*Journal of Chromatography*, 336 (1984) 189—198
*Biomedical Applications*
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMBIO. 2291

# PURIFICATION OF SYNTHETIC OLIGODEOXYRIBONUCLEOTIDES BY ION-EXCHANGE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

DENIS SCANLON*,*, JIM HARALAMBIDIS, CHRISTINA SOUTHWELL, JANICE TURTON and GEOFFREY TREGEAR

*Howard Florey Institute of Experimental Physiology and Medicine, University of Melbourne, Parkville 3052 (Australia)*

## SUMMARY

Synthetic oligodeoxyribonucleotides ranging from 11 to 37 nucleotides in length and with varying base compositions, prepared by both the phosphotriester and phosphite procedures, have been purified by ion-exchange high-performance liquid chromatography on Whatman Partisil 10/SAX columns using phosphate buffer gradients. The effects of different buffer systems on elution times and resolution have been evaluated. Oligomer composition and length had a marked effect on the resolution achieved. In general the use of formamide buffers gave the best results, particularly in the case of 2′-deoxyguanosine-rich sequences. These methods have also been successfully applied to the purification of mixtures of synthetic oligodeoxynucleotides.

## INTRODUCTION

Oligonucleotides have been shown to be extremely powerful tools in the field of recombinant DNA technology. They can be used as primers [1, 2] or probes [3, 4] to isolate cDNA coding for proteins or peptides of interest. They can be used as linkers and adapters for the modification of the ends of cDNA fragments, as in the generation of "sticky ends" for ligation reactions [5—7]. Large quantitities of extremely pure synthetic oligonucleotides (a few milligrams) can provide valuable DNA structural information by either X-ray crystallography [8] or nuclear magnetic resonance analysis of the samples [9]. Oligonucleotides can also be used for in vitro site-specific mutagenesis [10] or as aids in sequencing DNA [11].

The synthetic methodologies currently available to assemble these compounds are extremely elegant, rapid techniques. This is evidenced by the total synthesis of the gene encoding α-leukocyte interferon. The synthesis of this molecule, which contains 514 base pairs, was reported by Edge et al. in 1981 [12].

*Present address: Ludwig Institute for Cancer Research, Melbourne Tumour Biology Branch, Post Office Royal Melbourne Hospital, Melbourne, Victoria 3050, Australia.

0378-4347/84/$03.00 

Two types of chemistry are available to perform oligonucleotide synthesis: the phosphotriester method and the phosphite procedure. Each method is capable of assembling oligonucleotides up to 30 bases in length within a few days. A major problem is the purification of the desired oligomer from the mixture of truncated failure sequences which are produced during the synthesis and are present in the crude deprotected product. Methods which have been reported for the purification of these compounds include anion-exchange high-performance liquid chromatography (HPLC) [13—16], reversed-phase HPLC [17—20], gel electrophoresis [21] and preparative thin-layer chromatography (TLC) on glass-backed silica gel plates [22].

In this paper we describe the anion-exchange HPLC purification of synthetic oligonucleotides assembled by both phosphotriester and phosphite methodologies. The purification of single oligonucleotide sequences and a series of synthetic oligonucleotide mixtures of increasing complexity (to be used as probes) is described. A comparison is also made of buffer systems containing 5% ethanol, 30% acetonitrile, 30% formamide and 60% formamide.

## MATERIALS AND METHODS

Oligonucleotides synthesised by the phosphotriester method were assembled on 1% cross-linked polystyrene resin by a dimer block coupling approach [17]. The crude products were deprotected and released from the resin support (50 mg, 0.19 mmol/g starting functionality) by standard means [13] and dissolved in water (5 ml) ready for HPLC.

Oligonucleotides synthesised by the phosphite method were assembled on silical gel (Fractosil 500) by a monomer coupling approach [23]. The crude products were deprotected and released from the resin (100 mg 0.07 mmol/g starting functionality) by standard methods [18] and dissolved in water (5 ml) ready for HPLC.

The HPLC columns used were Whatman Partisil 10/SAX columns (25 × 0.46 cm, 10 $\mu$m particle size, 20,000 theoretical plates per m). Chromatography was performed on an Altex modular system (two Model 110A pumps, Model 420 system controller, a Hitachi 100-40 ultraviolet (UV) detector and Curken chart recorder). Partisil 10/SAX columns were eluted at a flow-rate of 1 ml/min with potassium dihydrogen phosphate combined with either ethanol, acetonitrile or formamide as follows: (i) 0.001—0.2 *M* $KH_2PO_4$ (pH 6.5, 5% ethanol) over 40 min; (ii) 0.001—0.2 *M* $KH_2PO_4$ (pH 6.5, 30% acetonitrile) over 40 min; (iii) 0.001—0.3 *M* $KH_2PO_4$ (pH 6.3, 30% formamide) over 60 min; or (iv) 0.001—0.3 *M* $KH_2PO_4$ (pH 6.3, 60% formamide) over 60 min.

All buffers were filtered through Millipore Type FH 0.5-$\mu$m filters prior to use. The crude oligonucleotide samples were filtered using a BioAnalytical Systems MF-1 filter kit (0.2-$\mu$m nitrocellulose filters).

Product peaks were collected from the anion-exchange HPLC runs, dialysed in Spectrapor 6 membrane tubing (obtained from Spectrum Medical Industry, molecular weight cut off 2000, four water changes) and lyophilized to dryness. The samples were then resuspended in sterile water and quantified by reading the absorbance at 260 nm. Each of the single sequences were characterised by the Maxam—Gilbert sequencing technique [24]. An example of the

characterisation technique is detailed in Fig. 2 where the sequencing data for an oligonucleotide 37 bases long are shown (space limitations prevent all sequencing data to be shown). The mixture sequences were characterised by $^{32}$P-radiolabelling the 5′ end of the oligonucleotide ([$\gamma^{32}$P]ATP and T4 polynucleotide kinase) and sizing with respect to oligonucleotides of known length on 18% polyacrylamide gels in the presence of 7 *M* urea (data not shown).

## RESULTS AND DISCUSSION

In the crude oligonucleotide mixture the desired product is the longest species and is usually contaminated with shorter truncated sequences. The

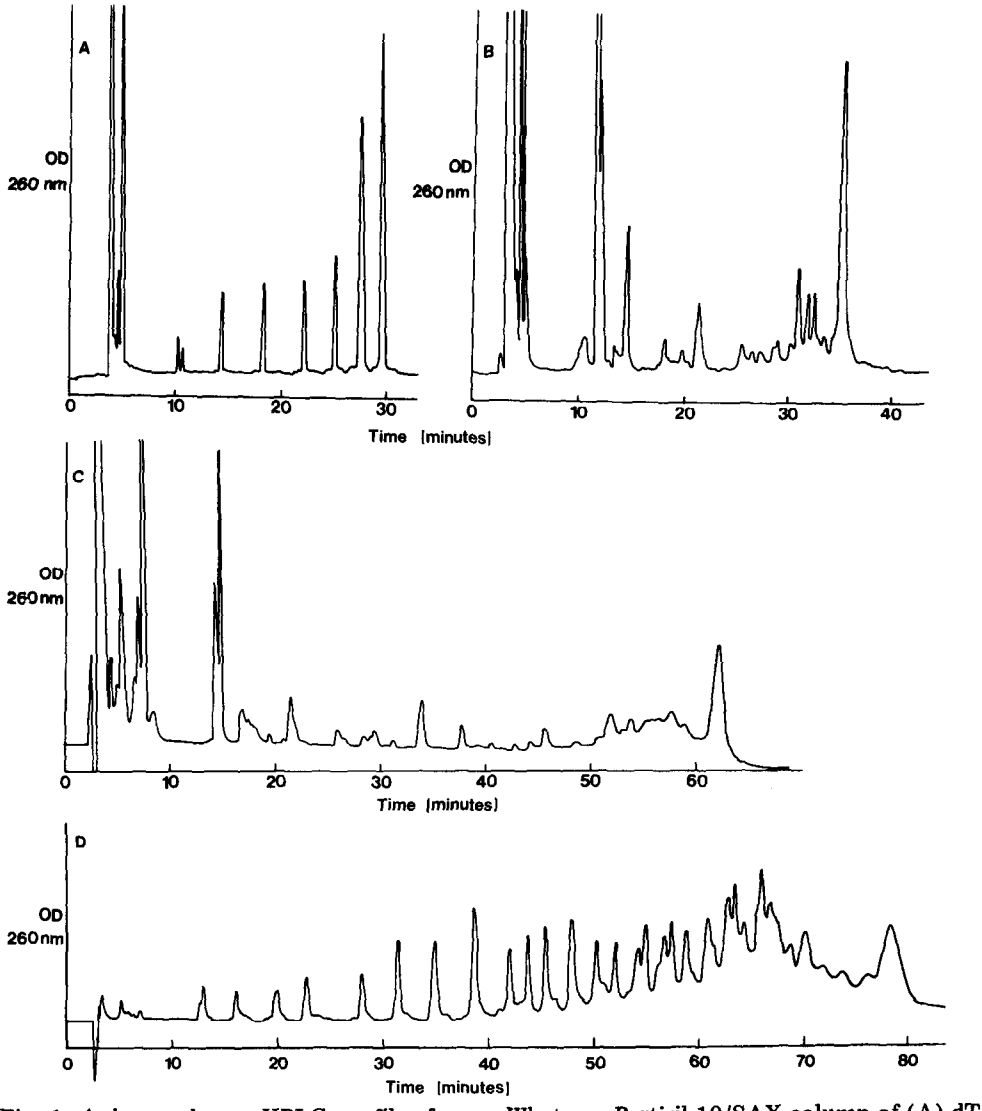


Fig. 1. Anion-exchange HPLC profiles from a Whatman Partisil 10/SAX column of (A) $dT_{15}$ with buffer i as eluent; (B) the 11′ mer d(ATTTCAACCCA) with buffer i as eluent; (C) the 17′ mer d(ATACTGCAGCTGCTGTA) with buffer iv as eluent; and (D) the 37′ mer d(CCTCCCATTTCCCTGGATCTCACCTTCCATCTCCTCC) with buffer iv as eluent.

product, therefore, contains more negatively charged phosphate moieties than any other species in the crude product mixture and will bind most strongly to the anion-exchange column. This means that the desired product should be the last peak to elute.

The following examples illustrate the utility of anion-exchange HPLC for the purification of a range of different synthetic oligonucleotide sequences.

*Purification of single sequences: $dT_{15}$, d(ATTTCAACCCA), d(ATACTGCAGCTGCTGTA) and d(CCTCCCATTTCCCTGGATCTCACCTTCCATCTCCTCC)*

Pentadecathymidilic acid ($dT_{15}$) was synthesized by a phosphotriester dimer block approach and was purified by the conditions detailed for buffer i above.

The effectiveness of the chromatography method is evidenced by this synthesis. The product $dT_{15}$ elutes as the major peak and the truncated sequences $dT_{13}$, $dT_{11}$, $dT_9$, $dT_7$, $dT_5$ and $dT_3$ are resolved into sharp peaks eluting sequentially earlier in the HPLC profile (Fig. 1A).

In other examples the 11′ mer, d(ATTTCAACCCA), and the 17′ mer, d(ATACTGCAGCTGCTGTA), eluted as the major product with very little truncated by-product (Fig. 1B and C). The 11′ mer was purified by the ethanol-based buffer i and the 17′ mer by the formamide-based buffer iv. The different properties and the advantages of each of these buffer systems is discussed in more detail below.

The power of the anion-exchange chromatography method is demonstrated by the purification of the oligonucleotide d(CCTCCCATTTCCCTGGATCTCACCTTCCATCTCCTCC) which contains 37 residues. The oligomer was purified using buffer iv and while there is evidence of a number of truncated products in the HPLC profile the product peak can be seen to be a major component of the crude mixture (Fig. 1D).

Radiolabelling the product peak from the HPLC purification with $^{32}P$ and analysing it on an 18% polyacrylamide gel showed the sample to be homogenous (Fig. 2A). Maxam–Gilbert sequencing analysis confirmed the structure of the 37-base oligonucleotide (Fig. 2B).

Sequences 30 bases long are now routinely synthesised and characterised in our laboratory by the methods described above.

*Purification of oligonucleotide mixtures: d(TTZGTCATCTC), d(GGCTTYTTXTC), d(GGXTTYTTXTC) and d(GCXTTNGTCATXTC) (where X=T/C, Y=A/G, Z=A/C and N=T/C/A/G)*

A series of oligonucleotide mixtures was synthesized for use as primers and probes. The applications of the mixtures will not be discussed in detail here, but a sample of the mixture syntheses containing a range of mixture complexities will be presented. Fig. 3 shows the HPLC profiles (buffer ii) of mixtures of two (Fig. 3A), four (Fig. 3B), eight (Fig. 3C) and sixteen (Fig. 3D) components. The oligonucleotide components, in the case of the mixtures containing two and four sequences [d(TTZGTCATCTC) and d(GGCTTYTTX-TC)] could be resolved into individual sequences. This was generally true for each two-component mixture synthesized, but other four-component mixtures (data not shown) could not be totally resolved. The eight- and sixteen-com-

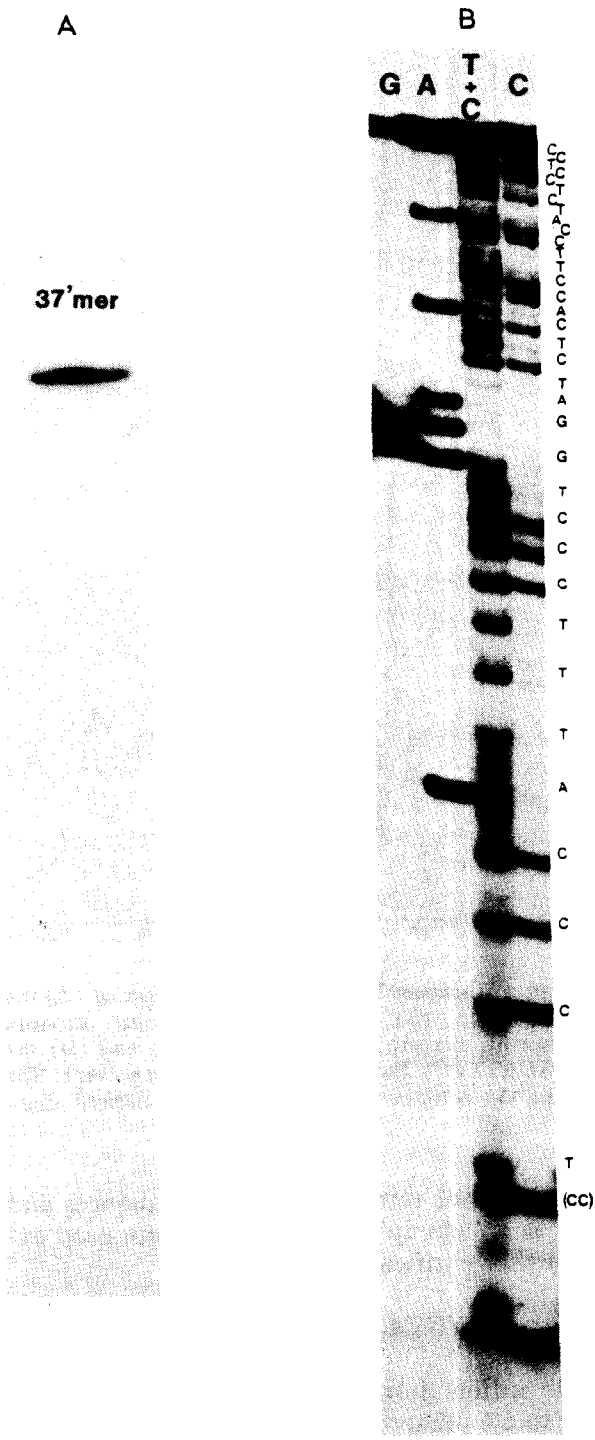


Fig. 2. (A) HPLC-purified 37′ mer radiolabelled with $^{32}P$ and analysed on an 18% polyacryl amide gel; (B) Maxam—Gilbert sequencing of the 37-base oligonucleotide.

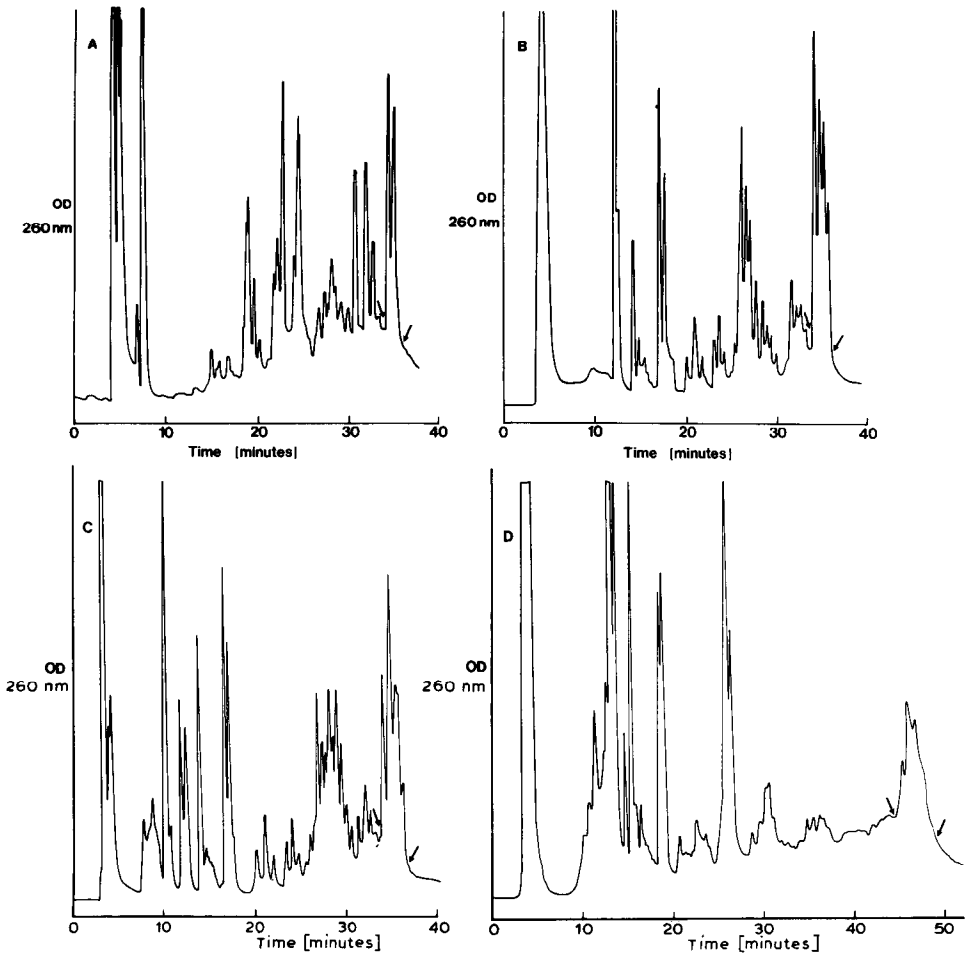


Fig. 3. Anion-exchange HPLC profiles from a Whatman Partisil 10/SAX column of (A) the two-component mixture d(TTZGTCATCTC); (B) the four-component mixture d(GGCTTYTTXTC); (C) the eight-component mixture d(GGXTTYTTXTC) and (D) the sixteen-component mixture d(GCXTTNGTCATXTC). Buffer ii was the eluting solvent. The arrows indicate the fractions collected from the column which contained the desired oligonucleotide mixtures.

ponent mixture HPLC purifications did not resolve the single sequences and hence the products were collected as a "clump" of peaks and were used experimentally in that form without further purification.

*Comparison of buffer conditions: d(GTAAAACGACGGCCAGT) and d(CAGGGGTTTTGGGCCAAAG)*

It was found that 30% acetonitrile buffers (buffer ii), which were used in the mixture purifications, were superior to 5% ethanol systems (buffer i), especially if the theoretical plate count of the Partisil 10/SAX column was low (in the range 16,000 to 20,000 plates per m). This, therefore, was the buffer of choice for the purification of oligonucleotides up to 15 bases long. The use of this

buffer was limited, however, as the maximum possible phosphate concentration was 0.2 *M*. HPLC profiles of sequences such as the 17′ mer d(GTAAAACGAC-GGCCAGT) when purified in buffer ii showed diminished resolution and poor peak shape (Fig. 4A). Substituting 60% formamide for 30% acetonitrile allowed higher phosphate concentrations (0.3 *M*) and the resolution was substantially improved (Fig. 4B). There was a significant difference in the yields of product obtained from 60% formamide and 30% acetonitrile buffer systems. In the example shown in Fig. 4 approx. 50% more solute product was obtained from the formamide-based system.

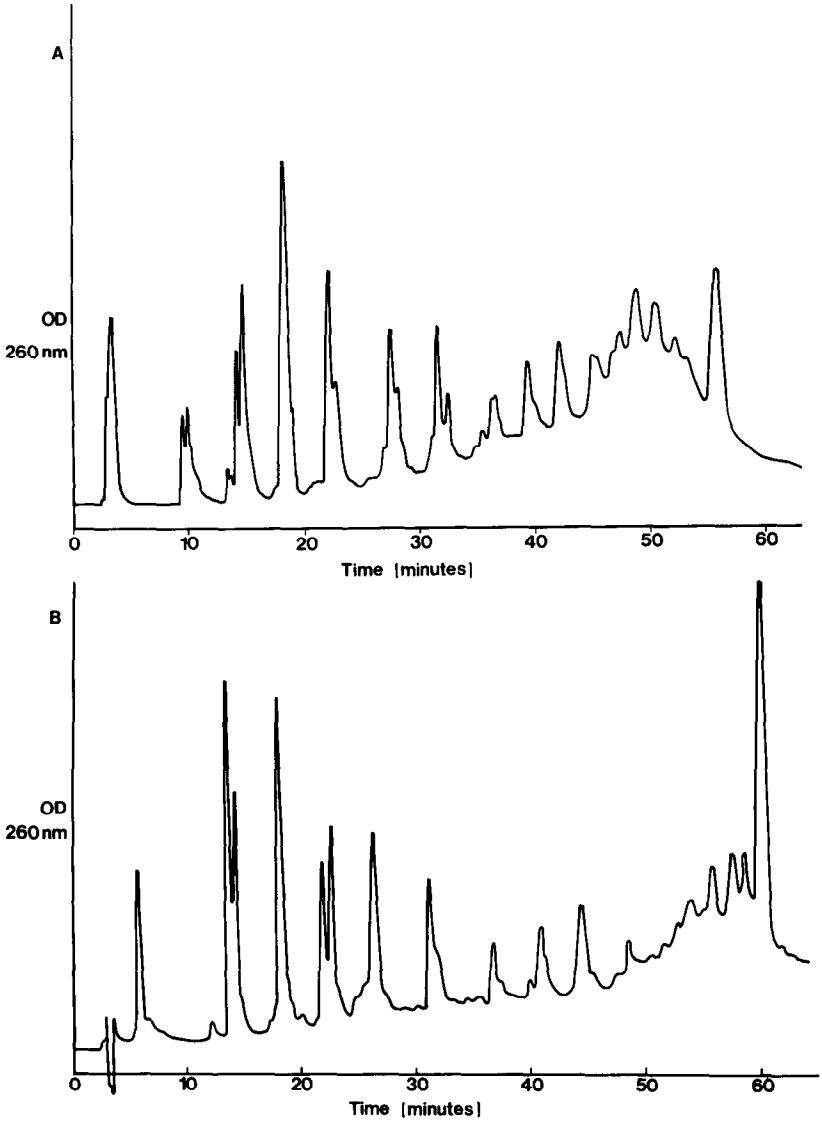


Fig. 4. Anion-exchange HPLC profiles from a Whatman Partisil 10/SAX column of the 17′ mer d(GTAAAACGACGGCCAGT) using (A) buffer ii as eluent and (B) buffer iv as eluent.

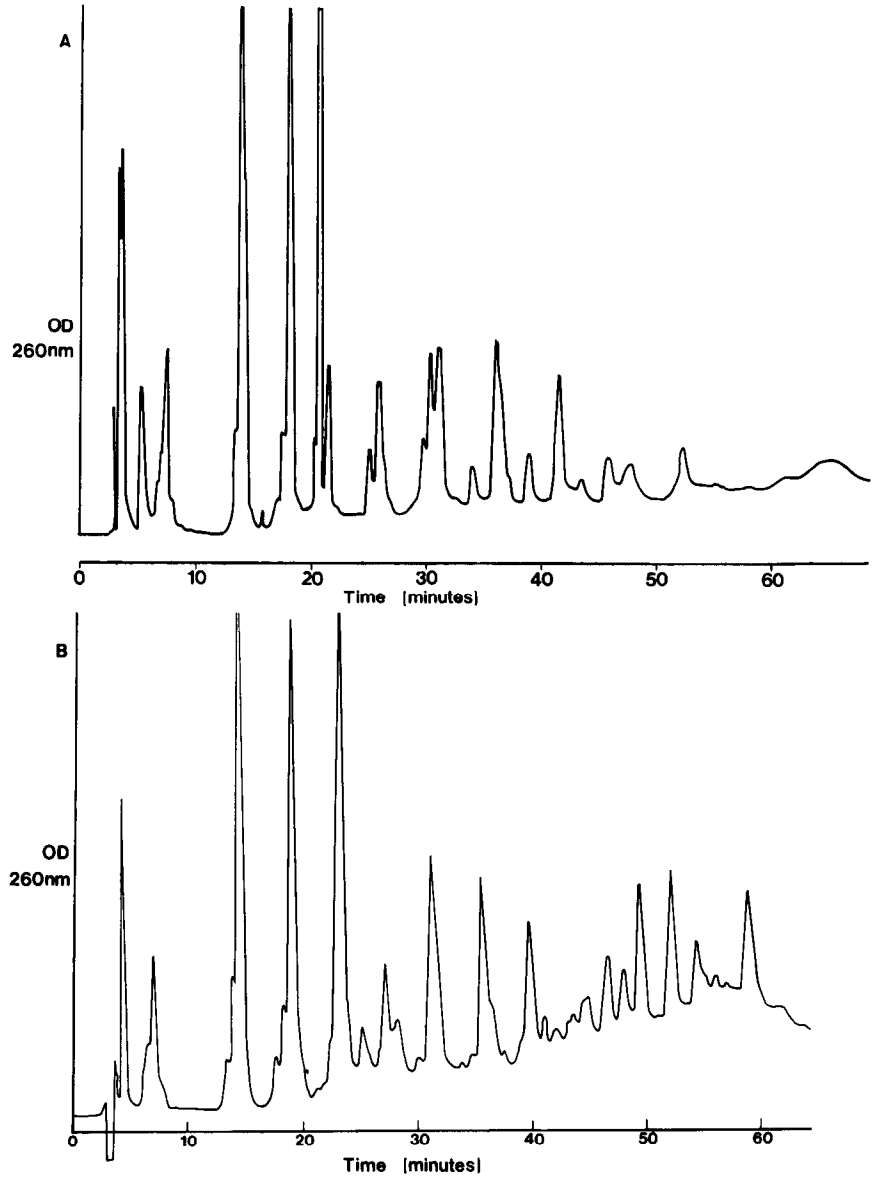


Fig. 5. Anion-exchange HPLC profiles from a Whatman Partisil 10/SAX column of the 19′ mer d(CAGGGGTTTTGGGCCAAAG) using (A) buffer iii as eluent and (B) buffer iv as eluent.

The effectiveness of differing concentrations of formamide in the HPLC eluting buffer was tested in the purification of the 2′-deoxyguanosine-rich sequence d(CAGGGGTTTTGGGCCAAAG). It is well documented that G-rich sequences, and especially those containing extended tracts of three or more consecutive G residues, are difficult to purify by anion-exchange HPLC [15, 25]. The 19′ mer is extremely G-rich and contains a tract of four consecutive 2′-deoxyguanosine residues. HPLC purifications of this oligomer in buffer iii (30% formamide) and buffer iv (60% formamide) are shown in Fig. 5. The buffer

iii HPLC profile showed a very broad product peak while the buffer iv HPLC profile showed the 19′ mer eluting as a sharp peak and to be a major component of the crude synthetic mixture. Maxam—Gilbert sequencing [24] confirmed the sequence of 19′ mer. The solute recovery yield of guanosine-rich sequences such as d(CAGGGGTTTTGGGCCAAG) was 60—70% of the expected value whereas for non-guanosine-rich sequences such as d(GTAAAACG-ACGGCCAGT) the value was greater than 90%.

## CONCLUSIONS

For the routine purification of synthetic oligonucleotides we recommend the use of anion-exchange HPLC in 60% formamide buffer systems containing potassium phosphate. This system has been found to be superior to ethanol- and acetonitrile-based systems especially for the purification of longer nucleotide sequences (greater than 30 residues) and can be used for the purification of both single oligonucleotide sequences and oligonucleotide mixtures. One practical disadvantage of the use of formamide buffers is that the life of the Partisil 10/SAX column is generally decreased. Ethanol- or acetonitrile-based solvent systems allow a Partisil 10/SAX column lifetime of approx. four to six months. Formamide-based systems, however, reduce the column lifetime to less than two months.

The methods described here detail a rapid technique which can be used by itself or in combination with gel electrophoresis, reversed-phase HPLC and preparative TLC, for the chromatographic purification of oligonucleotides. The method is particularly useful for the purification of synthetic oligonucleotides prepared by automated solid-phase procedures.

## ACKNOWLEDGEMENTS

This work was supported by the N.H. & M.R.C. of Australia and by N.I.H. Grant HD 11908.

## REFERENCES

1 M. Houghton, A.G. Stewart, S.M. Doel, J.S. Emtage, M.A.W. Eaton, J.C. Smith, T.P. Patel, H.M. Lewis, A.G. Porter, J.R. Birch, T. Cartwright and N.H. Carey, Nucleic Acids Res., 8 (1980) 1913.
2 J. Haley, P. Hudson, D. Scanlon, M. John, M. Cronk, J. Shine, G. Tregear and H. Niall, DNA, 1 (1982) 155.
3 R.B. Wallace, M.J. Johnson, T. Hirose, T. Miyake, E.H. Kawashima and K. Itakura, Nucleic Acids Res., 9 (1981) 879.
4 D.E. Woods, A.F. Markham, A.T. Ricker, G. Goldberger and H.R. Colten, Proc. Nat. Acad. Sci. U.S., 79 (1982) 5661.
5 C.P. Bahl, R. Wu, R. Brousseau, A.K. Sood, H.M. Hsiung and S.A. Narang, Biochem. Biophys. Res. Commun., 81 (1978) 695.
6 E. Ohtsuka, R. Fukumoto and M. Ikehara, Chem. Pharm. Bull., 28 (1980) 80.
7 Yu A. Berlin, N.M. Zvonok and A.L. Kayushin, Bioorg. Khim., 6 (1980) 1182.
8 A.H.-H. Wang, S. Fujii, J.H. van Boom and A. Rich, Proc. Nat. Acad. Sci. U.S., 79 (1982) 3968.
9 E.R.P. Zuiderweg, R.M. Scheek, G. Veeneman, J.H. van Boom, R. Kaptein, H. Rüterjans and K. Beyreuther, Nucleic Acids Res., 9 (1981) 6553.

10 M. Smith and S. Gillam, in J.K. Setlow and Hollaender (Editors), Genetic Engineering, Vol. 3, Plenum Press, New York, 1981, p. 1.
11 M.L. Duckworth, M.J. Gait, R. Goelet, G.F. Hong, M. Singh and R.C. Titmas, Nucleic Acids Res., 9 (1981) 1691.
12 M.D. Edge, A.R. Greene, G.R. Heathcliffe, P.A. Meacock, W. Schuch, D.B. Scanlon, T.C. Atkinson, C.R. Newton and A.F. Markham, Nature, 292 (1981) 756.
13 M.J. Gait, H.W.D. Matthes, M. Singh, B.S. Sproat and R.C. Titmas, Nucleic Acids Res., 10 (1982) 6243.
14 K. Miyoshi, R. Arentzen, T. Huang and K. Itakura, Nucleic Acids Res., 8 (1980) 5507.
15 C.R. Newton, A.R. Greene, G.R. Heathcliffe, T.C. Atkinson, D. Holland, A.F. Markham and M.D. Edge, Anal. Biochem., 129 (1983) 22.
16 J.E. Marugg, M. Tromp, P. Jhurani, C.F. Hoyng, G.A. van der Marel and J.H. van Boom, Tetrahedron, 40 (1984) 73.
17 Y. Ike, S. Ikuta, M. Sato, T. Huang and K. Itakura, Nucleic Acids Res., 11 (1983) 477.
18 M.D. Matteucci and M.H. Caruthers, J. Amer. Chem. Soc., 103 (1981) 3185.
19 R. Frank, W. Heikens, G. Heisterberg-Moutsis and H. Blöcker, Nucleic Acids Res., 11 (1983) 4365.
20 H. Köster, J. Biernat, J. McManus, A. Wolter, A. Stumpe, Ch.K. Narang and N.D. Sinha, Tetrahedron, 40 (1984) 103.
21 S.P. Adams, K.S. Kauka, E.J. Wykes, S.B. Holden and G.R. Galluppi, J. Amer. Chem. Soc., 105 (1983) 661.
22 G. Alvarado-Urbina, G.M. Sathe, W-C. Liu, M.F. Gillen, P.D. Duck, R. Bender and K.K. Ogilvie, Science, 214 (1981) 270.
23 L.J. McBride and M.H. Caruthers, Tetrahedron Lett., 24 (1983) 245.
24 A.M. Maxam and W. Gilbert, Methods Enzymol., 65 (1980) 499.
25 A.F. Markham, M.D. Edge, T.C. Atkinson, A.R. Greene, G.R. Heathcliffe, C.R. Newton and D. Scanlon, Nucleic Acids Res., 8 (1980) 5193.