

ВЫЧИСЛИТЕЛЬНЫЙ РЕАКТОР ДЛЯ ПЛАНИРОВАНИЯ ХИМИЧЕСКОГО СИНТЕЗА: ГЕНЕРАЦИЯ ВОЗМОЖНЫХ РЕАКЦИЙ С УЧАСТИЕМ ЗАДАННОГО ХИМИЧЕСКОГО СОЕДИНЕНИЯ

Фатыхова А.А.^a, Нугманов Р.И.^a, Мухаметгалеев Р.Н.^a, Маджидов Т.И.^a, Варнек А.^b

^a Химический институт им. А.М.Бутлерова, Казанский Федеральный Университет, Россия

^b Страсбургский университет, Страсбург, Франция

e-mail: adelik21979@gmail.com

Ключевой задачей в планировании химического синтеза является генерация новых соединений из набора имеющихся молекул. В рамках данной работы был разработан алгоритм, который позволяет получать новые химические структуры из исходных молекул, используя в качестве библиотеки трансформаций извлеченные из базы данных химические реакции. Основным преимуществом разработанного алгоритма является сохранение информации о механизме трансформации в новых реакциях. Эта информация представлена в виде атом-атомного отображения.

Была создана база данных возможных реагентов с использованием нашей собственной разработанной системой CGR-DB. Данные о реагентах были извлечены из баз данных коммерчески доступных химических строительных блоков. Химические структуры были проиндексированы на наличие в них определенных функциональных групп. Правила трансформаций были классифицированы на одно-, двух-, мульти-компонентные реакции. Функциональные группы, отвечающие за механизм реакции, были идентифицированы на основе извлеченных правил трансформаций.

На этапе генерации продуктов алгоритм определяет функциональные группы в исходной молекуле и извлекает из базы данных только те трансформации, которые могут быть применены к данной молекуле. Если выбранная трансформация содержит два или более реагентов, второй реагент выбирается также на основе анализа функциональных групп. Таким образом, алгоритм не пытается применить ненужные трансформации к неподходящим молекулам и с высокой скоростью генерирует новые соединения и реакции. Тестирование алгоритма показало высокое качество и скорость.

Работа финансируется в рамках проекта, поддержанного Министерством образования молодежи и спорта Чешской республики, соглашение MSMT-5727/2018-2, а также Министерством науки и высшего образования РФ, соглашение 14.587.21.0049 (уникальный идентификатор проекта RFMEFI58718X0049).