

Физическая информатика.

Информационные характеристики молекул

Аннотация

Взаимопроникновение физики и теории информации в ходе развития информатики сформировало синтетическую дисциплину «Физическая информатика». Информационные законы совместно с физическими законами могут служить эффективным инструментом познания физических систем и Вселенной в целом. Взаимосвязь между физическими и информационными характеристиками систем – массой, энергией, энтропией и информацией дает возможность использовать информационные оценки и методы исследования физических характеристик систем.

Методика оценки объема информации в молекулах. Объем информации в молекуле I_{ml} равен сумме объема информации в атомах $\sum_i N_i I_{iat i}$ и информации в структуре молекулы

$I_{str ml} : I_{ml} = \sum_i N_i I_{iat i} + I_{str ml}$, где N_{iat} – количество атомов типа i в молекуле, I_{iat} – объем информации в атоме типа i .

Оценка объема информации в структуре молекул. Предлагается использовать оценку объема информации в структуре графа, соответствующей структуре молекулы

$I_{gr} = \sum_{i=1}^m n_i \log_2 v_i$ приведенную в работах автора [1-6]. Здесь где m – количество классов топологически эквивалентных вершин графа (вершин заданной степени v_i), n_i – число вершин графа степени v_i . Данная оценка является оценкой сверху. Рассматриваемое

количество направлений выхода из каждой вершины обеспечивает обход графа с использованием всех возможных путей (направлений перехода) от атома к атому. Оценка объема информации в структуре графа, соответствующей структуре молекулы, использующая степени вершин, уменьшенные на единицу, $I'_{gr} = \sum n_j \log_2 (v_j - 1)$, является

оценкой снизу. Для сравнения рассмотрим оценки N. Rashevsky [7] и E. Trusso [8]. Оценка

N. Rashevsky: $I_{gr} = - \sum_{i=1}^m w_i \log_2 w_i$, где m – количество классов топологически

эквивалентных вершин графа, $w_i = \frac{n_i}{n}$ – вероятность принадлежности вершины графа к

классу i . Объем информации в структуре однородного графа ($w_i = 1$) равен нулю. Оценка

E. Trusso: $I_{gr} = - \sum_{i=1}^m w_i \log_2 w_i - \sum_{i=1}^m \gamma_i \log_2 \gamma_i$, где $\gamma_i = 1 - w_i$. Объем информации в структуре

однородного графа ($w_i = 1$) равен нулю.

Сравнение оценок объемов информации в структурах молекул по N. Rashevsky, E. Trusso и Гуревича И.М. Показано, что оценки объемов информации по N. Rashevsky и E. Trusso в структурах молекул, описываемых однородными графами, равны нулю. Оценки объемов информации по N. Rashevsky и E. Trusso в графах, описывающих структуры молекул типа бутана практически не зависят от количества в молекуле атомов углерода. Это представляется неадекватным реальности. Следовательно, объем информации в структурах молекул, молекулах в целом целесообразно оценивать по методике Гуревича.

Рассмотрены несколько примеров. Приведены оценки объема информации в молекулах метана CH_4 , бутана C_4H_{10} , молекулах лекарственных препаратов (ацетилсалициловая кислота, золедроновая кислота, индапамид, кларитромицин,

рифампицин, аторвастатин, резвератрол), в молекулах оснований и в триплетах оснований, кодирующих аминокислоты. Показано, что кодирование аминокислот и, соответственно белков, структурами молекул оснований в 76 раз экономнее, чем кодирование собственно молекулами оснований. Избыточность генетического кодирования при кодировании структурой молекул равна 2,9769, при кодировании молекулами равна 2,7244. Таким образом, избыточность генетического кодирования равна почти трем, причем кодирование структурами молекул оснований примерно на 9% эффективнее, чем кодирование самими молекулами оснований (0,093-0,085) [5,6].

На основании вышеизложенного сделаны следующие основные выводы.

1. Наряду с общепринятыми физическими и химическими характеристиками молекулы описываются информационными характеристиками, в частности, объемом информации в структуре молекул и молекулах.
2. Предложенная методика позволяет просто оценивать объем информации в структуре молекул на основании общепринятых структурных формул.
3. **Следует дополнить общепринятое определение химии: химия это наука, изучающая наличие и изменение информации в структуре молекул.**

Литература

1. Гуревич И.М. Оценка объема неопределенности (информации) в элементарных частицах, атомах и молекулах/И.М. Гуревич//Вестн. СевНТУ. Сер. Физика и математика: сб.научн.тр.- Севастополь, 2009. – Вып. 99. – С. 121-129.
2. Гуревич И.М. Информационные характеристики физических систем. / И.М.Гуревич.– «Кипарис». Севастополь. – 2010. – 260 с.
3. Гуревич И.М. Оценка объема информации в структурах молекул и молекулах/И.М. Гуревич//Материалы международной научно-технической конференции «Актуальные вопросы биологической физики и химии БФФХ-2011». Севастополь, 2011. – С. 185-187.
4. Гуревич И.М. Информация – всеобщее свойство материи. Характеристики. Оценки. Ограничения. Следствия./И.М. Гуревич, А.Д.Урсул.// – М.: Книжный дом «ЛИБРОКОМ». – 2012. – 312 с.
5. Гуревич И.М. Автоматизация оценки объема информации в структурах молекул и молекулах/И.М. Гуревич, М.П. Евстигнеев, М.А. Пучков. //Материалы международной научно-технической конференции «Актуальные вопросы биологической физики и химии БФФХ-2012». Севастополь, 2012. – С. 93-95.
6. Гуревич И.М. Оценка объема информации в соединениях цепей ДНК/И.М. Гуревич //Материалы международной научно-технической конференции «Актуальные вопросы биологической физики и химии БФФХ-2012». Севастополь, 2011. – С. 90-92.
7. Rashevsky N. «Same Theorems in Topology and a Possible Biological Implication». / N. Rashevsky //Bulletin of mathematical biophysics. – 1956. – Volume 17, – P. 111-126.
8. Trucco E. «On the information content of graphs: compound symbols; different states for each point». / E. Trucco //Bulletin of mathematical biophysics. – 1956. – Volume 16. – P. 237-258.

Информация о целях и месте проведения семинара – см. сайт
www.soc-phys.chem.msu.ru