CROATICA CHEMICA ACTA CCACAA 59 (2) 359—367 (1986)

YU ISSN 0011-1643
CCA-1648 UDC 541
Original Scientific Paper

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 Rugjer 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.

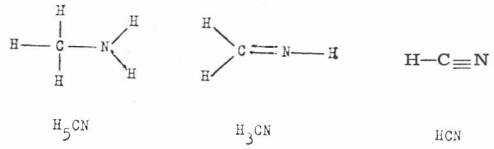
In this work we wish to present the enumeration procedure and the generation procedure for non-ionic/radical members of the cyanopolyyne family. There is an abundance of the linear cyanopolyynes observed in interstellar space.[2,3] The astrophysicists and astrochemists would like to know how many cyanopolyynes 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

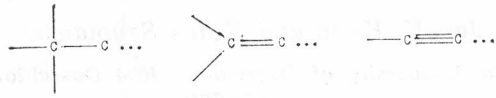
The bare skeleton of non-ionic/radical members of the cyanopolyyne family may be considered as a non-branched (hydrogen-suppressed) tree with single, double, and triple bonds.[10] We will first enumerate these structures of general formula:

$$H_h C_n N;\ h = 1, 3, 5, \ldots$$
$$n = 1, 2, 3, \ldots$$

This enumeration can be done in an inductive way. For *one* carbon atom there are possible 3 structures shown below.

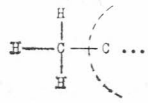


Let the structure consist of a chain of n carbon atoms. We now consider the possible bonds starting at the first carbon atom (whereas the last carbon atom is linked to the nitrogen atom). The first carbon atom may be linked to the neighbouring carbon atom with a single, double, or triple bond.



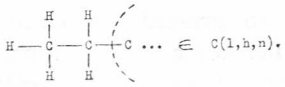
The set of considered structures may be divided into disjoined classes C (b, h, n) where b ($1 \leq b \leq 3$) represents a single, double, or triple bond between the first carbon atom and the adjacent carbon atom in the considered structure, h is the total number of hydrogen atoms in the structure,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

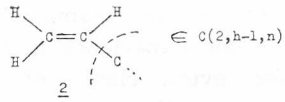


1

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

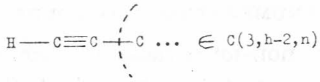


If we connect a carbon atom to the chain with a double bond, we obtain

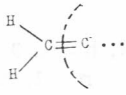


2

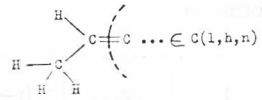
Finally, if we connect a carbon atom to the chain with a triple bond, we obtain



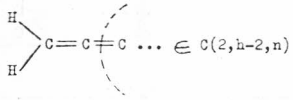
Let the structure be an element of C (2, h — 2, n — 1) if it ends as



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

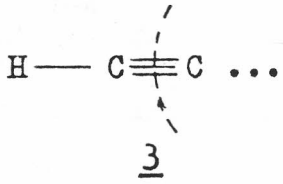


If we connect a carbon atom to a chain with a double bond, we obtain



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 C (3, h — 2, n — 1) if it ends as



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

Now it can be easily seen that

$$| C (1, h, n) | = | C (1, h—2, n—1) | + | C (2, h—2, n—1) | + | C (3, h—2, n—1) |$$

$$| C (2, h, n) | = | C (1, h, n—1) | + C | (2, h, n—1) |$$

$$| C (3, h, n) | = | C (1, h+2, n—1) |$$

Since we know that

$$| C (3, 1, 1) | = | C (2, 3, 1) | = | C (1, 5, 1) | = 1$$

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

b \ h	—1	1	3	5
1	0	0	0	1
2	0	0	1	0
3	0	1	0	0
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

b \ h	—1	1	.....	h—1	h	h+1
1			.....	a	d	f
2			.....	b	e	
3			.....	c		
Number of structures			.....			

(II) Number of carbon atoms: n + 1

b \ h	—1	1	.....	h—1	h	h+1	...
1			.....		a+b+c		...
2			.....		d+e		...
3			.....		f		...
Number of structures			.....				...

(2) Number of carbon atoms: n = 2

b \ h	—1	1	3	5	7
1	0	0	1	1	1
2	0	0	1	1	0
3	0	0	1	0	0
Number of structures	0	0	3	2	1

(3) Number of carbon atoms: n = 3

b \ h	—1	1	3	5	7	9
1	0	0	0	3	2	1
2	0	0	2	2	1	0
3	0	1	1	1	0	0
Number of structures	0	1	3	6	3	1

(4) Number of carbon atoms: n = 4

b \ h	—1	1	3	5	7	9	11
1	0	0	1	3	6	3	1
2	0	0	2	5	3	1	0
3	0	0	3	2	1	0	0
Number of structures	0	0	6	10	10	4	1

(5) Number of carbon atoms: n = 5

b \ h	—1	1	3	5	7	9	11	13
1	0	0	0	6	10	10	4	1
2	0	0	3	8	9	4	1	0
3	0	1	3	6	3	1	0	0
Number of structures	0	1	6	20	22	15	5	1

(6) Number of carbon atoms: n = 6

b \ h	—1	1	3	5	7	9	11	13	15
1	0	0	1	6	20	22	15	5	1
2	0	0	3	14	19	14	5	1	0
3	0	0	6	10	10	4	1	0	0
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_hC_nN.

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

(a) $N(1, 1) = N(3, 1) = N(5, 1) = 1$
$N(h, 1) = 0$ for $h > 5$ and $h < 1$.

(b) $N(1,2) = 0$
$N(3, 2) = 3$
$N(5, 2) = 2$
$N(7, 2) = 1$
$N(h, 2) = 0$ 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:

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

Proof

We define

$$N(h, n) = |C(1, h, n)| + |C(2, h, n)| + |C(3, h, n)|.$$

We know the following relationships

$$|C(1, h, n)| = |C(1, h-2, n-1)| + |C(2, h-2, n-1)| + |C(3, h-2, n-1)| \quad (1)$$

$$|C(2, h, n)| = |C(1, h, n-1)| + |C(2, h, n-1)| \quad (2)$$

$$|C(3, h, n)| = |C(1, h+2, n-1)| \quad (3)$$

TABLE

Cyanopolyynes and Related Chains H_hC_nN (h = 1, 3, 5, ..., 2n + 3)

n/h	1	3	5	7	9	11	13	15	17	19
1	1	1	1							
2	0	3	2	1						
3	1	3	6	3	1					
4	0	6	10	10	4	1				
5	1	6	20	22	15	5	1			
6	0	10	30	49	40	21	6	1		
7	1	10	50	91	100	65	28	7	1	
8	0	15	70	168	216	181	98	36	8	1
9	1	15	105	280	444	441	301	140	45	9
10	0	21	140	462	840	1001	812	470	192	55
11	1	21	196	714	1530	2101	2016	1386	699	255
12	0	28	252	1092	2640	4191	4628	3732	2232	1000
13	1	28	336	1596	4422	7931	10023	9276	6471	3432
14	0	36	420	2310	7128	14443	20566	21645	17280	10648
15	1	36	540	3234	11220	25311	40404	47755	43164	30360
16	0	45	660	4488	17160	43043	76258	100525	101728	80740
17	1	45	825	6072	25740	71071	139139	202596	228140	202180
18	0	55	990	8151	37752	114543	246064	394808	489600	480700
19	1	55	1210	10725	54483	180323	423488	743240	1010820	1091740

By virtue of (1) and the above definition the following holds:

$$N(h-2, n-1) = |C(1, h, n)| \tag{4}$$

Then we see that

$$\begin{aligned}
|C(2, h, n)| &= |C(2, h, n-1)| + |C(1, h, n-1)| \text{ (by virtue of (2))} \\
&= |C(2, h, n-1)| + N(h-2, n-2) \text{ (by virtue of (4))} \\
&= |C(2, h, n-2)| + |C(1, h, n-2)| + N(h-2, n-2) \text{ (by virtue of (2))} \\
&= |C(2, h, n-2)| + N(h-2, n-3) + N(h-2, n-2) \text{ (by virtue of (4))} \\
&\;\;\vdots \\
&= |C(2, h, 1)| + |C(1, h, 1)| + N(h-2, 1) + \ldots + N(h-2, n-2).
\end{aligned}$$

Since we know that

$$|C(2, 3, 1)| = |C(1, 5, 1)| = 1$$

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

we can define

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

and get

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

$$\begin{aligned}
|C(3, h, n)| &= C(1, h+2, n-1)| \text{ (by virtue of (3))} \\
&= N(h, n-2) \text{ (by virtue of (4)).}
\end{aligned}$$

TABLE (continued)

21	23	25	27	29	31	33	35	37	39	41
1										
10	1									
66	11	1								
330	78	12	1							
1386	418	91	13	1						
5082	1871	520	105	14	1					
16786	7293	2470	637	120	15	1				
50842	25532	10192	3199	770	136	16	1			
143284	81796	37674	13923	4075	920	153	17	1		
379922	243321	127192	54159	18648	5116	1088	171	18	1	
952666	679107	397995	192075	76112	24548	6341	1275	190	19	1

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

$$N—C_1—C_2—C_3—\ldots—C_n$$

Let b_1 be the number of bonds between N and C_1, and b_i the number of connections between C_i and C_{i-1}. Then the following relationships hold:

$$\text{If } b_i = 1 \text{ then } b_{i+1} \in \{1, 2, 3\} \tag{5}$$

$$\text{If } b_i = 2 \text{ then } b_{i+1} \in \{1, 2\} \tag{6}$$

$$\text{If } b_i = 3 \text{ then } b_{i+1} \in \{1\} \tag{7}$$

Now we can generate all sequences

$$(b_1, b_2, b_3, \ldots, b_k)$$

with

$$b_1 \in \{1, 2, 3\}$$

and

$$b_i \ (2 \leqslant i \leqslant k)$$

obeying conditions (5)—(7). These sequences represent all non-ionic/radical cyanopolyyne 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 $(1 \leqslant i \leqslant k)$ be the number of hydrogen atoms linked to the carbon atom labelled by i. Then,

$$h_0 = 3 - b_1 \text{ (the number of hydrogen atoms linked to N)}$$

$$h_i = 4 - b_i - b_{i+1} \ (1 \leqslant i \leqslant k - 1)$$

$$h_k = 4 - b_k$$

and

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

The structure belongs to the chemical formula H_hC_nN.

REMARK

The sequences can be generated by a rooted generating tree[11,12] with the weighted vertices.[13–17] The weights are the numbers b_i. The rooted generating tree is given in the Figure 1.

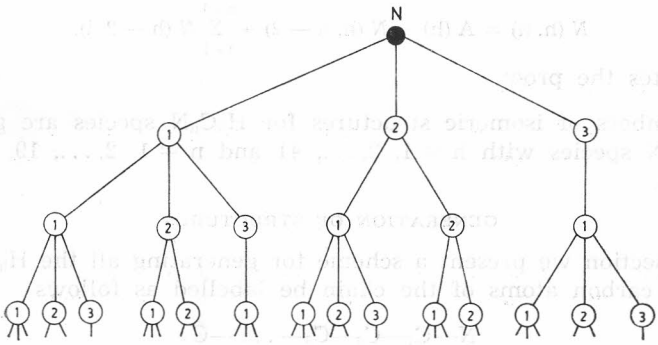
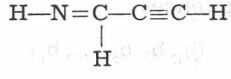


Figure 1. The rooted generating tree

The paths of length k + 1, starting at the root and ending at terminal vertices, represent sequences (or in translation H_hC_nN structures). For example, a sequence

(N, 2, 1,3)

is identical to the structure given below:



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 cyanopolyyne family.[18] This algorithm, of course, also enumerates these structures.

REFERENCES

1. Part VI: P. Křivka, R. B. Mallion, and N. Trinajstić, *Int. J. Quantum Chem.*, submitted.
2. H. W. Kroto, *Int. Revs. Phys. Chem.* **1** (1981) 309.
3. H. W. Kroto, *Chem. Soc. Revs.* **11** (1982) 435.
4. J. August, H. W. Kroto, and N. Trinajstić, work in preparation.
5. J. V. Knop, W. R. Müller, Ž. Jeričević, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **21** (1981) 91.
8. J. V. Knop, K. Szymanski, L. Klasinc, and N. Trinajstić, *J. Comput. Chem.* **4** (1983) 23.
7. N. Trinajstić, Ž. Jeričević, J. V. Knop, W. R. Müller, and K. Szymanski, *Pure Appl. Chem.* **55** (1983) 379.
8. J. V. Knop, K. Szymanski, L. Klasinc, and N. Trinajstić, *Comput. Chem.* **8** (1984) 107.
9. J. V. Knop, W. R. Müller, K. Szymanski, and N. Trinajstić, *Computer Generation of Certain Classes of Molecules*, SKTH, Zagreb, 1985.
10. N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, Florida, 1983.
11. D. E. Knuth, *The Art of Computer Programming*, Addison-Wesley, Reading, Mass., 1968, Vol. 1.
12. B. Džonova-Jerman-Blažič and N. Trinajstić, *Comput. Chem.* **6** (1982) 121.
13. R. B. Mallion, A. J. Schwenk, and N. Trinajstić, *Croat. Chem. Acta* **46** (1974) 71.
14. R. B. Mallion, N. Trinajstić, and A. J. Schwenk, *Z. Naturforsch.* **29a** (1974) 1481.
15. R. B. Mallion, A. J. Schwenk, and N. Trinajstić, in *Recent Advances in Graph Theory*, Edited by M. Fiedler, Academia, Prague, 1975, p. 345.
16. A. Graovac, O. E. Polansky, N. Trinajstić, and N. Tyutyulkov, *Z. Naturforsch.* **30a** (1975) 1696.
17. M. J. Rigby, R. D. Mallion, and A. C. Day, *Chem. Phys. Lett.* **51** (1977) 178; erratum *ibid.* **53** (1978) 418.
18. The computer program may be obtained on request.

SAŽETAK

Kemijska teorija crteža. VII. Prebrojavanje i generiranje ne-ionskih radikala iz obitelji cijanopoliina

J. V. Knop, K. Szymanski, H. W. Kroto i N. Trinajstić

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