

УДК 547.12:541.14(083.73)

## ЛИНЕЙНО-ЦЕПНОЕ КОДИРОВАНИЕ ФОРМУЛ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ IX. ПОЛНАЯ ИНФОРМАЦИЯ О ЦИКЛАХ В УГЛЕВОДОРОДЕ

© Я.Э. Брюске

Brüske Y.E. The line-chain notation of the organic compounds formulae. IX. Complete information about cycles in hydrocarbon. Such information is given by introduction in a code of data about the general number of atoms of carbon and about the least number of carbon atoms of each of the simple condensed cycles, which have no other cycles inside themselves. The continuous numbering of all the carbon atoms of a molecule with appropriation of less numbers to atoms of the greatest external cycle is applied. The possibility of submission of this information is carried out by complication of the rules of univalent numbering of tops of the corresponding molecular graph. The riches and value of the new information justify this complication.

Возврат к автономной нумерации атомов углерода дал возможность дополнительно получить явную информацию о длинах главной и всех боковых цепей в коде алифатического углеводорода. Для циклического углеводорода начало нумерации атомов углерода по наибольшему внешнему циклу определяет его условную форму (габарит). В этой же работе [1] была отмечена возможность предоставления явной информации о числе и величине каждого из его конденсированных циклов. Эта возможность осуществлена в настоящей работе.

В основу положена идея построения номенклатурных названий ароматических углеводородов, по которой положение шестичленного цикла в молекуле определяется наименьшим номером его атома [2]. Она была уже использована автором настоящей работы в линейно-цепных кодах ароматических углеводородов [3], а здесь она применена для обозначения в коде положения и длины (число атомов углерода в цикле) любого произвольного цикла.

**Возврат к сплошной нумерации и R-кодирование.** Циклический углеводород нумеруют от углеродных атомов наибольшего внешнего цикла и продолжают нумерацию по внутренней цепи, как это предложено в [1]. В коде цикл любой длины обозначают буквой R (от Ring – кольцо<sup>1</sup>). Число перед этой буквой является наименьшим номером атома обозначаемого цикла, а число после нее обозначает длину этого цикла. В отдельном сообщении обозначают внешний цикл, признаком которого является отсутствие номера перед буквой R, после которой обозначают длину внешнего цикла.

Наименьшие номера атомов циклов являются первыми номерами нецепных связей кода, а вторых номеров в коде нет; о них имеется косвенная информация, задаваемая числом атомов в каждом цикле. Поэтому фрагмент, цепь которого не является продолжением нумерации внешнего цикла, в отличие от [1], целесообразно нумеровать автономно. Следует

возвратиться к сплошной нумерации всех атомов углерода молекулы и в коде явно обозначать те нецепные связи начальных атомов, в которых эти атомы являются вторыми. Завершает код номер последнего атома в отдельном сообщении; если этот атом начальный, обозначают его нецепные связи. Эту разновидность линейно-цепного кода целесообразно назвать R-кодом<sup>2</sup>. На рис. 1 показаны углеводороды, приведенные в работе [1]. На рис. 2 показан углеводород, имеющий несистематическое название коронен (а) [4], его неароматический (б) и неполные ароматические аналоги (в, г). Буква А обозначает теперь шестичленный ароматический цикл, поэтому после нее длину цикла не обозначают.

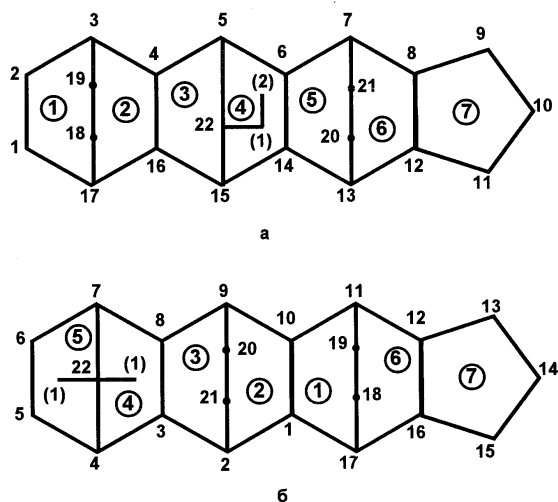
Объединение сообщений с одинаковостями приводят по общим правилам [5, 6].

**Другие определения.** *Собственная нумерация циклов.* Поскольку в R-коде явно обозначен наименьший номер каждого цикла, целесообразно упорядочить их по возрастанию этих номеров. Так, в углеводороде рис. 1а циклу с наименьшим номером 1 дают первый собственный номер, с наименьшим номером 5 – четвертый и т. д. У углеводорода рис. 1б циклов с одним и тем же наименьшим номером (здесь 1) два. Меньший собственный номер присваивают тому из них, в котором к атому с наименьшим номером присоединен следующий в порядке нумерации атом с большим номером, т. е. циклу 1, 10, 11, 19, 18, 17, а цикл 1, 2, 21, 20, 9, 10 получает последующий (второй) номер. Собственные номера циклов приведены в кружках на рис. 1.

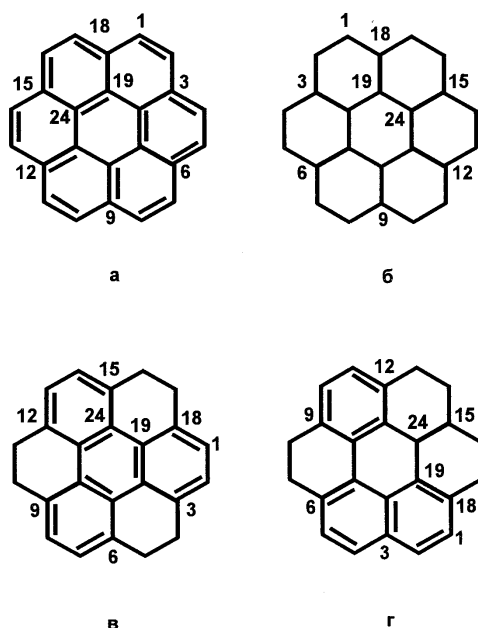
*Простые и составные циклы.* В [1] циклы, имеющие внутри себя другие циклы, определены как циклы, не имеющие одной грани. Целесообразнее определять такие циклы как *составные*, т. к. на самом деле они имеют столько граней, сколько содержат *простых* циклов, каждый из которых имеет одну грань.

<sup>1</sup> Буква R уже была использована в [3] для такого же обозначения ароматического цикла, однако для обозначения последнего целесообразнее использовать, вместо R, букву A: от слова арен (ароматический углеводород – англ. arene, нем. Aрен).

<sup>2</sup> Для укорочения терминологии целесообразно код по внешнему циклу, предложенный в [1], назвать C-кодом (Cycle), а код по внешнему циклу со сплошной нумерацией – CR-кодом.



**Рис. 1.** Углеводороды  $C_{24}H_{36}$ . С-коды: а) 12,8-13(2),7-14,6-15(1(2)),5-16,4-17,1-19,3; б) 7(1(1)(1)),4 -8,3-9(2),2-10,17,1-16,12-19,11. Названия: а) 5,15-этилметано-7,13-этан-8,3-10,1-16,12-гептациклогептадека-3-нонадекан; б) 4,7-диметилметано-2,9-этан-8,3-10,1-16,12-гептациклогептадека-11-нонадекан. CR-коды: а) 12,8-14,6-16,4-17,1-19,3-020,13-21,7-022(2),5,15; б) 8,3-10,17,1-16,12-19,11-020,9-21,2-022(1)(1),4,7. R-коды: а) 1,3,6,7R6-4,5,8R5-R17-020,13-022(2),5,15. б) 1,1,2,11R6-3,4,12R5-R17-020,9-022(1)(1),4,7. Названия: а) 22-этил-4,8,5-пента-1,3,6,7-гексагептациклогептадека-20,13-5,15-доэйкозан; б) 22-диметил-3,4,12-пента-1,1,2,11-гексагептациклогептадека-20,9-4,7-доэйкозан



**Рис. 2.** Коронен (а:  $C_{24}H_{12}$ ) и его аналоги (б:  $C_{24}H_{18}$ , в:  $C_{24}H_{18}$ , г:  $C_{24}H_{20}$ ). R-коды: а) 1,3,6,9,12,15,19A-R18-24; б) 1,3,6,9,12,15,19R6-R18-24; в) 1,6,12,19A-3,9,15R6-R18-24; г) 1,3,9A-6,12,15,19R6-R18-24. Названия: а) октадекатетракозагептаарен-1,3,6,9,12,15,19; б) 1,3,6,9,12,15,19-гексагептациклооктадекатетракозан; в) 3,9,15-гексатрициклооктадекатетракозатетраарен-1,6,12,19; г) 6.12.15.19-гексатетрациклооктадекатетракозатриарен-1,3,9

*Циклы с одной нецепной связью.* У такого цикла номера составляют одну числовую последовательность. Одно сообщение о нецепной связи в коде все-

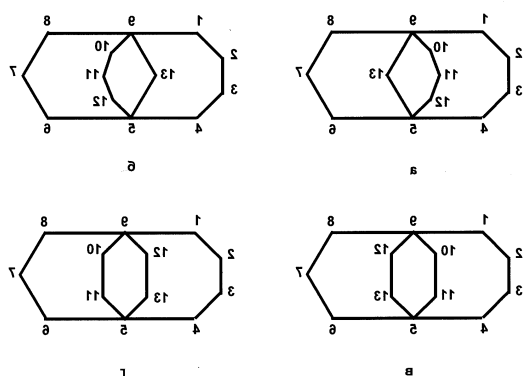
гда замыкает цикл с такой связью. Его длину легко определить путем вычитания первого номера этой связи из второго и прибавления единицы. Однако большей частью такой цикл содержит внутри себя другие циклы, т. е. является составным, а простой цикл чаще содержит несколько нецепных связей, и его номера состоят из нескольких числовых последовательностей.

*Сложноконденсированные углеводороды.* К ним относятся углеводороды, имеющие атомы, которые являются общими у трех или четырех (больше четырех быть не может) внутренних циклов. Углеводороды, у которых нет таких атомов, определяются как *простые*. По этой классификации углеводороды рис. 1а, б, 5а, б, в, г и 6 относятся к простым, а остальные являются сложноконденсированными. У углеводородов рис. 3а, б атомы 5 и 9 являются общими для всех трех внутренних циклов.

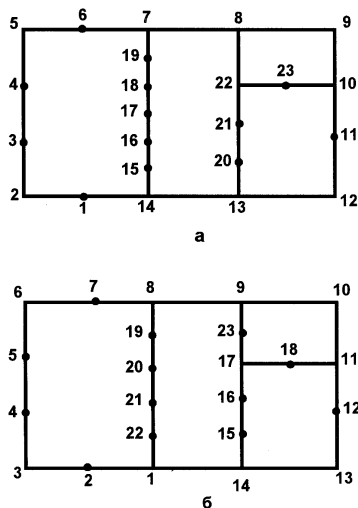
*Некоторые особенности R-кода. Неполнота.* Линейно-цепной код, в котором даная явная информация обо всех номерах нецепных связей, называется *полным*. Такими являются коды цепной канонической нумерации [6], линейно-цепной нумерации алифатических углеводородов [7], С- и CR-коды. R-код является неполным, т. к. в нем отсутствует явная информация о вторых номерах нецепных связей. Наименьшие номера циклов являются первыми номерами, а явно обозначенное число атомов в каждом цикле даёт неявную (косвенную) информацию о вторых номерах нецепных связей атомов, не являющихся начальными. Однако превратить эту информацию в явную при отсутствии нумерованного МГ непросто.

*Спирогидроуглеводороды.* Если в углеводороде имеются разрывы цепи внешнего цикла у атомов с меньшими номерами, чем у последнего, эти номера помещают в отдельных сообщениях кода, при этом сразу становится ясным, что последний атом внешнего цикла с первым также не связан (рис. 5г). Если же нет только этой связи (цепь внешнего цикла не разорвана, только не замкнута в цикл), её отсутствие обозначают нулем перед буквой R в сообщении о числе атомов во внешнем цикле (рис. 5в), а если последний атом начальный, ноль ставят после буквы R (рис. 5а, б). Нецепные связи начальных атомов, разрывающих внешний цикл, не обозначают.

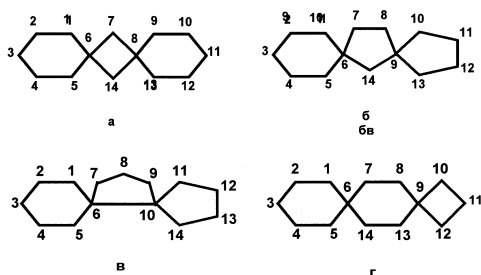
*Неоднозначность.* Различные проекции одного и того же углеводорода (изоморфные плоские [1, 8] МГ) с проекциями одного и того же наибольшего внешнего цикла могут иметь различные, внутренние циклы. Поэтому, в отличие от других линейно-цепных кодов, R-коды таких проекций различные, и здесь необходимо ввести дополнительный критерий однозначности проектирования. Две таких проекции показаны на рис. 3а, б. Из двух возможных выбирают ту, у которой имеется более длинный цикл с меньшим собственным номером (I), т. е. 3а. На рис. 3в, г приведены две проекции, которые не дают неоднозначности R-кода, однако при нумерации появляется неоднозначность номеров простых циклов. Правильной считают ту нумерацию, при которой меньшие номера получают цикл с меньшим собственным номером (II), т. е. 3в. Эти виды неоднозначности обусловлены возможностью поворота одного (иногда больше – рис. 7) цикла вокруг оси, без разрыва связей на  $180^\circ$ , при этом изменяются не поворачивающийся цикл, а циклы, с ним сконденсированные. Поворот



**Рис. 3.** Углеводороды  $C_{13}H_{24}$ . Коды а, б: С-код: 9,1-9(1),5-12,5 CR-код: 9,1-12,5-013,5,9 R-коды: а) 1R9-5,5R6-R9-013,5,9; б) 1R7-5R6R8-R9-013,5,9 название по а: 5,5-гекса-1-нонтрициклонона-5,9-тридекан. Коды в,г: С-код: 9,1-9(2),5-11,5 CR-код: 9,1-11,13,5-012,9 R-код: 1R8-5R6R7-R9-012,9-13. Название: 5-гексагепта-1-октатрициклонона-12,9-тридекан



**Рис. 4.** Сложноконденсированные углеводороды  $C_{23}H_{40}$ . CR-коды: а) 14,1-19,7-020,13-22,8-23,10; б) 14,22,1-18,11-019,8-023,9,17 R-коды: а) 1R13-7R12-8R5-10R8-R14-020,13-23; б) 1,1R12-9R6-11R8-R14-019,8-023,9,17 названия: а) 8-пента-10-окта-7-додека-1-тридекатетрациклотетрадека-20,13-трикозан б) 9-гекса-11-окта-1,1-додекатетрациклотетрадека-19,8-9,17-трикозан



**Рис. 5.** Спироуглеводороды  $C_{14}H_{26}$ . R-коды: а) 1,8R6-6R4-R014; б) 1R6-6,9R5-R014; в) 1R6-6,10R5-0R14; г) 1,6R6-9R4-013-R14 названия: а) 6-бута-1,8-гекса трицикло-00-тетрадекан; б) 6,9-пента-1-гекса трицикло-00-тетрадекан; в) 6,10-пента-1-гекса трицикло-0-тетрадекан; г) 9-бута-1,6-гекса трицикло-13-тетрадекан. Второй ноль в названиях а, б показывает, что последний атом внешнего цикла не соединен не только с первым, но и с предшествующим ему атомом

возможен, когда этот цикл присоединен только к двум атомам внешнего (внешневнутреннего – рис. 8) цикла.

Другие виды неоднозначности, когда две различные нумерации одного и того же МГ (рис. 9вг) или два различных МГ (рис. 10) дают один и тот же R-код, рассмотрены ниже.

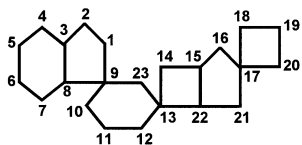
**Углеводороды с полными внутренними циклами.** Их в сложноконденсированных углеводородах необходимо выделить, т. к. они имеют ряд отличительных особенностей. Основное отличие: у этих циклов нет атомов, общих с атомами внешнего цикла. Такими углеводородами являются коронен и его аналог (рис. 2). Число полных внутренних циклов определяется числом их наименьших номеров, которые больше наибольшего номера внешнего цикла.

**Формы полных внутренних циклов.** Две основные структуры показаны на рис. 7 и 8. На первой из них имеется два разобщенных цикла внутри внешнего, тогда как на второй один из них находится внутри другого. Поэтому определяют *уровень (глубину) вложения* полного внутреннего цикла во внешний цикл. На рис. 7 оба шестичленных полных внутренних цикла имеют первую глубину вложения, а на рис. 8а два шестичленных цикла занимают первый и второй уровни. Следует отметить, что эти весьма сложные углеводородные структуры рассматриваются в [2], а углеводород рис. 7 синтезирован в 30-х годах и описан в каталоге циклических углеводородных структур [9]. Углеводород рис. 8а интересен тем, что длина одного из его внутренних циклов (11, 12, 13, 22, 21, 20, 19, 18, 17, 16) равна длине внешнего. Если перепроектировать его МГ, сделав внешним этот внутренний цикл, то получится МГ, приведенный на рис. 8бв. Чтобы показать соответствие проекций, на рис. 8б сохранена нумерация МГ рис. 8а, а на рис. 8в нумерация начата с нового внешнего цикла.

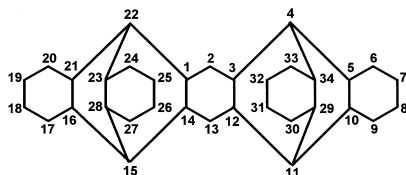
На рис. 9 приведены МГ, содержащие два (9аб) и три (9вг) полных внутренних цикла. В каждом из этих МГ можно выделить сложный полный внутренний цикл, в котором находятся простые полные внутренние циклы. Этот цикл, играющий роль внешнего по отношению к входящим в него внутренним циклам, целесообразно назвать «противоречивым» термином *внешневнутренний* цикл. Его отличительный признак: он сконденсирован по всем своим связям с неполными внутренними циклами.

В МГ рис. 8а имеется два полных шестичленных внешневнутренних цикла разных уровней. Цикл первого уровня (11, 12, 13, 14, 15, 16) составной, он содержит внутри себя десятичленный цикл (11, 12, 13, 22, 21, 20, 19, 18, 17, 16) и два шестичленных (13, 14, 15, 16, 17, 22 и 17, 18, 19, 20, 21, 22) цикла. Второй из этих шестичленных циклов (простой) является внешневнутренним циклом второго уровня. Его отличительные признаки: он не имеет атомов, общих с атомами цикла предшествующего уровня, а по всем своим связям он сконденсирован с внешневнутренними циклами предшествующего уровня. В МГ рис. 8в, вместо них, образовались два цикла первого уровня из шести и десяти атомов, которые входят во внешневнутренний цикл из десяти атомов.

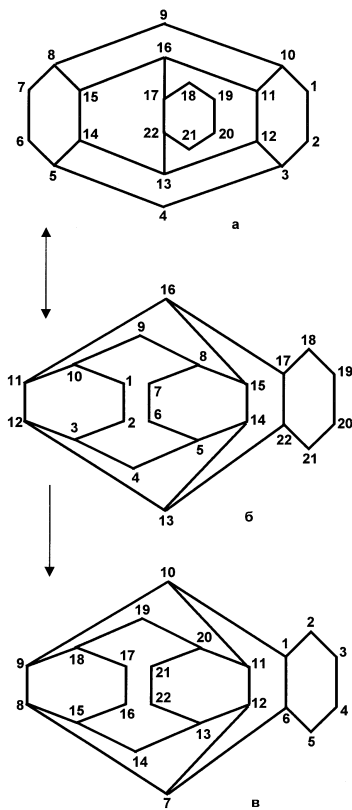
В *нумерацию атомов* МГ с полными внутренними циклами введены три важные дополнительные правила. Первое: нумерацию атомов полных внутренних циклов начинают с внешневнутреннего цикла до его



**Рис. 6.** «Спироконденсированный» углеводород  $C_{23}H_{36}$  R-код: 1,15R5-3,9R6-13,17R4-021-R023. Название: 13,17-бута-1,15-пента-3,9-гексагексацикло-21,0-трикозан. Ноль после числа 21 показывает то же, что и второй ноль в названиях рис. 6аб



**Рис. 7.** Углеводород  $C_{34}H_{48}$  с двумя полными внутренними циклами первого уровня. CR-код: 10,5-12,3-14,22,1-21,16-28,15,23-029,11-34,4,29 R-код: 1R10R6-4,5,15,16,23,29R6-3R10-R22-029,11-34. Название: 4,5,15,16,23,29-гекса-1-декагекса-3-деканонациклодокоза-29,11-тетратриаконтан



**Рис. 8.** Углеводород  $C_{22}H_{32}$ . а: проекция с внешневнутренними циклами двух уровней вложения; б: перепроектирование по новому внешнему циклу с сохранением нумерации; в: перенумерация проекции б по новому внешнему циклу CR-коды: а,б: 10,1-12,3-14,5-15,8-16,11-22,13,17; в: 6,10,1-12,7-15,8-18,9-20,11-22,13 R код и название по а: 1,3,5,8,13,17R6-11R10-R10-22 1,3,5,8,13,17-гекса-11-декагептациклодекадокозан. Выбор между проекциями а и в определяет п. 7 правил однозначной нумерации, т.к. ни один из предшествующих ему пунктов не дает преимущества ни одной из них. Первая нецепная связь CR-кода а (10,1) старше первой нецепной связи CR-кода в (6,1): 10>6. Поэтому углеводород надо проектировать в форме а

замыкания, а затем нумеруют другие атомы (III). На рис. 9 вначале пронумерованы атомы внешнего восьмичленного цикла, затем – шестичленного внешне-внутреннего, а потом получает номер атом внутри внешневнутреннего цикла. На рис. 8 после нумерации атомов внешнего цикла нумеруют атомы шестичленного внешневнутреннего цикла первого уровня, а затем – шестичленного цикла второго уровня.

Теперь необходимо обратить внимание на круговое направление нумерации атомов циклов. Номера внешних циклов МГ рис. 9аб последовательно обходят атомы по часовой стрелке, тогда как нумерация атомов обоих внешних циклов рис. 9вг осуществляется против часовой стрелки. Круговое направление нумерации атомов внешневнутренних циклов МГ рис. 9а и 9в одинаковое с направлением нумерации атомов соответствующих внешних циклов (по часовой стрелке на рис. 9а и против – на рис. 9в), тогда как круговые направления нумерации соответствующих внешних и внешневнутренних циклов МГ рис 9б и 9г противоположные. Второе правило: круговое направление нумерации атомов всех внешневнутренних циклов должно быть одинаковым с круговым направлением нумерации атомов внешнего цикла (IV). Нумерации МГ рис. 9б и 9г противоречат этому правилу и являются неправильными.

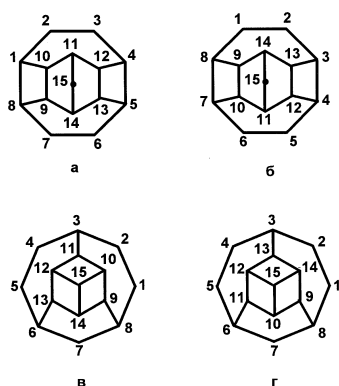
Третье правило применяется, когда в МГ находится начальный атом, входящий во внешневнутренний цикл, и он может получить номер в любом месте этого цикла (рис. 7). В этом случае действует правило выбора начального атома, аналогичное правилу выбора новой цепи в алифатическом углеводороде [7]: начальный атом выбирают так, чтобы он имел связь с атомом, пронумерованным раньше (V).

**Неоднозначность.** Теперь можно рассмотреть упомянутые выше другие виды неоднозначности. На рис. 9вг две различных нумерации одного и того же углеводорода дают одинаковые R-коды и, соответственно, одинаковые названия. Неоднозначность устраняется тем, что правильной считается только нумерация МГ рис. 9в. На рис. 10 приведен специально подобранный пример неоднозначности, при которой два различных, хотя и близких по структуре углеводорода имеют один и тот же R-код и название. Неоднозначность здесь устраняется путем вывода информации о циклах различной длины, имеющих один и тот же атом с наименьшим номером, в одно сообщение кода. Обозначения этих циклов помещают в нем в порядке увеличения собственных номеров (VI) этих циклов, а не их длин.

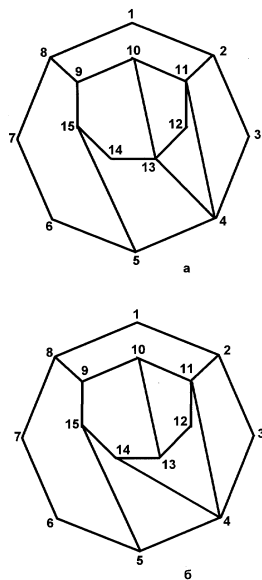
**Правила однозначной нумерации углеродных атомов молекулы.** Правильность нумерации атомов молекулярной структуры углеводорода (нумеруют атомы углерода структурной формулы, обычно это вершины соответствующего МГ) проверяют по CR-коду, т.к. по R-коду этого сделать нельзя, ввиду его неполноты. Затем по проверенному пронумерованному МГ составляют R-код.

Разработанные правила (они обозначены выше и, соответственно, будут обозначены ниже римскими числами в скобках) необходимо собрать вместе, чтобы расположить их в порядке убывающего старшинства [10, 11] так же, как это сделано для разработанных раньше видов линейно-цепного кодирования [1, 6, 7].

1. Меньшие номера дают атомам наибольшего внешнего цикла.



**Рис. 9.** Углеводороды  $C_{15}H_{20}$ . Круговое направление нумерации циклов. CR-коды: а: 8,10,1-12,4-13,5-14,9-15,11; б: 8,1-10,7-12,4-13,3-15,11; в: 8,1-11,3-13,6-14,9-15,10,12; г: 8,1-13,3-11,6-14,9-15,10,12 R-коды; а: 1R4R7-4R4-5R7-9,11R5-R8-15; б: 1,4R7-3,7R4-9,11R5-R8-15; в,г: 1,3R7-6R6-9,10,12R4-R8-15. Названия: а: 4-бута-1-бутагепта-9,11-пентагексациклооктапентадекан в,г: 9,10,12-бута-6-гекса-1,3-гептагексациклооктапентадекан



**Рис. 10.** Углеводороды  $C_{15}H_{18}$ . Иллюстрация одного из видов неоднозначности. CR-коды; а: 8,1-11,2,4-13,4,10-15,5,9; б: 8,1-11,2,4-13,10-14,4-15,5,9 R-коды; а, б: 1,5R6-2,4,10R4-4,9R5-R8-15. Названия: а, б: 2,4-бута-4,9-пента-1,5-гексагептациклооктапентадекан. Устранение неоднозначности. R-коды; а: 1,5R6-2,10R4-4R4R5-9R5-R8-15; б: 1,5R6-2,10R4-4R5R4-9R5-R8-15. Названия; а: 2,10-бута-4-бутапента-9-пента-1,5-гексагептациклооктапентадекан; б: 2,10-бута-4-пентабута-9-пента-1,5-гексагептациклооктапентадекан

2. Нумерацию атомов полных внутренних циклов начинают с атомов внешневнутренних циклов (III).
3. Круговое направление нумерации атомов всех внешневнутренних циклов должно быть одинаковым с круговым направлением нумерации атомов внешнего цикла (IV).
4. Число начальных атомов должно быть минимальным.
5. Каждый начальный атом должен быть соединен с атомом, имеющим меньший, чем у него, номер (V).
6. Множество номеров начальных атомов должно быть лексикографически [6, 7] возможно старше.

7. Множество нецепных связей должно быть лексикографически возможно старше.

Эти 7 пунктов дают однозначный CR-код. Три последующих пункта правил предназначены для устранения неоднозначности R-кода и для обеспечения однозначности нумерации некоторых простых циклов, которые не устраняются действием этих пунктов.

8. При неоднозначности проектирования циклов на плоскость выбирают проекцию, в которой большую длину имеет цикл с меньшим собственным номером (I).

9. При неоднозначности нумерации простых циклов меньшие номера дают атомам цикла с меньшим собственным номером (II).

10. Циклы различной длины, имеющие один и тот же атом с наименьшим номером, выделяют в R-коде в отдельное сообщение, располагая информацию о них в порядке увеличения собственных номеров (VI).

Здесь так же, как и в правилах других видов линейно-цепного кодирования [1, 10, 13], каждый пункт должен быть применен, если его выполнение не ухудшает результата выполнения старших (с меньшими номерами) пунктов.

Пункты 4, 6 и 7 правил обязательны для обеспечения однозначности нумерации во всех разновидностях линейно-цепного кодирования, начиная с канонического [6]; для обеспечения однозначности последнего этих трех пунктов достаточно, а для однозначности C- и CR-кода к ним достаточно добавить пункт 1. Дополнительные пункты в правилах для CR- и R-кодирования обусловлены неполнотой R-кода и необходимостью однозначного выделения и нумерации простых циклов. Без них приведенное ниже определение номеров простых циклов во многих случаях оказывается невозможным ни по R-, ни по CR-коду.

Пункты 1, 2, 3 и 5 этих правил ухудшают действие пункта 7, если бы этого не было, их можно было бы не вводить. В связи с этим возникает мысль о возможности снятия этого противоречия путем изменения пункта 7 на противоположный, сделав множество нецепных связей лексикографически младше. Не говоря о том, что это изменение уменьшит вероятность замещения атомов водорода у углеродных атомов с меньшими номерами, что явно нежелательно [1], оно и не поможет, т. к. внесет дополнительные неблагоприятные изменения в структуру нумерации, вследствие которых упомянутая невозможность определения номеров простых циклов останется.

**Определение номеров углеродных атомов каждого простого цикла.** Часто бывает, что нет МГ углеводорода, а только линейно-цепной код. В этом случае для определения номеров атомов каждого простого цикла необходима довольно сложная процедура. Можно составить алгоритм такого определения, однако рассмотрение этой процедуры на конкретных примерах займет значительно меньше места, чем описание общего алгоритма.

*Определение номеров по CR-коду.* В нем первые номера нецепных связей являются наименьшими номерами циклов, поэтому процедуру начинают с переписывания CR-кода с расположением сообщений в порядке увеличения этих номеров. При этом объединенные сообщения разделяют, а номера, не входящие в циклическую систему (в скобках), не пишут. Затем определяют число конденсированных циклов ( $c$ ) по формуле  $c = s - r$  [6], где  $s$  – число нецепных связей,  $r$  – число начальных атомов.

Пример 1. МГ рис. 1а.

17,1- 19,3- 16,4-022,5-14,6-21,7-12,8-020,13-22,15  
 $c = 9 - 2 = 7$ .

В переписанном CR-коде наименьшие номера циклов расположены в порядке увеличения собственных номеров этих циклов. Рассматривая каждое сообщение, обозначают наименьший номер  $i$ -го цикла ( $N_i$ ). Затем записывают последовательность его номеров, начиная с наименьшего, и до наименьшего номера  $i + 1$ -го цикла (связи 19,3, 16,4 и 022,5). От второго номера этой связи записывают последовательность, заканчивающуюся вторым номером нецепной связи с  $N_i$  (17, 19, 16). По числу номеров ставят длину  $i$ -го цикла ( $R_1 = 6, R_2 = 6, R_3 = 5$ ).

$N_1 = 1$ ; номера: 1, 2, 3, 19, 18, 17;  $R_1 = 6$

$N_2 = 3$ ; номера: 3, 4, 16, 17, 18, 19;  $R_2 = 6$ .

$N_3 = 4$ ; номера: 4, 5, 22, 15, 16;  $R_3 = 5$ . Здесь № 22 – начальный и последний, поэтому от него последовательность не может быть продолжена. Находится его связь с другим атомом (22,15), от которого и заканчивается последовательность.

$N_4 = 5$ ; номера: 5, 6, 14, 15, 22;  $R_4 = 5$  (связь 22, 5, 15).

$N_5 = 6$ ; номера: 6, 7, 21, 20, 13, 14;  $R_5 = 6$ . 20 – начальный атом, поэтому не 19..., а 13, 14 (020,13).

$N_6 = 7$ ; номера: 7, 8, 12, 13, 20, 21;  $R_6 = 6$ . 13 – первый номер следующей нецепной связи, поэтому надо продолжать запись по ней.

$N_7 = 8$ ; номера: 8, 9, 10, 11, 12;  $R_7 = 5$ . Цикл замкнулся на «своей» нецепной связи (12 < 13).

Пример 2. Углеводород рис. 7. Переписанный CR-код: 14,1-22,1-12,3-34,4-10,5-029,11-28,15-21,16-28,23-34,29

Число циклов:  $c = 10 - 1 = 9$ .

$N_1 = 1$ ; номера: 1, 14, 15, 28, 27, 26, 25, 24, 23, 22;  $R_1 = 10$ . Первый номер следующего сообщения также равен единице, поэтому сразу пишут второй номер рассматриваемого сообщения и от него – свою последовательность, которая обрывается на первом номере связи 28,15 и от № 28 последовательность заканчивается на втором номере связи 22,1.

$N_2 = 1$ ; номера: 1, 2, 3, 12, 13, 14;  $R_2 = 6$ .

$N_3 = 3$ ; номера: 3, 4, 34, 33, 32, 31, 30, 29, 11, 12;  $R_3 = 10$ . 29 – начальный атом и от него надо идти по его нецепной связи, а не к 28.

$N_4 = 4$ ; номера: 4, 5, 10, 11, 29, 34;  $R_4 = 6$  (по другой нецепной связи начального атома, т. к. 29 – первый номер этой связи).

$N_5 = 5$ ; номера: 5, 6, 7, 8, 9, 10.  $R_5 = 6$ .

$N_6 = 15$ ; номера: 15, 16, 21, 22, 23, 28.  $R_6 = 6$ . Больше, чем в двух циклах на плоскости ( $i = 3$  и 4), связь 29, 11 быть не может, поэтому она из рассмотрения исключена.

$N_7 = 16$ ; номера: 16, 17, 18, 19, 20, 21.  $R_7 = 6$ .

$N_8 = 23$ ; номера: 23, 24, 25, 26, 27, 28.  $R_8 = 6$ .

$N_9 = 29$ ; номера: 29, 30, 31, 32, 33, 34.  $R_9 = 6$ .

*Определение по R-коду.* Это определение в принципе сходно с определением по CR-коду, но немного сложнее. Определение начинают с цикла с наибольшим собственным номером, т. к. необходимая информация о нумерации атомов рассматриваемого цикла содержится в циклах с большими собственными номерами. Затем проводят это определение для остальных циклов в порядке уменьшения их собственных номеров.

Пример 3 (МГ рис. 1а):

$N_7 = 8, R_7 = 5$ ; номера: 8, 9, 10, 11, 12. Номера пишут последовательно от наименьшего номера цикла ( $N_7$ ) до окончания его длины ( $R_7$ ).

$N_6 = 7, R_6 = 6$ ; номера: 7, 8, 12, 13, 20, 21. Если длина цикла не закончена после номера  $N_{i+1}$ , её продолжают от номера, закончившего длину просмотренного перед ним цикла (12), до конца его длины или, как здесь, до появления первого номера нецепной связи начального атома, обозначенного в R-коде (20,13). В этом случае нумерацию продолжают от второго номера этой связи до окончания длины цикла.

$N_5 = 6, R_5 = 6$ ; номера: 6, 7, 21, 20, 13, 14. После номера  $N_{i+1}$  нумерацию ведут от последнего номера цикла  $i + 1$ , но в обратном порядке, т. к. атом 21 – конечный атом отрезка цепи (атом 22 начальный) и не может продолжать возрастающую нумерацию. Убывающую нумерацию продолжают до конца длины цикла или, как здесь, до начального атома с предшествующим собственным номером (по собственной нумерации начальных атомов) и тогда заканчивают от первого номера этой нецепной связи (020,13).

$N_4 = 5, R_4 = 5$ ; номера: 5, 6, 14, 15, 22 (связи 022,5,15).

$N_3 = 4, R_3 = 5$ ; номера: 4, 5, 22, 15, 16 (связи 022,5,15).

$N_2 = 3, R_2 = 6$ ; номера: 3, 4, 16, 17, 18, 19.

$N_1 = 1, R_1 = 6$ ; номера: 1, 2, 3, 19, 18, 17. Здесь 19 также конечный атом цепи (20 – начальный атом) и поэтому от него нумерацию атомов цикла 1 продолжают в убывающем порядке до окончания его длины.

Пример 4 (МГ рис. 1б).

$N_7 = 12, R_7 = 5$ ; номера: 12, 13, 14, 15, 16.

$N_6 = 11, R_6 = 6$ ; номера: 11, 12, 16, 17, 18, 19.

$N_5 = 4, R_5 = 5$ ; номера: 4, 5, 6, 7, 22 (связи 022,4,7).

$N_4 = 3, R_4 = 5$ ; номера: 3, 4, 22, 7, 8 (связи 022,4,7).

$N_3 = 2, R_3 = 6$ ; номера: 2, 3, 8, 9, 20, 21 (связь 020,9).

$N_2 = 1, R_2 = 6$ ; номера: 1, 2, 21, 20, 9, 10 (обратный порядок: 022, а после связи 020,9 прямой порядок).

$N_1 = 1, R_1 = 6$ ; номера: 1, 10, 11, 19, 18, 17. 11 является наименьшим номером цикла 6, поэтому дальше продолжать эту последовательность номеров нельзя. Необходимо вернуться к циклу 6 и закончить нумерацию атомов цикла 1 от конечного номера шестого цикла в обратном порядке.

Пример 5. Сложноконденсированный углеводород (рис. 4а).

$N_4 = 10, R_4 = 8$ ; номера: 10, 11, 12, 13, 20, 21, 22, 23. Здесь цикл с наибольшим собственным номером имеет более чем одну нецепную связь (связь 020,13).

$N_3 = 8; R_3 = 5$ ; номера: 8, 9, 10, 23, 22.

$N_2 = 7, R_2 = 12$ ; номера: 7, 8, 22, 21, 20, 13, 14, 15, 16, 17, 18, 19 (связь 020,13). Нумерацию от 22 продолжают в обратном порядке, как в цикле 3 ( $i + 1$ ).

$N_1 = 1, R_1 = 13$ ; номера: 1, 2, 3, 4, 5, 6, 7, 19, 18, 17, 16, 15, 14. Обратный порядок номеров: 20 не допускает прямого порядка нумерации после  $N_2 = 7$ .

Пример 6. Углеводород рис. 4б.

$N_4 = 11, R_4 = 8$ ; номера: 11, 12, 13, 14, 15, 16, 17, 18. Номер 17 соединен с 23, но к нему переходить нельзя, т. к. 23 не соединен с 11 и цикл 4 не замкнется.

$N_3 = 9, R_3 = 6$ ; номера: 9, 10, 11, 18, 17, 23. А здесь перейти к 23 нужно: 23 соединен с 9 (связи 023,9,17).

$N_2 = 1, R_2 = 12$ ; номера: 1, 2, 3, 4, 5, 6, 7, 8, 19, 20, 21, 22. Нецепная связь 019,8 появилась раньше, чем  $N_3$ .

$N_1 = 1, R_1 = 12$ ; номера: 1, 22, 21, 20, 19, 8, 9, 23, 17, 16, 15, 14 (связи 019,8 и 023, 9,17).

Пример 7: «Спироконденсированный» углеводород (МГ рис. 6).

$N_6 = 17, R_6 = 4$ ; номера: 17, 18, 19, 20.

$N_5 = 15, R_5 = 5$ ; номера: 15, 16, 17, 21, 22. Внешний цикл здесь разорван (21), поэтому номер конечного атома (20) не повторяют.

$N_4 = 13, R_4 = 4$ ; номера: 13, 14, 15, 22.

$N_3 = 9, R_3 = 6$ ; номера: 9, 10, 11, 12, 13, 23. И здесь разрыв внешнего цикла (23), поэтому номер 22 также не повторяют.

$N_2 = 3, R_2 = 6$ ; номера: 3, 4, 5, 6, 7, 8.

$N_1 = 1, R_1 = 5$ ; номера: 1, 2, 3, 8, 9.

Номера, заканчивающие длину каждого цикла, являются вторыми номерами нецепных связей с их наименьшими номерами. Нецепные связи начальных атомов, где эти атомы являются вторыми, входить сюда не могут, но все эти связи в коде обозначены явно. Номер, заканчивающий длину первого цикла, если это не спироуглеводород, всегда равен наибольшему номеру внешнего цикла; иное свидетельствует о неправомерности нумерации или ее воспроизведения.

Этих семи примеров достаточно для иллюстрации процедуры поиска номеров всех простых циклов любой молекулы по CR- и R-коду, а имея эти номера, можно построить и соответствующий МГ.

**Составление названия углеводорода по R-коду.** В современной номенклатуре имеется подходящий способ составления таких названий, по которому для обозначения числа атомов в конденсированном цикле от названия соответствующего углеводорода отнимают последнюю букву: пропан – пропа, пентан – пента и т. д. [4]. Перед каждым таким словом ставят наименьший номер его цикла из R-кода и отделяют его от слова дефисом. Циклы называют в порядке увеличения длины. Затем называют общее число конденсированных циклов и называют внешний цикл. После этого ставят номера начальных атомов без нулей перед ними и их нецепные связи с меньшими номерами. В корне пишут название последнего номера молекулы. Если последний номер начальный, перед названием корня пишут номера атомов, связанных с последним.

Названия спироуглеводородов составляют так же, при этом перед корнем названия пишут ноль, обозначающий наличие в нем спироатомов (рис. бабв). Если

имеются еще разрывы цепи атомов внешнего цикла, их обозначают полностью и тогда ноль перед корнем названия не пишут (рис. бг).

Не разделенные номерами термины пишут одним словом. При наличии двусмысленности ставят дефис, например, пентадека – пятнадцать, пента-дека – пять, десять.

Все названия приведены в подписях к рисункам.

Таким образом, применение теории графов к исследованию МГ углеводородов обусловило возможность предоставления полной информации обо всех простых циклах углеводорода. Эта возможность осуществлена в R-коде путем усложнения правил однозначной нумерации вершин МГ углеводорода. Богатство и ценность новой информации, о чем мечтали авторы фундаментальной работы по номенклатуре органических соединений ([2], с. 69, 71, 140), вполне оправдывает это усложнение.

## ЛИТЕРАТУРА

1. Брюске Я.Э. Линейно-цепное кодирование формул органических соединений. VIII. Увеличение явной информативности в кодах углеводородов. // Вестн. ТГУ. Сер. Естеств. и технич. науки. Тамбов, 2000. Т. 5. Вып. 1. С. 38-43.
2. Терентьев А.П., Кост А.Н., Цукерман А.М., Потанов В. М. Номенклатура органических соединений. М.: Изд-во АН СССР, 1955. 304 с.
3. Брюске Я.Э. Линейно-цепное кодирование формул органических соединений. IV. Кодирование связей в углеводородах // Вестн. ТГУ. Сер. Естеств. и технич. науки. Тамбов, 1997. Т. 2. Вып. 1. С. 57-59.
4. Номенклатурные правила ИЮПАК по химии. М.: ВИНТИ, 1979. Т. 2. 896 с.
5. Брюске Я.Э. Линейно-цепное кодирование формул органических соединений. VI. Симметрия. // Вестн. ТГУ. Сер. Естеств. и технич. науки. Тамбов, 1998. Т. 3. Вып. 4. С. 383-388.
6. Брюске Я.Э. Цепная нумерация и кодирование циклических углеводородов // Ж. структурной химии. 1995. Т. 36. № 4. С. 729-734.
7. Брюске Я.Э. Линейно-цепное кодирование и названия алифатических углеводородов // Вестник ТГУ. Сер. Естеств. и технич. науки. Тамбов, 1996. Т. 1. Вып. 1. С. 34-38.
8. Харари Ф. Теория графов. М.: Мир, 1973. 302 с.
9. Patterson A.M., Capell L.T. The Ring Index. N. Y., 1940.
10. Брюске Я.Э. Линейно-цепное кодирование формул органических соединений. VII. Старшинство // Вестн. ТГУ. Сер. Естеств. и технич. науки. Тамбов, 1998. Т. 3. Вып. 4. С. 389-393.
11. Лунский В. Комбинаторика для программистов. М.: Мир, 1988. 216 с.

Поступила в редакцию 23 ноября 1999 г.