

# ПРЕДСКАЗАНИЕ ТОКСИЧНОСТИ НА ОСНОВАНИИ АНАЛИЗА ВИРТУАЛЬНЫХ ВЗАИМОДЕЙСТВИЙ НИЗКОМОЛЕКУЛЯРНЫХ СОЕДИНЕНИЙ С БЕЛКАМИ ОРГАНИЗМА

В.С. Скворцов, А.В. Веселовский, А.В. Захаров, Е.В. Бовина, О.А. Раевский.

Данный проект нацелен в первую очередь на создание компьютерного подхода, позволяющего оценивать токсичность низкомолекулярных органических соединений, механизм действия которых основан на взаимодействии токсикантов с макромолекулой-мишенью. Кроме того, отдалённой целью можно считать создание комплементарной экспериментальной системы тестирования. Подобная система могла бы существенно упростить отработку потенциально токсичных вариантов при разработке фармакологических препаратов.

Подавляющее большинство биологических эффектов низкомолекулярных органических соединений проявляется в результате избирательного связывания с макромолекулами одного или разных типов. Кроме того, обычно этот процесс является видоспецифичным, тканеспецифичным или органоспецифичным, что пока вообще трудно поддаётся оценке, разве что придётся создать виртуального грызуна. Но даже для идеальных данных по LD50 для конкретного биологического вида, получение прямых корреляций между структурой и токсичностью низкомолекулярных соединений может привести к удовлетворительным результатам лишь с родственными соединениями и/или в тех случаях, когда эффект обусловлен не избирательным связыванием, а физико-химическими параметрами лиганда. В тоже время биологические эффекты любого соединения, в том числе и токсичность, могут быть описаны через совокупность его взаимодействий с макромолекулами клеток. Априори неясно будет ли эффект кумулятивным или будет срабатывать самое слабое место, вероятнее всего спектр вариантов достаточно широк. Однако, есть надежда что определённые закономерности могут быть выявлены.

В ходе работы требуется решить целый набор важных задач.

- 1) Подобрать и отработать скоростные алгоритмы и технологии докинга, позволяющие осуществить вычисления в приемлемое время.
- 2) Разработать максимально унифицированные методы оценки энергии лиганд-белкового взаимодействия и их «настройки» для индивидуальных белков с помощью наборов известных ингибиторов, субстратов и т.п.
- 3) Создать выборки из баз данных, включающие пространственные структуры белков, связывание с которыми может обуславливать токсических эффект; и наборы токсикантов, действие которых определяется взаимодействием с макромолекулами-мишенями.
- 4) Подобрать наиболее адекватный способ описания «профиля» лиганд-белковых взаимодействий и «настройки» корреляционных уравнений для LD50 с помощью данных об известных и хорошо изученных токсикантах.
- 5) Увязать все элементы последовательных вычислений в систему позволяющую проводить симуляцию и анализ фактически в автоматическом режиме, с минимальным вмешательством оператора.
- 6) Рассчитать полный набор данных для 200 оригинальных белков и 5000-7000 низкомолекулярных соединений с известными экспериментальными оценками токсичности.
- 7) Создать и проверить математических моделей предсказания токсичности.

Часть из них уже решена в той или иной степени. На настоящий момент уже отобрано более сотни белков (а конечное их количество должно быть не менее чем в два раза больше) - ключевых ферментов и рецепторов с известной трехмерной структурой и положением в ней активных и регуляторных центров. Для части из них подобраны репрезентативные выборки известных лигандов с экспериментальными оценками аффинности. Последнее очень важно с точки зрения определения вариантов, когда реакция приводящая к токсическому эффекту носит триггерный характер, т.е. запускается при связывании достаточно долгом для запуска необходимой последовательности событий, но раз запущенная уже не зависит от прочности связывания с мишенью.

Как это не парадоксально, применимость подхода может быть продемонстрирована даже на небольших выборках. Например, на семействе ядерных рецепторов. В работе было задействовано 12 лиганд-связывающих доменов ядерных рецепторов и 109 токсикантов (25 имели стероидную природу, про остальные было известно, что они связываются с эстрогеновым рецептором). Использование данных по виртуальному связыванию позволило построить модель с использованием простой линейной регрессии с квадратом коэффициента корреляции 0.71 (0.63 в процедуре скользящего контроля). А применение методов самосогласованной регрессии и нейронных сетей непрямого распространения даёт ещё лучший результат. В тоже время сложность предложенного подхода в том, что, несмотря на отдельные успехи в некоторых группах, реальный результат может быть получен только на всей сово-

купности данных. Причём чем размер матрицы для анализа будет больше, тем вероятнее будет получение качественных предсказаний.

С этим связана и самая большая проблема работы. В ходе работы для моделирования комплекса низкомолекулярное соединение – молекула мишень на настоящий момент используется следующая схема вычислений:

- 1) Докинг лиганда программой Dock 6.0. И отбор не более 3 вариантов по встроенным оценочным функциям этой программы.
- 2) Проведение цикла симуляции молекулярной динамики комплекса в водном окружении – минимизация энергии системы; разогрев системы; выравнивание давления; уравнивание системы; собственно симуляция молекулярной динамики (Amber 9.0).
- 3) На основе траекторий рассчитанных молекулярной динамикой оценить изменения свободной энергии в комплексе методом MMPBSA (Amber 9.0) и по этим величинам отобрать окончательное решение. Эти же величины могут использоваться и как мера аффинности в конечной матрице, если нет возможности предсказать величину аффинности для данного конкретного рецептора исходя из структуры полученного комплекса.

Такая процедура расчета требует существенных затрат компьютерного времени. Например, для рецептора прогестерона на вычислительном кластере ГУ НИИ БМХ РАМН на базе процессоров OPTERON около 4 часов на один процессор. А таких вычислений в работе требуется провести около миллиона. В настоящее время мы ищем пути снижения объёма вычислений.

Однако, если полученные модели будут давать качественные предсказания, даже такие большие затраты на вычисления будут заведомо дешевле набора экспериментальных проверок и быстрее. Хотя, последнее слово всё равно должно оставаться за экспериментом.

Ранее в Институте физиологически активных веществ РАН и в Институте биомедицинской химии РАМН был разработан единого сервисного комплекса с общим интерфейсом, предназначенный для предсказания физико-химических свойств органических соединений, прогнозирования их токсичности и других биологических активностей. Создан набор утилит, позволяющих запустить вычисления в рамках этой системы для любого произвольного химического соединения. В настоящее время эта возможность может быть использована для пополнения данными результирующей матрицы. Однако в перспективе планируется, что по данному пути будут запускаться вычисления для предсказания токсичности целевых соединений. К сожалению пока вопрос о представлении возможности обращаться к запуску подобных вычислений любому желающему не стоит, из-за огромных вычислительных затрат. Однако по мере развития вычислительных систем, а также по мере анализа пространства виртуальных взаимодействий ситуация может и измениться.

Данный проект выполняется в рамках гранта МНТЦ 3777.

Собственно вычислительный комплекс является развитием разработок в рамках Государственного контракта ФЦНТП "Исследования и разработки по приоритетным направлениям развития науки и техники на 2002-2006 годы" (02.434.11.1014), МНТЦ 888 и FPG (LSHB-СТ-2007-037590).