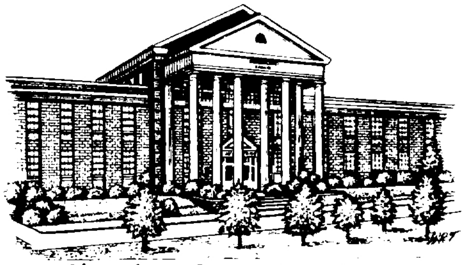

Physics Department

Southern College

Collegedale, Tennessee 37315
Telephone 615-238-2111

April 9, 1986

Joshua Lederberg, President
Rockefeller University
1230 York Ave.
New York, NY 10021

Dear Dr. Lederberg:

Some years ago when you were at Stanford University, you led out in what appears to me to be a landmark contribution in Fundamental Chemistry. I refer to the calculation of the number of isomers which a molecule might have given its constituent atoms. I have been able to obtain some information about the results of this work, including the article "The Scope of Structural Isomerism," by Dennis H. Smith, and two articles by L. M. Masinter* from a series on artificial intelligence.

I have a desperate need to find information on the number of possible isomers for as many molecules as possible. This need exists in order to clarify an important feature for the periodic system which we believe can be constructed for molecules with any given number of atoms. The periodic system for diatomic molecules has been extensively demonstrated to agree with the data and to allow the prediction of new data. Work on the periodic system for triatomic molecules has begun. As the number of atoms in the molecule increase, the problem of isomers grows more and more important - hence this letter.

The articles by Masinter* give the number of isomers for a few molecules. The article by Smith gives the number of isomers for a large number of molecules, but

*et al.

unfortunately the numbers of hydrogen atoms are not listed (the degree of unsaturation is listed). I know that, in principle, it should be possible to reconstruct the molecular formula including the number of hydrogens with the tabulated information. However, our time and expertise are limited and so I am writing to see whether you might have the original print-outs in which the number of isomers are listed for the specific molecules including the exact number of hydrogen atoms. I venture to take your time for this question because of the invitation at the end of Dr. Smith's article.

Your attention to this question and any assistance you might be able to render (such as to whom to write for the necessary information) are deeply appreciated.

Respectfully yours,

Ray Hefferlin

Ray Hefferlin,
Professor of Physics

RH'lh

such plots display a slightly decreasing slope with increasing carbon number for each given unsaturation.

As unsaturation increases, the slopes of the curves increase sharply (see Figure 1 as U increases from zero to four). However, curves representing higher unsaturation counts frequently begin, and remain for a short time below curves representing lower unsaturation counts, for example, the curves for U = 3 and U = 4, Figure 1. It seems clear that the higher the degree of unsaturation the greater the possible number and combinations of multiple bonds and ring systems. This factor is responsible for increasing slope with increasing U observed in Figure 1. But, carrying this argument to its extreme, why should not the number of isomers for a given combination of atoms always be greater for higher unsaturation counts, as long as the total valence of the composition list can support the specified unsaturation? The next section discusses this point in more detail.

The Number of Structural Isomers vs. Unsaturation Count. In Figure 2 some selected data from Table I are plotted including various combinations of six atoms. This figure illustrates the typical rise and fall of the number of isomers with increasing unsaturation. The value of unsaturation at which the number of isomers reaches a maximum is a complex function of the number and identity of atoms. The value shifts slowly to higher values of unsaturation as the number of atoms increases. Most composition lists of four atoms show maxima at U = 1 or 2. Several composition lists of six atoms have maxima at U = 3. However, most still display maxima at U = 1 or 2.

It is interesting to note that there is a smooth variation of number of isomers with unsaturation. No example presented in Table I shows a number of isomers which declines, then rises again with increasing unsaturation.

The curves presented in Figure 2 serve to introduce another set of questions and observations. These are concerned with the effects on the number of isomers caused by replacement of one type of atom with another type.

The Number of Structural Isomers vs. Atom Type. Most chemists intuitively expect that a given number of atoms and unsaturation will yield a larger number of isomers with a more diverse collection of atoms. As a first approximation this is true, as indicated in Figure 2. The curve of C_5N_1 lies above the curve of C_6 for all values of U; the curve of C_5O_1 lies above that of C_6 for $0 \le U < 4$. It is reasonable to expect that in a comparison of the results of exchange of atom type (e.g., either one N or one O for one C), the exchange yielding a higher total valence results in a group of atoms capable of yielding a greater diversity of structures. To a first approximation this is also true as the curve of C_5N_1 lies above that of C_5O_1 for all values of U (Figure 2).

Results presented previously, however, indicate the subtle interplay of atom type and valence. Intuition often fails to predict the relative numbers of isomers resulting from replacing an arbitrary number of atoms by the same number of atoms which differ in name and valence. Some representative data selected from Table I are presented in Table II to illustrate the influence of variation of atom type on the number of isomers.

For a given unsaturation count, the ratio between numbers of isomers resulting from N-substitution vs. O-substitution increases with increasing substitution, e.g., N_1/O_1, N_2/O_2, Table II.

DISCUSSION

The variations of numbers of isomers with increasing numbers of atoms, with unsaturation count or with substitution of one atom type with another can be rationalized on the basis of knowledge of an algorithm for constructing structural isomers such as that used by the structure gen-

Table I. Structural Isomers of $C_xN_yO_z$, $x + y + z \le 6$, for Allowed Values of the Unsaturation, U

No. of atoms	Unsaturation 0	1	2	3	4	5	6	7
C_1	1							
C_2	1	1	1	1				
C_3	1	2	3	2	1			
C_4	2	5	9	11	7	3		
C_5	3	10	26	40	40	21	6	
C_6	5	25	77	159	217	185	85	19
O_1	1							
C_1O_1	1	1						
C_2O_1	2	3	3	1				
C_3O_1	3	9	13	9	2			
C_4O_1	7	26	55	62	36	7		
C_5O_1	14	74	205	337	318	151	21	
O_2	1	1						
C_1O_2	2	2	1					
C_2O_2	5	10	9	3				
C_3O_2	11	34	52	34	7			
C_4O_2	28	122	263	301	163	25		
O_3	1	1						
C_1O_3	3	4	1					
C_2O_3	10	22	20	5				
C_3O_3	28	102	152	98	16			
O_4	1	1						
C_1O_4	5	6	2					
C_2O_4	20	48	41	10				
O_5	1	1						
C_1O_5	6	9	2					
O_6	1	1						
N_1	1							
C_1N_1	1	1	1					
C_2N_1	2	4	5	2				
C_3N_1	4	12	21	19	7			
C_4N_1	8	35	85	116	87	27		
C_5N_1	17	100	313	[illegible]	[illegible]	[illegible]	[illegible]	
N_2	1	1	1					
C_1N_2	2	4	4	1				
C_2N_2	6	18	27	19	5			
C_3N_2	14	62	136	155	86	11		
C_4N_2	38	218	[illegible]	[illegible]	1005	[illegible]	64	
N_3	1	2	1					
C_1N_3	4	17	13	6				
C_2N_3	14	58	110	99	34			
C_3N_3	45	[illegible]	651	969	706	134		
N_4	2	4	4	2				
C_1N_4	8	29	47	37	6			
C_2N_4	37	189	[illegible]	512	272	42		
N_5	2	7	9	4				
C_1N_5	15	73	145	151	42			
N_6	4	15	25	15	6			
N_1O_1	1	1						
$C_1N_1O_1$	3	5	3					
$C_2N_1O_1$	8	22	26	11				
$C_3N_1O_1$	21	84	154	136	46			
$C_4N_1O_1$	56	299	[illegible]	[illegible]	[illegible]	[illegible]		
N_2O_1	2	3	1					
$C_1N_2O_1$	8	21	18	4				
$C_2N_2O_1$	31	115	177	114	29			
$C_3N_2O_1$	102	527	1194	[illegible]	704	[illegible]		
N_3O_1	3	7	4					
$C_1N_3O_1$	21	71	86	24				
$C_2N_3O_1$	101	481	935	576	256			
N_4O_1	6	18	17	4				
$C_1N_4O_1$	52	225	361	207	32			
N_5O_1	11	43	55	22				
N_1O_2	2	2						
$C_1N_1O_2$	8	15	8					
$C_2N_1O_2$	28	84	95	40				
$C_3N_1O_2$	90	391	732	641	202			
N_2O_2	5	8	3					
$C_1N_2O_2$	28	75	65	12				
$C_2N_2O_2$	132	521	807	506	76			
N_3O_2	10	23	13					
$C_1N_3O_2$	56	306	369	137				
N_4O_2	24	72	66	14				
N_1O_3	3	3						
$C_1N_1O_3$	17	34	18					
$C_2N_1O_3$	76	246	288	110				
N_2O_3	9	15	5					
$C_1N_2O_3$	73	207	175	28				
N_3O_3	24	58	30					
N_1O_4	4	4						
$C_1N_1O_4$	33	68	34					
N_2O_4	17	27	8					
N_1O_5	5	5						