Ultrafast sequencing of oligodeoxyribonucleotides by FAB-mass spectrometry

Lutz Grotjahn, Ronald Frank and Helmut Blöcker

GBF (Gesellschaft für Biotechnologische Forschung mbH), Mascheroder Weg 1, D-3300 Braunschweig, FRG

Received 14 June 1982; Accepted 6 July 1982

ABSTRACT

A fully instrumental method is described for the bidirectional sequencing of oligodeoxyribonucleotides. The method makes use of the negative ion fragmentation patterns of fast atom bombardment mass spectrometry. It is less time consuming than any other sequencing procedure known to date. Since one sequencing run takes as little as one hour, this new method is anticipated to cut down considerably the time required for the controlled synthesis of oligodeoxyribonucleotides of (currently) up to ten nucleotide units in length.

INTRODUCTION

Synthetic oligonucleotides have proved extremely useful in modern biosciences. In particular the chemical synthesis of pre-defined nucleotide sequences has supported and accelerated the recent development of gene technology. This has stimulated the development of new and quicker synthetic techniques such as better coupling procedures and highly selective and mild agents for deprotection, and the use of suitable solid supports. These syntheses can be performed in a matter of days and as a consequence the longer part of the work is now the careful analysis of the deprotected oligonucleotides. Sequencing is usually done by chemical degradation or 'wandering spot' procedures after radioactive labelling. Here we describe a completely instrumental method for a much faster (less than one hour) and unambiguous sequence analysis of oligodeoxyribonucleotides up to at least ten units in length using the negative ion fragmentation patterns of fast atom bombardment mass spectrometry (FAB-MS).

MATERIALS AND METHODS

Oligodeoxyribonucleotides were synthesized following, in gene-
ral, established phosphate triester procedures. Benzoyl, anisoyl and isobutyryl groups were used for the protection of the heterocyclic bases, 4-methoxy-trityl for the protection of the 5'-OH, and 2-chlorophenyl and 2,2,2-tribromoethyl for the protection of the phosphate functions. Condensation reactions were performed in solution with 2,4,6-tri-isopropylbenzene-sulfonyl-3-nitro-1,2,4-triazolide as condensing agent. For the final work-up the unprotected oligomers were formed by treatment with 10 % pyridine in concentrated ammonia followed by 80 % acetic acid. Homogeneity higher than 99 % was routinely achieved after ion exchange chromatography on Sephadex A-25 in the presence of 7 M urea and subsequent desalting. This was carefully checked by reverse phase HPLC, polyacrylamide-gel electrophoresis in the presence of chain length markers and two-dimensional finger printing.

The mass spectra were recorded on a Kratos MS 50 S with a high field magnet (mass range ca. 3000 at 8 kV) and a Kratos FAB source. The atom gun used xenon and produced a beam of neutral atoms at 8-9 keV. A solution of the triethylammonium salt of each oligodeoxyribonucleotide (4-5 μl, containing 1-1.5 OD_{260}^{260} ca. 10 nmol) was injected into the glycerol matrix (ca. 2 μl) present on the FAB copper probe tip. The water was removed in the direct insertion lock and the spectra were recorded at a magnet scan rate of 300 s/decade.

RESULTS AND DISCUSSION

The Figures 1a, 2a and 3a show the negative ion FAB mass spectra of the oligodeoxyribonucleotides d(A-C-T-C-G-A-T-G), d(G-C-G-A-T-C-G-C), and d(G-A-A-G-A-T-C-T-C-T). The highest mass is the (M-H)^- ion. The corresponding doubly charged ion is found at half mass. The mass difference between two markers is indicative of a particular nucleotide. Those above the spectra mark fragment ions with 5'-phosphate ends (5'-P sequence ions) and those below mark fragment ions with 3'-phosphate ends (3'-P sequence ions). This is also represented by the shorthand structures in the Figures 1b, 2b and 3b. The dotted lines show the main fragmentations with attached mass, those above corresponding to the 5'-P sequence ions and those below to the 3'-P sequence ions.

A similar but more complex fragmentation behaviour was recent-
Figure 1a: Relevant regions of the negative ion FAB mass spectrum of d(A-C-T-C-G-A-T-G) excluding the glycerol matrix.

Figure 1b: The corresponding shorthand structure with those bonds marked that on breakage give rise to the main fragment ions.
Figure 2a: Relevant regions of the negative ion FAB mass spectrum of d(G-C-G-A-T-C-G-C) excluding the glycerol matrix.

Figure 2b: The corresponding shorthand structure with those bonds marked that on breakage give rise to the main fragment ions.
Figure 3a: Relevant regions of the negative ion FAB mass spectrum of d(G-A-A-G-A-T-C-T-T-C) excluding the glycerol matrix.

Figure 3b: The corresponding shorthand structure with those bonds marked that on breakage give rise to the main fragment ions.
ly described for negative ion $^{252}$Cf-plasma desorption mass spectrometry of fully protected oligonucleotides$^{12}$. These authors also state, that oligonucleotides containing the naturally occurring phosphodiester internucleotidic links exhibit a still more complex fragmentation pathway, and that molecular ion yields are severely attenuated. On the contrary in our FAB mass spectra we only observe sequence ions accompanied by some minor -18 peaks. This exclusive formation of specific 5'-P and 3'-P sequence ions gives rapid sequence information simply by counting mass differences between consecutive sequence ions of the same type. The discrimination of these two types of sequence ions is possible, because those ions with the same number of nucleotide units show a consistent difference in peak intensities. Since the bond involving the 3'-O atom, connected to a secondary carbon of the sugar moiety, is more labile than the one involving the 5'-O atom, connected to the primary carbon of the sugar, the 5'-P sequence ions are without exception more intense than the corresponding 3'-P sequence ions.

In practice the determination of the nucleotide sequence proceeds as follows: Starting from the (M-H)$^-$ peak all sequence ions are classified as either 5'-P ions or 3'-P ions according to their intensities. Then the exact mass differences between neighbouring peaks of the same family are determined by counting to give their corresponding nucleotides. The complete base sequence of the oligonucleotide can then be read independently twice, starting either from the 5'- or from the 3'-end.

It is obvious, however, that all fragments with the same base composition have the same mass for both the 5'-P and the 3'-P sequence ions. Thus, the spectrum of d(G-C-G-A-T-C-G-C) shows only one peak for both terminal dinucleotide ions. However, despite any ambiguity in this mass area, the evaluation of the actual sequence does not create any problem. It is clear from the spectrum as a whole that only one peak shows up in a wide mass range. Masses and peak intensities of the higher oligonucleotide ions clearly confirm that, reading the sequence from the 5'-end, position 2 is C and position 7 is G and the single peak at mass 635 represents both dinucleotide ions.

The FAB-MS negative ion fragmentation pattern can well be com-
pared to the result of the simultaneous action of two exonucleases partially digesting oligonucleotides from either end. However, it is unlikely that there will ever be any enzymatic procedure of comparable speed and resolving power.

We have performed mass spectrometric sequence analysis of various oligonucleotides. The method is of general applicability and is now routinely used in our laboratory. Thus, negative ion FAB-MS is a quick and reliable method for the bidirectional sequence analysis of unprotected oligodeoxyribonucleotides up to at least ten base units in length.

Future instrumental technology should enable mass ranges up to and above 5000 to be exploited with the corresponding possibility of sequencing oligonucleotides of more than 16 base units. Current improvements to the mass spectrometer will allow us to sequence our longer oligonucleotides by this method. We anticipate that this method will help to speed up considerably the reliable synthesis of gene and other DNA fragments.

ACKNOWLEDGMENT

We thank A. Van Aerschot, M. Becker, W. Heikens, and I. Holllatz for their support in the synthetic work, B. Foitor and I. Schweer for artwork. Part of this work was supported by BMFT grant BCT 0339.

REFERENCES

10. Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N. Na-