# Applications of Degree Distribution. 2. Construction and Enumeration of Isomers in the Alkane Series[1]

Mark D. Jackson* and Theodore I. Bieber*

Department of Chemistry, Florida Atlantic University, Boca Raton, Florida 33431

Received March 12, 1993*

We report a new method utilizing degree distribution (DD) domains and subdomains to generate and enumerate constitutional isomers in the alkane series. For a given number of carbon atoms per molecule, $n$, various DD domains exist, each of which has a specific DD $[a,b,c,d]$, meaning that all its isomers possess $a$ carbons of degree 1, $b$ of degree 2, $c$ of degree 3, and $d$ of degree 4. Obviously, $a + b + c + d = n$ must be true for each domain. In most cases a DD domain is divisible into subdomains based on embryo skeletons of different structures with one and the same embryo composition $[c,d]$, where $[c,d]$ is identical with the last two terms of the degree distribution $[a,b,c,d]$ of that domain. An embryo skeleton is an incomplete skeleton consisting of carbons of degree 3 and 4 only. The simplest complete skeleton derivable from an embryo skeleton is the corresponding infant skeleton. A supplement of $a = c + 2d + 2$ carbons of degree 1 is required for the conversion of an embryo skeleton to its infant skeleton, no matter which embryo skeleton of composition $[c,d]$ is submitted to this conversion. Hence all embryo skeletons of composition $[c,d]$ produce infant skeletons with the same degree distribution $[a,b=0,c,d]$. Each of these infant skeletons must be "expanded", or "diluted", with $b$ carbons of degree 2 in all possible nonredundant ways to generate all the isomers of the corresponding subdomain in the $[a,b,c,d]$ domain. In this process *every* bond of the infant skeleton serves as a "box" for the insertion of zero to $b$ carbons of degree 2. The indicated task may be accomplished manually or by a computer program that is specific for a given infant skeleton but applicable to any value of $b$.

## A PERSPECTIVE ON ISOMER ENUMERATION

In the method of Cayley[2] for enumerating all of the constitutionally different alkanes with a given number $n$ of carbon atoms, the isomers are grouped into domains, each of which is distinct with respect to the number of carbon atoms in the longest chain (LC) contained by every one of its isomers. Hence, any two isomers belonging to different domains have different LC's. The domains are best referred to as LC domains. In filling up these domains, one commences with the domain whose LC is $n$. This domain contains only one alkane, the normal (unbranched) alkane. One then moves to domains whose LC is $n-1$, $n-2$, etc., until the minimum LC tolerated by $n$ carbons is reached.

Any LC is treated as a "V" of ordinary shape or one of modified shape, in which case the bottom point is replaced by a short horizontal line. Any LC with an odd number of carbon atoms can be represented by an ordinary V whose bottom point denotes the central carbon of the LC. Any LC with an even number of carbon atoms can be represented by a modified V whose bottom horizontal line denotes the bond between the two central carbon atoms of the LC. Consequently, any odd-numbered LC is said to be monocentric, and any even-numbered LC is said to be bicentric.

Imagine that the LC, in the shape of an ordinary or modified V, is fastened to a ladder in such a manner that all its carbon atoms are located on rungs of the ladder and all its bonds, except the central bond of a bicentric LC, connect carbon atoms on successive rungs. If, as is usually the case, branches emanate from the LC, they must be directed upward. All their carbon atoms must be located on rungs of the ladder, and all their bonds must join carbon atoms on successive rungs. In a given LC domain the difference between the number of carbon atoms of the LC ($n_{LC}$) and the total number of carbon atoms ($n$), hence $n - n_{LC}$, equals the number of carbon atoms contributed by branches. All possible combinations of branches and all possible location patterns must be utilized in order for the complete LC domain to be generated. In the course of this process, one must make sure that no branch reaches a rung above the one containing the end carbons of the LC. If that happens, the skeleton has a larger LC than intended and one would be mixing different LC domains. Furthermore, it is essential that redundancies be avoided. This means that careful attention must be paid to symmetry, a task that is facilitated by the shape of the V, ordinary or modified.

The method of LC domains started out as a constructionist one. In other words, the isomers in each domain were generated and counted. The sum of the counts for all LC domains is, of course, the total isomer count. In the case of large $n$'s, it became desirable to develop a mathematical enumeration based on the same physical model but obviating the need for generating structures. Such efforts were started by Cayley[2] and continued by Henze and Blair.[3,4] The recursion formulas developed by the latter two investigators have in more recent times been translated into computer programs, and these have been used to obtain the number of constitutional isomers in the alkane series for values of $n$ up to 57.[5] Balaban, Kennedy, and Quintas[6] have applied Pólya's theorem[7] to determine the number of isomers in each LC domain for a given $n$.

Various computer-oriented methods for isomer construction and enumeration not based on LC domains have been described.[8-15] These methods are based on rules (canons) that codify molecular structures. The method based on the $N$-tuple code[14] has been applied to the enumeration of constitutionally different alkanes with $n$ up to 100.

In the present paper we propose and illustrate a new method for isomer enumeration: Constitutionally different alkanes with a given $n$ are grouped into domains which are distinct not with respect to longest chain (LC) but with respect to degree distribution (DD).[1] They shall therefore be referred

* Abstract published in *Advance ACS Abstracts*, August 15, 1993.

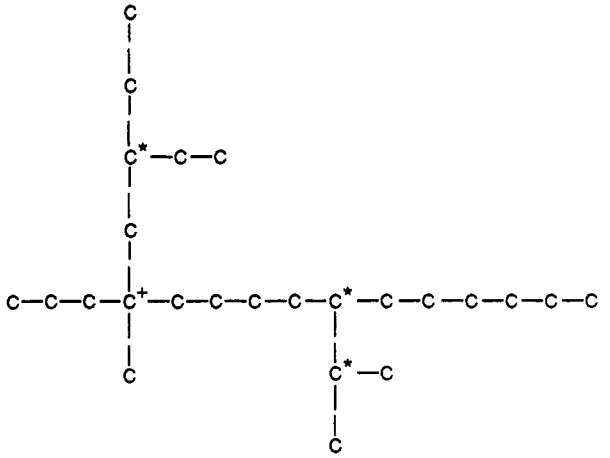
0095-2338/93/1633-0701$04.00/0 

to as DD domains. All of the isomers in a given DD domain have the same DD and are therefore, by definition, valence isomers of one another.

## STRATEGY OF ISOMER ENUMERATION USING DEGREE DISTRIBUTIONS

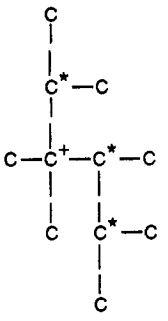
To develop a strategy for the enumeration of constitutionally different alkanes with a given number ($n$) of carbon atoms on the basis of DD domains, we find it convenient to distinguish between the following three types of skeletons: the *embryo skeleton*, the *infant skeleton*, and the *expansion skeleton*. These terms will be explained in retro-order.

Any skeleton that contains one or more carbons of degree 2 is an expansion skeleton. An example is shown as follows:



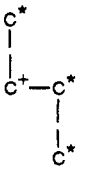
For reasons that will become apparent shortly, each carbon of degree 3 is provided with a certain marker, such as the asterisk (*), and each carbon of degree 4 is provided with a different marker, such as the cross (+).

We now excise all carbons of degree 2 and allow the remainder of the skeleton to collapse by bond formation between carbons previously separated by the excised atoms.



This results in an infant skeleton, which is defined as a skeleton that contains no carbons of degree 2. In the $N$-tuple coding system, such skeletons are called homeomorphic irreducible skeletons.[16]

Next, we remove every carbon of degree 1 from the infant skeleton, whereby we obtain



This is an embryo skeleton. While it looks like an ordinary carbon skeleton, one must keep in mind that it consists entirely of carbons of degree 3 and 4. It is really a germinal skeleton, germinal with respect to its infant skeleton and germinal with respect to all its expansion skeletons. The embryo skeleton, applied here to acyclic systems but equally applicable to other systems, is akin to the vertex graph of Masinter, Sridharan, Lederberg, and Smith[17] to denote the bridge head atoms of condensed ring systems; the vertex graph provides assistance in organizing the allocation of "secondary nodes" among all the bridges sharing the same pair of bridge head atoms.

There is a one-to-one relationship between an embryo skeleton and an infant skeleton. The embryo skeleton contains all the instructions one needs for building it up to the infant skeleton. One can justly claim that there is no "free will" between the embryo skeleton and the infant skeleton.

On the other hand, there is plenty of free will between the infant skeleton and the expansion skeletons. Every bond of the infant skeleton can serve as a "box" into which any number of carbons of degree 2 may be inserted. These carbons can therefore be regarded as fillers, extenders, expanders, or diluents. An infant skeleton may be "diluted" to yield any one of countless expansion skeletons.

One may distinguish between peripheral and nonperipheral dilution of an infant skeleton, even though both types of dilution can be mixed at will. In peripheral dilution all the bonds used as boxes are between carbons of degree 1 and embryo carbons. In nonperipheral dilution only bonds from one embryo carbon to another are used as boxes. Nonperipheral dilution leads to dispersion of the embryo skeleton, but, no matter how extensive this dispersion may be, the identity of the embryo skeleton is never lost.

Embryo size and embryo composition are useful terms when dealing with embryo skeletons. Embryo size, denoted by $q$, is the number of carbons in an embryo skeleton. Embryo composition, expressed as $[c,d]$, gives the number of carbons of degree 3 ($c$) and of degree 4 ($d$) composing an embryo skeleton. Because $q = c + d$, it follows that for a given embryo size, $q$, there must be $q + 1$ embryo compositions, which range from $[c=q,d=0]$ to $[c=0,d=q]$.

For any embryo composition whose $q$ exceeds 3, the existence of a variety of embryo skeletons is the rule. If $q = 3$, two of the four embryo compositions correspond to two embryo skeletons each, the remaining two embryo compositions to one embryo skeleton each. Every embryo composition whose $q = 2$, 1, or 0 corresponds to one embryo skeleton.

If $q = 0$ and hence $[c=0,d=0]$, signifying that the embryo skeleton is naught, the infant skeleton is C–C and all expansion skeletons are unbranched chains. In this case there is only one expansion skeleton for a given $n$, namely, C–$C_{n-2}$–C.

Table I shows the embryo skeletons for $q = 1$, 2, and 3. The infant skeleton for each embryo skeleton is also shown.

Table II, which applies to $q = 4$, shows for each embryo composition how many embryo skeletons are of the butane type and how many are of the 2-methylpropane type.

Table III, which applies to $q = 5$, is patterned after Table II. Now, however, embryo skeletons may be of three types, namely, pentane, 2-methylbutane, and 2,2-dimethylpropane.

With the embryo sizes so far considered, all embryo skeletons with a given embryo composition are easily found by inspection. With larger embryo sizes, this task increases in difficulty and a computer program becomes desirable. Such a program will be the subject of future work.

If any embryo skeleton having a given embryo composition $[c,d]$ is converted to its infant skeleton, the number of carbons of degree 1, $a$, to be appended is readily shown to be $c + 2d + 2$. Thus the DD of any infant skeleton with embryo

**Table I.** Embryo Skeletons and Infant Skeletons for $q$ = 1, 2, and 3

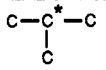
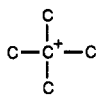
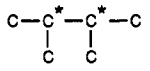
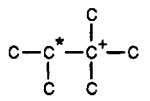
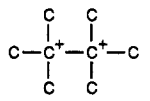
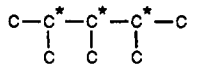
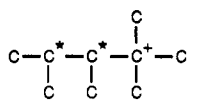
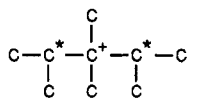
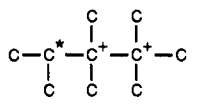
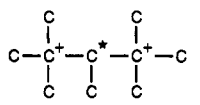
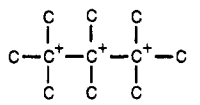
| $q$ | [$c$,$d$] | embryo skeleton(s) | infant skeleton(s) |
|---|---|---|---|
| 1 | [1,0] | C* |  |
| 1 | [0,1] | C+ |  |
| 2 | [2,0] | C*—C* |  |
| 2 | [1,1] | C*—C+ |  |
| 2 | [0,2] | C+—C+ |  |
| 3 | [3,0] | C*—C*—C* |  |
| 3 | [2,1] | C*—C*—C+ |  |
| | | C*—C+—C* |  |
| 3 | [1,2] | C*—C+—C+ |  |
| | | C+—C*—C+ |  |
| 3 | [0,3] | C+—C+—C+ |  |

**Table II.** Embryo Skeleton Multiplicities in the $q$ = 4 Group[a]

| | skeletal type | | |
|---|---|---|---|
| [$c$,$d$] | butane | 2-methylpropane | total |
| [4,0] | 1 | 1 | 2 |
| [3,1] | 2 | 2 | 4 |
| [2,2] | 4 | 2 | 6 |
| [1,3] | 2 | 2 | 4 |
| [0,4] | 1 | 1 | 2 |

[a] The number of embryo skeletons having a stated [$c$,$d$] is shown for each $C_4$ skeleton type.

**Table III.** Embryo Skeleton Multiplicities in the $q$ = 5 Group[a]

| | skeletal type | | | |
|---|---|---|---|---|
| [$c$,$d$] | pentane | 2-methylbutane | 2,2-dimethylpropane | total |
| [5,0] | 1 | 1 | 0 | 2 |
| [4,1] | 3 | 4 | 1 | 8 |
| [3,2] | 6 | 6 | 1 | 13 |
| [2,3] | 6 | 6 | 1 | 13 |
| [1,4] | 3 | 4 | 1 | 8 |
| [0,5] | 1 | 1 | 1 | 3 |

[a] The number of embryo-skeletons having a stated [$c$,$d$] is shown for each $C_5$ skeleton type.

composition [$c$,$d$] is [$a$=$c$+2$d$+2,$b$=0,$c$,$d$]. It also follows that the total number of carbons, $n_i$, of any infant skeleton with embryo composition [$c$,$d$] is $2c + 3d + 2$. Any of these infant skeletons is dilutable with some number, $b$, of carbons of degree 2 to yield expansion skeletons having a DD of [$a$=$c$+2$d$+2,$b$,$c$,$d$].

**Table IV.** Number of Primary Carbon Atoms ($a$) and the Total Number of Carbon Atoms ($n_i$) in Infant Skeletons Whose Embryo Skeletons Have the Indicated Embryo Compositions [$c$,$d$]

| $q = c + d$ | [$c$,$d$] | $a = c + 2d + 2$ | $n_i = 2c + 3d + 2$ |
|---|---|---|---|
| 0 | [0,0] | 2 | 2 |
| 1 | [1,0] | 3 | 4 |
| | [0,1] | 4 | 5 |
| 2 | [2,0] | 4 | 6 |
| | [1,1] | 5 | 7 |
| | [0,2] | 6 | 8 |
| 3 | [3,0] | 5 | 8 |
| | [2,1] | 6 | 9 |
| | [1,2] | 7 | 10 |
| | [0,3] | 8 | 11 |
| 4 | [4,0] | 6 | 10 |
| | [3,1] | 7 | 11 |
| | [2,2] | 8 | 12 |
| | [1,3] | 9 | 13 |
| | [0,4] | 10 | 14 |
| 5 | [5,0] | 7 | 12 |
| | [4,1] | 8 | 13 |
| | [3,2] | 9 | 14 |
| | [2,3] | 10 | 15 |
| | [1,4] | 11 | 16 |
| | [0,5] | 12 | 17 |
| 6 | [6,0] | 8 | 14 |
| | [5,1] | 9 | 15 |
| | [4,2] | 10 | 16 |
| | [3,3] | 11 | 17 |
| | [2,4] | 12 | 18 |
| | [1,5] | 13 | 19 |
| | [0,6] | 14 | 20 |
| 7 | [7,0] | 9 | 16 |
| | [6,1] | 10 | 17 |
| | [5,2] | 11 | 18 |
| | [4,3] | 12 | 19 |
| | [3,4] | 13 | 20 |
| | [2,5] | 14 | 21 |
| | [1,6] | 15 | 22 |
| | [0,7] | 16 | 23 |

In Table IV embryo compositions are grouped according to embryo size, from $q$ = 0 to $q$ = 7, and for each embryo composition are shown $a$ and $n_i$ of the corresponding infant skeleton(s).

The following regularities should be noted: Within any $q$ group the value of $n_i$ goes up by 1 each time $c$ is reduced by 1 and hence $d$ is increased by 1. The $n_i$ corresponding to [$c$=$q$,$d$=0], the $d$-poorest composition for a given $q$, is $2q$ + 2. The $n_i$ corresponding to [$c$=0,$d$=$q$], the $d$-richest composition for the same $q$, is $3q$ + 2. The jump in the value of $n_i$ between two $q$ groups, i.e. between [$c$=0,$d$=$q$], and [$c$=$q$+1,$d$=0] is $2-q$. Obviously, Table IV is readily extended as far as one desires.

Table IV (in extended form if necessary) provides a method for finding all the degree distributions applicable to isomeric alkanes with a chosen $n$. This method is an alternative to the one utilizing the computer program described in our previous paper.[1] One simply finds all the embryo compositions for which $n_i \leq n$. We shall call these the "approved" embryo compositions. Each such embryo composition corresponds to a certain degree distribution and hence to a certain DD domain. For each approved [$c$,$d$] the DD is [$a$=($c$+2$d$+2),$b$=($n$–$n_i$),$c$,$d$], which may also be written as [$a$=($c$+2$d$+2),$b$={$n$–(2$c$+3$d$+2)},$c$,$d$]. The relationships that appear in these expressions have already been deduced by Hendrickson and Parks.[18]

Having obtained all the DD domains for the chosen $n$, we have to find the subdomains for each domain. To achieve this

objective for a given domain, we focus on the [c,d] part of its DD and write all the embryo skeletons having this [c,d]. The number of embryo skeletons that can be written equals the number of subdomains. Each embryo skeleton is then converted into its infant skeleton, and the latter is diluted with $b = n - n_i$ carbons of degree 2 in all conceivable nonredundant ways to produce all the expansion skeletons comprising that subdomain. If the [c,d] of a DD domain is such that $b = n - n_i = 0$, then each subdomain contains only one skeleton, namely, the infant skeleton.

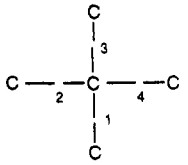
The isomer count for a DD domain is obtained by totaling the isomer counts for its subdomains, and the full isomer count is obtained by totaling the isomer counts for all DD domains.

Dilution of an infant skeleton with $b = n - n_i$ secondary carbon atoms to all possible expansion skeletons can in simple cases be carried out by hand. Generally, however, it is necessary to devise a computer program, to be dubbed a *dilution program*, that is specific for every infant skeleton. Our dilution programs are operative with *any* value of $b$, so that no further programming is required if the same infant skeleton is to be diluted with a larger $b$. Thus, if one wishes to perform an isomer count for a larger $n$, one can concentrate on the following tasks: Approving all additional embryo compositions that become effective between the old $n$ and the new $n$; finding all new embryo skeletons and hence all the corresponding infant skeletons, elaborating dilution programs for those new infant skeletons that require dilution and for former infant skeletons not previously in need of dilution.

## EXAMPLES OF DILUTION PROGRAMS

Programs of increasing complexity will be presented for three infant skeletons to be diluted to $n = 11$.

**Case 1: Embryo Composition [c=0,d=1].** The embryo composition [c=0,d=1] corresponds to DD [4,0,0,1] for the infant skeleton(s) and to DD [4,6,0,1] for the domain whose $n = 11$. Since only one embryo skeleton, $C^+$, has the indicated embryo composition, only one [4,0,0,1] infant skeleton exists



and hence only one subdomain makes up the [4,6,0,1] domain.

To obtain all the isomers of the DD domain, the infant skeleton above must be diluted with $b = 6$ carbons of degree 2. To the four boxes of this infant skeleton arbitrary numbers are assigned, as shown. All of the boxes are exchangeable. Hence, they must be ranked (hierarchically ordered) to avoid redundancy in isomer generation. For example: box 1 ≥ box 2 ≥ box 3 ≥ box 4. This is to be translated as follows: The occupancy of box 2 cannot exceed that of box 1, the occupancy of box 3 cannot exceed that of box 2, and the occupancy of box 4 cannot exceed that of box 3. It is in this sense that box 1 outranks box 2, box 2 outranks box 3, and box 3 outranks box 4.

The constraints on the occupancy of the boxes of this infant skeleton are as follows: (1) the hierarchy of exchangeable boxes described in the previous paragraph must be maintained, and (2) the total occupancies of the four boxes must equal $b$, in this case $b = 6$.

On the basis of these constraints the six carbons of degree 2 are distributed among the four boxes to yield unique occupancies and therefore unique expansion skeletons. For example, we may place all six carbons of degree 2 in a single box. In that event they must be placed in box 1. Placing them in any of the other boxes produces a redundant structure.

The constraints can be built into a computer program which generates expansion skeletons from unique arrangements of box occupancies. Table V presents the dilution program for this infant skeleton. Table VI shows the computer printout giving a depiction of the box occupancies for each expansion skeleton. Nine such expansion skeletons, hence nine isomers, are obtained. In the printout (Table VI) each of the four boxes of the infant skeleton is denoted with a number from 1 to 4. The program varies the occupancy of each box within the previously described constraints. Iterations covering all possible box occupancies are handled with DO loops. Since the total occupancies of the boxes must be monitored by the program, the DO loops which assign box occupancies must be nested. The upper limit of the outer loop, giving the occupancy of box 1, is $b$ (in this case 6). The next loop sets the occupancy of box 2 to be no larger than the occupancy

**Table V.** Dilution Program Generating the [4,b,0,1] Domain for any $b$ (Hence for $n = 5 + b$) from the [4,0,0,1] Infant Skeleton[a]

```
c
c
      write(6,*)'  +   subdomain (embryo-skeleton)'
      write(6,*)'                '
      write(6,*)' assignment of box numbers in infant skeleton
      write(6,*)'                '
      write(6,*)'          #       '
      write(6,*)'          3       '
      write(6,*)'      # 2 + 4 #    '
      write(6,*)'          1       '
      write(6,*)'          #       '
      write(6,*)'        '
      write(6,*)'          box number     '
      write(6,13)one,two,three,four
      write(6,*)'          occupancy '
c
      itmp=0
      DO 621 I1=0,ib
      iocc(1)=i1
      DO 621 I2=0,I1
      iocc(2)=i2
      DO 621 I3=0,I2
      iocc(3)=i3
      DO 621 I4=0,I3
      iocc(4)=i4
c
      ISUM=Iocc(1)+Iocc(2)+Iocc(3)+Iocc(4)
      IF(ISUM.NE.Ib)GO TO 621
      write(6,13)Iocc(1),Iocc(2),iocc(3),Iocc(4)
      IINT=IINT+1
      itmp=itmp+1
  621 CONTINUE
      write(6,*)'         isomers of this domain = ',itmp
      write(6,*)' ************************************************'
```

[a] The embryo skeleton is $C^+$.

**Table VI.** Dilution Program Output: Depiction of the Isomers of the [4,6,0,1] Domain

```
     +   domain (embryo-skeleton)

   assignment of box numbers in infant skeleton

          #
          3
      # 2 + 4 #
          1
          #

        box number
                                  1 2 3 4
        occupancy
                                  2 2 1 1
                                  2 2 2 0
                                  3 1 1 1
                                  3 2 1 0
                                  3 3 0 0
                                  4 1 1 0
                                  4 2 0 0
                                  5 1 0 0
                                  6 0 0 0


        isomers of this domain =          9

  ************************************************
```

**Table VII.** Dilution Program Generating the [5,*b*,2,1] Subdomain for Any *b* (Hence for $n = 8 + b$) from that [5,0,2,1] Infant Skeleton Whose Embryo Skeleton Is C*–C*–C+

```
c
      write(6,*)' *  -  *  -  +   subdomain (embryo-skeleton)'
      write(6,*)'                 '
      write(6,*)' assignment of box numbers in infant skeleton
      write(6,*)'                 '
      write(6,*)'           #     #     #  '
      write(6,*)'           2     4     8  '
      write(6,*)'      # 1 *  3  *  5  + 7 #'
      write(6,*)'                        6  '
      write(6,*)'                        #  '
      write(6,*)'       '
      write(6,*)'         box number     '
      write(6,13)one,two,three,four,five,six,
     1seven,eight
      write(6,*)'         occupancy '
c
      itmp=0
      DO 1621 I1=0,ib
      iocc(1)=i1
      DO 1621 I2=0,I1
      iocc(2)=i2
c
      DO 1621 I3=0,Ib
      iocc(3)=i3
      DO 1621 I4=0,Ib
      iocc(4)=i4
      do 1621 i5=0,ib
      iocc(5)=i5
c
      do 1621 i6=0,ib
      iocc(6)=i6
      do 1621 i7=0,i6
      iocc(7)=i7
      do 1621 i8=0,i7
      iocc(8)=i8
c
      ISUM=Iocc(1)+Iocc(2)+Iocc(3)+Iocc(4)+Iocc(5)+iocc(6)+iocc(7)+
     1iocc(8)
      IF(ISUM.NE.Ib)GO TO 1621
      write(6,13)Iocc(1),Iocc(2),iocc(3),Iocc(4),
     1Iocc(5),iocc(6),iocc(7),iocc(8)
      IINT=IINT+1
      itmp=itmp+1
 1621 CONTINUE
      write(6,*)'         isomers of this subdomain = ',itmp
      write(6,*)' ************************************************'
```

**Table VIII.** Dilution Program Output: Depiction of the Isomers of the First Subdomain of [6,2,2,1]

```
   *  -  *  -  +   subdomain (embryo-skeleton)

   assignment of box numbers in infant skeleton

            #     #     #
            2     4     8
      # 1 *  3  *  5  + 7 #
                        6
                        #

         box number
                                  1 2 3 4 5 6 7 8
         occupancy
                                  0 0 0 0 0 1 1 0
                                  0 0 0 0 0 2 0 0
                                  0 0 0 0 1 1 0 0
                                  0 0 0 0 2 0 0 0
                                  0 0 0 1 0 1 0 0
                                  0 0 0 1 1 0 0 0
                                  0 0 0 2 0 0 0 0
                                  0 0 1 0 0 1 0 0
                                  0 0 1 0 1 0 0 0
                                  0 0 1 1 0 0 0 0
                                  0 0 2 0 0 0 0 0
                                  1 0 0 0 0 1 0 0
                                  1 0 0 0 1 0 0 0
                                  1 0 0 1 0 0 0 0
                                  1 0 1 0 0 0 0 0
                                  1 1 0 0 0 0 0 0
                                  2 0 0 0 0 0 0 0

         isomers of this subdomain =          17

   ************************************************
```

of box 1. Continuing in a similar fashion, we structure the loops which assign the occupancies of boxes 3 and 4.

**Cases 2 and 3: Embryo Composition [*c*=2,*d*=1].** The embryo composition [*c*=2,*d*=1] corresponds to DD [6,0,2,1] for infant skeletons. Hence this embryo composition yields DD [6,2,2,1] for the domain of expansion skeletons whose $n = 11$. Since two embryo skeletons, C*–C*–C+ and C*–C+–C*, have the indicated embryo composition, two [6,0,2,1] infant skeletons exist and hence two subdomains make up the [6,2,2,1] domain.

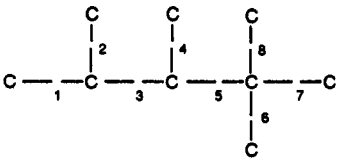
**Table IX.** Dilution Program Generating the [5,*b*,2,1] Subdomain for Any *b* (Hence for $n = 8 + b$) from That [5,0,2,1] Infant Skeleton Whose Embryo Skeleton is C*–C+–C*

```
c
      write(6,*)' *  -  +  -  *   subdomain (embryo-skeleton)'
      write(6,*)'                 '
      write(6,*)' assignment of box numbers in infant skeleton
      write(6,*)'           #     #     #  '
      write(6,*)'           2     8     3  '
      write(6,*)'      # 1 *  6  +  7  * 4 #'
      write(6,*)'                  5        '
      write(6,*)'                  #        '
      write(6,*)'       '
      write(6,*)'         box number     '
      write(6,13)one,two,six,five,eight,
     1seven,three,four
      write(6,*)'         occupancy '
      itmp=0
      DO 2621 I1=0,ib
      iocc(1)=i1
      DO 2621 I2=0,I1
      iocc(2)=i2
      DO 2621 I3=0,I1
      iocc(3)=i3
      DO 2621 I4=0,I3
      iocc(4)=i4
      do 2621 i5=0,ib
      iocc(5)=i5
      do 2621 i8=0,i5
      iocc(8)=i8
      do 2621 i6=0,ib
      iocc(6)=i6
      do 2621 i7=0,ib
      iocc(7)=i7
      ISUM=Iocc(1)+Iocc(2)+Iocc(3)+Iocc(4)+Iocc(5)+iocc(6)+iocc(7)
     1+iocc(8)
      IF(ISUM.NE.Ib)GO TO 2621
      isym1=0
      if((iocc(1).eq.iocc(3)).and.(iocc(2).eq.iocc(4)))isym1=1
      ilsum=IOCC(2)+iocc(1)
      irsum=IOCC(4)+iocc(3)
      IF((irsum.GT.ilsum).and.(iocc(1).eq.iocc(3)))GO TO 2621
      IF((IOCC(7).GT.IOCC(6)).and.(isym1.eq.1))GO TO 2621
      write(6,13)Iocc(1),Iocc(2),iocc(6),iocc(5),
     1iocc(8),iocc(7),Iocc(3),Iocc(4)
      IINT=IINT+1
      itmp=itmp+1
 2621 CONTINUE
      write(6,*)'         isomers of this subdomain = ',itmp
      write(6,*)'    '
      write(6,*)' ************************************************'
```

**Table X.** Dilution Program Output: Depiction of Isomers of the Second Subdomain of [6,2,2,1]

```
   *  -  +  -  *   subdomain (embryo-skeleton)

   assignment of box numbers in infant skeleton

            #     #     #
            2     8     3
      # 1 *  6  +  7  * 4 #
                  5
                  #

         box number
                                  1 2 6 5 8 7 3 4
         occupancy
                                  0 0 1 0 0 1 0 0
                                  0 0 2 0 0 0 0 0
                                  0 0 1 1 0 0 0 0
                                  0 0 0 1 1 0 0 0
                                  0 0 0 2 0 0 0 0
                                  1 0 0 0 0 1 0 0
                                  1 0 1 0 0 0 0 0
                                  1 0 0 1 0 0 0 0
                                  1 0 0 0 0 0 1 0
                                  1 1 0 0 0 0 0 0
                                  2 0 0 0 0 0 0 0

         isomers of this subdomain =          11

   ************************************************
```

*Case 2: Embryo Skeleton C*–C*–C+.* The corresponding infant skeleton is




This must be diluted with $b = 2$ carbons of degree 2 to generate all the isomers of the first subdomain of the [6,2,2,1] domain. The boxes are indexed arbitrarily as shown above.

**Table XI.** Tabulation of Isomeric Alkanes, for $n$ = 4–12, by DD Domains and Subdomains

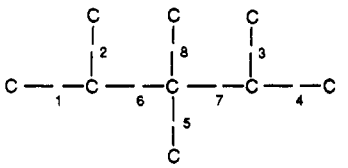
| | | subdomain | | domain | | | | subdomain | | domain | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | DD domain [$a$,$b$,$c$,$d$] | embryo skeleton | no. of isomers | no. of isomers | no. of isomers | $n$ | DD domain [$a$,$b$,$c$,$d$] | embryo skeleton | no. of isomers | no. of isomers | no. of isomers |
| 4 | | | | | 2 | 11 | | | | | 159 |
| | [2,2,0,0] | naught | 1 | 1 | | | [2,9,0,0] | naught | 1 | 1 | |
| | [3,0,1,0] | C* | 1 | 1 | | | [3,7,1,0] | C* | 8 | 8 | |
| 5 | | | | | 3 | | [4,6,0,1] | $C^+$ | 9 | 9 | |
| | [2,3,0,0] | naught | 1 | 1 | | | [4,5,2,0] | C*–C* | 27 | 27 | |
| | [3,1,1,0] | C* | 1 | 1 | | | [5,4,1,1] | C*–$C^+$ | 33 | 33 | |
| | [4,0,0,1] | $C^+$ | 1 | 1 | | | [6,3,0,2] | $C^+$–$C^+$ | 10 | 10 | |
| 6 | | | | | 5 | | [5,3,3,0] | C*–C*–C* | 24 | 24 | |
| | [2,4,0,0] | naught | 1 | 1 | | | [6,2,2,1] | C*–C*–$C^+$ | 17 | 28 | |
| | [3,2,1,0] | C* | 2 | 2 | | | | C*–$C^+$–C* | 11 | | |
| | [4,1,0,1] | $C^+$ | 1 | 1 | | | [7,1,1,2] | C*–$C^+$–$C^+$ | 5 | 8 | |
| | [4,0,2,0] | C*–C* | 1 | 1 | | | | $C^+$–C*–$C^+$ | 3 | | |
| 7 | | | | | 9 | | [8,0,0,3] | $C^+$–$C^+$–$C^+$ | 1 | 1 | |
| | [2,5,0,0] | naught | 1 | 1 | | | [6,1,4,0] | C*–C*–C*–C* | 4 | 6 | |
| | [3,3,1,0] | C* | 3 | 3 | | | | C*(–C*)$_3$ | 2 | | |
| | [4,2,0,1] | $C^+$ | 2 | 2 | | | [7,0,3,1] | C*–C*–C*–$C^+$ | 1 | 4 | |
| | [4,1,2,0] | C*–C* | 2 | 2 | | | | C*–C*–$C^+$–C* | 1 | | |
| | [5,0,1,1] | C*–$C^+$ | 1 | 1 | | | | $C^+$–C*(–C*)$_2$ | 1 | | |
| 8 | | | | | 18 | | | $C^+$(–C*)$_3$ | 1 | | |
| | [2,6,0,0] | naught | 1 | 1 | | 12 | | | | | 355 |
| | [3,4,1,0] | C* | 4 | 4 | | | [2,10,0,0] | naught | 1 | 1 | |
| | [4,3,0,1] | $C^+$ | 3 | 3 | | | [3,8,1,0] | C* | 10 | 10 | |
| | [4,2,2,0] | C*–C* | 5 | 5 | | | [4,7,0,1] | $C^+$ | 11 | 11 | |
| | [5,1,1,1] | C*–$C^+$ | 3 | 3 | | | [4,6,2,0] | C*–C* | 43 | 43 | |
| | [6,0,0,2] | $C^+$–$C^+$ | 1 | 1 | | | [5,5,1,1] | C*–$C^+$ | 58 | 58 | |
| | [5,0,3,0] | C*–C*–C* | 1 | 1 | | | [6,4,0,2] | $C^+$–$C^+$ | 20 | 20 | |
| 9 | | | | | 35 | | [5,4,3,0] | C*–C*–C* | 55 | 55 | |
| | [2,7,0,0] | naught | 1 | 1 | | | [6,3,2,1] | C*–C*–$C^+$ | 46 | 73 | |
| | [3,5,1,0] | C* | 5 | 5 | | | | C*–$C^+$–C* | 27 | | |
| | [4,4,0,1] | $C^+$ | 5 | 5 | | | [7,2,1,2] | C*–$C^+$–$C^+$ | 18 | 28 | |
| | [4,3,2,0] | C*–C* | 9 | 9 | | | | $C^+$–C*–$C^+$ | 10 | | |
| | [5,2,1,1] | C*–$C^+$ | 8 | 8 | | | [8,1,0,3] | $C^+$–$C^+$–$C^+$ | 3 | 3 | |
| | [6,1,0,2] | $C^+$–$C^+$ | 2 | 2 | | | [6,2,4,0] | C*–C*–C*–C* | 17 | 24 | |
| | [5,1,3,0] | C*–C*–C* | 3 | 3 | | | | C*(–C*)$_3$ | 7 | | |
| | [6,0,2,1] | C*–C*–$C^+$ | 1 | 2 | | | [7,1,3,1] | C*–C*–C*–$C^+$ | 7 | 21 | |
| | | C*–$C^+$–C* | 1 | | | | | C*–C*–$C^+$–C* | 7 | | |
| 10 | | | | | 75 | | | C*–C*(–C*)$_2$ | 4 | | |
| | [2,8,0,0] | naught | 1 | 1 | | | | $C^+$(–C*)$_3$ | 3 | | |
| | [3,6,1,0] | C* | 7 | 7 | | | [8,0,2,2] | C*–C*–$C^+$–$C^+$ | 1 | 6 | |
| | [4,5,0,1] | $C^+$ | 6 | 6 | | | | C*–$C^+$–C*–$C^+$ | 1 | | |
| | [4,4,2,0] | C*–C* | 17 | 17 | | | | C*–$C^+$–$C^+$–C* | 1 | | |
| | [5,3,1,1] | C*–$C^+$ | 17 | 17 | | | | $C^+$–C*–C*–$C^+$ | 1 | | |
| | [6,2,0,2] | $C^+$–$C^+$ | 5 | 5 | | | | C*–C*(–$C^+$)$_2$ | 1 | | |
| | [5,2,3,0] | C*–C*–C* | 10 | 10 | | | | $C^+$–$C^+$(–C*)$_2$ | 1 | | |
| | [6,1,2,1] | C*–C*–$C^+$ | 5 | 8 | | | [7,0,5,0] | C*–C*–C*–C*–C* | 1 | 2 | |
| | | C*–$C^+$–C* | 3 | | | | | C*–C*–C*(–C*)$_2$ | 1 | | |
| | [7,0,1,2] | C*–$C^+$–$C^+$ | 1 | 2 | | | | | | | |
| | | $C^+$–C*–$C^+$ | 1 | | | | | | | | |
| | [6,0,4,0] | C*–C*–C*–C* | 1 | 2 | | | | | | | |
| | | C*(–C*)$_3$ | 1 | | | | | | | | |

This infant skeleton exhibits five kinds of boxes. Boxes 1 and 2 are exchangeable, boxes 3–5 are unique and boxes 6–8 are exchangeable. We choose the hierarchies to be as follows: box 1 ≥ box 2 and box 6 ≥ box 7 ≥ box 8.

In the assignment of carbons of degree 2 to the boxes, the hierarchies must be obeyed and the total occupancy must be $b$, where $b$ = 2 in this case.

Table VII presents the dilution program for this infant skeleton. Table VIII shows the computer printout with the depiction of box occupancies for each of the seventeen expansion skeletons (isomers). As in case 1, the dilution program achieves each of the desired box hierarchies by nesting the DO loops, with the limits of the inner loops determined by the occupancies of higher-ranking boxes. The program checks that "ISUM", the total occupancy of all boxes in the infant skeleton, is equal to $b$ = 2, the number of carbons of degree 2 in the DD domain.

*Case 3*: *Embryo Skeleton C*–C$^+$–C**. The corresponding infant skeleton is



This must be diluted with $b$ = 2 carbons of degree 2 to generate all the isomers of the second subdomain of the [6,2,2,1] domain.

Boxes 1 and 2 are exchangeable; let box 1 ≥ box 2. Boxes 3 and 4 are exchangeable; let box 3 ≥ box 4. Boxes 5 and 8 are exchangeable; let box 5 ≥ box 8.

If the occupancy of box 1 equals the occupancy of box 3 *and* the occupancy of box 2 equals the occupancy of box 4, then boxes 6 and 7 are exchangeable, in which case we choose the hierarchy to be box 6 ≥ box 7. In the absence of the aforementioned condition, boxes 6 and 7 are unique.

Table IX presents the dilution program for this infant skeleton. Table X shows the printout depicting box occu-

**Table XII.** Tabulation of LC (the Longest Chain) for the 355 $C_{12}$ Alkanes Arranged by DD Domains and Subdomains

| DD domain | subdomain | | no. of isomers for given LC | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| [a,b,c,d] | embryo skeleton | isomers | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| [2,10,0,0] | naught | 1 | | | | | | | | 1 |
| [3,8,1,0] | C* | 10 | | | | | 2 | 3 | 5 | |
| [4,7,0,1] | C+ | 11 | | | 1 | 3 | 3 | 4 | | |
| [4,6,2,0] | C*–C* | 43 | | | 1 | 9 | 17 | 16 | | |
| [5,5,1,1] | C*–C+ | 58 | | 1 | 14 | 22 | 21 | | | |
| [6,4,0,2] | C+–C+ | 20 | | 4 | 7 | 9 | | | | |
| [5,4,3,0] | C*–C*–C* | 55 | | | 10 | 26 | 19 | | | |
| [6,3,2,1] | C*–C*–C+ | 46 | | 3 | 23 | 20 | | | | |
| | C*–C+–C* | 27 | | 4 | 13 | 10 | | | | |
| [7,2,1,2] | C*–C+–C+ | 18 | 1 | 7 | 10 | | | | | |
| | C+–C*–C+ | 10 | | 4 | 6 | | | | | |
| [8,1,0,3] | C+–C+–C+ | 3 | 1 | 2 | | | | | | |
| [6,2,4,0] | C*–C*–C*–C* | 17 | | 1 | 7 | 9 | | | | |
| | C*(–C*)$_3$ | 7 | | 1 | 6 | | | | | |
| [7,1,3,1] | C*–C*–C*–C+ | 7 | | 2 | 5 | | | | | |
| | C*–C*–C+–C* | 7 | | 2 | 5 | | | | | |
| | C+–C*(–C*)$_2$ | 4 | | 4 | | | | | | |
| | C+(–C*)$_3$ | 3 | 1 | 2 | | | | | | |
| [8,0,2,2] | C*–C*–C+–C+ | 1 | | 1 | | | | | | |
| | C*–C+–C*–C+ | 1 | | 1 | | | | | | |
| | C*–C+–C+–C* | 1 | | 1 | | | | | | |
| | C+–C*–C*–C+ | 1 | | 1 | | | | | | |
| | C*–C*(–C+)$_2$ | 1 | 1 | | | | | | | |
| | C+–C+(–C*)$_2$ | 1 | 1 | | | | | | | |
| [7,0,5,0] | C*–C*–C*–C*–C* | 1 | | | 1 | | | | | |
| | C*–C*–C*(–C*)$_2$ | 1 | | 1 | | | | | | |
| Total | | | 5 | 42 | 109 | 108 | 62 | 23 | 5 | 1 |

pancies for each of the eleven expansion skeletons (isomers). The occupancies of the exchangeable boxes (box 1/box 2, box 3/box 4, and box 5/box 8) are tabulated using nested iterative loops, as in the previous cases. For each possible expansion skeleton, the program monitors the occupancy of all the boxes. If the occupancy of box 1 equals that of box 3 and the occupancy of box 2 equals that of box 4, then the program sets the variable ISYM1 to 1; otherwise ISYM1 is set to 0. If ISYM1 = 1, the program forces boxes 6 and 7 to be exchangeable and enforces the hierarchy box 6 ≥ box 7.

A dilution program can be written for any infant skeleton. The cases discussed here illustrate exchangeable and unique boxes as well as boxes that are either exchangeable or unique depending on conditions. Clearly, our dilution programs are operative with any number, *b*, of carbons of degree 2.

## RESULTS OF ISOMER ENUMERATION

Table XI lists the results obtained by our method for alkanes with *n* up to 12. For each *n* we show the isomer counts for subdomains, the isomer counts for domains, and the total isomer count. The total isomer counts agree with literature values.[4,6,9a,10b,14a,c]

Inspection of Table XI enables one to make semiquantitative statements which, from the vantage point of the DD method, account for the ever-increasing steepness of the growth curve relating total isomer count and *n*.

(1) It is evident from Table XI as well as from Table IV (or its easily constructed extension) that an increase in *n* produces a large increase in the number of permitted embryo compositions, hence in the number of DD domains.

(2) Since the additional embryo compositions are of increasing embryo size, there is a dramatic increase in the number of embryo skeletons per embryo composition, hence in the number of subdomains per DD domain.

(3) The larger the value of *n*, the larger is the number of carbons of degree 2 (*b*) required in the dilution of an infant skeleton. With a large *b* there are many more "dilution" possibilities and hence many more expansion skeletons, i.e. isomers, than with a small *b*. Table XI shows that the greater the number of unique boxes contained in the infant skeleton, the more dramatic is the mushrooming of expansion skeletons. The table also shows that this effect is especially strong if boxes are unconditionally unique.

## LONGEST-CHAIN LENGTH

An important characteristic of a molecule is its longest chain, LC. The number of carbon atoms from one end of the LC to the other is featured in the systematic name of the molecule. The number of bonds in the LC is called the diameter.[6] This is one less than the number of carbon atoms.

The LC of any isomer generated by our method can be determined either by manual means or by an extension of the dilution program responsible for the subdomain to which the isomer belongs. For each pair of carbons of degree 1 there exists a certain pathway through the molecule. All the pathways are measured so as to yield LC. If the isomer represents an infant skeleton (i.e. *b* = 0 so that all boxes are vacant), then the LC traverses carbons that are exclusively of degree 3 and/or 4. If the isomer represents an expansion skeleton, then the nature of the underlying infant skeleton, the number of carbons of degree 2 (*b*), and the distribution of these carbons among the boxes influence the identity of LC. An LC in an expansion skeleton either traverses carbons that are exclusively of degree 3 and/or 4 or it also includes one or more boxes containing carbons of degree 2. The latter situation is considerably more common.

We have determined LC for all isomeric alkanes with *n* up to 12. For each *n* the LC distribution found by our method matches that reported by Balaban and co-workers.[6] The results for *n* = 12 are shown in Table XII.

## CONCLUDING REMARKS

We have demonstrated the utility of degree distribution for the generation, visualization, and enumeration of isomers in the alkane series. Our work is being extended in a number of directions.

(a) The codification of structures by linear notations is always a worthwhile objective. Since the DD method develops all of the isomers of a subdomain by expansion of a certain infant skeleton, it would be most expedient if the notations for the expansion skeletons could be easily related to the notation for the infant skeleton and hence also to each other. In a future paper we shall show that the $N$-tuple code[14] is well suited for this purpose.

(b) It should be noted that in any structure generated by our method only the carbon atoms originating from its embryo skeleton can possibly be asymmetric. This fact will greatly facilitate the counting of stereoisomers.

(c) Self-developing dilution programs are desirable to make our method more efficient, especially when applied to infant skeletons that are large and therefore complex.

(d) Starting with infant skeletons, we shall apply Pólya's theorem[7] to the nonconstructive isomer enumeration for the corresponding subdomains.

(e) Since all the isomers of a subdomain are based on the same infant skeleton, they show the same branching pattern. Hence a comparative study of certain properties of such isomers should be of interest.

(f) Our work will, of course, be extended to other classes of compounds.

**Acknowledgment** is made to the Donors of the Petroleum Research Fund, administered by the American Chemical Society, for the partial support of this research.

## REFERENCES AND NOTES

(1) Bieber, T. I.; Jackson, M. D. Applications of Degree Distribution. 1. (a) General Discussion and Computer Generation of Degree Distributions. (b) Maximal Degree Distributions. *J. Chem. Inf. Comput. Sci.*, previous paper in this issue.

(2) (a) Cayley, A. On the Theory of Analytical Forms Called Trees. *Philos. Mag.* **1857**, *13*, 172–176. [*Collected Mathematical Papers of A. Cayley*; Cambridge University: Cambridge, U.K., 1889–1898; Vol. 3, pp 242–246.] (b) Cayley, A. On the Analytical Forms Called Trees-Part II. *Philos. Mag.* **1859**, *18*, 374–378. [*Collected Mathematical Papers of A. Cayley*; Cambridge University: Cambridge, U.K., 1889–1898; Vol. 4, pp 112–115.] (c) Cayley, A. On the Mathematical Theory of Isomers. *Philos. Mag.* **1874**, 447, 444–447. [*Collected Mathematical Papers of A. Cayley*; Cambridge University: Cambridge, U.K., 1889–1898; Vol. 9, pp 202–204.] (d) Cayley, A. On the Analytical Figures Called Trees in Mathematics and Their Applications to the Theory of Chemical Combinations. *Ber. Dtsch. Chem. Ges.* **1875**, *8*, 1056–1059. (e) Cayley, A. On the Analytical Forms Called Trees, with Application to the Theory of Chemical Combinations. *Rep. Brit. Assoc. Adv. Sci.* **1875**, 257–305. [*Collected Mathematical Papers of A. Cayley*; Cambridge University: Cambridge, U.K., 1889–1898; Vol. 9, pp 427–460.] (f) Cayley, A. On the Analytical Forms Called Trees. *Am. J. Math.* **1881**, *4*, 266–268. [*Collected Mathematical Papers of A. Cayley*; Cambridge University: Cambridge, U.K., 1889–1898; Vol. 11, pp 365–367.]

(3) (a) Henze, H. R.; Blair, C. M. The Number of Structurally Isomeric Alcohols of the Methanol Series. *J. Am. Chem. Soc.* **1931**, *53*, 3042–3046. (b) Henze, H. R.; Blair, C. M. The Number of Isomeric Hydrocarbons of the Methane Series. *J. Am. Chem. Soc.* **1931**, *53*, 3077–3085. (c) Blair, C. M.; Henze, H. R. The Number of Stereoisomeric and Non-Stereoisomeric Mono-substituted Products of the Paraffins. *J. Am. Chem. Soc.* **1932**, *54*, 1098–1106. (d) Blair, C. M.; Henze, H. R. The Number of Stereoisomeric and Non-Stereoisomeric Paraffin Hydrocarbons. *J. Am. Chem. Soc.* **1932**, *54*, 1538–1545. (e) Henze, H. R.; Blair, C. M. The Number of Structurally Isomeric Hydrocarbons of the Ethylene Series. *J. Am. Chem. Soc.* **1933**, *55*, 680–686. (f) Henze, H. R.; Blair, C. M. The Number of Structural Isomers of the More Important Types of Aliphatic Compounds. *J. Am. Chem. Soc.* **1934**, *56*, 157. (g) Coffman, D. D.; Blair, C. M.; Henze, H. R. The Number of Structurally Isomeric Hydrocarbons of the Acetylene Series. *J. Am. Chem. Soc.* **1933**, *55*, 252–253.

(4) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1983; Vol. II, pp 147–154.

(5) (a) Davis, C. C.; Cross, K.; Ebel, M. Computer Calculation of Alkane Isomers. *J. Chem. Educ.* **1971**, *48*, 675. (b) Ebert, K.; Ederer, H.; Isenhour, T. L. *Computer Applications in Chemistry*; VCH Publishers: New York, 1989; pp 256–264.

(6) Balaban, A. T.; Kennedy, J. W.; Quintas, L. V. The Number of Alkanes Having n Carbons and a Longest Chain of Length d. *J. Chem. Educ.* **1988**, *65*, 304–313.

(7) (a) Pólya, G.; Read, R. C. *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*; Springer-Verlag: New York, 1987. Contains an English translation of Pólya's paper: Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.* **1937**, *68*, 145–254. (b) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1983; Vol. II, pp 152–159.

(8) For a discussion of computer programs for the generation of isomers and for additional references, see: Balaban, A. T. In *Chemical Graph Theory*; Bonchev, D., Rouvray, D. H., Eds.; Gordon and Breach: New York, 1991; Vol. 1, Chapter 5, pp 219–222.

(9) (a) Lederberg, J.; Sutherland, G. L.; Buchanan, B. G.; Feigenbaum, E. A.; Robertson, A. V.; Duffield, A. M.; Djerassi, C. Applications of Artificial Intelligence for Chemical Inference. I. The Number of Possible Organic Compounds. Acyclic Structures Containing C, H, O, and N. *J. Am. Chem. Soc.* **1969**, *91*, 2973–2976. (b) Masinter, L. M.; Sridharan, N. S.; Lederberg, J.; Smith, D. H. Applications of Artificial Intelligence for Chemical Inference. XII. Exhaustive Generation of Cyclic and Acyclic Isomers. *J. Am. Chem. Soc.* **1974**, *96*, 7702–7714.

(10) (a) Hendrickson, J. B.; Toczko, A. G. Unique Numbering and Cataloguing of Molecular Structures. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 171–177. (b) Hendrickson, J. B.; Parks, C. A. Generation and Enumeration of Carbon Skeletons. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 101–107.

(11) Kvasnička, V.; Pospichal, J. Canonical Indexing and Constructive Enumeration of Molecular Graphs. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 99–105.

(12) (a) Schultz, H. P. Topological Organic Chemistry. I. Graph Theory and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 227–228. (b) Schultz, H. P.; Schultz, T. P. Topological Organic Chemistry. 3. Graph Theory, Binary and Decimal Adjacency Matrices, and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 144–147.

(13) Burden, F. R. Molecular Identification Number for Substructure Searches. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 225–227.

(14) (a) Knop, J. V.; Müller, W. R.; Jeričević, Ž.; Trinajstić, N. Computer Enumeration and Generation of Trees and Rooted Trees. *J. Chem. Inf. Comput. Sci.* **1981**, *21*, 91–99. (b) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. Molecular Topological Index. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 160–163. (c) Trinajstić, N.; Nikolić, S.; Knop, J. V.; Müller, W. R.; Szymanski, K. *Computational Chemical Graph Theory*; Ellis Horwood: New York, 1991; Chapters 2 and 3. See especially Table III.8, pp 140–141.

(15) Davies, R. E.; Freyd, P. J. $C_{167}H_{336}$ Is the Smallest Alkane with More Realizable Isomers than the Observed Universe Has "Particles". *J. Chem. Educ.* **1989**, *66*, 278–281.

(16) See: Reference 14c, p 27.

(17) See: Reference 9b, p 7703.

(18) See: Reference 10b, p 103.