovskaya E.V., Kulikova L.E., Kamernitzky A.V., Kachala V.V., Smirnov A.N. (2008) Steroids, **73**, 815-827.

Поступила: 15. 07. 2013.

MOLECULAR DOCKING AND 3D-QSAR  
ON 16 $\alpha$ ,17 $\alpha$ -CYCLOALKANOPROGESTERONE ANALOGUES  
AS PROGESTERONE RECEPTOR LIGANDS

*I.V. Fedyushkina<sup>1</sup>, V.S. Skvortsov<sup>1</sup>, I.V. Romero Reyes<sup>1</sup>, I.S. Levina<sup>2</sup>*

<sup>1</sup>Orekhovich Institute of Biomedical Chemistry of Russian Academy of Medical Sciences,  
Pogodinskaya ul., 10, Moscow, 119121 Russia; tel.: 499-245-0768; e-mail: ivf@ibmc.msk.ru

<sup>2</sup>N.D. Zelinsky Institute of Organic Chemistry of Russian Academy of Sciences,  
Leninsky prosp., 47, 119991 Moscow, Russia

A series of 42 steroid ligands was used to predict a binding affinity to progesterone receptor. The molecules were the derivatives of 16 $\alpha$ ,17 $\alpha$ -cycloalkanoprogesterones. Different methods of prediction were used and analyzed such as CoMFA and artificial neural networks. The best result ( $Q^2=0.91$ ) was obtained for a combination of molecular docking, molecular dynamics simulation and artificial neural networks. A predictive power of the model was validated by a group of 8 pentarans synthesized separately and tested in vitro ( $R^2_{test}=0.77$ ). This model can be used to determine the affinity level of the ligand to progesterone receptor and accurate ranking of binding compounds.

**Key words:** pentaranes, progesterone receptor, ligand-binding domain, affinity, QSAR, computational methods, COMFA, COMSIA.