

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Кафедра физической химии и хроматографии

С. В. Курбатова, Е. А. Колосова, Е. Е. Финкельштейн

**ТОПОЛОГИЧЕСКИЕ ИНДЕКСЫ
В ХИМИЧЕСКИХ РАСЧЕТАХ**

*Утверждено редакционно-издательским советом университета
в качестве учебного пособия*

Самара
Издательство «Самарский университет»
2014

УДК 577.1
ББК 28.9
К 93

Рецензенты: д-р хим. наук, проф. П. П. Пурыгин,
д-р хим. наук, проф. А. В. Буланова

Курбатова, С. В.

К 93 Топологические индексы в химических расчетах : учебное пособие /
С. В. Курбатова, Е. А. Колосова, Е. Е. Финкельштейн. – Самара : Изда-
тельство «Самарский университет», 2014. – 32 с.

Данное учебное пособие включает основную часть теоретического мате-
риала, читаемого в курсах лекций «Физико-химические методы исследова-
ния» и «Физико-химические основы методов исследования веществ и мате-
риалов».

Предназначено для бакалавров и магистров направления «Химия», мо-
жет быть полезно при выполнении курсовых и выпускных квалификацион-
ных работ.

УДК 577.1
ББК 28.9

© Курбатова С. В., Колосова Е. А.
Финкельштейн Е. Е., 2014
© ФГБОУ ВПО «Самарский
государственный университет», 2014

ПРЕДИСЛОВИЕ

Одной из важных задач теоретической химии является определение свойств химических соединений на основании их молекулярной структуры. Свойства молекулы закодированы в ее структуре, поэтому знание количественного соотношения «строение – свойство» помогает наиболее точно расшифровать эти свойства и, таким образом, понять, как строение соединения влияет на его способность вступать в межмолекулярные взаимодействия. Вопрос о понимании взаимоотношения «строение – свойство» стал наиболее актуален в последнее время, так как ежедневно в мире синтезируют огромное количество новых соединений, имеющих широкий спектр применения. Именно поэтому необходимо решить задачу по установлению связи между строением и свойствами, чтобы иметь возможность прогнозировать свойства гипотетических молекул, а, следовательно, вести синтез соединений с программируемыми (заданными) свойствами. Решение этой задачи даже для одного класса соединений значительно расширяет возможности теоретической химии.

Для нахождения взаимосвязи между строением и свойствами соединений используются различные методы математического моделирования, выбор которых зависит от природы исследуемых химических соединений и анализируемых свойств. Среди этих методов особое место занимают топологические подходы, использующие только информацию, содержащуюся в структурной формуле исследуемого соединения. При решении задач «структура – свойство» топологический подход является весьма перспективным, быстрым и удобным, так как обладает высокой информативностью и широкими аналитическими и прогнозирующими возможностями. Топологическая оценка структуры химических соединений представляет собой универсальный метод описания особенностей химического строения.

При изучении взаимосвязи «строение – свойство» важными являются различные способы перевода молекулярных структур химических соединений в единые характеристические структурные дескрипторы, выраженные численными (топологическими) индексами. Топологические индексы (ТИ) используют для кодирования химической информации, оценки реакционной способности молекул, при планировании химического эксперимента, для количественного описания химических структур при анализе связи между строением молекулы, ее свойствами и биологической активностью. Способность различать очень близкие по строению и составу соединения делает их одними из наиболее распространенных молекулярных дескрипторов в корреляционном анализе.

Преимущество ТИ состоит в том, что они не требуют сложных экспериментальных определений или квантово-химических расчетов, а вычисляются непосредственно из структурной формулы соединений. Это особенно важно, когда речь идет об оценке свойств гипотетических структур, для которых отсутствуют различные данные. К недостаткам ТИ можно отнести невозможность учета всех особенностей молекулярного строения, а также сложность их физико-химической интерпретации. Тем не менее, корреляционные соотношения, полученные с их помощью, оказываются весьма полезными в решении различных практических вопросов.

Топологические индексы коррелируют со многими физико-химическими характеристиками (температура кипения и плавления, плотность, теплоты образования и атомизации, молекулярная поляризуемость и рефракция, растворимость, молярный объем, диамагнитная восприимчивость, октановое число и др.) и величинами, характеризующими биологическую активность соединений.

Топологический подход к установлению взаимосвязи «структура-свойство»

Строение молекулы может быть описано с различной степенью детализации. В литературе изложены различные качественные представления об иерархии уровней описания молекул органических соединений. Так, брутто-формула показывает, сколько и каких атомов содержится в молекуле, структурная формула содержит информацию о том, каким образом эти атомы связаны между собой. Наконец, задание координат ядер атомов в пространстве позволяет строить трехмерные модели молекул. С помощью квантовомеханических расчетов можно получить дополнительную информацию о распределении электронной плотности в молекуле. В ряде случаев для характеристики молекул используют физико-химические параметры, отражающие строение косвенным образом. Современные экспериментальные методы исследования позволяют достаточно надежно определять геометрию и различные физико-химические характеристики молекул. Но подобные методы применимы лишь к реально существующим веществам, причем не ко всяким. В тех случаях, когда эксперимент затруднен, могут быть использованы квантовохимические расчеты геометрических и энергетических характеристик молекул. Кроме того, как уже было сказано, для решения проблемы «структура – свойство» используют и более формальные и простые методы, основанные на описании строения молекулы структурной формулой. При таком топологическом способе описания молекулярного строения учитывают лишь типы атомов и характер их связывания, но пренебрегают

метрическими характеристиками молекулы. Однако, при кажущейся простоте топологический способ описания строения может вполне адекватно охарактеризовать соединение, так как известно, что многие свойства, особенно в сопряженных π -электронных системах, являются, в первую очередь, следствиями схемы связанности атомов в молекуле. Следовательно, топологический подход к строению молекул предполагает, прежде всего, исследование порядка взаимодействия атомов в молекуле. Этот факт является основой большинства применений теории графов для изучения молекулярной структуры, к которым можно отнести, в частности, такую модель, как теория молекулярных орбиталей Хюккеля. Поэтому в современной теоретической химии наблюдается тенденция к более широкому использованию математического аппарата для описания молекулярных структур и химических превращений, а теоретико-графовые и топологические представления приобретают все возрастающую роль в исследовании строения и свойств химических соединений.

Известно, что существует достоверное соответствие между первичными концепциями, лежащими в основе понятия молекулярной структуры, и топологическими элементами молекулярного зарядового распределения. Полная теория основывается на единственном наиболее важном свойстве зарядового распределения – оно имеет локальные максимумы лишь в точках расположения ядер. С топологической точки зрения положения ядер соответствуют критическим точкам в зарядовом распределении $\rho(r, X)$. В результате ядра представляют собой точечные аттракторы в векторном поле градиента $\rho(r, X)$ – поле $\nabla\rho(r, X)$. Поэтому атом определяют как объединение аттрактора и его бассейна. В результате конкуренции двух соседних аттракторов за находящуюся между ними зарядовую плотность возникает критическая точка, а бассейны соседних аттракторов разделяются межатомной поверхностью. Траектория, начинающаяся у критической точки и оканчивающаяся у одного из соседних ядер, определяет линию межатомного взаимодействия – путь через пространство, связывающий соседние ядра, вдоль которых зарядовая плотность максимальна по сравнению с любым другим соседним путем. Полная сеть связевых путей для данной ядерной конфигурации (\bar{q}) характеризует молекулярный граф, другими словами, молекулярный граф определяется как объединение линий взаимодействия и является однозначной характеристикой соответствующего векторного поля градиента $\nabla\rho(r, X)$. Таким образом, само понятие «молекулярная структура» определяется как эквивалентный класс молекулярных графов, то есть одно и то же число связевых путей связывает одни и те же ядра в каждом молекулярном графе, принадлежащем данной молекулярной структуре. Все

конфигурационное пространство ядерных координат (\bar{q}) можно подразделить на конечное число структурных областей, каждой из которых отвечает своя молекулярная структура. На границах раздела таких областей происходит скачкообразное изменение структурного типа, то есть переход от одной молекулярной формы к другой. Таким образом, понятие топологии вполне адекватно отражает электронную структуру молекул и атомов.

Молекулярный граф

Химическое строение молекул обладает основным свойством топологических структур – сохранением целостности и непрерывности взаимодействия атомов в молекуле при всех изменениях геометрии: межатомных расстояний, валентных углов.

Молекулы различаются атомами, из которых они состоят, и способом связывания атомов друг с другом, то есть, своей топологией. Строение молекул может быть удобно изображено на языке теории графов. Важную роль молекулярные графы играют при описании внутри- и межмолекулярных взаимодействий, взаимопревращений изомеров, способствуют упорядочиванию молекулярных свойств, механизмов химических реакций. Графы дают возможность конкретных описаний многих абстрактных величин, применяемых в химии. В органической химии графы применяются для представления структурных формул молекул (молекулярные графы).

Молекулярный граф определяется как объединение линий взаимодействия, и такой граф может быть отнесен к каждой точке X в пространстве ядерных конфигураций.

Удобно определить молекулярный граф связей, как конструкцию, состоящую из точек (ядер) и ребер (связей), в которых разные типы ядер определяют разнотипные точки, а различные типы связей (например, двойная и простая) – различающиеся ребра. Понятие молекулярный граф практически соответствует понятию химический граф.

Структурные диаграммы, с помощью которых изображают молекулы, являются химическими графами, указывая атомный состав и связи.

Таким образом, молекулярные структуры фактически являются химическими графами, в которых атомы представляются вершинами, а ковалентные химические связи – ребрами. Такой граф описывает связанность атомов в молекулярном скелете независимо от метрических свойств, то есть равновесного межъядерного расстояния, валентных углов, соответствующих данной химической структуре.

Часто рассматриваются графы, которые описывают “голые” молекулярные скелеты, то есть молекулярные структуры, при изображении которых атомы водорода опускаются. Такие молекулярные графы называют графами со стертыми атомами водорода. Следовательно, химический граф – это граф с конечным числом помеченных ребер и вершин.

Веса вершин выбирают таким образом, чтобы учесть различие в типах атомов и связей. Кроме того, задавая по-разному веса на молекулярном графе, можно получать инварианты, отражающие различные особенности не только топологического, но и электронного, а также пространственного строения молекул (например, разветвленность, связанность, сложность, информационное содержание). Лучше всего для химических графов подошел бы термин “структурные графы”, поскольку молекулы, имеющие неизоморфные графы, классифицируются как структурные изомеры.

В молекулярном графе вершины соответствуют атомам и могут быть обозначены соответствующими химическими символами, а ребра, то есть линии химических графов, также часто отмечаются обозначениями кратности (простая, двойная, тройная и т.д.) или типа (σ -, π -, ионная и т.д.) связей. При этом более простые представления молекул (в виде конечных множеств, последовательностей, векторов) получаются из графового представления, а меры, определяемые для этих представлений, могут рассматриваться как меры на графах.

Топологические индексы как структурные дескрипторы

Как уже отмечалось, существуют различные способы перевода молекулярных структур химических соединений в единые характеристические структурные дескрипторы, выраженные численными индексами. Под дескрипторами понимают независимые переменные, характеризующие структурные особенности соединений или их частей (заместителей, фрагментов), и используют в количественном описании молекулярной структуры химических соединений в рамках компьютерного дизайна веществ.

Возможны различные способы классификации дескрипторов. Как правило, рассматривают три типа классификации.

В первом варианте классификации дескрипторы делят на чистые (описывающие какой-то один эффект межатомных взаимодействий) и композитные (описывающие два и более эффекта). Среди композитных дескрипторов, в свою очередь, выделяются уникompозитные (описывающие эффекты одного и того же типа) и мультикомпозитные (отображающие сумму эффектов различных типов).

Второй тип классификации основывается на способах оценки дескрипторов – экспериментальных или теоретических.

Основу третьего типа классификации составляет учет эффекта, который описывает данный дескриптор. Предлагается рассматривать три категории эффектов: электронные (электрические), стерические и межмолекулярные.

Одной из основных проблем при установлении зависимости между структурой и свойствами веществ является выбор молекулярных дескрипторов и вида функциональной зависимости, наилучшим образом описывающих связь между этими дескрипторами и свойством. Выбор дескрипторов производится, как правило, из заранее заданного конечного множества путем перебора вариантов. При этом требуется, чтобы число дескрипторов было меньше числа соединений в исходной выборке. Существенное преимущество структурных дескрипторов по сравнению с многочисленными экспериментальными характеристиками состоит в том, что они определяются непосредственно из химической формулы (структуры) вещества и могут быть легко вычислены как для реальных химических соединений, так и для гипотетических структур.

Таким образом, анализ структурной формулы, который приходится проводить в рамках проблемы «структура – свойство», сводится к обработке методами теории графов соответствующей ей абстрактной математической структуры – молекулярного графа (МГ). Как уже отмечалось выше, наиболее распространенными дескрипторами, используемыми при исследовании корреляций структура – свойство, являются инварианты взвешенных молекулярных графов, описывающих химические структуры. Молекулярный граф G – это двумерное отображение молекулы (вершины соответствуют атомам, а ребра – химическим связям молекулы). При этом обычно рассматриваются скелетные атомы (со “стертыми” атомами водорода) и связи между ними.

Каждый молекулярный граф можно представить матрицей смежности $A(G)$ и расстояния $D(G)$, полиномом или числовым индексом (топологическим). Большинство современных индексов основано либо на матрице расстояний, либо на матрице смежности (еще в конце 19 века появились работы немецкого химика Коппа, в которых была высказана идея о том, что физико-химические свойства отдельных членов ряда родственных соединений могут быть предсказаны по их положению в матрице). При этом каждый элемент матрицы расстояний представляет собой число ребер, соединяющих две вершины наикратчайшим путем. Элементы матрицы смежности равны либо единице, либо нулю в зависимости от того, связана ли ребром одна вершина графа с другой.

Обычно в структурной химии используют взвешенные МГ. Взвешенный граф определяется своей матрицей смежности $A = (a_{ij})$, $i, j = 1, \dots, n$, где a_{ij} – вес i -й вершины; n – число вершин графа; $a_{ij} = a_{ji}$ – вес ребра (i, j) ; $a_{ij} = a_{ji} = 0$, если вершины i, j не связаны ребром. Вид матрицы смежности зависит от способа нумерации вершин графа. Но каждому графу можно приписать различные числа, характеризующие его строение и не зависящие от способа нумерации вершин. Такие числа называются инвариантами графа или структурными константами. Примерами инвариантов графа являются число вершин, число ребер, число путей заданной длины, определитель матрицы смежности и т.д. Инварианты взвешенных графов и называют топологическими индексами. Они строятся путем преобразования химического графа в числовое значение. ТИ, как правило, рассчитываются в результате формального анализа структуры молекулярного графа. В ряде случаев такой анализ дополняется соображениями качественного характера в выборе критериев для оценки степени разветвленности, компактности молекулы или наличия системы сопряженных двойных и тройных связей. Если молекулярная система содержит достаточно протяженный плоский фрагмент, то это дает возможность разбить ее электронную систему на две части – σ и π - составляющие и исследовать каждую систему независимо. При этом, как правило, рассматривают только систему π - электронов, предполагая, что она находится в некотором эффективном поле, которое создается атомными ядрами, электронами внутренних оболочек и σ - электронами, принимающими участие в образовании молекулярных двухцентровых σ - связей. В случае сопряженных углеводородов число π - электронов обычно совпадает с числом атомов углерода, образующих сопряженную часть молекулы. Структура сопряженной части может быть описана молекулярным графом в упрощенной форме. Его получают из полного МГ путем удаления вершин, соответствующих одновалентным заместителям. Такие величины, как собственные числа МГ, элементы матрицы плотности и другие характеристики молекулярных систем могут быть легко вычислены, однако их трудно интерпретировать непосредственно в структурных терминах. Между тем такая интерпретация представляет интерес для понимания источников возникновения определенных функциональных связей между структурой молекулы и ее свойствами. Поэтому в последние годы достаточно интенсивно развивалось направление, связанное с поиском соотношений между результатами расчетов по методу Хюккеля и наглядными характеристиками МГ.

Возможен и другой подход к проблеме «структура – свойство», связанный с ТИ, суть которого заключается в том, что на основе ТИ вводят понятие «рас-

стояния» между структурами, позволяющее оценивать степень близости структур, и соотносить ее с изменением рассматриваемого свойства. В настоящее время разработаны и применяются разновидности топологических индексов – центрические, составные, теоретико- и тополого-информационные (использование которых базируется на функциях Шэннона), а также тополого-конфигурационные, значения которых определяются не только структурой, но и относительным расположением атомов рассматриваемого соединения на плоскости (т.е. плоской укладкой молекулярного графа). Описывается применение так называемых трехмерных индексов, при расчете которых используют реальные расстояния между атомами в молекуле.

Для установления связи между структурными параметрами x_1, \dots, x_k (в данном случае - ТИ) и числовой характеристикой некоторого свойства «Y», как правило, ищут функциональную зависимость вида $Y = a_0 + \sum a_i x_i$, где a_i – неизвестные константы, определяемые методом наименьших квадратов.

Применение индексов связанности для характеристики структуры молекул

Одним из наиболее популярных топологических индексов, достаточно широко используемых в теоретической химии, является индекс Рэндиша, или индекс связанности (ИС). ИС представляет собой численное описание структуры молекулы и рассчитывается путем математической обработки параметров, приписываемых атомам молекулярного скелета с учетом степени разветвления скелета, циклизации, ненасыщенности, наличия гетероатомов, и, таким образом, характеризующий молекулярную способность к образованию связей и описывающий топологические свойства молекул. Различают индексы связанности разных порядков ($^n\chi$).

Индекс связанности первого порядка ($^1\chi$) учитывает парные взаимодействия соседних атомов в молекуле. К его недостаткам можно отнести невозможность различения мета- и пара-изомеров ароматических соединений. Индексы связанности высоких порядков, учитывающие не только ближайшие парные, но и отдаленные взаимодействия атомов в молекуле, обеспечивают более полное описание структурных особенностей соединений.

Индексы связанности третьего и последующих порядков могут быть представлены в виде индексов цепи (линейный индекс) χ_p и разветвления (узловой индекс) χ_c .

Линейный индекс третьего порядка учитывает атомы молекулы, образующие цепь (линию) из трех соседних связей.

Узловой индекс третьего порядка рассматривает атомы трех соседних связей, образующие узел (структуру типа треноги).

Аналогичный индекс четвертого порядка ${}^4\chi_{рс}$ называется линейно-узловым индексом (в отличие от линейного индекса ${}^4\chi_{рл}$) и учитывает количество узлов (разветвлений), включающих четыре соседние связи (изопентановая структура).

Таким образом, модель молекулярной связанности позволяет соотносить как общие, так и локальные структурные свойства молекул.

Показано, что в ряду индексов цепи ${}^1\chi, {}^2\chi, {}^3\chi_{р}, {}^4\chi_{р}$ дескриптор, ${}^3\chi_{р}$ является главным структурным фактором, оказывающим существенное влияние на уровень корреляции уравнений регрессии. Характеризуя гибкость алкильных цепей, индекс ${}^3\chi_{р}$ объясняет, в частности, большую часть изменений индексов хроматографического удерживания.

Важную роль играет дескриптор V_w (Ван-дер-ваальсов объем). Его дополнительный ввод в структурные модели молекулярной связанности или вместо отдельных индексов связанности способствует повышению значимости уравнений и снижению величины стандартного отклонения. Рассчитанные значения V_w и индексы связанности цепи и разветвления различных порядков применяются для построения структурных моделей органических соединений и определения их корреляционной связи с индексами удерживания путем решения всех возможных линейно-регрессионных уравнений вида:

$$Y = f (V_w , {}^1\chi \dots {}^n\chi_m)$$

Весьма заметна роль линейно-узлового индекса ${}^4\chi_{рс}$. Характеризуя тонко-структурные особенности строения, дескриптор ${}^4\chi_{рс}$ входит практически во все уравнения, описывающие корреляции ИС с физико-химическими параметрами.

Из литературы известно, что эти ИС коррелируют с самыми разнообразными физико-химическими параметрами, такими, например, как теплота образования, молярная рефракция, поляризуемость, хроматографические характеристики и пр. В литературе показано, что ИС могут служить мерой внутримолекулярных стерических взаимодействий, особенно для алифатических заместителей.

Одним из важнейших приложений индексов связанности является их использование для установления количественных соотношений между структурой и величинами, характеризующими биологическую активность соединений (неспецифическая наркотическая активность, ингибирование ферментов, бактерицидное действие, LD_{50} и др.).

Достаточно широко в последнее время используются так называемые сорбционно-структурные корреляции, описывающие зависимости между хро-

матографическим удерживанием (сорбционными свойствами) и структурными (топологическими) характеристиками анализируемых веществ. Такие корреляции применяют для прогнозирования хроматографического удерживания веществ (например, трудно хроматографируемых высококипящих изомеров), прогнозирования свойств гипотетических молекул, поиска новых соединений с заданными свойствами и т.д. Наиболее часто для установления таких корреляций используют индекс связанности.

Использование ТИ для оценки физико-химических свойств соединений

При оценке ряда физико-химических свойств соединений, например, температуры кипения или температуры плавления, существенное значение имеют конкретные оценки размеров молекул и степени их разветвленности, так как именно от этих параметров зависит в значительной степени энергия межмолекулярных взаимодействий. В топологическом приближении для решения этой задачи используют в основном индексы, характеризующие размеры и форму графа. Часть из них непосредственно связана с матрицей расстояний D . Такой подход впервые был реализован Винером, который предложил использовать для сравнения углеродных скелетов насыщенных углеводородов два топологических индекса – W и P , определенных в терминах расстояний между вершинами соответствующих молекулярных графов. Число W , называемое обычно индексом Винера, выражается через элементы дистанционной матрицы D по формуле: $W = \frac{1}{2} \sum_{i,j} d_{ij}$. Индекс W

может быть также записан в виде $W = \sum_j g_j \cdot j$, где g_j - число пар вершин, находящихся друг от друга на расстоянии j . Для того, чтобы учесть наличие гетероатомов и кратных связей в молекуле, вводится матрица расстояний D , построенная с учетом весов, приписываемых вершинам и ребрам молекулярного графа. В таком случае элементы матрицы D определяют по формулам:

$$D_{ij} = 1 - \frac{6}{Z_i} \quad \text{и} \quad D_{ij} = \sum \frac{1}{b_{kl}} \cdot \frac{36}{z_k \cdot z_l},$$

где Z_j – заряд j -ого атома, b_{kl} – кратность связи (k, l), равная 1, 2, 3 или 3/2, в зависимости от того, является ли связь соответственно простой, двойной, тройной или ароматической. Суммирование распространяют на все связи (k, l), образующие кратчайший путь между i -ой и j -ой вершинами.

Индексы, основанные на инвариантном множестве степеней вершин, легко вычислять как для простых, так и для взвешенных графов. Поэтому они относятся к числу наиболее популярных топологических индексов в теоретиче-

ской химии. Если для построения индексов в качестве подграфов графа G выбраны ребра, то используя различные симметричные функции от двух переменных $F(x_1, x_2)$, можно определить ряд ТИ, зависящих от степеней вершин графа G :

G : $f = \sum (v_i - v_j - 2)$ – индекс Платта, $M_2 = \sum v_i \cdot v_j$ – индекс Гутмана,

$\chi = \sum (v_i \cdot v_j)^{-1/2}$ – индекс молекулярной связанности, или индекс Рэндича.

Используют также индекс молекулярной связанности порядка h , определяемый по формуле:

$${}^h\chi = \sum (v_{i1}, v_{i2}, \dots, v_{ih+1})^{-1/2} \quad (\text{по путям длины } h),$$

где $h \geq 1$ – фиксированное целое число, v_{ik} – степень вершины i_k на данном пути длины h , $k=1, \dots, h$.

Предложено также определение топологического индекса B_c , связанное с понятием центра графа. К таким центрическим индексам относят центрический индекс Балабана, определяемый как: $B_c = \sum r_i^2$, где r_i – расстояние от i -ой вершины до центра.

Один из способов сравнения молекулярных структур в топологическом приближении состоит в подсчете для каждой из них числа структурных фрагментов определенного типа и последующем сравнении этих величин. Так как структурные фрагменты одного типа могут различаться между собой по какому-либо признаку, их можно классифицировать с различной степенью детализации. В настоящее время в структурной химии используют ряд топологических индексов, получивших название информационных и основанных на применении теории информации. К их числу можно отнести хроматические индексы, индексы электропии, информационные разновидности индексов Хосойи, Рэндича, Балабана.

Таким образом, к настоящему времени определено большое число разнообразных ТИ, многие из которых находят применение при анализе структуры молекул и построении корреляций между структурой и свойством. Некоторые из ТИ связаны между собой простыми соотношениями с достаточно высоким коэффициентом корреляции, в таком случае индексы этой группы несут примерно одинаковую информацию о структуре соединений определенных классов. Были получены подробные соотношения между некоторыми ТИ, определяемыми в методе Хюккеля, и индексами, непосредственного отношения к этому методу не имеющими (такими, например, как индексы связанности, числа Винера и другие). Для сопряженных углеводородов, например, индекс связан-

ности ${}^1\chi$ первого порядка, определяемый по формуле ${}^1\chi = \sum_{i,j} (\delta_i \cdot \delta_j)^{-1/2}$ достаточно хорошо коррелирует с полной π - электронная энергией.

Традиционные способы оценки физико-химических свойств соединений основаны на аддитивных схемах. В классической теории химического строения постулируется, что некоторые физико-химические свойства органических соединений (например, теплота образования, температура кипения, поляризуемость, дипольный момент и другие) могут быть вычислены как сумма вкладов $\epsilon_i, \epsilon_{ij}, \epsilon_{kl}$, отвечающих атомам i , связям (i,j) и парам валентно несвязанных атомов (k, l) соответствующих молекул. Эта аддитивная схема расчета свойств молекул допускает ряд модификаций. Так, одна из разновидностей аддитивных схем сводится к представлению численного значения изучаемого свойства химического соединения в виде линейной функции $y = \sum a_i \cdot k_i$, определенной на соответствующем молекулярном графе. Здесь i - характеризует фрагменты молекулярного графа определенного типа, которые могут иметь сложную структуру, k_i – топологический индекс, равный числу таких фрагментов, a_i – вклад, который вносит каждый из фрагментов i -го класса в данное свойство и не зависящий от выбора класса соединений.

Следует отметить, что сам факт существования корреляционной зависимости физико-химических свойств и топологических параметров молекул представляется вполне логичным. Например, индексы удерживания по определению пропорциональны свободным энергиям взаимодействия сорбатов с сорбентами или неподвижными фазами. Следовательно, физико-химическая природа выявленной корреляции может быть сформулирована как пропорциональность структурных вкладов сорбционного взаимодействия соответствующим изменениям внутренних энергий молекул органических соединений при вариациях их структуры. Именно с одинаковой энергетической природой обеих сопоставляемых величин связан исключительно простой линейный характер зависимости $I = a \cdot \chi + b$.

Винер предложил описывать такие свойства изомеров алканов как молярную рефракцию, теплоту образования и другие с помощью уравнений вида $y - y_0 = a(W - W_0) + b(P - P_0)$, где y – числовое значение данного свойства; W - индекс Винера для МГ, соответствующего углеродному скелету молекулы; P - число полярности, a, b - константы, зависящие от исследуемого свойства

Среди других ТИ, используемых для оценки физико-химических свойств соединений, наибольшей популярностью пользуется индекс связанности ${}^1\chi$

и его разновидности ${}^2\chi$, ${}^{1v}\chi$, ${}^{2v}\chi$. Топологический индекс χ оказался полезным в хроматографии для оценки индексов удерживания. Однако иногда при этом наряду с χ приходится использовать и другие топологические характеристики молекулы, описывающие ее размеры, например, число путей длины или число шестичленных циклов в ароматических углеводородах.

Индекс Моровитца I_{MOR} коррелирует с энтропией и теплотой образования органических соединений различных классов. Уравнение вида:

$$y = a + bn + cI_D^W + dI_D^W n,$$

где a , b , c , d – константы, использованы для оценки молекулярного объема и таких свойств углеводородов, как теплоты образования и окисления, парообразования, температуры кипения.

Среди ТИ, используемых при исследовании биологической активности, наиболее широкое распространение получили индексы молекулярной связанности, в частности индекс Рэндича χ .

Для ряда анестетиков (алифатические спирты, кетоны, эфиры, амины, фенолы, гетероциклы, углеводороды) найдена корреляция между $\lg(\text{MBC})$ и индексом χ : $\lg(\text{MBC}) = 3.55 - 0.762\chi$ ($r=0.983$; $s=0.390$; $N=36$), где MBC – минимальная блокирующая концентрация анестетиков для изолированных мускульных и нервных волокон. Существуют корреляционные соотношения между χ и параметрами таких видов биоактивности, как неспецифическая наркотическая активность, ингибирование ферментов, бактерицидное действие, токсичность, анестезирующая активность. Обсуждена возможность оценки вкуса и запаха веществ с помощью топологических характеристик структуры молекул. В литературе предложено характеризовать форму молекул параметрами χ/n , при этом предполагается, что вещества, имеющие одинаковый запах, должны иметь близкие значения χ/n . Проанализирована связь между строением и токсичностью эфиров, характеризуемой параметром LD_{50} .

Таким образом, роль ТИ в решении проблемы «структура – свойство» неуклонно возрастает. Введение новых типов ТИ, обладающих значительной дискриминирующей способностью для изомерных структур, позволяет практически однозначно описывать большинство органических соединений. Большой прогресс достигнут в топологическом представлении молекул с гетероатомами. Разработаны и внедряются в практику разновидности ТИ, позволяющие с максимальной точностью оценивать структуру соединений и прогнозировать свойства веществ на основании корреляционного топологического анализа. В то же время в связи с большим разнообразием используемых ТИ для конкретного применения их, как

правило, выбирают методом проб и ошибок, поэтому сравнение и оценка их дискриминирующей и прогностической способности является весьма важной проблемой, решение которой позволит унифицировать систему ТИ при описании структуры и свойств молекул, а также корреляций между ними.

Алгоритм расчета топологических индексов

Индекс связанности. Это аддитивно-конструктивный параметр, характеризующий молекулярную способность к образованию связей и описывающий топологические свойства органических молекул. Индекс связанности представляет собой математически закодированную информацию о числе атомов в молекуле, их связях между собой, о степени разветвления молекулы.

В топологическом понимании структурную формулу органического соединения составляет граф (углеродный скелет без атомов водорода), вершины которого соответствуют атомам, а ребра - связям между атомами. Для углеводородов существует прямое соотношение между валентностью атома углерода (δ_i) и числом атомов водорода (h_i), наличие которых подразумевается при данном атоме i : $\delta_i = 4 - h_i$

Таким образом, для метильной группы (первичный атом водорода) $\delta = 4 - 3 = 1$, для метиленовой группы (вторичный атом водорода) $\delta = 4 - 2 = 2$, для третичного атома углерода $\delta = 4 - 1 = 3$, четвертичного атома углерода $\delta = 4$.

Для молекул, содержащих гетероатомы, при расчете ИС используется понятие о валентной молекулярной связанности, которая включает учет валентных электронов, занимающих σ и π -орбитали, и орбитали неподеленных пар каждого атома, за исключением связей с водородом. Валентная величина δ^v равна:

$$\delta_i^v = Z_i^v - h_i,$$

где Z_i^v - число электронов валентной оболочки, h_i - число атомов водорода, связанных с атомом i . Тогда для атома кислорода в гидроксильной группе значение $\delta_i^v = 5$, для азота пиридина и хинолина $\delta_i^v = 5$, для азота первичных аминов $\delta_i^v = 3$, вторичных $\delta_i^v = 4$ и т.д. Таким образом, величина δ_i^v содержит информацию, как об объемных, так и об электронных характеристиках молекулы. Этот вывод основан на том, что кодирование σ - и π -орбиталей и орбиталей неподеленных пар с помощью величины δ^v предполагает возможность описания пространства, занятого этими орбиталями. Орбитальное пространство, в свою очередь, определяет размеры атома в соответствующем валентном состоянии. Электронные свойства зависят от числа σ - и π -электронов и не-

поделённых электронных пар, которые определяют величину δ^v . Величины δ^v для отдельных групп с гетероатомами приведены в табл. 1.

Таблица 1

Значения валентных величин для гетероатомов и отдельных групп атомов

Группа	δ^v	Группа	δ^v
--NH ₂	3	-ОН	5
>NH	4	О (в фуране)	6
>N-	5	>СО	6
N (в пиридине, хиолине)	5	F	-20
N (в нитрогруппе)	6	Cl	0,690
-CN	5	Br	0,254
-O-	6	I	0,085

Величина χ может быть рассчитана для различных уровней связанности атомов молекулы между собой.

Индекс связанности нулевого порядка, отражающий лишь валентное состояние атомов в рассматриваемой молекуле, определяется по формуле:

$${}^0\chi = \sum_{g=1}^k (\delta_i)^{-\frac{1}{2}} \quad (1)$$

Наиболее распространенным является индекс связанности первого порядка, вычисляемый согласно следующему уравнению:

$${}^1\chi = \sum_{g=1}^k (\delta_i \delta_j)^{-\frac{1}{2}} \quad (2)$$

где δ^- – разность между числом валентных электронов и количеством атомов водорода, связанных с данным атомом; i и j – номера атомов молекулы, формально связанных рассматриваемой связью; δ_i , δ_j – значения δ для двух атомов связи; k – общее число связей в молекуле].

Аналогичным образом рассчитываются индексы связанности второго порядка:

$${}^2\chi = \sum_{g=1}^k (\delta_i \delta_j \delta_k)^{-\frac{1}{2}} \quad (3)$$

третьего и более высоких порядков, учитывающих количество двух, трех, четырех и более соседних связей:

$${}^3\chi = \sum_{g=1}^k (\delta_i \delta_j \delta_k \delta_l)^{-\frac{1}{2}} \quad (4)$$

Индексы связанности третьего и последующих порядков могут быть представлены в виде индексов цепи (линейный индекс) χ_p и разветвления (узловой индекс) χ_c .

Линейный индекс третьего порядка равен

$${}^3\chi_p = \sum_{g=1}^M (\delta_i \delta_j \delta_k \delta_l)^{-1/2} \quad (5)$$

где i, j, k, l – атомы молекулы, образующие цепь (линию) из трех соседних связей; M – число таких цепочек из трех соседних связей в молекуле.

Узловой индекс третьего порядка рассчитывают по формуле

$${}^3\chi_c = \sum_{g=1}^N (\delta_i \delta_j \delta_k \delta_l)^{-1/2} \quad (6)$$

где i, j, k, l – атомы трех соседних связей, образующие узел (структуру типа треноги); N – число таких структур в молекуле.

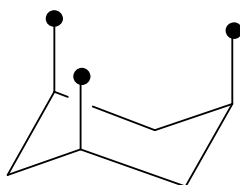
Аналогичный индекс четвертого порядка называется линейно-узловым индексом (в отличие от линейного индекса ${}^4\chi_p$) и учитывает количество узлов (разветвлений), включающих четыре соседние связи (изопентановая структура).

$${}^4\chi_{pC} = \sum_{g=1}^P (\delta_i \delta_j \delta_k \delta_l \delta_m)^{-1/2} \quad (7)$$

где i, j, k, l, m – атомы четырех соседних связей, включающие узел типа изопентана; P – число таких структур в молекуле.

Индекс связанности первого порядка учитывает парные взаимодействия соседних атомов в молекуле. К его недостаткам можно отнести невозможность различения мета- и пара-изомеров ароматических соединений. Индексы связанности высоких порядков, учитывающие не только ближайшие парные, но и отдаленные взаимодействия атомов в молекуле, обеспечивают более полное описание структурных особенностей соединений. Таким образом, модель молекулярной связанности позволяет соотносить как общие, так и локальные структурные свойства молекул.

Рассмотрим пример расчета индексов связанности для 1,3,5-триметилциклогексана:



Величины δ_i равны: $\delta_1 = 3$, $\delta_2 = 2$, $\delta_3 = 3$, $\delta_4 = 2$, $\delta_5 = 3$, $\delta_6 = 2$, $\delta_7 = 1$, $\delta_8 = 1$, $\delta_9 = 1$.

Согласно уравнению (1) ${}^0\chi$ для 1,3,5-триметилциклогексана

$${}^0\chi = \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{1}} = 6,853$$

Согласно уравнению (2) индекс ${}^1\chi$ для 1,3,5-триметилциклогексана

$${}^1\chi = \frac{1}{\sqrt{1 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 1}} + \frac{1}{\sqrt{3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 1}} + \frac{1}{\sqrt{3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3}} = 4,182$$

Согласно уравнению (3) индекс ${}^2\chi$ 1,3,5-триметилциклогексана

$${}^2\chi = \frac{1}{\sqrt{1 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 1}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 1}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 1}} = 4,002$$

Как уже было сказано, индексы связанности третьего порядка делятся на два вида: индекс цепи (линейный индекс) и индекс разветвления (узловой).

Согласно уравнению (4) индекс цепи χ_p 1,3,5-триметилциклогексана

$${}^3\chi = \frac{1}{\sqrt{3 \cdot 2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{3 \cdot 2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{2 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{3 \cdot 2 \cdot 3 \cdot 2}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2 \cdot 3}} + \frac{1}{\sqrt{1 \cdot 3 \cdot 2 \cdot 3}} = 2,414$$

Аналогичным образом рассчитываются индексы связанности четвертого и последующих порядков.

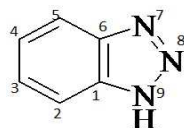
Ниже приведен пример расчета индексов связанности 0 – 2 порядков для молекулы хинолина:

$${}^0\chi = \frac{1}{\sqrt{5}} + 2 \cdot \frac{1}{\sqrt{4}} + 7 \cdot \frac{1}{\sqrt{3}} = 5,488$$

$${}^1\chi = \frac{1}{\sqrt{5 \cdot 3}} + \frac{1}{\sqrt{5 \cdot 4}} + 5 \cdot \frac{1}{\sqrt{3 \cdot 3}} + 3 \cdot \frac{1}{\sqrt{3 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 4}} = 3,258$$

$${}^2\chi = \frac{1}{\sqrt{5 \cdot 3 \cdot 3}} + 2 \cdot \frac{1}{\sqrt{5 \cdot 4 \cdot 3}} + \frac{1}{\sqrt{5 \cdot 4 \cdot 4}} + 3 \cdot \frac{1}{\sqrt{3 \cdot 3 \cdot 3}} + 4 \cdot \frac{1}{\sqrt{3 \cdot 3 \cdot 4}} + 3 \cdot \frac{1}{\sqrt{3 \cdot 4 \cdot 4}} = 2,189$$

Рассмотрим расчет индексов связанности различных порядков для бензотриазола.



Величины δ_i равны:

$\delta_1=4-0=4$	$\delta_6=4-0=4$
$\delta_2=4-1=3$	$\delta_7=5-0=5$
$\delta_3=4-1=3$	$\delta_8=5-0=5$
$\delta_4=4-1=3$	$\delta_9=5-1=4$
$\delta_5=4-1=3$	

Индекс связанности нулевого порядка (${}^0\chi$):

$${}^0\chi = \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{4}} = 4,704$$

Индекс связанности первого порядка (${}^1\chi$):

$${}^1\chi = \frac{1}{\sqrt{4 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 4}} = 2,474$$

Индекс связанности второго порядка (${}^2\chi$):

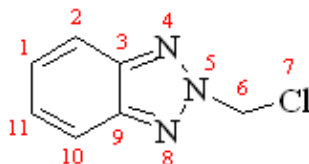
$${}^2\chi = \frac{1}{\sqrt{4 \cdot 3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3 \cdot 4}} + \frac{1}{\sqrt{3 \cdot 3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 4 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 4 \cdot 3}} +$$

$$+ \frac{1}{\sqrt{3 \cdot 4 \cdot 5}} + \frac{1}{\sqrt{4 \cdot 4 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 4 \cdot 5}} + \frac{1}{\sqrt{4 \cdot 4 \cdot 5}} + \frac{1}{\sqrt{4 \cdot 5 \cdot 5}} + \frac{1}{\sqrt{4 \cdot 5 \cdot 5}} = 1,829$$

Индексы связанности более высоких порядков, учитывающие количество двух, трёх, четырёх и более соседних связей, рассчитываются подобным же образом.

Аналогично индексы связанности рассчитываются для производных бензотриазола, содержащих гетероатомы (кислород, серу, азот и т.д.).

Рассмотрим пример расчета индексов связанности для 2-(хлорметил)-2*H*-бензотриазола.



Величины δ_i равны: $\delta_1 = 3$, $\delta_2 = 3$, $\delta_3 = 4$, $\delta_4 = 5$, $\delta_5 = 5$, $\delta_6 = 2$, $\delta_7 = 0,69$, $\delta_8 = 5$, $\delta_9 = 4$, $\delta_{10}=3$, $\delta_{11}=3$.

Согласно уравнению (1) ${}^0\chi$ для 2-(хлорметил)-2*H*-бензотриазола:

$${}^0\chi = \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{0,69}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} = 6,562$$

Согласно уравнению (2) индекс ${}^1\chi$ для 2-(хлорметил)-2*H*-бензотриазола:

$${}^1\chi = \frac{1}{\sqrt{3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 5}} + \frac{1}{\sqrt{4 \cdot 4}} + \frac{1}{\sqrt{5 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 0,69}} + \frac{1}{\sqrt{5 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3}} = 3,842$$

Согласно уравнению (3) индекс ${}^2\chi$ для 2-(хлорметил)-2*H*-бензотриазола:

$${}^2\chi = \frac{1}{\sqrt{3 \cdot 3 \cdot 4}} + \frac{1}{\sqrt{3 \cdot 3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 4 \cdot 5}} + \frac{1}{\sqrt{3 \cdot 4 \cdot 4}} + \frac{1}{\sqrt{4 \cdot 5 \cdot 5}} + \frac{1}{\sqrt{4 \cdot 4 \cdot 3}} + \frac{1}{\sqrt{4 \cdot 4 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 5 \cdot 2}} + \frac{1}{\sqrt{5 \cdot 5 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 4 \cdot 4}} + \frac{1}{\sqrt{5 \cdot 2 \cdot 0,69}} + \frac{1}{\sqrt{5 \cdot 5 \cdot 4}} + \frac{1}{\sqrt{2 \cdot 5 \cdot 5}} + \frac{1}{\sqrt{5 \cdot 4 \cdot 3}} + \frac{1}{\sqrt{4 \cdot 3 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 3 \cdot 3}} = 2,442$$

Аналогичным образом рассчитываются индексы связанности третьего и последующих порядков.

Индекс Винера (W) является инвариантами топологической матрицы расстояний МГ (D).

Алгоритм нахождения индекса Винера включает в себя построение топологической матрицы расстояний $D = (d_{ij})$ в соответствии с аксиомами метрики для связного r -вершинного графа G . Матрица расстояний представляет собой квадратную симметричную $r \times r$ матрицу с нулевыми диагональными элементами.

Составными элементами этой матрицы являются кратчайшие расстояния от i вершины к j вершине (d_{ij}). Вид матрицы расстояний в общем виде зависит от способа нумерации вершин молекулярного графа, однако сумма матричных элементов не зависит от нумерации вершин. Индексом или числом Винера $W(G)$ называют полусумму элементов матрицы расстояний или сумму ее наддиагональных элементов, поскольку матрицы расстояний симметричны относительно своих нулевых диагоналей:

$$W = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N d_{ij}$$

где d_{ij} - наикратчайшее расстояние между вершинами i и j ; N - число вершин молекулярного графа.

Недиагональные элементы матрицы определяют по формуле :

$$d_{ij} = \sum \frac{1}{b} \cdot \frac{36}{z_i \cdot z_j}$$

где z_i, z_j - заряд ядра (числа всех электронов) атомов i и j , соединенных данной связью; b - величина, характеризующая порядок (кратность) связи (1, 2, 3, или 3/2 в зависимости от того является ли связь одинарной, двойной, тройной или ароматической).

Диагональные элементы матрицы расстояний определяют следующим образом:

$$d_{ii} = 1 - \frac{6}{z_i},$$

Итак, индекс Винера для молекулы 2-(2'-тиенил)хинолина рассчитывают следующим образом:

- 1) Представляют молекулу в виде молекулярного графа со «стертыми» атомами водорода. Нумерацию вершин можно принять произвольной.
- 2) Недиагональные элементы определяют из соотношений:

$$d_{(C-N)} = \frac{1}{1,5} \cdot \frac{36}{6 \cdot 7} = 0,57$$

$$d_{(C-C)} = \frac{1}{1,5} \cdot \frac{36}{6 \cdot 6} = 0,66$$

$$d_{(C2-C3)} = \frac{1}{1} \cdot \frac{36}{6} = 1$$

$$d_{(C-S)} = \frac{1}{1,5} \cdot \frac{36}{6 \cdot 16} = 0,25$$

Диагональные элементы:

$$d_{(N-N)} = 1 - \frac{6}{7} = 0,143$$

$$d_{(S-S)} = 1 - \frac{6}{16} = 0,625$$

$$d_{(C-C)} = 1 - \frac{6}{6} = 0$$

3) Матрица расстояний имеет вид:

D =

0,07	0,57	1,57	1,82	2,07	2,73	2,23	1,23	1,89	1,23	1,89	2,55	1,89	1,23	0,07
	0	1	1,25	1,5	2,16	1,66	0,66	1,32	1,8	2,46	3,12	2,46	1,8	1,14
		0	0,25	0,5	1,16	0,66	1,66	3,32	2,8	3,46	4,12	3,46	2,8	2,14
			0,31	0,25	0,91	0,91	1,91	2,57	3,05	3,71	4,37	3,71	3,05	2,39
				0	0,66	1,16	2,16	2,82	3,3	3,96	4,62	3,96	3,3	2,64
					0	0,66	2,82	3,48	3,96	4,62	5,28	4,62	3,96	3,3
						0	2,32	2,98	3,46	4,12	4,78	4,12	3,46	2,8
							0	0,66	1,32	1,98	2,64	3,12	2,46	1,8
								0	0,66	1,32	1,98	2,64	1,98	1,32
									0	0,66	1,32	1,98	1,32	0,66
										0	0,66	1,39	1,98	1,32
											0	0,66	1,32	1,98
												0	0,66	1,32
													0	0,66
														0

4) Суммируя наддиагональные элементы, получаем $W^* = 229,36$.

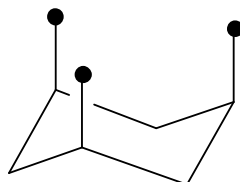
В качестве элементов матрицы расстояний можно использовать реальные длины связей, выраженные, например, в ангстремах. Получаемые таким образом индексы Винера известны как трехмерные или модифицированные. В данном примере в качестве элементов матрицы использованы величины длин связей, рассчитанные с помощью программы HyperChem полуэмпирическим квантово-химическим методом PM3 с полной оптимизацией геометрии молекулы. Модифицированная матрица расстояний для молекулы 2-(2'-тиенил)хинолина имеет вид:

0	1,35	2,80	4,56	6,29	5,62	4,20	2,78	4,16	2,81	4,22	5,60	4,20	2,81	1,38
	0	1,46	3,20	4,93	4,26	2,84	1,43	2,81	4,16	5,64	6,95	5,54	4,16	2,73
		0	1,76	3,48	2,81	1,39	2,88	4,25	5,61	7,02	8,40	6,99	5,61	4,18
			0	1,72	3,10	3,14	4,63	6,01	7,36	8,78	10,15	8,74	7,36	5,94
				0	1,37	2,80	6,36	7,73	9,08	10,56	11,87	10,46	9,08	7,66
					0	1,43	5,69	7,07	8,42	9,84	11,21	9,80	8,42	6,99
						0	4,27	5,64	6,99	8,41	9,78	8,37	6,99	5,57
							0	1,38	2,80	4,21	5,59	6,97	5,59	4,16
								0	1,42	2,86	4,22	5,63	4,27	2,84
									0	1,42	2,80	4,21	2,85	1,42
										0	1,38	2,79	4,17	2,84
											0	1,41	2,79	4,21
												0	1,32	2,80
													0	1,42
														0

Модифицированный индекс Винера $W = 515,751$.

$$W/W^* = 515,751 / 229,36 = 2,248$$

Как видно из полученных данных, модифицированный индекс более, чем в два раза превышает значение традиционного индекса Винера. Следовательно, алгоритм расчета модифицированного индекса Винера позволяет увеличить дискриминирующую способность получаемых значений индекса.



Рассмотрим пример расчета индекса Винера 1,3,5-триметилциклогексана

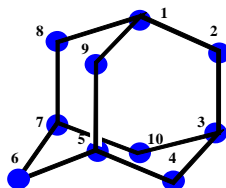
Составим матрицу расстояний D для данной молекулы.

$$D = \begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 3 & 2 & 1 & 1 & 3 \\ 1 & 0 & 1 & 2 & 3 & 4 & 3 & 2 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 & 2 & 3 & 3 & 1 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 & 2 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 1 & 3 \\ 3 & 4 & 3 & 2 & 1 & 0 & 1 & 2 & 2 & 2 \\ 2 & 3 & 2 & 3 & 2 & 1 & 0 & 1 & 3 & 1 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 & 0 & 2 & 2 \\ 1 & 2 & 3 & 2 & 1 & 2 & 3 & 2 & 0 & 4 \\ 3 & 2 & 1 & 2 & 3 & 2 & 1 & 2 & 4 & 0 \end{pmatrix}$$

Согласно уравнению (8), индекс W 1,3,5-триметилциклогексана

$$W = (1/2) \cdot 168 = 84$$

Рассмотрим пример расчета индекса Винера молекулы адамантана:



При этом матрица расстояний для молекулы адамантана имеет вид:

$$D = \begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 3 & 2 & 1 & 1 & 3 \\ 1 & 0 & 1 & 2 & 3 & 4 & 3 & 2 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 & 2 & 3 & 3 & 1 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 & 2 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 1 & 3 \\ 3 & 4 & 3 & 2 & 1 & 0 & 1 & 2 & 2 & 2 \\ 2 & 3 & 2 & 3 & 2 & 1 & 0 & 1 & 3 & 1 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 & 0 & 2 & 2 \\ 1 & 2 & 3 & 2 & 1 & 2 & 3 & 2 & 0 & 4 \\ 3 & 2 & 1 & 2 & 3 & 2 & 1 & 2 & 4 & 0 \end{pmatrix}$$

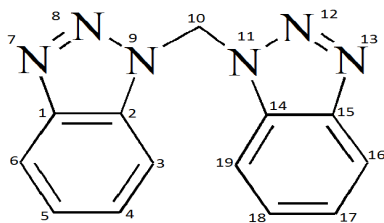
Согласно уравнению (8), индекс W адамантана $W = (1/2) \cdot 168 = 84$

При анализе производных адамантана учитываются особенности структуры соответствующих молекулярных графов, связанные с наличием гетероатомов и кратных связей. Для таких соединений недиагональные элементы матрицы расстояний определяют по формуле:

$$d_{ij} = \sum \frac{1}{b_r} \frac{36}{z_i \cdot z_j},$$

при этом суммирование проводится по r ($r = 1, 2, 3 \dots$) взвешенным и невзвешенным связям, а параметр b определяет кратность связи и равен 1, 2, 3 или $3/2$ для одинарной, двойной, тройной или ароматической связи соответственно, z_i, z_j - заряд ядра i -го или j -го атомов, соединенных такой связью.

Рассмотрим пример расчета индекса Винера для 1,1'-метилдиилбис(1*H*-бензотриазола). Молекулу бензотриазола представляют в виде молекулярного графа, который соответствует его структурной формуле со «стертыми» атомами водорода.



Расчет диагонального элемента D_{i-i} проводят по формуле (9):

$$D_{i-i} = 1 - 6 / Z_i,$$

где Z_i – порядковый номер элемента в периодической системе. В данном соединении диагональным элементом является углерод, поэтому:

$$D_{C-C} = 1 - 6/6 = 0$$

Атом азота в данном соединении также является диагональным элементом. Рассчитаем значения для него:

$$D_{N-N} = 1 - 6/7 = 0,14$$

Для гетероатомов при построения матрицы расстояний берем половину D_{N-N} , т.е.:

$$0,14 / 2 = 0,07$$

Далее вычисляем значение недиагонального элемента матрицы расстояний $d_{i,j}$, применяя формулу (8):

$$d_{ij} = \sum \frac{1}{b_r} \cdot \frac{36}{z_i \cdot z_j}$$

b – кратность связи в молекуле. В нашем случае кратность равна 1.

Z_i и Z_j – атомы между которыми осуществляется связь.

$$d_{C-C} = \frac{1}{1} \cdot \frac{36}{6 \cdot 6} = 1$$

Найдем значение d_{C-C} для атома углерода, находящемся при двойной связи, учитывая, что параметр b определяется кратностью связи и равен 1, 2, 3 или 3/2 для одинарной, двойной, тройной или ароматической связи соответственно, z_i, z_j – заряд ядра i -го или j -го атомов, соединенных такой связью.

$$d_{C-C} = \frac{3}{2} \cdot \frac{36}{6 \cdot 6} = 1,5$$

Найдем значение d_{C-N} , т.е. расстояние между атомами углерода и азота:

$$d_{C-N} = \frac{1}{1} \cdot \frac{36}{6 \cdot 7} = 0,86$$

Затем определим значение d_{C-N} , т.е. расстояние между атомами углерода и азота при двойной связи:

$$d_{C-N} = \frac{1}{2} \cdot \frac{36}{6 \cdot 7} = 0,43$$

Аналогичным образом рассчитываем значения d_{N-N} и d_{N-N} . Получаем соответственно 0,73 и 0,37.

С использованием выше посчитанных значений построим матрицу расстояний Винера для молекулы 1,1'-метилдиилбис(1*H*-бензотриазола). Главная диагональ нашей матрицы D является нулевой т.к. по определению $D_{C-C} = 0$.

Практически произвольно пронумеруем атомы от 1 до 19 и записываем полученные значения в таблицу.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1	0	1,5	3	4,5	3	1,5	0,86	1,23	1,86	2,22	3,08	3,81	4,18	3,94	5,44	6,94	8,44	6,94	5,44	
2		0	1,5	3	4,5	3	2,63	1,59	0,86	1,72	2,58	3,31	3,68	3,44	4,94	6,44	7,94	6,44	4,94	
3			0	1,5	3	4,5	3,86	3,09	2,36	3,22	4,08	4,81	5,18	5,67	7,17	8,67	10,17	8,67	7,17	
4				0	1,5	3	5,36	4,59	3,86	4,72	5,58	6,31	6,68	7,17	8,67	10,17	11,67	10,17	8,67	
5					0	1,5	3,86	4,23	5,36	6,22	7,08	7,81	8,18	7,94	9,44	10,94	12,44	10,94	9,44	
6						0	2,36	2,73	3,86	4,72	5,58	6,31	6,68	6,44	7,94	9,44	10,94	9,44	7,94	
7							0,07	0,37	1,1	1,96	2,82	3,55	3,92	3,68	5,18	6,68	8,18	6,68	5,18	
8								0,07	0,73	1,59	2,45	3,18	3,55	3,31	4,81	6,31	7,81	6,31	4,81	
9									0,07	0,86	1,72	2,45	2,82	2,58	4,08	5,58	7,08	5,58	4,08	
10										0	0,86	1,53	1,96	1,72	3,22	4,72	6,22	4,72	3,22	
11											0,07	0,73	1,1	0,86	2,36	3,86	5,36	3,86	2,36	
12												0,07	0,37	1,59	3,09	4,59	6,09	4,59	3,09	
13													0,07	2,36	0,86	2,36	3,86	5,36	3,86	
14														0	1,5	3	4,5	3	1,5	
15															0	1,5	3	4,5	3	
16																0	1,5	3	4,5	
17																	0	1,5	3	
18																		0	1,5	
19																				0

Суммируя столбцы и строки по всем элементам нашей матрицы, получаем число Винера для 1,1'-метилдиилбис(1*H*-бензотриазола), применяя формулу (7).

$$W = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N d_{ij}, \quad (7)$$

Получаем значение $W=764,03$.

Очевидно, что чем сложнее строение молекулы, т.е. чем больше число атомов, входящих в ее состав, тем больше значение индекса Винера.

В отличие от индексов связанности, индекс Винера может различать изомеры, так как учитывает порядок связей в молекуле и характер их взаимодействия.

Библиографический список

1. Химические приложения топологии и теории графов / под ред. Р. Кинга. М.: Мир, 1987. 560 с.
2. Минкин В.И., Симкин Б.Я., Миняев Р.М. Теория строения молекул. Р.-Д.: «Феникс». 1997. 556 с.
3. Randić M. In search of structural invariants // J. Math. Chem. 1992. V. 9. P. 97–146.
4. Л.А. Грибов, В.А. Дементьев, И.В. Михайлов. Матрицы смежности и графы химических превращений // Журн. структ. химии. 2008. Т.49. №2. С.211-214.
5. Balaban A.T. Visual Chemistry: Three-Dimensional Perception of Chemical Structures // J. Sci. Edu. & Technol. 1999. V. 8. № 4. P.251-255.
6. Станкевич М.И., Станкевич И.В., Зефирова Н.С. Топологические индексы в органической химии // Успехи химии. 1988. Т. 57. №3. С. 337–366.
7. Galvez J., Garcia-omenech R., Gregorio-lapont C. Indices of differences of path lengths: novel topological descriptors derived from electronic interferences in graphs // J. Comp.- Aid. Mol. Des. 2000. № 14. P. 679–687.
8. Francisco Torrens. A new topological index to elucidate apolar hydrocarbons // J. Comp.- Aid. Mol. Des. 2001. V. 15. P. 709–719.
9. Gutman I., Vukicevica D., Zerovnik J. // Croatica Chemica Acta. 2004. V. 77. № 1,2. P. 103-109.
10. Estrada E. Physicochemical Interpretation of Molecular Connectivity Indices // J. Phys. Chem. 2002. V. 106. P. 9085-9091.
11. Скворцова М.И., Станкевич И.В., Палюлин В.А., Зефирова Н.С. Концепция молекулярного подобия и её использование для прогнозирования свойств химических соединений // Успехи химии. 2006. Т.75. №11. С.1074-1093.
12. Стьюпер Э., Брюггер У., Джурс П. Машинный анализ связи химической структуры и биологической активности. М.: Мир, 1982. 321 с.
13. Станкевич И.В. Применение теории графов в химии. Новосибирск: Наука, 1988. 69 с.
14. Красных Е.Л. Прогнозирование энтальпии испарения на основе индексов Рэндиша. I. Алканы // Журн. структ. химии. 2008. Т.49. №. 6. С.1026-1032.
15. Randić M. The connectivity index 25 years after // J. of Mol. Graph. and Model. 2001. № 20. P. 19–35.
16. Randić M., Zupan J. On interpretation of well-known topological indices // J. Chem. Inf. Comput. Sci. 2001. V.41. P.550-560.
17. Применение теории графов в химии / под ред. Н.С. Зефирова. Новосибирск: Наука, 1988. 305 с.
18. Набивач В.М., Васильев В.Э. Корреляционные зависимости газохроматографических индексов удерживания от физико-химических свойств и структуры ароматических углеводородов // Журн. физ. химии, 1988. Т.62. №7. С.1873-1877.
19. Малышева Ю.А., Папулов Ю.Г., Виноградова М.Г., Ботов А.Б., Смоляков В.М. Свойства и строение органических молекул. I. Трёхмерные топологические индексы алканов // Журн. структ. химии. 1998. Т.39. №3. С.484-492.

20. Красных Е.Л. Прогнозирование энтальпии испарения на основе индексов Рэндиша. III. Карбоновые кислоты // Журн. структ. химии. 2010. Т.51. №. 2. С.231-236.
21. Randic M., Zupan J. On interpretation of well-known topological indices // J. Chem. Inf. Comput. Sci. 2001. V.41. P.550-560.
22. Trinajstić N. Chemical graph theory. London: CRC Press. 1992. 322 p.
23. Randic M. In search of structural invariants // J. Math. Chem. 1992. V. 9. P. 97-146.
24. Кузьмин В.Е., Коновороцкий Ю.П. Взаимосвязь структурных и топологических характеристик молекул // Журн. структур. химии. 1985. Т.26. №4. С.14-21.
25. Hall L.H., Mohney B., Kier L.B. The electrotopological state: structure information at the atomic level for molecular graphs // J. Chem. Inf. Comput. Sci. 1991. V.31. P.76-82.
26. Kier L.B. Shape indices of orders one and three from molecular graphs // Quant. struct.-act. relat. 1986. V.5. P.1-7.
27. Виноградова М.Г., Папулов Ю.Г., Смоляков В.М., Салтыкова М.Н. Корреляция структура-свойство с использованием теории графов // Журн. физ. химии. 1996. Т.70. №4. С.675-680.
28. Папулов Ю.Г., Чернова Т.И., Смоляков В.М., Поляков М.Н. Использование топологических индексов при построении корреляции структура - свойство // Журн. физ. химии. 1993. Т.67. №2. С.203-209.
29. Торопов А.А., Торопова А.П., Исмаилов Т.Т., Воропаева Н.Л., Рубан И.Н. Корреляции индексов, вычисляемых по матрице математических топологических расстояний с температурами кипения алкилбензолов // Журн. структ. химии. 1997. Т.38. №1. С.167-172.
30. Трач С.С., Девдариани Р.О., Зефирова Н.С. Комбинаторные модели и алгоритмы в химии. Тополого-конфигурационные аналоги индекса Винера // Журн. органич. химии. 1990. Т.26. №5. С.921-932.
31. Diudea M.V., Gutman I. Wiener-Type Topological Indices // Croatica Chemica Acta. 1998. V.71. P. 21-51.
32. Зефирова Н.С., Палюлин В.А., Радченко Е.В. Метод анализа топологии молекулярного поля в исследованиях количественной связи между структурой и активностью органических соединений // Доклады академии наук. 1997. Т.352. №5. С.630-633.
33. Вилков Л.В., Матрюков В.С., Садова Н.И. Определение геометрического строения свободных молекул. Л.: Химия. 1978. 223 с.
34. Трач С.С. Математические модели в стереохимии. I. Комбинаторные характеристики состава, связанности и конфигурации органических молекул // Журн. органич. химии. 1995. Т.31. №9. С.1320-1351.
35. Макаров Л.И. Методика нахождения информативного набора индексов молекулярных графов для прогноза свойств химических соединений // Журн. структ. химии. 1997. Т.38. №4. С.795-802.
36. Торопов А.А., Торопова А.П., Исмаилов Т.Т., Воропаева Н.Л., Рубан И.Н. Расширенная молекулярная связность: прогноз температур кипения алканов // Журн. структ. химии. 1997. Т.38. №6. С.1154-1159
37. Герасименко В.А., Набивач В.М. Сорбционно-структурные корреляции алкиланилинов // Журн. физич. химии. 1991. Т.65. № 8. С.2261-2266.

38. Набивач В.М. Корреляционные зависимости термодинамических функций растворения от структурных характеристик углеводов // Журн. физич. химии. 1992. Т.66. №3. С.824-826.
39. Chunhui Lu, Weimin Guo, Xiaofang Hu, Yang Wang, Chunsheng Yin. A novel Lu index to QSPR studies of aldehydes and ketones// J. Math Chem. 2004.V. 40. P.379-388.
40. Bo Zhou, Nenad Trinajstić. On resistance-distance and Kirchhoff index// J. Math Chem. 2009.V. 46. P.283-289.
41. Barysz M., Jashari G., Lall R.S., Srivastava V.K., Trinajstić N. // Chemical Applications of Topology and Graph Theory / King R.B. Amsterdam: Elsevier Science Publishers B.V. 1983. P.222.
42. Heuberger C. Wagner S.G. Chemical trees minimizing energy and Hosoya index// J Math Chem. 2009.V. 46. P.214–230.
43. Kier L.B., Hall L.H./ Molecular Connectivity in Chemistry and Drug Research. N.Y.: Acad. Press., 1981. V.70. №6. P.583.
44. Зенкевич И.Г. Принцип структурной аналогии при оценке газохроматографических индексов удерживания // Журн. структур. химии.1996. Т.37. №4.С.784-795.
45. Раевский О.А. Дескрипторы молекулярной структуры в компьютерном дизайне биологически активных веществ // Успехи химии. 1999. Т.68. №6. С.555-575.

ОГЛАВЛЕНИЕ

Предисловие	3
Топологический подход к установлению взаимосвязи «структура-свойство»	4
Молекулярный граф	6
Топологические индексы как структурные дескрипторы	7
Применение индексов связанности для характеристики структуры молекул	10
Использование ТИ для оценки физико-химических свойств соединений	12
Алгоритм расчета топологических индексов	16
Библиографический список	28

Учебное издание

Курбатова Светлана Викторовна,
Колосова Елена Александровна,
Финкельштейн Елена Евгеньевна

ТОПОЛОГИЧЕСКИЕ ИНДЕКСЫ В ХИМИЧЕСКИХ РАСЧЕТАХ

Учебное пособие

Публикуется в авторской редакции
Титульное редактирование *Т. И. Кузнецовой*
Компьютерная верстка, макет *Н. П. Бариновой*

Подписано в печать 29.12.14. Формат 60x84/16. Бумага офсетная. Печать оперативная.
Усл.-печ. л. 1,86; уч.-изд. л. 2,0. Гарнитура Times.
Тираж 100 экз. Заказ № 2583.
Издательство «Самарский университет», 443011, г. Самара, ул. Акад. Павлова, 1.
Тел. 8 (846) 334-54-23.
Отпечатано на УОП СамГУ.