A NEW METHOD FOR SEQUENCE DETERMINATION OF LARGE OLIGONUCLEOTIDES

Robert W. Holley, James T. Madison and Ada Zamir

U. S. Plant, Soil and Nutrition Laboratory, Soil and Water Conservation Research Division, Agricultural Research Service, U. S. Department of Agriculture, and Department of Biochemistry, Cornell University, Ithaca, N. Y.

Received August 27, 1964

The most promising approach to the determination of the nucleotide sequence of an RNA involves, as an initial step, the enzymatic cleavage of the RNA, and the determination of the structures of the various oligonucleotides that are formed. With present methods, the determination of the structures of the smaller oligonucleotides, particularly the di- and trinucleotides, is straightforward, but methods are inadequate for the determination of the structures of many larger oligonucleotides. The present communication describes a method which is applicable to these larger oligonucleotides. The method has been applied successfully to oligonucleotides containing from five to eight nucleotide residues.

The oligonucleotide is first treated with alkaline phosphatase to remove the 3'-phosphate, since this phosphate group inhibits the action of snake venom phosphodiesterase (Privat de Garilhe, et al., 1957). Then, partial degradation with snake venom phosphodiesterase, an exonuclease, degrades the oligonucleotide stepwise from the 3'-end (Razzell and Khorana, 1959) and gives a series of successively smaller degradation products. These products are separated by chromatography, and the 3'-end of each stepwise degradation product is determined by identifying the nucleoside formed by alkaline hydrolysis. This gives the identity of the nucleotide in successive positions down the chain and establishes the sequence of the oligonucleotide.

EXPERIMENTAL

Approximately 0.5 mg of oligonucleotide is dissolved in 0.10 ml of 0.1 M pH 8.3 Tris (chloride) buffer, 0.10 ml (0.4 mg per ml) of alkaline phosphatase (Worthington[1/], rechromatographed on DEAE-cellulose according to Garen and Levinthal, 1960) is added, and the solution is incubated 1 hr at 37°. The solution is then diluted with 0.6 ml of water and is extracted three times with 0.8 ml of phenol (previously saturated with water). The last traces of phenol are removed from the aqueous layer by at least five extractions with ether, and the ether remaining with the aqueous layer is evaporated. The final volume of the aqueous solution is approximately 0.5 ml. (The extractions with phenol and ether lower the pH to 6 to 7. The pH can be left low, or it can be raised by the addition of 0.1 ml of 0.1 M pH 7.5 Tris buffer, in which case less snake venom phosphodiesterase may be required.) To this solution of the dephosphorylated oligonucleotide is added 0.1 ml of 0.1 M magnesium chloride and approximately 0.02 ml (2 mg per ml) of snake venom phosphodiesterase (Worthington, purified according to Keller, 1964), and the solution is left at room temperature for 15 minutes. (To determine the exact amount of enzyme needed, a small trial incubation and column may be required, or, if there is sufficient oligonucleotide, the digestion can be carried to approximately 30 percent completion in an automatic titrator.) After the digestion, the solution is mixed with 0.4 g of urea and 3 ml of 7 M urea, and the solution is added immediately to a 0.35 X 30-cm column of DEAE-cellulose packed in 7 M urea (Tomlinson and Tener, 1962). The degradation products are eluted from the column with a linear gradient formed from 60 ml each of 7 M urea and 0.6 M pH 7.5 sodium acetate in 7 M urea, and

---

1/Trade names and company names are included for the benefit of the reader and do not infer any endorsement or preferential treatment of the product listed by the U. S. Department of Agriculture.

fractions of approximately 1.2 ml are collected. Appropriate fractions are combined, after the peaks are located by ultraviolet absorption, the solutions are "desalted" on small DEAE-cellulose columns (Rushizky and Sober, 1962), and the material is recovered by evaporation. The separated stepwise degradation products are hydrolyzed with 0.05 ml of 0.5 N potassium hydroxide at 37° for 18 hours, and the hydrolysates are applied directly to S & S 589 Green Ribbon paper and subjected to two-dimensional paper chromatography, using isopropyl alcohol-water-ammonia [isopropyl alcohol, 60; water, 40; ammonia in the vapor phase (cf. Hershey, Dixon and Chase, 1953)] in the first direction, and isopropyl alcohol-HCl [isopropyl alcohol, 170; concentrated hydrochloric acid, 41; water to 250 (Wyatt, 1941)] in the second direction. The nucleosides and nucleotides are identified from their positions on the chromatograms as well as from spectra after elution of the spots from the paper.

RESULTS

Application of the method to the determination of the sequences of a hexanucleotide and an octanucleotide, obtained from the alanine RNA (Apgar, Holley and Merrill, 1962), will be described.

The hexanucleotide contained one Ap, two Cp's, one Gp and two Up's, and, since it was isolated from a takadiastase T1 RNase digest, it was known that Gp was the 3'-terminal nucleotide. It was also known that complete degradation of the hexanucleotide with snake venom phosphodiesterase gave adenosine, establishing that Ap was the 5'-terminal nucleotide. Degradation with pancreatic RNase gave ApUp, indicating that the hexanucleotide could be represented as ApUp(Cp,Cp,Up)Gp. The present method was used to establish the sequence of the remaining three nucleotides. Partial degradation of the hexanucleotide with snake venom phosphodiesterase, followed by chromatography on DEAE-cellulose, according to the above procedure, gave the pattern shown in

Fig. 1. The peaks numbered 5, 4, 3 and 2 contained material with five, four, three and two negative charges, respectively, and represented the

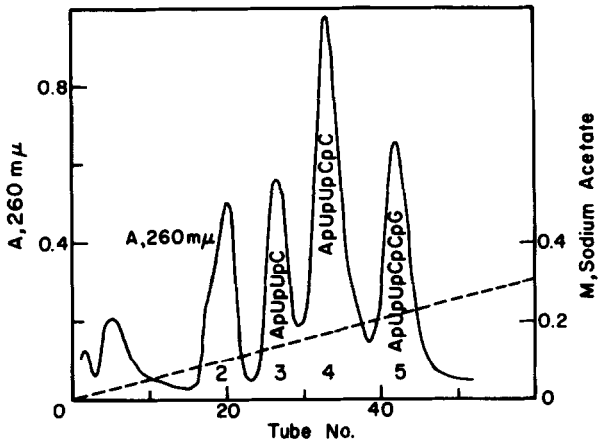

Figure 1. Chromatography of products of partial snake venom phosphodiesterase digestion of ApUpUpCpCpG.

dephosphorylated hexanucleotide, carrying five negative charges, and products of the degradation, carrying four, three and two negative charges, respectively. Peak 2 contained primarily mononucleotides formed by the snake venom phosphodiesterase. Peaks 3 and 4 each gave, after alkaline hydrolysis, cytidine as the only nucleoside, indicating that both of these stepwise degradation products had 3'-terminal cytidine. The only structures consistent with all the data are ApUpUpC for peak 3 and ApUpUpCpC for peak 4, and complete nucleotide analyses of the peaks were consistent with these structures. The sequence of the hexanucleotide was therefore ApUpUpCpCpGp. (The smaller stepwise degradation products ApUpU and ApU were lost in peak 2 under the chromatographic conditions used. If the identifications of these smaller products had been required to establish the sequence of the oligonucleotide, separation of these products would have required further fractionation of peak 2.)

The octanucleotide contained two Ap's, five Gp's and one Up, and it was known that one Gp was at the 5'-end and the Up was at the 3'-end.

The complete sequence was established using the present method. Partial degradation with snake venom phosphodiesterase followed by chromatography of the products gave the pattern shown in Fig. 2. Peak 9, with nine negative charges, was a small amount of the original octanucleotide. Peak 7, with seven negative charges, was the dephosphorylated

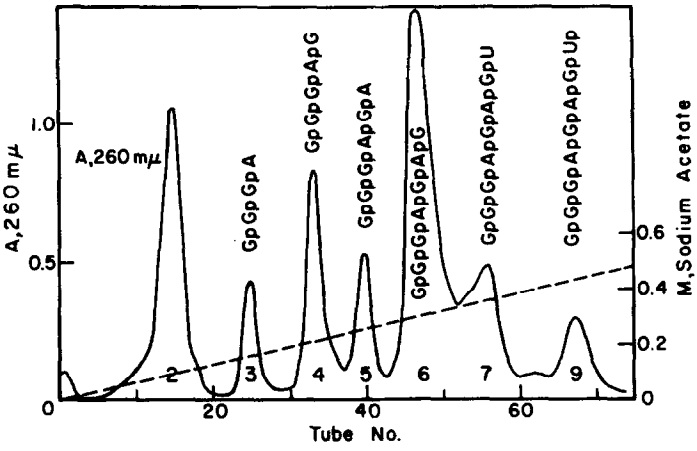

Figure 2. Chromatography of products of partial snake venom phosphodiesterase digestion of GpGpGpApGpApGpU.

---

octanucleotide, and peaks 6, 5, 4 and 3 were successive stepwise degradation products carrying six, five, four and three negative charges, respectively. Only one nucleoside was obtained on alkaline hydrolysis of each peak, and these were, for peaks 6, 5, 4 and 3, respectively, guanosine, adenosine, guanosine and adenosine. The nucleosides represent the 3'-end of each stepwise degradation product, and establish that the sequence of the octanucleotide was GpGpGpApGpApGpUp.

SUMMARY

Partial degradation of large oligonucleotides with snake venom phosphodiesterase gives a series of stepwise degradation products, and identification of the terminal nucleosides obtained from these products may be used to establish the sequence of the oligonucleotide.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation and by the National Institutes of Health.

REFERENCES

Apgar, J., Holley, R. W., and Merrill, S. H., J. Biol. Chem. 237, 796 (1962).
Garen, A., and Levinthal, C., Biochim. Biophys. Acta, 38, 470 (1960).
Hershey, A. D., Dixon, J., and Chase, M., J. Gen. Physiol. 36, 777 (1953).
Keller, E. B., Biochem. Biophys. Research Commun. 17, 412 (1964)(This issue
Privat de Garilhe, M., Cunningham, L., Laurila, U.-R., and Laskowski, M., J. Biol. Chem. 224, 751 (1957).
Razzell, W. E., and Khorana, H. G., J. Biol. Chem. 234, 2114 (1959).
Rushizky, G. W., and Sober, H. A., Biochim. Biophys. Acta, 55, 217 (1962).
Tomlinson, R. V., and Tener, G. M., J. Am. Chem. Soc. 84, 2644 (1962).
Wyatt, G. R., Biochem. J., 48, 584 (1951).