Journal of Chromatography, 584 (1992) 135–139
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 6600

Behavior of large nucleic acids in reversed-phase high-performance liquid chromatography

Jean-Pierre Liautard

INSERM U-65, Departement Biologie Santé, Université Montpellier II, Place E. Bataillon, 34095 Montpellier (France)

ABSTRACT

Large nucleic acids can be separated by reversed-phase high-performance liquid chromatography. Analysis shows that the retention time depends not only on the chain length but also on the base composition and the secondary structure of the molecule. A model is proposed to interpret their behavior. This model, called "multiple-point interaction theory" is based on the observation that macromolecules are flexible and very large compared to the hydrophobic phase (octadecylsilane) of the column. It explains the behavior of large nucleic acids in terms of an equilibrium of the macromolecule between the two phases through a multiple-point attachment to the chromatographic matrix, the parameters of the equilibrium being both the hydrophobicity of the base and the number of attachment points. This model fits the experimental data and can be applied to all types of flexible macromolecules, especially proteins and nucleic acids, when they are chromatographed on reversed-phase columns. The model is used to explain the separation of nucleic acids of importance in molecular biology.

INTRODUCTION

The nucleic acids of biological origin (RNA or DNA) are mixtures of macromolecules that must often be fractionated to obtain information about their function. Numerous methods have been employed to obtain a good separation, *i.e.* electrophoresis, centrifugation [1] and more recently gel permeation chromatography [2] and anion-exchange high-performance liquid chromatography (HPLC) [3]. However, all these methods separate the nucleic acid according to a single criterion, the size. In some specific cases the spatial structure influences the electrophoretic mobility or the centrifugation [1]. The need for a method that can separate nucleic acids according to other criteria, as for example base composition or secondary structure, is obvious. For this purpose we have developed a reversed-phase chromatographic method for large macromolecular nucleic acids, and successfully used reversed-phase columns to analyze and isolate nucleic acids [4–6]. As HPLC methods are easy to perform and the results reproducible, we have used them to facilitate some of the burdensome steps involved in genetic engineering [7,8].

Large nucleic acids are macromolecules and their behavior is not as simple as that of small molecules. Such complex behavior has also been demonstrated for proteins [9,10] as well as for nucleic acids [4]. The aim of the present paper is to develop a model to explain the behavior of large nucleic acids. This model enables us to understand the mechanisms involved in the retention and the elution characteristics observed. The model is described, and potential applications to genetic engineering presented and discussed.

Correspondence to: J. P. Liautard, INSERM U-65, Departement Biologie Santé, C.P. 100, Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex, France.

0378-4347/92/$05.00

MULTIPLE-POINT INTERACTION MODEL

The reversed-phase chromatography of small organic molecules can be considered as liquid–liquid chromatography. The small molecules distribute themselves between the organic stationary phase (such as octadecylsilane) and the mobile phase (namely the aqueous phase). There is a rapid exchange between the two phases that is correctly described by the distribution coefficient [11].

However, this simple model cannot be applied to large macromolecules such as RNA or DNA. Indeed, the size of these molecules is far too large for the octadecylsilane organic phase. By comparing the size of a 1000-base RNA and the size of the organic stationary phase, it is obvious that RNA cannot enter the stationary phase. The octadecylsilane has a length of 21 Å when extended whereas a 1000-base RNA is around 3400 Å long.

In the case of a nucleic acid this phenomenon is amplified by the chemical nature of the nucleic acid. A nucleic acid can be divided into two structures: the bases and the phosphosugar backbone (see Fig. 1). The bases are roughly hydrophobic at neutral pH whereas the backbone is charged and largely hydrophilic. One may thus assume that only the bases enter the organic stationary phase while the backbone stays in the aqueous mobile phase. However, each individual base will behave as a small molecule and partitions between the two phases. Because there are numerous bases linked together the partition coefficient for n bases of a large nucleic acid will be:

$$K_n = K^n \tag{1}$$

where K is the partition coefficient for one base and n the number of bases involved. Using this hypothesis the retention volume will be:

$$V_r = K^n V_s + V_m \tag{2}$$

where V_r is the retention volume (retention time multiplied by flow), V_s is the volume of the stationary phase and V_m the dead volume of the column.

Experiments have shown [4,12] that it is necessary to elute large macromolecules with a gradient. According to Snyder *et al.* [13], if acetonitrile is used, the dependence of k' (capacity factor) on the fraction of acetonitrile (f) is approximated by

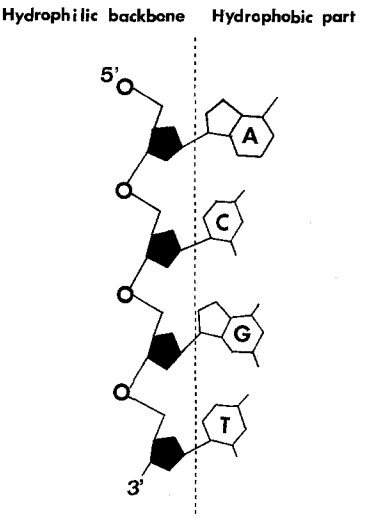

Fig. 1. Amphiphilic nature of the nucleic acids. The nucleic acids are composed of a hydrophilic backbone comprising phosphate groups, negatively charged, and sugar. Riboses are more hydrophilic than deoxyriboses. However, the bases formed a hydrophobic zone that can partition between the aqueous mobile phase and the stationary organic phase.

$$k' = k_0 \cdot 10^{(-Sf)} \tag{3}$$

where k_0 is the capacity factor at the beginning of the gradient, and S is the solvent strength factor. Using eqns. 2 and 3 and the definition of capacity factor it follows that:

$$V_r = (k_0 \cdot 10^{-Sf})^n \cdot V_m + V_m \tag{4}$$

where k_0 is the partition coefficient of the base without acetonitrile, f the fraction of acetonitrile in the elution buffer and S a constant called the elution strength of acetonitrile [13].

Simulations of this equation are presented in Fig. 2. The behavior of large nucleic acids is well simulated by this simple equation. When n, the number of bases, is large the molecule binds to the column, and the retention volume changes dramatically for a very small concentration change of acetonitrile. This means that adsorbed molecules are desorbed at a defined concentra-

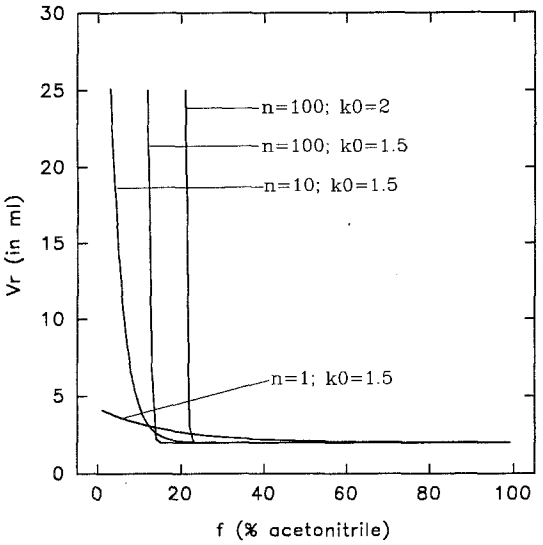

Fig. 2. Theoretical simulation of the retention behavior as a function of acetonitrile. Computer simulation of the behavior of nucleic acid according to eqn. 4. n represents the number of bases that interact with the stationary phase, k_0 represents the affinity of the base for the stationary phase when pure water is used as mobile phase, $S = 3$ is the elution strenth, f the fraction of acetonitrile, and $V_m = 2$ is the dead volume of the column.

tion of acetonitrile. This concentration depends only on the k_0 factor, *i.e.* the nature and the hydrophobicity of the base. When n is very large (>100) the number of bases does not influence the concentration of acetonitrile required for elution. Thus a heterogeneous-size homopolymer should elute as a single homogeneous peak. This is shown in Fig. 3. The consequence of the multiple-point model is that it should be possible to separate large nucleic acids according to their base composition, as shown in Fig. 3. Furthermore, the retention time depends only on the nature of the base; the more hydrophobic the base is, the longer the retention time.

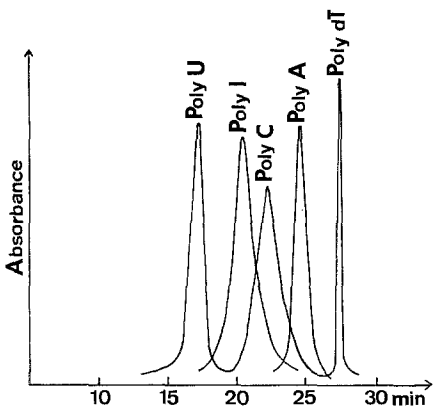

Fig. 3. Elution profile of large homopolymer nucleic acids. The polymers were obtained from Boehringer (Mannheim, Germany). They are synthetized using enzymatic methods and thus exhibit a very heterogeneous size; 90% of their size lies between 10^5 and $3 \cdot 10^5$ Da. Analysis was performed on a LiChrosorb RP18 column equilibrated with 100 mM ammonium acetate and eluted with acetonitrile-water (1:1, v/v) from 0 to 40% in 40 min, using a flow-rate of 1.0 ml/min.

SECONDARY STRUCTURE AND RETENTION BEHAVIOR

Table I compares the retention times of different RNAs. tRNAs are rapidly eluted but they are heterogeneous [4,5]. The snRNAs elute first, then rRNA 18S and 28S and finally poly(A)$^+$-globin messenger RNA. The elution of a type of RNAs (tRNAs, snRNA, rRNA) around the same retention time is striking. If the separation was

TABLE I

RETENTION TIMES OBSERVED FOR SOME NUCLEIC ACIDS

Nucleic acids were applied to the LiChrosorb RP18 column (250 mm × 40 mm I.D., Merck) equilibrated with 0.1 M ammonium acetate, pH 6.6 and eluted by increasing the acetonitrile–water (1:1, v/v) to 60% in 120 min at a flow-rate of 1.0 ml/min.

Nucleic acid	Retention time (min)	Base number[b]
poly(U)	24	$5 \cdot 10^3$–$2 \cdot 10^4$ b
poly(A)	45	$5 \cdot 10^3$–$3 \cdot 10^4$ b
tRNAs	20–25	70–90 b
18S rRNA	32	1750 b
28s rRNA	34	4850 b
snRNAs	36–38	104–196 b
glob. mRNA poly(A)+	44	900 b
glob. mRNA poly(A)−	33	700 b
ssDNA	56	$5 \cdot 10^2$–$3 \cdot 10^3$ b
dsDNA	26	$5 \cdot 10^2$–$3 \cdot 10^3$ bp
S.C. plasmid[a]	36	3103 bp
R. plasmid[a]	26	3103 bp

[a] S.C. plasmid = supercoiled plasmid; R. plasmid = relaxed plasmid.

[b] b = bases; bp = base pairs.

achieved only according to base composition, as it is for homopolymer, the separation should be far better. It must therefore be concluded that another property is involved in the retention behavior of nucleic acids. This may be specified by analyzing the retention of DNA. Single-stranded DNA is more retained that double-stranded DNA. This means that bases involved in pairing interaction do not interact as efficiently as the free bases. There is probably a competition between the interaction with stationary phase and with hydrogen binding to the complementary base. However, the fact that double-stranded DNA binds to the column implies that some of the bases are able to interact with the stationary phases. This can be explained by dynamic interaction between strands. This hypothesis is in agreement with the behavior of supercoiled plasmid. Indeed, supercoiling is equivalent to a partial denaturation in a relaxed plasmid. The supercoiled plasmid is thus in equilibrium between two forms, a true supercoiled form and a partially fused form. Thus the supercoiled plasmid should be more retained than the relaxed one, as shown in Table I.

APPLICATION OF REVERSED-PHASE CHROMATOGRAPHY TO GENETIC ENGINEERING

We will briefly discuss here the main applications of reversed-phase chromatography to genetic engineering.

mRNA can be chromatographed by reversed-phase chromatography [5,8]. However, the separation is not satisfactory because of the heterogeneity of the poly(A) tail. A single messenger RNA can elute over many fractions. However, the ribosomal RNAs are separated from the bulk of the mRNA. Furthermore, we have observed that mRNAs purified by reversed-phase chromatography are always more active in a cell-free translation system. Another application of this purification is the preparation of mRNA to realize a cDNA library. Contaminating chromosomal DNA present in the mRNA preparation will be cloned giving rise to unspliced transcripts. The purification by reversed-phase chromatography eliminates this DNA.

We have shown that plasmid purification can be achieved by reversed-phase chromatography [7,8]. The supercoiled plasmid elutes after cellular RNA and before cellular DNA. Crude plasmid preparation was obtained according to Birboin and Doly [14]. We have shown that the plasmid can be purified in two rounds of chromatography of 1 h each [7]. This method is relatively rapid and gives high-quality pure plasmid. However, due to the presence of denatured protein in the Birboin and Doly extracts, the life-time of the column is relatively short under these conditions. The elimination of denatured proteins is difficult by conventional methods (*i.e.* phenol extraction), where the extraction must be repeated three or four times in order to ensure complete elimination.

The rapid purification of single-stranded DNA has been used to isolate M13 DNA. Under certain conditions the M13 phages disrupt the bacteria, resulting in the appearance in the culture supernatant of rRNA and DNA from the bacteria. The phages were prepared as recommended by Maniatis *et al.* [1] and nucleic acids were purified by phenol extraction. The method is based on the difference of retention times exhibited by single-

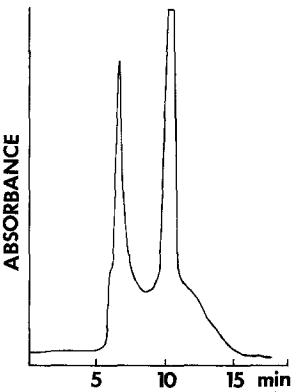

Fig. 4. Purification of single-stranded DNA from M13 phage. M13 was prepared as recommended by Maniatis *et al.* [1]. Nucleic acids were extracted with phenol and precipitated with ethanol. The sample was applied to a LiChrosorb RP18 column equilibrated with 100 m*M* ammonium acetate (pH 6.5) and eluted with acetonitrile–water (1:1, v/v) from 0 to 40% in 15 min, using a flow-rate of 1.0 ml/min. Fractions were collected and analyzed by agarose gel electrophoresis in order to identify the materials.

stranded DNA and double-stranded DNA [6]. The single-stranded genome from M13 can be separated from *E. coli* DNA or RNA contaminating some preparations in 10 min (Fig. 4). Electrophoretic analysis of the first-eluted peak reveals rRNA and double-stranded DNA and the second peak exhibits pure M13 single-stranded DNA. The DNA purified in this way is used for sequencing with a very low background.

REFERENCES

1 T. Maniatis, E. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988.
2 J. Kruppa, L. Graeve, A. Bauche and P. Földi, *LC Mag.*, 2 (1984) 848–853.
3 A. Pingoud, A. Fliess and V. Pingoud, in R. W. A. Oliver (Editor), *HPLC of Macromolecules*, IRL Press, Oxford, 1989, pp. 183–208.
4 S. Garcia and J. P. Liautard, *J. Chromatogr. Sci.*, 21 (1983) 398–404.
5 S. Garcia and J. P. Liautard, *J. Chromatogr.*, 296 (1984) 355–362.
6 J. P. Liautard, *J. Chromatogr.*, 285 (1984) 221–225.
7 S. Colote, C. Ferraz and J. P. Liautard, *Anal. Biochem.*, 154 (1986) 15–20.
8 S. Colote, C. Ferraz, J. Sri Widada and J. P. Liautard, *Biosciences*, 8 (1989) 42–45.
9 R. Barford, B. Sluwinski, A. Breyer and J. Rothbait, *J. Chromatogr.*, 235 (1982) 281–288.
10 F. Regnier, *Science*, 222 (1983) 245–252.
11 R. E. Majors, in C. F. Simpson (Editor), *Practical High Performance Liquid Chromatography*, Heyden & Son, London, 1978, pp. 109–118.
12 P. H. Coran, in R. W. A. Oliver (Editor), *HPLC of Macromolecules*, IRL Press, Oxford, 1989, pp. 127–157.
13 L. R. Snyder, J. W. Dolan and J. R. Gant, *J. Chromatogr.*, 165 (1979) 3–30.
14 H. Birboin and J. Doly, *Nucleic Acids Res.*, 7 (1979) 1513–1524.