

УДК 541.62/.636

**ХЕМОИНФОРМАТИКА В АНАЛИЗЕ И ПРЕДСКАЗАНИИ ХАРАКТЕРИСТИК
ХИМИЧЕСКИХ РЕАКЦИЙ****Т.И. Маджидов¹, Р.И. Нугманов¹, Т.Р. Гимадиев¹, В.А. Афонина¹, А.А. Фатыхова¹, Р.Н. Мухаметгалеев¹,
А.Б. Рахимбекова¹, С.И. Мусаева¹, Ф. Боначера², И.И. Баскин³, И.С. Антипин¹, А.А. Варнек²**

¹НИЛ «Хемоинформатика и молекулярное моделирование», Химический институт им. А.М. Бутлерова, Казанский федеральный университет, Казань, Россия; ²Лаборатория хемоинформатики, Страсбургский университет, Страсбург, Франция; ³Химический факультет, МГУ им. М.В. Ломоносова, Москва, Россия

Синтез химических соединений является одним из основных элементов разработки лекарственных препаратов. Соединения с требуемой химической структурой необходимо синтезировать для широкого- или мелкомасштабного биологического скрининга, в ходе процедуры оптимизации фармакологических свойств, при наработке вещества для доклинических испытаний. Отдельную проблему представляет разработка лабораторного или промышленного регламента синтеза. Несмотря на то, что технологии анализа данных и вычислительного моделирования широко используются в разработке лекарства, они практически не используются при разработке синтеза химического соединения. В рамках данного доклада будут представлены разработки, направленные на облегчение принятия решений в органическом синтезе. В основе всех предложенных нами подходов находится концепция Конденсированного графа реакции (КГР). Данная концепция позволяет закодировать в одном графе химические изменения, протекающие в ходе химической реакции. Конденсированный граф реакции позволяет рассчитать дескрипторы химической реакции, что предоставляет возможности последующего моделирования и визуализации, кроме того, в нем закодирована информация о реакционном центре. В докладе будут рассмотрены: 1. Поиск реакций по сходству реакционных центров¹ с использованием концепции КГР, эффективные алгоритмы структурного поиска в реакционных базах. 2. Предсказание характеристик химических реакций, таких как константы скорости и равновесия химических процессов². Показано, что в ряде случаев точность превышает точность предсказания методами квантовой химии. 3. Предсказание оптимальных условий проведения химических реакций с использованием методов машинного обучения. 4. Визуализация химического пространства реакций с помощью методов понижения размерности и анализ изменения интереса к реакциям различного типа во времени. *Исследования проведены при поддержке гранта Российского научного фонда N14-43-00024.*

Ключевые слова: химические реакции, хемоинформатика, машинное обучение, синтез.

Литература

1. Lin A.I. et al. Automatized Assessment of Protective Group Reactivity: A Step Toward Big Reaction Data Analysis // J. Chem. Inf. Model. 2016. Vol. 56, № 11. P. 2140–2148.
2. Gimadiev T.R. et al. Assessment of tautomer distribution using the condensed reaction graph approach // J. Comput. Aided. Mol. Des. 2018. Vol. 32, № 3. P. 401–414.