Chemical Graph Theory. VII.1 Enumeration and Generation of the Non-ionic/Radical Members of Cyanopolyyne Family

Jan V. Knop and Klaus Szymanski

Computer Centre, The University of Düsseldorf, 4000 Düsseldorf, Federal Republic of Germany,

and

Harold W. Kroto

School of Molecular Sciences, The University of Sussex, Brighton BN1 9QJ, England, UK,

and

Nenad Trinajstić

The Ruđer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia

Received July 8, 1985

Algorithms for the enumeration and generation of non-ionic/radical members of the cyanopolyyne family are developed. In this work we wish to present the enumeration procedure and the generation procedure for non-ionic/radical members of the cyanopolyyne family. There is an abundance of the linear cyanopolyynes observed in interstellar space.2,3 The astrophysicists and astrochemists would like to know how many cyanopolyynes and related chains can be derived for a given composition.4 Since we have some experience in developing the enumeration and generation algorithms for several classes of molecules, we set ourselves to devise schemes for the enumeration and generation of non-ionic/radical cyanopolyynes and related chains.5,9

ENUMERATION PROCEDURE

The bare skeleton of non-ionic/radical members of the cyanopolyyne family may be considered as a non-branched (hydrogen-suppressed) tree with single, double, and triple bonds.10 We will first enumerate these structures of general formula:

\[ H_nC_6N; \ h = 1, 3, 5, \ldots \]
\[ n = 1, 2, 3, \ldots \]

This enumeration can be done in an inductive way. For one carbon atom there are possible 3 structures shown below.
Let the structure consist of a chain of $n$ carbon atoms. We now consider the possible bonds starting at the first carbon atom (whereas the last carbon atom is linked to the nitrogen atom). The first carbon atom may be linked to the neighbouring carbon atom with a single, double, or triple bond.

The set of considered structures may be divided into disjoined classes $C(b, h, n)$ where $b (1 \leq b \leq 3)$ represents a single, double, or triple bond between the first carbon atom and the adjacent carbon atom in the considered structure, $h$ is the total number of hydrogen atoms in the structure, and $n$ is the number of carbon atoms.

Let the structure be an element of $C(1, h - 2, n - 1)$, if it ends as

If we connect one carbon atom to the chain with the above fragment with a single bond, we obtain

If we connect a carbon atom to the chain with a double bond, we obtain

Finally, if we connect a carbon atom to the chain with a triple bond, we obtain

Let the structure be an element of $C(2, h - 2, n - 1)$ if it ends as

If we connect one carbon atom to the chain, containing the above fragment, with a single bond, we obtain
If we connect a carbon atom to a chain with a double bond, we obtain

\[
\text{H} \quad \text{C}=\text{C} \quad \ldots \quad \in \text{C}(2, h-2, n)
\]

In this case it is not possible to connect a carbon atom to a chain with fragment 2 with a triple bond.

Let the structure be an element of \( \text{C} (3, h-2, n-1) \) if it ends as

\[
\text{H} \quad \text{C} \equiv \text{C} \quad \ldots
\]

We can connect a carbon atom to a chain containing the above fragment only with a single bond

\[
\text{H} \quad \text{C}=\text{C} \quad \ldots \quad \in \text{C}(1, h, n-1).
\]

Now it can be easily seen that

\[
\begin{align*}
| \text{C} (1, h, n) | &= | \text{C} (1, h-2, n-1) | + | \text{C} (2, h-2, n-1) | + | \text{C} (3, h-2, n-1) | \\
| \text{C} (2, h, n) | &= | \text{C} (1, h, n-1) | + | \text{C} (2, h, n-1) | \\
| \text{C} (3, h, n) | &= | \text{C} (1, h+2, n-1) |
\end{align*}
\]

Since we know that

\[
| \text{C} (3, 1, 1) | = | \text{C} (2, 3, 1) | = | \text{C} (1, 5, 1) | = 1
\]

we can compute the number of elements in the other classes by a simple algorithm. Let us first present the above result in a different way.

(1) Number of carbon atoms: \( n = 1 \)

<table>
<thead>
<tr>
<th>b</th>
<th>h</th>
<th>-1</th>
<th>1</th>
<th>3</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Number of structures

| 0 | 1 | 1 | 1 |

All other classes can be enumerated by an algorithm schematically presented below.
(I) Number of carbon atoms: \( n \)

<table>
<thead>
<tr>
<th>( b )</th>
<th>(-1)</th>
<th>1</th>
<th>( h-1)</th>
<th>( h)</th>
<th>( h+1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>a</td>
<td>d</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>b</td>
<td>e</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>c</td>
<td>f</td>
</tr>
</tbody>
</table>

Number of structures: 

(II) Number of carbon atoms: \( n+1 \)

<table>
<thead>
<tr>
<th>( b )</th>
<th>(-1)</th>
<th>1</th>
<th>( h-1)</th>
<th>( h)</th>
<th>( h+1)</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>a+b+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>d+e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>f</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of structures: 

(2) Number of carbon atoms: \( n = 2 \)

<table>
<thead>
<tr>
<th>( b )</th>
<th>(-1)</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of structures: 0 0 3 2 1

(3) Number of carbon atoms: \( n = 3 \)

<table>
<thead>
<tr>
<th>( b )</th>
<th>(-1)</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of structures: 0 1 3 6 3 1

(4) Number of carbon atoms: \( n = 4 \)

<table>
<thead>
<tr>
<th>( b )</th>
<th>(-1)</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Number of structures: 0 0 6 10 10 4 1
(5) Number of carbon atoms: \( n = 5 \)

<table>
<thead>
<tr>
<th>( h )</th>
<th>(-1)</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>8</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of structures: 0, 1, 6, 20, 22, 15, 5, 1

(6) Number of carbon atoms: \( n = 6 \)

<table>
<thead>
<tr>
<th>( h )</th>
<th>(-1)</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>13</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>20</td>
<td>22</td>
<td>15</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>14</td>
<td>19</td>
<td>14</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of structures: 0, 0, 10, 30, 49, 40, 21, 6, 1

The regularities in the above results allowed us to derive the recursive formula for counting the structures with chemical formula \( H_3C_5N \).

Let \( N(h, n) \) be the number of structures for the cyanopolyyne family with \( h \) hydrogen atoms and \( n \) carbon atoms, respectively. Then

(a) \( N(1, 1) = N(3, 1) = N(5, 1) = 1 \)
\[ N(h, 1) = 0 \text{ for } h > 5 \text{ and } h < 1. \]

(b) \( N(1, 2) = 0 \)
\[ N(3, 2) = 3 \]
\[ N(5, 2) = 2 \]
\[ N(7, 2) = 1 \]
\[ N(h, 3) = 0 \text{ for } h > 7. \]

In general,

\[ A(h) = \begin{cases} 1 & \text{for } h = 3, 5 \\ 0 & \text{otherwise} \end{cases} \]

Then the recursive formula for counting is given by:

\[ N(h, n) = A(h) + N(h, n - 2) + \sum_{i=1}^{n-1} N(h - 2, i) \text{ for } n > 2, \ h = 2k + 1, \ 0 \leq k \leq n + 1. \]

**Proof**

We define

\[ N(h, n) = |C(1, h, n)| + |C(2, h, n)| + |C(3, h, n)|. \]

We know the following relationships

\[ |C(1, h, n)| = |C(1, h - 2, n - 1)| + |C(2, h - 2, n - 1)| + |C(3, h - 2, n - 1)| \]
\[ |C(2, h, n)| = |C(1, h, n - 1)| + |C(2, h, n - 1)| \]
\[ |C(3, h, n)| = |C(1, h + 2, n - 1)| \]
By virtue of (1) and the above definition the following holds:

\[ N(h-2, n-1) = |C(1, h, n)| \quad (4) \]

Then we see that

\[ |C(2, h, n)| = |C(2, h, n-2)| + |C(1, h, n-1)| \quad (by \, virtue \, of \, (2)) \]

\[ = |C(2, h, n-2)| + N(h-2, n-2) \quad (by \, virtue \, of \, (4)) \]

\[ = |C(2, h, n-2)| + N(h-2, n-2) \quad (by \, virtue \, of \, (2)) \]

\[ = |C(2, h, n-2)| + N(h-2, n-2) + N(h-2, n-2) \quad (by \, virtue \, of \, (4)) \]

\[ \vdots \]

\[ \vdots \]

\[ = |C(2, h, 1)| \quad (by \, virtue \, of \, (3)) \]

Since we know that

\[ |C(2, 3, 1)| = |C(1, 5, 1)| = 1 \]

\[ |C(2, h, 1)| = |C(1, h, 1)| = 0 \quad \text{for all other values of} \, h \]

we can define

\[ A(h) = \begin{cases} 1 & \text{for} \, h = 3, 5 \\ 0 & \text{otherwise} \end{cases} \]

and get

\[ |C(2, h, n)| = A(h) + \sum_{i=1}^{n-2} N(h-2, i) \]

\[ |C(3, h, n)| = C(1, h + 2, n - 1) \quad (by \, virtue \, of \, (3)) \]

\[ = N(h, n - 2) \quad (by \, virtue \, of \, (4)). \]
Finally we obtain
\[ N(h, n) = N(h - 2, n - 1) + A(h) + \sum_{i=1}^{n-2} N(h-2, i) + N(h, n - 2) \]
or
\[ N(h, n) = A(h) + N(h, n - 2) + \sum_{i=1}^{n-1} N(h-2, i) . \]
This completes the proof.

The numbers of isomeric structures for \( H_hC_nN \) species are given in the Table. \( H_hC_nN \) species with \( h = 1, 2, \ldots, 41 \) and \( n = 1, 2, \ldots, 19 \), respectively are reported.

**GENERATION OF STRUCTURES**

In this section we present a scheme for generating all the \( H_hC_nN \) species. Let the carbon atoms of the chain be labelled as follows

\[ \text{N-C}_1-\text{C}_2-\text{C}_3-\ldots-\text{C}_n \]

Let \( b_i \) be the number of bonds between \( \text{N} \) and \( \text{C}_i \), and \( b_i \) the number of connections between \( \text{C}_i \) and \( \text{C}_{i-1} \). Then the following relationships hold:

If \( b_i = 1 \) then \( b_{i+1} \in \{1, 2, 3\} \)

If \( b_i = 2 \) then \( b_{i+1} \in \{1, 2\} \)

If \( b_i = 3 \) then \( b_{i+1} \in \{1\} \)

Now we can generate all sequences

\[(b_1, b_2, b_3, \ldots, b_n)\]
with 

\[ b_i \in \{1, 2, 3\} \]

and 

\[ b_i (2 \leq i \leq k) \]

obeying conditions (5)—(7). These sequences represent all non-ionic/radical cyanopolyene structures with \( n \) carbon atoms and all possible bonds between them. The next we have to decide how many hydrogen atoms belong to each structure, and then we will know how to assign the appropriate chemical formula to a generated structure.

Let \( h_i (1 \leq i \leq k) \) be the number of hydrogen atoms linked to the carbon atom labelled by \( i \). Then,

\[ h_k = 3 - b_i \quad \text{(the number of hydrogen atoms linked to N)} \]

\[ h_i = 4 - b_i - b_{i+1} \quad (1 \leq i \leq k - 1) \]

\[ h_k = 4 - b_k \]

and

\[ h = \sum_{j=0}^{k} h_j. \]

The structure belongs to the chemical formula \( \text{H}_h\text{C}_n\text{N} \).

**REMARK**

The sequences can be generated by a rooted generating tree with the weighted vertices. The weights are the numbers \( b_i \). The rooted generating tree is given in the Figure 1.

![Figure 1. The rooted generating tree](image)

The paths of length \( k + 1 \), starting at the root and ending at terminal vertices, represent sequences (or in translation \( \text{H}_h\text{C}_n\text{N} \) structures). For example, a sequence 

\[(N, 2, 1, 3)\]

is identical to the structure given below:

\[ \text{H} - \text{N} = \text{C} - \text{C} = \text{C} - \text{H} \]

\[ \text{H} \]
which is obtained after the appropriate translation rules from above were applied. A FORTRAN program is written which allows the generation of all non-ionic/radical members of cyanopolyne family. This algorithm, of course, also enumerates these structures.

REFERENCES
18. The computer program may be obtained on request.

SAŽETAK
Kemijska teorija crteža. VII. Prebrojavanje i generiranje ne-ionskih radikala iz obitelji cijanopolina
J. V. Knop, K. Szymanski, H. W. Kroto i N. Trinajstić

Prikazani su algoritmi za prebrojavanje i generiranje ne-ionskih radikala iz obitelji cijanopolina.