

УДК 548.31

НЕЧЕТКИЙ НАБОР ОБРАЗУЮЩИХ КОНТАКТОВ В МОЛЕКУЛЯРНОМ АГЛОМЕРЕТЕ

А.М. Банару

(Московский государственный университет имени М.В. Ломоносова, химический факультет, кафедра физической химии; e-mail: banaru@phys.chem.msu.ru)

Показано, что с позиции теории нечетких множеств набор образующих (ключевых) контактов в любом молекулярном агломерате можно интерпретировать как нечеткий. Рассмотрены и систематизированы наборы образующих контактов в моносистемных цепочечных молекулярных агломератах. Проведено сравнение нечетких наборов образующих контактов по степени нечеткости.

Ключевые слова: нечеткое множество, молекулярный агломерат, межмолекулярный контакт.

В процессе формулирования теории плотной упаковки А.И. Китайгородский ввел понятие молекулярного координационного числа (МКЧ) [1] как числа молекул, касающихся исходной молекулы своей ван-дер-ваальсовой поверхностью. В органической кристаллохимии отдельно от ординарных ван-дер-ваальсовых контактов рассматривают специфические взаимодействия, например Н-связи. А.И. Китайгородский полагал, что в отсутствие таковых МКЧ обычно принимает значение 12. По мнению П.М. Зоркого, МКЧ = 14 реализуется все же чаще, чем МКЧ = 12 [2]. Однако подсчет МКЧ с помощью системы ван-дер-ваальсовых радиусов имеет очевидный недостаток: длина невалентного контакта может отклоняться в большую или меньшую сторону на 5% от суммы радиусов даже для однотипных пар атомов внутри одного кристалла. Принципиально иной подход к расчету МКЧ предложили авторы [3]. В их работе вокруг органической молекулы по оригинальной методике строился полиэдр Вороного–Дирихле, а число граней у полученного полиэдра приравнивалось к МКЧ.

Согласно данным этих авторов, в выборке моносистемных органических кристаллов из 33,5 тыс. структур МКЧ = 14 явно преобладало над МКЧ = 12. Совсем недавно были проанализированы данные, депонированные в CSD, по 166,3 тыс. гомомолекулярных кристаллических структур с небольшими органическими молекулами [4]. Более чем у 50% структур МКЧ = 14, около 20% структур имеют МКЧ = 16 и лишь у 5% МКЧ = 12. При этом чаще всего МКЧ = 14 реализуется в виде топологического эквивалента ОЦК-мотива с ромбододекаэдрическим полиэдром (примерно у 30% моносистемных

кристаллов, где молекулы занимают одну систему симметрически эквивалентных позиций, и у 18% бисистемных кристаллов, где таких систем занято две). Параллельно в этой работе анализировались кристаллические структуры белков, и самым распространенным для них оказалось МКЧ = 8.

Еще в 70-е годы прошлого века советскими кристаллографами было доказано [5], что периодичность кристалла следует из одинаковости локального окружения его структурных единиц. В их доказательстве использовалась так называемая правильная система точек Делоне, или (r, R) -система. Эта система состоит из таких точек пространства, что r -окрестность каждой точки не содержит других точек системы, и R -окрестность каждой точки пространства содержит как минимум одну точку. Позже было доказано, что для формирования двумерного (плоского) кристалла достаточно, чтобы каждая точка (r, R) -системы имела одинаковое окружение в окрестности радиуса $4R$ [6]. В той же работе было сделано предположение, что для формирования трехмерного кристалла будет достаточно одинаковости локального окружения в сфере радиуса $6R$. Однако совсем недавно Н.П. Долбилин, один из авторов первоначальной теоремы [5], показал, что для трехмерного случая достаточно одинаковых $10R$ -окрестностей и что эту границу в будущем едва ли удастся уменьшить [7]. Однако для одного особого случая, когда локальное окружение каждой точки центросимметрично, эта граница уменьшается до $2R$ [7–8]. Возникает вопрос, насколько с этими границами соотносится МКЧ. Содержит ли первая координационная сфера молекулы больше или, наоборот, меньше молекул, чем попадает в ее $10R$ или

2R-окрестность? Наши предварительные расчеты для серии кристаллических структур centrosимметричных α,ω -диолов показывают, что МКЧ примерно равно числу ближайших молекул внутри сферы соответствующего радиуса, но может отличаться от него как в большую, так и в меньшую сторону.

В связи с этим следует анализировать МКЧ не только с чисто геометрической точки зрения, но и с других позиций. Так, О.В. Гринева ввела понятие энергетического координационного числа молекулы (ЭКЧ), равного числу межмолекулярных контактов, у которых энергия отличается не более чем на 3% от самого сильного [9]. С учетом симметрической эквивалентности контактов ЭКЧ чаще всего бывает равно 1, 2, 4 или 8, т.е. заметно меньше, чем МКЧ. В работе [10] недавно был проведен поиск критерия, позволяющего с помощью расчета энергии найти значение самого МКЧ. Расчет по случайной выборке 1000 структур Кембриджского банка (CSD) показал, что этим критерием может служить «сильное» (не менее 40%) отличие в энергии последовательных невалентных контактов, ранжированных по убыванию энергии. Однако лишь примерно у трети анализированных структур первое «сильное» уменьшение энергии отделяло МКЧ от МКЧ+1, тогда как у остальных структур оно наблюдалось хотя бы на одном из более сильных контактов [10].

Недостаток МКЧ и ЭКЧ заключается в том, что эти характеристики включают симметрически эквивалентные контакты и несут мало информации о том, какие именно контакты (ключевые) необходимы для формирования данного кристалла. МКЧ, как правило, больше числа ключевых контактов, а ЭКЧ, наоборот, меньше. В [11] нами было определено критическое координационное число (ККЧ) как наименьшее число межмолекулярных контактов, достаточное для формирования данного кристалла. У гомомолекулярных моносистемных кристаллов (с $Z' = 1$) ККЧ равняется мощности минимального набора образующих элементов пространственной группы и может принимать значения от 2 до 6 включительно. Впрочем, у других кристаллов ККЧ может выходить за рамки этого интервала. В [12] приведены примеры структур класса $Fdd2$ ($Z = 8$), у которых ККЧ = 1, например, структура диметилсульфата (CSD-рефкод VEGNAX01) с коротким взаимодействием $\text{Me}\dots\text{O}=\text{S}$ ($\text{H}\dots\text{O}$ 2,4 Å), дихлорфторметана (SOPYUR) с агрегатом коротких взаимодействий $\text{Cl}\dots\text{F}$ 3,1 Å и другие. Отметим,

что в кристалле такие ключевые контакты далеко не всегда обладают самой большой энергией. Например, в структурах замещенных пролинов [13] среди самых сильных контактов почти всегда присутствует необязательный с точки зрения образующих элементов контакт трансляционно-связанных молекул, объединяющий их в стопки, в то время как контакт вдоль оси 2_1 в том же направлении проявляется слабее. Первый из этих контактов может быть симметрически порожден вторым, но не наоборот.

Образующие контакты

Всякая кристаллографическая группа является конечно-порожденной, поэтому некоторое конечное число ее элементов может симметрически породить все остальные элементы. У каждой кристаллографической группы свое число порождающих (образующих) элементов. Минимальным набором образующих элементов группы называется такой набор, который не содержит никакого другого набора образующих данной группы. Минимальные наборы с наименьшим возможным числом элементов мы называем существенно минимальными и далее в тексте под минимальными наборами подразумеваем именно их. У трехмерных точечных групп G_0^3 в таком наборе 2–3 элемента [14, с. 107], у двумерных пространственных (плоских) групп (G_2^2) в наборе имеются 2–4 элемента [15, с. 64–79], у трехмерных пространственных (федоровских) групп (G_3^3) от 2 до 6 элементов [11]. В теории групп эту величину иногда называют рангом группы. Минимальный набор образующих может быть выбран большим числом способов, у конечных групп оно конечно, у плоских и федоровских – бесконечно. Однако даже у федоровских групп число комбинаторно-различных вариантов все же ограничено (например, если рассматривать элементы симметрии в пределах одной элементарной ячейки кристалла). Набор образующих элементов группы можно трактовать как нечеткий. Пусть U – множество элементов u_i и $\mu_V: U \rightarrow [0,1]$.

Нечетким подмножеством V в U называется множество вида [16] $\{(u, \mu_V(u)): u \in U\}$, а $\mu_V(u)$ называют степенью принадлежности u к V ; $\mu_V(u) = 0$ означает полную непринадлежность u к V , а $\mu_V(u) = 1$ – его полную принадлежность, и в этом случае u называют четким элементом в V . Сумму $M = \sum_{i=1}^n \mu_V(u_i)$, где $n = |U|$, называют мощностью, или массой нечеткого подмножества V . По устоявшейся традиции нечеткое под-

множество обычно называют просто нечетким множеством. Утверждают, что нечеткие множества V и W равны, если для всех u из U $\mu_V(u) = \mu_W(u)$.

Степень нечеткости множества может быть охарактеризована несколькими величинами [17], в первую очередь энтропией, рассчитываемой по формуле:

$$S(V) = 1/n \sum_{i=1}^n (-\mu_i \ln \mu_i - (1-\mu_i) \ln(1-\mu_i)). \quad (1)$$

Энтропия монотонно растет в интервале (0, 1/2] и монотонно убывает в интервале [1/2, 1), а у четкого множества считается нулевой. Нормировка $1/n$ необходима для того, чтобы сравнивать энтропию нечетких множеств с разным числом элементов. Альтернативная мера нечеткости рассчитывается через расстояние между нечеткими множествами:

$$\xi = 1 - 2 \sigma(V; V_{0,5}), \quad (2)$$

где σ – нормированное расстояние между подмножеством V и максимально нечетким подмножеством $V_{0,5} = \{(u_i, 0,5)\}$:

$$\sigma(V; W) = \frac{1}{n} \sum_{i=1}^n |\mu_V(u_i) - \mu_W(u_i)|.$$

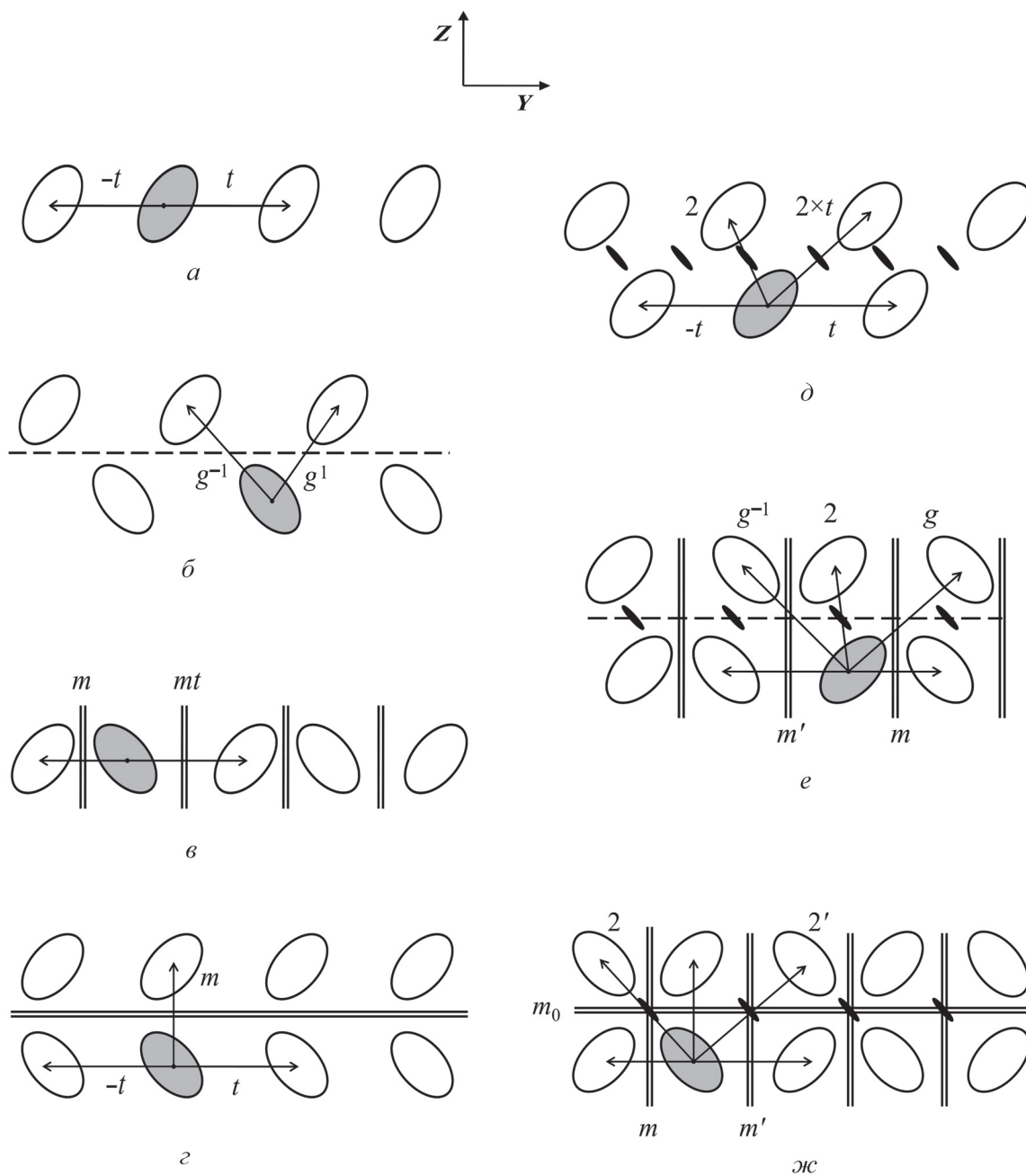
В настоящей работе мы применяем элементы теории нечетких множеств к межмолекулярным контактам. Хотя термин «контакт» используется в кристаллохимии в разных значениях, здесь мы будем понимать под ним упорядоченную пару молекул. Каждый межмолекулярный контакт отвечает некоторой операции симметрии в кристалле. Любой минимальный набор образующих группы симметрии молекулярного агломерата отвечает некоторому набору контактов. Цель настоящей работы заключалась в поиске нечетких наборов образующих контактов у простейших молекулярных агломератов. Наше рассмотрение ограничено наиболее распространенными из «первичных» [18] молекулярных агломератов в кристалле – молекулярными цепями, которые описываются группами симметрии бордюра G_1^2 (всего таких групп 7). Степень принадлежности образующего контакта к нечеткому набору мы рассчитываем как долю наборов, содержащих данный контакт, от общего числа комбинаторно-различных минимальных наборов. Масса нечеткого набора образующих контактов совпадает с рангом группы G_1^2 . Более сложные нечеткие наборы образующих у молекулярных слоев (группы G_2^2) обсуждаются нами отдельно [19].

Образующие контакты в общем случае не обязаны быть ни эффективными (т.е. давать наибольший вклад в энергию межмолекулярного взаимодействия в кристалле [18]), ни опорными (т.е. включаться в расчет МКЧ), хотя второе условие нарушается реже, чем первое. Среди образующих контактов иногда с необходимостью оказывается один из довольно слабых контактов [13], при отсутствии которого кристаллическая структура заданной симметрии просто бы не сформировалась.

Цепочечный агломерат

Заметим, что взаимно обратные операции симметрии u и u^{-1} всегда задают один и тот же межмолекулярный контакт, поскольку контакт между первой и второй молекулами – это тот же самый контакт, что между второй и первой молекулами. Поэтому u и u^{-1} объединяются в один образующий элемент. Цепь симметрии $p1$ (рисунок, а) может быть построена с помощью единственного образующего контакта между молекулами, связанными кратчайшей трансляцией (на рисунке он обозначен стрелкой, с обозначением t). В этом случае мы имеем четкий порождающий набор $\{(t; 1)\}$. У цепи симметрии $p11g$ (рисунок, б) порождающий набор тоже четкий и состоит из одного контакта $\{(g; 1)\}$, порожденно-го скольжением вдоль плоскости. Здесь контакт между трансляционно-связанными молекулами не может быть порождающим ($t = g^2$).

В цепи симметрии $p1m1$ (рисунок, в) контакты через плоскости зеркального отражения, расположенные по разные стороны от исходной молекулы, симметрически не эквивалентны. Если один из них обозначить m , то второй, расположенный по направлению t , соответствует симметрической операции mt . (Здесь и далее порядок выполнения операций принят слева направо.) Контакт через трансляцию в этом случае может быть порождающим элементом, но может быть и порожден ($t = mt \cdot m^{-1}$). В данном случае m^{-1} можно заменить на m , поскольку $m^2 = e$. Подчеркнем, что тройная композиция контактов – не то же самое, что тройная композиция операций симметрии. Контакт m^{-1} в обсуждаемом агломерате у третьей молекулы происходит не через ту же зеркальную плоскость, что связывает первые две молекулы, а через другую, отстоящую от первой на t (рисунок, в). Контакт именно через эту плоскость эквивалентен контакту первой пары молекул, тогда как при композиции операций было бы необходимо отразить третью



Цепочечные молекулярные агломераты ($Z' = 1$) симметрии $p1$ (а), $p11g$ (б), $p1m1$ (в), $p11m$ (з), $p2$ (д), $p2mg$ (е) и $p2mm$ (ж). Исходная молекула закрашена серым цветом

молекулу в первоначальной плоскости m , что означало бы $t = (mtm)^{-1}$.

Итак, получаются три контакта, претендующие на два вакантных места в минимальном наборе образующих. Возможны порождающие наборы $\{m, t\}$, $\{m, mt\}$ и $\{t, mt\}$, что эквивалентно нечеткому порождающему набору $\{(m; 2/3), (t; 2/3), (mt; 2/3)\}$.

Агломерат симметрии $p11m$ (рисунок, з) может быть порожден двумя контактами $\{m, t\}$.

Вместо m или t в порождающем наборе можно было бы задействовать скольжения tm или $-tm$. Вместе с тем контакты, отвечающие скольжениям, оказываются длиннее, чем t (если сравнивать расстояния между центрами масс молекул). Это означает, что энергия таких контактов в общем случае будет недостаточно велика для того, чтобы оказывать существенное влияние на формирование молекулярного агломерата. Исходная молекула как бы не «дотягивается» до дальних

соседей. По этой причине здесь и далее мы не включаем контакты длиной $d > t$ в нечеткий набор образующих. В результате у группы $p11m$ набор образующих контактов получается четким – $\{(m; 1), (t; 1)\}$.

В группе $p2$ ситуация иная (рисунок, d). Исходная молекула расположена относительно близко ($d < t$) к соседним молекулам, симметрически связанным с ней посредством двух разных поворотных осей, и каждый такой контакт симметрически уникален. Порождающим набором может быть t в паре с любым из этих контактов ($\{t, 2\}, \{t, 2t\}$), а также контакты через поворотные оси, взятые вместе ($\{2, 2t\}$). Запись $2t$ в данном случае означает не арифметическое произведение, а композицию поворота второго порядка с кратчайшей трансляцией вдоль Y (рисунок, d). Получается нечеткий набор $\{(t; 2/3), (2; 2/3), (2t; 2/3)\}$.

Порождающий набор группы $p2mg$ (рисунок, e) состоит из двух элементов. Контакт вдоль плоскости g может быть порожден любой парой ближайших друг к другу m и 2 . Однако, как в предыдущем случае, не всякая соседняя молекула, симметрически связанная с исходной, расположена близко ($d < t$). Разумеется, было бы неправильно говорить, что кратчайший контакт обязан быть самым сильным, потому что аттрактивно взаимодействовать могут выступающие части далеко отстоящих друг от друга молекул. Однако для однотипных элементов симметрии, отстоящих от молекулы через равные промежутки все дальше и дальше, энергия соответствующего контакта будет, несомненно, убывать. Так что одновременное попадание таких контактов в набор ключевых возможно лишь при попадании в него самого ближнего. Если рассматривать отдельно плоскости зеркального отражения, то образующие контакты у молекулы могут быть только с отражениями в плоскостях m и m' (рисунок, e). Возможны образующие наборы $\{2, m\}, \{2, m'\}$. Кроме того, с каждым из перечисленных контактов в отдельности образующий набор дает g ($\{g, m\}, \{g, m'\}, \{g, 2\}$). Всего получается 5 вариантов образующего набора, или нечеткий набор $\{(g; 3/5), (2; 3/5), (m; 2/5), (m'; 2/5)\}$.

Самым громоздким оказывается вывод образующих наборов в группе $p2mm$ (рисунок, $ж$), поэтому в отличие от остальных групп образующих контактов здесь не два, а три. Заметим, что на рисунке видны изображения не всех элементов симметрии, через которые у молекулы может

быть образующий контакт: среди них отсутствует та же g , которая была в предыдущей группе, но теперь ее скрывает зеркальная плоскость m_0 . Общее число возможных наборов по три образующих контакта равно 24, причем g опять имеет наибольшую из всех степень принадлежности к нечеткому набору: $\{(g; 1/2), (m_0; 5/12), (t; 5/12), (m; 5/12), (m'; 5/12), (2; 5/12), (2'; 5/12)\}$.

Расчет степени нечеткости наборов образующих контактов по формулам (1) и (2) показывает (таблица), что самым нечетко-порожденным является самый симметричный агломерат ($p2mm$). Агломераты симметрии $p1m1$ и $p2$ порождены одинаково нечетко, хотя их строение принципиально разное.

Если молекулы будут занимать частные позиции, то мощность набора образующих контактов будет уменьшаться на число занимаемых молекулой элементов симметрии. Так, при формировании цепи симметрии $p1m1$ из зеркально-симметричных молекул в минимальном наборе образующих останется единственный контакт t , и такой набор будет четким. То же самое получится у агломерата симметрии $p2$, в котором молекулы расположены на поворотных осях. Лишние образующие контакты при этом становятся элементами группы симметрии самой молекулы, что приводит к повышению симметрии ее потенциального поля [20] и позволяет сформировать агломерат из меньшего числа симметрически независимых контактов.

Заключение

Молекулярные структуры содержат множество межмолекулярных контактов, определяющих локальное окружение молекулы и координационные числа (МКЧ, ЭКЧ, ККЧ). Контакты отличаются друг от друга по кратности, энергии межмолекулярного взаимодействия, а также по необходимости присутствовать в порождающем (обра-

Степень нечеткости минимального набора образующих контактов в цепочечном молекулярном агломерате ($Z' = 1$)

Симметрия агломерата	S	ξ
$p1, p11g, p11m$	0	0
$p1m1, p2$	0,637	0,667
$p2mg$	0,673	0,800
$p2mm$	0,681	0,857

зующем) подмножестве конкретной молекулярной структуры. Подмножество (набор) структурообразующих межмолекулярных контактов в любом молекулярном агрегате может быть выбран бесконечно большим числом способов,

однако число комбинаторно-различных способов выбора конечно. Распределение принадлежностей образующих контактов к комбинаторно-различным порождающим наборам определяет меру нечеткости порождающего набора как целого.

СПИСОК ЛИТЕРАТУРЫ

1. *Китайгородский А.И.* Органическая кристаллохимия. М., 1955.
2. *Зоркий П.М.* // Журн. физ. химии. 1994. Т. 68. С. 966.
3. *Peresyupkina E.V., Blatov V.A.* // Acta Cryst. 2000. Vol. B56. P. 1035.
4. *Carugo O., Blatova O.A., Medrish E.O., Blatov V.A., Proserpio D.M.* // Sci. Rep. 2017. Vol. 7. P. 13209.
5. *Делоне Б.Н., Долбиллин Н.П., Штогрин М.И., Галуллин Р.В.* // Докл. АН СССР. 1976. Т. 227. С. 19.
6. *Senechal M.* // Comp. & Maths. with Appls. 1986. Vol. 12 B. P. 565.
7. *Dolbilin N.* // Struct. Chem. 2016. Vol. 27. P. 1725.
8. *Baburin I.A., Bouniaev M., Dolbilin N., Erokhovets N.Yu., Garber A., Krivovichev S.V., Schulte E.* // Acta Cryst. 2018. Vol. A74. P. 616.
9. *Гринева О.В.* // Журн. структ. химии. 2017. Т. 58. С. 390.
10. *Motherwell W.D.S.* // CrystEngComm. 2017. Vol. 19. P. 6869.
11. *Банару А.М.* // Вестн. Моск. ун-та. Сер. 2. Химия. 2009. Т. 50. С. 100.
12. *Лорд Э.Э., Банару А.М.* // Вестн. Моск. ун-та. Сер. 2. Химия. 2012. Т. 53. С. 81.
13. *Banaru A., Kochnev A.* // Studia UBB Chemia. 2017. Vol. LXII. P. 121.
14. *Галуллин Р.В.* Кристаллографическая геометрия. М., 2005.
15. *Коксетер Г.С.М., Мозер У.О.Дж.* Порождающие элементы и определяющие соотношения дискретных групп. М., 1980.
16. *Рыжов А.П.* Элементы теории нечетких множеств и ее приложений. М., 1998.
17. *De Luca A., Termini S.* // Information and Control. 1972. Vol. 20. P. 301.
18. *Зоркий П.М., Зоркая О.Н.* // Журн. структ. химии. 1998. Т. 39. С. 126.
19. *Banaru A.M.* // Cryst. Rep. 2018. Vol. 63. P. 1071.
20. *Zorky P.M.* // J. Mol. Struct. 1996. Vol. 374. P. 9.

Поступила в редакцию 08.11.2018
Получена после доработки 05.11.2018
Принята к публикации 15.12.2018

A FUZZY SET OF GENERATING CONTACTS IN A MOLECULAR AGGLOMERATE

A.M. Banaru

(*Lomonosov Moscow State University, Division of Physical Chemistry; e-mail: banaru@phys.chem.msu.ru*)

In the article it is shown that from the attitude of fuzzy sets theory a set of generating (key) intermolecular contacts of any molecular agglomerate can be considered as a fuzzy one. Generating sets of contacts are investigated and systematized for monosystemic molecular chains. A comparison of fuzziness for generating fuzzy sets of contacts is performed.

Key words: fuzzy set, molecular agglomerate, intermolecular contact.

Сведения об авторе: *Банару Александр Михайлович* – ст. преподаватель кафедры физической химии химического факультета МГУ имени М.В. Ломоносова, канд. хим. наук (banaru@phys.chem.msu.ru).