CALCULATION OF MOLECULAR FORMULAS IN HIGH-RESOLUTION MASS SPECTROMETRY

By: Joshua Lederberg
Professor of Genetics
Stanford University
School of Medicine

An appendix to

STRUCTURE ELUCIDATION OF NATURAL PRODUCTS BY MASS SPECTROMETRY

Volume II: Steroids, terpenoids, sugars, and miscellaneous classes

Herbert Budzikiewicz
Research Associate
Stanford University

Carl Djerassi
Professor of Chemistry
Stanford University

Dudley H. Williams
Research Associate
Stanford University

HOLDEN-DAY, INC.
San Francisco, London, Amsterdam
1964

Reprinted with the permission of Holden-Day, Inc. for use as an interim report to the National Aeronautics and Space Administration under grant No. NsG 81-60.

The clipped pages are obsolete.

Appendix

Calculation of Molecular Formulas in High-resolution Mass Spectrometry

by J. Lederberg

The advent of commercial instruments for high resolution work introduces a new dimension into mass spectrometry, as pioneered by Beynon.[1] The high resolution measurement of a molecular mass must then be interpreted as a consistent formula. For example, given a reading of 718.3743 ± .0060, what formulas should be considered to account for it? Few chemists will have the patience and calculating skill to wish to do this by arithmetic trial and error; accordingly, extensive tables[2] and other algorithms[3,4,5] have been presented to aid in this task. At one extreme, one might imagine a complete directory in which every formula is listed — and the chemist need only look up his answer — but he might also need a large library in which to store the necessary number of volumes. To greater advantage, the calculations can be programmed on a computer[3] and this may be the most constructive direction of future work. Alternatively, procedures involving a modest amount of arithmetic can be applied with the help of an abbreviated table. A related procedure and more extensive tables have been presented more fully elsewhere.[3] The present account is confined to compounds containing C, H, O and N, and may be used independently, although reference 3 might be consulted for further clarification if necessary.

Instead of listing all formulas one by one, we note that ^{12}C by definition has a mass of precisely 12. Hence, the fractional part of the mass number is not attributable to C, but only to H, O, and N. A comprehensive table (see for example, ref. 3, Table 2) can be seen to go through a repeating cycle every 12H atoms. Table 3 in this appendix is, in fact, just the basic block also covering the ranges of oxygen up to 11, nitrogen up to 7, and any value of carbon.

To encompass larger values of H in a compressed version of the tables for the purposes of this appendix, the calculation includes a step (Table 2) of figuratively extracting some multiple of 12H from the molecule, as necessary.

For example, to analyze an intact molecule whose mass is determined as 718.3743 ± .0060, we follow these steps:

1. According to the fifth entry of Table 2, a decimal of .30000 - .40000 calls for the subtraction of 36H (=36.28170). We therefore calculate the molecule as (718.37430 - 36.28170) =682.09260.

2. Divide 682 by 12 through the use of Table 1.

   Quotient = 56

   Integer residue = 10

   Decimal = .09260 ± .00600 (expressed as 5 decimal places)

3. Look up in Table 3 integer residue class 10 for 09260 ± 00600, i.e., values in the range 08660 to 09860. Since the molecule is stated to be intact, ignore lines marked with an asterisk (*), which refer to free radicals or protonated species.

4. The following values will be noted as candidates:

DECML	H	N	O	=C	WMIN
08731	14	6	8	18	262
09000	18	0	10	14	274
09134	14	4	6	13	214
09536	14	2	4	8	166
09669	10	6	0	7	106
09721	18	4	11	20	322

Other criteria will have to be used to choose among them, as illustrated in steps 5 and 6.

5. Check that the formula weight, i.e., 682, is not less than the WMIN given in each case.

6. To illustrate further interpretation, suppose that analytical data call for 4 to 5 nitrogens and 6 to 8 oxygens. Then the only solution is listed as

DECML	H	N	O	=C
09134	14	4	6	13

=C stands for the mass of the H+N+O expressed in units of carbon mass, i.e., $H_{14}N_4O_6 \equiv C_{13}$. Thus, 13 is subtracted from the quotient of step 2, (56-13=43) to give the value of the answer: $C_{43}H_{14}N_4O_6$ = 682.09134.

7. Restore 36H (=36.28170 — see Table 2) subtracted in step 1 to give the final answer:

$C_{43}H_{14}N_4O_6$	682.09134
H_{36}	36.28170
$C_{43}H_{50}N_4O_6$	718.37304

REFERENCES

1. J. H. Beynon, Mass Spectrometry and Its Applications to Organic Chemistry, Elsevier, Amsterdam, 1960.

2. J. H. Beynon and A. E. Williams, Mass and Abundance Tables for Use in Mass Spectrometry, Elsevier, Amsterdam, 1963.

3. J. Lederberg, Computation of Molecular Formulas for Mass Spectrometry, Holden-Day, San Francisco, 1964.

4. J. Lederberg, Tables and An Algorithm for Calculating Functional Groups of Organic Molecules in High Resolution Mass Spectrometry, NASA Sci. Techn. Aerosp. Rep., N64-21426 (1964).

5. E. Kendrick, A Mass Scale Based on CH_2 = 14.0000 for High-Resolution Mass Spectrometry of Organic Compounds, Analytical Chemistry, 35, 2146 (1963).