Journal of Chromatography, 177 (1979) 380–384

CHROM. 12,002

Note

Automated sequence analysis of a synthetic DNA fragment by anion-exchange high-performance liquid chromatography

J. F. M. DE ROOIJ, W. BLOEMHOFF and J. H. VAN BOOM*

Department of Organic Chemistry, Gorlaeus Laboratories, State University Leiden, P.O. Box 9502, 2300 RA Leiden (The Netherlands)

(First received March 28th, 1979; revised manuscript received May 4th, 1979)

Previously[1] we developed a practical and simple method to establish directly the sequence of synthetically prepared RNA and DNA fragments. The method was based on the identification and quantitation by anion-exchange high-performance liquid chromatography (HPLC) of the products obtained during the stepwise degradation of nucleic acids with the enzyme venom phosphodiesterase (PDE). This enzyme starts at the 3′-end of a nucleic acid fragment ($^{5'}$NpN----pNp$N^{3'}$)** and proceeds towards the 5′-end producing nucleoside 5′-phosphates and the 5′-terminal nucleoside.

In the previous procedure, samples were withdrawn from the enzymatic digest at appropriate intervals of time. The enzyme was deactivated by the addition of 0.2 *M* KH_2PO_4 (pH 3.5) to the samples, which were stored at −20° until they were analyzed by anion-exchange HPLC on Permaphase AAX[2]. The enzymatic products were identified from their retention times, and quantified with respect to an internal standard (Gp). This method enabled us to establish unambiguously, without recourse to prior ^{32}P-labelling of the substrate[3], the sequence of synthetic oligonucleotides up to a length of six units.

Nonetheless we felt the need to try to make the procedure less laborious and time consuming by automating the sequence analysis process.

This paper reports the rapid automated sequence analysis of a synthetic deoxydecanucleotide, d-TpApTpCpApApGpTpTpG.

MATERIALS AND METHODS

The decamer d-TpApTpCpApApGpTpTpG was prepared according to a phosphotriester approach[4]. Enzyme and chemicals used in this study were purified as described earlier[1].

* To whom correspondence should be addressed.

** Abbreviations used in this paper: N = nucleoside, pN = nucleoside 5′-phosphate, Np = nucleoside 3′-phosphate, dT = deoxythymidine, dC = deoxycytidine, dA = deoxyadenosine, dG = deoxyguanosine, G = guanosine. The letter "p" between two nucleoside initials stands for a 3′–5′-linked phosphodiester bond.

Apparatus

The sequence analysis was performed with a system consisting of a Spectra-Physics 740B pump, a Spectra-Physics SP8200 UV detector operating at 254 nm and a Linear 112 flat-back recorder. The enzymatic digestion was carried out in a Glenco 19925-1 syringe, inserted into a Perfusor infusion syringe pump. An air-actuated Valco injection valve equipped with a 10-μl sample loop served as the injection system. A small proportion (12 μl) of the reaction mixture was, at intervals of 12.5 min, pumped into the injection valve and from this amount 10 μl was immediately applied to the column. Anion-exchange HPLC was performed with the strong anion-exchange resin Permaphase AAX[2] (DuPont, Wilmington, Del., U.S.A.) dry-packed into a stainless-steel column (50 cm $\times$ 2.1 mm I.D.). Isocratic elution of the nucleoside and nucleotides was effected by a buffer containing 0.0025 *M* KH_2PO_4 (pH 4.5). A flow-rate of 1 ml/min at 20° was standard at a pressure of 500 p.s.i. The peak areas were recorded by a LDC computing integrator model 308.

RESULTS AND DISCUSSION

The automation of the sequence analysis had additional advantages over the previously described method[1]. For instance, withdrawal of samples at appropriate intervals of time and subsequent inactivation of the enzyme was unnecessary. No accumulation of the enzyme on the stationary phase was observed. The latter was corroborated as follows: a small portion (10 μl) of the buffer solution (see *Sequence analysis*) containing the enzyme was applied to the column and eluted under isocratic conditions. Analysis of the eluate indicated that the enzyme was not retarded on the column: it was eluted from the column within 1 min (*e.g.*, the peak at $t = 1$ min in Fig. 3a). Further, owing to the very good reproducibility of the injections and the accuracy of product quantitation no internal standard was necessary. The reproducibility and accuracy of the method were substantiated by the low value for the average standard deviation ($\sigma = 0.9\%$) when measuring the peaks of a mixture containing the four common d-nucleoside 5'-phosphates over ten injections.

The products obtained after partial digestion with PDE were identified from their retention times[1]. The retention times of the four common d-nucleoside 5'-phosphates are recorded in Table I.

TABLE I

BASE COMPOSITION OF THE DECANUCLEOTIDE AND RETENTION TIMES OF THE PDE DIGESTION PRODUCTS

Expected composition of the decanucleotide	*Products after PDE digestion*		*Calculated composition of the decanucleotide*
	Molar fractions found	*Retention time (min)* *	
d-($CG_2A_3T_4$)	dT 1.0	0.0	d-T-($CG_2A_3T_3$)
	pdC 0.9	1.3	
	pdT 3.3	3.7	
	pdA 2.9	5.4	
	pdG 1.9	8.8	

* Relative to the injection peak (*i.e.*, the peak at $t = 1$ min in Fig. 3) which contains the unretarded components of the reaction mixture such as the enzyme and the deoxynucleoside dT.

Sequence analysis

The determination of the base composition, which comprises the first step in the sequence analysis, could easily be corroborated by HPLC analysis of the nucleoside and nucleoside 5'-phosphates liberated by a complete digestion of the decanucleotide by PDE[1]. The base composition reveals not only the molar fractions (see Table I) of the nucleoside 5'-phosphates but also, because of the mode of action of the enzyme, the identity of the nucleoside at the 5'-end of the decamer.

The sequence analysis was performed as follows: 5 μl of buffer (0.025 *M* Tris·HCl, 0.005 *M* $MgCl_2$, pH 8.5) containing 0.1 μg of PDE was added to a solution of one absorbance unit of the decanucleotide (*ca.* 10 nmole) in 0.25 ml of the same Tris·HCl buffer. The reaction mixture was taken up into the syringe, which was then inserted into the infusion syringe pump. The first injection, which was performed immediately, was necessary to fill the capillary between syringe pump and injection valve. The other injections were performed automatically at intervals of 12.5 min.

During the analysis process trinucleotides and longer oligomers are not eluted from the column under the conditions. Oligonucleotide material, which accumulated on the column during the total sequence analysis process, was removed afterwards by elution with a buffer containing 0.05 *M* KH_2PO_4 and 0.5 *M* KCl (pH 4.5).

After 20 injections (total time 4 h), the peak areas recorded by the computing integrator were divided by the extinction coefficients[1] of the corresponding nucleoside 5'-phosphates, and the values thus obtained were plotted against time (see Fig. 1).

Plotting the sum of the liberated nucleoside 5'-phosphates against time (Fig. 2) gave, up to the time of 135 min, a straight line: at this point the remaining oligo-

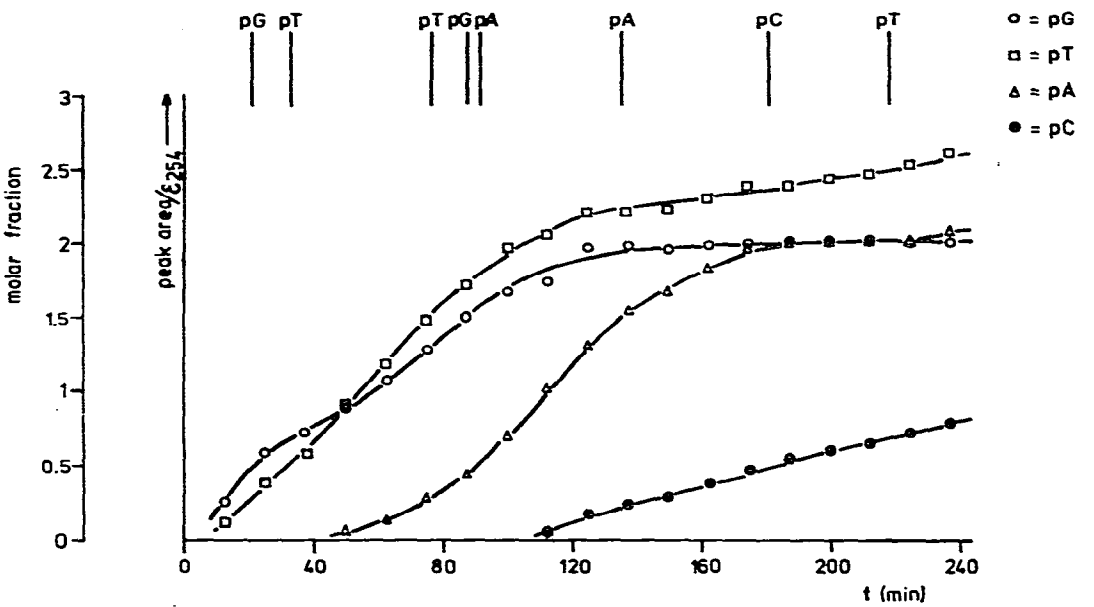

Fig. 1. Curves of the time of appearance of the nucleoside 5'-phosphates (pdG, pdT, pdA and pdC) from PDE action on the synthetic decamer d-TpApTpCpApApGpTpTpG. The vertical bars headed by the symbols for the nucleoside 5'-phosphate represent the half-times of the corresponding nucleoside 5'-phosphates.

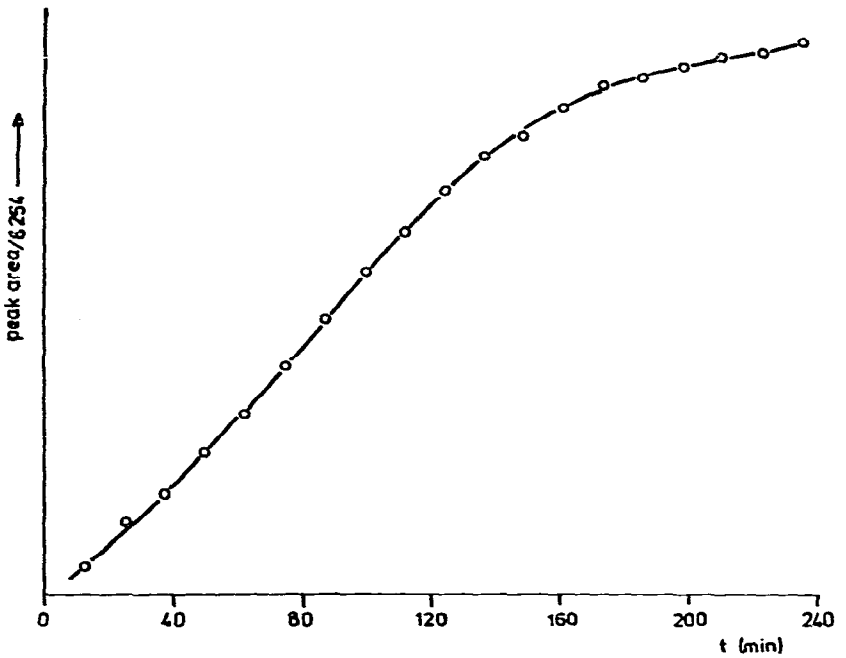

Fig. 2. Relationship between time and the total amount of nucleoside 5′-phosphates released (*i.e.*, the sum of the peak areas divided by the corresponding extinction coefficients at 254 nm (ε_{254}) of the nucleoside 5′-phosphates).

nucleotide reached an average length of four units. The reason for this levelling off is that the reaction decelerates at the stage of a tetramer and especially so when the enzymatic digestion reaches the dimer level.

The very low reaction rate at the dimer level can be used as an indication that the enzymatic reaction is near completion after 240 min (see Fig. 3). Furthermore, the total amount of deoxyguanosine 5′-phosphate (pdG), which was released during

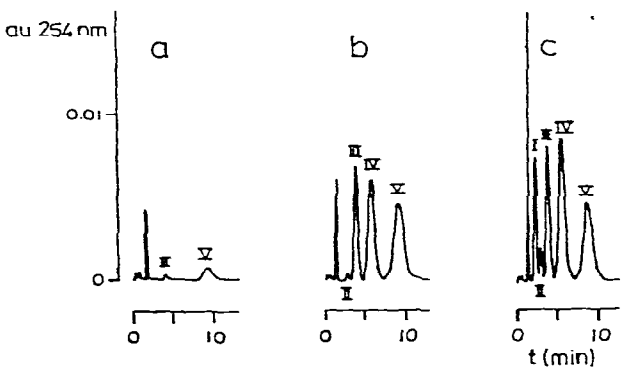

Fig. 3. Separation of the PDE digestion products on Permaphase AAX at several stages of the sequence analysis. I = d-TpA; II = pdC; III = pdT; IV = pdA; V = pdG. (a) Injection 2; t = 12.5 min. (b) Injection 10; t = 112.5 min. (c) Injection 20; t = 237.5 min. Accumulation of the dimer I indicates that the reaction is near completion.

the enzymatic digestion, reached a constant value after 120 min (see Fig. 1). From these observations, together with the base composition of the decamer | d-T-($CG_2A_3T_3$) |, we may conclude that this amount of pdG must correspond to a molar fraction of 2. The latter number was used as an intrinsic internal standard to normalize all values, which were obtained by dividing all peak areas by the corresponding extinction coefficients at 254 nm, to molar fractions (see Fig. 1).

Inspection of the curves in Fig. 1 reveals the correct sequence of the decamer. The first nucleoside 5'-phosphate to appear was pdG, followed by two successive pdT's. The sequence can therefore be written as d-T-(CGA_3T)-T-T-G. The next nucleoside 5'-phosphate released was another pdG, followed by two successive pdA's, thus affording the sequence d-T-(CAT)-A-A-G-T-T-G. At this stage the reaction slowed down. Thus, after 183 min pdC reached a molar fraction of 0.5 while pdT attained the level of 2.5 molar fractions after 215 min. These two molar fractions indicate that pdC precedes pdT, leaving pdA next to the nucleoside at the 5'-end, thus affording the sequence: d-TpApTpCpApApGpTpTpG.

Another way to derive the right sequence of the decamer is based on the use of the successive half-times of release of the nucleoside 5'-phosphates in the enzymatic process. Reading off the half-times, presented as vertical bars in Fig. 1, gives, together with the known base composition, the sequence of the decamer.

CONCLUSION

The data presented here show that automation of the sequence analysis of synthetic DNA fragments, which is based on HPLC in combination with PDE action, promises to be of general use to organic chemists working in this particular field.

ACKNOWLEDGEMENTS

This research was supported by The Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from The Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

REFERENCES

1 J. H. van Boom and J. F. M. de Rooij, *J. Chromatogr.*, 131 (1977) 169.
2 R. A. Henry, J. A. Schmidt and R. C. Williams, *J. Chromatogr. Sci.*, 11 (1973) 358.
3 F. Sanger, G. G. Brownlee and B. G. Barrell, *J. Mol. Biol.*, 13 (1965) 373.
4 R. Arentzen, C. A. A. van Boeckel, G. van der Marel and J. H. van Boom, *Synthesis*, (1979) 137.