

УДК 54:51-7

РАСЧЕТ И СВОЙСТВА ФРАКТАЛЬНЫХ ДЕСКРИПТОРОВ В РЯДУ АЛКАНОВ $C_2 \dots C_9$

В.Ю. Григорьев¹, Л.Д. Григорьева²

(¹Институт физиологически активных веществ РАН; ²Факультет фундаментальной физико-химической инженерии МГУ имени М.В. Ломоносова)

Для расчета молекулярных дескрипторов в ряду из 74 алканов использована дискретная дифференциальная функция распределения (гистограмма) межатомных расстояний. На ее основе вычислены фрактальные размерности (дескрипторы) четырех типов. Выявлена связь фрактальных дескрипторов между собой, а также с числом атомов и точечной группой симметрии молекул.

Ключевые слова: КССА, КССС, фрактальная размерность, фрактальный дескриптор.

Междисциплинарные исследования играют важную роль при получении нового знания об окружающем нас мире. Одним из инструментов для проведения такого рода исследований является интегральный подход «количественная связь – структура – активность/структура – свойство» (QSAR/QSPR) [1]. При этом для описания структуры химических соединений используются численные характеристики (дескрипторы). В настоящее время их число достигает нескольких тысяч [2], однако поиск новых структурных параметров продолжается.

Новым направлением в этой области является использование идей фрактальной геометрии, включая понятие о фрактальной размерности, которые широко проникли в различные области знания [3]. Для характеристики поверхности больших молекул (биополимеров) и оценки ее «шероховатости» (roughness) в публикациях [4, 5] предложено использовать величину фрактальной размерности D . Теоретически эта величина должна находиться в пределах от 2 до 3. Для лизоцима, рибонуклеазы А и супероксиддисмутазы были рассчитаны средние значения $D = 2,4$. В дальнейшем этот подход был модифицирован другими исследователями путем расширения числа рассматриваемых молекул и детализации их поверхности [6–8]. При этом установлено, что активные центры биополимеров и остальные области молекул имеют разную «шероховатость» (разные величины D).

Представляется, что методы оценки фрактальной размерности поверхности больших молекул выглядят неподходящими для обычных малых молекул из-за того, что размеры пробных и исследуемых молекул в этом случае сопоставимы между собой. В литературе предложены другие подходы для оценки фрактальных параметров. Так, в работе [9] использован метод размножения молекул

для получения молекулярного ансамбля, обладающего фрактальными свойствами. На основе рассчитанных фрактальных параметров получены количественные модели «структура – противоопухолевая активность» в ряду макроциклических пиридинофанов и их ациклических аналогов. Авторы публикации [10] предложили другой подход для расчета фрактальных дескрипторов. Он основан на использовании дифференциальных функций распределения (гистограмм) межатомных расстояний и расчете соответствующих фрактальных размерностей.

Цель настоящей работы – дальнейшее развитие предложенного подхода, расчет новых молекулярных фрактальных дескрипторов и исследование их свойств в ряду алканов $C_2 \dots C_9$.

Расчет фрактальной размерности

Дескрипторы должны быть инвариантными по отношению к молекулярному сдвигу и вращению, а также к обозначению и нумерации молекулярных атомов [11]. Дифференциальные функции распределения межатомных расстояний соответствуют этому требованию, поэтому они могут использоваться для создания новых молекулярных дескрипторов. Для их расчета необходимо вычислить межатомные расстояния (r) и определить число межатомных расстояний ($f(r)$), попадающих в определенный интервал. В качестве примера расчета рассмотрим молекулу пропана. На основе декартовых координат атомов, полученных из трехмерной структуры, вычисляли расстояния между атомами. На рис. 1 представлена гистограмма межатомных расстояний, рассчитанная с шагом 0,01 и 0,01 Å по осям абсцисс и ординат соответственно, т.е. с размером квадратного пикселя 0,01.

Для дальнейших расчетов гистограмму оцифровывали, т.е. представляли в виде квадратной

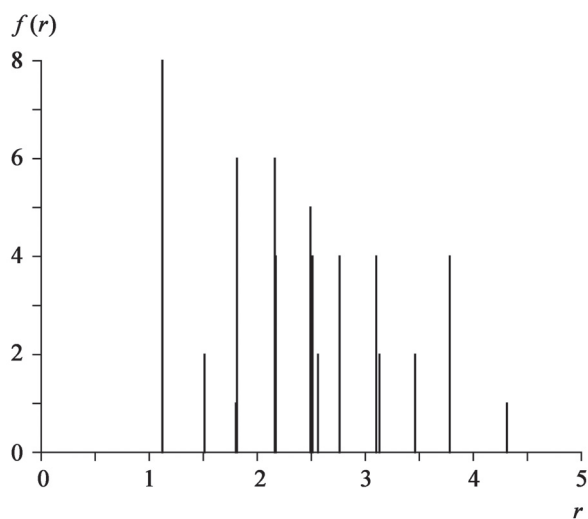


Рис. 1. Дифференциальная функция распределения (гистограмма) межатомных расстояний пропана

бинарной фрактальной матрицы, где окрашенным и неокрашенным пикселям соответствовали величины 1 и 0. Размер матрицы определяли на основе максимума из двух величин: $X = (r_{\text{макс}} - r_{\text{мин}})/0,01$ и $Y = (f(r)_{\text{макс}} - f(r)_{\text{мин}})/0,01$, где X – максимальный размер гистограммы по оси r ; $r_{\text{макс}}$ – максимальное межатомное расстояние; $r_{\text{мин}}$ – минимальное межатомное расстояние; Y – максимальный размер гистограммы по оси $f(r)$; $f(r)_{\text{макс}}$ – максимальная величина $f(r)$; $f(r)_{\text{мин}}$ – минимальная величина $f(r)$.

Для пропана

$$X = (4,31 - 1,12)/0,01 = 319,$$

$$Y = (8 - 0)/0,01 = 800,$$

т.е. размер фрактальной матрицы составляет $800 \times 800 = 640\,000$ пикселей.

Фрактальную размерность определяли «клеточным» методом путем покрытия фрактальной матрицы квадратами переменной величины и подсчетом числа квадратов, которые пересекали окрашенные пиксели [12]. Локальную фрактальную размерность рассчитывали по уравнению (1):

$$\log(M) = \log C - D \log(L), \quad (1)$$

где M – минимальное число квадратов со стороной L , необходимых для покрытия фрактала; D – «клеточная» фрактальная размерность (со знаком минус); C – константа.

Для определения величины D проводили измерения при разных величинах L , меняя этот параметр от 1 до 10 с шагом 1, при этом соответствующая площадь квадратов варьировала от 1 до 100 пикселей. В результате для молекулы пропана получены данные, которые хорошо описываются прямой линией (рис. 2). Рассчитанная величина D составила 1,08.

Геометрические параметры и точечные группы симметрии молекул рассчитывали с помощью компьютерной программы HyperChem [13]. Для этого использовали полуэмпирический квантово-химический метод AM1 с полной оптимизацией структуры молекул. Для каждого исследуемого соединения на основе оптимизированных 3D-структур рассчитали по 3224 дескриптора с помощью программы DRAGON [2]. Для проведения регрессионного анализа применяли модифицированную компьютерную программу SVD [14]. В качестве статистических характеристик уравнений использовали следующие параметры: n – число точек, R^2 – квадрат коэффициента линейной корреляции, s – стандартное отклонение, Q^2 – квадрат коэффициента линейной корреляции в условиях скользящего контроля с выбором по одному. При необходимости проводили автошкалирование дескрипторов путем вычитания средних значений и деления на стандартное отклонение.

Результаты и их обсуждение

В таблице представлены результаты расчета фрактальных дескрипторов.

Исследуемый ряд состоит из 74 соединений. При этом минимальное число атомов в молекуле равно 8, а максимальное – 29. Анализируемые молекулы относятся к одной из восьми точечных групп симметрии (C_1 , C_p , C_s , C_{2h} , C_{2v} , C_{3v} , D_{3d} , T_d). Величина общего фрактального дескриптора D изменяется в достаточно больших пределах: от 1,00 (этан) до 1,48 (2,4-диметил-3-этилпентан). Теоретически для рассматриваемых геометрических объектов фрактальная размерность может меняться от 1 (гладкая кривая) до 2 (плоскость).

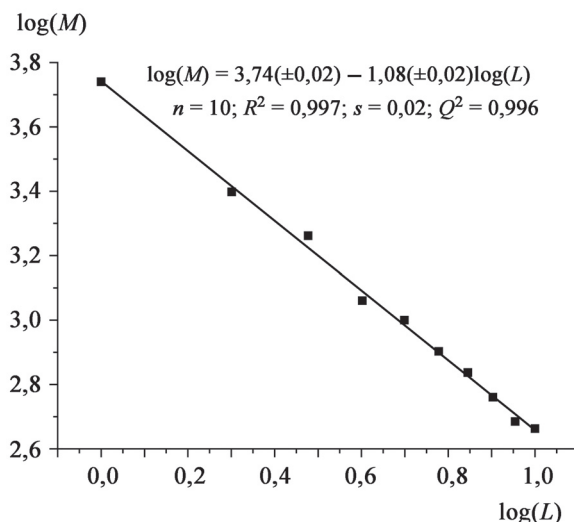


Рис. 2. Зависимость числа квадратов (M), покрывающих фрактальную матрицу пропана, от размера квадрата (L)

Фрактальные дескрипторы 74 алканов

| п/п | Соединение | CAS RN | <i>N</i> | <i>S</i> | <i>D</i> ±Δ | <i>D_{val}</i> ±Δ | <i>D_{vDw}</i> ±Δ | <i>D_{unb}</i> ±Δ |
|-----|----------------------|-----------|----------|-----------------------|-------------|---------------------------|---------------------------|---------------------------|
| 1 | Этан | 74-84-0 | 8 | <i>D_{3d}</i> | 1,00±0,01 | 1,00±0,01 | 1,00±0,01 | 1,00±0,01 |
| 2 | Пропан | 74-98-6 | 11 | <i>C_{2v}</i> | 1,08±0,02 | 1,00±0,01 | 1,08±0,03 | 1,17±0,03 |
| 3 | Бутан | 106-97-8 | 14 | <i>C_{2h}</i> | 1,11±0,02 | 1,00±0,01 | 1,13±0,03 | 1,15±0,02 |
| 4 | 2-Метилпропан | 75-28-5 | 14 | <i>C_{3v}</i> | 1,09±0,02 | 1,03±0,01 | 1,13±0,04 | 1,13±0,03 |
| 5 | Пентан | 109-66-0 | 17 | <i>C_{2v}</i> | 1,11±0,02 | 1,00±0,01 | 1,11±0,02 | 1,16±0,02 |
| 6 | 2-Метилбутан | 78-78-4 | 17 | <i>C₁</i> | 1,27±0,01 | 1,04±0,02 | 1,24±0,01 | 1,37±0,03 |
| 7 | 2,2-Диметилпропан | 463-82-1 | 17 | <i>T_d</i> | 1,02±0,01 | 1,00±0,01 | 1,00±0,01 | 1,00±0,01 |
| 8 | Гексан | 110-54-3 | 20 | <i>C_{2h}</i> | 1,11±0,01 | 1,00±0,01 | 1,15±0,02 | 1,11±0,01 |
| 9 | 2-Метилпентан | 107-83-5 | 20 | <i>C₁</i> | 1,31±0,02 | 1,03±0,02 | 1,27±0,03 | 1,34±0,02 |
| 10 | 3-Метилпентан | 96-14-0 | 20 | <i>C_s</i> | 1,23±0,02 | 1,04±0,03 | 1,22±0,03 | 1,26±0,03 |
| 11 | 2,2-Диметилбутан | 75-83-2 | 20 | <i>C_s</i> | 1,27±0,02 | 1,03±0,01 | 1,26±0,03 | 1,31±0,03 |
| 12 | 2,3-Диметилбутан | 79-29-8 | 20 | <i>C_{2h}</i> | 1,20±0,02 | 1,05±0,02 | 1,19±0,03 | 1,16±0,02 |
| 13 | Гептан | 142-82-5 | 23 | <i>C_{2v}</i> | 1,13±0,01 | 1,00±0,01 | 1,13±0,02 | 1,18±0,02 |
| 14 | 2-Метилгексан | 591-76-4 | 23 | <i>C₁</i> | 1,33±0,02 | 1,03±0,01 | 1,28±0,03 | 1,39±0,02 |
| 15 | 3-Метилгексан | 589-34-4 | 23 | <i>C₁</i> | 1,34±0,02 | 1,04±0,03 | 1,28±0,03 | 1,39±0,02 |
| 16 | 2,2-Диметилпентан | 590-35-2 | 23 | <i>C_s</i> | 1,28±0,01 | 1,04±0,03 | 1,29±0,03 | 1,27±0,01 |
| 17 | 2,3-Диметилпентан | 565-59-3 | 23 | <i>C₁</i> | 1,40±0,01 | 1,07±0,03 | 1,36±0,01 | 1,45±0,02 |
| 18 | 2,4-Диметилпентан | 108-08-7 | 23 | <i>C_s</i> | 1,27±0,02 | 1,05±0,03 | 1,25±0,03 | 1,29±0,02 |
| 19 | 3,3-Диметилпентан | 562-49-2 | 23 | <i>C_{2v}</i> | 1,21±0,02 | 1,06±0,02 | 1,21±0,02 | 1,19±0,02 |
| 20 | 3-Этилпентан | 617-78-7 | 23 | <i>C₁</i> | 1,36±0,02 | 1,05±0,03 | 1,28±0,04 | 1,44±0,02 |
| 21 | 2,2,3-Триметилбутан | 464-06-2 | 23 | <i>C_s</i> | 1,31±0,02 | 1,06±0,02 | 1,22±0,01 | 1,38±0,04 |
| 22 | Октан | 111-65-9 | 26 | <i>C_{2h}</i> | 1,14±0,02 | 1,00±0,01 | 1,11±0,02 | 1,18±0,02 |
| 23 | 2-Метилгептан | 592-27-8 | 26 | <i>C₁</i> | 1,35±0,02 | 1,02±0,01 | 1,29±0,03 | 1,43±0,02 |
| 24 | 3-Метилгептан | 589-81-1 | 26 | <i>C₁</i> | 1,33±0,02 | 1,03±0,02 | 1,28±0,03 | 1,38±0,02 |
| 25 | 4-Метилгептан | 589-53-7 | 26 | <i>C_s</i> | 1,29±0,02 | 1,03±0,02 | 1,27±0,03 | 1,32±0,03 |
| 26 | 3-Этилгексан | 619-99-8 | 26 | <i>C₁</i> | 1,38±0,02 | 1,05±0,03 | 1,32±0,02 | 1,44±0,03 |
| 27 | 2,2-Диметилгексан | 590-73-8 | 26 | <i>C_s</i> | 1,27±0,01 | 1,04±0,03 | 1,29±0,01 | 1,29±0,01 |
| 28 | 2,3-Диметилгексан | 584-94-1 | 26 | <i>C₁</i> | 1,38±0,01 | 1,07±0,04 | 1,39±0,03 | 1,43±0,03 |
| 29 | 2,4-Диметилгексан | 589-43-5 | 26 | <i>C₁</i> | 1,35±0,02 | 1,06±0,03 | 1,32±0,02 | 1,40±0,02 |
| 30 | 2,5-Диметилгексан | 592-13-2 | 26 | <i>C_i</i> | 1,27±0,02 | 1,05±0,03 | 1,23±0,02 | 1,28±0,02 |
| 31 | 3,3-Диметилгексан | 563-16-6 | 26 | <i>C_s</i> | 1,26±0,01 | 1,06±0,02 | 1,30±0,01 | 1,24±0,02 |
| 32 | 3,4-Диметилгексан | 583-48-2 | 26 | <i>C_i</i> | 1,31±0,02 | 1,06±0,03 | 1,27±0,01 | 1,31±0,02 |
| 33 | 3-Этил-2-метилпентан | 609-26-7 | 26 | <i>C₁</i> | 1,43±0,02 | 1,06±0,03 | 1,33±0,01 | 1,50±0,03 |
| 34 | 3-Этил-3-метилпентан | 1067-08-9 | 26 | <i>C₁</i> | 1,41±0,02 | 1,07±0,02 | 1,38±0,02 | 1,47±0,02 |
| 35 | 2,2,3-Триметилпентан | 564-02-3 | 26 | <i>C₁</i> | 1,42±0,01 | 1,08±0,02 | 1,38±0,02 | 1,45±0,02 |
| 36 | 2,2,4-Триметилпентан | 540-84-1 | 26 | <i>C₁</i> | 1,39±0,01 | 1,07±0,02 | 1,33±0,02 | 1,45±0,03 |
| 37 | 2,3,3-Триметилпентан | 560-21-4 | 26 | <i>C₁</i> | 1,39±0,01 | 1,06±0,01 | 1,30±0,01 | 1,46±0,02 |
| 38 | 2,3,4-Триметилпентан | 565-75-3 | 26 | <i>C_s</i> | 1,36±0,02 | 1,08±0,02 | 1,33±0,02 | 1,38±0,02 |

Окончание таблицы

| п/п | Соединение | CAS RN | <i>N</i> | <i>S</i> | <i>D</i> ±Δ | <i>D_{val}</i> ±Δ | <i>D_{vDw}</i> ±Δ | <i>D_{unb}</i> ±Δ |
|-----|--------------------------|------------|----------|-----------------------|-------------|---------------------------|---------------------------|---------------------------|
| 39 | 2,2,3,3-Тетраметилбутан | 594-82-1 | 26 | <i>D_{3d}</i> | 1,14±0,03 | 1,02±0,01 | 1,10±0,02 | 1,12±0,02 |
| 40 | Нонан | 111-84-2 | 29 | <i>C_{2v}</i> | 1,14±0,01 | 1,00±0,01 | 1,06±0,01 | 1,19±0,02 |
| 41 | 2-Метилотан | 3221-61-2 | 29 | <i>C₁</i> | 1,37±0,01 | 1,02±0,01 | 1,27±0,02 | 1,44±0,02 |
| 42 | 3-Метилотан | 2216-33-3 | 29 | <i>C₁</i> | 1,36±0,02 | 1,03±0,02 | 1,28±0,02 | 1,41±0,02 |
| 43 | 4-Метилотан | 2216-34-4 | 29 | <i>C₁</i> | 1,35±0,01 | 1,03±0,02 | 1,29±0,03 | 1,40±0,02 |
| 44 | 3-Этилгептан | 15869-80-4 | 29 | <i>C₁</i> | 1,38±0,02 | 1,04±0,02 | 1,32±0,03 | 1,46±0,02 |
| 45 | 4-Этилгептан | 2216-32-2 | 29 | <i>C₁</i> | 1,39±0,02 | 1,04±0,02 | 1,29±0,02 | 1,44±0,03 |
| 46 | 2,2-Диметилгептан | 1071-26-7 | 29 | <i>C_s</i> | 1,30±0,01 | 1,04±0,03 | 1,34±0,02 | 1,31±0,02 |
| 47 | 2,3-Диметилгептан | 3074-71-3 | 29 | <i>C₁</i> | 1,37±0,02 | 1,06±0,03 | 1,39±0,03 | 1,41±0,03 |
| 48 | 2,4-Диметилгептан | 2213-23-2 | 29 | <i>C₁</i> | 1,38±0,01 | 1,06±0,03 | 1,31±0,02 | 1,45±0,03 |
| 49 | 2,5-Диметилгептан | 2216-30-0 | 29 | <i>C₁</i> | 1,38±0,01 | 1,05±0,03 | 1,31±0,01 | 1,44±0,02 |
| 50 | 2,6-Диметилгептан | 1072-05-5 | 29 | <i>C_s</i> | 1,29±0,02 | 1,04±0,02 | 1,30±0,01 | 1,31±0,03 |
| 51 | 3,3-Диметилгептан | 4032-86-4 | 29 | <i>C_s</i> | 1,30±0,01 | 1,06±0,02 | 1,30±0,01 | 1,31±0,02 |
| 52 | 3,4-Диметилгептан | 922-28-1 | 29 | <i>C₁</i> | 1,39±0,01 | 1,06±0,03 | 1,36±0,03 | 1,45±0,02 |
| 53 | 3,5-Диметилгептан | 926-82-9 | 29 | <i>C_s</i> | 1,31±0,02 | 1,07±0,03 | 1,27±0,02 | 1,34±0,03 |
| 54 | 4,4-Диметилгептан | 1068-19-5 | 29 | <i>C_{2v}</i> | 1,21±0,02 | 1,04±0,02 | 1,26±0,02 | 1,19±0,02 |
| 55 | 3-Этил-2-метилгексан | 16789-46-1 | 29 | <i>C₁</i> | 1,42±0,01 | 1,08±0,04 | 1,41±0,02 | 1,48±0,02 |
| 56 | 4-Этил-2-метилгексан | 3074-75-7 | 29 | <i>C₁</i> | 1,41±0,02 | 1,06±0,03 | 1,33±0,03 | 1,48±0,02 |
| 57 | 3-Этил-3-метилгексан | 3074-76-8 | 29 | <i>C₁</i> | 1,41±0,02 | 1,06±0,01 | 1,33±0,02 | 1,49±0,03 |
| 58 | 3-Этил-4-метилгексан | 3074-77-9 | 29 | <i>C₁</i> | 1,42±0,01 | 1,06±0,03 | 1,35±0,02 | 1,49±0,02 |
| 59 | 2,2,3-Триметилгексан | 16747-25-4 | 29 | <i>C₁</i> | 1,42±0,01 | 1,05±0,01 | 1,41±0,02 | 1,46±0,03 |
| 60 | 2,2,4-Триметилгексан | 16747-26-5 | 29 | <i>C₁</i> | 1,40±0,01 | 1,06±0,02 | 1,34±0,02 | 1,45±0,02 |
| 61 | 2,2,5-Триметилгексан | 3522-94-9 | 29 | <i>C₁</i> | 1,38±0,01 | 1,05±0,03 | 1,33±0,03 | 1,41±0,03 |
| 62 | 2,3,3-Триметилгексан | 16747-28-7 | 29 | <i>C₁</i> | 1,40±0,01 | 1,06±0,02 | 1,38±0,02 | 1,44±0,02 |
| 63 | 2,3,4-Триметилгексан | 921-47-1 | 29 | <i>C₁</i> | 1,42±0,01 | 1,10±0,03 | 1,41±0,02 | 1,47±0,03 |
| 64 | 2,3,5-Триметилгексан | 1069-53-0 | 29 | <i>C₁</i> | 1,39±0,01 | 1,08±0,04 | 1,32±0,02 | 1,42±0,03 |
| 65 | 2,4,4-Триметилгексан | 16747-30-1 | 29 | <i>C₁</i> | 1,40±0,01 | 1,06±0,02 | 1,35±0,02 | 1,46±0,02 |
| 66 | 3,3,4-Триметилгексан | 16747-31-2 | 29 | <i>C₁</i> | 1,43±0,01 | 1,06±0,02 | 1,38±0,02 | 1,47±0,02 |
| 67 | 3,3-Диметилпентан | 1067-20-5 | 29 | <i>C_s</i> | 1,40±0,02 | 1,13±0,04 | 1,35±0,02 | 1,43±0,03 |
| 68 | 2,2-Диметил-3-этилпентан | 16747-32-3 | 29 | <i>C₁</i> | 1,45±0,02 | 1,07±0,02 | 1,38±0,01 | 1,53±0,03 |
| 69 | 2,3-Диметил-3-этилпентан | 16747-33-4 | 29 | <i>C₁</i> | 1,45±0,01 | 1,07±0,02 | 1,40±0,02 | 1,50±0,03 |
| 70 | 2,4-Диметил-3-этилпентан | 1068-87-7 | 29 | <i>C₁</i> | 1,48±0,01 | 1,12±0,03 | 1,40±0,02 | 1,53±0,03 |
| 71 | 2,2,3,3-Тетраметилпентан | 7154-79-2 | 29 | <i>C_s</i> | 1,37±0,02 | 1,04±0,01 | 1,46±0,01 | 1,42±0,03 |
| 72 | 2,2,3,4-Тетраметилпентан | 1186-53-4 | 29 | <i>C₁</i> | 1,45±0,01 | 1,08±0,02 | 1,43±0,02 | 1,50±0,02 |
| 73 | 2,2,4,4-Тетраметилпентан | 1070-87-7 | 29 | <i>C_{2v}</i> | 1,30±0,03 | 1,08±0,03 | 1,30±0,02 | 1,28±0,04 |
| 74 | 2,3,3,4-Тетраметилпентан | 16747-38-9 | 29 | <i>C_s</i> | 1,38±0,02 | 1,07±0,02 | 1,40±0,02 | 1,38±0,04 |

Обозначения. CAS RN – регистрационный номер; *N* – число атомов в молекуле; *S* – точечная группа симметрии; *D* – общий фрактальный дескриптор, рассчитанный на основе всех межатомных расстояний; *D_{val}* – фрактальный дескриптор, рассчитанный на основе межатомных расстояний между валентно связанными атомами; *D_{vdw}* – фрактальный дескриптор, рассчитанный на основе межатомных расстояний между атомами, связанными Ван-дер-Ваальсовыми силами; *D_{unb}* – фрактальный дескриптор, рассчитанный на основе межатомных расстояний между несвязанными атомами; Δ – стандартная ошибка.

В целом, с ростом числа атомов величина общего фрактального дескриптора увеличивается (рис. 3). Этот рост обусловлен тем, что с увеличением числа атомов быстро возрастает число межатомных расстояний, как $N \cdot (N - 1) / 2$. Это приводит к увеличению числа окрашенных пикселей во фрактальной матрице и росту D . Однако при одном и том же числе атомов в молекуле наблюдаются довольно значительные изменения величин фрактального дескриптора. Так, в случае изомеров C_9H_{20} ($N = 29$) величина D меняется от 1,14 (нонан) до 1,48 (2,4-диметил-3-этилпентан). Объяснить наблюдаемое явление можно, обратившись к рис. 4.

Представленные на рис. 4 данные свидетельствуют о том, что между фрактальной размерностью и точечной группой симметрии (при одинаковом числе атомов) существует связь в виде

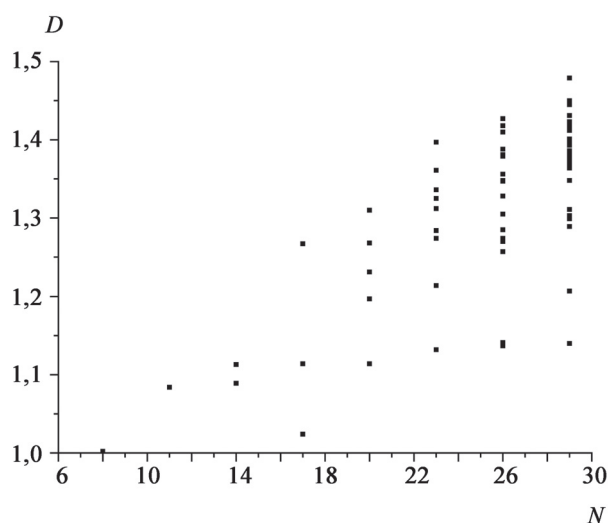


Рис. 3. Зависимость общего фрактального дескриптора (D) от числа атомов (N)

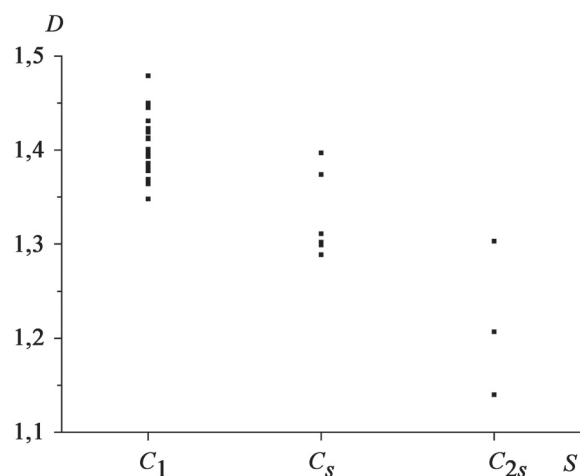


Рис. 4. Зависимость общего фрактального дескриптора (D) от точечной группы симметрии (S) для изомеров C_9H_{20}

тенденции. Отметим при этом, что группа C_1 включает в качестве операции симметрии $\{E\}$, C_s содержит $\{E, \sigma_h\}$, а C_{2v} состоит из $\{E, C_2, \sigma_v\}$, т.е. с увеличением числа элементов симметрии в группе величина D уменьшается. Это связано с уменьшением числа уникальных межатомных расстояний (при сохранении их общего числа), что приводит, в свою очередь, к упрощению структуры фрактальной матрицы и, как следствие, к уменьшению фрактальной размерности.

Однако даже при одинаковом числе атомов и в рамках одной группы симметрии величина D изомеров изменяется в некотором интервале. Чтобы объяснить это явление, проанализируем структуру соединений на основе дифференциальной функции распределения межатомных расстояний (рис. 1). При этом все межатомные расстояния в молекуле можно разбить на три группы. В первую группу включены межатомные расстояния между валентно связанными атомами. Ко второй группе относятся межатомные расстояния между атомами, которые связаны Ван-дер-Ваальсовыми силами, третья группа образуется из межатомных расстояний между несвязанными атомами. Типы пар атомов для каждого расстояния определяли путем анализа матриц связности и расчета суммы Ван-дер-Ваальсовых радиусов. Так, в молекуле пропана $CH_3-CH_2-CH_3$ общее число межатомных расстояний равно $11 \cdot 10 / 2 = 55$. Из них 10 межатомных расстояний относятся к первой группе (8 C–H; 2 C–C), 22 – ко второй группе (7 H...H; 14 C...H; 1 C...C) и 23 – к третьей группе. Для каждой группы были рассчитаны соответствующие величины локальных фрактальных дескрипторов (таблица):

$$D_{val} = 1,00 \pm 0,01;$$

$$D_{vdw} = 1,08 \pm 0,03;$$

$$D_{unb} = 1,17 \pm 0,03.$$

Из таблицы видно, что локальные фрактальные дескрипторы имеют разные значения и, следовательно, изменение величины общего фрактального дескриптора D в некотором интервале можно связать с существованием различных типов межатомных расстояний.

Таким образом, каждое из 74 исследованных соединений может быть охарактеризовано с помощью четырех фрактальных дескрипторов. Рассмотрим соотношения между ними на примере ряда изомеров C_9H_{20} .

$$D = -0,11(\pm 0,07) + 0,31(\pm 0,08) D_{val} + 0,20(\pm 0,04) D_{vdw} + 0,63(\pm 0,03) D_{unb}. \quad (2)$$

При использовании автошкалированных величин D_{val} , D_{vdw} , D_{unb} :

$$D = 1,37(\pm 0,02) + 0,008(\pm 0,002) D_{val} + 0,014(\pm 0,003) D_{vdw} + 0,053(\pm 0,002) D_{unb}, \quad (3)$$

$$n = 35; R^2 = 0,978; s = 0,01; Q^2 = 0,966.$$

Полученная регрессионная модель имеет хорошие статистические характеристики. Общий фрактальный дескриптор может быть представлен в виде линейной комбинации локальных фрактальных дескрипторов. Из них наибольший вклад вносит величина D_{unb} . Вклад величин D_{vdw} и D_{val} менее значим.

Аналогичная модель может быть получена для всего ряда исследованных соединений.

$$D = -0,33(\pm 0,08) + 0,46(\pm 0,09) D_{val} + 0,28(\pm 0,04) D_{vdw} + 0,60(\pm 0,03) D_{unb}. \quad (4)$$

При использовании автошкалированных величин D_{val} , D_{vdw} , D_{unb} :

$$D = 1,32(\pm 0,01) + 0,013(\pm 0,003) D_{val} + 0,028(\pm 0,004) D_{vdw} + 0,076(\pm 0,004) D_{unb}, \quad (5)$$

$$n = 74; R^2 = 0,984; s = 0,01; Q^2 = 0,981.$$

В моделях (4) и (5), так же как и в моделях (2) и (3), наибольший вклад в общий фрактальный дескриптор вносит локальный фрактальный дескриптор, рассчитанный на основе расстояний между несвязанными атомами. Вклады D_{vdw} и D_{val} имеют меньшее значение.

Для того чтобы выяснить связь новых фрактальных дескрипторов с другими известными молекулярными дескрипторами, были рассчитаны и проанализированы коэффициенты линейной корреляции между 4 дескрипторами (D , D_{val} , D_{vdw} , D_{unb}) и 3224 дескрипторами, рассчитанными с помощью компьютерной программы DRAGON для исследуемого ряда алканов из 74 соединений. В результате было установлено, что наибольшая линейная связь наблюдается между парами

$$D \text{ и } R5v (R^2 = 0,707),$$

$$D_{val} \text{ и } R8m \text{ с } R^2 = 0,659,$$

$$D_{vdw} \text{ и } RTv \text{ с } R^2 = 0,666,$$

$$D_{unb} \text{ и } H0u \text{ с } R^2 = 0,658,$$

где $R5v$, $R8m$, RTv и $H0u$ представляют собой GETAWAY-дескрипторы [15], которые рассчитываются на основе 3D-структур. Однако величины коэффициентов корреляции не позволяют говорить о тесной линейной связи между анализируемыми молекулярными дескрипторами.

Рассмотренный в настоящей работе метод расчета фрактальных дескрипторов не имеет каких-либо принципиальных ограничений на раз-

мер анализируемых молекул. В частности, мы рассчитали величины дескрипторов для членов гомологического ряда C_nH_{2n+2} :

пентаконтан ($C_{50}H_{102}$)

$$D = 1,20 \pm 0,02,$$

$$D_{val} = 1,00 \pm 0,01,$$

$$D_{vdw} = 1,02 \pm 0,01,$$

$$D_{unb} = 1,25 \pm 0,02;$$

гектан ($C_{100}H_{202}$)

$$D = 1,32 \pm 0,02,$$

$$D_{val} = 1,00 \pm 0,01,$$

$$D_{vdw} = 1,01 \pm 0,01,$$

$$D_{unb} = 1,37 \pm 0,02.$$

Сравнивая эти величины с данными таблицы (№ 1–3, № 5, № 8, № 13, № 22, № 40), можно отметить, что в ряду исследованных соединений с ростом числа атомов величины D и D_{unb} в основном увеличиваются, D_{val} не меняется, а D_{vdw} проходит через максимум.

Таким образом, для расчета новых молекулярных дескрипторов использована дифференциальная функция распределения (гистограмма) межатомных расстояний. С помощью «клеточного» алгоритма вычислены фрактальные размерности 74 алканов. Эти размерности были использованы в качестве четырех новых фрактальных дескрипторов. Они рассчитаны на основе различных типов межатомных расстояний – все расстояния, расстояния между валентно связанными атомами, расстояния между атомами, находящимися в ван-дер-ваальсовом контакте, и расстояния между несвязанными атомами. Показано, что общий фрактальный дескриптор может быть представлен в виде линейной комбинации локальных фрактальных дескрипторов. Установлено, что максимальный линейный коэффициент корреляции R^2 фрактальных дескрипторов с известными 3224 DRAGON-дескрипторами не превышает величину 0,707. Разработанные в настоящей работе фрактальные дескрипторы могут быть использованы для описания структуры как малых, так и больших молекул при конструировании QSAR/QSPR-моделей, включая количественные модели «структура–свойства» полимерных наноконпозиций. Фрактальные дескрипторы можно рассматривать в качестве количественной меры сложности геометрической структуры молекул. В этом качестве они могут найти применение при изучении процессов и состояний, в которых происходит изменение этой характеристики, например, при исследовании фазовых переходов «кристалл–газ», при анализе влияния структуры полимеров на их свойства и др.

Проект выполняется в МГУ имени М.В. Ломоносова в рамках Соглашения с Минобрнауки РФ от 05.06.14 № 14.607.21.0002 (уникальный идентификатор проекта RFMEFI60714X0002). Исследования выполнены при финансовой поддержке Минобрнауки РФ.

СПИСОК ЛИТЕРАТУРЫ

1. Roy K., Kar S., Das R.N. A Primer on QSAR/QSPR Modeling. Fundamental Concepts. Springer International Publishing, 2015.
2. <http://www.taletе.mi.it/>.
3. Мандельброт Б. Фрактальная геометрия природы. М., 2002. 656 с.
4. Pfeifer P., Welz U., Wippermann H. // Chem. Phys. Lett. 1985. Vol. 113. P. 535.
5. Lewis M., Rees D.C. // Science. 1985. Vol. 230. P. 1163.
6. Aqvist J., Tapia J. // J. Mol. Graph. 1987. Vol. 5. P. 30.
7. Pettit F.K., Bowie J.U. // J. Mol. Biol. 1999. Vol. 285. P. 1377.
8. Todoroff N., Kunze J., Schreuder H., Hessler G., Baringhaus K.-H., Schneider G. // Mol. Inf. 2014. Vol. 33. P. 588.
9. Artemenko A.G., Kovdienko N.A., Kuz'min V.E., Kamalov G.L., Lozitskaya R.N., Fedchuk A.S., Lozitsky V.P., Dyachenko N.S., Nosach L.N. // Experimental Oncology. 2002. Vol. 24. P. 123.
10. Григорьев В.Ю., Раевский О.А. // Журн. общ. химии. 2011. Т. 81. С. 353.
11. <http://www.molecular-descriptors.eu/>.
12. Кроновер Р.М. Фракталы и хаос в динамических системах. Основы теории. М., 2000.
13. <http://www.hyper.com/>.
14. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. М., 1980.
15. Consonni V., Todeschini R., Pavan M. // J. Chem. Inf. Comput. Sci. 2002. 42. P. 682.

Поступила в редакцию 08.12.15

CALCULATION AND PROPERTIES OF FRACTAL DESCRIPTORS FOR C₂...C₉ ALKANES

V.Yu. Grigorev¹, L.D. Grigoreva²

(¹*Institute of Physiologically Active Compounds RAS;* ²*Department of Fundamental Physical-Chemical Engineering*)

For calculation of molecular descriptors among 74 alkanes is used discrete differential distribution function (histogram) of interatomic distances. On its basis are calculated fractal dimensions (descriptors) of 4 types. Relation of fractal descriptors between themselves, and also with number of atoms and point group of symmetry of molecules is revealed.

Key words: QSAR, QSPR, fractal dimension, fractal descriptor.

Сведения об авторах: Григорьев Вениамин Юрьевич – вед. науч. сотр. отдела компьютерного молекулярного дизайна ИФАВ РАН, докт. хим. наук (beng@ipac.ac.ru); Григорьева Людмила Дмитриевна – доцент факультета фундаментальной физико-химической инженерии МГУ имени М.В. Ломоносова, канд. физ.-матем. наук (ldg@physchem.msu.ru).