

АНАЛИЗ ФРАГМЕНТНОГО СОСТАВА ГУМИНОВЫХ КИСЛОТ ТОРФОВ И ПОЧВ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНОЙ СИСТЕМЫ «ИК-ЭКСПЕРТ» И КЛАСТЕРИЗАЦИИ СТРУКТУР ПОИСКОВОГО ОТВЕТА

Т.Ф. Богданова, В.Н. Пиоттукх-Пелецкий, В.Д. Тихова, В.П. Фадеева

Федеральное государственное бюджетное учреждение науки
Новосибирский институт органической химии им. Н.Н. Ворожцова СО РАН
630090, г. Новосибирск, просп. Лаврентьева, 9. bogd@nioch.nsc.ru

Установление строения гуминовых кислот (ГК), представляющих собой смеси полимеров нерегулярного строения, до сих пор является нерешенной задачей, полная структуры исследуемого объекта практически не определена, все предлагаемые структурные формулы ГК имеют оценочный характер. Возможно лишь достаточно обоснованное исследование фрагментного состава ГК. Поэтому представляет интерес поиск новых подходов к определению структурных фрагментов ГК.

В настоящей работе изучение химической структуры ГК, содержащихся в 28 образцах сибирских торфов и почв, проводится по спектрам инфракрасного поглощения (ИК) исследуемых образцов с использованием компьютерной системы ИК-ЭКСПЕРТ и процедуры кластеризации структур поискового ответа (ПО).

Разработанная в НТЦ ХИ система «ИК-ЭКСПЕРТ» использует базу данных, содержащую около 32 тыс. записей о структурах органических соединений и соответствующих ИК спектрах. Информация о строении каждого соединения представлена как топологическим описанием (поатомный код химического графа в виде списка связей с перечислением координат вершин), так и полным набором неизоморфных связанных структурных фрагментов заданного размера (в данной работе использовались фрагменты размером от 2 до 10 остовных атомов). Разбиение структур всех соединений БД на фрагменты производится один раз в процессе формирования БД.

Процедура определения наиболее вероятных структурных фрагментов гуминовых кислот проводится в два этапа. Сначала спектр исследуемого образца сравнивается с каждым спектром из БД и в ПО отбираются соединения, ИК спектры которых имеют коэффициент спектрального подобия со спектром изучаемого соединения выше заданного значения. Поскольку ИК спектры образцов ГК торфов не содержат ярко выраженных характеристических полос, сравнение спектральных кривых проводится по евклидовой метрике.

На втором этапе из структур ПО формируются группы (кластеры) структурно-подобных соединений. Используемое в системе «ИК-ЭКСПЕРТ» описание структурных формул в виде полного набора фрагментов помогает легко определять расстояние между молекулярными графами и проводить быструю кластеризацию структур ПО [1]. Анализ структурных формул соединений, составляющих отдельный кластер, позволяет распознать крупный связный фрагмент исследуемого соединения, наиболее характерный для данного кластера. Обсуждается соответствие выявленных структурных фрагментов и гипотез о строении гуминовых кислот, выдвинутых на основе других методов анализа.