

УДК 577.152.3.

©Коллектив авторов

КОМПЬЮТЕРНОЕ КОНСТРУИРОВАНИЕ ЛЕКАРСТВ НА ОСНОВЕ АНАЛИЗА СТРУКТУРЫ МАКРОМОЛЕКУЛЫ-МИШЕНИ: I. ПОИСК И ОПИСАНИЕ МЕСТА СВЯЗЫВАНИЯ ЛИГАНДА В БЕЛКЕ-МИШЕНИ

А.С.Иванов, А.В.Дубанов, В.С.Скворцов, А.И. Арчаков.

ГУ НИИ биомедицинской химии им. В. Н. Ореховича РАМН, 119832, Москва,
Погодинская ул., 10; факс: (095) 245-0857; эл.почта ivanov@ibmh.msk.su

Оценена возможность применения метода молекулярного докинга для нахождения места связывания лиганда в белке-мишени. В основе подхода лежит генерация гипотез об образовании белок-лигандных комплексов путем докинга трехмерной структуры низкомолекулярного лиганда к поверхности белка с последующей оценкой качества гипотез по величине площади контактирующих поверхностей и расчетной функции, характеризующей энергию образования комплекса. Процедуру докинга выполняли с помощью оригинальной компьютерной программы DockSearch, работающей как на PC Pentium, так и на компьютерах SGI. Программа генерирует и оценивает набор вероятных позиций лиганда на поверхности белка. Минимизацию энергии комплексов выполняли с помощью программного комплекса Sybyl 6.5 (Tripos Inc.) на сервере Silicon Graphics Origin200. В конечную выборку отбирали модели с наибольшей площадью контактирующих молекулярных поверхностей и величиной энергии образования комплекса. Данный подход был проверен на 26 различных белках, для которых известны структуры их комплексов с лигандами. Была произведена оценка границ применимости описанного подхода. Показано, что для большинства белков-мишеней предлагаемый метод позволяет правильно найти место связывания лиганда и получить его описание с достаточной точностью для последующего поиска веществ-прототипов в базах данных структур химических соединений. В работе также проанализированы причины негативных результатов, полученных при использовании данного подхода в некоторых случаях.

Ключевые слова: компьютерное конструирование лекарств, молекулярное моделирование, структура белка-мишени, молекулярный докинг, молекулярная графика, низкомолекулярный лиганд.

ВВЕДЕНИЕ. Поиск новых веществ-прототипов в базах данных химических соединений является одним из основных подходов в компьютерном конструировании лекарств [1-3]. Как правило в данном методе используется процедура автоматизированного докинга низкомолекулярных соединений к заданному месту связывания лиганда в белке-мишени с последующей оценкой результатов моделирования межмолекулярного комплекса [4]. Процедура докирования выполняется для каждого низкомолекулярного соединения, содержащегося в химической базе данных. Полученные гипотезы белок-лигандных комплексов оценивают с помощью различных оценочных функций (scoring functions) включая пространственную комплементарность, гидрофобные и электростатические взаимодействия, взаимодействия Ван-дер-Ваальса и образование водородных связей.

Существует около десятка различных академических и коммерческих программ докинга [4]. Для поиска в базах данных могут быть использованы такие программы как Dock, Caveat, Clix и ограниченно Ludi. Так как строгое решение

задачи докинга требует большого объема вычислений, связанных с расчетами как энергии молекулярных комплексов, так и огромного числа конформационных и конфигурационных вариантов молекул, то в алгоритмах докинга используется ряд упрощений, таких как моделирование воды с помощью диэлектрического континуума, использование энтальпии вместо свободной энергии и трактовка гибких лигандов и белков как жестких объектов. В зависимости от используемых упрощений методы докинга сгруппируются в четыре семейства: 1) дескрипторные методы, 2) методы координатной сетки, 3) методы фрагментов и 4) кинетические методы.

Ранее нами была разработана оригинальная программа DockSearch [5], предназначенная для быстрого поиска потенциальных лигандов в базах данных. Программа основана на поиске по координатной сетке с размещением лиганда в белке путем его дискретного вращения и перемещения при сохранении жесткого белка-мишени в неподвижном состоянии. В отличие от упомянутых выше программ, DockSearch работает как генератор гипотез о возможной структуре молекулярных комплексов “белок-лиганд”, первоначальная селекция которых осуществляется непосредственно в рамках работы DockSearch по степени пространственной комплементарности. Полученные результаты переносятся в электронные молекулярные таблицы пакета молекулярного моделирования Sybyl и дальнейшая селекция гипотез осуществляется после выполнения их энергетической оптимизации с учетом всех видов взаимодействий, вычисляемых с помощью методов молекулярной механики.

Данный подход был ранее успешно использован нами для поиска в базах данных новых ингибиторов цитохромов P450 [6].

Обязательным условием реализации описанного подхода является наличие геометрического описания места связывания лиганда в белке-мишени. Эту информацию получают путем анализа структуры комплекса белка-мишени с каким-либо известным низкомолекулярным лигандом. В настоящее время банк данных Protein Data Bank (PDB) [7] содержит несколько сотен белковых структур, которые могут рассматриваться как потенциальные мишени для новых лекарств. К сожалению, трехмерные структуры многих белков в комплексе с известными лигандами, как правило, отсутствуют.

Целью данной работы была оценка возможности применения метода молекулярного геометрического докинга для нахождения с достаточной точностью места связывания лиганда в белках с известной трехмерной структурой путем генерации гипотез о формировании комплексов “белок-лиганд” с последующей их оценкой и отбором по величине площади контактирующей поверхности и расчетной величине, характеризующей стабильность комплекса.

МЕТОДИКА. В качестве молекулярных тест-объектов из базы данных PDB [7] были выбраны белки, для которых известна трехмерная структура их комплексов со специфичными низкомолекулярными лигандами. Предпочтение отдавалось тем белкам, для которых известна не только структура комплекса, но и структура белка без лиганда. Среди лигандов предпочтение отдавалось малым молекулам (как полярным, так и неполярным). В список тест-объектов было также включено несколько комплексов белков с пептидными лигандами. В результате были отобраны 26 комплексов и 12 свободных белков. Список всех молекулярных тест-объектов приведен в таблице 1.

Основные вычисления выполняли с помощью программ DockSearch [5] и Sybyl 6.5 [8] на рабочей графической станции Silicon Graphics O2 (R5000) и сервере Silicon Graphics Origin200 (2xR10000). Ряд вычислений и визуализация результатов были выполнены с помощью оригинальной программы Onix [9-11] на PC Pentium II под управлением ОС Windows98.

Этапы выполненных вычислительных экспериментов показаны на рис. 1. Трехмерные координаты структур тестовых белков-мишеней и комплексов “белок-лиганд” были взяты из базы данных PDB. Координаты структур лигандов были получены из известных структур комплексов “белок-лиганд” и/или с помощью

Таблица 1. Молекулярные тест-объекты

No	Код PDB*	Название белка **	Индекс лиганда в PDB **
1	lahi (lahh)	7- α -Hydroxysteroid dehydrogenase (0)	CHO
2	lbmm	α -Thrombin (H + L)	I
3	ldwc	α -Thrombin (H + I)	I
4	lbtx	β -Trypsin (A)	H
5	lwat (lwas)	Aspartate receptor (B)	S (ASP)
6	latp	cAMP-dependent protein kinase (E + ATP)	I
7	lcil	Carbonic anhydrase II (0)	ETS
8	lcin (lca2)	Carbonic anhydrase II (0 + ZN)	MTS
9	lkel (lkem)	Catalytic antibody FAB fragment (L + H)	AAH
10	lmnc	Collagenase (0)	PLH
11	lphg (lphc)	Cytochrome P450 101 (0 + HEM)	MYT
12	2cpp (lphc)	Cytochrome P450 101 (0 + HEM)	CAM
13	loxa	Cytochrome P450 107 (0 + HEM)	DEB
14	leat (llvy)	Elastase (0)	TFI
15	lelb	Elastase (A)	B
16	lfgk (lffk)	Fk506 Binding Protein (0)	SB3
17	lgsf (lgd)	Glutathione transferase a1-1 (A)	EAA
18	lcp	HIV-1 protease (A + B)	C
19	lhpx	HIV-1 protease (A + B)	KNI
20	lmmp	Matrilysin (0)	RSS
21	lmb	Matrix metalloproteinase (0)	BAT
22	lpge (lprh)	Prostaglandin H2 synthase I (A + HEM)	ISF
23	lcbi (lcbi)	Retinoic-acid-binding protein (A)	REA
24	lrlc (lrlc)	Ribulose-1,5-bisphosphate carboxylase (L)	S
25	ldil	Sialidase (0)	EQP
26	6tmn (lfn)	Thermolysin (E)	I

Примечание: * - В скобках указаны коды белков без лигандов; ** - названия белков и лигандов цитируются по файлам PDB. Одной буквой обозначены пептидные цепи, двумя или тремя – небелковые структуры. Если в PDB файле имеется только одна белковая цепь, то она обозначена цифрой 0.

молекулярного моделирования средствами программного комплекса Sybyl. Для небольших гибких лигандов, состоящих не более чем из 20 тяжелых (не водородных) атомов, осуществляли поиск наиболее энергетически выгодных конформеров путем нахождения глобального минимума энергии с использованием процедуры RandomSearch (Sybyl).

Всего было выполнено четыре группы экспериментов (группы *A*, *B*, *C* и *D*, Таблица 2), отличающихся подготовкой молекулярных объектов перед процедурой докинга. Структуры белков из комплексов “белок-лиганд” были использованы в качестве мишеней в группах *A* и *B*, а структуры белков, не содержащие лигандов, – в группах *C* и *D*. Эксперименты группы *D* служили контрольными тестами для объектов, которые дали неудовлетворительные результаты в экспериментах группы *C*.

На первом этапе производили разделение молекулярных комплексов на отдельные структуры белка и лиганда. Затем их ориентацию в пространстве изменяли случайным образом.

На втором этапе осуществляли генерацию гипотез молекулярных комплексов “белок-лиганд” с помощью программы DockSearch (см. общую схему применения программы DockSearch для скрининга баз данных на рис. 2). Для каждого анализируемого низкомолекулярного соединения генерировался набор возможных позиций на поверхности белка. Далее отбирали лучшие гипотезы по стерической комплементарности (см. пример селекции гипотез комплексов по

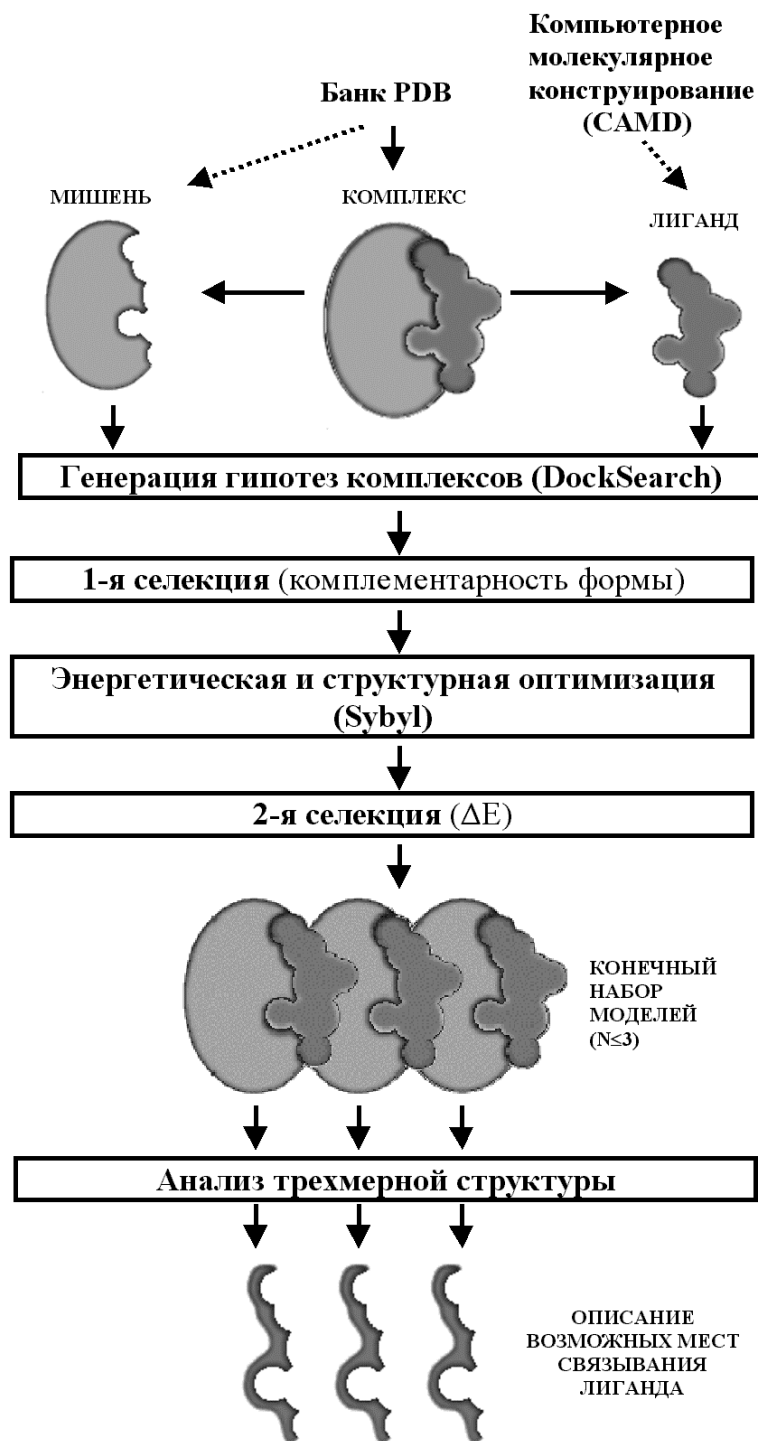


Рисунок 1.

Основные этапы выполненных вычислительных экспериментов по нахождению места связывания лиганда в белке-мишени.

Таблица 2. Группы вычислительных экспериментов и этапы их выполнения

Этап	Действие	Группы экспериментов			
		A	B	C	D
1	Белок-мишень	Из комплекса	Из комплекса	Свободный	Свободный
2	Лиганд	Из комплекса	Из комплекса	Из комплекса	Из комплекса
3	Минимизация мишени	НЕТ	ДА	ДА	ДА
4	Поиск конформеров лиганда	НЕТ	НЕТ	ДА	НЕТ
5	Минимизация лиганда	НЕТ	ДА	ДА	ДА

Таблица 3. Сравнительный анализ компьютерных моделей и экспериментальных данных о строении комплексов белок-лиганд.

<i>Эксперименты группы А (13 экспериментов)</i>					
Номер объекта (табл.1)	PDB код	RMS* (Е)	Отклонение** (Е)	Остатки*** (%)	Оценка
#2	1bmm	1.1	-	80	+
#3	1dwc	0.2	-	100	+
#4	1btx	1.4	-	80	+
#6	1atp	11.6	-	-	-
#10	1mnc	0.9	-	90	+
#15	1elb	0.2	-	100	+
#17	1gsf	22.1	20.9	0	-
#18	1cpi	0.6	-	90	+
#19	1hpx	0.3	-	100	+
#20	1mmp	0.7	-	90	+
#21	1mmb	0.3	-	100	+
#22	1pge (1prh)	12.2	10.8	0	-
#25	1dil	1.9	1.5	80	+
<i>Эксперименты группы В (2 эксперимента)</i>					
#7	1cil	4.9	0.8	90	+
#13	1oxa	5.8	1.4	100	+
<i>Эксперименты группы С (13 экспериментов)</i>					
#1	1ahi (1ahh)	8.2	0.4	70	+
#5	1wat (1was)	15.7	15.2	25	-
#8	1cin (1ca2)	6.3	3.4	64	±
#9	1kel (1kem)	33.4	33.1	0	-
#11	1phg (1phc)	16.3	10.7	0	-
#12	2cpp (1phc)	2.4	0.3	100	+
#14	1eat (1lvy)	7.7	3.0	58	±
#16	1fkg (1fkk)	9.8	4.4	64	±
#17	1gsf (1gsd)	12.2	10.9	19	-
#22	1pge (1prh)	12.7	11.0	21	-
#23	1cbr (1cbi)	10.0	2.2	75	+
#24	1rlc (1rld)	24.9	24.1	0	-
#26	6tmn (1lnf)	5.7	2.2	63	±
<i>Эксперименты группы D (2 эксперимента)</i>					
#17	1gsf(1gsd)	13.0	11.4	0	-
#22	1pge(1prh)	11.7	11.2	0	-

Примечание: * - RMS для тяжелых атомов лиганда, ** - отклонение геометрического центра лиганда, *** - процент идентичных аминокислотных остатков в месте связывания.

величине площади контактирующей молекулярной поверхности на рис. 3).

Площадь контактирующих поверхностей белка и лиганда оценивали с помощью программы DockSearch. В этой программе поверхность сфер, соответствующих атомам белка и лиганда, представлялась в виде набора точек. Для каждой из точек оценивается ее доступность для молекулы растворителя, аппроксимированной тестовой сферой. Площадь контакта рассчитывалась как разность числа точек, доступных для растворителя в свободных белке и лиганде и числа точек, доступных для растворителя в комплексе «белок-лиганд». Для аппроксимации молекулы воды использовалась сфера радиусом 1,5 Е и для выявления «плотных» контактов белка с лигандом – сфера с радиусом 0,7 Е. Для дальнейшей оптимизации выбиралось по 20 комплексов с наибольшими значениями

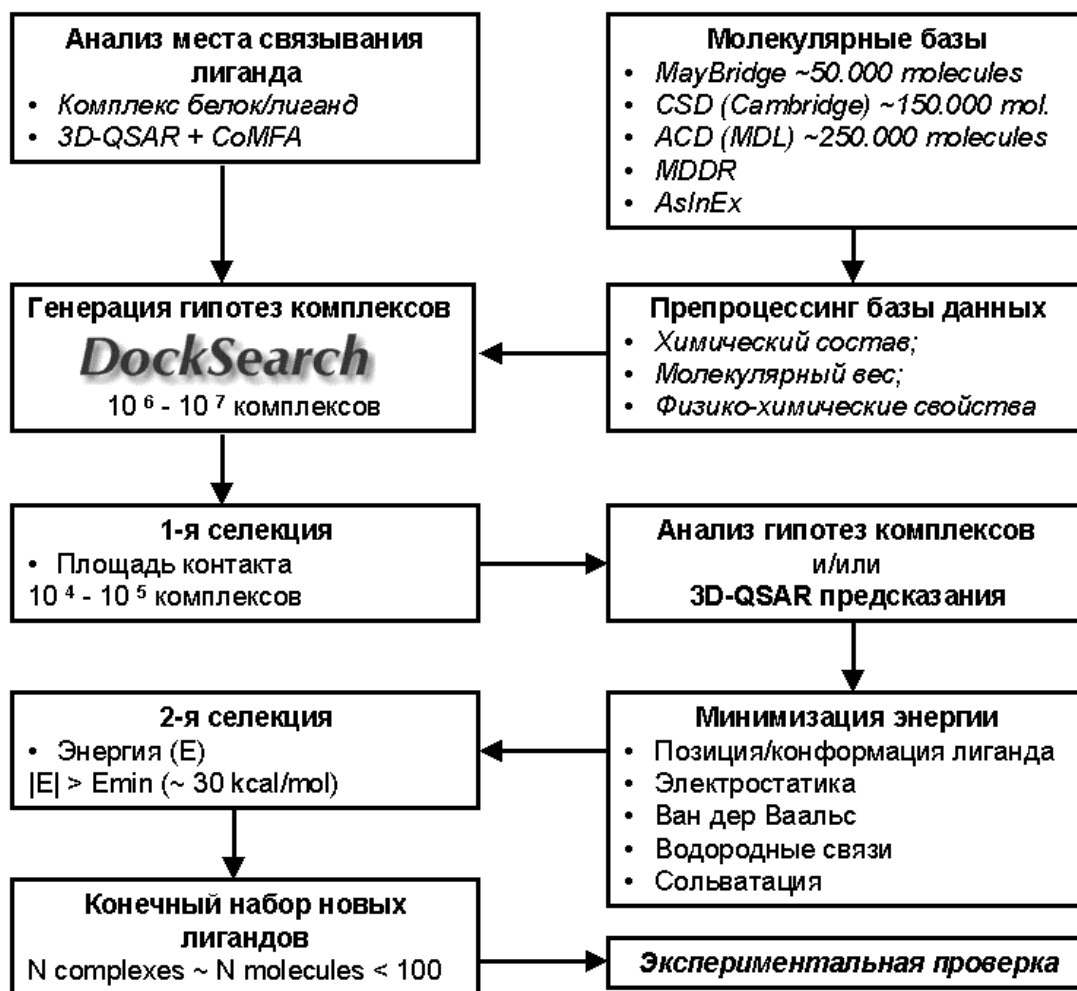


Рисунок 2.
Схема работы программы DockSearch.

площади контакта, рассчитанной с использованием тестовых сфер 1,5 Е и 0,7 Е.

Оптимизацию структуры гипотетических комплексов “белок-лиганд” и минимизацию их энергии выполняли с помощью программного комплекса Sybyl 6.5 для лигандов и элементов белковых структур в пределах 12Е от лиганда.

На следующем (третьем) этапе выполняли вторую селекцию гипотез по величине энергии образования комплекса, рассчитанной с помощью метода молекулярной механики с использованием силового поля Tripos программного комплекса Sybyl. При этом учитывались следующие составляющие: конформационные компоненты (изменения энергий связей, углов и торсионных углов в белке и лиганде при образовании комплекса), внутри- и межмолекулярные взаимодействия (электростатические и Ван-дер-Ваальса), а также образование водородных связей.

В результате получали конечный набор гипотез (не более 3 гипотез на пару “белок-лиганд”) и анализировали возможные места связывания лиганда. Определяли позиции лигандов и места их контактов с белком.

Полученные результаты сравнивали с экспериментальными данными из базы данных PDB с использованием следующих параметров: среднеквадратичное отклонение (RMS) для всех тяжелых атомов лиганда; отклонение геометрического центра лиганда; список аминокислотных остатков белка, образующих место связывания и процент идентичных остатков.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ. Результаты сравнительного анализа построенных моделей и известных структур комплексов представлены в табл. 3 и 4. Как можно видеть, в 10 из 13 экспериментов группы А (т.е. в 77 % случаев) место

Таблица 4. Основные результаты сравнительного анализа компьютерных моделей и экспериментальных данных о строении комплексов белок-лиганд.

Тесты группы А

	Средняя величина	Лучшая величина	Худшая величина	% ошибки
RMS (E)	4,1	0,2	20,9	9%
Отклонение (E)	-	-	-	
Остатки (%)	91	100	80	

Тесты группы В

	Средняя величина	Лучшая величина	Худшая величина	% ошибки
RMS (E)	5,5	4,9	5,8	-
Отклонение (E)	1,1	0,8	1,4	
Остатки (%)	95	100	80	

Тесты группы С

	Средняя величина	Лучшая величина	Худшая величина	% ошибки
RMS (E)	12,7	2,4	33,4	46%
Отклонение (E)	9,3	0,3	33,1	
Остатки (%)	36	100	0	

Тесты группы D

	Средняя величина	Лучшая величина	Худшая величина	% ошибки
RMS (E)	12,4	11,7	13,0	-
Отклонение (E)	11,3	11,4	11,2	
Остатки (%)	-	-	-	

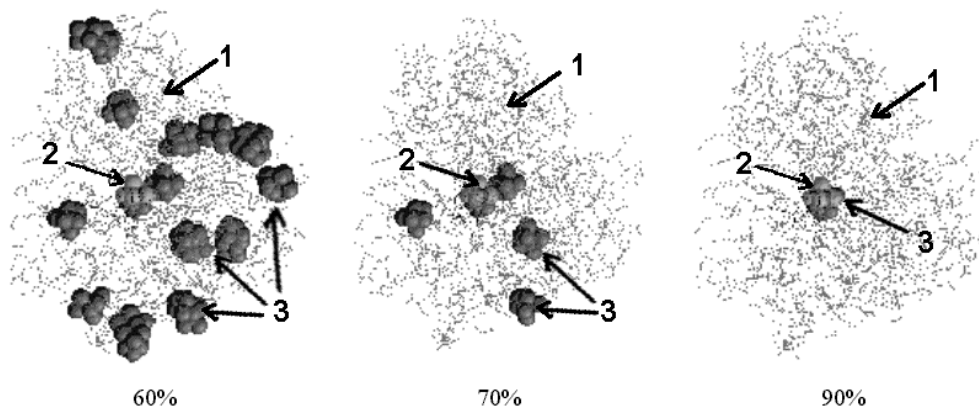


Рисунок 3.

Пример селекции гипотез комплексов по величине площади контактирующей молекулярной поверхности. Тест-объект - цитохром P450 101 (P450cam) (см. #12 в Таблице 1).

1 - цитохром P450cam, 2 - позиция лиганда (камфара) в белок-лигандном комплексе из PDB файла, 3 - DockSearch гипотезы возможных мест связывания лиганда. Внизу на рисунке цифрами указан порог селекции гипотез по величине площади контакта (в процентах от максимального значения, полученного в вычислительном эксперименте)

связывания лиганда в белке-мишени было найдено с достаточной точностью ($RMS \leq 1,9E$) и с идентификацией не менее 80% аминокислотных остатков, образующих место связывания. В 7 экспериментах (в 54 % случаев) точность была заметно выше ($RMS < 1E$). Данные результаты свидетельствуют о том, что как место связывания лиганда, так и его ориентация были определены нами правильно.

Лиганды в этих комплексах представляют собой модифицированные пептиды, то есть являются относительно большими молекулами с разветвленной структурой и характерным "рельефом" молекулярной поверхности. Так как в тестах группы А конформация докинуемого лиганда строго соответствовала конформации лиганда в комплексе, то его место связывания может быть найдено с достаточной

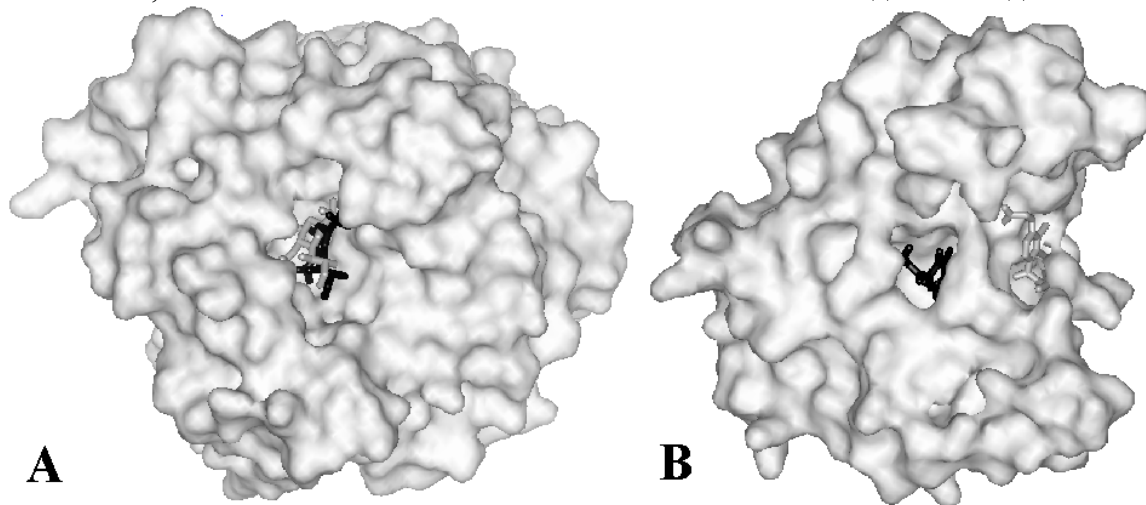


Рисунок 4.

Положительный (А) и отрицательный (В) примеры нахождения места связывания лиганда в экспериментах группы А.

А - комплекс сиалидазы (#25) с ингибитором (светлая молекула - экспериментальные данные, темная - компьютерная модель). В - комплекс глутатионтрансферазы А1-1 (#17) с лигандом - этакриновой кислотой (светлая молекула - экспериментальные данные, темная - компьютерная модель).

точностью уже на стадии первичного отбора по величине площади контактирующей молекулярной поверхности.

В трех случаях (комплексы № 6, № 17 и № 22) были отмечены значительные различия между моделями комплексов и экспериментальными данными. Анализ этих комплексов выявил следующие возможные причины ошибок. В случае цАМФ-зависимой протеинкиназы (№ 6) ошибка обусловлена особенностями как структуры лиганда, так и структуры белка. Лиганд представляет собой пептидный ингибитор (20 аминокислотных остатков) в основном с нерегулярной структурой (имеется только одна короткая спираль из 6 остатков). Место связывания лиганда на поверхности одной из субъединиц фермента имеет вид неглубокой канавки и не обеспечивает высокое стерическое соответствие молекулярных поверхностей. Это обстоятельство обуславливает высокую вероятность ошибки нахождения места связывания лиганда при использовании величины площади контактирующих молекулярных поверхностей в качестве критерия первичного отбора гипотез комплексов.

В двух других случаях (№ 17 и № 22) лиганды являются небольшими полярными молекулами, у которых стерическое соответствие молекулярных поверхностей с белком не играет ведущей роли в процессе связывания. Соответственно, первичный отбор по величине контактирующих поверхностей приводит к выбору гипотез комплексов, где позиции лигандов не являются оптимальными. Положительный и отрицательный примеры из группы А показаны на рис. 4.

В двух экспериментах группы В место связывания лиганда было идентифицировано правильно, о чем свидетельствуют все параметры, приведенные в Таблице 3). Необходимо однако отметить, что ориентация лигандов в обоих случаях не полностью соответствует экспериментальным данным. По-видимому, отклонение обусловлено тем, что ориентация лиганда в комплексе значительной степени определяется взаимодействием зарядов.

Таким образом, тесты групп А и В в целом показали применимость проверяемого подхода для нахождения места связывания низкомолекулярных лигандов в белковых молекулах.

Эксперименты группы С отражают ситуацию, приближенную к реальной, когда исследователю необходимо найти место связывания лиганда в молекуле белка. В качестве мишени были использованы структуры белков, не содержащие лигандов. Предполагалось, что в реальном исследовании пространственная структура лиганда будет моделироваться и оптимизироваться с использованием процедуры минимизации энергии. Поэтому для приближения к реальной ситуации структуры лигандов, полученные из структур известных комплексов, были также минимизированы. Для молекул, содержащих в своем составе крупные гибкие фрагменты, дополнительно была выполнена процедура поиска конформеров. В экспериментах с докингом были использованы конформеры с наименьшей величиной энергии.

В тестах группы С в целом отмечено заметное снижение точности нахождения положения лиганда в составе комплекса по сравнению с группами А и В (см. Таблицу 3). Ориентация лиганда в группе С практически во всех случаях значительно отличалась от экспериментальных данных. Как и в группах А и В, наихудшие результаты (№№ 5, 9 и 17) были получены в случаях, когда лиганд является небольшой и сильно полярной молекулой. Причина ошибок все та же – использование комплементарности поверхностей в качестве критерия первичного отбора комплексов.

При сравнении модельных и экспериментальных комплексов, где лиганды являются гибкими молекулами (например, №№ 17 и 24), было обнаружено, что отличия в конформации лиганда также могут стать причиной неправильного нахождения места связывания. Теоретически, это затруднение может быть преодолено оптимизацией структуры каждого комплекса, генерируемого программой докинга для каждого конформера лиганда. Однако такой подход

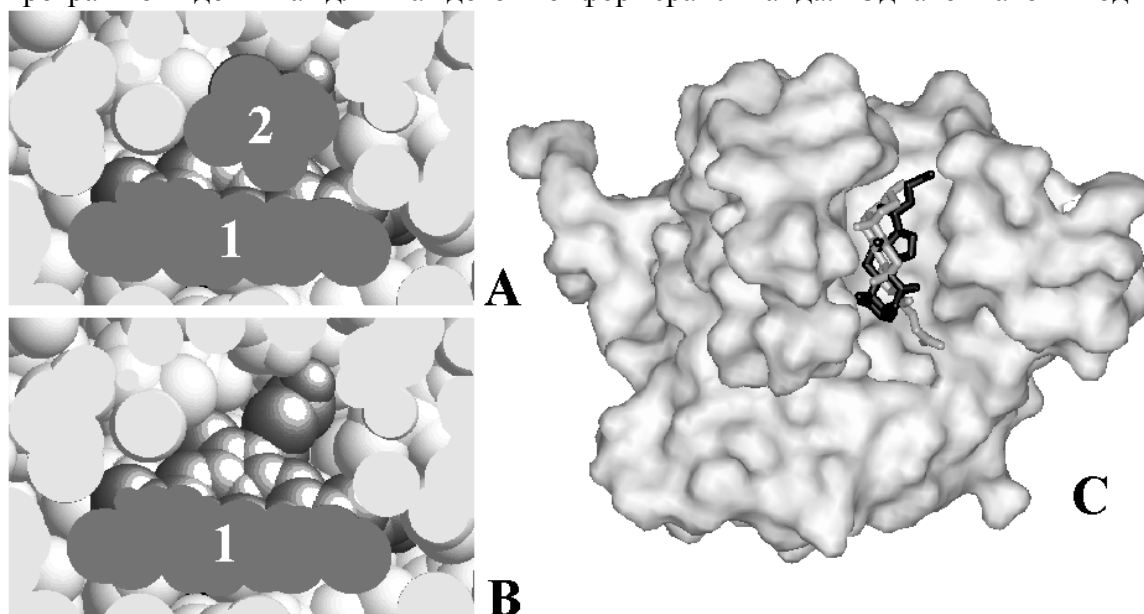


Рисунок 5.

Положительные примеры нахождения места связывания лиганда в экспериментах группы С. А - структура комплекса цитохрома Р450 101 с камфарой (#12), разрез через активный центр фермента (1 - камфара, 2 - гем). В - тоже без камфары. Видно, что место связывания камфары представляет собой практически замкнутую полость, а лиганд находится в тесном контакте с окружающими ее элементами места связывания. С - комплекс 7α-гидроксистероиддегидрогеназы (#1) с 7-оксогликохенодиоксихоловой кислотой (светлая молекула - экспериментальные данные, темная - компьютерная модель). В обоих случаях ориентация лиганда в месте связывания отличается от экспериментальных данных.

требует чрезвычайно больших вычислительных затрат, так как необходимо будет оценить очень большое число (сотни тысяч) вариантов.

Более высокая точность была получена в комплексах белков, в которых место связывания лиганда представляет собой узкую щель (№№ 1 и 23) или практически замкнутую полость (№ 12), а сами лиганды являются молекулами с жестким гидрофобным ядром из конденсированных циклов (№№ 1 и 12) или число энергетически выгодных конформеров которых ограничено (№ 23).

Ориентация лигандов в модельных комплексах во всех случаях имела отличия от экспериментальных данных, что обусловлено по крайней мере в случаях № 1 и № 23 взаимодействием полярных групп белков и лигандов, которое не учитывалось при первичном отборе гипотез комплексов.

Позитивные примеры для группы *C* показаны на рис. 5. Для некоторых объектов, для которых были получены неудовлетворительные результаты в тестах группы *C*, были выполнены дополнительные тесты (группа *D*), которые повторяли эксперименты группы *C* с единственным отличием – для докинга были взяты

Таблица 5. Соотношение электростатических и ван-дер-ваальсовых взаимодействий в комплексах белок-лиганд, для которых не удалось правильно идентифицировать место связывания лиганда в структуре белка-мишени.

№.	5	9	17	22	24
Код PDB	1wat	1kel	1gsf	1pge	1rlc
E_{VDW} / E_{elst}	1 / 9	1,2 / 8,7	3,6 / 6,4	4,5 / 5,5	0,3 / 9,7

исходные конформации лигандов из структур комплексов. Данные тесты выполняли для выяснения причин ошибки в нахождении места связывания лиганда. Предполагалось, что ошибка может быть обусловлена либо отличием конформации лиганда от конформации в составе экспериментального комплекса, либо свойствами лиганда и/или белка в месте связывания лиганда (т.е. стерическая комплементарность не является ведущим фактором молекулярного узнавания).

Сопоставление результатов экспериментов групп *A*, *C* и *D* позволило получить ответ на вопрос о том, какая из нижеследующих причин является основным источником ошибок в нахождении места связывания лиганда описанным подходом:

- отличие конформации свободного лиганда от его конформации в составе комплекса;
- отличие конформации свободного белка от его конформации в составе комплекса;
- ограничение применимости данного подхода для лигандов с определенными свойствами.

Анализ показал, что последняя причина является наиболее существенным фактором, приводящим к ошибке. В предлагаемом нами подходе отсутствует учет электростатических взаимодействий полярных лигандов с белком при первичном отборе гипотез комплексов. Этот вывод основан на примерах в комплексах, показавших неудовлетворительные результаты в тестах *A*, *C* и *D*, в энергию образования которых основной вклад вносят именно электростатические взаимодействия (табл. 5).

Таким образом, рассматриваемый подход достаточно эффективен для нахождения места связывания лиганда в тех случаях, когда большая часть молекулы лиганда является объемной, жесткой и неполярной структурой, а место связывания лиганда в структуре белка-мишени имеет вид "кармана" или полости. Подход практически не пригоден для поиска места связывания небольших или гибких полярных лигандов. Определить точную ориентацию лиганда в месте связывания также не представляется возможным без использования дополнительных методов оптимизации структуры комплекса. Такие ограничения на применение данного подхода обусловлены первичным отбором гипотез комплексов только по величине контактирующих молекулярных поверхностей, без учета дальнедействующих видов межмолекулярных взаимодействий. Использование такого упрощения в программе DockSearch оправдано, так как продиктовано необходимостью иметь быстрый и эффективный алгоритм

скрининга низкомолекулярных баз данных, когда место связывания лиганда уже описано. Однако при поиске места связывания лиганда это упрощение оказывается существенным ограничивающим фактором.

ЗАКЛЮЧЕНИЕ. 1. Оценены границы применения метода молекулярного геометрического докинга для нахождения места связывания лиганда в белках с известной трехмерной структурой: метод достаточно эффективен в случае относительно больших лигандов с жесткой структурой и не применим для гибких или небольших полярных лигандов. Метод не позволяет точно определить ориентацию лиганда в месте его связывания.

2. В пределах применимости метод позволяет получить описание места связывания лиганда, удовлетворяющее требованиям поиска веществ-прототипов в химических базах данных низкомолекулярных соединений.

3. Основным фактором, ограничивающим применение метода молекулярного геометрического докинга, является принцип первичного отбора гипотез комплексов по величине площади контактирующих молекулярных поверхностей.

Следует отметить, что данный подход позволяет лишь частично решить проблему поиска места связывания лиганда. Вопрос о создании простых, эффективных и универсальных методов нахождения места связывания низкомолекулярных лигандов в белковых молекулах остается открытым.

Данная работа была частично поддержана Российским Фондом Фундаментальных Исследований (гранты РФФИ 99-04-48754 и 01-04-06544) и Минпромнауки РФ (ФЦНТП «Исследования и разработки по приоритетным направлениям развития науки и техники гражданского назначения» направление «Технологии живых систем» подпрограмма "Создание новых лекарственных средств методами химического и биологического синтеза").

Авторы выражают благодарность фирмам Tripos GmbH и NC Group за научную и техническую поддержку.

ЛИТЕРАТУРА

1. Burkhard P., Taylor P., Walkinshaw M.D. (1998) J. Mol. Biol. **277**, 449-66.
2. Shi L.M., Myers T.G., Fan Y., O'Connor P.M., Paull K.D., Friend S.H., Weinstein J.N. (1998) Mol. Pharmacol. **53**, 241-51.
3. Zheng W., Cho S.J., Waller C.L., Tropsha A. (1999) J. Chem. Inf. Comput. Sci. **39**, 738-46.
4. Shoichet B.K. Docking ligands to proteins. In book "Protein Structure Prediction - A Practical Approach" Ed. M.J.E. Sternberg, Oxford University Press, London, 1996, 263-290.
5. DockSearch (<http://imgdd.ibmh.msk.su/lab/originalsoftware/DS/DS.htm>)
6. Белкина Н.В., Скворцов В.С., Иванов А.С., Арчаков А.И. (1998) Вопр. Мед. Хим. **44**, 464-473.
7. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. (2000) Nucleic Acids Research, **28**, 235-242. WWW address: <http://www.rcsb.org/pdb/>
8. Sybyl 6.5, Tripos Inc., 1699 South Hanley Road, St. Louis, Missouri, 63144, USA.
9. Ivanov A.S., Rumjantsev A.B., Skvortsov V.S., Archakov A.I. (1997) Comput. Appl. Biosci., **13**, 111-113.
10. Ivanov A.S., Rumjantsev A.B., Skvortsov V.S., Archakov A.I. (1996) J. Chem. Inf. Comput. Sci., **36**, 660-663.
11. Onix (<http://imgdd.ibmh.msk.su/lab/originalsoftware/ONIX95.htm>).

Поступила 31.10.01

**COMPUTER AIDED DRUG DESIGN BASED ON STRUCTURE OF MACROMOLECULAR
TARGET: I. SEARCH AND DESCRIPTION OF LIGAND BINDING SITE IN TARGET
PROTEIN.**

A.S.Ivanov, A.V.Dubanov, V.S.Skvortsov, A.I.Archakov

Orekhovich Institute of Biomedical Chemistry of RAMS, Pogodinskaya str. 10, Moscow, 119832,
Russia, Fax: 007 (095) 245-0857; e-mail: ivanov@ibmh.msk.su)

The applicability of molecular docking method for finding of ligand binding site in target protein was tested. The basic principle of tested approach is the generation of hypotheses of protein-ligand complexes by molecular docking of small ligand to all surface of protein. The subsequent scoring of these hypotheses utilizes the values of contact surfaces and complex formation energy. The docking procedure was executed using the original software DockSearch running on PC Pentium or SGI computers. A set of hypotheses of probable ligand positions on the protein surface was created and evaluated with the help of DockSearch. Energy minimization of molecular complexes was done using Sybyl 6.5 (Tripos Inc.) running on SGI server Origin200. The final set of the best hypotheses of complexes were selected by the values of contact surfaces and complex formation energy. The applicability and the limits of this approach was tested using known 3D structures of different proteins in free state and in complexes with ligands. For most target proteins the method allows to find the ligand binding site correctly. The accuracy of description of ligand binding site is adequate for subsequent searching of lead compounds by database mining. The reasons of some negative results obtained in testing of this approach are also discussed.

Keywords: computer aided drug design, molecular modelling, structure of target protein, molecular docking, molecular graphics, small ligands