Quantitation of the common components of deoxyribonucleic acids by mass spectrometry:
application to the analysis of DNAs of unusual composition

D.M.Hawley and J.L.Wiebers

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

Received 17 July 1978

ABSTRACT

A mass spectral method for the quantitation of the percentages of deoxy-adenosine, deoxyguanosine, deoxycytidine, and thymidine in intact DNAs has been devised. Standard curves for each nucleoside have been constructed which are based upon the observation that a direct correlation exists between the heights (% deflection) of diagnostic peaks from these nucleosides in a mass spectrum and the published percent composition of specific DNAs. Analyses of DNA from Clostridium perfringens, Micrococcus luteus, Escherichia coli, Bacillus subtilis, Pseudomonas fluorescens, Drosophila melanogaster, salmon sperm, and bacteriophage lambda were used to determine standard curves. The validity of the method was demonstrated by comparison of the results from the mass spectral procedure with results from the chemical analyses of the DNAs from calf thymus and wheat germ. Analysis of ϕ X-174 DNA yielded values consistent with the published values obtained via sequence analysis and indicated that the method is applicable to both single and double-stranded DNAs. Results from T2 DNA, which contains no cytidine, exhibited artificially high values for adenosine, guanosine and thymidine with concomitant alteration in the A/T and G/C molar ratios. Such skewed results are useful in predicting the presence of modified nucleosides. The extreme sensitivity of the method has been exploited in the analysis of subnanogram quantities of restriction endonuclease fragments from DNA.

INTRODUCTION

Previous studies ¹⁻⁷ have demonstrated that the technique of pyrolysis-electron impact-mass spectrometry (PyEIMS) can be applied to intact DNA yielding information concerning (a) the qualitative composition of DNA, (b) the base sequence of DNA fragments, and (c) the detection and identification of modified nucleosides in the whole DNA molecule. This communication shows how the same mass spectral method used in the previous studies for qualitative results can be extended to the quantitative compositional analysis of DNA.

The need for a precise and highly sensitive quantitative method becomes apparent in such investigations as the following wherein (a) the nucleoside

composition of a series of restriction endonuclease fragments from a DNA molecule are being compared, (b) the dissimilarities in composition of two DNAs are being examined -- e.g., main band DNA versus satellite DNA or, (c) the presence and the extent of modified nucleosides in two DNAs is being contrasted -- e.g., host DNA versus viral DNA. In all of these instances the available amounts of DNA are usually very small and, thus, are not always amenable to analysis by conventional methods. The advantages of the mass spectral procedure are that the DNA is directly fragmented in the spectrometer without prior chemical or enzymatic treatment and that the analysis can be carried out on subnanogram quantities of DNA.

The rationale for the development of a quantitative mass spectral method has been the following; mass spectra of DNA exhibit major peaks at the mass values for the purine and pyrimidine bases -- i.e., at m/e 111 for cytosine, m/e 126 for thymine, m/e 135 for adenine and m/e 151 for guanine. These diagnostic ion products derive from the electron impact cleavage of the nucleosidic bond that occurs subsequent to the initial pyrolytic cleavage of the DNA at the phosphodiester linkages². Therefore, measurement of the base peaks reflects the amount of nucleoside present. In this case, the principal factors which affect the height of a specific base peak are (a) the amount of nucleoside present, and (b) the volatility of that particular nucleoside. In other words, the mass spectrum of a sample that contains equimolar amounts of the four common nucleosides does not show diagnostic peaks of equal intensity for all four components, but rather reflects the volatility of the base peaks which is in the order of $Cty > Thy > Ade > Gua$. The relationship between the amount of a given component and the percent deflection of its diagnostic peak in the mass spectrum was established by analyzing multiple samples of specific DNAs whose compositions had been determined by other methods. The DNAs chosen differed over a wide range in their base composition. For example, Cl. perfringens DNA is known to have a low guanosine and cytidine content and a high adenosine and thymidine value, while the reverse situation occurs in M. luteus DNA. By selecting appropriate DNAs in this respect for the study, and by analyzing multiple samples of each DNA, it was possible to demonstrate a linear relationship between the percent deflection of the diagnostic peaks for each component and the amount of that component as previously calculated and published⁸. Consequently, standard curves which were reproducible and statistically valid could be constructed for each of the common nucleosides in DNA. This communication describes how these standard curves can be used in the quantitative analysis of the composition of any

DNA or DNA fragment, and how the presence of modified residues affects the character of the results.

MATERIALS AND METHODS

The DNAs used were obtained from the following sources: M. luteus, bacteriophages ØX-174, T2, and lambda, Miles Laboratories, Elkhart, IN; Cl. perfringens, B. subtilis, E. coli K-12, Ps. fluorescens, and salmon sperm, Calbiochem, San Diego, CA; calf thymus and wheat germ, Dr. J. H. Spencer, McGill University; D. melanogaster, Dr. E. H. Cohen, Princeton University.

Mass Spectral Analysis (PyEIMS).

Solutions containing 0.001 to 1.0 $A_{260\text{nm}}$ unit of DNA were dried in capillary sample tubes. The samples were introduced into the spectrometer (DuPont 21-490B) by direct probe, heated to about 250° (exact temperature not critical), and analyzed at 70 eV, 3×10^{-6} Torr, and at a source temperature of 200°C . The spectra were recorded in two seconds (UV Oscillograph) when the ion monitor indicated that the maximum number of ions had been generated. All ions were measured from a single channel.

Construction of Standard Curves for Deoxyadenosine, Deoxyguanosine, Deoxycytidine, and Thymidine Content in DNA. Mass spectral analysis as described above was carried out on ten individual samples of 1.0 $A_{260\text{nm}}$ unit each of DNA from the following sources: M. luteus, Cl. perfringens, salmon sperm, B. subtilis, E. coli K-12, Ps. fluorescens, D. melanogaster, and bacteriophage lambda. From the mass spectrum, the peak heights of the appropriate diagnostic peaks (Cyt, m/e 111; Thy, m/e 126; Ade, m/e 135; Gua, m/e 151) were measured in millimeters. These values were then summed and each peak height was divided by this total to calculate the percent deflection. An example of a typical calculation is as follows:

Salmon sperm DNA

<u>Nucleoside</u>	<u>m/e</u>	<u>Peak Height (cm)</u>	<u>% Deflection</u>	<u>% Composition</u>
Cytosine	111	380	20.4	21.6
Thymine	126	620	33.3	27.1
Adenine	135	730	39.2	28.9
Guanine	151	130	7.6	22.4
		Total	1860	

The relationship between the percent deflection and the percent composition from the literature was then explored in the following way: Values of the percent deflection from ten individual analyses for each nucleoside were taken and compared with the known composition figure for that nucleoside (Table I) for each of the eight different DNAs analyzed. This procedure established the reproducibility of the relationship between percent deflection and percent composition. Subsequently, the ten values for percent deflection for each nucleoside in each DNA were averaged and the relationship between the percent deflection and the percent composition was shown by regression analysis to be linear.

Table I. Values* Used for Construction of Standard Curves for % Composition of Nucleosides in DNA

DNA	CYD	THD	ADO	GUO
<u>Cl. perfringens</u>	12.8	36.3	36.9	14.0
<u>D. melanogaster</u>	20.2	29.4	30.7	19.6
<u>B. subtilis</u>	21.4	28.7	28.9	21.0
Salmon sperm	21.6	27.1	28.9	22.4
Bacteriophage λ	24.3	25.8	26.0	23.8
<u>E. coli</u> K-12	25.6	24.3	24.6	25.6
<u>Ps. fluorescens</u>	30.0	18.8	18.2	33.0
<u>M. luteus</u>	37.1	13.9	13.6	35.4

* Values listed as moles of nitrogenous constituents per 100 moles of total constituents⁸.

RESULTS AND DISCUSSION

Standard Curves for Quantitation of Nucleosides in DNAs. The results of the relationship between the percent deflection and the percent composition in

each case for deoxyadenosine, deoxyguanosine, deoxycytidine and thymidine from eight different DNAs are shown in Figure 1. It may be observed that in all four cases there is a linear relationship between the percent deflection and the percent composition. The slope of the curve is different for each nucleoside and this result reflects the volatility differences of the four nucleosides in the mass spectrometer.

Compositional Analysis of Other DNAs Using the Standard Curves. Other DNAs were analyzed by the mass spectral method and the percent composition of the four nucleosides was calculated from the standard curves shown in Figure 1. Table II is a compilation of the percent composition values obtained from the mass spectral method as compared with values found in the literature⁸ for these DNAs. In all cases, the assessments from the mass spectral determination fall within the limits of the range of the published values determined by other methods. For example, the results from the analysis of bacteriophage ϕ X-174 DNA are of particular interest in that the percent composition estimations obtained by the mass spectral method are in closer agreement with the values that were found as a result of sequence studies⁹ than to the values previously given for this DNA⁸. This analysis also showed that the mass spectral procedure is applicable to single-stranded as well as double-stranded DNA.

The results (Table II) from the analysis of bacteriophage T2 DNA demonstrated another useful facet of the mass spectral method. This DNA is unusual in that all of the cytidine residues are hydroxymethylated; the mass spectrum reveals this difference by the presence of the base peak for hydroxymethylcytosine at m/e 141⁴. Since the diagnostic peak from cytidine at m/e 111 is only at background level in the mass spectrum, the percent deflection values for deoxyadenosine, deoxyguanosine, and thymidine appear to be exaggerated. Consequently, when these values are converted to percent composition values from the standard curve, the results are abnormally high for these nucleosides and do not represent their actual values in the DNA. However, this result is meaningful because it shows that, since reasonable A/T and G/C molar ratios cannot be obtained from these percent composition values, a substantial proportion of the deoxycytidine residues are modified in this DNA. We believe that this situation should allow the detection of base modification any time that a sufficient percentage of one type of base is modified such that agreement of the A/T or G/C ratios cannot be obtained. A reasonable limit on the sensitivity of this method is that, if five percent of a given base were modified, the base ratios would show significant deviation from normal.

Figure 1. Standard plot for quantitation of (a) deoxycytidine, (b) thymidine, (c) deoxyadenosine, and (d) deoxyguanosine in DNA. Each point represents the average percent deflection values from analyses of ten individual samples of DNA. Numbers adjacent to the points indicate the specific DNA: 1) *Cl. perfringens*, 2) *D. melanogaster*, 3) *B. subtilis*, 4) salmon sperm, 5) bacteriophage lambda, 6) *E. coli* K-12, 7) *Ps. fluorescens*, 8) *M. luteus*. Presumably, curves (b) and (c) do not begin at the origin because of minor contributions to the peak height from other fragment ions of the same mass.

Table II. Comparison of Mass Spectral Quantitation Method for Nucleoside Composition in DNAs with Published Values Obtained by Alternative Methods.

DNA	% COMPOSITION								
	MASS SPECTRAL METHOD				PUBLISHED VALUES				
	Cyd	Thd	Ado	Guo	Cyd	Thd	Ado	Guo	Ref
Calf thymus *	20	28	29	23	20.0- 21.1	26.6- 31.2	27.1- 30.3	19.5- 23.9	8
Wheat germ **	17	29	28	21	16.6- 19.6	27.0- 28.6	25.4- 29.0	20.2- 25.0	8
ØX-174	21	31	24	24	18.5 21.5	32.7 31.2	24.6 23.9	24.1 23.3	8 9
T2 ***	1	41	37	21	-	36.0	32.0	17.0	8

* DNA contains 1.3-1.9 moles of 5-methyldeoxycytidine per 100 grams of DNA phosphorous⁴.

** DNA contains 5.6 moles 5-methyldeoxycytidine per 100 grams of DNA phosphorous⁴.

*** DNA has 100% hydroxymethylation of deoxycytidine residues.

CONCLUSION

Although other methods exist that have been employed for the calculation of the percent composition of the common nucleosides in DNA, the mass spectral method is particularly applicable in cases where very limited amounts of DNA or DNA fragments must be analyzed. The forte of the mass spectral procedure lies in its extreme sensitivity. In our laboratory, we are currently exploiting this advantage in the analysis of restriction endonuclease fragments. Quantitative compositional analysis of 0.001 A_{260nm} units of fragments from 40 to 900 base pairs in length have been carried out. In the cases where particular fragments contain modified nucleosides, the percent composition values for the four nucleosides yield subsequent A/T or G/C ratios which are inconsistent, and thus reflect the presence of modification. The concomitant presence of diagnostic ion products in the mass spectrum of these fragments

for methylated deoxyadenosine, methylated deoxycytidine, or deoxyuridine residues explains the imbalance of the common residues^{10,11}. At the present time, the amount of modified nucleosides present cannot be calculated because of the lack of appropriate DNAs in which a sufficient range of percent modification is known and, therefore, standard curves cannot be constructed as yet. Theoretically, it would be possible to prepare DNA with varying percentages of methylated (or otherwise substituted) bases by enzymatic polymerization of modified triphosphates, and the appropriate standard curves for the estimation of modified nucleosides could then be determined.

In summary, the mass spectral quantitation method (a) permits the calculation of the percentages of the four common nucleosides in intact DNA and in DNA fragments, (b) indicates from such calculations the presence of unusually high amounts of purines or pyrimidines, and (c) discloses alterations in the A/T and G/C ratios which suggests the presence of modified nucleosides in the DNA.

ACKNOWLEDGEMENT

This investigation was supported by NSF Grant PCM 76-21554 and NIH Grant GM 19395.

REFERENCES

1. Wiebers, J. L. (1973) Anal. Biochem. **51**, 542-556.
2. Wiebers, J. L. and Shapiro, J. A. (1977) Biochemistry **16**, 1044-1050.
3. Burgard, D. R., Perone, S. P. and Wiebers, J. L. (1977) Biochemistry **16**, 1051-1057.
4. Wiebers, J. L. (1976) Nucleic Acids Res. **3**, 2959-2970.
5. Burgard, D. R., Perone, S. P. and Wiebers, J. L. (1977) Anal. Chem. **49**, 1444-1446.
6. Armbruster, M. A. and Wiebers, J. L. (1977) Anal. Biochem. **83**, 570-579.
7. Wiebers, J. L. (1978) in High Performance Mass Spectrometry: Chemical Applications, Chapter 13, 248. In Press.
8. Shapiro, H. S. (1976) in Handbook of Biochemistry and Molecular Biology, Nucleic Acids, 3rd Edition, Volume II, G. D. Fasman, Ed., CRC Press, Cleveland, Ohio, pp. 241-281.
9. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A. III, Slocumb, P. M. and Smith, M. (1977) Nature **265**, 687-695.
10. Wiebers, J. L. and Ho, N. W. Y. (1977) Federation Proc. **36**, 899.
11. Ho, N. W. Y. and Wiebers, J. L. (1978) Federation Proc. **37**, 1737.