

© КОЛЛЕКТИВ АВТОРОВ, 1997  
УДК

### УЧЕТ КОНФОРМАЦИОННОЙ ГИБКОСТИ МОЛЕКУЛ В ИССЛЕДОВАНИЯХ ЗАВИСИМОСТЕЙ СТРУКТУРА-СВОЙСТВО

О. Г. МЕКЕНЯН\*, А. Е. ТАРХОВ\*\*, В. И. ШВЕЦ\*\*

\* Университет «Ас. Златаров», 8010, Бургас, Болгария

\*\* Московская Государственная Академия Тонкой Химической Технологии им. М.В.Ломоносова,  
177571, Москва, пр. Вернадского, 86

Описаны методы исследования зависимостей структура-свойство, учитывающие конформационную гибкость молекул. Исчерпывающе анализируя конформационную гибкость молекул, методы позволяют определить: активные соединения для специфического биологического взаимодействия, сходство между трехмерными структурами молекул и общие образы реакционной способности. Приведен пример применения метода определения общего образа реакционной способности для определения структурных требований, предъявляемых к молекулам ретиналя и его аналогов, необходимых для их успешного встраивания в молекулу бактериоопсина. Найденный образ общей реакционной способности, выраженный через совокупность интервалов варьирования стереоэлектронных параметров, соответствует предполагаемому механизму взаимодействия.

**Ключевые слова:** Конформационная гибкость, QSAR, выбор активного конформера, OASIS-COREPA, ретиноиды.

**ВВЕДЕНИЕ.** Большинство методов анализа зависимостей «структура-свойство», учитывающих трехмерную геометрию молекул, обычно основаны на представлении исследуемого вещества наиболее энергетически выгодной структурой. Однако встречается множество примеров построения моделей активности различных веществ, основанных не на энергетически самых выгодных конформерах [1]. Это особенно часто справедливо для биологически активных веществ, так как в сложной реакционной среде в биологических системах естественно ожидать, что молекула может существовать в форме того или иного конформера на разных этапах взаимодействия, например, при проникновении через мембрану или образовании комплекса лиганд-рецептор. Более того, специфичность процесса заставляет разные конформеры быть активными для проявления разных свойств [2]. Идентификация этих конформеров является сложной задачей, особенно если они могут легко взаимопревращаться. Энергия таких процессов, как сольватация, химическое связывание или другие «внешние» процессы в сложной реакционной среде, может компенсировать энергетические затраты превращения энергетически наиболее выгодного конформера в другие, энергетически менее выгодные, но являющиеся активными в данном случае. С другой стороны, известно, что у энергетически близких конформеров молекул данного соединения может наблюдаться большой интервал варьирования стереоэлектронных индексов [2]. Отсюда ясно, что выбор активного конформера является столь же важным для анализа зависимостей структура-свойство, как и отбор подходящих молекулярных дескрипторов. Это можно проиллюстрировать на примере распределения конформеров гидроксифутамида по значениям энергии НСМО - потенциала ионизации (рассчитанного

с помощью программы МОРАС 7 с использованием гамильтониана AM1). Потенциал ионизации варьируется в интервале более чем 1эВ (рис. 1) для всех рассматриваемых конформеров в интервале теплот образования  $|\Delta H_f| \leq 20$  ккал/моль.

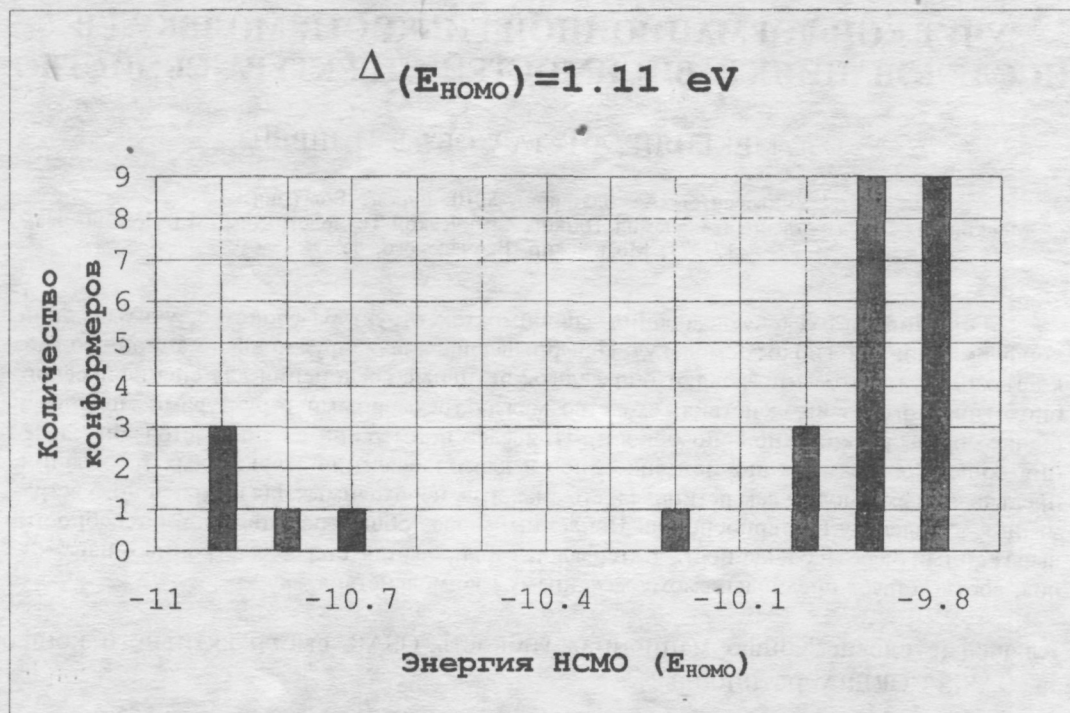


Рис. 1. Распределение конформеров гидроксифлутамида по энергии НСМО.

Далее рассмотрены три группы методов исследования зависимостей структура-свойство с учетом конформационной гибкости молекул и приведен пример определения образа общей реакционной способности для моделирования способности ретиналя и его аналогов встраиваться в бактериоопсин [3].

#### Выбор активного конформера для изучения зависимостей структура-свойство.

Недавно предложенный метод отбора активных конформеров [4] был назван "динамическим", так как он пытается имитировать множественность конформеров, способных участвовать в различных этапах реакции, с использованием набора "статических" конформеров. Реализованный в компьютерной системе OASIS [4], метод комбинирует генерацию конформеров с последующим скринингом конформеров. Сначала генерируется исчерпывающее множество конформаций [5], которое затем сокращается за счет применения иерархического набора экспертных правил отбора сгенерированных конформеров в соответствии с химической экспертизой [6]. Стереoeлектронная структура каждого конформера из исходного набора оценивается с помощью квантово-химических методов, обычно используется программа МОРАС в комбинации со специфическими модулями OASIS'а для вычисления стереoeлектронных индексов [4]. Затем на основании рабочих гипотез можно отобрать различные наборы конформеров.

Система OASIS включает топологические, такие как индекс Винера, индекс Рандича, индекс Балабяна и некоторые другие, физико-химические и стерические индексы, учитывающие объемные геометрические характеристики, такие как Ван-дер-Ваальсов объем и поверхность молекулы, максимальное межатомное расстояние, суммы межатомных расстояний, а также стереоэлектронные индексы, отражающие пространственную геометрию и локальные электронные характеристики атомов молекулы. Сюда входят глобальные молекулярные характеристики, такие как теплота образования, общая электронная энергия, энергии граничных орбиталей, разница энергий НСМО-ВЗМО\*, дипольный момент, объемная поляризуемость, и локальные (атомные) параметры - атомные заряды, порядки связей, атомные поляризуемости, донорные и акцепторные суперделокализуемости и т.п. Этот набор индексов позволяет учитывать как топологические особенности молекулы, так и ее электронные свойства в целом и свойства отдельных атомов молекулы.

Конформеры могут быть отобраны на основании их частотного распределения по специфическим геометрическим и/или стереоэлектронным параметрам. Весь интервал варьирования параметра разделяется на определенное число равных отрезков и количество конформеров, попадающих внутрь каждого отрезка, дает высоту столбца в получаемой гистограмме (рис. 1). Существует несколько схем селекции, которые могут применяться для отбора конформеров:

1. По всему интервалу варьирования - из каждого отрезка отбирается по одному конформеру
2. По наиболее заселенному отрезку - выбираются все конформеры из наиболее заселенного отрезка
3. По экстремальным значениям - выбираются конформеры из отрезка с минимальными (максимальными) значениями исследуемого параметра
4. Конформеры, заселяющие определенный энергетический уровень - в соответствии с распределением Больцмана отбираются конформеры, заселяющие определенный энергетический уровень
5. Конформеры со «взвешенными» значениями исследуемых параметров. Структурные дескрипторы для одного соединения могут «взвешиваться». При этом значения исследуемого параметра умножаются на весовые коэффициенты конформеров, равные заселенностям энергетических уровней, соответствующих этим конформерам.

Возможность приложения динамического подхода для выбора активного конформера можно проиллюстрировать моделированием QSAR взаимодействия 14 полихлорбифенилов (ПХБ) с арилуглеводородным рецептором (AhR) [7]. Система OASIS сгенерировала 103 конформера для 14 исследуемых соединений. Так как в эксперименте было обнаружено, что при взаимодействии ПХБ с рецептором получались комплексы с переносом заряда со смещением заряда в сторону ПХБ, была выдвинута гипотеза о возможной корреляции сродства ПХБ к рецептору с энергией граничных орбиталей, в особенности с энергией НСМО. Последняя была выбрана как молекулярный дескриптор.

---

\* НСМО - низшая свободная молекулярная орбиталь, ВЗМО - высшая занятая молекулярная орбиталь



Геометрии всех конформеров оптимизировались с использованием гамильтониана PM3. При включении всех конформеров в корреляционный анализ получаемая модель имела коэффициент корреляции  $r^2=0.38$ . Модель также не сильно улучшалась, если выбирать только энергетически наиболее выгодные конформеры:  $r^2=0.43$ . Однако если отбирались самые планарные конформеры, коэффициент корреляции возрастал до  $r^2=0.72$ . Аналогичные результаты получались и без оптимизации геометрии конформеров:  $r^2=0.78$ . Таким образом, отбор наиболее планарных конформеров позволяет получить модель, наилучшим образом соответствующую экспериментальным данным.

#### **Оценка подобия соединений с учетом конформационной гибкости молекул.**

Существуют методы, оценивающие сходство молекул на основании «плоских» формул без учета их пространственного строения, см. например [8-10]. Однако в последнее время много внимания уделяется анализу подобия трехмерных (3D) структур молекул на основании различных рейтингов или метрик подобия. В некоторых из этих методов подобие определяется путем наложения трехмерных структур друг на друга. В этом случае рейтингами подобия являются такие критерии, как сумма квадратов расстояний между парами соответствующих атомов разных молекул [11]. Сюда же относятся методы, сравнивающие квантово-механические характеристики молекул, такие как перекрывание электронных облаков, поверхностей потенциальных энергий и т.п. [12]. К этой же группе методов можно отнести и CoMFA [13], в котором в качестве рассматриваемой характеристики анализируемых молекул используются стерические и электростатические поля наложенных друг на друга трехмерных структур.

В других методах сравниваются трехмерные структуры молекул на основании спектров межатомных расстояний [14]. В этом случае для определения схожести молекул сравниваются частотные распределения всех внутримолекулярных межатомных расстояний для исследуемой серии соединений.

Наконец есть методы, реализующие алгоритмы поиска сходства на основании сопоставления свойств атомов в молекулах [15].

Существующие методы анализа трехмерных структур, однако, обычно сравнивают структуры с жесткой геометрией. Вычисляемые таким образом меры сходства не учитывают конформационной гибкости молекул. Так, два конформера сравниваемых соединений могут сильно отличаться по своей стереоэлектронной структуре, выражаемой через молекулярные дескрипторы, в то время как другие два конформера этих же соединений могут иметь очень близкие значения тех же самых дескрипторов. Кроме того, при наложении трехмерных структур возникает проблема выбора реперных точек для наложения (alignment) [16]. Существуют попытки учитывать конформационную гибкость молекул, например в [17] рассматривается метод кодирования конформационной гибкости, используемый при поиске новых базовых структур для лекарственных соединений. Недавно был предложен еще один метод - так называемого «динамического подобия» [2], который учитывает конформационное разнообразие сравниваемых соединений. На основании предположений о влиянии того или иного стереоэлектронного параметра на анализируемую активность, отбираются параметры для проверки подобия. Далее значения этих параметров определяются для всех конформеров сравниваемой пары соединений. Получающиеся интервалы варьирования параметров затем делятся на равные отрезки и строятся частотные распределения конформеров по каждому параметру. За частоту принимается число конформеров, попадающих внутрь данного отрезка всего интервала. В случаях, когда параметр является атомным, или



локальным стереоэлектронным параметром, таким как атомный заряд, поляризуемость, суперделокализуемость и т.д., частотное распределение основывается на числе атомов всех конформеров, заселяющих отрезок. Межатомные расстояния также считаются локальными индексами и в этом случае распределение базируется на числе внутримолекулярных расстояний конформеров внутри определенного интервала межатомных дистанций.

Сходство оценивается путем сравнения конформационных распределений по отобранным параметрам. При наложении распределений конформеров двух соединений друг на друга в каждом конкретном отрезке могут оказаться конформеры только одного или другого соединения или обоих соединений. Различия в частотных распределениях для двух молекул могут быть оценены с помощью различных коэффициентов - рейтингов или мер подобия, таких как коэффициент Танимото, коэффициент Ходжкинга-Ричардса и других [2], которые и служат мерой подобия сравниваемых соединений.

#### **Определение общих образов реакционной способности соединений.**

Метод определения общих образов реакционной способности (COREPA - COmmon REactivity Pattern) основан на предположении, что вещества, проявляющие сходную активность (биологическую активность, токсичность и т.п.) по одному и тому же механизму воздействия должны иметь общность в строении и электронных свойствах, выражаемую через стереоэлектронные параметры.

Поиск общего образа реакционной способности состоит из трех этапов:

1. Определенная группа из всех исследуемых соединений отделяется как обучающая. Эта группа может включать или наиболее активные, или наименее активные вещества. Отбор наиболее и наименее активных производится путем введения порогового значения активности.

2. Отбор стереоэлектронных параметров, предположительно связанных с моделируемой активностью, определяется на основании рейтингов подобия между каждой парой молекул обучающего набора. Стереоэлектронный параметр, получающий максимальные значения рейтингов подобия для отобранных соединений, считается связанным с исследуемой активностью и анализируется дальше.

3. Для каждого стереоэлектронного параметра, выявленного на предыдущем этапе, строится частотное распределение конформеров отобранных соединений и эти распределения накладываются друг на друга с целью выявить такие интервалы перекрывания распределений, в которых оказываются по крайней мере по одному конформеру от каждого соединения. Совокупность таких общих интервалов и определяет образ общей реакционной способности самых активных (неактивных) соединений.

#### **Пример моделирования способности ретиналя и его аналогов встраиваться в бактериоопсин**

Ретиналь входит в состав природных фоточувствительных мембранных белков, таких как родопсин и бактериородопсин. Эти белки являются протонными насосами, управляемыми светом, и ретиналь ответственен за выполнение ими фоточикла. Также ретиноиды (ретиналь, его метаболиты и синтетические аналоги) представляют интерес и как малотоксичные противоопухолевые препараты, антиоксиданты и хемопротекторы [18].

В [19] исследовалась способность all-E и 13-цис неизомеризующихся аналогов ретиналя встраиваться в бактериоопсин. В экспериментах определялась способ-

ность аналога встраиваться в апобелок, и если такая реконструкция была успешной, то проверялась фоточувствительность полученного пигмента. В результате экспериментов было выявлено, что некоторые аналоги не способны к реконструкции или реконструкция протекает не до конца, давая нековалентный неустойчивый комплекс, разрушающийся при действии света. Другие же аналоги и сам ретиналь быстро встраивались в апобелок, давая устойчивые ковалентные комплексы - пигменты. Схема взаимодействия ретиналя или его аналогов представлена на рис. 2.

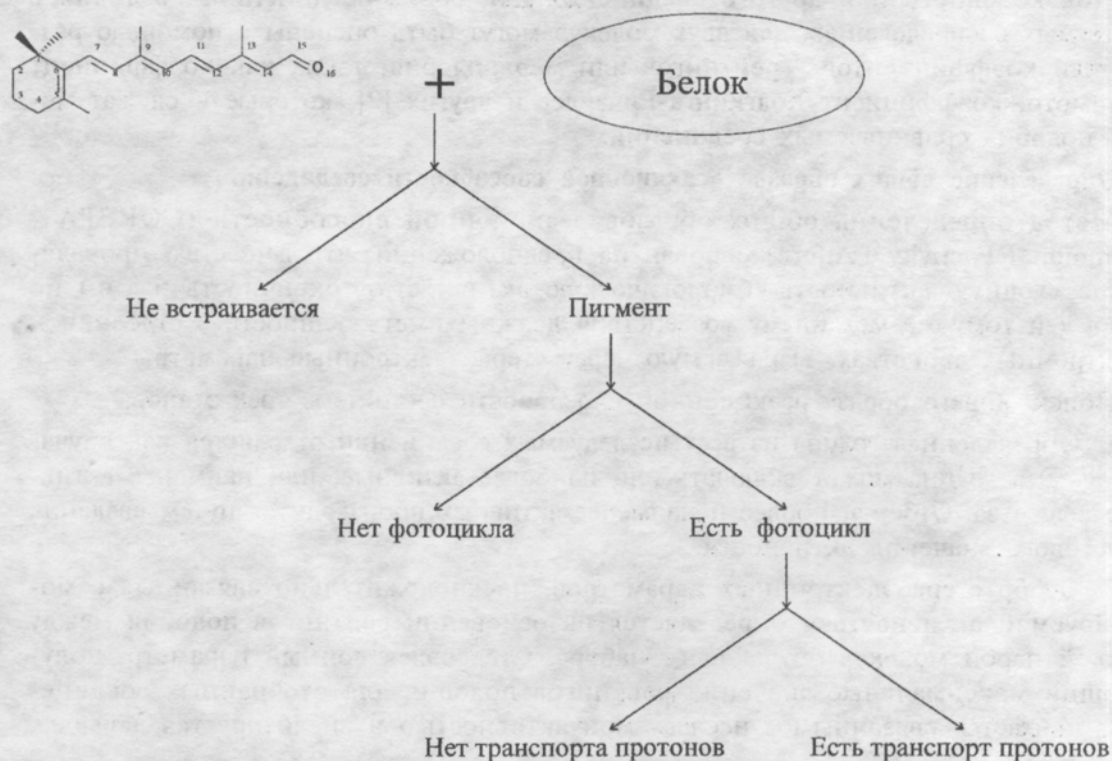


Рис. 2. Схема взаимодействия аналогов ретиналя с бактериоопсином

На основании данных [19] было сделано предположение, что взаимодействие молекул ретиналя (или его аналогов) с апобелком протекает в две стадии: первая-проникновение молекулы ретиналя внутрь полости молекулы апобелка, и вторая-образование ковалентной связи. Логично предположить, что на первой стадии активность обуславливается одним конформером, а на второй - другим. Значит, чтобы молекула в целом была активна - могла встраиваться в апобелок - она должна иметь устойчивые конформеры, отвечающие требованиям и первой и второй стадий процесса. Для первой стадии лимитирующими предположительно являются геометрические ограничения, налагаемые размером полости в молекуле апобелка, а для второй стадии - свойства альдегидной группы как реакционного центра, так как ковалентное связывание происходит путем образования основания Шиффа (нуклеофильное присоединение).

В [18] для моделирования, например, противоопухолевой активности ретиноидов применяются методы QSAR, основанные на моделировании свойств соедине-

ний с использованием характеристик только одного конформера, энергетически наиболее выгодного. Однако, принимая во внимание тот факт, что несмотря на очень маленькую разницу в теплотах образования различных конформеров одного и того же соединения ( $|\Delta\Delta H_f| \leq 10$  ккал/моль), эти конформеры сильно различаются по значениям своих стереоэлектронных параметров (рис. 3), при моделировании свойств ретиналя и ему подобных соединений необходимо рассматривать не один энергетически наиболее выгодный конформер, а совокупность всех энергетически приемлемых конформеров. Поэтому для моделирования способности ретиналя и его аналогов встраиваться в бактериоопсин был применен метод COREPA.

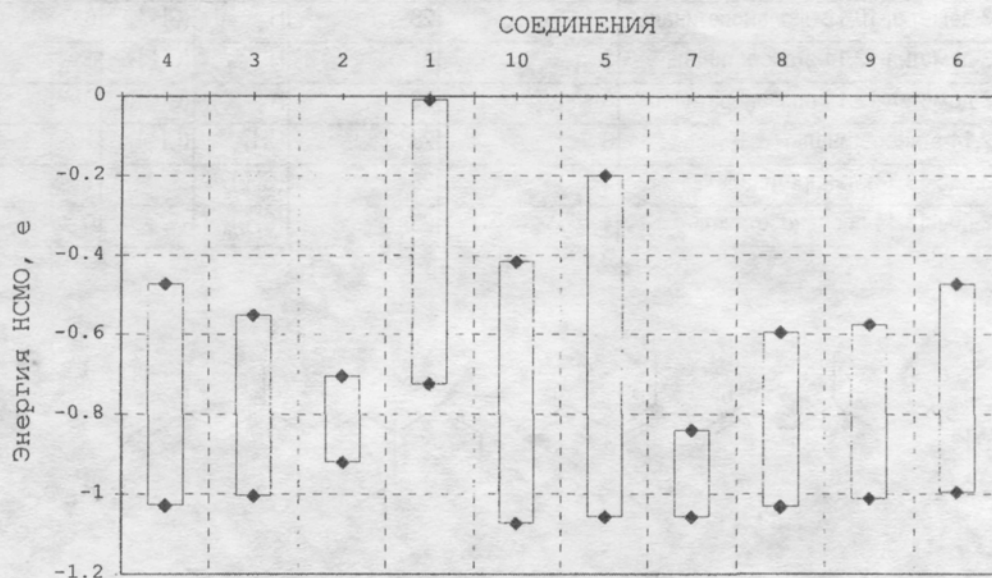


Рис. 3. Интервалы значения энергии НСМО для исследованных соединений (таблица 1)

В таблице 1 представлены рассматриваемые в данном примере соединения. Под активностью понимается способность встраиваться в бактериоопсин с образованием устойчивого пигмента.

Из десяти рассматриваемых соединений естественным образом образуются две группы — активные и неактивные.

Для проведения анализа подобия молекул и определения образа общей реакционной способности необходимо построить набор конформеров для каждого соединения, оптимизировать их геометрии и вычислить стереоэлектронные индексы, из которых можно построить образ реакционной способности исследуемых соединений.

Для каждого из рассматриваемых соединений была произведена генерация конформеров с использованием модуля 3DGEN системы OASIS [5]. Алгоритм 3DGEN производит исчерпывающее исследование конформационного пространства, соответствующего заданной молекулярной топологии с учетом ограничений, задаваемых пользователем. В рассматриваемом случае для молекул ретиналя и его аналогов при генерации были зафиксированы торсионные углы, соответствующие транс-конфигурациям двойных связей. Вращение вокруг одинарных связей производилось с разрешением  $60^\circ$ .



## Исследуемые соединения

Таблица 1

Соединение	Номер формулы (рис. 4)	Число конформеров	R <sub>1</sub>	R <sub>2</sub>	Активность
1. Ретиналь	1	50	-	-	1
2. 10,12-этаноретиналь	2	30	CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>2</sub>	1
3. 10,12-пропаноретиналь	2	27	CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>3</sub>	0
4. 13-деметил-10,12-этаноретиналь	2	29	H	(CH <sub>2</sub> ) <sub>2</sub>	1
5. 13-деметил-10,12-пропаноретиналь	2	28	H	(CH <sub>2</sub> ) <sub>3</sub>	0
6. 13-деметил-12,14-этаноретиналь	3	31	H	(CH <sub>2</sub> ) <sub>2</sub>	1
7. 13-деметил-12,14-пропаноретиналь	3	29	H	(CH <sub>2</sub> ) <sub>3</sub>	0
8. 12,14-этаноретиналь	3	28	CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>2</sub>	1
9. 13-цис-13,14-этаноретиналь	4	28	(CH <sub>2</sub> ) <sub>2</sub>	-	1
10. 13-цис-13,14-пропаноретиналь	4	25	(CH <sub>2</sub> ) <sub>3</sub>	-	0

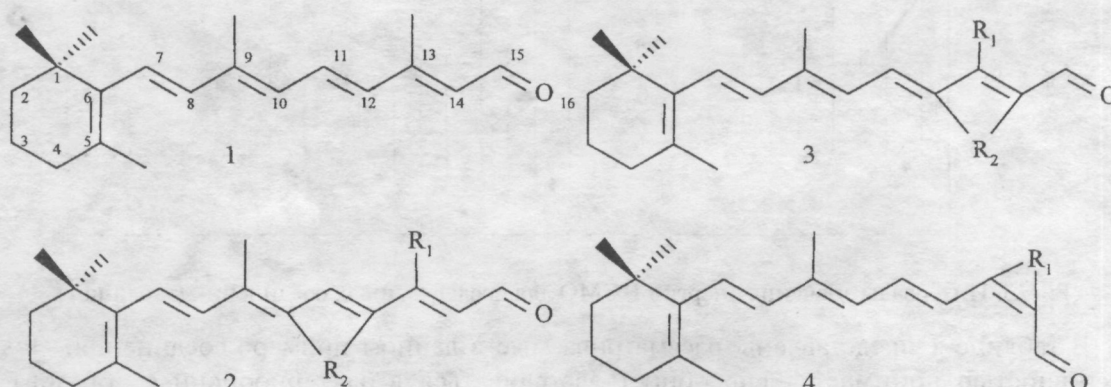


Рис. 4. Исследуемые соединения.

После генерации была произведена оптимизация геометрии структур конформеров с использованием программы МОРАС 7, входящей в состав системы OASIS. Оптимизация проводилась с использованием гамильтониана PM3. После оптимизации часть конформеров была отброшена вследствие того, что процедура оптимизации не достигала сходимости. Оставшиеся конформеры все оказались приемлемыми с энергетической точки зрения, так как все они попадали в интервал  $0 \leq \Delta H_f \leq 10$  Ккал/моль. Количество конформеров, получившееся в результате для каждого соединения, приведено в табл. 1.

Затем для исследуемых молекул были вычислены глобальные квантово-химические индексы, такие как расчетная теплота образования, энергии НСМО (LUMO) и ВЗМО (HOMO), объемная поляризуемость, дипольный момент, а также локальные стереоэлектронные индексы, такие как атомные заряды, электронные плотности и др.

Для определения общих структурных черт, присущих по отдельности активным и неактивным соединениям, необходимо определить набор индексов, которые

лучше всего отражают сходство внутри групп соединений отдельно с высокой и с низкой активностью. Для этого были вычислены несколько рейтингов подобия - Танимото (T), 3DTанимото (T3D), среднее квадратичное расстояние (D), информационную функцию (IF) и коэффициент Ходжкина-Ричардса (C). Для функций T, IF, T3D и C чем больше значение рейтинга, тем лучше исследуемый индекс отражает сходство между молекулами внутри группы, а для D - наоборот, чем больше рейтинг, тем хуже. Таким образом, можно отобрать только те индексы, которые лучше отражают сходство молекул отобранных соединений.

В результате вычисления рейтингов подобия для неактивных (группа 1) и активных (группа 2) соединений были отобраны индексы с наибольшими значениями рейтинга T ( $0.42 \leq T \leq 0.05$ ). Ими оказались: полярность, донорно-акцепторные свойства, специфические межатомные расстояния, заряды, энергия НСМО, дипольный момент, планарность и объемная поляризуемость. Другие рейтинги также дали для этих индексов высокие значения, отличался только порядок расположения индексов. И у группы 1, и у группы 2 наибольшие рейтинги подобия получили одни и те же индексы, но интервалы значений этих индексов для групп 1 и 2 оказались смещены относительно друг друга.

Далее для каждой из исследуемых молекул в пределах одной группы были построены гистограммы в координатах интервалы значений индекса - количество конформеров, попадающих в этот интервал. Количество конформеров, использованных при построении гистограмм, указано в таблице 1. Полученные гистограммы затем накладывались друг на друга с тем, чтобы найти интервалы перекрывания, то есть интервалы, в которых оказывается по крайней мере по одному конформеру от каждой молекулы. Те индексы, которые показывают наибольшее сходство в распределениях внутри группы, считаются индексами, обуславливающими активность, и совокупность интервалов перекрывания этих индексов принимается как образ общей реакционной способности.

В таблице 2 представлены только те индексы, которые имеют уникальные интервалы перекрывания для групп 1 и 2 (то есть в интервале перекрывания индексов для группы 1 не оказывается ни одного конформера соединений из группы 2 и наоборот) или статистически различимые интервалы для групп 1 и 2 соответственно. Под статистической различимостью в данном случае подразумевается различие в распределениях конформеров соединений групп 1 и 2 в пределах выявленных общих интервалов значений индексов. Для проверки статистической различимости был использован непараметрический тест Колмогорова-Смирнова (КС-тест). В таблице 2 представлены полные интервалы варьирования значений индексов для всех конформеров молекул групп 1 и 2, а также приведены доверительные интервалы значений этих индексов, рассчитанные относительно середины интервалов перекрывания, общих по отдельности для групп 1 и 2.

На рис. 5 проиллюстрированы сравнительные гистограммы распределений конформеров молекул групп 1 и 2 для следующих индексов: заряд на атоме C<sub>15</sub>, планарность, энергия НСМО, и максимальное внутримолекулярное расстояние.

Из таблицы 2 видно, что группа 2 (активные соединения) отличается от группы 1 (неактивные соединения) по следующим параметрам:

- акцепторные свойства для атома C<sub>15</sub> молекул группы 2 больше, чем у группы 1
- атом C<sub>15</sub> имеет больший положительный заряд для соединений группы 2
- энергия НСМО для группы 2 ниже, чем для группы 1

- молекулы соединений группы 2 более плоские (значение индекса планарности для группы 2 меньше, чем для группы 1)
- молекулы соединений группы 2 имеют больший дипольный момент
- молекулы соединений группы 2 более полярны
- молекулы соединений группы 2 более компактны - имеют меньшую максимальную длину и специфические расстояния  $C_2-O_{16}$  и  $C_2-C_{15}$ .

Таблица 2

## Интервалы значений индексов с наибольшими "рейтингами подобия"

Индекс	Атомы	Группа 1	Группа 2	Доверительный* интервал $\pm 95\%$	КС- тест**
Акцепторная суперделокализуемость	$C_{15}$	0.2731÷0.2749	0.2715÷0.2770	0.272÷0.273 0.274÷0.275	1<2
Энергия НСМО	все	-1.0278÷-0.0158	-1.0948÷-0.1987	-0.755÷-0.668 -0.856÷-0.804	1>2
Полярность	$C_{15}$	0.0438÷0.0446	0.0436÷0.0445	0.0441÷0.0442 0.0440÷0.0441	1<2
Дипольный момент	все	3.1074÷3.7042	3.2757÷4.0281	3.330÷3.459 3.543÷3.632	1<2
Планарность	все	27.35÷28.55	21.07÷22.57	27.873÷28.191 21.486÷21.770	1>2
Заряд	$C_{15}$	0.3053÷0.3386	0.3031÷0.3514	0.317÷0.320 0.323÷0.326	1<2
Межатомное расстояние	$C_{2,3,4}-O_{16}$ $C_{2,3,4}-C_{15}$	11.7991÷14.8870	11.6604÷14.718	12.725÷12.803 12.600÷12.654	1>2
Максимальное внутримолекулярное расстояние	все	15.4071÷15.5572	14.5556÷15.1587	15.450÷15.510 14.817÷14.877	1>2

Часть этих индексов, а именно планарность и специфические и максимальные расстояния, отвечает за облегчение диффузии внутрь полости апобелка, другая - акцепторные свойства атома  $C_{15}$  и его заряд, полярность и дипольный момент, энергия и заселенность НСМО - за протекание реакции образования основания Шиффа. По полученным данным видно, что активные соединения (группа 2) более компактны и имеют более реакционноспособную альдегидную группу.

Выявленные характерные структурные особенности, отличающие производные ретиналя, встраивающиеся в бактериоопсин, от не способных встраиваться, позволяют прогнозировать наличие или отсутствие этой способности у еще не синтезированных производных ретиналя. Для проверки наличия или отсутствия активности у вновь предлагаемого вещества необходимо для него построить набор приемлемых с энергетической точки зрения конформеров. Затем для этих конформеров надо рассчитать значения тех индексов, которые входят в образ реакционной способности, построить для этих индексов распределения конформеров и

\* Верхний интервал для группы 1, нижний - для группы 2

\*\* 1&lt;2 - группа 1 имеет распределение, смещенное относительно группы 2 влево, 1&gt;2 - группа 1 имеет распределение, смещенное относительно группы 2 вправо



сравнить эти распределения с распределениями конформеров веществ из группы «активных» и «неактивных». В результате сравнения распределений конформеров вещество может быть отнесено к «активной» или «неактивной» группе. Так как в образ входят несколько разных индексов, то возможно две ситуации:

1. По всем индексам вещество отнесено к активным (неактивным)
2. По части индексов вещество признается активным, а по другой части - неактивным.

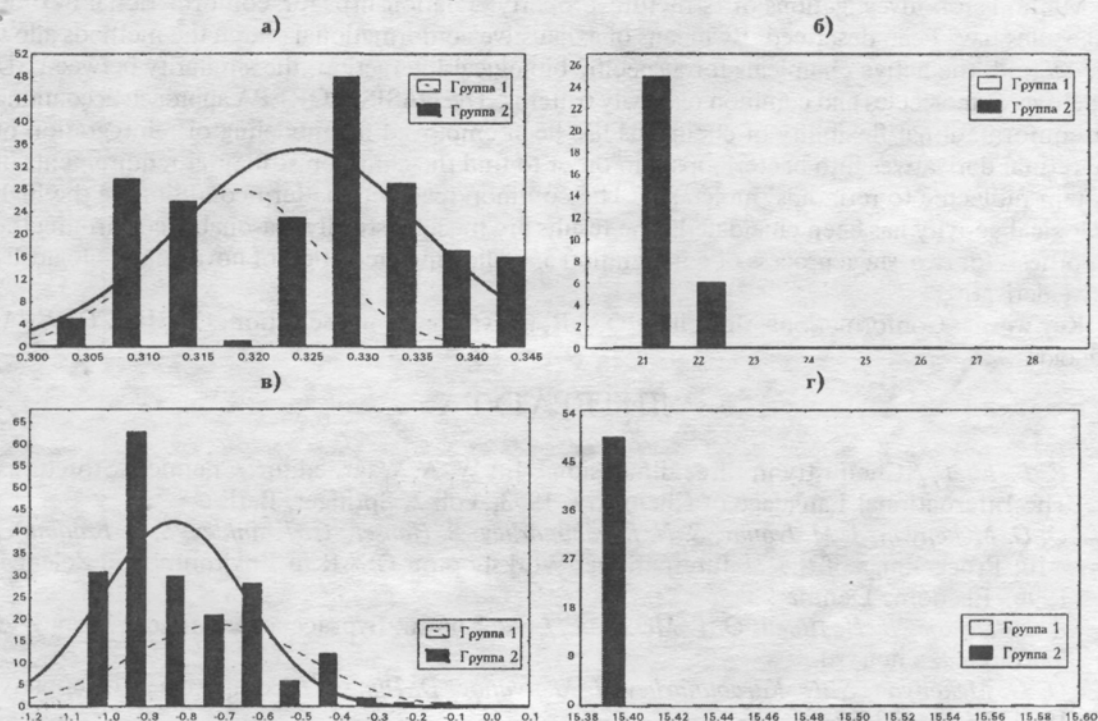


Рис. 5. Частотные распределения для зарядов на атоме  $C_{15}$  (а), планарности (б), энергии НСМО (в) и максимального внутримолекулярного расстояния (г).

Если в первом случае не возникает вопроса о классификации нового соединения, то во втором случае можно отнести вещество с некоторой вероятностью  $P_{акт}$  к активным ( $P_{неакт}$  к неактивным):

$$P_{акт} = \frac{N_{акт}}{N} \quad P_{неакт} = \frac{N_{неакт}}{N} \quad (1)$$

где  $N_{акт}$  - число индексов, по которым вещество классифицируется как активное,  $N_{неакт}$  - число индексов, по которым вещество классифицируется как неактивное,  $N$  - общее число индексов ( $N = N_{акт} + N_{неакт}$ )

Вещество следует отнести к той группе, у которой вероятность больше. Кроме того, возможно построение экспертных систем, которые на основании данных образа реакционной способности смогут принять решение о том, к какой группе следует отнести вновь предлагаемое соединение.

## ACCOUNTING FOR CONFORMATIONAL FLEXIBILITY OF MOLECULES IN SAR/ QSAR STUDIES

*O. G. Mekenyan\*, A. E. Tarkhov\*\*, V. I. Shvets\*\**

\* «As. Zlatarov» University, 8010, Bourgas, Bulgaria

\*\* Moscow State Academy of Fine Chemical Technology named after M.V. Lomonosov,  
177571, Moscow, Vernadskogo ave, 86

Methods for investigations of «structure-property» relationships for conformational flexible molecules have been described. By means of exhaustive conformational search the methods allow one to find: the active chemicals for a specific biological interaction, the similarity between 3D structures of molecules and common reactivity patterns. The OASIS-COREPA approach accounting for conformational flexibility of chemicals has been employed for modeling of reintegration of the retinal derivatives into bacterioopsin in order to find the common structural requirements of protein molecule to retinoids' molecules. The common reactivity pattern conditioning the high biological activity has been elucidated. The results are mechanistically reasonable confirming the hypothesis for two-stage process of reintegration and allowing prediction of novel and biologically active derivatives.

**Key words:** Conformational flexibility, QSAR, active conformer selection, OASIS-COREPA, retinoids.

## ЛИТЕРАТУРА

1. E. L. Eliel // Chemistry in three dimensions, In: W. A. Waar, editor. Chemical Structures. The International Language of Chemistry, 1993, Vol. 2, Springer, Berlin.
2. O. G. Mekenyan, J. M. Ivanov, S. N. Karabunarliev, B. Hansen, G. T. Ankley, S. P. Bradbury / / In: Proceedings of the 7<sup>th</sup> International Workshop on QSAR in Environmental Science, 1996, Elsinore, Denmark.
3. А. Е. Тархов, В. И. Швеи, О. Г. Мекенян // Ежегодник Бургасского Университета «Ас. Златаров», в печати.
4. O. G. Mekenyan, S. N. Karabunarliev, J. M. Ivanov, D. Dimov // Computers & Chemistry, 1994, 18, p. 173-178.
5. J. M. Ivanov, S. N. Karabunarliev, O. G. Mekenyan // J. Chem. Inf. Comput. Chem., 1994, 34, 234-243.
6. O. G. Mekenyan, J. M. Ivanov, G. D. Veith, S. P. Bradbury // Quant. Struct.-Act. Relat., 1994, 13, 302-307.
7. O. G. Mekenyan, G. T. Ankley, G. D. Veith, D. J. Call // Environ. Health Perspect. (in press).
8. В. В. Авидон // Хим.-Фарм. Журнал, 1974, 8, 22-25.
9. В. Е. Голендер, А. Б. Розенблит // Журнал ВХО им. Д.И. Менделеева, 1980, 25, №1, 28-35.
10. D. A. Filimonov, V. V. Poroikov // Design of Bioactive Compounds for Industrial Use, Oxford: Bio. Sci. Publishers, 1996, 47-56.
11. M. T. Bakarat, P. M. Dean // J. Computer-Aided Mol. Des., 1990, 4, 295-316 и 317-330.
12. R. Carbo, L. Leyda, M. Arnau // Int. J. Quant. Chem., 1980, 17, 1185-1189.
13. 3D QSAR in Drug Design. Theory, Methods and Applications // Ed. By H. Kubinyi, ESCOM, Leiden, 1993.
14. C. A. Pepperrell, P. Willett, // In: W. A. Waar, editor. Chemical Structures. The International Language of Chemistry, 1993, Vol. 2, Springer, Berlin, 377-382.
15. F. O. Guner, D. W. Hughes, L. M. DuMont // J. Chem. Inf. Comput. Sci. 1991, 31, 408-414.
16. B. D. Silverman, D. E. Platt // J. Med. Chem., 1996, 39, 2129-2140.
17. J. S. Mason // In: Edited by V. Claassen, Trends In Drug Design, 1993, Elsevier, Amsterdam, 147-156.
18. D. Bonchev., C. F. Mountain, W. A. Seitz, A. T. Balaban // J. Med. Chem., 1993, 36, 1562-1568.
19. L. V. Khitrina, Yu. G. Kirillova, A. A. Khodonov // Abstract Volume 2<sup>nd</sup> IUBMB Conference Biochemistry of Cell Membranes, Bari, Italy, 1993, P. 286.