

БИОИНФОРМАТИКА

УДК 577.152.3.
©Коллектив авторов

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ ЦИТОХРОМОВ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

А.С.Иванов, В.С.Скворцов, А.А.Сеченых, А.В.Дубанов, А.В.Лисица

ГУ НИИ биомедицинской химии им. В.Н.Ореховича РАМН
119121, Москва, Погодинская ул. 10;
факс: (095) 245-0857; эл.почта: ivanov@ibmh.msk.su

Обзор посвящен описанию основных подходов, проблем и перспектив компьютерного моделирования трехмерной структуры цитохромов P450 (P450). В нем рассмотрены следующие вопросы: (1) место компьютерного моделирования в исследованиях P450; (2) основные этапы моделирования пространственной структуры белка по гомологии; (3) цели и задачи молекулярного моделирования; (4) проблемы моделирования P450 по гомологии; (5) критерии оценки успешного моделирования P450 по гомологии; (6) методы уточнения моделей P450, проверки их правильности и достоверности.

Ключевые слова: цитохром P450, трехмерная структура белка, молекулярное моделирование, молекулярная графика

ВВЕДЕНИЕ. В настоящее время в биомедицинских исследованиях большое внимание уделяется цитохромам P450 (P450), так как они играют ключевую роль в окислении различных ксенобиотиков (в том числе и лекарств), а также ряда эндогенных веществ [1]. Совершенно очевидно, что для детального изучения молекулярного механизма функционирования P450, требуется исследование их пространственной молекулярной структуры.

Как известно, современные подходы в анализе пространственной структуры белков базируются на комбинации методов молекулярной биологии, биохимии и белковой кристаллографии. Схематично это выглядит как ряд последовательных этапов:

- 1) Получение чистого нативного или чаще всего рекомбинантного водорастворимого белка.
- 2) Получение белковых кристаллов, пригодных для рентгеноструктурного анализа (РСА) (подбор условий кристаллизации и выращивание кристаллов).
- 3) Рентгеноструктурный анализ белковых кристаллов, компьютерный анализ дифракционных данных и реконструкция трехмерной структуры белка.

К сожалению, данный подход пригоден только для водорастворимых белков, методы экспрессии и кристаллизации которых хорошо разработаны [2]. В то же время около 30-40% белков живых организмов связаны с биологическими мембранами [3]. Тот факт, что пока исследованы только единичные структуры

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

мембранных белков (в белковом банке данных PDB [4] имеется всего порядка 30 таких структур [5]), обусловлен существующими проблемами их экспрессии и кристаллизации [6].

Все вышесказанное в полной мере относится и к цитохромам P450, большинство из которых является мембранными белками [1]. Это обусловлено наличием гидрофобного мембранного якоря и гидрофобных кластеров на поверхности белковой глобулы.

В настоящее время известно более тысячи последовательностей P450 [7] и в тоже время только 12 кристаллических структур водорастворимых P450 доступны в PDB (табл. 1). Этот очевидный дефицит данных о молекулярном строении P450 заставляет исследователей прибегать к методам компьютерного моделирования их пространственной структуры [8].

Таблица 1. Список цитохромов P450 с известной пространственной структурой, доступной в белковом банке PDB (март 2003 г.).

| № | Название | | | Организм | Число структур в PDB |
|-----|------------|----------|--|---------------------------------------|----------------------|
| 1. | P450cam | CYP101A | Camphor Monooxygenase | <i>Pseudomonas putida</i> | 38 |
| 2. | P450 bm-3 | CYP102A | Fatty Acid Hydroxylase | <i>Bacillus megaterium</i> | 8 |
| 3. | P450 terp | CYP108A | Terpineol Oxidoreductase | <i>Pseudomonas sp.</i> | 1 |
| 4. | P450 Nor | CYP55 | Fungal Nitric Oxide Reductase | <i>Fusarium oxysporum</i> | 14 |
| 5. | P450 Eryf | CYP107A1 | Polycyclic Aromatic Oxidoreductase | <i>Saccharopolyspora erythraea</i> | 6 |
| 6. | P450 2C5 | CYP2C5 | Mammalian Progesterone 21-hydroxylase | <i>Oryctolagus cuniculus</i> (кролик) | 1 + (2) * |
| 7. | P450 119 | CYP119 | Thermophilic Oxidoreductase | <i>Sulfolobus solfataricus</i> | 5 |
| 8. | P450 MT1 | CYP51 | 14 alpha-sterol demethylase | <i>Mycobacterium tuberculosis</i> | 2 + (1) |
| 9. | P450 165 | CYP165B | OxyB | <i>Amycolatopsis orientalis</i> | 3 |
| 10. | P450 MT2 | CYP121A | - | <i>Mycobacterium tuberculosis</i> | 2 |
| 11. | P450 154C1 | CYP154C1 | 12- and 14-carbon macrolactone monooxygenase | <i>Streptomyces coelicolor</i> | 1 |
| 12. | P450 175A1 | CYP175A1 | Thermophilic Cytochrome P450 | <i>Thermus thermophilus</i> HB27 | 1 |

Ряд исследователей пытается решить проблему кристаллизации P450 с помощью методов генной инженерии - модифицировать молекулу путем удаления (транкирования) мембранного якоря белка (N-концевой фрагмент) и выполнения ряда точечных мутаций на поверхности белка. В результате такая модифицированная молекула должна стать более гидрофильной, растворяться в воде в виде мономера и легко кристаллизоваться. Использование данного подхода позволило получить первый кристалл P450 млекопитающего (модифицированный CYP2C5 кролика) [9, 10]. Проблемой в реализации подхода является идентификация гидрофобных кластеров на поверхности P450 и выбор оптимального варианта точечных мутаций для уменьшения общей гидрофобности поверхности P450 без изменения пространственной структуры белка. В связи с этим представляется целесообразным построение трехмерных компьютерных моделей P450 с оптимизацией молекулярной поверхности. На основе анализа структуры гидрофобных кластеров и участков взаимодействия с белками-партнерами могут быть выполнены компьютерные эксперименты по модификации

белка для уменьшения его гидрофобности. Анализ различных вариантов замены гидрофобных аминокислот на поверхности белка позволит спрогнозировать минимальное число точечных мутаций для последующего получения методами генной инженерии модифицированных водорастворимых кристаллизующихся Р450 [11].

Основным подходом в трехмерном моделировании белков является метод моделирования по гомологии, где для предсказания структуры моделируемого белка используются данные об аминокислотных последовательностях и о пространственной структуре известных белков-гомологов. Существует несколько компьютерных программ, позволяющих осуществлять моделирование структуры белков по гомологии (Sybyl [12], Insight II [13], Quanta [14], Modeller [15, 16] и др.). Так как в нашей работе мы используем пакет молекулярного моделирования Sybyl (Tripos Inc.), то при написании данного обзора мы основывались на опыте моделирования структуры белка по гомологии с использованием программы Composer [12, 17] из данного пакета.

1. Основные этапы моделирования пространственной структуры белка по гомологии.

1.1. Поиск белков-гомологов с известной пространственной структурой в банке данных PDB. Поиск гомологов с известной трехмерной структурой осуществляется путем сканирования белковой базы данных PDB с целью нахождения белков с последовательностью, похожей на последовательность целевого белка. Для моделирования Р450 этот этап поиска пока не нужен, так как в PDB имеется на момент написания данного обзора (март 2003 г.) всего 12 структур-гомологов (табл. 1).

1.2. Идентификация ключевых аминокислотных остатков ("seed residues").

Поиск и идентификация ключевых аминокислотных остатков ("seed residues" или "seeds") осуществляется на основе гомологии между белковыми последовательностями. "Seeds" представляют собой топологически эквивалентные, абсолютно консервативные (инвариантные) остатки в последовательностях и используются как "опорные точки" для первоначального пространственного совмещения структур - структурного выравнивания (structural alignment). Так как возможные ошибки и несоответствия в выборе ключевых аминокислотных остатков могут быть компенсированы на следующем этапе, то исследователи, как правило, не тратят много времени на данный этап.

1.3. Структурное выравнивание с использованием ключевых остатков для выявления структурно-консервативных областей (Structurally Conserved Regions, SCR).

Программа Composer выполняет оптимизированное структурное выравнивание (пространственное совмещение) набора известных гомологов с использованием ключевых аминокислотных остатков. Для совмещения структур в трехмерном пространстве используются координаты только α атомов. Критерием оптимального совмещения структур является минимальная величина среднеквадратичного отклонения расстояний между α атомами (RMS). Если есть только два известных белка-гомолога, то структурное выравнивание представляет собой простую процедуру совмещения по методу наименьших квадратов. В тех случаях, когда имеется три и более структур-гомологов, осуществляется множественное совмещение путем выполнения серии парных взвешенных среднеквадратичных совмещений. В результате структурного выравнивания определяются структурно-консервативные области (SCR), т.е. те участки последовательностей, в которых три и более аминокислотных остатка совместились в пространстве с достаточной точностью (расстояние между α атомами $< 3,5 \text{ \AA}$). Процедура структурного выравнивания позволяет также определить усредненные структуры консервативных областей для всего набора гомологов. Полученные таким образом данные используются в дальнейшем для конструирования модели целевого белка.

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

1.4. Определение участков SCR в моделируемом белке.

После нахождения SCR в белках-гомологах определяют соответствующие участки в последовательности моделируемого белка. Для этого выполняется локальное парное выравнивание (алгоритм Needleman-Wunsch [18, 19]) последовательности целевого белка с последовательностями SCR из белков-гомологов. Особенностью данного выравнивания является назначение различной величины штрафа за разрыв (gap) моделируемой последовательности в различных областях. Для предотвращения разрыва областей SCR для них назначается крайне высокий штраф, в то время как для участков между SCR назначается низкая величина штрафа, разрешающая их разрывы. В результате такого выравнивания получается набор оценок сходства между последовательностью моделируемого белка и соответствующим SCR каждого из гомологов. Так как автоматические методы выравнивания пока далеки от совершенства, то исследователь должен вручную проверять результаты полученного выравнивания и выбирать более разумные варианты. В случае P450 часто выбирают участки SCR из наиболее близкого гомолога (белка-образца). На рис. 1 показан в качестве примера результат выравнивания последовательностей CYP2B4 (модель) и CYP2C5 (образец) [20].

| | | | | | | | | | |
|---|---|---|---------|-------------|--------------|----------|-----|----------|--|
| Sequence 1: CYP2B4 (model) | | | | | | | | | |
| Sequence 2: CYP2C5 (1DT6) | | | | | | | | | |
| Helix A | | Helix B | | b11 | | b12 | | helix B' | |
| ----- | | ----- | | ----- | | ----- | | ----- | |
| 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | | |
| PPGPSPLPVLGNLLQMDRKGLLRSFLRLREKYGDVFTVYLGSRPVVVLGCTDAIREALVDQAEAFSGRGKIAVVDPIFQG | | | | | | | | | |
| PPGPTFPFIIGNILQIDAKDISKSLTKFSECYGPFVFTVYLGMRKPTVVLLHGYEAVKEALVDLGEFAGRGSVPILEKVSQG | | | | | | | | | |
| <u>SCR1</u> | | <u>SCR2</u> | | <u>SCR3</u> | | | | | |
| helix C | | | helix D | | | b21 | | helix E | |
| ----- | | | ----- | | | ----- | | ----- | |
| 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | | |
| YGVIFANGERWRALRRFSLATMRDFGMRKRSVEERIQEEARCLVEELRKS K GALLDNTLLFHSITSNIICSIIVFGKRFDY | | | | | | | | | |
| LGIAFNAKTKENRRRSLMTRLNRFMGMRKRSIEDRIQEEARCLVEGLRKTNASPCDPTFILGCAPCNVICSVIFHNRFDY | | | | | | | | | |
| <u>SCR4</u> | | <u>SCR5</u> | | <u>SCR6</u> | | | | | |
| Helix F | | | helix G | | | helix H | | | |
| ----- | | | ----- | | | ----- | | | |
| 200 | 210 | 220 | 230 | 240 | 250 | 260 | 270 | | |
| KDPVFLRLLDLFFQSFLSISSFSSQVFEFLFPGFLKHPGTHRQIYRNLOEINTFICQSVEKHRATLDPSNPRDFIDVYLL | | | | | | | | | |
| KDEEFLKLMESLHENVLLGT | P | LDYFPGIHKTLKKNADYIKNFIMEKVKHEQKLLDVNNPRDFIDCFLI | | | | | | | |
| <u>SCR7</u> | | <u>SCR8</u> | | | <u>SCR9</u> | | | | |
| helix I | | | helix J | | | helix J' | | | |
| ----- | | | ----- | | | ----- | | | |
| 280 | 290 | 300 | 310 | 320 | 330 | 340 | 350 | | |
| RMEKDKSDPSSEFHQNLILTLVLSLFFAGTETTSTTLRYGFLMLKYPHTERVQKEIEQVIGSHRPPALDDRAKMPYTD | | | | | | | | | |
| KMEQ | ENNLEFTLESLVIAVSDLEFGAGTETTSTTLRYSLLLLLLKHPEVAARVQEEIERVIGRHRSPCMQDRSRMPYTD | | | | | | | | |
| <u>SCR10</u> | | | | | | | | | |
| helix K | | b14 | | b13 | | helix K' | | | |
| ----- | | ----- | | ----- | | ----- | | | |
| 360 | 370 | 380 | 390 | 400 | 410 | 420 | 430 | | |
| AVIHEIQRLGDLIFPGVPHVTVKDTQFRGYVIPKNTVEFPVLSALEDPRYFETPNTFNPGHFLDANGALKRNEGFMPPFS | | | | | | | | | |
| AVIHEIQRFIDLLFTNLPHAVTRDRVFRNYFIPKGTDIITSLTSLVHLDEKAFPNPKVFDPGHFLDESGNFKKSDFMPPFS | | | | | | | | | |
| <u>SCR11</u> | | <u>SCR12</u> | | | <u>SCR13</u> | | | | |
| helix L | | b23 | | b22 | | | | | |
| ----- | | ----- | | ----- | | | | | |
| 440 | 450 | 460 | 470 | 480 | 490 | 500 | 510 | | |
| LGKRICLGEGI?RTEFLFFFTTILQNFASVPVPEIDILTPRESGVGNVPPSYQIRFLAR | | | | | | | | | |
| AGKRMVCGEGLARMEFLFLTSLILQNFKLSLVEPKDLDTAVVNGFVSVPPSYQLCFIPIHH | | | | | | | | | |
| <u>SCR14</u> | | <u>SCR15</u> | | | <u>SCR16</u> | | | | |

Рисунок 1.

Выравнивание последовательностей моделируемого белка (CYP2B4) и белка-образца (CYP2C5) с учетом участков SCR [18].

1.5. Конструирование участков SCR модели.

Базовая ("скелетная") структура SCR (только координаты $C\alpha$ атомов), полученная на этапе 1.3, представляет собой усредненную пространственную структуру консервативных областей белков-гомологов и лишена реальной белковой геометрии. Процедура построения SCR модели ("Build SCR") создает скелет каждого участка SCR путем совмещения фрагмента одного из известных гомологов и соответствующего участка базовой структуры. Затем определяется конформация боковых цепей аминокислотных остатков моделируемого белка с учетом вторичной структуры основной цепи и геометрии боковых цепей соответствующих остатков в каждом из гомологов. Для этой цели используется специальный подход, основанный на знании (knowledge-based), и специальная база данных, содержащая набор соответствующих правил.

1.6. Конструирование структурно-вариабельных областей модели.

Процедура "Add Loops" завершает моделирование белковой модели путем конструирования структурно-вариабельных участков (Structurally Variable Regions, SVR), часто называемых также участками петель (loop regions). Для этого используются структуры белковых петель из гомологов и/или из общей белковой базы данных. Для каждой петли программа пытается найти фрагменты из известных структур гомологов, которые будут совместимы с уже построенными частями модели. Затем она использует информацию об аминокислотных последовательностях для выбора наилучшего фрагмента для построения конечной модели. Если один из гомологов с известной структурой имеет в соответствующей области петлю той же длины, то это считается хорошим вариантом. В противном случае находят петлевые фрагменты, геометрия которых совместима с геометрией консервативных областей, граничащих с петлей в структуре модели. Далее на основе информации о геометрии и аминокислотных последовательностях выбирают наилучший петлевой фрагмент среди найденных вариантов.

1.7. Моделирование структуры гема и его встраивание в модель белка.

Так как P450 является гем-содержащим белком, то при моделировании его пространственной структуры необходимо решать вопрос о корректном моделировании гема и его позиционировании в модели. Известно, что ион железа в геме связан с помощью двух ионных и двух координационных связей. В структуре цитохрома P450 атом железа связан также с остатком цистеина координационной связью Fe-S, которая не может рассматриваться как простая ковалентная связь. Необходимо также иметь в виду, что система колец в молекуле гема имеет планарную структуру [21]. В тоже время принципы молекулярной механики, используемые во всех программах молекулярного моделирования белков, не позволяют работать с координационными связями. При чтении молекулярных файлов, содержащих гем, эти программы (в том числе и Sybyl) безуспешно пытаются описать его структуру в виде ароматической системы. Таким образом, для относительно корректного представления структуры гема в рамках молекулярной механики с использованием стандартных типов атомов и связей, описанных в силовом поле, а также для стабилизации геометрии гема во время выполнения процедур минимизации энергии и моделирования молекулярной динамики, необходимо вводить ряд приближений и упрощений. Они заключаются в следующем: 1) структура гема представляется в виде системы конъюгированных двойных связей; 2) все связи рассматриваются как ковалентные; 3) атом железа представляется типом атома "Fe", используемым в молекулярной механике; 4) атомы азота представляются как плоские аминные. В результате таких упрощений гем сохраняет планарную структуру при выполнении процедуры минимизации энергии, что подтверждается минимальными оценками искажения планарности структуры типа "out-of-plane distortions".

Другой проблемой, связанной с моделированием гема в P450, является его встраивание в структуру модели белка и позиционирование. Проблема решается с помощью выполнения следующих процедур: 1) структура моделируемого P450

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

совмещается в пространстве со структурой белка-образца (в качестве опорных точек для совмещения используются $C\alpha$ атомы аминокислотных остатков гем-связывающего домена); 2) структура гема копируется из белка-образца в модель; 3) в структуре модели формируется связь Fe-S(Cys).

1.8. Уточнение модели (*model refinement*).

На завершающем этапе моделирования полученная модель P450 должна быть уточнена и оптимизирована с помощью последовательного выполнения следующих процедур: 1) стерическая оптимизация, представляющая собой нахождение и разрешение стерических конфликтов (сильное перекрытие сфер Ван дер Ваальса атомов); 2) энергетическая оптимизация - нахождение молекулярной конформации с минимальной величиной потенциальной энергии (энергетическая минимизация и моделирование молекулярной динамики в водном окружении).

2. Цели и задачи молекулярного моделирования.

Любые виды моделирования имеют своей целью представление сложных объектов или явлений в упрощенном виде (модели) для выяснения особенностей строения, свойств и функций. Молекулярное моделирование пространственной структуры белка не является исключением и также используется для достижения подобных целей.

В случае моделирования P450 можно выделить, по крайней мере, 4 основных цели:

- 1) моделирование фолда белка;
- 2) моделирование структуры активного центра;
- 3) моделирование функционально важных участков поверхности белка;
- 4) моделирование мембранного "якоря" белка.

В зависимости от цели может создаваться модель с детализацией тех участков структуры, которые необходимы для решения определенных задач. Так, например, может быть построена модель P450 без части поверхностных петель, но с уточненной структурой активного центра фермента. Такая модель вполне достаточна для изучения молекулярных механизмов катализа и взаимодействия с низкомолекулярными лигандами (субстратами и ингибиторами), а также для поиска новых лигандов и конструирования лекарств [22-24]. В тоже время, модель без точного строения активного центра, но с уточненной оптимизированной структурой поверхности белковой глобулы может быть использована для изучения взаимодействия P450 с белками-партнерами [25, 26], моделирования точечных мутаций на поверхности белка с целью получения его водорастворимой формы [27], поиска мест связывания низкомолекулярных лигандов [28, 29] и др.

3. Проблемы моделирования P450 по гомологии.

3.1. Недостаточное число гомологов.

Как уже было сказано выше, существует огромный разрыв между числом известных последовательностей P450 (более тысячи последовательностей) [7] и числом известных кристаллических структур (11 структур, табл. 1). То есть число структур составляет всего около 1% от числа последовательностей P450. Так как в надсемействе P450 наблюдается большое разнообразие белков (250 семейств и 337 подсемейств [7]), то совершенно очевидно, что число гомологов с известной трехмерной структурой недостаточно для моделирования многих представителей P450.

3.2. Низкая идентичность известных гомологов.

Проблемой является не только недостаток числа известных трехмерных структур P450, но и низкая идентичность их последовательностей между собой. Табл. 2 содержит результаты выравнивания последовательностей с помощью программы ALN [30]. Низкая идентичность значительно ухудшает результаты сравнительного анализа как последовательностей (парное и множественное выравнивание), так и пространственных структур (структурное выравнивание).

3.3. Проблема оптимального структурного выравнивания.

Как было описано выше, при моделировании структуры белка по гомологии обязательным этапом является выявление структурно-консервативных участков (SCR) с помощью совмещения известных структур-гомологов в пространстве

(процедура структурного выравнивания). При низкой идентичности последовательностей гомологов (как в случае с P450) возникает проблема оптимального совмещения структур. Вариабельность структур приводит также к уменьшению числа и значительному укорочению выявляемых SCR и, соответственно, росту участков петель (рис. 2).

Таблица 2. Матрица попарных сравнений между последовательностями аминокислотных остатков цитохромов P450 с известной пространственной структурой, доступных в PDB. Сходство между последовательностями выражено в процентах идентичности, рассчитанной на основе выравнивания, выполненного с помощью программы ALN [30] с матрицей замен PAM250.

| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----------------------------------|-----------------------------|------|------|------|------|------|------|------|------|------|------|------|
| Название P450 PDB индекс | 1DT6 | 1E9X | 1JFB | 1CP4 | 1JPZ | 1EGY | 1CPT | 1IO7 | 1N40 | 1GWI | 1LFK | 1N97 |
| | 1 P450 2C5 (CYP2C5) 1DT6 | - | 23 | 22 | 23 | 25 | 23 | 22 | 22 | 21 | 25 | 24 |
| 2 P450 51 (CYP51) 1E9X | 23 | - | 24 | 22 | 29 | 27 | 26 | 25 | 24 | 25 | 27 | 26 |
| 3 P450 Nor (CYP55A1) 1JFB | 22 | 24 | - | 28 | 22 | 31 | 30 | 27 | 27 | 29 | 29 | 25 |
| 4 P450cam (CYP101) 1CP4 | 23 | 22 | 28 | - | 23 | 25 | 27 | 24 | 25 | 27 | 25 | 25 |
| 5 P450 bm-3 (CYP102) 1jz | 25 | 29 | 22 | 23 | - | 23 | 27 | 23 | 23 | 24 | 23 | 29 |
| 6 P450 Eryf (CYP107A1) 1EGY | 23 | 27 | 31 | 25 | 23 | - | 28 | 32 | 30 | 37 | 33 | 26 |
| 7 P450 terp (CYP108) 1CPT | 22 | 26 | 30 | 27 | 27 | 28 | - | 28 | 26 | 30 | 29 | 24 |
| 8 P450 119 (CYP119) 1IO7 | 22 | 25 | 27 | 24 | 23 | 32 | 28 | - | 28 | 27 | 25 | 25 |
| 9 P450 MT2 (CYP121) 1N40 | 21 | 24 | 27 | 25 | 23 | 30 | 26 | 28 | - | 29 | 28 | 25 |
| 10 P450 154C1 (CYP154C1) 1GWI | 25 | 25 | 29 | 27 | 24 | 37 | 30 | 27 | 29 | - | 30 | 30 |
| 11 P450 OxyB (CYP165B) 1LFK | 24 | 27 | 29 | 25 | 23 | 33 | 29 | 25 | 28 | 30 | - | 26 |
| 12 P450 175A1 (CYP175A1) 1N97 | 26 | 26 | 25 | 25 | 29 | 26 | 24 | 25 | 25 | 30 | 26 | - |

Участки SCR располагаются преимущественно в N-концевой части белка. Большинство участков SCR соответствуют элементам вторичной структуры P450. По-видимому, эти участки формируют фолд цитохрома P450. Большинство консервативных элементов локализованы в области α -спиралей, т.к. именно в этой области происходит фиксация гема. В структурно-консервативное ядро P450 входят следующие элементы начиная с N-конца: α -спираль I и K, участок β -листа 1, "меандр" (извилины), гем-пептид.

Меандр представляет собой неупорядоченную структуру, локализованную на поверхности белка и выступающую из глобулы. Мутации в этой области, приводят к нарушению взаимодействия P450 с цитохром P450-редуктазой.

Гем-пептид является самым консервативным участком при сравнении множества различных форм цитохромов P450. В гем-пептид входит петля и участок α -спирали L. Именно в петле находится остаток цистеина, являющийся пятым аксиальным лигандом гема. Гем размещен наподобие сэндвича между гем-пептидом и α -спиралью I.

Спираль I является наиболее протяженной из всех спиралей и несет не только структурную нагрузку по фиксации гема, но и функциональную - остаток тирозина участвует в активации кислорода в реакционном центре фермента.

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

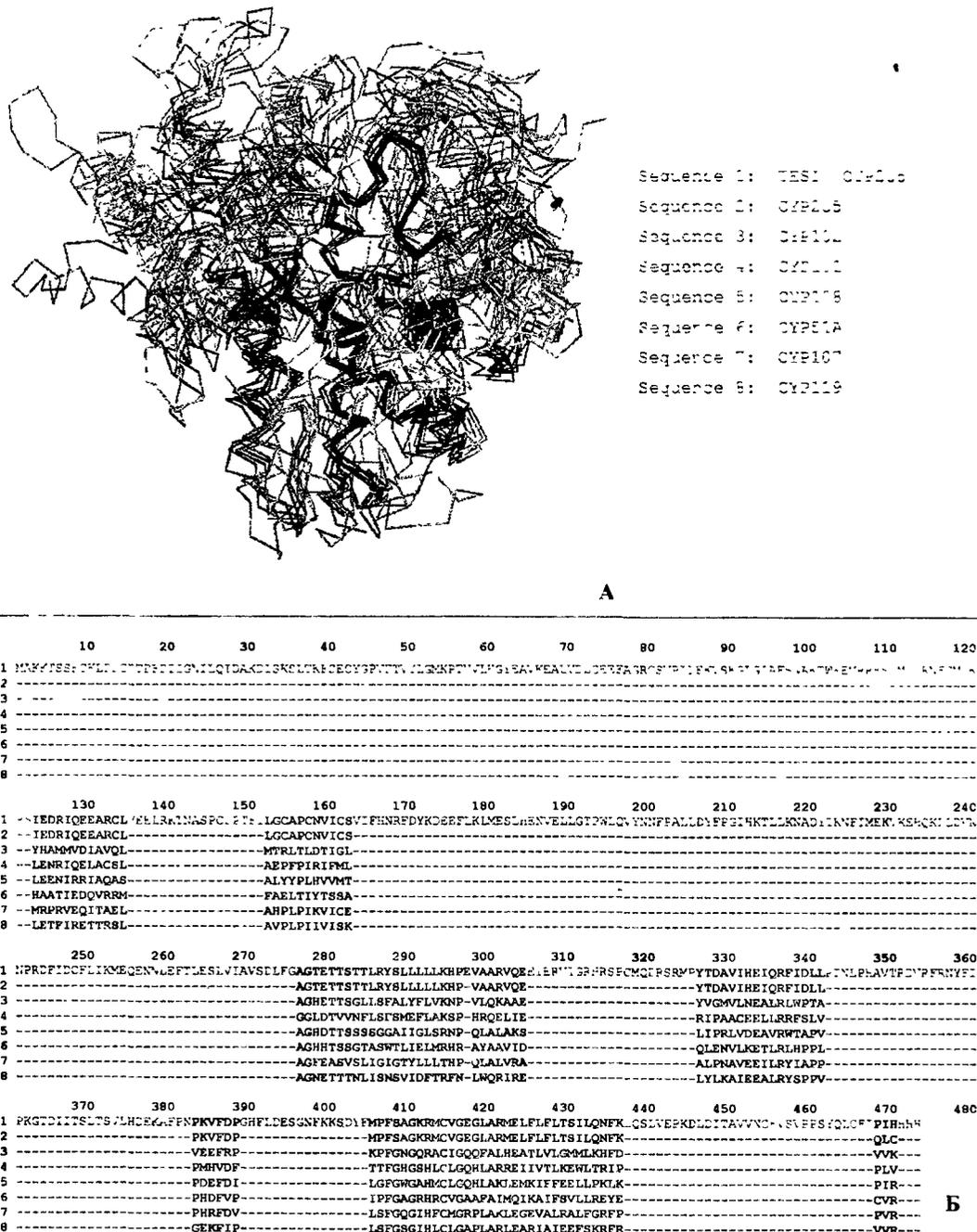


Рисунок 2.

Структурное выравнивание всех гомологов P450 с известной структурой.

А - Результат пространственного совмещения структур (области SCR показаны гемным цветом).

Б - Выравнивание последовательностей с учетом усредненных областей SCR.

Таким образом, консервативное ядро P450 несет структурную нагрузку по формированию фолда, а также отвечает за обеспечение базовых функций катализа: активацию кислорода и взаимодействие с белком-партнером.

3.4. Моделирование петель.

Непростой сам по себе этап моделирования свободных петель значительно усложняется из-за уменьшения числа и длины SCR, что приводит к необходимости

моделирования чрезмерно длинных петель. Анализ найденных в белках из PDB различных фрагментов требует дополнительной экспертной оценки, основанной на экспериментальных данных о структурных особенностях P450, учете расположения и взаимного влияния петель друг на друга. В настоящее время появилась возможность построения участков петель с помощью методов моделирования структуры белка *de novo*, основанных на выборе подходящего варианта петли среди набора случайно построенных конформаций с учетом расчетных энергетических параметров [31]. Однако данный подход также требует тщательной экспертной оценки и глубокой оптимизации методами энергетической минимизации и моделирования молекулярной динамики [32 -34].

3.5. Моделирование мембранного якоря.

Проблема моделирования мембранного якоря является, несомненно, самой сложной. Это обусловлено следующими причинами:

1) Среди немногочисленных известных структур мембранных белков (около 30 структур в PDB [5]) нет мембранных участков, имеющих высокую гомологию с мембранными якорями P450. Этот факт значительно ограничивает применимость метода моделирования по гомологии для решения данной задачи.

2) Возможно создание "мозаичной" модели мембранного якоря в виде комбинации небольших фрагментов, построенных по гомологии с различными гидрофобными фрагментами из разных белков в PDB (преимущественно фрагменты, расположенные внутри белковых глобул). Однако при этом подходе, как правило, возникает неоднозначность модели, обусловленная различными вариантами комбинирования фрагментов. Таким образом, результат моделирования в значительной степени зависит от процедур поиска и выбора гомологичных фрагментов. Получаемые варианты модели являются практически равновероятными и окончательный выбор должен базироваться на сверке модели с экспериментальными данными о мембранной топологии моделируемого P450. Это могут быть данные о локализации белка в мембране (одностороннее или трансмембранное расположение мембранной части), данные об аминокислотной последовательности мембранного якоря, данные о местах действия протеаз, данные о химических модификациях, результаты точечного мутагенеза и т.д.

3) Конформация боковых групп аминокислотных остатков в мембранном якоря обусловлена их взаимодействием с липидным бислоем. Следовательно, для их корректного моделирования требуется сложная молекулярная система в виде модели бислоя, разделяющего две водные фазы. Создание подобных модельных систем представляет собой отдельную сложную задачу, для решения которой требуются специальные программные средства и большие вычислительные мощности.

Для решения этих проблем в некоторых случаях используется ряд упрощений, что иногда может быть оправдано. Так, например, нами была успешно выполнена работа по реконструкции полноразмерного цитохрома *b5 (b5)* [35] путем моделирования его мембранной части в виде "мозаичной" модели с последующим ее объединением с кристаллической структурой трипсинового водорастворимого фрагмента *b5*, доступной в PDB. Первоначальная конформация мембранного якоря *b5* была собрана из отдельных коротких преимущественно гидрофобных фрагментов, найденных в различных белках в PDB на основе гомологии последовательностей. Для оптимизации структуры мембранного якоря *b5* была использована упрощенная двухфазная система растворителей (вода/бензол) имитирующая границу фаз (вода/мембрана).

3.6. Глубокая оптимизация модели (*refinement*).

Первоначально построенная модель является предварительной и содержит много неточностей. Поэтому для получения более реалистичной модели требуется выполнение нескольких этапов ее "доводки" (уточнения). Два первых этапа упоминались выше (п. 2.8) - это ликвидация стерических конфликтов и энергетическая оптимизация. В результате получается формально правильная модель, не имеющая грубых ошибок. Однако точность построения поверхностных

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

петель остается довольно низкой. Данная проблема является общей для моделирования различных белков. Считается, что для окончательного уточнения модели необходимо выполнить ее глубокую оптимизацию путем моделирования длительной молекулярной динамики в водном окружении с многократной энергетической оптимизацией. Эти процедуры позволяют получить более стабильную, сбалансированную со средой структуру. Однако существует ряд трудностей как технического, так и теоретического характера. Так как общее число атомов в системе может достигать 10^5 (с учетом полноатомной модели белка и окружающей ее молекул воды), то для выполнения подобных процедур требуются большие вычислительные мощности и очень много вычислительного времени. В ходе моделирования молекулярной динамики конформация белка удаляется от стартовой (в основном за счет "дрейфа" петель и их конформационных изменений). Изменения контролируются по средней величине среднеквадратичных отклонений (RMS) позиций $\text{C}\alpha$ атомов относительно их стартовых положений. В дальнейшем траектория изменения структуры выходит на плато, что свидетельствует о ее относительной стабилизации, однако наблюдаются тепловые колебательные движения некоторых ее подвижных элементов (преимущественно тех же петель). В результате модель может быть представлена в виде набора равновероятных структур и возникает проблема выбора одной из них в качестве конечной.

Необходимо отметить, что остается открытым вопрос об адекватности глубоко оптимизированной модели и структуры нативного белка в растворе. В кристалле у белка возникают значительные ограничения для ряда подвижных элементов. Более того, упаковка белка в кристалле может приводить к тому, что некоторые петли принимают позиции и конформации, не характерные для них в растворе. Так в контрольных экспериментах по моделированию молекулярной динамики в водной фазе также наблюдается дрейф, когда в качестве стартовой структуры была взята кристаллическая структура белка из PDB. В тех случаях, когда структуру белка удается исследовать непосредственно в растворе с помощью метода ЯМР, как правило, наблюдается набор резонансных структур, отличающихся друг от друга позициями подвижных элементов.

4. Критерии оценки успешного моделирования P450 по гомологии.

4.1. Эксперименты CASP (*Critical Assessment of protein Structure Prediction*).

В основе Интернет-экспериментов CASP лежит идея проверки алгоритмов и точности методов молекулярного моделирования структуры белков путем сравнения кристаллических структур с компьютерными моделями одних и тех же белков [36]. Для обеспечения объективности и достоверности экспериментов была предложена следующая схема:

1) экспериментаторы, у которых имеются еще неопубликованные данные о структуре новых белков, информируют об этом других участников эксперимента и размещают описание объектов (название, источник и аминокислотную последовательность белка) в Интернете на сервере организаторов эксперимента;

2) группы ученых, занимающиеся компьютерным моделированием, используя предоставленную информацию о последовательностях, создают компьютерные модели целевых белков и отправляют их на тот же сервер до момента публикации экспериментальных результатов;

3) группа независимых экспертов сравнивает модели целевых белков с экспериментальными данными о трехмерной структуре с использованием набора различных методов проверки и делает заключения.

В экспериментах CASP было показано, что модели белков, полученные с помощью методов моделирования по гомологии, имеют хорошую точность, если идентичность последовательностей моделируемого белка и белка-образца превышает 40%. При этом точность позиций большинства $\text{C}\alpha$ атомов (до 90%), оцененная по ошибке средней (RMSD), оказывается менее 1 Å, что соответствует кристаллической структуре с низким разрешением (~ 2,5 Å). Если идентичность

последовательностей находится в пределах 30-40%, точность ухудшается до ~ 1,5 Å для 80% остатков. В тех случаях, когда идентичность последовательностей менее 30% (типичная ситуация для большинства цитохромов P450), точность падает до 4-6 Å и основным фактором, ограничивающим использование данного подхода, является проблема оптимального локального выравнивания последовательностей моделируемого белка и белка-образца с учетом участков SCR. Доля неправильно выровненных аминокислотных остатков может достигать 20% и это число быстро растет при меньших значениях идентичности последовательностей. Это ограничивает применимость методов моделирования по гомологии, так как в настоящее время ни одна из существующих технологий моделирования не может компенсировать ошибки некорректного первоначального выравнивания последовательностей [37-39].

4.2. Оценка результатов выравнивания.

Если две аминокислотных последовательности имеют длину более 100 остатков и идентичность более 25%, то существует большая вероятность, что они гомологичны. Если идентичность равна 15-25%, последовательности еще могут иметь сходство, но для подтверждения этого необходимо выполнение процедуры "смешивания" (jumbling) для определения статистической значимости выравнивания (сравнение оценки реального выравнивания с усредненной оценкой выравниваний случайных последовательностей, полученных путем случайного перемешивания текста оригинальных последовательностей). Если идентичность последовательностей менее 15%, то, по всей вероятности, они не имеют сходства.

Исходя из вышесказанного, можно сделать заключение, что основным критерием успешности моделирования P450 по гомологии может являться оценка результатов выравнивания последовательностей моделируемого белка и белка-образца (локальное выравнивание с учетом участков SCR) с помощью параметра "значимость" (Significance Score) [40, 41]. В табл. 3 приведены основные принципы оценки результатов выравнивания последовательностей с помощью этого параметра.

Таблица 3. Оценка выравнивания с помощью параметра "значимость" (Significance Score, X) [40, 41].

| Определение | |
|---|--------------------|
| <p>Если идентичность двух последовательностей находится в диапазоне 15-25%, то необходимо проверять статистическую значимость результатов выравнивания последовательностей. Для этого применяется процедура "смешивания" (jumbling), представляющая собой сравнение результата реального выравнивания (S_0) с усредненным результатом выравниваний случайных последовательностей (S), полученных путем случайного перемешивания текста оригинальных последовательностей:</p> $X = \frac{S - S_0}{D},$ <p>где X – параметр «значимость», D – стандартное отклонение результатов выравниваний смешанных (случайных) последовательностей.</p> <p>Качество будущей модели трехмерной структуры белка, построенной по гомологии, может быть оценено по величине X:</p> | |
| «значимость» (X) | Качество модели |
| < 20 | плохое |
| 20 - 30 | удовлетворительное |
| 30 - 50 | среднее |
| 50 - 80 | хорошее |
| > 80 | очень хорошее |

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

4.3. Параметр "значимость" для предварительной оценки возможности моделирования P450.

На основе гомологии с известными пространственными структурами P450 (табл. 1) может быть осуществлено моделирование ряда белков из различных семейств и подсемейств P450. Нами были рассчитаны соответствующие величины параметра "значимость" (X) для моделирования белков из различных подсемейств P450 (табл. 4). Как уже отмечалось выше, все последовательности P450 с известной трехмерной структурой имеют низкое сходство друг с другом, поэтому каждый из этих белков может выступать в качестве индивидуального белка-образца для моделирования индивидуального набора представителей разных подсемейств P450. Эти индивидуальные данные были суммированы для получения имеющейся на начало 2003 г. общей картины для всего надсемейства P450 (рис. 3А). Необходимо отметить, что в последние три года данная картина довольно быстро менялась в связи с появлением новых структур (рис. 3Б). Естественно предполагать, что в дальнейшем при появлении в PDB новых структур P450 ситуация будет улучшаться и дальше.

Таблица 4. Распределение 287 подсемейств P450 по группам оценки возможности их моделирования на основе гомологии с известными P450 (см. табл. 3) по величине параметра "значимость" (X).

| № | P450 | Значимость (X) (Качество моделей) | | | |
|-----|----------|--------------------------------------|---------------------|----------------------|----------------------|
| | | < 20 (неудовл.) | 20 – 30 (удовл.) | 30 – 50 (среднее) | 50 – 80 (хорошее) |
| 1. | CYP101A | 284 | 1 | 1 | 1 |
| 2. | CYP102A | 283 | 4 | - | 1 |
| 3. | CYP108A | 280 | 7 | - | - |
| 4. | CYP55A | 280 | 5 | 1 | 1 |
| 5. | CYP107A1 | 259 | 21 | 6 | 1 |
| 6. | CYP2C5 | 251 | 22 | 9 | 5 |
| 7. | CYP119 | 281 | 6 | - | - |
| 8. | CYP51 | 286 | - | 1 | - |
| 9. | CYP165B | 270 | 14 | 3 | - |
| 10. | CYP121A | 274 | 13 | - | - |
| 11. | CYP154C1 | 271 | 15 | 1 | - |
| 12. | CYP175A1 | 287 | - | - | - |

До 2001 года лучшим известным гомологом для моделирования P450 эукариот был P450 102 (P450 bm-3), однако недостаточно высокий уровень идентичности его последовательности с последовательностями большинства моделируемых белков (менее 30%) не позволял создавать модели с хорошей точностью. В 2000 году был получен первый кристалл эукариотического цитохрома P450 (модифицированный P450 2C5 кролика) [9, 10] и в настоящее время его структура является наилучшим образцом для моделирования других P450 эукариот по гомологии (идентичность между некоторыми последовательностями моделируемых белков и CYP2C5 достигает 40-60%).

5. Методы проверки правильности моделей P450.

5.1. Структурная и топологическая проверка.

Самая первоначальная проверка пространственной модели белка представляет собой принципиальную оценку корректности его фолда. Существуют три основных признака некорректного фолда модели [42]: (1) экспозиция гидрофобных остатков в растворитель, т.е. расположение гидрофобных аминокислотных остатков на поверхности белка и их контакт с водой; (2) присутствие во внутренней структуре модели заряженных или полярных боковых цепей, которые недоступны для воды и не участвуют в образовании солевых мостиков или водородных связей; (3)

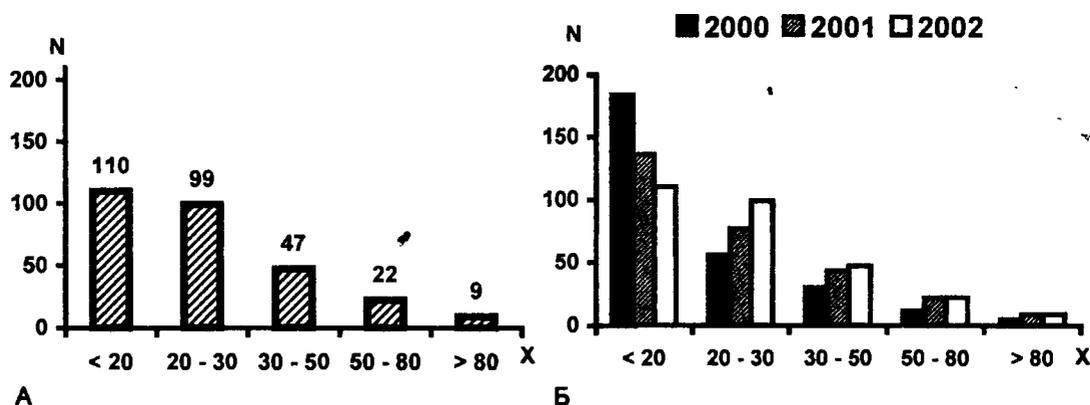


Рисунок 3.

Оценка возможности моделирования цитохромов Р450 из различных подсемейств по гомологии с известными пространственными структурами, имеющимися в PDB на начало 2003 г.

А - Гистограмма распределения числа (N) типичных представителей из 337 подсемейств цитохромов Р450 от величины параметра "значимость" (X), рассчитанной путем сравнения их последовательностей с последовательностями 9 цитохромов Р450 с известной пространственной структурой. Б - Изменение общей картины для всего надсемейства Р450 в возможности моделирования в течение последних 3 лет в связи с появлением в PDB новых кристаллических структур.

положительная величина свободной энергии сольватации, что, собственно, является прямым результатом первых двух признаков.

5.2. Моделирование молекулярной динамики.

Одной из серьезных проверок модели является моделирование молекулярной динамики (МД) в водном окружении. При этом структура модели должна быть стабильной - сохранять фолд и элементы вторичной структуры. Не редко построенная модель Р450 не выдерживает такой проверки и "разрушается" в ходе МД - происходит увеличение ее объема и "разрыхление" белковой глобулы за счет входа в нее воды, разрушение вторичных структур и выход белковых петель в водную фазу далеко от поверхности молекулы.

5.3. Статистические методы проверки.

Стандартным общепринятым подходом является проверка модели с помощью различных программ, базирующихся на статистических данных, полученных из анализа пространственных структур всех белков в белковом банке PDB.

Пакет программ *ProCheck* [43] может быть использован для проверки качества модели белка путем анализа геометрии аминокислотных остатков по сравнению со "стерео-химическими" параметрами, полученными из кристаллических структур белков с высоким разрешением. Результаты анализа модели могут быть представлены в виде диаграмм Рамачандрана. Области белка, отмечаемые *ProCheck* как запрещенные (необычные), не обязательно являются ошибками моделирования. Такие области можно обнаружить и при анализе с помощью *ProCheck* кристаллических структур. Так, например, на диаграммах для кристаллических структур CYP2C5 и CYP102 (рис. 4) можно видеть отдельные точки в запрещенной области (на светлом фоне). В любом случае точки в неблагоприятной области диаграммы Рамачандрана, обнаруживаемые программой *ProCheck* как в модели, так и в кристалле, требуют особого внимания и разумного объяснения.

Программа *ProTable* [44] является одним из модулей программного комплекса Sybyl [12] и предназначена для анализа и проверки качества белковой структуры по различным критериям: геометрия и стереохимия, площадь доступной для растворителя поверхности, конформационные вероятности боковых цепей, конформация скелета и боковых цепей, вторичные структуры, совместимость аминокислотных остатков с конкретным структурным окружением

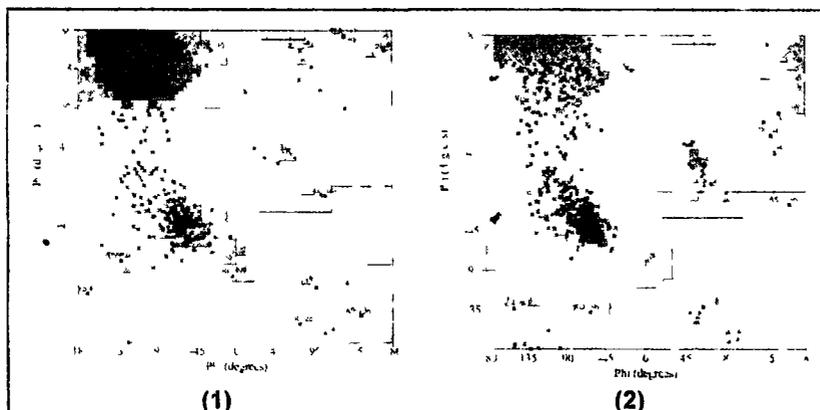


Рисунок 4

Диаграммы Рамачандрана, рассчитанные для кристаллических структур CYP2C5 (1) и CYP102 (2) с помощью программы ProCheck

("энергия", вычисляемая другим модулем Sybyl - программой MatchMaker, см. ниже) ProTable представляет собой интерактивную электронную таблицу, в которой каждая строка содержит различную информацию об отдельном аминокислотном остатке, отражающую как локальную структурную информацию о каждом остатке, так и качество белковой структуры в данном месте. Важными дополнительными свойствами ProTable являются возможность выборочной раскраски остатков, получение графов, отражающих свойства отдельных остатков, интерактивный анализ корреляции табличных данных с трехмерной структурой. Все параметры, рассчитанные в ProTable, могут быть визуальным образом отражены на трехмерной структуре белка. Таким образом, программа ProTable может быть использована в качестве аналитического инструмента для проверки качества белковой модели, а также для ее сверки с известными экспериментальными данными (см. ниже).

Программа MatchMaker [45], являющаяся модулем программного комплекса Sybyl [12], использует подход, известный как "обратный фолдинг". Она предназначена для анализа соответствия аминокислотной последовательности белка его трехмерной структуре. Путем сравнения программы определяет, в какую структуру может быть свернут новый белок. Рассчитываемые программой "энергии MatchMaker" являются мерой оценки совместимости аминокислотных остатков с их локальным структурным окружением. MatchMaker может быть использован для независимой проверки качества и относительной корректности модели путем определения соответствия фолда модели белка и аминокислотной последовательности.

Программа WHAT IF [46] проверяет соответствие различных параметров модели существующим статистическим данным: длины и углы связей, доноры водородных связей внутри глобулы, *cis*-конформации пептидных связей, конформации боковых цепей и их планарность, упаковка остатков пролина, качество упаковки белка, ротамеры боковых цепей, симметрия, торсионные углы, кластеры воды, занятость пространства атомами и т.д.

Программа Prosa II [47] позволяет проверить качество пространственной структуры белка. Если оценка модели выходит за характерный диапазон природных белковых фолдов, то это означает, что в структуре модели есть проблемы. Так называемые энергетические графы позволяют выявить проблемные области.

Программа Profiles-3D [48] ищет в базе данных совместимые с моделью трехмерные профили с использованием классификации аминокислот по их окружению. Программа позволяет предсказывать пространственную структуру белка путем измерения совместимости его последовательности с известными белковыми структурами с последующим использованием метода обратного

фолдинга. Программа *Profiles-3D* может быть использована для проверки качества построенной модели путем сравнения ее структуры со своей собственной последовательностью.

5.4. Проверка соответствия модели косвенным экспериментальным данным.

Для проверки корректности модели белка может быть использован ряд непрямых экспериментальных данных, таких как спектроскопические исследования и результаты анализа структурных элементов на поверхности белка. В частности, модель должна полностью соответствовать или, по крайней мере, не противоречить данным о химической модификации аминокислотных остатков на поверхности белка, результатам анализа доступности специфических участков белковой цепи для действия протеолитических ферментов, результатам экспериментов по точечному мутагенезу, анализу антигенных детерминант (антигенное картирование) и т.д. Такого рода проверки осуществляются в интерактивном режиме методами молекулярной графики: раскраска участков модели в соответствии с экспериментальными данными, построение доступной для воды молекулярной поверхности и визуальный анализ соответствия модели.

ЗАКЛЮЧЕНИЕ. В заключение этого краткого обзора необходимо отметить, что основным фактором, ограничивающим моделирование P450 по гомологии, является проблема корректного выравнивания последовательностей моделируемого белка и белка-образца при низкой идентичности их последовательностей (типично - ниже 30%). Появление в PDB новых пространственных структур P450 улучшает общую ситуацию в моделировании P450. Особенно это касается тех случаев, когда появляется новая структура P450, последовательность которой имеет высокую идентичность с последовательностями других P450, моделирование которых представляет интерес.

Данный обзор написан при благотворительной поддержке Регионального Общественного Фонда содействия отечественной медицине.

Авторы благодарны фирме Tripos GmbH (Германия) за научную и техническую поддержку.

ЛИТЕРАТУРА

1. Poulos T.L. (1995) *Current Opinion in Structural Biology*, **5**, 767-774.
2. Jan Drenth (1994) *Principles of Protein X-ray Crystallography*. Springer-Verlag
3. Wallin E., von Heijne G. (1998) *Protein Sci.*, **7**, 1029-1038.
4. 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. (<http://www.rcsb.org/pdb>)
5. Ulmschneider M.B., Sansom M.S.P. (2001) *BBA*, **1512**, 1-14.
6. Grisshammer R., Tate C.G. (1995) *Q. Rev. Biophys.*, **28**, 315-422.
7. Cytochromes P450 Database (CPD) (<http://cpd.ibmh.msk.su>).
8. Chang Y.T., Stiffelman O.B., Loew G.H. (1996) *Biochimie*, **78**, 771-779.
9. Williams P.A., Cosme J., Sridhar V., Johnson E.F., McRee D.E. (2000) *Mol. Cell.*, **5**, 121-131.
10. Williams P.A., Cosme J., Sridhar V., Johnson E.F., McRee D.E. (2000) *J. Inorg. Biochem.*, **81**, 183-190.
11. Skvortsov V.S., Belkina N.V., Ivanov A.S. (2000) *Molecular Simulation*, **24**, 369-378.
12. SYBYL 6.7.1, Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA. (<http://www.tripos.com>).
13. Insight II, Accelrys (subsidiary of Pharmacoepia Inc.). <http://www.accelrys.com/insight/>
14. Quanta, Accelrys (subsidiary of Pharmacoepia Inc.). <http://www.accelrys.com/quanta/>
15. Sanchez R., Sali A. (2000) *Methods Mol. Biol.*, **143**, 97-129.
16. Sanchez R., Sali A. (1997) *Proteins.*, Suppl **1**, 50-58.

МОДЕЛИРОВАНИЕ P450: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

- 17 *Blundell T.L., Carney D.P., Gardner S., Hayes F.R.F., Howlin B., Hubbard T.J.P., Overington J.P., Singh D.A., Sibanda B.L. and Sutcliffe M.J.* (1988) *Eur. J. Biochem.*, **172**, 513-520.
- 18 *Needleman S., Wunsch C.* (1970) *J. Mol. Biol.*, **48**, 443-453.
- 19 *Fredman, M.* (1984) *Bull. of Mathematical Biology*, **46**, 553-566.
- 20 *Сеченых А.А., Дубанов А.В., Скворцов В.С., Иванов А.С., Арчаков А.И.* (2002) *Вопр. мед. хим.*, **48**, 526-538.
- 21 Database PROMISE (The Prosthetic groups and Metal Ions in Protein Active Sites Database Version 2.0), Chapter "P450 proteins". (<http://metallo.scripps.edu/PROMISE/P450.html>).
- 22 *Belkina N.V., Lisurek M., Ivanov A.S., Bernhardt R.* (2001) *J. Inorg. Biochem.*, **87**, 197-207.
- 23 *Белкина Н.В., Скворцов В.С., Иванов А.С., Арчаков А.И.* (1998) *Вопр. мед. хим.*, **44**, 464-473.
- 24 *Ivanov A.S., Sechenykh A.A., Skvortsov V.S., Shkrob A.M., Archakov A.I.* (1997) *Proceedings of the 2nd Int. Conf. Mol. Str. Biology, Vienna*, 126/P62.
- 25 *Bridges A., Gruenke L., Chang Y.T., Vakser I.A., Loew G., Waskell L.* (1998) *J. Biol. Chem.*, **273**, 17036-17049.
- 26 *Skvortsov V.S., Sechenykh A.A., Ivanov A.S., Ivanov Y.D., Archakov A.I.* (1997) *FASEB J.*, **11**, A786/P85.
- 27 *Skvortsov V.S., Belkina N.V., Ivanov A.S.* (2000) *Molecular Simulation*, **24**, 369-378.
- 28 *Иванов А.С., Дубанов А.В., Скворцов В.С., Арчаков А.И.* (2002) *Вопр. мед. хим.*, **48**, 304-315.
- 29 *Ivanov A.S., Dubanov A.V., Skvortsov V.S., Archakov A.I.* (1999) *Proceedings of the 3rd Int. Conf. Mol. Str. Biology, Vienna*, 148/P92.
- 30 *Gotoh O.* (1999) *Adv. Biophys.* **36**, 159-206.
- 31 *Huang X.* (1994) *Comp. App. Biosciences* **10**, 227-235.
- 32 *Koehl P., Levitt M.* (1999) *J. Mol. Biol.* **293**, 1161-1181.
- 33 *Fiser A., Do R.K., Sali A.* (2000) *Protein Sci.* **9**, 1753-1773.
- 34 *Das B., Meirovitch H.* (2001) *Proteins* **43**, 303-314.
- 35 *Иванов А.С., Скворцов В.С., Арчаков А.И.* (2000) *Вопр. мед. хим.*, **46**, 615-625.
- 36 CASP experiment (Protein Structure Prediction Center, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California, USA) (<http://predictioncenter.llnl.gov>).
- 37 *Sanchez R., Sali A.* (1997) *Current Opinion in Structural Biology*, **7**, 206-214.
- 38 *May A.C.W., Blundell T.L.* (1994) *Current Opinion in Biotechnology*, **5**, 355-360.
- 39 *Al-Lazikani B., Jung J., Xiang Z., Honig B.* (2001) *Current Opinion in Chemical Biology*, **5**, 51-56.
- 40 *Barton G.J.* (1990) *Methods in Enzymology*, **183**, 403-428.
- 41 *Henikoff S.* (1994) *In Informatics and Genome Projects*, (Smith D.W. ed.), Academic Press, pp 87-117.
- 42 *Novotny J., Rashin A.A., Bruccoleri R.E.* (1988) *Proteins*, **4**, 19-30.
- 43 *Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J.M.* (1993). *J. Appl. Cryst.*, **26**, 283-291. (<http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html>)
- 44 *Sybyl's ProTable* (<http://www.tripos.com/sciTech/inSilicoDisc/media/LitCtr/protable.pdf>).
- 45 *Godzik A., Kolinski A., Skolnick J.* (1993) *J. Comp. Aid. Mol. Design*, **7**, 397-438. (<http://www.tripos.com/admin/LitCtr/matchmaker.pdf>)
- 46 *Vriend G.* (1990) *J. Mol. Graph.* **8**, 52-56. (<http://www.cmbi.kun.nl/whatif/>)
- 47 *Sippl M.J.* (1993) *Proteins*, **17**, 355-362. (<http://lore.came.sbg.ac.at/Services/prosa.html>)
- 48 *Luthy R., Bowie J.U., Eisenberg D.* (1992) *Nature*, **356**, 83-85. (http://www.accelrys.com/insight/Profiles-3D_page.html)

Поступила 21.03.03

**COMPUTER MODELLING OF 3D STRUCTURE OF CYTOCHROMES P450:
PROBLEMS AND FUTURE TRENDS**

A.S.Ivanov, V.S.Skvortsov, A.A.Sechenykh, A.V.Dubanov, A.V.Lisitsa

Orekhovich Institute of Biomedical Chemistry of Russian Academy of Medical Sciences
Pogodinskaya str. 10, Moscow, 119121, Russia; fax: (095) 245-0857; e-mail: ivanov@ibmh.msk.su;

The review is devoted to the description of basic approaches, problems and prospects of computer modelling of 3D structures of cytochromes P450 (P450s). The following questions have been considered in the review: (1) the role of computer molecular modelling in P450s research; (2) basic stages of homology modelling of protein 3D structures; (3) purposes and tasks of molecular modelling; (4) problems of P450s homology modelling; (5) criteria for successful P450s homology modelling; (6) methods of P450s models refinement, verification of correctness and reliability.

Key words: cytochrome P450, computer molecular modelling, homology modeling.