

DOI: 10.15643/mmcct-2023-21

Информационная энтропия молекулярного ансамбля: неаддитивность и применение к анализу химических реакций

Д. Ш. Сабиров

Институт нефтехимии и катализа УФИЦ РАН

Россия, Республика Башкортостан, 450075 г. Уфа, пр. Октября, 141.

Email: diozno@mail.ru

В лаборатории математической химии ИНК УФИЦ РАН было найдено, что информационная энтропия молекулярного ансамбля складывается не только из значений информационной энтропии молекул – и включает дополнительное слагаемое, названное корпоративной энтропией $H\Omega$, которая возникает при объединении молекул в ансамбль. Корпоративная энтропия не зависит от строения молекул и определяется только их размером. Эта особенность описания строения молекулярных ансамблей открывает новые возможности для цифрового описания химических реакций органических соединений.

Ключевые слова: информационная энтропия, кооперативная энтропия, молекулярный ансамбль, химическая реакция, энтропия перераспределения, энтропия реорганизации.

Широкое распространение для описания строения молекул получили структурные дескрипторы теоретико-множественного типа на основе информационной энтропии (h), являющиеся инвариантами молекулярных графов [1]. Информационная энтропия молекулы рассчитывается с учётом числа типов неэквивалентных атомов в её структуре (n) и их заселённостей (N_j ; $N = \sum N_j$ – общее число атомов в молекуле):

$$h = - \sum_{j=1}^n \frac{N_j}{\sum_{j=1}^n N_j} \log_2 \frac{N_j}{\sum_{j=1}^n N_j} \quad (1)$$

Расчет значений h для молекул подробно рассматривался ранее. Здесь приводится пример такого расчета для двух изомеров C_5H_{12} (рис. 1). Молекуле ставится в

соответствие граф, затем проводится сортировка его вершин, т.е. разбиение молекулы (множества вершин-атомов) на подмножества атомных типов. При этом учитывается неэквивалентность атомов, т.е. их принадлежность к тому или иному химическому элементу и положение в молекулярном графе. Если молекула «разбивается» на относительно большие атомные типы (подмножества с высокими кардинальными числами), ее информационная энтропия меньше, чем значение h для изомера, например, с равномерным разбиением. Соответственно, наличие высоких точечных групп симметрии в молекулярной структуре является фактором, понижающим значение h , и наоборот. В целом, значения h не зависят от размера молекул (N) и определяются только структурой [2, 3].

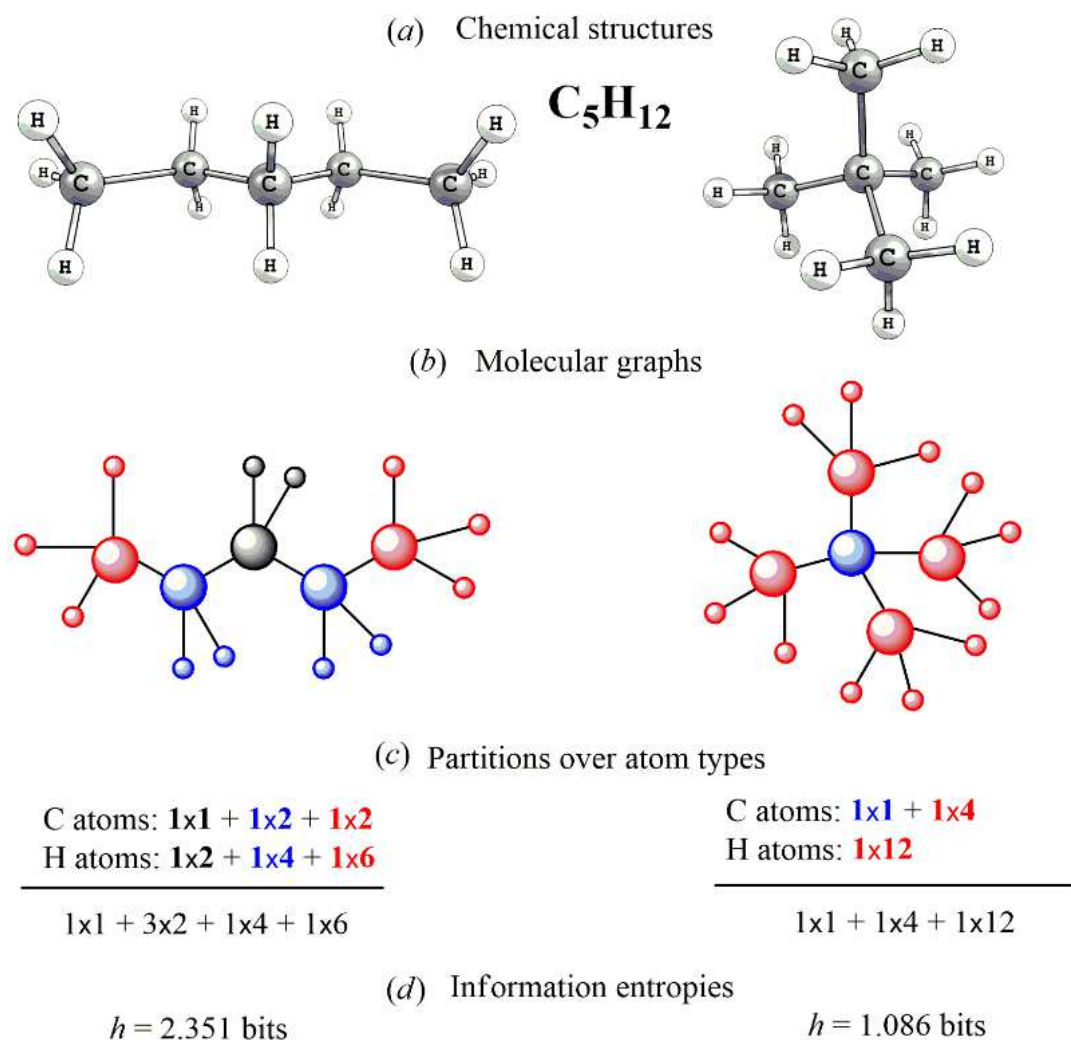


Рис. 1. Схема вычисления информационной энтропии двух изомеров пентана: (а) молекулы; (б) соответствующие молекулярные графы с выделенными вершинами в зависимости от принадлежности тому или иному атомному типу; (с) подсчет неэквивалентных вершин и представление результата в виде «формулы разбиения»; (д) вычисление значений h по уравнению 1.

Молекулярным ансамблем называют группу молекул, рассматриваемых в качестве единого целого при вычислении информационной энтропии. В общем случае расчёт информационной энтропии ансамбля h_{ME} из m молекул выполняется аналогично формуле (1):

$$h_{ME} = -\sum_{i=1}^m \sum_{j=1}^n \frac{N_{ij}}{N} \log_2 \frac{N_{ij}}{N} \quad (2)$$

Последнее выражение мы преобразовали к виду [4]:

$$h_{ME} = H_{\Omega} + \sum_{i=1}^m \omega_i h_i \quad (3)$$

с учетом следующих обозначений. Доля атомов, приходящаяся на каждый тип рассматриваемых молекул:

$$\omega_i = \frac{N_i}{\sum_{i=1}^m N_i} \quad (4)$$

Кооперативная энтропия H_{Ω} – слагающее, отражающее изменение информационной энтропии при объединении молекул в ансамбль:

$$H_{\Omega} = -\sum_{i=1}^m \omega_i \log_2 \omega_i \quad (5)$$

В терминах теории самоорганизации кооперативная энтропия является эмерджентным параметром, то есть H_{Ω} – характеристика системы (молекулярного ансамбля), но не ее элементов (отдельных молекул). Этот параметр обращается в нуль в случае мономолекулярного ансамбля (когда $\omega = 1$) и достигает максимального значения $H_{\Omega} = 1$, если количество частиц равно количеству типов атомов $m = N$ и $N_i = 1$. Последний случай соответствует ансамблю, составленному из разных атомов.

Формула (3) указывает на неаддитивность информационной энтропии молекулярного ансамбля – она не может быть вычислена как простая сумма значений h_i молекул. Простое суммирование приводит к результатам, противоречащим химическому смыслу – и только использование уравнения (3) позволяет получать оценки, соответствующие основным понятиям структурной химии [4,5]. Примечательно, что в информационной энтропии молекулярного ансамбля можно выделить члены, которые связаны только со строением молекул (h_i) либо только с их размером (ω_i).

В основе применения развиваемого подхода к химическим реакциям лежит представление последних как процессов превращения одного молекулярного ансамбля (исходных веществ) в другой (продукты) [6]. В рамках этого формализма изменение информационной энтропии в химической реакции рассчитывается как

$$\Delta h_R = h_{ME}^{prod} - h_{ME}^{react} \quad (6)$$

Подстановка в уравнение (6) выражений, полученных с помощью уравнения (3), и простые математические преобразования позволяют разделить Δh_R на два вклада:

$$\Delta h_R = H_{redistr} + H_{reorg} \quad (7)$$

где первое слагаемое – **энтропия перераспределения**, связанная с изменением размеров молекул в ходе химического превращения:

$$H_{redistr} = H_{\Omega}^{prod} - H_{\Omega}^{react} \quad (8)$$

Второй вклад, названный **энтропией реорганизации**, зависит как от h_i так и от ω_i , т.е. и от строения и от размера молекул:

$$H_{reorg} = \sum_i^{prod} \omega_i h_i - \sum_j^{react} \omega_j h_j \quad (9)$$

В последнем выражении можно выделить член, зависящий только от химического строения (H_{reorg}^{str}), но зависимость оставшейся части ($H_{reorg}^{str+size}$) одновременно от строения и размера не устранима:

$$H_{reorg} = H_{reorg}^{str} + H_{reorg}^{str+size} \quad (10)$$

$$H_{reorg}^{str} = \sum_i^{prod} h_i - \sum_j^{react} h_j \quad (11)$$

$$H_{reorg}^{str+size} = \sum_j^{react} (1 - \omega_j) h_j - \sum_i^{prod} (1 - \omega_i) h_i \quad (12)$$

Только первый член энтропии реорганизации (уравнения (11)) определяется исключительно строением молекул и рассчитывается аналогично изменениям термодинамических функций в химических процессах.

Таким образом, изложенная схема расчета позволяет рассчитывать не только полное изменение информационной энтропии в химической реакции Δh_R , но и ее составляющие, связанные с изменением размера и/или строения молекулы:

$$\Delta h_R = H_{redistr}^{size} + H_{reorg}^{str} + H_{reorg}^{str+size} \quad (13)$$

Важно, что выделение отдельных вкладов – результат строгих математических преобразований. Предложенный формализм согласуется с оценками других информационно-энтропийных подходов. Для реакций определённых классов можно получить аналитические зависимости Δh_R (табл. 1).

Таблица 1. Аналитические выражения для информационной энтропии типичных химических реакций

Реакция	Формальное уравнение	h_{ME}^{prod}	h_{ME}^{react}	H_{reorg}	$H_{redistr}$
Диссоциация	$D \rightarrow A + B + \dots + C$	$\sum \omega_i h_i + H_{\Omega}^{prod}$	h_D	$\sum \omega_i h_i - h_D$	H_{Ω}^{prod}
Присоединение	$A + B + \dots + C \rightarrow D$	h_D	$\sum \omega_i h_i + H_{\Omega}^{react}$	$h_D - \sum \omega_i h_i$	$-H_{\Omega}^{react}$
Атомизация	$A_a B_b \dots C_c \rightarrow aA + bB + \dots + cC$	H_{Ω}^{prod}	$h(A_a B_b \dots C_c)$	$-h(A_a B_b \dots C_c)$	H_{Ω}^{prod}
Изомеризация	$A \rightarrow B$	h_B	h_A	$h_B - h_A$	0

Обсуждаются количественные оценки, полученные в рамках развиваемого подхода, для реакций Дильса – Альдера, этерификации, циклоприсоединения к фуллеренам и др. [7]

Было изучено заполнение фуллереновых клеток атомами X с образованием эндофуллеренов $X + C_N \rightarrow X@C_N$. Такие процессы представляют собой удобную модель для тестирования структурных дескрипторов, поскольку происходят с минимальными изменениями химического строения. В зависимости от типа фуллерена и инкапсулируемого атома возможны несколько вариантов процесса, различающихся соотношением точечных групп симметрии образующегося эндофуллерена и исходного фуллерена. Сохранению, увеличению и понижению симметрии соответствуют нулевое, отрицательные и положительные значения Δh_R . Помимо этого, вклады в Δh_R зависят от размера и строения участвующего в процессе фуллерена. Это позволяет предложить вариант алгоритмизированной классификации, в которой по расчетным информационно-энтропийным параметрам заполнения C_N делается вывод о строении C_N (рис. 2), то есть по параметрам химического процесса делается вывод о структуре, которая в нем участвует. Последнее представляет интерес для развития математических методов в структурной химии. Одной из основных задач классической химии является установление строения химического соединения по его химическим свойствам. Таким образом, развиваемый подход способен решить «цифровой» аналог такой задачи: из количественных оценок химических реакций, в которой участвует молекула – сделать заключение о ее структурных особенностях [8].

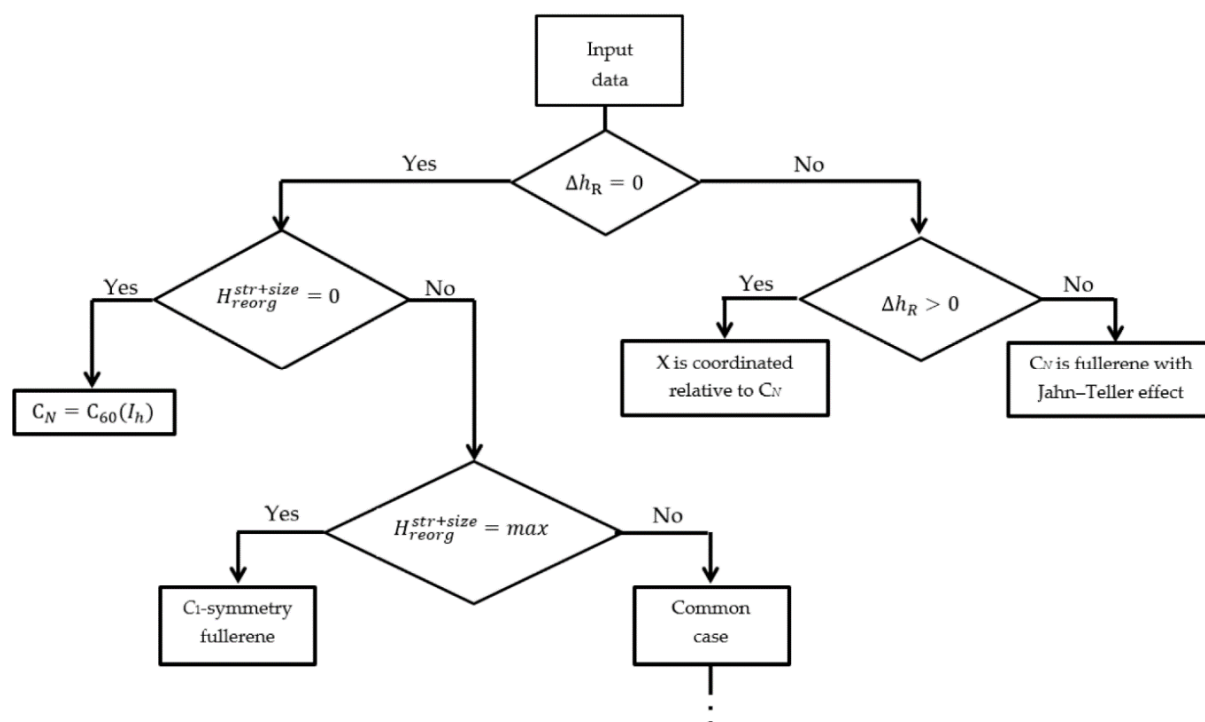


Рис. 2. Алгоритм определения размера и типа фуллерена, участвующего в процессе $X + C_N \rightarrow X@C_N$, исходя из информационно-энтропийных параметров процесса заполнения. Входные параметры – полное изменение информационной энтропии в процессе Δh_R и его составляющие.

Работа выполнена при поддержке Российского научного фонда (проект №22-13-20095).

Литература

1. Barigye S.J., Marrero-Ponce Y., Pérez-Giménez F., Bonchev, D. // *Mol. Divers.* **2014**. V. 18. P. 673.
2. Sabirov D.S., Osawa E. // *J. Chem. Inf. Model.* **2015**. V. 55. P. 1576.
3. Sabirov D., Koledina K. // *EPJ Web Conf.* **2020**. V. 244. P. 01016.
4. Sabirov D.S. // *Comput. Theor. Chem.* **2018**. V. 1123. P. 169.
5. Sabirov D.S. // *Comput. Theor. Chem.* **2020**. V. 1187. P. 112933.
6. Sabirov D.S., Tukhbatullina A.A., Shepelevich I.S. // *J. Mol. Graph. Model.* **2022**. V. 110. P. 108052.
7. Sabirov D.S., Shepelevich, I.S. // *Entropy.* **2021**. V. 23. P. 1240.
8. Sabirov D.S., Tukhbatullina A.A., Shepelevich I.S. // *Symmetry.* **2022**. V. 14. P. 1800.