

ИММУНОСЕТЕВОЕ МОДЕЛИРОВАНИЕ ЗАВИСИМОСТИ «СТРУКТУРА – СВОЙСТВО» ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ

Г.А.Самигулина, С.В.Чебейко

Разработан иммуносетевой подход к моделированию зависимостей «структура-свойство» лекарственных препаратов. Предложенная интеллектуальная технология позволяет уменьшить погрешности энергетических оценок и повысить достоверность прогноза зависимости «структура-свойство» химических соединений.

Актуальнейшим вопросом биоинформатики в настоящее время является целенаправленный поиск новых веществ и материалов с заранее заданными свойствами, в том числе разработка новых лекарственных средств на основе компьютерного молекулярного дизайна. Применение последних достижений вычислительной техники, искусственного интеллекта, инновационных суперкомпьютерных технологий открывает широкие возможности для успешного решения данной проблемы. Вполне реальным становится создание систем по типу САПР (напомним, что в технике это означает «системы автоматического проектирования») для проектирования и конструирования молекулярных устройств с заданными функциями [1].

Разработка новейших интеллектуальных технологий позволит существенно сократить временные и финансовые расходы при производстве новых лекарств. Даже простая предварительная информация о возможных способах посадки предполагаемой молекулы – лекарства на соответствующий рецепторный участок позволяет резко сократить расходы на создание новых фармакологических препаратов.

Время на создание нового лекарственного препарата составляет 12-15 лет, а общие затраты денег более 800 млн. долларов [2]. Из 100 тысяч потенциальных лекарственных соединений только 1000 могут стать основой для нового препарата, из которых только 3 смогут выйти на фармакологический рынок. С помощью компьютерного молекулярного дизайна, интеллектуальных технологий возможно создание принципиально новых компьютерных алгоритмов, программ поиска и отбора активных веществ целевого назначения на основе концепции взаимосвязи молекулярной структуры и биологической активности химических соединений.

Процесс создания нового лекарства включает в себя поиск мишени действия нового препарата; поиск биологически активного вещества, обладающего нужным фармакологическим действием; изучение этого соединения и т.д.

Количественное описание молекулярной структуры химических соединений в компьютерном молекулярном дизайне осуществляется с помощью дескрипторов [3]. Дескриптор

- это математический параметр, который характеризует структуру органического соединения, отмечая наиболее важные черты этой структуры. Существует проблема создания дескрипторов наиболее полно характеризующих рассматриваемое соединение и позволяющих в удобной форме использовать их в вычислительном процессе.

В современном компьютерном дизайне выделяется три основных этапа исследований:

- формирование обучающей выборки соединений с заданным свойством (активностью);
- описание молекулярной структуры исследуемых соединений (дескрипторов);
- установление взаимосвязи «структура – свойство (биологическая активность)» с последующим созданием устойчивых прогностических математических моделей.

Среди методов прогнозирования зависимости «структура – свойство» следует отметить рост исследований по искусственным нейронным сетям [4]. Наиболее популярна многослойная нейронная сеть прямого распространения, обучающаяся по методу обратного распространения ошибки.

Моделирование биологической активности органических соединений также возможно с помощью нового биологического направления искусственного интеллекта – искусственных иммунных систем [5] (ИИС). Под ИИС понимаются информационные методологии, использующие понятия теоретической иммунологии для решения различных прикладных задач.

Интерес к таким задачам объясняется их сложной структурой, неполнотой и зашумленностью данных, невозможностью успешного использования классических математических подходов.

Постановка задачи формулируется следующим образом: необходимо разработать эффективную интеллектуальную информационную технологию для компьютерного молекулярного дизайна (моделирования и предсказания свойств новых лекарственных препаратов с заданными свойствами) на основе биологического подхода искусственных иммунных систем [6, 7]. Специфика состоит в том, что при построении математических моделей необходимо одновременно использовать как распознающие, так и оптимизационные способности иммунной сети [8]. Задачу прогнозирования можно рассматривать как задачу оптимизации: необходимо построить оптимальную модель (структуру), которая бы наиболее точно описывала требуемую биологическую активность (свойство) органического соединения.

Реальные данные почти всегда содержат нежелательные составляющие, которые называются шумом. Существуют различные виды неопределенностей данных, например случайные погрешности (неточности при сборе и формировании дескрипторов); систематические (причинные) погрешности; ошибки при моделировании из-за использования упрощенных моделей и т.д. Шум и избыточность данных проявляются в корреляции между пе-

ременными. За исключением простых случаев, искажения в данных не могут быть устранены полностью. Основной целью при построении систем на основе иммунносетевого моделирования является уменьшение ошибки обобщения, поскольку малая ошибка обучения гарантирует адекватность модели лишь в заранее выбранных точках.

Возможна ситуация когда почти идентичные по структуре соединения имеют диаметрально противоположные свойства и следовательно биологическую активность, что совершенно недопустимо при конструировании лекарственных соединений. Актуально применение свойств реальных белков для решения этой проблемы.

Таким образом, решение задачи минимизации ошибки обобщения позволяет повысить прогностическую способность модели и является наиболее трудной при построении данных систем.

Литература

1. Шайтан К.В. Молекулярная динамика пептидов. <http://www.bioeng.ru/doc/stat/moldyn.htm>.
2. Кубиньи Г. В поисках новых соединений-лидеров для создания лекарств //Российский химический журнал. 2006, № 2. С. 5-17.
3. Раевский О.А. Дескрипторы водородной связи в компьютерном молекулярном дизайне //Российский химический журнал. 2006, № 2. С. 97-108.
4. Гальберштам Н.М., Баскин И.И., Палюлин В.А., Зефиоров Н.С. Нейронные сети как метод поиска зависимостей структура – свойство органических соединений //Успехи химии. 2003, № 72(7). С. 706-727.
5. Тараканов А.О. Математические модели биомолекулярной обработки информации: формальный пептид вместо формального нейрона // Проблемы информатизации. 1998. С.65-70.
6. Самигулина Г.А., Чебейко С.В. Прогнозирование зависимости структура-свойство органических соединений на основе иммунносетевого моделирования // Химический журнал Казахстана. –Алматы. 2010, № 3. С. 164-172.
7. Самигулина Г.А., Чебейко С.В. Технология иммунносетевого моделирования для компьютерного молекулярного дизайна лекарственных препаратов // Вестник Харьковского Университета. Тематический выпуск: информатика и моделирование. –Харьков. 2011, № 17. С.142-148. http://www.nbu.gov.ua/Portal/natural/vcpi/iim/2011_17/20pdf
8. Бидюк П.И., Литвиненко В.И., Фефелов А.А., Баклан И.В. Гибридная иммунная сеть для решения задач структурной идентификации // Искусственный интеллект. -2004, - № 3. –С.89-99.