

## Интегральная платформа «От гена до прототипа лекарства» *in silico* и *in vitro*

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

**АЛЕКСЕЙ СЕРГЕЕВИЧ ИВАНОВ** — доктор биологических наук, профессор, заведующий лабораторией молекулярно-графического конструирования лекарств ГУ НИИ биомедицинской химии им. В.Н. Ореховича РАМН (ГУ НИИ БМХ РАМН). Область научных интересов: биоинформатика (компьютерное молекулярное моделирование, молекулярный докинг, поиск новых мишеней для создания лекарств, компьютерное конструирование лекарств), биохимия (белок-лигандные взаимодействия, олигомеризация белков), биофизика (биологические и модельные мембраны).

**АЛЕКСАНДР ВЛАДИМИРОВИЧ ВЕСЕЛОВСКИЙ** — доктор биологических наук, ведущий научный сотрудник лаборатории молекулярно-графического конструирования лекарств ГУ НИИ БМХ РАМН. Область научных интересов: биофизика, компьютерное молекулярное моделирование, компьютерное конструирование лекарств, белок-лигандные взаимодействия, олигомеризация белков.

**АЛЕКСАНДР ВЯЧЕСЛАВОВИЧ ДУБАНОВ** — кандидат биологических наук, ведущий научный сотрудник лаборатории молекулярно-графического конструирования лекарств ГУ НИИ БМХ РАМН. Область научных интересов: фармакология, биоинформатика (поиск новых мишеней для создания лекарств, компьютерное конструирование лекарств).

**ВЛАДЛЕН СТАНИСЛАВОВИЧ СКВОРЦОВ** — кандидат биологических наук, старший научный сотрудник лаборатории молекулярно-графического конструирования лекарств ГУ НИИ БМХ РАМН. Область научных интересов: биоинформатика (компьютерное молекулярное моделирование, молекулярный докинг), создание компьютерных программ для решения задач в области биоинформатики.

**АЛЕКСАНДР ИВАНОВИЧ АРЧАКОВ** — академик Российской академии медицинских наук, доктор биологических наук, профессор, директор ГУ НИИ БМХ РАМН. Область научных интересов: протеомика, геномика, биоинформатика, биохимия (цитохром P450, белок-лигандные взаимодействия, олигомеризация белков).

119121 Москва, Погодинская ул., 10, ГУ НИИ БМХ РАМН, тел. (495)246-36-93, факс (495)245-08-57,  
E-mail alexei.ivanov@ibmc.msk.ru

### Введение

Общий путь создания нового лекарства («от идеи до аптеки») включает семь основных этапов: 1) выбор болезни, для лечения которой создается лекарство, 2) выбор молекулярной мишени для действия лекарства, 3) нахождение базовой структуры нового лекарства, 4) оптимизация базовой структуры, 5) доклинические испытания, 6) клинические испытания, 7) производство препарата. В течение длительного периода основными источниками соединений для экспериментального тестирования на физиологическую активность были живые организмы (растения, животные, микроорганизмы) и классический химический синтез. При экспериментальном тестировании подавляющее большинство соединений отбрасывалось как неперспективные прототипы лекарств из-за низкой целевой активности или полного ее отсутствия, высокой токсичности, канцерогенности, сложности синтеза и т.д. И только одно из 100 тысяч исследованных соединений может стать препаратом с выходом на фармацевтический рынок. Общие затраты времени и денег на создание нового препарата могут достигать 12–15 лет и более 800 млн. долларов [1]. Сокращение времени и финансовых расходов на последних этапах (клинические испытания и производство препарата) фактически невозможно по причине строгих государственных стандартов и законов. Поэтому главные усилия разра-

ботчиков лекарств, направленные на повышение эффективности процесса создания новых препаратов, касаются его более ранних стадий.

В последнее десятилетие современные компьютерные технологии, биоинформатика и новые экспериментальные методы в области медицинской химии были объединены в общий подход, часто называемый «рациональное конструирование лекарств», обеспечивающий ускорение и оптимизацию процесса нахождения новых биологически активных соединений — базовых структур новых лекарств. Кроме того, экстенсивный процесс расшифровки геномов различных организмов, включая человека, позволил использовать методы биоинформатики для предсказания ряда новых потенциальных мишеней для действия лекарств.

Наконец, интеграция перечисленных выше подходов привела к созданию единой платформы «От гена до прототипа лекарства», которая охватывает первые 4 из 7 этапов создания новых лекарств (рис. 1). Многие из блоков этой платформы включают компьютерное моделирование, виртуальный скрининг и предсказание свойств веществ, и эти подходы позволяют во много раз снизить затраты времени и денег. Методы компьютерного конструирования лекарств могут сократить на два порядка количество соединений, которые необходимо синтезировать и проверить на наличие целевой биологической активности. Однако эти подходы не могут полностью заменить реальные экс-

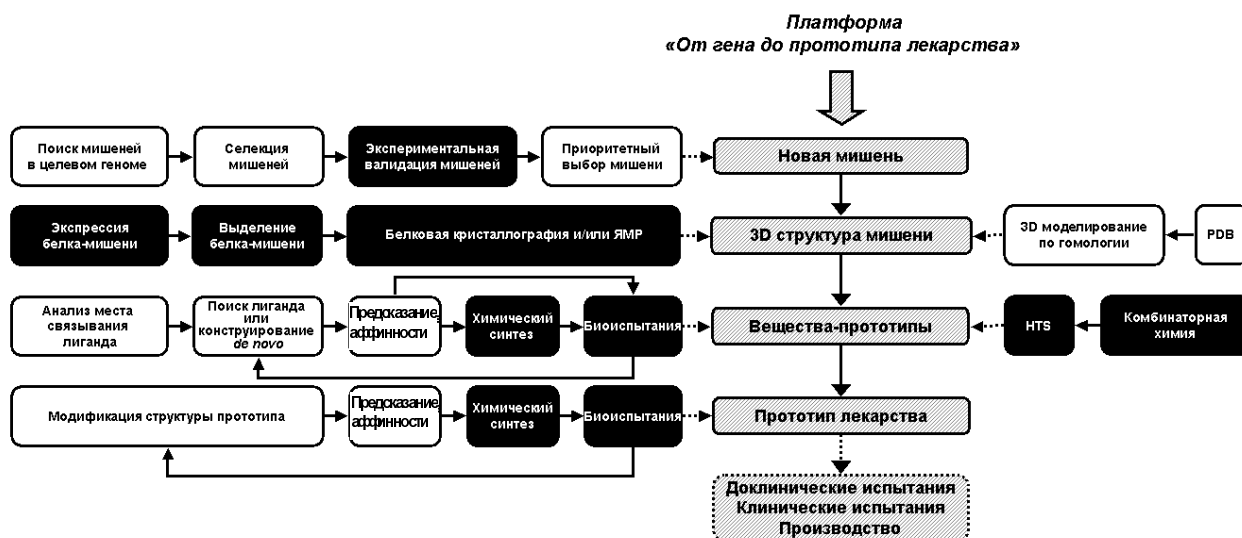


Рис. 1. Общая схема платформы «От гена до прототипа лекарства».

Экспериментальные блоки показаны черным цветом, компьютерные — белым

перименты. Цель компьютерных методов — это генерация высоковероятных гипотез о новых мишенях для действия лекарств и лигандах, взаимодействующих с мишенью (основа для будущих лекарств), которые должны быть проверены позже в прямых экспериментах.

Данный обзор посвящен описанию главных стратегий создания платформы «От гена до прототипа лекарства», основных компьютерных и экспериментальных подходов, используемых в различных блоках платформы, и их взаимосвязи. Последняя часть обзора содержит примеры выполнения некоторых этапов данной платформы.

### Информационные ресурсы и вычислительные средства биоинформатики

#### Геномные базы данных

Для автоматической селекции потенциальных мишеней для действия лекарства могут быть использованы следующие базы данных.

- Базы с первичными геномными данными (полные геномы, плазмиды и последовательности белков): NCBI GenBank [2], EBI-EMBL [3], DDBJ [4].

- Базы данных с аннотированными белковыми последовательностями: Swiss-Prot и TrEMBL [5], PIR [6].

- Базы данных с результатами кросс-геномных сравнительных анализов: COG/KOG (Clusters of Orthologous groups of proteins) [7], KO (KEGG orthologies) [8].

- Базы данных, содержащих информацию о белковых семействах и классификацию белков: Pfam и SUPFAM [9], TIGRFAMs [10].

- Веб-сервисы для кросс-геномного анализа (включая средства селекции мишеней для создания антимикробных средств): CMR [11], MBGD [12].

- Базы данных о белок-белковых взаимодействиях (как экспериментальные, так и предсказанные данные): DIP [13], BIND [14], InterDom [15], FusionDB [16].

- Базы данных метаболических и регуляторных путей: KEGG [8], PathDB [17].

- Базы данных о трехмерной структуре белков: Protein Data Bank (PDB) [18], вторичные базы и их компиляции, например PDB-REPRDB [19].

- Интегрированные ресурсы: PEDANT [20].

Полную информацию о доступных в настоящее время через Интернет геномных базах данных можно найти в специальном выпуске журнала Nuclear Acid Research [21]. Список баз данных, пригодных для поиска новых мишеней для создания антимикробных средств, опубликован в [22].

#### Молекулярные базы данных

Решающее значение в создании новых лекарств имеет доступ к базам данных, которые содержат пространственные структуры макромолекул и низкомолекулярных соединений. Сведения о структурах белков депонируются в банке PDB [18]. Что касается структур низкомолекулярных веществ, то их можно найти в многочисленных базах данных. Кристаллографические структуры собраны в Cambridge Structural Database (CSD) [23]. Эта база данных является самой большой и уникальной коллекцией экспериментальных данных о трехмерных структурах низкомолекулярных соединений (более 300000 веществ). Огромный размер CSD можно считать некоторым недостатком, так как она, с одной стороны, включает много структур, которые не привлекательны для проектов по созданию новых лекарств (например, много металлоорганических структур), а с другой стороны, она содержит много веществ, которые фактически отсутствуют и должны быть синтезированы, если они выбраны в качестве перспективных прототипов. Фармацевтические ком-

пании часто предпочитают использовать свои собственные корпоративные базы данных, собранные из различных открытых и корпоративных источников, а также коммерчески доступные базы, например NCI [24], MDDR [25] и СМС [26].

На практике более удобно использовать специализированные небольшие базы данных, содержащие структуры коммерчески доступных образцов химических соединений, которые можно быстро приобрести для выполнения экспериментальных проверок. Использование таких баз данных позволяет пропускать этап химического синтеза. Примеры таких баз данных — ASINEX [27], ChemBridge [28], MayBridge [29] и др. В некоторых случаях эти базы данных могут быть исследователями заменены собственной базой данных, содержащей виртуальные структуры, сконструированные по специальной компьютерной программе, например CombiFlexX из пакета Sybyl [30].

#### **Программные средства и вычислительная техника**

Представляем специализированные программные средства, с помощью которых может быть выполнен автоматизированный поиск потенциальных мишеней для создания новых антибактериальных средств.

- Оригинальные академические программы, разработанные исследователями для себя и применительно к конкретным задачам. Этот путь был характерен для ранних стадий эволюции методов автоматического поиска мишеней [31, 32].

- Коммерческие программные средства и базы данных, разработанные для решения рутинных биоинформационных задач. Например, это пакет программ GeneData Phylosopher [33], содержащий модуль для поиска биомишеней с целью создания новых антимикробных средств.

- Пользовательские сценарии (scripts, macros), написанные на программных языках-интерпретаторах типа Perl или Python с использованием бесплатных библиотек, классов и функций для решения задач биоинформатики, таких как BioPerl и BioPython. С нашей точки зрения, этот путь наиболее предпочтителен, так как делает возможным поиск мишеней в строгом соответствии с требованиями проекта. Все расчеты и поиск мишеней могут быть выполнены как локально, так и с использованием Веб-сервиса, например Entrez Utilities [34] и NCBI BLAST service [35]. Такой подход представляет разумную альтернативу коммерческим программам, поскольку обеспечивает необходимую гибкость по сравнению с другими подходами.

- Веб-сервисы для кросс-геномных анализов, такие как Comprehensive Microbial Resource [11]. Этот Веб-сервис является преимущественно демонстрационным, однако пригоден для предварительного поиска биомишеней. Во многих случаях кросс-геномный анализ может быть выполнен локально, например, когда анализируются неопубликованные геномы.

Компьютерное конструирование лекарств требует различных вычислительных методов, реализованных в специализированных компьютерных программах, вводимых в высокопроизводительные компьютеры. До недавнего времени подобные вычислительные операции выполнялись преимущественно на мощных мультипроцессорных серверах и графических станциях под управлением UNIX. В настоящее время для такого

вида вычисления становятся все более популярными кластеры компьютеров под управлением LINUX. Наиболее подходят для компьютерного конструирования лекарств два основных коммерческих пакета программного обеспечения — SYBYL (Tripos Inc.) [31], Insight II и Quanta (Accelrys) [36], которые позволяют осуществить почти полный процесс проектирования базовых структур лекарств. Мы предпочитаем использовать SYBYL, и представленная в данном обзоре наша работа основана на использовании различных модулей этого пакета. Для решения ряда конкретных задач мы также применяем несколько дешевых академических программ, например AMBER [37], и условно бесплатное программное обеспечение или свободно распространяемые программы (например, довольно популярная программа для моделирования молекулярной динамики GROMACS [38]).

#### **Поиск мишеней для создания новых противоионфекционных средств на основе анализа геномных данных**

##### **Биоинформационные подходы**

Первый шаг в создании новых противоионфекционных средств — это, как правило, автоматический поиск биологических мишеней, поскольку большинство требований к целевому белку-мишени в патогенных микробах может быть легко формализовано. Эти требования можно сформулировать, исходя из желательных свойств «идеального» противоионфекционного препарата (табл. 1) [32]. Согласно современной концепции выбора мишени все белки, закодированные в целевом геноме микроба, должны быть исследованы на соответствие этим требованиям, и лучшие кандидаты могут быть отобраны как потенциальные биологические мишени.

Поиск мишеней обычно разделяется на два этапа: 1) предварительный автоматический поиск потенциальных мишеней; 2) конечная ручная селекция мишеней, основанная на данных о свойствах белков. Первый этап сокращает число потенциальных мишеней с тысяч (типичный геном бактерии содержит около 1500—4000 белков) до десятков (порядка 10—20). На втором этапе выбираются наиболее перспективные мишени и далее осуществляется их приоритизация. Мы рекомендуем пошаговую селекцию мишеней [32] вместо схемы одновременных их оценок на соответствие медико-биологическим требованиям [31], так как значимость ряда критериев может варьировать в различных исследованиях.

##### **Основные критерии селекции мишеней для создания антимикробных средств**

*Спектр действия и селективность белка-мишени.* Каждый белок из микробного генома может быть оценен на соответствие медико-биологическим требованиям путем сравнительного анализа этого белка и белков-гомологов в геномах других организмов. Обычно из группы анализируемых микробных геномов один наиболее изученный геном выбирается в качестве целевого, а другие геномы разделяются на две группы: 1) геномы, в которых присутствие белка-мишени (или ее гомолога) необходимо, а вся группа (различные штаммы целевого вида и другие патогенные виды) отражает спектр антимикробного действия будущего лекарства; 2) геномы, в которых присутствие

Таблица 1

Требования к «идеальному» антимикробному средству и его молекулярной мишени

Препарат	Мишень
<b>Биомедицинские требования</b>	
Эффективно подавляет рост и размножение микроорганизмов	Мишень важна для роста и размножения микроорганизмов
Летальное действие на патоген	Мишень важна для выживания микроорганизма
Имеет нужный антимикробный спектр	Имеется у всех видов и штаммов целевых микроорганизмов
Селективен:	
минимальная токсичность для организма-хозяина	Мишень отсутствует у хозяина (человек)
минимальное действие на нормальную микрофлору	Отсутствует у микроорганизмов-симбионтов человека
Низкий риск появления лекарственной резистентности	Структурно консервативна у всех штаммов целевых микроорганизмов
<b>Технологические требования</b>	
Возможность конструирования лекарств на основе структуры белка-мишени	Доступна (известна или может быть сконструирована) трехмерная структура мишени
Определен механизм действия	Известна функция мишени

мишени (или ее гомолога) нежелательно, а вся группа (геномы человека, млекопитающих, бактерий-симбионтов человека) отражает селективность антимикробного действия будущего лекарства. Белки целевого генома, которые имеют гомологи в первой группе и не имеют их во второй, составляют первичный список потенциальных мишеней. Для подобного отбора могут быть использованы карты кросс-геномной гомологии, а также базы данных генов-ортологов [22]. Однако нахождение похожих белков в геномах различных видов не гарантирует их идентичность в плане специфичности будущего лекарства, так как на данном этапе не определяется место связывания лекарства в структуре белка-мишени.

**Пространственная структура белка-мишени.** Наиболее рациональный путь нахождения соединения-прототипа нового лекарства с новым механизмом действия — это компьютерное конструирование лекарства на основе пространственной структуры белка-мишени (Structure-Based Drug Design, SBDD). Для реализации данного подхода необходима трехмерная структура мишени (полученная экспериментально или компьютерная модель). Возможность моделирования структуры белка по гомологии зависит от наличия близкого гомолога (образца) с известной трехмерной структурой в базе данных PDB. Требуется не менее 40% идентичности последовательностей моделируемого белка и известного гомолога.

Ранее было показано, что наличие пространственной структуры белка-мишени или его близкого гомолога является очень жестким ограничением, которое может отсеять большую часть потенциальных мишеней, удовлетворяющих другим требованиям [32]. Нахождение близкого гомолога с известной структурой делает более простыми, быстрыми и дешевыми последующие компьютерные и экспериментальные шаги в выборе белка-мишени, так как может быть использована существующая практика экспрессии белков, очистки, кристаллизации и т.д.

Все критерии относительно спектра действия и селективности мишени, а также ее пространственной

структуры, легко формализуются и вычисляются на основе анализа сходства аминокислотной последовательности. Поэтому эти критерии используются для автоматического скрининга мишеней на первом этапе создания лекарства.

**Белок-белковые взаимодействия.** Помимо технологии создания лекарства, основанной на подборе лиганда, специфически взаимодействующего с молекулой-мишенью, в последнее время большое внимание уделяется разработке новых антимикробных средств, действующих как ингибиторы белок-белковых взаимодействий [39, 40]. Интерфейс взаимодействия белков друг с другом наиболее консервативен, чем реакционная способность активных центров ферментов и других функционально важных составляющих биомолекул. Поэтому возможно значительное снижение риска появления лекарственной резистентности, обусловленной мутациями белка-мишени. Считается, что предпочтительно выбирать в качестве мишеней ферменты, функционирующие в гомо-, гетеро- или олигомерной форме. Жизнеспособная мутация в области контакта между субъединицами белка-фермента может происходить крайне редко в случае второй синхронной мутации, компенсирующей искажение интерфейса.

Существует ряд различных биоинформационных методов (*in silico*) для предсказания белок-белковых взаимодействий. Обычно эти методы используются для аннотации новых генов, кодирующих белки. Выявление взаимодействующих белков базируется на результатах кросс-геномных анализов, для чего могут быть использованы как давно известные подходы, так и принципиально новые. Рассмотрим эти подходы.

**Аннотация по сходству** основана на том, что ряд белков из целевого организма могут показывать значительное сходство с белками, вовлеченными в белок-белковые взаимодействия у других организмов. Большинство аннотаций, указывающих на участие белка в белок-белковых взаимодействиях, были получены именно таким способом.

Другой подход — *использование филогенетических паттернов (шаблонов)* для нахождения белков, вовле-

ченных в общие метаболические пути или в образование комплексов (функционально связанные белки), основан на предположении, что такие белки совместно присутствуют (или отсутствуют) в ряде организмов. Филогенетический паттерн описывает встречаемость конкретного белка в ряде организмов. Если два или более белков имеют идентичные филогенетические паттерны, можно предположить, что эти белки функционально связаны [41]. Этот подход дает только общую информацию о функциональной взаимосвязи между белками, но не о предполагаемых их комплексах.

Метод *хромосомных соседей* заключается в следующем. Если в ряде геномов гены, кодирующие конкретные белки, расположены рядом (соседи) на хромосомной последовательности ДНК, то эти белки могут быть функционально связаны [22, 41]. Например, было показано, что выявление хромосомных соседей корректно идентифицирует функционально связанные восемь ферментов биосинтеза аргинина у *Mycobacterium tuberculosis* [23]. Этот метод наиболее применим для анализа геномов прокариот, где имеются общие опероны<sup>1</sup>, хотя он также применим и для анализа геномов эукариот [6, 41].

Еще один метод — *слияние доменов (метод Розеттского камня, Rosetta Stone)* основан на предположении, что в результате эволюции гены соседних функционально связанных белков могли слиться с образованием одного двухдоменного белка. Возможна и обратная ситуация: два или более белков в одном организме функционально связаны и соответствуют разным доменам одного белка в другом организме [41]. Таким образом, если в ряде геномов находят гены, кодирующие белки Розеттского камня, то эти белки могут быть функционально связаны.

Метод *коррелированных мутаций* исходит из того, что если конкретная пара аминокислотных остатков в отдельных белках важна для образования белок-белкового комплекса, то будет наблюдаться корреляция мутаций данных остатков. Таким образом, предсказание формирования белковых комплексов может быть основано на нахождении таких мутаций.

Отметим еще метод *взаимодействующих доменов*. Мотивы в последовательностях взаимодействующих доменов могут быть найдены в аннотированных белках. Вероятность взаимодействия для каждой пары мотивов может быть использована для поиска взаимодействующих пар в других белках, в которых находят эти мотивы.

Ряд методов, упомянутых выше, могут предсказывать образование гетеро-олигомерных комплексов белков. Нахождение взаимодействующих доменов может также указывать на возможность формирования гомо-олигомерных комплексов, хотя обычно такое заключение делается с помощью методов анализа сходства последовательностей.

Методы, описанные выше, обычно дают некоторое количество ложных положительных (также как и ложных отрицательных) предсказаний. Для повышения достоверности предсказаний эти методы рекомендуются использовать совместно [41, 42]. Неполный анализ геномов и ошибки в распознавании генов также могут

приводить к ложным предсказаниям белок-белковых взаимодействий. Конечно, результаты компьютерных предсказаний должны быть проверены в реальных экспериментах.

*Конечная селекция и приоритезация потенциальных мишеней.* В настоящее время этот самый критический этап процесса создания новых лекарств может быть выполнен только вручную. На данном этапе производится оценка возможных финансовых и временных затрат и ряда плохо формализуемых критериев типа функции белка и его метаболической роли. Преимущественно это функция белка и его метаболическая роль.

К предпочтительным мишеням относят ферменты клеточной стенки, компоненты синтеза витаминов, компоненты систем транскрипции и трансляции. Предпочтительны также ферменты, катализирующие ключевые стадии метаболизма. Что касается структурных и транспортных белков, а также ферментов альтернативных метаболических путей, то их лучше исключить из списка потенциальных мишеней [31]. С технологической точки зрения более предпочтительны водорастворимые белки по сравнению с мембранными. Информация о возможной растворимости белка и его внутриклеточной локализации может быть найдена в базах данных аннотированных геномов.

Существует три основных подхода для предсказания возможной внутриклеточной локализации белка [43]: 1) прямое предсказание на основании сходства последовательностей анализируемого белка и уже аннотированных белков; 2) предсказание на основе филогении (концепция эволюции клеточных компонент); 3) предсказание на основе распознавания мотивов сортирующих сигнальных последовательностей. Этот подход наиболее точен, хотя наши знания о сортирующих сигналах пока ограничены. Комбинация всех перечисленных выше подходов является оптимальным путем.

В последнее время активно разрабатываются высокопроизводительные технологии экспериментальной проверки — валидации потенциальных мишеней (high-throughput target validation). Они позволяют проводить проверку примерно 10% всех белков, кодируемых типичным бактериальным геномом [44].

#### Экспериментальные технологии валидации потенциальных мишеней

В настоящее время число известных биологических мишеней, на которые действуют известные лекарства, равно нескольким сотням, в то время как число новых перспективных мишеней, которые предсказываются на основе геномных данных, может достигать нескольких тысяч. И уже это свидетельствует о том, что очень трудно решить, какая новая мишень является «правильной».

В фармацевтической индустрии выбор новой мишени — очень ответственная задача, поскольку затраты на поиск прототипов нового лекарства и проведение доклинических и особенно клинических испытаний крайне велики. Именно поэтому выбор «правильной» мишени становится критическим шагом в создании новых лекарств [45], и вероятность финансовых потерь должна быть по возможности уменьшена на

<sup>1</sup> Генетическая функциональная единица, состоящая из регуляторных и структурных генов, функционирующих координированно.

ранних этапах осуществления проекта. В такой ситуации без сомнения возрастает роль валидации потенциальных мишеней. Эта процедура — самое узкое место в реализации платформы «От гена до прототипа лекарства», так как конечный выбор мишени определяет все последующие этапы создания лекарства.

Основной задачей валидации мишеней является максимальное сокращение их числа (по крайней мере до десятка) и получение дополнительной информации для приоритизации оставшихся мишеней. Конечный выбор «правильной» мишени может быть сделан только супервизором проекта с учетом всех факторов, обеспечивающих успех реализации проекта.

Разрабатываемые в настоящее время технологии экспериментальной валидации потенциальных мишеней довольно разнообразны — от экспериментов *in vitro* до моделей на животных.

### Протеомные методы валидации потенциальных мишеней

Протеомные методы [46] незаменимы для предварительной валидации потенциальных мишеней, позволяющей исключить из списка явно «неправильные» мишени. Первоначально необходимо проверить, существует ли реально в целевом организме белок, выбранный в качестве потенциальной мишени чисто теоретически, на основе геномной информации. Для этой цели полностью подходит стандартная схема протеомного картирования белков: 1) получение биологического образца; 2) экстракция белков; 3) разделение белков (двумерный электрофорез или хроматография); 4) трипсинизация белков (частичный протеолиз белков под действием трипсина); 5) масс-спектрометрический анализ; 6) биоинформационная идентификация белков на основе сравнительного анализа

экспериментальных и теоретических масс-спектров. В случае создания антимикробного средства с широким спектром действия необходимо проверить экспрессию целевого белка у всех микроорганизмов целевой группы.

В результате предварительной валидации белков-мишеней должны быть получены протеомные карты для различных штаммов целевых микроорганизмов. Если обнаружится вариабельность экспрессии белка, то такой белок может быть исключен из списка потенциальных мишеней по соображению возможного проявления лекарственной резистентности.

Как было отмечено выше, в настоящее время активно обсуждается вопрос о создании новых лекарств, действующих на уровне белок-белковых взаимодействий [40]. При этом наиболее часто рассматриваются варианты ингибиторов этих взаимодействий и, в первую очередь, ингибиторов димеризации [39, 47]. Новейшие протеомные методы пригодны для анализа белок-белковых взаимодействий [48] и использования этой информации для выбора белков-мишеней, работающих в составе надмолекулярных комплексов [49].

### Геномные методы валидации потенциальных мишеней

Последующая валидация мишеней, успешно прошедших проверку протеомными методами, может быть выполнена различными геномными методами [49, 50]. Большинство этих подходов основано на одной общей идее — проверить функциональную значимость потенциальной белковой мишени путем остановки ее экспрессии на одном из этапов «от гена до белка» (рис. 2). Естественно, наиболее эффективны геномные методы блокирования экспрессии целевого белка в быстро растущих клетках. В тех случаях, когда потенциальная мишень существует длительное время

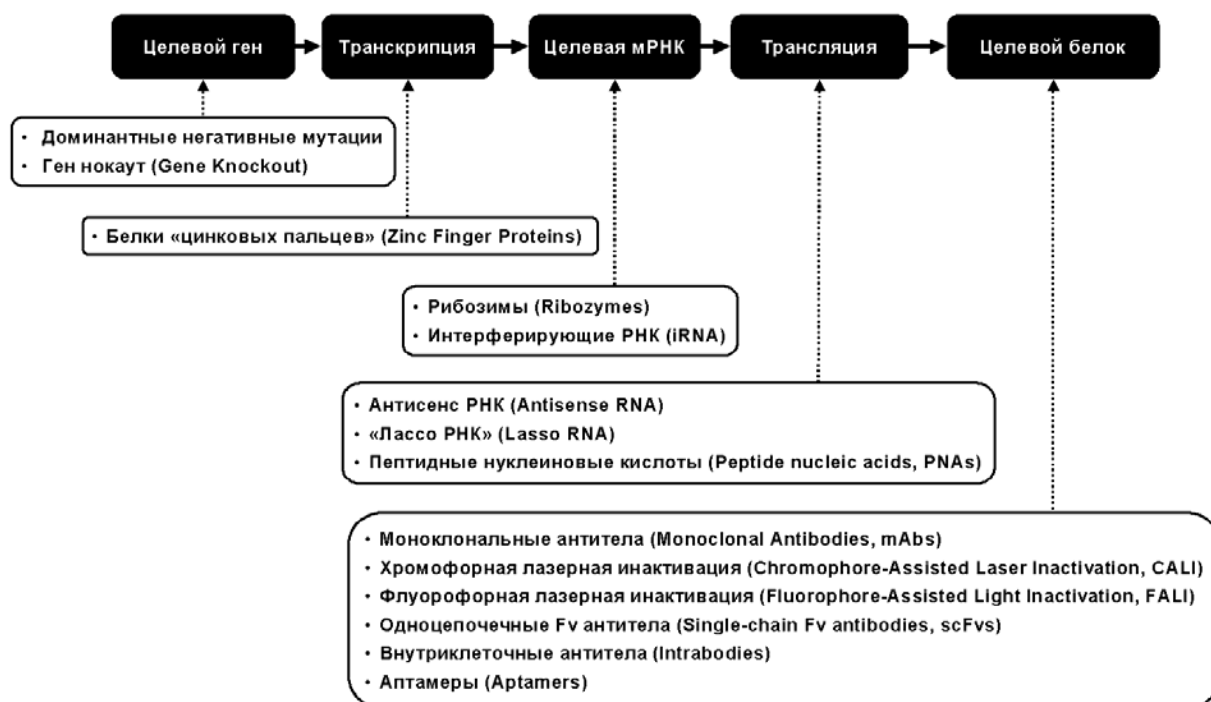


Рис. 2. Блокирование экспрессии целевого гена на различных этапах пути «от гена до белка»

и ее обновление происходит крайне медленно, более предпочтительны методы прямой инактивации белка (см. ниже разд. «Инактивация целевого белка»).

#### Инактивация целевого гена

В этих целях используют два метода. Метод доминантных негативных мутаций основан на создании таких мутаций, которые не только блокируют функцию мутантного белка, но и одновременно блокируют функцию экспрессируемого нормального белка [51]. Этот подход наиболее эффективен в случае работы белка в составе мультимерных комплексов.

Другой подход — нокаут гена (Gene Knockout) заключается в создании методами генной инженерии мутантных организмов, которые несут один или более генов с полностью нарушенной функцией [52].

#### Подавление транскрипции гена

Данный метод основан на действии белков «цинковых пальцев» (Zinc-Finger Proteins). Эти белки относятся к группе транскрипционных факторов и составляют большое семейство (более 1000 последовательностей) [53]. В их структуре присутствует несколько характерных мотивов — «цинковых пальцев» в виде короткого двутяжевого антипараллельного  $\beta$ -листа и  $\alpha$ -спирали, связанных с ионом цинка. Белки охватывают «цинковыми пальцами» ДНК с высоким уровнем узнавания между  $\alpha$ -спиралью и ДНК (рис. 3). Селективность данных белков изменяется путем вариации числа «цинковых пальцев» и мутаций в них, что вызывает блокирование трансляции их генов.

#### Разрушение целевой мРНК

Разрушение целевой мРНК способны вызывать рибозимы. Рибозимы (Ribozymes) — это специфические молекулы РНК (олигонуклеотиды), которые могут катализировать различные биохимические реакции без участия каких-либо белков-ферментов [54]. Известно много различных природных и искусственных

рибозимных мотивов. Самый маленький мотив «головка молотка» может быть использован для конструирования и синтеза препаратов, специфически разрушающих целевые мРНК. Этот подход особенно перспективен для валидации мишеней [55].

Разрушение целевой мРНК может происходить также под действием малых интерферирующих РНК (Small interfering RNA, siRNA). Интерференция РНК — одно из новых революционных открытий [56]. Этот эффект основан на действии siRNA (двунитевая молекула РНК длиной примерно в 20 азотистых оснований) на целевую мРНК. Воздействие осуществляется следующим образом. Когда молекула siRNA оказывается внутри клетки, она связывается с комплексом из двух ферментов — хеликазы и нуклеазы. Далее хеликаза расплетает и разъединяет нити siRNA, и одна из нитей в комплексе с нуклеазой специфически взаимодействует с целевой мРНК.

#### Подавление трансляции

Этот метод основан на действии антисенс РНК и их аналогов. Антисенс РНК (Antisense RNA) — это олигонуклеотиды (~ 20 оснований), которые комплементарно связываются с целевой мРНК и блокируют экспрессию гена [57], не позволяя считывать информацию с мРНК. Антисенсы активно используются для валидации потенциальных мишеней [58].

Комплексы антисенс РНК с мРНК часто разрушаются специфическими РНКазами, что значительно снижает эффективность применения антисенсов. Для решения этой проблемы фирма «SomaGenics» (США) разработала подход, названный «РНК лассо» [59]. Метод основан на использовании модифицированных антисенс РНК, концы которой могут взаимодействовать друг с другом, образуя узел, что делает комплекс с целевой мРНК более прочным.

Еще одна альтернатива технологии антисенсов — пептидные нуклеиновые кислоты (Peptide nucleic

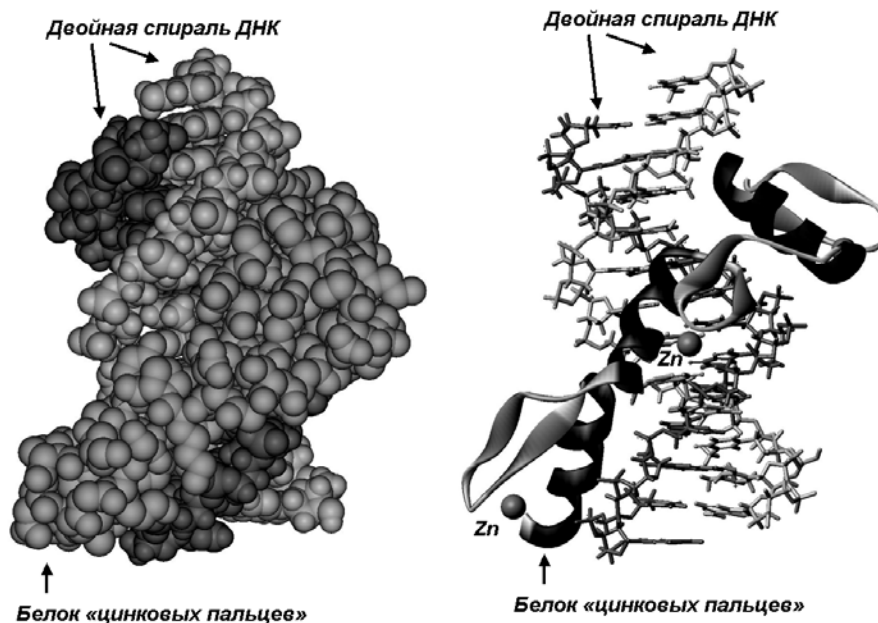


Рис. 3. Пример комплекса белка «цинковых пальцев» и ДНК.

Данные взяты из PDB, индекс файла 1MEY

acids, PNAs). PNAs — искусственные аналоги олигонуклеотидов, имеющие полиамидную основную цепь (подобно полипептидам) и боковые основания (подобно нуклеиновым кислотам) [60]. PNAs взаимодействуют подобно антисенсам с ДНК или РНК с высокой специфичностью, не подвержены ферментативной деградации и могут заменить антисенсы в валидации мишеней [61].

#### Прямая инактивация целевого белка

Моноклональные антитела взаимодействуют с целевыми белками с высокой специфичностью [62], однако только ~ 0,5% белков при этом теряют свою функцию [63]. Поэтому для валидации потенциальных белковых мишеней путем блокирования их функции антитела в чистом виде не пригодны. Они должны быть в комплексе с другими веществами, нейтрализующими функцию мишени, например с хромофорным веществом для необратимой инактивации целевых белков при поглощении лазерного излучения.

Другой подход — хромофор-опосредованная лазерная инактивация (Chromophore-Assisted Laser Inactivation, CALI) — основан на взаимодействии целевого белка с конъюгатом моноклонального антитела и хромофора (обычно это малахитовый зеленый) с последующей инактивацией белка при поглощении хромофором лазерного излучения [64]. Этим методом можно превратить более 90% антител в средство инактивации целевых белков [63].

Флуорофор-опосредованная световая инактивация (Fluorophore-Assisted Light Inactivation, FALI) [65] является аналогом CALI, где вместо хромофора применяются флуоресцирующие вещества. Метод позволяет использовать рассеянный свет для обработки многих проб одновременно, поэтому FALI пригоден для автоматических высокопроизводительных систем скрининга.

Эффективно также действие одноцепочечных антител (Single-chain Fv antibodies, scFvs). Это искусственные малые антитела, сконструированные в виде

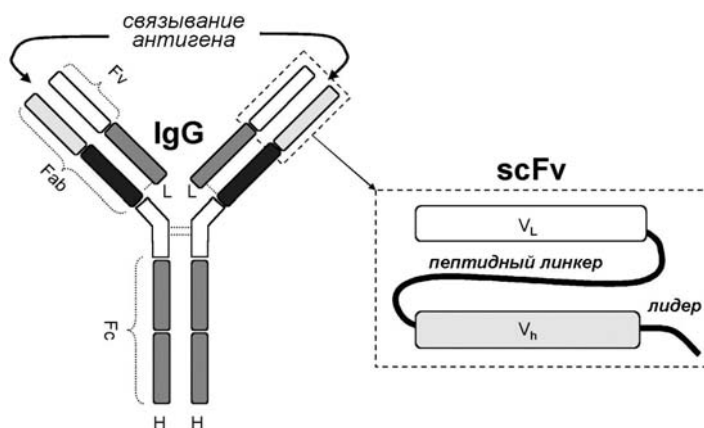


Рис. 4. Схема структуры моноклонального антитела (IgG) и одноцепочечного антитела (scFvs)

одной белковой цепи на основе двух участков из Fab фрагмента обычных антител [66] (рис. 4). Основное преимущество scFvs — их гомогенность, малая масса (~ 30 кДа), высокая проникающая способность и низкая иммуногенность. Они могут быть использованы во всех технологиях, где применяются моноклональные антитела.

Описанные выше scFv могут быть закодированы в генетическом материале — внутриклеточные антитела (Intrabodies) — и экспрессироваться непосредственно в целевой клетке [67].

Синтетические РНК или ДНК (олигонуклеотиды, 5—25 кДа) — аптамеры — подобно антителам способны связываться с целевыми белками с высокой специфичностью [68]. Аптамеры нетоксичны и неиммуногенны. Химическая стабильность аптамеров делает их крайне удобным инструментом для валидации потенциальных мишеней (рис. 5). В настоящее время аптамеры и интерферирующие РНК (siRNA) часто используются совместно, так как технологии их применения хорошо совместимы.

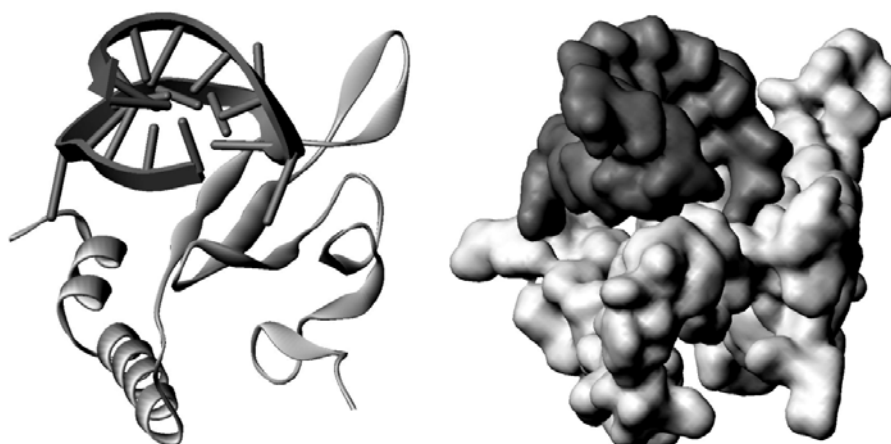


Рис. 5. Пример комплекса РНК-аптамера (темный цвет) и белка (светлый цвет).

Данные взяты из PDB, индекс файла 5MSF



### **Высокопроизводительные технологии поиска и валидации мишеней (high-throughput technologies)**

В последнее десятилетие многие фармацевтические компании занялись разработкой новых высокопроизводительных технологий поиска новых мишеней и их валидации на основе роботизированных мультипараллельных процессов [69]. Эти высокотехнологичные системы могут выполнять до тысячи операций в день. Многие из описанных выше геномных и протеомных технологий были адаптированы к условиям работы этих систем [68, 70, 71]. Разрабатываются также технологии поиска мишеней и их валидации на основе ДНК-полей, ДНК-чипов и микрожидкостных технологий [72].

### **Валидация потенциальных мишеней *in vivo***

При рассмотрении методов по экспериментальной валидации мишеней нельзя не упомянуть о методах *in vivo*. Последние достижения в геномных исследованиях показали, что небольшие многоклеточные организмы (нематоды, дрозофила, рыбка-зебра и др.) содержат гены, близкие к тем, которые ответственны за возникновение болезней у человека. Поэтому в целях валидации мишеней для действия лекарств при различных заболеваниях человека могут быть использованы разнообразные модели *in vivo* — от беспозвоночных до «гуманизированных» мышей [73]. К таким системам относятся, например, трансгенные мыши и мыши с нокаут-генами [74], рыбка-зебра (*Danio rerio*) [75], нематоды, в том числе хорошо известная *Caenorhabditis elegans* [76], плодовая мушка дрозофила (*Drosophila melanogaster*) [77], платформа VITA (Validation In vivo of Targets and Assays for antiinfectives) на мышах [78], паразитарные организмы [79].

### **Пространственная структура мишеней**

Очевидно, что пространственная структура мишеней играет ключевую роль в методах конструирования лекарств, основанных на структуре мишени (SBDD). Информация о структуре мишени может быть получена с помощью как экспериментальных, так и биоинформационных методов.

### **Экспериментальные методы расшифровки пространственных структур белков**

Для исследования пространственной структуры мишени применяют в основном два метода: рентгеноструктурный анализ (РСА) [80] и многомерный вариант ЯМР [81]. Экспериментальные измерения требуют сравнительно больших количеств чистого нативного белка, причем в случае РСА, объектом исследования которого являются кристаллы, возникает проблема кристаллизации белка.

Информация о пространственной структуре белков депонируется в базе данных PDB [18] и доступна через сеть Интернет для свободного пользования. Отметим, что на ноябрь 2005 года в банке содержалось около 34000 структур (примерно 29000 были решены методом РСА и порядка 5000 — методом ЯМР). Однако уникальных структур белков значительно меньше (порядка 6—7 тысяч). Остальные представляют собой мутанты этих белков или их комплексы с различными лигандами.

Широкомасштабные геномные исследования последних лет привели к взрывообразному увеличению количества выявленных белковых последовательностей. Однако исследовательские возможности совре-

менных методов изучения пространственных структур не поспевают за скоростью выявления новых последовательностей. К этому надо добавить, что оба основных метода хорошо зарекомендовали себя в выяснении структур водорастворимых белков, тогда как более 40% белков являются мембранными или мембранно-связанными [82]. Кроме того, метод ЯМР не способен расшифровать структуру белков массой более 30 кДа. В итоге к настоящему времени экспериментально раскрыты структуры только нескольких десятков мембранных белков, и в основном затруднения вызваны проблемами, связанными с необходимостью их экспрессии. Современные методы молекулярно-графического конструирования открывают большие возможности исследования молекулярных структур, а огромный дефицит потенциально возможных пространственных структур белков стимулирует использование компьютерных технологий трехмерного моделирования.

### **Моделирование пространственных структур белков**

В настоящее время существует единственный метод моделирования пространственной структуры белка, дающий достаточно точные модели, пригодные для технологии SBDD. Это метод моделирования структуры белка по гомологии, основанный на схожести первичных и пространственных структур исследуемого белка и его гомологов, для которых пространственная структура известна [84]. Ряд компьютерных программ предоставляют возможности такого моделирования: Sybyl [30], Insight II [36], Quanta [36], Modeller [85] и др. Мы предпочитаем использовать программу Composer [86] из пакета Sybyl (Tripos Inc.).

*Критерии для моделирования по гомологии.* Широкомасштабные эксперименты CASP (Critical Assessment of protein Structure Prediction) [87] показали, что точность модели белка, построенной по гомологии, зависит от схожести последовательностей моделируемого белка и белка-гомолога. Как правило, хороший результат получается, когда идентичность последовательностей превышает 40%. При идентичности менее 30% возникают проблемы из-за ошибок в выравнивании последовательностей белков, достигающих 20%. Число ошибок быстро растет по мере уменьшения сходства последовательностей. Это основное ограничение данного метода, поскольку не существует подходов, способных компенсировать ошибки, допущенные в выравнивании последовательностей.

*Основные этапы моделирования пространственной структуры белка по гомологии.*

1. Выбор гомологичного белка с известной пространственной структурой в PDB на основании первичной последовательности (первичной структуры) моделируемого белка.

2. Идентификация «зародышевых» аминокислотных остатков (*seed residues*) на основании гомологии последовательностей. Эти остатки используются только для предварительного пространственного выравнивания (совмещения) структур гомологов, возможные неточности могут быть исправлены на последующих этапах.

3. Пространственное выравнивание структур белков-гомологов для определения структурно-консервативных участков (*Structurally Conserved Regions, SCRs*). Для совмещения структур используются координаты только  $C_{\alpha}$

атомов аминокислотных остатков. Данная процедура позволяет выявить структурно-консервативные участки и получить их усредненную структуру для последующего моделирования целевого белка.

4. Нахождение структурно-консервативных участков в последовательности моделируемого белка. Для этого используется процедура попарного выравнивания последовательностей моделируемого белка и его гомологов.

5. Моделирование структуры консервативных участков целевого белка. На этом этапе создается основная цепь для каждого такого участка. Боковые радикалы аминокислотных остатков достраиваются на основании вторичных структур консервативных участков и положения боковых радикалов в гомологичных белках.

6. Моделирование структурно-вариабельных участков. Построение этих участков выполняется путем нахождения подходящих петель в структуре белков-гомологов и/или всех белков с известной пространственной структурой. Для каждой петли программа ищет фрагмент, который будет использован для построения конечной модели. Моделирование наружных петель требует дополнительной экспертной оценки и выбора наиболее приемлемой петли из возможных вариантов, найденных в других белках.

7. Оптимизация (уточнение) построенной модели путем ликвидации стерических противоречий и минимизации энергии молекулярной системы с последующим моделированием молекулярной динамики.

*Валидация трехмерной модели белка.* Структурная и топологическая валидация модели осуществляется по следующим критериям: а) избыточное выделение гидробоных остатков во внешнюю среду; б) избыточное

наличие заряженных остатков внутри белковой глобулы; в) положительная энергия сольватации.

При моделировании молекулярной динамики модель белка тестируется на устойчивость структуры, компактность белковой глобулы и на устойчивость вторичных структур в динамике в водном окружении.

Валидация трехмерной модели проводится по статистическим критериям. Для этого используется ряд программ, оценивающих достоверность модели на основе статистических данных, полученных при анализе пространственных структур белков из PDB: ProCheck [88], ProTable [89], MatchMaker [90], WHAT IF [91], Prosa II [92], Profiles-3D [93].

Для оценки качества модели проверяется ее соответствие известным экспериментальным данным. Это могут быть химическая модификация поверхностных аминокислотных остатков, места действия протеаз, точечные мутации, антигенные детерминанты и др.

### Стратегия компьютерного конструирования лекарств (Computer-Aided Drug Design, CADD)

Выбор метода реализации CADD определяется в зависимости от типа начальных доступных данных. Следовать стратегии CADD можно только при знании структуры мишени или набора известных лигандов к ней. В противном случае для конструирования можно использовать только экспериментальные методы, например высокоэффективный скрининг и комбинаторную химию [94].

Магистральный путь конструирования биологически активных соединений состоит из нескольких этапов (рис. 6): 1) анализ структуры мишени и выбор места

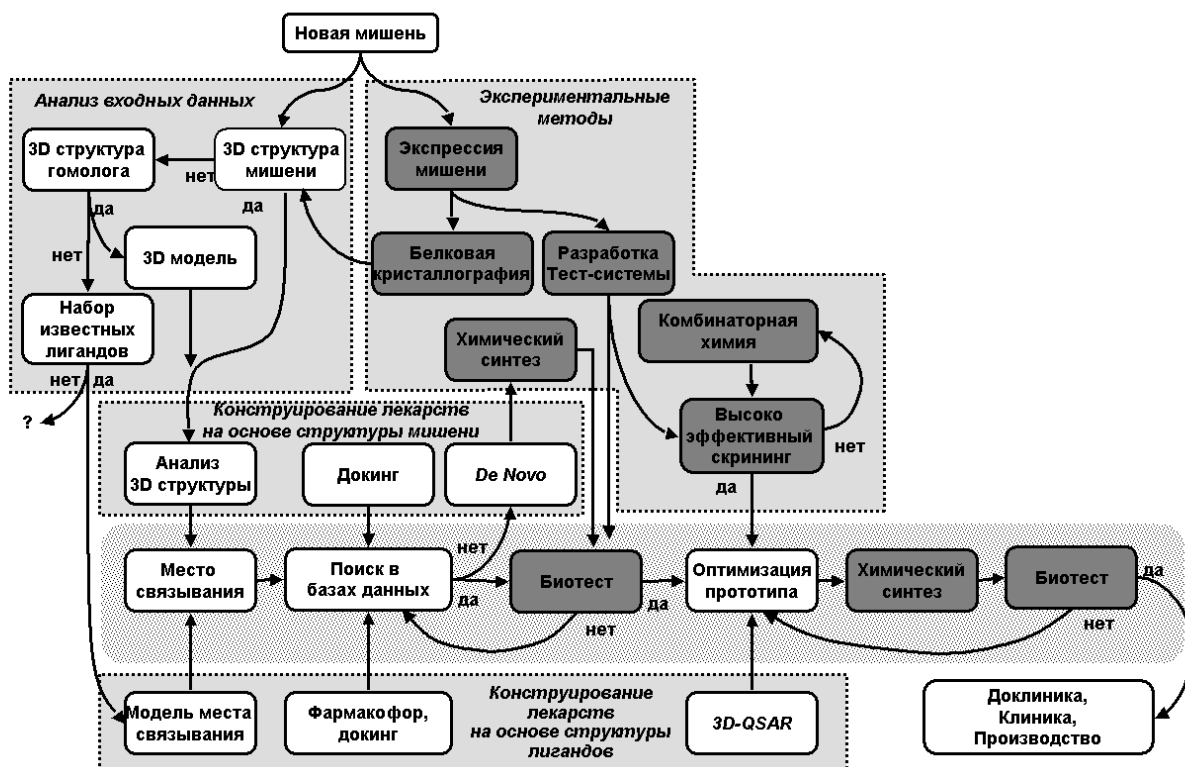


Рис. 6. Общая схема стратегии компьютерного конструирования лекарств

связывания лиганда; 2) предсказание структур базовых биологически активных соединений и их экспериментальное тестирование; 3) оптимизация базовой структуры с последующим тестированием на требуемую активность; 4) доклинические и клинические испытания выбранного биологически активного вещества.

Методы компьютерного моделирования лекарств направлены на ускорение и оптимизацию поиска новых биологически активных соединений. Цель CADD — генерация гипотез о возможных новых лигандах и их взаимодействии с мишенью. Однако эти подходы ни в коей мере не могут заменить экспериментального тестирования. Каждый этап CADD должен заканчиваться экспериментальным тестированием отобранных соединений. Если предсказанные базовые соединения не проявляют искомого уровня активности, проводится повторный этап компьютерного моделирования, при этом принимаются во внимание полученные отрицательные результаты (например, для поиска используется другая база низкомолекулярных веществ, перестраиваются фармакофорная или QSAR модели, выполняется дополнительный анализ структуры активного центра соединения с оценкой возможных конформационных изменений его структуры при связывании лигандов, учитывается участие в связывании молекул воды и т.д.). Если базовое соединение найдено, то проводится по циклам оптимизация его структуры для повышения биологической активности.

На этапе оптимизации синтезируются близкие по структуре соединения с последующим тестированием на биологическую (фармакологическую) активность. Когда требуемая активность оптимизированной структуры базового соединения будет достигнута, найденное соединение может быть передано на последующие доклинические испытания (испытания активности *in vivo*, на токсичность, канцерогенность и т.д.). По результатам этих тестов возможен повторный цикл компьютерного моделирования для оптимизации структуры с целью улучшения фармакокинетических свойств (адсорбционные, в процессах метаболизма, экскреции и т.п. — оптимизация ADME свойств) [95]. Считается, что CADD может уменьшить количество соединений, которые необходимо синтезировать и проверить на биологическую активность при разработке лекарственного препарата, примерно в 100 раз. Это позволяет резко сократить время и финансовые затраты на создание лекарств.

Как было указано выше, ключевым моментом в выборе метода CADD является наличие информации о пространственной структуре мишени. Если пространственная структура белка-мишени известна, то применяется группа методов конструирования лекарств на основе структуры мишени (Structure-Based Drug Design, SBDD), часто называемые прямыми методами. По этим методам можно сконструировать или найти соединения, которые комплементарны целевому участку на поверхности белка-мишени с учетом структурных и физико-химических свойств поверхности. При отсутствии информации о трехмерной структуре мишени используется другая группа методов — конструирование лекарств на основе структур лигандов (Ligand-Based Drug Design, LBDD), так называемые не прямые методы. В этом случае выполняется

анализ известных лигандов к белку для выявления общих свойств лигандов, которые коррелируют с биологической активностью, и на основании полученных данных строится модель активного центра белка-мишени.

#### **Конструирование лекарств на основе структуры мишени (технология SBDD)**

Первый этап SBDD состоит в анализе структуры мишени и выявлении места связывания низкомолекулярного лиганда. Оптимальным вариантом является ситуация, когда доступна структура комплекса мишени с известным лигандом (субстрат или конкурентный ингибитор). В противном случае потребуется поиск возможного места связывания лиганда. Самый простой подход — это анализ полостей (углублений) в структуре белка-мишени (фермента), и самая большая из них будет представлять с высокой вероятностью активный центр фермента. В рамках более точного подхода определяются ключевые аминокислотные остатки, участвующие в ферментативном катализе (например, триада сериновых протеаз), или кофактора, тогда окружающий регион будет представлять собой активный центр. Однако в этом случае остается неопределенность пространственного положения лиганда в активном центре мишени. Она решается путем докирования по всей поверхности белка-мишени известных субстратов и/или ингибиторов функции мишени и на основании результатов докинга определяется точное место связывания лиганда [96].

Наиболее популярной стратегией нахождения новых базовых структур биологически активных соединений является поиск низкомолекулярных соединений в молекулярных базах данных (виртуальный скрининг). Этот подход основан на предположении, что соединение с требуемой активностью уже существует, но оно не было проверено на наличие этой активности. Отметим, что в настоящее время в различных базах данных химических соединений собрано несколько миллионов разнообразных низкомолекулярных структур.

Основным методом виртуального скрининга соединений из химических баз данных является метод молекулярного докинга, когда к макромолекуле-мишени «подстраиваются» лиганды по принципу стерической комплементарности. В настоящее время известно много программ, ряд из которых оптимизирован для поиска лигандов в больших базах данных. Все программы докинга генерируют набор гипотез о вероятном положении потенциальных лигандов в месте связывания в макромолекуле-мишени. Достоверность этих гипотез оценивается с помощью различных оценочных функций (виртуальная энергия связывания, величина контактирующих поверхностей, число возможных водородных связей и т.д.). На основании этих оценок выбираются соединения для дальнейших исследований.

Преимуществом методов докинга является быстрота получения результатов, возможность проверки огромного количества молекулярных структур. Однако этот подход практически не позволяет сразу найти соединения с высокой активностью (ожидаемые значения констант диссоциации комплекса белка-мишени с лигандом  $K_d$  не превышают  $10^{-4}$  М).

В случае неудачного поиска базовых соединений в молекулярных базах данных можно использовать программы, реализующие методы конструирования лекарств *de novo*. Эти же методы используются для модифицирования найденных методом докинга базовых структур с целью повышения их активности, а также на последующих стадиях доклинических исследований для оптимизации фармакокинетических свойств и снижения побочных эффектов.

Эффективность связывания лиганда со своей мишенью оценивается по величинам виртуальной энергии связи взаимодействия. Основная проблема всех методов SBDD заключается в неспособности предсказать истинную величину энергии связывания и, соответственно, перевести ее в экспериментально определяемые параметры (константа диссоциации комплекса мишень/лиганд  $K_d$ , концентрация ингибитора, блокирующая 50% активности фермента,  $IC_{50}$ , константа ингибирования фермента  $K_i$ ). Хотя в ряде случаев значимую корреляцию между этими параметрами и удается обнаружить, но найденную взаимосвязь невозможно перенести на другие системы [97].

#### Конструирование лекарств на основе структур лигандов (технология LBDD)

Методы LBDD применяются в тех случаях, когда неизвестна пространственная структура мишени. Эти методы основаны на анализе наборов известных лигандов с требуемой биологической активностью. Поскольку структура места связывания лиганда в мишени неизвестна, первоначально необходимо построить его модель. Существует два типа моделей: фармакофорные модели [98] и различные модели «псевдо-рецептора» [99, 100]. Фармакофорная модель представляет собой набор точек в пространстве с определенными физико-химическими свойствами места связывания и расстояниями между ними. Модели «псевдо-рецептора» описывают в основном геометрические формы и размер места связывания в мишени. Указанные модели позволяют провести поиск новых лигандов в молекулярных базах данных. Виртуальный скрининг с использованием фармакофорной модели предполагает отбор молекул, удовлетворяющих требованиям данной модели относительно функциональных групп и расстояний между ними. Модели «псевдо-рецептора» позволяют искать новые лиганды методом молекулярного докинга. Метод «псевдо-рецептора» предпочтителен в том случае, когда фармакофорная модель содержит относительно мало фармакофорных точек, что приводит к отбору слишком большого количества соединений для экспериментального тестирования.

Для предсказания активности отобранных соединений, а также для оптимизации структур базовых соединений используются методы QSAR, основанные на регрессионном анализе взаимосвязи между биологической активностью набора гомологичных соединений и набором параметров (дескрипторов), рассчитанных по структурам этих соединений. Такие корреляционные уравнения дают возможность предсказывать активность новых аналогов из данного набора гомологов [101]. В настоящее время активно применяются методы трехмерного QSAR (3D-QSAR), основанные на описании пространственного распределения свойств лигандов, например, методы срав-

нительного анализа молекулярных полей (Comparative Molecular Field Analysis, CoMFA) и сравнительного анализа подобия молекул (Comparative Molecular Similarity Indices Analysis, CoMSIA) [102]. Эти методы позволяют охарактеризовать области пространства вокруг молекул лигандов (стерические, электростатические, гидрофобные и другие свойства), определяющие взаимодействие макромолекулы-мишени с лигандами из данного набора.

#### Виртуальный скрининг потенциальных лигандов

Существует три основных метода скрининга лигандов: 1) поиск лигандов на основе сходства двухмерных структурных формул; 2) поиск в молекулярных базах данных по фармакофорным моделям; 3) поиск в базах данных методом молекулярного докинга.

Обычно до начала проведения виртуального скрининга выполняется ряд подготовительных процедур — препроцессинг базы данных (удаление из базы соединений, не отвечающих некоторым параметрам: слишком большая или малая молекулярная масса, отсутствие определенных химических групп, несоответствие диапазону физико-химических свойств, таких как гидрофобность, растворимость, проницаемость и др.). Кроме того, часто необходимо конвертировать двухмерные химические структуры в трехмерные, рассчитать конформеры и т.п.

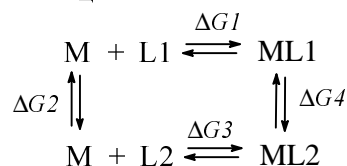
Поиск в базах данных позволяет найти соединения с требуемой биологической активностью, однако величины этих активностей обычно низкие. Найденные соединения могут быть использованы для дальнейшей оптимизации структуры или как структурные блоки для конструирования соединений *de novo*.

#### Предсказание аффинности лигандов

Предсказание аффинности лиганда (химического сродства к биологической мишени) — очень важная, но вместе с тем и самая проблематичная процедура в SBDD. Большинство программ молекулярного докинга имеют свою собственную, как правило, довольно грубую систему для ее оценки. Поэтому часто используются другие дополнительные методы оценки. Наиболее популярны оценочные функции, получаемые путем статистической обработки данных об известных структурах комплексов белок-лиганд и параметров их устойчивости [103—106]. Такие методы характеризуются высокой скоростью предсказания и независимостью от типа мишени. Можно проводить предсказания аффинности одновременно по нескольким оценочным функциям и осуществлять отбор по совокупности результатов предсказаний.

Наиболее точным методом предсказания аффинности лигандов является метод пертурбации свободной энергии (Free Energy Perturbation, FEP) [107].

Метод основан на использовании замкнутого термодинамического цикла:



Если известна свободная энергия  $G1$  связывания лиганда  $L1$  с мишенью  $M$ , то можно предсказать энер-

гию связывания  $G3$  лиганда  $L2$  с той же мишенью. Так как в замкнутом термодинамическом цикле изменение внутренней энергии суммарно равно нулю, то  $G3 = G1 + G4 - G2$ . Энергия  $G1$  известна из эксперимента;  $G2$  можно рассчитать из моделирования превращения  $L1 \rightarrow L2$  в воде;  $G4$  можно рассчитать из моделирования превращения  $L1 \rightarrow L2$  в комплексе с мишенью  $M$ .

Реализация метода требует больших объемов вычислений – длительное моделирование молекулярной динамики ( $\approx 100$  пс). Поэтому метод хорошо зарекомендовал себя только для небольших наборов гомологов.

Процесс предсказания аффинности потенциальных лигандов состоит из нескольких этапов (переход от грубых оценок к точным), направленных на последовательное сокращение количества отобранных кандидатов. В общем случае это следующие этапы: 1) первичный отбор гипотетических комплексов белок-лиганд при докинге; 2) предсказание аффинности быстрыми, грубыми методами; 3) по более точным оценочным функциям; 4) предсказание структуры комплексов методом молекулярной динамики и предсказание аффинности методом пертурбации свободной энергии. Конкретно, какие этапы нужно провести, зависит от количества соединений, которое предполагается экспериментально проверить на требуемую активность.

#### Модификация базовых соединений и дизайн *de novo*

Модификация базовых соединений и дизайн *de novo* проводятся с целью повышения требуемой активности, улучшения физико-химических и фармакокинетических свойств базовых соединений, а также для уменьшения побочных эффектов. Основным недостатком соответствующих методов является необходимость синтеза всех построенных соединений. Из программ, обеспечивающих модификацию структур и дизайн, отметим LUDI [108], CLIX [109], CAVEAT [110] и LeapFrog [111]. Все эти программы основаны примерно на одинаковых принципах моделирования новых молекул и оптимизации их структур с учетом пространственного расположения участков в активном центре, стерической комплементарности, электростатических взаимодействий, водородного связывания и др. На первом этапе работы таких программ осуществляется поиск химических групп, специфически взаимодействующих с мишенью. Далее следует этап поиска и оптимизации структурных блоков, связывающих эти группы в единую молекулу.

#### Предсказание фармакокинетических свойств

Для предсказания фармакокинетических свойств разработан ряд программ, предсказывающих ADME свойства соединений (адсорбция, распределение, метаболизм, экскреция). Эти предсказания могут быть использованы как на этапе препроцессинга баз данных для скрининга, так и после отбора соединений методом молекулярного докинга. Такие системы позволяют предсказать неблагоприятные свойства и вероятные побочные эффекты соединений. Примером такой программы может служить компьютерная система прогноза спектра биологической активности PASS (Prediction of Activity Spectra for Substances) [112].

#### Экспериментальное тестирование предсказанных соединений

Технология SBDD дает набор молекулярных структур, ранжированных по неким виртуальным параметрам, но не может ответить на вопрос о возможном биологическом эффекте данных соединений. Поэтому любое компьютерное предсказание необходимо проверять в реальных экспериментах по оценке биологической активности найденных или сконструированных и синтезированных соединений. На первом этапе обычно применяют различные тесты *in vitro*, которые позволяют быстро оценить способность потенциальных лигандов взаимодействовать с макромолекулой-мишенью. Для этих целей предпочтительно использовать высокоэффективные системы, основанные на прямом измерении взаимодействия лигандов с мишенями в реальном времени с определением кинетических, равновесных и термодинамических параметров, например метод поверхностного плазмонного резонанса (surface plasmon resonance technology, SPR) [113, 114]. Удачным инструментальным воплощением данной технологии являются оптические биосенсоры Biacore 3000 и Biacore S51, которые позволяют регистрировать в режиме реального времени скорость взаимодействия иммобилизованной мишени с потенциальными лигандами, молекулярная масса которых может быть от 150 Да и выше. Следует отметить, что другие системы анализа, основанные на регистрации изменения функции молекулы-мишени, являются узко специализированными и плохо вписываются в реализацию единой платформы «От гена до прототипа лекарства». Более того, в случае поиска лигандов к новой мишени требуется долгая и кропотливая разработка специального метода биотестирования.

#### Примеры прохождения этапов платформы «От гена до прототипа лекарства»

##### Поиск мишеней для лекарств антитуберкулезного и антибактериального действия

Приведем примеры из нашей практики использования биоинформационных подходов для поиска новых потенциальных мишеней. При создании антитуберкулезных препаратов [32] с помощью оригинальной программы GenMesh [32] был выполнен сравнительный анализ генома *Mycobacterium tuberculosis* H37Rv путем сопоставления с геномами *Mycobacterium tuberculosis* CDC1551 и *Mycobacterium leprae*, а также всеми известными белками человека и всеми белками из PDB (рис. 7). В результате был получен список из 13 новых потенциальных мишеней. Последующий детальный анализ позволил сократить список до 8 белков (кодируемых генами *rpoA*, *rpsD*, *rpsE*, *prsH*, *kdtB*, *ruvA*, *kdtB*). Позже два белка из этого списка были подтверждены другими исследователями как перспективные мишени для создания антибактериальных средств широкого спектра действия [44, 115].

Нами выполнен также анализ, направленный на поиск мишеней для новых антибактериальных препаратов широкого спектра действия (включая микобактерий), но без воздействия на нормальную микрофлору человека. Геномы *B. subtilis*, *E. coli* K12 и *B. longum* NCC2705 были использованы как представители симбионтов человека. Схема поиска показана на рис. 8А.

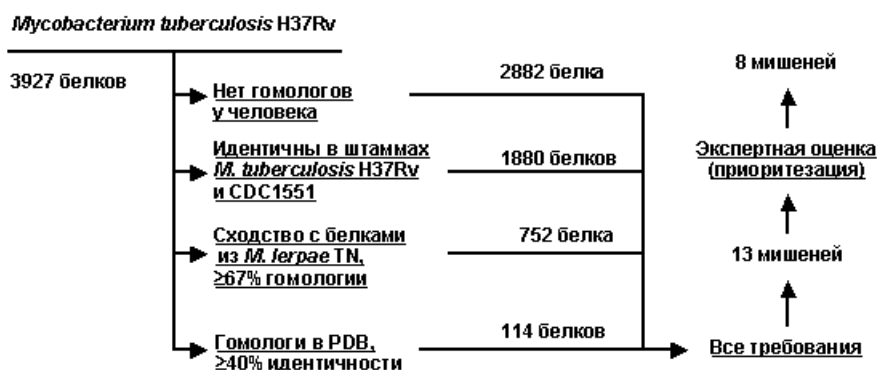


Рис. 7. Схема поиска новых мишеней в геноме *Mycobacterium tuberculosis* H37Rv

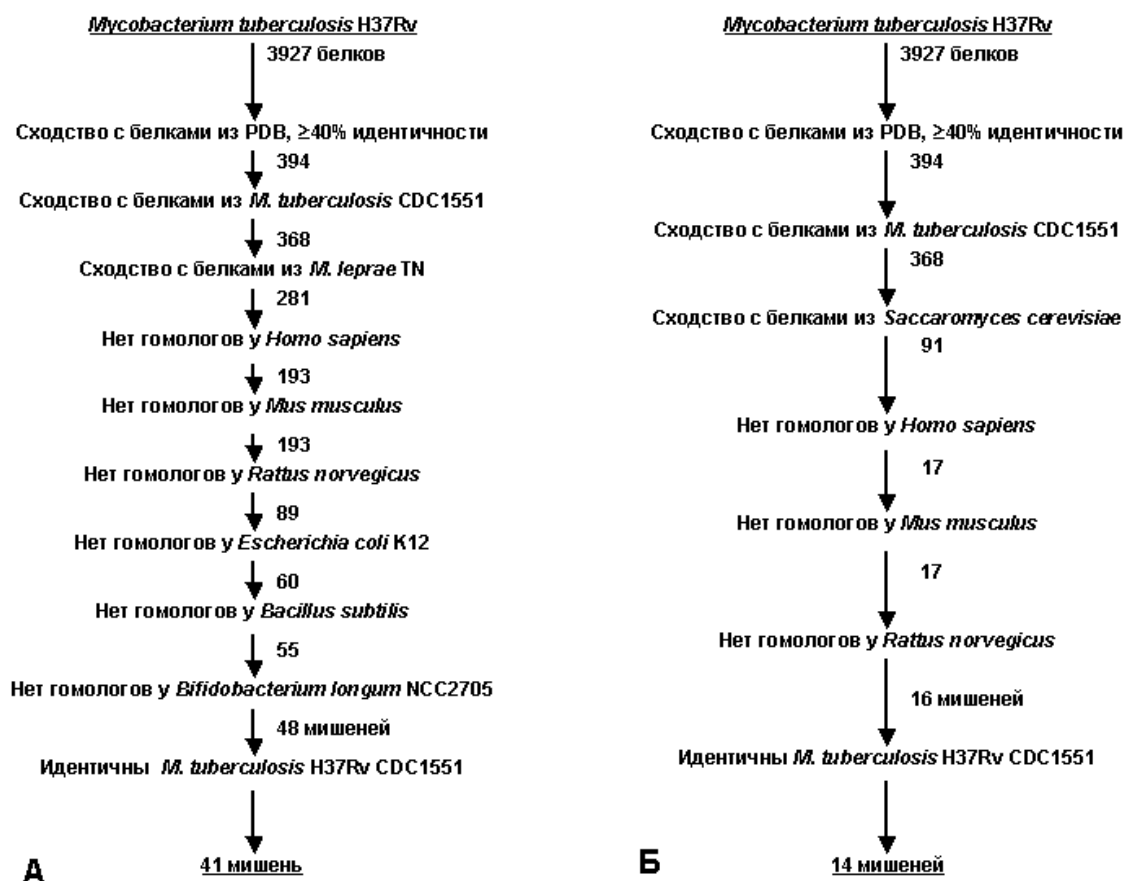


Рис. 8. Направленный поиск мишеней для новых антибактериальных препаратов с широким спектром действия (А) и общих мишеней у *Mycobacteria* и патогенных грибов (Б)

Полученный предварительный список включает 14 белков-потенциальных мишеней.

Сделана также попытка нахождения общих мишеней у *Mycobacteria* и патогенных грибов. Сопоставление геномов *Mycobacteria* с геномом *Saccharomyces cerevisiae* (модельный геном патогенных грибов) позволило выявить 14 других белков, лиганды которых могут служить основой для антибактериальных и антигрибковых препаратов (рис. 8Б).

**Моделирование пространственной структуры цитохрома P450 1A2 и виртуальный скрининг новых лигандов в молекулярных базах данных**

Цитохромы P450 являются гемсодержащими монооксигеназами — ферментами, присущими почти всем живым существам и окисляющими как эндогенные, так и экзогенные низкомолекулярные органические соединения. К этому обширному надсемейству принадлежит и цитохром P450 1A2 человека, который участвует в метаболизме стероидов и, кроме того, способен окислять многие лекарственные соединения и другие ксенобиотики, иногда образуя при этом канцерогенные вещества. По этой причине ингибиторы P450 1A2 могут найти применение для профилактики онкологических заболеваний, причиной которых служит лекарственная терапия или контакт с вредными химическими веществами на производстве, в быту или вследствие проживания в экологически неблагоприятных районах. Кроме того, окислительная модификация лекарственных соединений уменьшает их терапевтический эффект, и потому ингибирование P450 1A2 позволяет снизить обычно применяемые дозы.

Модель фермента цитохрома P450 1A2 была построена по гомологии с известными пространственными структурами других форм цитохрома P450 [116]. Структура активного центра модели оптимизирована в комплексе фермента с известными субстратами (кофеин и 7-этоксирезорифин) методом молекулярной динамики в водном окружении.

Предсказание аффинности новых лигандов проводили по 3D-QSAR модели, построенной на базе 24 известных лигандов к цитохрому P450 1A2. Новые лиганды (субстраты и/или ингибиторы) брали из объединенной базы, полученной в результате препроцессинга баз ASINEX [27], MayBridge [29] и CMC [26], в которую были добавлены 204 известных лиганда цитохрома P450 1A2 в качестве внутреннего

контроля. Новые лиганды цитохрома P450 1A2 подстраивали в модель фермента методом молекулярного докинга по программе DockSearch [117] с последующей оптимизацией положения лиганда в активном центре программой LeapFrog. В полученный список потенциальных лигандов вошли 185 (из 204) добавленных известных лигандов, что позволяет оценить эффективность поиска в 90%. В качестве примера в табл. 2 приведены семь контрольных соединений с предсказанными значениями констант диссоциации  $K_d$  комплекса фермент-лиганд.

В результате были выбраны 52 новых потенциальных лиганда цитохрома P450 1A2 для дальнейшего экспериментального тестирования.

**Поиск новых базовых структур ингибиторов для моноаминоксидазы при неизвестной структуре мишени**

Моноаминоксидаза (МАО; КФ 1.4.3.4) играет центральную роль в метаболизме важнейших нейромедиаторов — моноаминов. Изменение активности этого

Таблица 2

**Примеры контрольных соединений — лигандов цитохрома P450 1A2 и предсказанные значения констант диссоциации  $K_d$  комплексов фермент-лиганд**

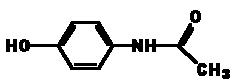
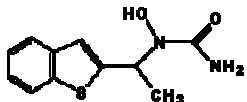
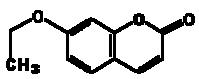
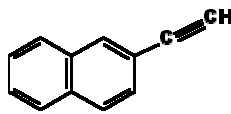
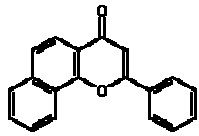
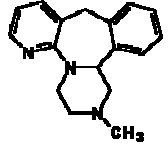
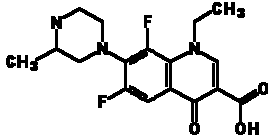
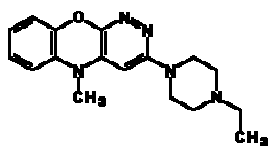
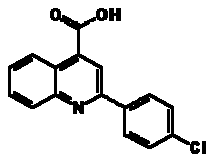
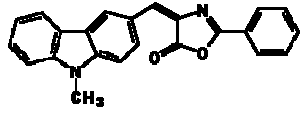
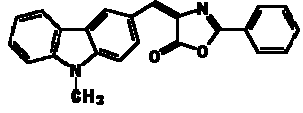
название	Лиганд структура	$K_d$ , мкМ	
		предсказано	эксперимент
Субстраты			
Парацетамол (Acetaminophen)		3000	3430
Зилеутон (Zileuton)		200	340
7-Этоксикумарин (7-Ethoxycoumarin)		20	21
Ингибиторы			
2-Этилнифталин (2-Ethynyl-naphthalene)		30	26
α-Нафтофлавоин (α-Naphthoflavone)		0,5	0,013
Миртазапин (Mirtazapine)		30	159
Ломефлоксацин (Lomefloxacin)		400	1230

Таблица 3

Результаты экспериментальной проверки найденных соединений на ингибиторную активность ( $IC_{50}$ , мкМ) для MAO A и MAO B

Индекс	Структура	MAO A	MAO B
BAS 0318949		316	>>100
BAS 0355758		>100	>>100
BAS 0370811		160	>>100
BAS 0442467		79	>>100

фермента при многих невропсихических расстройствах и возможность их коррекции ингибиторами MAO сделали этот фермент популярным объектом фундаментальных и клинических исследований. В организме млекопитающих MAO присутствует в двух формах — MAO A и MAO B, которые кодируются различными генами и отличаются по их чувствительности к ацетиленовым ингибиторам (хлоргилину и депренилу, соответственно) и по преимущественному окислению субстратов (серотонину и фенилэтилен-амину, соответственно).

MAO A и B — интегральные белки внешней мембраны митохондрий и в процессе очистки легко агрегируют. Поэтому долгое время исследователям не удавалось установить их пространственную структуру методами белковой кристаллографии. Отсутствие каких-либо белков-гомологов с известной пространственной структурой не позволяло использовать компьютерное молекулярное моделирование по гомологии для построения трехмерной структуры MAO.

На момент выполнения данной работы пространственная структура MAO была неизвестна. Для поиска новых ингибиторов MAO A и MAO B мы использовали подход, основанный на молекулярном докинге структур из базы данных в построенную модель полости активного центра фермента, так как фармакофорная модель, построенная на основе ингибиторов из разных химических классов, оказалась малоинформативной и содержала всего только две фармакофорные точки (ароматическое кольцо и рядом расположенный отрицательно заряженный атом).

Геометрические модели полости активного центра MAO A и MAO B («слепки») были построены путем пространственной суперпозиции структур соответствующих высокоэффективных конкурентных ингиби-

торов двух форм ферментов и объединения их объемных изображений [100]. В эти модели докировали структуры соединений, взятые из молекулярной базы данных коммерчески доступных образцов (примерно 8000 соединений). Предварительно для них были рассчитаны все варианты устойчивых конформеров — примерно 50000 структур. Все эти структуры были использованы для молекулярного докинга. В результате были отобраны 7000 структур в качестве предварительного списка потенциальных лигандов. Для всех отобранных структур были предсказаны величины ингибиторной активности по построенным ранее 3D-QSAR с CoMFA моделям. На основании предсказаний был осуществлен окончательный отбор потенциальных лигандов и четыре лучших были переданы для экспериментальной проверки, которая показала, что все соединения ингибируют активность MAO A (табл. 3). Лучшее соединение имело фармакологическую активность ( $IC_{50}$ ), равную 80 мкМ. Хотя ингибиторная активность этого соединения не очень высока, оно может быть использовано в качестве базовой структуры для последующей оптимизации методами CADD.

#### ЛИТЕРАТУРА

1. Lohse M.J. Trends Pharmacol. Sci., 1998, v. 19, p. 198.
2. NCBI GenBank (<http://www.ncbi.nlm.nih.gov>).
3. Kulikova T., Aldebert P., Althorpe N. e. a. Nucleic Acids Res., 2004, v. 32, Database issue, D27-D30, (<http://www.ebi.ac.uk/embl>).
4. DDBJ (<http://www.ddbj.nig.ac.jp>).
5. Boeckmann B., Bairoch A., Apweiler R. e. a. Nucleic Acids Res., 2003, v. 31, p. 365–370 (<http://www.expasy.org/sprot>).
6. Wu C.H., Huang H., Yeh L.-S.L., Barker W.C. Comput. Biol. Chem., 2003, v. 27, p. 37–47 (<http://pir.georgetown.edu>).
7. Tatusov R.L., Fedorova N.D., Jackson J.D. e. a. BMC Bioinformatics, 2003, v. 4, p. 41.
8. Kanehisa M., Goto S., Kawashima S., Okuno Y., Hattori M. Nucleic Acids Res., 2004, v. 32, Database issue, D277-D280.
9. Pandit S.B., Bhadra R., Gowri V.S., Balaji S., Anand B., Srinivasan N. BMC Bioinformatics, 2004, v. 5, p. 28 (<http://www.sanger.ac.uk/Software/Pfam/>).
10. Haft D.H., Selengut J.D., White O. Nucleic Acids Res., 2003, v. 31, p. 371–373 (<http://www.tigr.org/TIGRFAMs/>).
11. Peterson J.D., Umayam L.A., Dickinson T., Hickey E.K., White O. Ibid., 2001, v. 29, p. 123–125 (<http://www.tigr.org/CMR>).
12. Uchiyama I. Ibid., 2003, v. 31, p. 58–62 (<http://mbgd.genome.ad.jp>).
13. Xenarios I., Salwinski L., Duan X.J. e. a. Ibid., 2002, v. 30, p. 303–305 (<http://dip.doe-mbi.ucla.edu>).
14. Bader G.D., Betel D., Hogue C.W. Ibid., 2003, v. 31, p. 248–250 (<http://www.blueprint.org/bind/bind.php>).
15. Ng S.K., Zhang Z., Tan S.H. Bioinformatics, 2003, v. 19, p. 923–929 (<http://interdom.lit.org.sg>).
16. Suhre K., Claverie J.-M. Nucleic Acids Res., 2004, v. 32, Database issue, D273-D276 (<http://igs-server.cnrs-mrs.fr/FusionDB/>).
17. PathDB (<http://www.ncgr.org/pathdb/>).



18. *Berman H.M., Westbrook J., Feng Z. e. a.* Nucleic Acids Res., 2000, v. 28, p. 235–242 (<http://www.rcsb.org/pdb>).
19. *Noguchi T., Akiyama Y.* Ibid., 2003, v. 31, p. 492–493 (<http://www.cbrc.jp/pdbdrepdb>).
20. *Frishman D., Mokrejs M., Kosykh D. e. a.* Ibid., 2003, v. 31, p. 207–211 (<http://pedant.gsf.de>).
21. *Galperin M.Y.* Ibid., 2004, v. 32, Database issue, D3–D22.
22. *Freiberg C.* Drug Discovery Today, 2001, v. 6, p. S72–S80.
23. *Allen F.H.* Acta Cryst., 2002, v. B58, p. 380–388 (<http://www.ccdc.cam.ac.uk>).
24. National Cancer Institute: Pure Chemicals Repository ([http://www.dtp.nci.nih.gov/branches/dscb/repo\\_open.html](http://www.dtp.nci.nih.gov/branches/dscb/repo_open.html)).
25. MDL Drug Data Report, MDL Information Systems, Inc., 14600 Catalina Street San Leandro, CA 94577, USA (<http://www.mdl.com>).
26. Comprehensive Medicinal Chemistry, MDL Information Systems, Inc., 14600 Catalina Street San Leandro, CA 94577, USA (<http://www.mdl.com>).
27. ASINEX Ltd. 5 Gabriчевского St. Bldg 8, Moscow 123367, Russia (<http://www.asinex.com>).
28. ChemBridge Corporation 16981 Via Tazon, Suite G San Diego, CA 92127, USA (<http://www.chembridge.com>).
29. Maybridge, Trevillet, Tintagel, Cornwall PL34 OHW, England (<http://www.maybridge.com>).
30. SYBYL 6.7.1, Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA (<http://www.tripos.com>).
31. *Spaltmann F., Blunck M., Ziegelbauer K.* Drug Discovery Today, 1999, v. 4, p. 17–26.
32. *Дубанов А.В., Иванов А.С., Арчаков А.И.* Вopr. мед. химии., 2001, т. 47, с. 353–367.
33. GeneData Phylosopher (<http://www.genedata.com>).
34. Entrez Utilities ([http://www.ncbi.nlm.nih.gov/entrez/query/static/eutils\\_help.html](http://www.ncbi.nlm.nih.gov/entrez/query/static/eutils_help.html)).
35. NCBI BLAST service (<http://www.ncbi.nlm.nih.gov/BLAST/Doc/urlapi.html>).
36. Insight II и Quanta (<http://www.accelrys.com>).
37. *Case D.A., Darden T.A., Cheatham III T.E. e. a.* AMBER 8, University of California, San Francisco, 2004 (<http://amber.scripps.edu>).
38. *Berendsen H.J.C., van der Spoel D., van Drunen R.* Comp. Phys. Comm., 1995, v. 91, p. 43–56 (<http://www.gromacs.org>).
39. *Veselovsky A.V., Ivanov Yu.D., Ivanov A.S., Archakov A.I., Lewi P., Janssen P. J.* Molecular Recognition, 2002, v. 15, p. 405–422.
40. *Archakov A.I., Govorun V.M., Dubanov A.V., Ivanov Y.D., Veselovsky A.V., Lewi P., Janssen P.* Proteomics, 2003, v. 3, p. 380–391.
41. *Eisenberg D., Marcotte E.M., Xenarios I., Yeates T.O.* Nature, 2000, v. 405, p. 823–826.
42. *Jansen R., Lan N., Qian J., Gerstein M. J.* Struct. Funct. Genomics, 2002, v. 2, p. 71–81.
43. *Marcotte E.M., Xenarios I., van Der Blik A.M., Eisenberg D.* Proc. Natl. Acad. Sci. U.S.A., 2000, v. 97, p. 12115–12120.
44. *Thanassi J.A., Hartman-Neumann S.L., Dougherty T.J. e. a.* Nucleic Acids Res., 2002, v. 30, p. 3152–3162.
45. *Boguslavsky J.* Drug Discovery & Development, 2002, v. 5, p. 41–48.
46. *Walgren J.L., Thompson D.C.* Toxicol. Lett., 2004, v. 149, p. 377–385.
47. *Jones S., Thornton J.M.* Prog. Biophys. Mol. Biol., 1995, v. 63, p. 31–65.
48. *Wilkinson K.D.* Methods Mol. Biol., 2004, v. 261, p. 15–32.
49. *Pillutla R.C., Goldstein N.I., Blume A.J., Fisher P.B.* Expert Opin. Ther. Targets, 2002, v. 6, p. 517–531.
50. *Butcher S.P.* Neurochem. Res., 2003, v. 28, p. 367–371.
51. *Sheppard D.* Am. J. Respir. Cell. Mol. Biol., 1994, v. 11, p. 1–6.
52. *Homanics G.E., Quinlan J.J., Mihalek R., Firestone L.L.* Toxicol Lett., 1998, v. 100–101, p. 301–307.
53. *Luscombe N.M., Austin S.E., Berman H.M., Thornton J.M.* Genome Biology, 2000, v. 1, reviews001.1–001.10.
54. *Cech T.R.* Curr. Opin. Struct. Biol., 1992, v. 2, p. 605–609.
55. *Goodchild J.* Expert Opin. Ther. Targets, 2002, v. 6, p. 235–247.
56. *Jain K.K.* Drug Discov. Today, 2004, v. 9, p. 307–309.
57. *Baker B.F., Monia B.P.* Biochim. Biophys. Acta, 1999, v. 1489, p. 3–18.
58. *Ravichandran L.V., Dean N.M., Marcusson E.G.* Oligonucleotides, 2004, v. 14, p. 49–64.
59. SomaGenic (<http://www.somagenics.com/platform.html>).
60. *Pellestor F., Paulasova P.* Int. J. Mol. Med., 2004, v. 13, p. 521–525.
61. *Ray A., Norden B.* FASEB J., 2000, v. 14, p. 1041–1060.
62. *Banker D.D.* Indian J. Med. Sci., 2001, v. 55, p. 651–654.
63. *Henning S.W., Beste G.* Curr. Drug Discov., 2002, May, p. 17–21.
64. *Liao J.C., Roider J., Jay D.G.* Proc. Natl. Acad. Sci. USA, 1994, v. 91, p. 2659–2663.
65. *Eustace B.K., Jay D.G.* Methods Enzymol., 2003, v. 360, p. 649–660.
66. *Chowdhury P.S., Vasmatzis G.* Methods Mol. Biol., 2003, v. 207, p. 237–254.
67. *Tanaka T., Lobato M.N., Rabbitts T.H. J.* Mol. Biol., 2003, v. 331, p. 1109–1120.
68. *Burgstaller P., Girod A., Blind M.* Drug Discov. Today, 2002, v. 7, p. 1221–1228.
69. *Ilag L.L., Ng J.H., Beste G., Henning S.W.* Ibid., 2002, v. 7, p. S136–S142.
70. *Flook P.K., Yan L., Szalma S.* Targets, 2003, v. 2, p. 217–223.
71. *Xin H., Bernal A., Amato F.A. e. a.* J. Biomol. Screen., 2004, v. 9, p. 286–293.
72. *Sundberg S.A., Chow A., Nikiforov T., Wada, H.G.* Drug Discov. Today, 2000, v. 5, p. 92–103.
73. *Matthews D., Kopczyński J.* Ibid., 2001, v. 6, p. 141–149.
74. *Tornell J., Snaith M.* Ibid., 2002, v. 7, p. 461–470.
75. *Sumanas S., Lin S.* Ibid., 2004, v. 3, p. 89–96.
76. *Lee J., Nam S., Hwang S.B. e. a.* J. Biochem. Mol. Biol., 2004, v. 37, p. 107–113.
77. *Wassarman D.A., Therrien M., Rubin G.M.* Curr. Opin. Genet. Dev., 1995, v. 5, p. 44–50.
78. *Chopra I.* Microbiology Today, 2000, v. 27, p. 4–6 (<http://www.cubist.com/ar2000text/discovery.html>).
79. *Jackson L.K., Phillips M.A.* Curr. Top. Med. Chem., 2002, v. 2, p. 425–438.
80. Methods In Enzymology. Eds. C.W.Jr. Carter, R.M. Sweet. 2003, v. 368: Macromolecular Crystallography, Part C. San Diego: Acad. Press, CA.
81. *Downing A.K.* Protein NMR Techniques, Second Edition, 2004, Humana, Totowa, NJ.
82. *Wallin E., Von Heijne G.* Protein Sci., 1998, v. 7, p. 1029–1038.
83. *Grisshammer R., Tate C.G.* Q. Rev. Biophys., 1995, v. 28, p. 315–422.
84. *Eswar N., John B., Mirkovic N. e. a.* Nucleic Acids Res., 2003, v. 31, p. 3375–3380.
85. *Fiser A., Sali A.* Methods Enzymol., 2003, v. 374, p. 461–491.
86. *Topham C.M., Thomas P., Overington J.P. e. a.* Biochem. Soc. Symp., 1990, v. 57, p. 1–9.
87. *Moult J., Fidelis K., Zemla A., Hubbard T.* Proteins, 2003, v. 53, Suppl. 6, p. 334–339 (<http://predictioncenter.llnl.gov>).

88. *Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J.M.* J. Appl. Cryst., 1993, v. 26, p. 283–291 (<http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html>).
89. ProTable (<http://www.tripos.com/sciTech/inSilicoDisc/media/LitCtr/protable.pdf>).
90. *Godzik A., Kolinski A., Skolnick J. J.* Comp. Aid. Mol. Design, 1993, v. 7, p. 397–438 (<http://www.tripos.com/admin/LitCtr/matchmaker.pdf>).
91. *Vriend G. J.* Mol. Graph., 1990, v. 8, p. 52–56 (<http://www.cmbi.kun.nl/whatif/>).
92. *Sippl M.J.* Proteins, 1993, v. 17, p. 355–362 (<http://lore.came.sbg.ac.at/Services/prosa.html>).
93. *Luthy R., Bowie J.U., Eisenberg D.* Nature, 1992, v. 356, p. 83–85 ([http://www.accelrys.com/insight/Profiles-3D\\_page.html](http://www.accelrys.com/insight/Profiles-3D_page.html)).
94. *Entzeroth M.* Curr. Opin. Pharmacol., 2003, v. 3, p. 522–529.
95. *Clark D.E., Pickett S.D.* Drug Discov. Today, 2000, v. 5, p. 49–58.
96. *Иванов А.С., Дубанов А.В., Скворцов В.С., Арчаков А.И.* Вopr. мед. химии, 2002, т. 48, с. 304–315.
97. *Hubbard R. E.* Curr. Opin. Biotechnol., 1997, v. 8, p. 696–700.
98. *Flohr S., Kurz M., Kostenis E. e. a.* J. Med. Chem., 2002, v. 45, p. 1799–1805.
99. *Schleifer K.-J.* J. Comp.-Aided Mol. Design, 2000, v. 14, p. 467–475.
100. *Veselovsky A.V., Tikhonova O.V., Skvortsov V.S., Medvedev A.E., Ivanov A.S.* QSAR SAR Environm. Res., 2001, v. 12, p. 345–358.
101. *Kubinyi H.* Quant. Struct.-Act. Relat., 1994, v. 13, p. 285–294.
102. *Cramer III R.D., Petterson D.E., Bunce J.D.* J. Am. Chem. Soc., 1988, v. 110, p. 5959–5967.
103. *Jones G., Willett P., Glen R. e. a.* J. Mol. Biol., 1997, v. 267, p. 727–748.
104. *Muegge I., Martin Y.C.* J. Med. Chem., 1999, v. 42, p. 791–804.
105. *Ewing T.J.A., Kuntz I.D.* J. Comp. Chem., 1996, v. 18, p. 1175–1189 (<http://dock.compbio.ucsf.edu>).
106. *Eldridge M.D., Murray C.W., Auton T.R., Paolini G.V., Mee R.P.* J Comput Aided Mol Des., 1997, v. 11, p. 425–445.
107. *Pearlman D.A., Rao B.G.* Encyclopedia of Computational Chemistry. 1998, John Wiley, p. 1036–1061.
108. *Bohm H.J.* J. Comp.-Aided. Mol. Design, 1992, v. 6, p. 61–78.
109. *Lawrence M.C., David P.C.* Proteins: Struct. Funct. Genet., 1992, v. 12, p. 31–41.
110. *Bartlett P.A., Shea G.T., Telfer S.J., Waterman S.* Molecular recognition in chemical and biological problems. 1989, Royal Chem. Soc., London, v.78, p. 182–196.
111. LeapFrog: SYBYL® 6.9.2 Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA (<http://www.tripos.com>).
112. *Poroikov V.V., Filimonov D.A., Borodina Yu.V., Lagunin A.A., Kos A.* J. Chem. Inf. Comput. Sci., 2000, v. 40, p. 1349–1355 (<http://www.ibmh.msk.su/PASS/>).
113. Real-Time Analysis of Biomolecular Interactions: Applications of Biacore. Eds K. Nagata, H. Handa. Tokyo: Springer-Verlag, 2000.
114. *Rich R.L., Myszka D.G.* Curr. Opin. Biotechnol., 2000, v. 11, p. 54–61.
115. *Freiberg C., Wieland B., Spaltmann F. e. a.* J. Mol. Microbiol. Biotechnol., 2001, v. 3, p. 483–489.
116. *Белкина Н.В., Скворцов В.С., Иванов А.С., Арчаков А.И.* Вopr. мед. химии, 1998, т. 44, с. 464–473.
117. DockSearch (<http://imgdd.ibmh.msk.su/originalsoftware/DS/DS.htm>).