THERMAL ELUTION OF OLIGONUCLEOTIDES ON CELLULOSE
COLUMNS CONTAINING OLIGODEOXYRIBONUCLEOTIDES
OF DEFINED LENGTH AND SEQUENCE

by

Caroline R. Astell

B.Sc., The University of British Columbia, 1964.
M.Sc., The University of British Columbia, 1966.

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in the Department of
Biochemistry

We accept this thesis as conforming to
the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December, 1970.
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Biochemistry

The University of British Columbia
Vancouver 8, Canada

Date December 22, 1970
ABSTRACT

The high degree of specificity of base interactions of complementary polynucleotides is a fundamental property of biological systems. Through interactions of this type, genetic information is able to flow, to succeeding generations by replication of the genome, and from deoxyribonucleic acid and ribonucleic acids to proteins, by the steps of transcription and translation.

The experiments reported in this thesis are concerned with a study of the feasibility of developing a method suitable for the isolation of naturally occurring polynucleotides by hybridization of the polynucleotide with a complementary oligodeoxyribonucleotide. Homo-oligodeoxyribonucleotides (lengths up to the dodecanucleotides) of thymidylic and deoxyadenylic acids have been prepared by chemical procedures, as have oligodeoxyribonucleotides of the repeating, complementary trinucleotide base sequences, $d(pTpTpC)_n$, $d(pCpTpT)_n$, and $d(pApApG)_n$, ($n = 1$ to $4$). Various oligonucleotides have been linked covalently to cellulose, an insoluble, inert matrix, via the 5' phosphoryl group on the oligodeoxyribonucleotides, using the water-soluble carbodiimide, $N$-cyclohexyl-$N'\beta(4$-methylmorpholinium)ethylcarbodiimide p-toluenesulfonate. The resultant oligonucleotide-celluloses, in the form of small
columns, were examined for their ability to retain complementary oligomers.

The retained oligonucleotides can be eluted conveniently with a linear temperature gradient. The data indicate that at least all but one nucleotide, and possibly the entire oligonucleotide attached to the cellulose is capable of hydrogen bonding with its complementary sequence. The resolution obtained with these columns is such that oligonucleotides, differing in length by one nucleotide residue may be fractionated. The capacity is $1/3$ to $1/4$ of the theoretical amount.

Preliminary experiments suggest that these oligonucleotide-celluloses are capable of selectively removing a complementary sequence of nucleic acid from a mixture of nucleic acids.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xlii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xvii</td>
</tr>
<tr>
<td>ABBREVIATIONS USED</td>
<td>xviii</td>
</tr>
<tr>
<td>GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Part I: Chemical Synthesis of Oligodeoxyribonucleotides of Defined Length and Sequence</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>19</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>27</td>
</tr>
<tr>
<td>Nucleotides and Chemicals</td>
<td>27</td>
</tr>
<tr>
<td>Methods</td>
<td>29</td>
</tr>
<tr>
<td>1) Conversion of nucleotides to pyridinium salts</td>
<td>29</td>
</tr>
<tr>
<td>2) Anhydrous conditions</td>
<td>30</td>
</tr>
<tr>
<td>3) Stopping reactions</td>
<td>30</td>
</tr>
<tr>
<td>4) Condensation reaction workup</td>
<td>30</td>
</tr>
<tr>
<td>a) DCC reactions</td>
<td>30</td>
</tr>
<tr>
<td>b) sulfonyl chloride reactions.</td>
<td>31</td>
</tr>
<tr>
<td>c) removal of terminal protecting groups</td>
<td>31</td>
</tr>
<tr>
<td>Table of Contents (Continued)</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5) Anion exchange chromatography of protected nucleotides.</td>
<td>31</td>
</tr>
<tr>
<td>6) Concentration of protected nucleotides following anion exchange chromatography.</td>
<td>32</td>
</tr>
<tr>
<td>7) Removal of all protecting groups</td>
<td>32</td>
</tr>
<tr>
<td>8) Characterization of protected di- and trinucleotides.</td>
<td>32</td>
</tr>
<tr>
<td>9) Purification and characterization of longer oligonucleotides from the polymerization of protected tri-nucleotides</td>
<td>33</td>
</tr>
<tr>
<td>10) Paper chromatography</td>
<td>34</td>
</tr>
<tr>
<td>EXPERIMENTAL AND RESULTS.</td>
<td>35</td>
</tr>
<tr>
<td>1) Synthesis of oligodeoxythymidylylate polymers terminated with a 5'-phosphate and 3'-hydroxyl ( (d(pT)_n \quad n = 2 \text{ to } 12) ).</td>
<td>35</td>
</tr>
<tr>
<td>a) Polymerization of d-pT</td>
<td>35</td>
</tr>
<tr>
<td>b) Polymerization of the trinucleotide, d-pTpTpT</td>
<td>38</td>
</tr>
<tr>
<td>(i) Stepwise synthesis of (d(pT)_3)</td>
<td>38</td>
</tr>
<tr>
<td>(ii) Polymerization of the trinucleotide d-pTpTpT</td>
<td>42</td>
</tr>
<tr>
<td>Table of Contents (Continued)</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2) Synthesis of oligonucleotides of repeating sequence d(pTpTpC)&lt;sub&gt;n&lt;/sub&gt; (n = 1 to 4)</td>
<td>47</td>
</tr>
<tr>
<td>a) Synthesis of the protected trinucleotide, d-pTpTpC&lt;sup&gt;An&lt;/sup&gt;</td>
<td>47</td>
</tr>
<tr>
<td>b) Polymerization of the trinucleotide, d-pTpTpC&lt;sup&gt;An&lt;/sup&gt;</td>
<td>51</td>
</tr>
<tr>
<td>3) Synthesis of oligonucleotides of the sequence d(pCpTpT)&lt;sub&gt;n&lt;/sub&gt; (n = 1 to 4)</td>
<td>51</td>
</tr>
<tr>
<td>a) Synthesis of d-pC&lt;sup&gt;An&lt;/sup&gt;pTpT.</td>
<td>51</td>
</tr>
<tr>
<td>b) Polymerization of d-pC&lt;sup&gt;An&lt;/sup&gt;pTpT.</td>
<td>56</td>
</tr>
<tr>
<td>4) Synthesis of oligodeoxyadenylate polymers, d(pA)&lt;sub&gt;n&lt;/sub&gt; (n = 2 to 12)</td>
<td>58</td>
</tr>
<tr>
<td>5) Synthesis of oligodeoxyribonucleotides of repeating sequence d(pApApG)&lt;sub&gt;n&lt;/sub&gt; (n = 1 to 4)</td>
<td>58</td>
</tr>
<tr>
<td>a) Synthesis of the protected trinucleotide d&lt;sub&gt;pA&lt;/sub&gt;B&lt;sub&gt;z&lt;/sub&gt;p&lt;sub&gt;A&lt;/sub&gt;B&lt;sub&gt;z&lt;/sub&gt;p&lt;sub&gt;G&lt;/sub&gt;B&lt;sub&gt;z&lt;/sub&gt;</td>
<td>58</td>
</tr>
<tr>
<td>b) Polymerization of the protected trinucleotide, d&lt;sub&gt;pA&lt;/sub&gt;B&lt;sub&gt;z&lt;/sub&gt;p&lt;sub&gt;A&lt;/sub&gt;B&lt;sub&gt;z&lt;/sub&gt;p&lt;sub&gt;G&lt;/sub&gt;B&lt;sub&gt;z&lt;/sub&gt;</td>
<td>67</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>71</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

Part II: Thermal Elution of Oligonucleotides
On Cellulose Columns Containing
Oligonucleotides of Defined Length
And Sequence

INTRODUCTION ............................................. 74
MATERIALS AND METHODS ...................................... 87
A. Interaction of Complementary Oligo-
nucleotides in Solution .................................. 87
   1) Estimation of average extinction coefficient
      per base at 25° for oligonucleotides of
      mixed base sequence .................................. 87
   2) Thermal denaturation curves for
      complementary interactions of the
      type d(pT)ₙ.d(pA)ₙ and d(pT)ₙ.poly A ................. 89
   3) Tm and % hypochromicity .............................. 91
B. Synthesis and Properties of Oligo-
nucleotide-Celluloses ................................... 93
   1) Synthesis of oligonucleotide-celluloses...
      Reaction of oligonucleotides on
      cellulose paper with a water-soluble
      carbodiimide .......................................... 93
<table>
<thead>
<tr>
<th>Table of Contents (Continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Page</strong></td>
</tr>
<tr>
<td>2) Possible side reactions in the reaction of oligonucleotides with the carbodiimide.</td>
</tr>
<tr>
<td>3) Preparation of column matrix and procedure for packing column.</td>
</tr>
<tr>
<td>4) Procedure for running oligonucleotide-cellulose columns.</td>
</tr>
<tr>
<td>a) Thermal elution of oligonucleotides.</td>
</tr>
<tr>
<td>b) Other methods of elution of oligonucleotides.</td>
</tr>
<tr>
<td>5) Preparation of oligonucleotides of the series (d(A)_n) and (r(A)_n).</td>
</tr>
<tr>
<td>6) Isolation of Salmon Liver RNA.</td>
</tr>
<tr>
<td>RESULTS AND DISCUSSION.</td>
</tr>
<tr>
<td>A. Interaction of Complementary Oligonucleotides in Solution.</td>
</tr>
<tr>
<td>1) Estimation of the average extinction coefficient per base at (25^\circ) for oligonucleotides of mixed base sequence.</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

2) Thermal denaturation curves for
complementary oligonucleotide inter-
actions and poly A.d(pT)$_n$ interactions. . . . . 107
   a) Typical thermal denaturation curve . . . . 107
   b) Width of the thermal transition . . . . 109
   c) % Hypochromicity . . . . . . . . . 110
   d) Summary of data on Tm curves for
      various interactions of complementary
      oligonucleotides and polynucleotides . . . 110
   e) Thermal denaturation curves for
      longer oligonucleotides. . . . . . . . . 110
   f) Thermal denaturation curves for
      oligonucleotides of mixed,
      repeating base sequence. . . . . . . . . 115
   g) Summary of data from the thermal
      denaturation studies of complementary
      oligonucleotides . . . . . . . . . . . . . 118

B. Synthesis and Properties of Oligo-
nucleotide-Celluloses . . . . . . . . . . . . . 119

1) Synthesis of oligonucleotide-celluloses . . . . 119
   a) Incorporation of oligonucleotide
      by the water-soluble carbodiimide
      method . . . . . . . . . . . 119
Table of Contents (Continued)

b) Examination of the CMC-OTs
   reaction .................................................. 120

c) Summary of the synthesis of oligonucleotide-celluloses ........ 129

2) Oligonucleotide-cellulose columns .......................... 130

   a) Thermal elution of complementary
      oligonucleotides ........................................ 130
         (i) Stepwise thermal elution .......................... 130
         (ii) Linear temperature gradient
               elution ............................................. 131
         (iii) Typical elution profile ........................ 133

   b) Alternate methods for the elution
      of oligonucleotides from oligonucleotide-cellulose columns
         ...................................................... 137
         (i) Reverse salt gradient ............................. 138
         (ii) Formamide ......................................... 138
         (iii) N,N'-dimethylformamide ........................ 140
         (iv) Dimethylsulfoxide ................................. 140
         (v) Sodium perchlorate ................................ 140

c) Capacity of oligonucleotide-cellulose columns ............... 141
Table of Contents (Continued)

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>d) Resolution of oligodeoxyadenylylates on</td>
<td></td>
</tr>
<tr>
<td>cell-d(pT)$_8$, cell-d(pT)$<em>9$, and cell-d(pT)$</em>{12}$</td>
<td>147</td>
</tr>
<tr>
<td>e) Elution of d(pA)$_n$, d(A)$_n$ and r(A)$_n$</td>
<td></td>
</tr>
<tr>
<td>on cellulose-d(pT)$_9$</td>
<td>156</td>
</tr>
<tr>
<td>f) Characteristics of different preparations of oligonucleotide-</td>
<td></td>
</tr>
<tr>
<td>cellulose</td>
<td>159</td>
</tr>
<tr>
<td>g) Oligonucleotide-cellulose of</td>
<td></td>
</tr>
<tr>
<td>mixed repeating base sequences</td>
<td></td>
</tr>
<tr>
<td>and the elution of complementary oligomers</td>
<td>164</td>
</tr>
<tr>
<td>(i) Celluloses-d(pTpTpC)$_2$ and</td>
<td></td>
</tr>
<tr>
<td>-d(pCpTpT)$_2$</td>
<td>164</td>
</tr>
<tr>
<td>(ii) Celluloses-d(pTpTpC)$_3$ and</td>
<td></td>
</tr>
<tr>
<td>-d(pCpTpT)$_3$</td>
<td>169</td>
</tr>
<tr>
<td>a) Elution of the hexanucleotide,</td>
<td></td>
</tr>
<tr>
<td>d(pApApG)$_2$</td>
<td>169</td>
</tr>
<tr>
<td>b) Elution of the nonanucleotide,</td>
<td></td>
</tr>
<tr>
<td>d(pApApG)$_3$</td>
<td>171</td>
</tr>
<tr>
<td>h) Effect of GC base pairs on the stability</td>
<td></td>
</tr>
<tr>
<td>of oligonucleotide-cellulose: complementary oligomer interactions</td>
<td>174</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>i) Elution of non-complementary oligonucleotides on oligonucleotide-</td>
<td>175</td>
</tr>
<tr>
<td>celluloses</td>
<td></td>
</tr>
<tr>
<td>j) Purification of chemically synthesized</td>
<td>176</td>
</tr>
<tr>
<td>oligonucleotides on oligonucleotide-cellulose columns</td>
<td></td>
</tr>
<tr>
<td>k) Elution of RNA from an oligonucleotide-cellulose column, and the</td>
<td>177</td>
</tr>
<tr>
<td>selective retention of a complementary oligonucleotide</td>
<td></td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>183</td>
</tr>
<tr>
<td>LITERATURE CITED</td>
<td>188</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Minimum length non-repeating base sequences in genomes of $3 \times 10^3$ to $3 \times 10^9$ nucleotide base pairs</td>
<td>13</td>
</tr>
<tr>
<td>Table 2</td>
<td>Characterization of dephosphorylated oligonucleotides by venom phosphodiesterase digestion and estimation of molar ratios</td>
<td>36</td>
</tr>
<tr>
<td>Table 3</td>
<td>Yield of oligonucleotides $d(pT)_n$ from the polymerization of $d$-$pT$.</td>
<td>37</td>
</tr>
<tr>
<td>Table 4</td>
<td>Summary of $R_f$ values of non-protected and protected nucleotides</td>
<td>40</td>
</tr>
<tr>
<td>Table 5</td>
<td>Yield of oligonucleotide obtained by polymerization of suitably protected trinucleotides.</td>
<td>46</td>
</tr>
<tr>
<td>Table 6</td>
<td>Average molar extinction coefficient per base at 35° for oligonucleotides of mixed base sequence</td>
<td>106</td>
</tr>
<tr>
<td>Table 7</td>
<td>Tm values (°C) and percent hypochromicity for $d(pT)_n$:d(pA)$_n$ and $d(pT)_n$:poly A interactions</td>
<td>111</td>
</tr>
<tr>
<td>Table 8</td>
<td>Thermal stability of complementary oligonucleotides of mixed base sequence</td>
<td>116</td>
</tr>
<tr>
<td>Table 9</td>
<td>Synthesis of oligonucleotide-cellulose by the water-soluble carbodiimide method</td>
<td>121</td>
</tr>
<tr>
<td>Table 10</td>
<td>Characterization of byproducts produced on reaction of $d$-$pT$ with CMC-OTs at pH 6.0.</td>
<td>126</td>
</tr>
<tr>
<td>Table 11</td>
<td>Capacity of cell-$d(pT)_8$ column</td>
<td>144</td>
</tr>
<tr>
<td>Table 12</td>
<td>Elution of deoxyadenylates, $d(pA)<em>6$ to $d(pA)</em>{11}$, on cell-$d(pT)_8$, cell-$d(pT)<em>9$ and cell-$d(pT)</em>{12}$.</td>
<td>148</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>XIII</td>
<td>Elution of $d(pA)_n$ on cell-$d(pT)_9$ columns</td>
<td>154</td>
</tr>
<tr>
<td>XIV</td>
<td>Elution of oligoribo- and oligodeoxyribonucleotides on cell-$d(pT)_9$</td>
<td>158</td>
</tr>
<tr>
<td>XV</td>
<td>Elution of $d(pA)_9$ and $d(pA)_7$ on different preparations of cell-$d(pT)_9$</td>
<td>161</td>
</tr>
<tr>
<td>XVI</td>
<td>Elution of $d(pApApG)_n$ ($n = 2,3$) on oligonucleotide-celluloses of mixed, repeating base sequences</td>
<td>165</td>
</tr>
<tr>
<td>XVII</td>
<td>Thermal elution of RNA and $d(pApApG)_3$ on cellulose-$d(pCpTpT)_3$</td>
<td>181</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>N-acyl protecting groups for nucleoside 5' monophosphates.</td>
<td>22</td>
</tr>
<tr>
<td>Figure 2</td>
<td>General method for the synthesis of protected deoxyribotrinitrulucleotides.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Anion exchange chromatography of d-pTpT.</td>
<td>41</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Anion exchange chromatography of d-pTpTpT.</td>
<td>43</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Anion exchange chromatography. Separation of products from the polymerization of d-pTpTpT.</td>
<td>45</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Anion exchange chromatography of products from the synthesis of d-pTpTpC&lt;sub&gt;An&lt;/sub&gt;.</td>
<td>50</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Anion exchange chromatography of products from the polymerization of d-pTpTpC&lt;sub&gt;An&lt;/sub&gt;.</td>
<td>52</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Anion exchange chromatography purification of d-pC&lt;sub&gt;An&lt;/sub&gt;pTpT.</td>
<td>55</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Anion exchange chromatography of products from the polymerization of d-pC&lt;sub&gt;An&lt;/sub&gt;pTpT.</td>
<td>57</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Anion exchange chromatography purification of d-pA&lt;sub&gt;Bz&lt;/sub&gt;pA&lt;sub&gt;Bz&lt;/sub&gt;.</td>
<td>63</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Anion exchange chromatography of d-pA&lt;sub&gt;Bz&lt;/sub&gt;pA&lt;sub&gt;Bz&lt;/sub&gt;pG&lt;sub&gt;Bz&lt;/sub&gt;, at pH 8.5.</td>
<td>66</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Anion exchange chromatography of d-pA&lt;sub&gt;Bz&lt;/sub&gt;pA&lt;sub&gt;Bz&lt;/sub&gt;pG&lt;sub&gt;Bz&lt;/sub&gt;, at pH 5.5.</td>
<td>68</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Anion exchange chromatography of products from the polymerization of d-pA&lt;sub&gt;Bz&lt;/sub&gt;pA&lt;sub&gt;Bz&lt;/sub&gt;pG&lt;sub&gt;Bz&lt;/sub&gt;.</td>
<td>70</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

Figure 14. Definition of terms related to thermal denaturation curves .................................. 92

Figure 15. Dimensions (in centimeters) of jacketed column used for temperature controlled elution of oligonucleotides from columns of oligonucleotide-cellulose .................................. 98

Figure 16. Venom phosphodiesterase digestion of d(pApApG)$_2$ and d(pApApG)$_3$ .......... 105

Figure 17. Thermal denaturation profiles for the interactions d(pT)$_8$.poly A (A) and d(pT)$_8$.d(pA)$_8$ (B) ........................................ 108

Figure 18. Thermal denaturation profiles for the interactions d(pA)$_{11}$.d(pT)$_9$ and d(pA)$_{11}$.d(pT)$_{12}$ ........................................................................ 114

Figure 19. Reaction of d-pT with CMC-OTs ........................................................................ 124

Figure 20. Spectrum of peak b (CMCd-pT) ........................................................................ 127

Figure 21. Reaction of d-pT, d-pC, and d-pA with 100 mg (total) of CMC-OTs at pH 6.0 ........................................ 128

Figure 22. Elution of d(pA)$_8$ on cellulose-d(pT)$_8$ ......................................................... 132

Figure 23. Typical elution profiles for oligonucleotide-celluloses ......................................................... 134

Figure 24. Alternate methods for the elution of oligonucleotides on oligonucleotide-cellulose columns ........................................ 139

Figure 25. The capacity of a cell-d(pT)$_8$ column ...................................................................... 143

Figure 26. % A.U. d(pA)$_8$ not retained by a cell-d(pT)$_8$ column vs A.U. d(pA)$_8$ loaded ........................................................................ 145
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Composite graph elution profiles for elution of d(pA)$_n$ on cell-d(pT)$_9$, A; cell-d(pT)$<em>9$, B; and cell-d(pT)$</em>{12}$, C.</td>
<td>149</td>
</tr>
<tr>
<td>28</td>
<td>Tm$^c$ of d(pA)$_n$ on cellulose-d(pT)$_8$, -d(pT)$<em>9$, and -d(pT)$</em>{12}$ vs oligomer length n.</td>
<td>152</td>
</tr>
<tr>
<td>29</td>
<td>Composite elution profile for d(pA)$_n$ (n = 6 to 9) on cell-d(pT)$_9$ (0.9 cm x 15 cm, 150 A.U. d(pT)$_9$).</td>
<td>155</td>
</tr>
<tr>
<td>30</td>
<td>Tm$^c$ (°C) versus oligomer length for the elution of the series d(pA)$_n$, d(A)$_n$, and r(A)$_n$ on cell-d(pT)$_9$.</td>
<td>160</td>
</tr>
<tr>
<td>31</td>
<td>Elution of d(pApApG)$_2$$_3$ on celluloses-d(pTpTpC)$_2$ and -d(pCpTpT)$_2$.</td>
<td>166</td>
</tr>
<tr>
<td>34</td>
<td>Elution of RNA on cell-d(pCpTpT)$_3$.</td>
<td>179</td>
</tr>
<tr>
<td>35</td>
<td>Elution of d(pApApG)$_3$ and RNA, separately, and as a mixture on cell-d(pCpTpT)$_3$.</td>
<td>180</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

During the course of these experiments I have had the pleasure of working with a number of people whose discussions have been most helpful and to whom I am grateful. To Vivian, Deena, Nadine, Grant, Richard, Ron, Jerry, Yasuko, Bill, Paul, Pat, David, and Albert I express my appreciation. I would also like to thank Dr. Ian Gillam who has offered his advice on a number of problems which arose.

Mr. W. Morrison of the Department of Physics, The University of British Columbia, designed and built the gear box for the gear system necessary to run the linear temperature gradient.

Finally, most of all, the supervision of Dr. Michael Smith is gratefully acknowledged. His sincere interest in students, and his guidance of their research will be remembered.

I was the recipient of a National Research Council of Canada Scholarship (1966 - 1967) and a Killam Predoctoral Fellowship (1967 - 1970).
ABBREVIATIONS USED

The abbreviations used for the nucleotides are those suggested by the IUPAC-IUB Combined Commission on Biochemical Nomenclature (Revised Tentative Rules, 1965) (1) and used extensively in the papers on polynucleotide synthesis published by Dr. H.G. Khorana and his associates (see Schaller and Khorana, ref. 2, footnote 7). These abbreviations are summarized below:

\[
\begin{align*}
\text{dT, dC, dA, dG} & \quad \text{the deoxyribonucleosides of the four bases; thymine, cytosine, adenine, and guanine.} \\
\text{d-pT, d-pC, d-pA, d-pG} & \quad \text{the 5' deoxyribonucleoside monophosphates.} \\
\text{d-pA\text{\textsubscript{Bz}}} & \quad \text{N-benzoyl d-pA} \\
\text{d-pC\text{\textsubscript{An}}} & \quad \text{N-anisoyl d-pC} \\
\text{d-pG\text{\textsubscript{Bz}}} & \quad \text{N-benzoyl d-pG} \\
\text{CEd-pT} & \quad \text{the 8-cyanoethyl derivative of d-pT.} \\
\text{d-pT-OAc} & \quad \text{the 3'-0-acetyl derivative of d-pT.} \\
\text{d-pA\text{\textsubscript{Bz}}pA\text{\textsubscript{Bz}}pA\text{\textsubscript{Bz}}} & \quad \text{the protected deoxytrinucleotide, 5'-0-phosphoryl-N-benzoyldeoxyadenyllyl-(3'\rightarrow 5')-N-benzoyldeoxyadenyllyl-(3'\rightarrow 5')-N-benzoyldeoxyguanosine.} \\
\text{d(pN)\_n} & \quad \text{oligodeoxyribonucleotide of N,n nucleotides long, with a 5' phosphate group and 3' hydroxyl.}
\end{align*}
\]
**Abbreviations Used (Continued)**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>d(N)$_n$</td>
<td>d(pN)$_n$ which has been dephosphorylated.</td>
</tr>
<tr>
<td>r(N)$_n$</td>
<td>the oligoribonucleotide of N, n nucleotides long, with 5' and 3' hydroxyl termini.</td>
</tr>
<tr>
<td>d(pTpC)$_n$</td>
<td>the oligodeoxyribonucleotide of repeating sequence, d-pTpC, 3n nucleotides long.</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid.</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid.</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer ribonucleic acid.</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid.</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal ribonucleic acid.</td>
</tr>
<tr>
<td>Poly G</td>
<td>polyriboguanylic acid.</td>
</tr>
<tr>
<td>Poly (U, G)</td>
<td>polyribonucleotide with random sequence of uridylate and guanylate residues. Poly (A, G) and poly (I, G) are analogous structures.</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine 5' triphosphate.</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytosine 5' triphosphate.</td>
</tr>
<tr>
<td>dITP</td>
<td>deoxyinosine 5' triphosphate.</td>
</tr>
<tr>
<td>UDP</td>
<td>uridine 5' diphosphate.</td>
</tr>
</tbody>
</table>

Other abbreviations and definitions used are:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>polynucleotide-cellulose</td>
<td>cellulose containing covalently linked homo-oligonucleotides of random length.</td>
</tr>
</tbody>
</table>
Abbreviations Used (Continued)

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>oligonucleotide-cellulose</td>
<td>cellulose containing covalently linked oligonucleotides of defined length and sequence.</td>
</tr>
<tr>
<td>cell-(d(pT)_n)</td>
<td>cellulose to which has been covalently attached the oligodeoxyribonucleotide, (d(pT)_n), via a phosphodiester link between the 5' terminal nucleotide and an hydroxyl group on the cellulose.</td>
</tr>
<tr>
<td>(R_f)</td>
<td>mobility relative to the solvent front.</td>
</tr>
<tr>
<td>(R_{d-pT})</td>
<td>mobility relative to (d-pT).</td>
</tr>
<tr>
<td>(R_{d-pTpT})</td>
<td>mobility relative to (d-pTpT).</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer ((10^{-9} \text{ meter, 1 millimicron})).</td>
</tr>
<tr>
<td>A 260 nm</td>
<td>absorbance at 260 nm.</td>
</tr>
<tr>
<td>A.U.</td>
<td>one absorbance unit, is that amount of substance, which when dissolved in 1 ml of solvent has an absorbance of 1.0 in a cell with a 1 cm light path. (The absorbance value is read at the (\lambda_{\text{max}}).)</td>
</tr>
<tr>
<td>(\lambda_{\text{max}} \ (\lambda_{\text{min}}))</td>
<td>wavelength at absorbance maximum (minimum).</td>
</tr>
<tr>
<td>(E)</td>
<td>the molar extinction coefficient, equal to the absorbance of a 1 molar solution in a 1 cm light path.</td>
</tr>
</tbody>
</table>
Abbreviations Used (Continued)

U  unit of enzyme activity.
DCC  N,N'-dicyclohexylcarbodiimide.
MsSO₂Cl  2-mesitylenesulfonyl chloride.
TEA  n-triethylamine.
TEAB  triethylammonium bicarbonate.
TBA  n-tributylamine.
THA  n-trihexylamine.
Na⁺Mes⁻  sodium salt of 2-(N-morpholino)-ethane-sulfonic acid.
CMC-OTs  N-cyclohexyl-N′β(4-methylmorpholinum)-ethylcarbodiimide p-toluenesulfonate.
MBS  molar buffered saline; 0.01 M NaH₂PO₄, pH 7.0, 1 M in NaCl.
DEAE-cellulose  0-(diethylaminoethyl)cellulose.
SDS  sodium dodecyl sulfate.
Tm  temperature at one half the increase in absorbance of a thermal dissociation profile.
Tm°C  temperature of elution of an oligonucleotide from an oligonucleotide-cellulose column (defined as the temperature at the peak tube).
Abbreviations Used (Continued)

$\Delta T_m$ ($\Delta T_{mC}$)  
difference between two $T_m$ ($T_{mC}$) values.

Linear temperature gradient  
a linear increase in temperature of the eluting buffer in a column.
GENERAL INTRODUCTION

The discovery of nucleic acids has been attributed to Friedrich Meischer, in 1871 (3), who isolated "nuclein" both from pus cells and from sperm heads of the Atlantic salmon Salmo salar (4, 5). Although the chemical composition and wide distribution of nucleic acids became known largely in the next 30 to 40 years (4), the biological function of deoxyribonucleic acid was not appreciated until the work of Avery, McLeod and McCarty, in 1944 (6), who demonstrated that DNA was the factor responsible for bacterial transformation. A few years later, in 1952, Hershey and Chase (7) demonstrated that during T2 infection of Escherichia coli, the nucleic acid of the bacteriophage (and not the protein coat) entered the host bacterium, and therefore the nucleic acid must carry the genetic information. These two observations plus corroborative work of others led to the conclusion that DNA is the storehouse of genetic information. More recent work has demonstrated that RNA may also encode information from one generation to another (in RNA viruses). However, for the most part RNA is identified with structural (rRNA), informational (mRNA) and adaptor (tRNA) functions within the cell. All three classes of RNA are intimately involved

---

1 The use of "structural" does not imply an inactive role.
in the process of protein synthesis (8).

The significance of nucleic acids has led to an almost phenomenal assault on the problems of detailed structure and function. For example, methods have been developed for the sequencing of ribonucleic acids (9-11) and a number of complete sequences for transfer RNAs have been reported (see 10 for a recent summary). Partial sequences for several viral RNAs are also known (10) as well as for the two ribosomal RNAs from *E. coli* (12, 13). Lack of extensive chemical or specific enzymatic methods have severely restricted sequence studies in all but simple DNA molecules to sequences as long as 10 to 15 bases (14, 15). However, recent reports (16-18) have indicated that the determination of the sequence of a DNA molecule may be possible within the next few years. With this interest in structure and function of nucleic acids it is not surprising, then, that a great deal of effort has been directed toward isolation procedures for nucleic acids.

The fractionation of nucleic acids had its beginning both in cytochemical studies (19) and chemical composition studies of cell fractions (20-23). Thus, in eucaryotic cells, the nucleus was shown to be the primary source of DNA, while the cytoplasm was relatively rich in RNA. Many nucleic acid extraction procedures take advantage of this differential location of the molecules, and consequently numerous methods have been developed for the isolation of nuclei as well as other cell organelles (24).
Ribosomal RNA (including 5s RNA) is readily extracted from ribosomes or polysome fractions. Transfer RNA may be obtained from the high speed supernatant fraction during centrifugation of fragmented cell preparations, or is often obtained from the aqueous layer of phenol extracted cytoplasm, and is separated from other RNAs by various chromatographic procedures (25). More refined fractionation methods (for example, countercurrent separation, DEAE-Sephadex chromatography, ion-exchange chromatography, reversed-phase partition chromatography, and benzoylated DEAE-cellulose chromatography, 10) have been used to separate different tRNA molecules for sequence analyses. However, while these classes of nucleic acids (rRNAs and tRNAs) are readily fractionated, DNA and messenger RNAs are still prepared largely as heterogeneous fractions.

The isolation of a homogeneous population of DNA molecules (with the exception of special cases such as viral, bacterial or mitochondrial DNAs, etc.) has not been accomplished. The difficulty is obvious from the known content and complexity of cellular DNA in eucaryotic organisms. For example, in the calf genome, there are $3.2 \times 10^9$ nucleotide base pairs per haploid set (26). Although 45% of the calf genome belongs to the class of repetitious DNA molecules (27), there are still some $1.8 \times 10^9$ base pairs comprising unique DNA sequences. With the knowledge that at least several genes are located on
different chromosomes (26), it must be concluded that in higher organisms the genome consists of a large population of different DNA molecules. If one were attempting to isolate a particular region of the genome, for example, a cistron, occurring in the non-repetitious portion of the genome, one would be attempting to isolate approximately $10^{-4}\%$ of the genome!\(^2\)

Possibly a more feasible approach to the isolation of specific sequences of informational nucleic acid would be the isolation of a specific messenger RNA. This is theoretically a simpler problem in that in a differentiated cell expressing a particular gene, many copies of that messenger could be expected to be present, and only a fraction, possibly as few as $10^3$ genes occurring in the genome may be transcribed (28).\(^3\) Also, in the isolation of a mRNA there is the possibility of an in vitro assay, in which the RNA may be identified by its ability to promote the synthesis of a specific protein (31, 32), although coupled DNA-RNA protein synthesizing complexes have been demonstrated (33-35).

Techniques have been developed recently in which RNA (tRNAs from E. coli) has been hybridized with its homologous region of the genome (36, 37). When sheared DNA (molecular weight

---

\(^2\) I have assumed that a cistron is the length of DNA coding for a single protein, about 500 amino acids long and therefore equal to 1500 base pairs. If the number of pairs per haploid set is $1.8 \times 10^9$ (the non-repetitious portion of the genome) (27), then a cistron which occurs once represents about $10^{-4}\%$ of the genome.

\(^3\) In fact, there is no good evidence yet to suggest how much of the DNA in a eucaryote codes for proteins (29, 30).
1.25 \times 10^5$ daltons) was used, a region of the DNA coding for 4 to 5 tRNA molecules was obtained, suggesting a contiguous arrangement of tRNA cistrons in the *E. coli* genome (36). In other work, an endonuclease specific for single stranded DNA was used to digest away non-hybridized regions of the DNA. The size of the DNA fragment obtained was equivalent to the length of a single tRNA molecule, indicating that the DNA-tRNA hybrids formed must have at least a short gap, between two adjacent cistrons, which the endonuclease is able to recognize (37). It is conceivably possible that the deoxypolymer isolated in this way could then be copied using DNA polymerase. A preliminary report using a similar hybridization technique has shown that this method is also feasible for use with a messenger RNA (38). However, the rapid rate and extent of hybridization of mouse globin mRNA with its homologous DNA suggests that the globin messenger hybridized with a portion of the repetitious DNA sequences in the genome. This is further substantiated by the relatively low Tm of the globin mRNA:DNA hybrids as well as the extreme sensitivity of the hybrid to ribonuclease digestion.

It may also be possible to obtain the deoxyribo-strand complementary to an mRNA molecule using one of the recently discovered RNA dependent DNA polymerases associated with RNA tumor viruses (39-41).
From the consideration discussed above (less heterogeneity, potential in vitro assay of the product, and the availability of a method to obtain the complementary DNA sequence by isolation or enzymatic synthesis), it would seem logical to concentrate on a less formidable problem than the one of isolation of a gene, the isolation of a specific mRNA.

The first observation of an informational type RNA was made in 1956 by Volkin and Astrachan (42) who observed a rapid incorporation of $^{32}$P into RNA after infection of E. coli with bacteriophage T2. The name of messenger RNA was put forth by Jacob and Monod (43) in an hypothesis to explain the rapid synthesis of RNA after phage infection, as well as the kinetics of enzyme induction and repression. In the next 10 years after Volkin and Astrachan's observation, a number of properties of mRNA were defined. For example, it was believed that messenger RNA had a base composition and sequence similar to total DNA, it was heterogeneous in size, it stimulated amino acid incorporation, it was rapidly labelled in pulse experiments and had a short half-life, and it was present in polysomes. Unfortunately, these properties are not necessarily unique to messenger RNA. Singer and Leder (44) suggested that experiments with cell-free systems provide the most convincing evidence in support of the concept of mRNA. In the experiments

---

4 For a discussion on the evaluation of mRNA, see reference 44.
that these workers have referred to, an RNA (often a viral RNA) was used to program the \textit{in vitro} translation of a protein. A partial proteolytic digest of the protein was then compared with the \textit{in vivo} product, normally by two dimensional thin layer chromatography (45-49). Since then at least three reports of the \textit{in vitro} translation of mRNA into active enzyme molecules (phage T4 lysozyme, 31; phage T4 glucosyl transferase, 32; and \textit{E. coli} alkaline phosphatase, 50) have appeared in the literature, indicating a remarkable fidelity of the \textit{in vitro} translation process.

Isolation procedures for messenger RNAs are poorly developed compared with those for other RNAs, possibly because of the size heterogeneity, the small amount normally present in a cell,\footnote{The messenger RNA accounts for at most 5-10\% of the RNA in a cell, while ribosomal (80\%) and tRNA (10 to 15\%) make up the remainder (51).} as well as the extreme sensitivity of these molecules to ribonuclease (52). However, two general procedures for the isolation of messenger RNAs have been used.

The first of these methods involved the hybridization of RNA to homologous DNA. Bautz and Hall (53) reported the use of a bacteriophage T4 DNA-phosphocellulose matrix to isolate T4 specific mRNA, presumably by hybridization of the homologous RNA and DNA. Their procedure also permitted them to effect a purification of a specific portion of the T4 messenger by
elution of this material on a T4 deletion DNA-phosphocellulose matrix. The method is a potentially promising one, and is theoretically applicable to the isolation of any messenger RNA.

DNA cellulose has not been used extensively for this purpose, although alternate methods with DNA immobilized in agar or on nitrocellulose (54, 55) have been used. Recently a øX 174 DNA cellulose column has been used to prepare minus strands from the RF form of the virus (56). DNA celluloscs have also been used to isolate enzymes with DNA associated functions (57, 58).

Another method that has been used to purify messenger RNA is based on the concept that mRNA in the process of translation is associated with polysome fractions. Accordingly, a number of attempts have been made to isolate and characterize the mRNA for haemoglobin.6 The work related to the isolation of haemoglobin messenger up to 1967 has been summarized by Chantrenne et al. (52). Several more recent reports have appeared.

Laycock and Hunt (60) have reported a salt fractionation procedure for reticulocyte RNA. One fraction they obtained contained an 8S and 14S RNA species, and these workers were

6Many workers have chosen globin messenger as a model system. Their choice is probably related to the observation that the soluble ribosomes of the reticulocyte are concerned almost exclusively with the synthesis of the α and β chains of the globin molecule (59). Also, reticulocytes are readily lysed by controlled osmotic shock, making it possible to obtain intact polysomes (51).
able to demonstrate that the 8S species stimulated incorporation of radioactive amino acids into a globin type molecule, using an E. coli cell-free protein synthesizing system and the "initiator" N-acetyl-valyl-tRNA. However, recent work of Housman et al. (61) has demonstrated that both globin chains are initiated as NH₂-met-val in a reticulocyte cell-free system supplemented with yeast (eucaryotic) methionine tRNAs, and the methionine is subsequently removed from the amino terminus of the proteins. Thus, the globin synthesis observed by Laycock and Hunt (60) would appear to have been initiated abnormally. However, this does not mean that the 8S RNA is not the globin messenger.

Labrie (62) has reported a method for isolation of rabbit globin mRNA obtained from polysomes and purified by electrophoresis. The RNA (10S) was identified as putative globin messenger by its size, kinetics of labelling, sensitivity to ribonuclease and its apparent lack of homology with either ribosomal RNA species, determined from fingerprint patterns of ribonuclease T₁ digestion of the RNAs (62).

A number of other tRNAs have been prepared by the polysome procedure. For example, mRNA for myosin has been isolated from embryonic chicken skeletal muscle polysomes and purified by sucrose gradient sedimentation (63). The isolation of histone messengers has also been reported (64-66). An interesting and
potentially highly discriminating modification of the polysome method was that reported by Williams and Askonas (67). These workers isolated the polysomes which were in the process of making the heavy and light chains of immunoglobins, by precipitation with antisera prepared against the two chains. They observed a highly specific precipitation of the 300S polysomes with antiserum against the heavy chains, while a less discrete region centered on 120-180S polysomes was precipitated with antiserum against light chains.

It would seem that the application of these methods (salt fractionation, polysome extraction, sedimentation analysis and electrophoresis) to messenger RNA isolation, although reasonably successful in several cases, owe much of their success to the use of cells which are highly specialized and making very few proteins.

In looking for a method for the isolation of specific polynucleotides, possibly either modifications of hybridization procedures used by Bautz and Hall (53) or isolation of specific polysome classes by an immunological precipitation step (67) would be more generally useful.

Gilham has reported the synthesis of phosphodiester-linked polynucleotide-cellulose columns to isolate complementary oligonucleotides (68, 69). Thus, deoxythymidine polynucleotide-cellulose has been demonstrated to fractionate deoxyadenylates
(tri- through heptanucleotides, 69), and this procedure has been used to fractionate an enzymic digest of viral RNA (70). Edmonds and Abrams (71) have reported the use of deoxythymidine polynucleotide-cellulose to isolate a naturally occurring polyadenylate from calf thymus nuclei, while the same procedure has been used to isolate an adenine-rich polymer from Ehrlich ascites cells (72). The extension of the principle of polynucleotide-cellulose columns, containing homo-oligodeoxynucleotides of random length, to cellulosics containing phosphodiester-linked oligonucleotides of defined length and sequence might offer a feasible technique by which polynucleotides such as mRNA molecules may be isolated. The objective of the work described in this thesis is to examine the possibility of this extension. The procedure proposed would involve the isolation of a polynucleotide containing an oligonucleotide sequence complementary to a chemically-synthesized oligomer which is covalently attached to an insoluble matrix (see below).

\[
\begin{array}{c}
\text{cellulose} \quad TP\bar{TP}C \quad pC\bar{TP}A \\
\text{pAP} \quad \text{3' to 5'}
\end{array}
\]

In such a procedure, it is necessary to decide what is the length of the base sequence required in order for the sequence to be unique for a particular polynucleotide. Obviously, it must be
long enough to provide a stable (hydrogen bonded) interaction, and at the same time, be long enough such that on a random probability basis, that sequence occurs only once in the genome of the organism. The probability of observing a particular sequence of \( n \) nucleotides long is \( \frac{1}{4^n} \). Assuming a random distribution of bases, for a genome which contains \( X \) nucleotide base pairs, a unique sequence of bases is present in an oligomer \( n \) nucleotides long, where \( 4^n = X \).

Various groups (26, 27, 74) have summarized data on the minimum number of nucleotide base pairs per haploid set in different organisms. The data are combined and listed in Table I. The organisms are listed in order of increasing number of base pairs per haploid set (given in the brackets, column A), and are aligned in position with the corresponding value of the expansion \( 4^n \) (column B). Column C is the corresponding

---

7Thomas (73) has presented a more complicated calculation of the minimum, non-repeating length in a genome. He concludes that the length is \( \geq f \) where the number of nucleotides in the genome is approximately \( 1/2.4^f \). The two calculations actually differ by a factor of 4. Thomas's calculation is based on the total number of nucleotides, while the one described above is for the total number of nucleotide base pairs (i.e. 1/2 the total number of nucleotides). Thomas also excluded inverted repeating sequences. If one uses Thomas's relationship to determine the minimum unique chain length, the length obtained is one base longer than that obtained by the relationship derived above.
Table I. Minimum length non-repeating base sequences in genomes of $3 \times 10^3$ to $3 \times 10^9$ nucleotide base pairs.

<table>
<thead>
<tr>
<th>Organism (nucleotide base pairs)</th>
<th>$4^n$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>R17 (3300)$^a$; MS-2 (4000)$^a$; ØX 174 (5,100)$^a$; SV40 (7,000)</td>
<td>4,096</td>
<td>6</td>
</tr>
<tr>
<td>T7 (40,000); λ (50,000)</td>
<td>16,384</td>
<td>7</td>
</tr>
<tr>
<td>T2, T4 (2 $\times$ 10$^5$)</td>
<td>65,536</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>262,144</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1,048,576</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>4,194,304</td>
<td>11</td>
</tr>
<tr>
<td>E. coli (4.5 $\times$ 10$^6$)</td>
<td>16,777,216</td>
<td>12</td>
</tr>
<tr>
<td>Yeast (2 $\times$ 10$^7$)</td>
<td>67,108,864</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>268,435,456</td>
<td>14</td>
</tr>
<tr>
<td>Teleosts (4 $\times$ 10$^8$)</td>
<td>1,073,741,824</td>
<td>15</td>
</tr>
<tr>
<td>Amphibians (1 $\times$ 10$^9$)</td>
<td>4,294,967,296</td>
<td>16</td>
</tr>
<tr>
<td>Reptiles (1.5 $\times$ 10$^9$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammals (3.2 $\times$ 10$^9$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Number of base pairs in the double-stranded replicative form.
value of \( n \), or the minimum, unique sequence.

From the table it can be seen that a number of viral and bacterial polynucleotides should contain unique oligonucleotide sequences of remarkably short length (6 to 12 nucleotides), within the range that is convenient to prepare synthetically by existing chemical methods (75). Mushynski and Spencer (14, 15) have looked at the frequency of occurrence of pyrimidine tracts in two bacterial DNA viruses, T7 and \( \lambda \). They found that in T7, there are two tracts of oligodeoxythymidylates, \( d(pT)_6 \), per molecule of T7 DNA (both on the left strand). In \( \lambda \), there are five tracts of \( d(pT)_7 \) per molecule (two on the right strand and three on the left strand). Min Jou and Fiers (76) have reported that there are at least eight unique octanucleotides in the MS-2 genome, while Thirion and Kaesberg (77) have also reported similar results for the R17 and M12 genomes.

The unique length in a eucaryotic genome is longer (14 to 16 nucleotides) than one could conveniently synthesize by existing methods. However, the isolation of mammalian polynucleotides is still theoretically possible using two oligonucleotide-cellulose columns, each eight nucleotides long.\(^8\)

\(^8\) The probability of the occurrence of two sequences each \( \frac{n}{2} \) nucleotides long is equal to the probability of occurrence of one sequence \( n \) nucleotides long (i.e., \( \frac{1}{4^n} \times \frac{1}{4^n} = \frac{1}{4^n} \)).
The isolation of random sequence polynucleotides by hybridization to a relatively short complementary oligonucleotide would therefore seem to be a feasible approach to the isolation of specific polynucleotides, providing interactions between such short tracts of oligonucleotides are sufficiently stable.

The final problem is that of how one determines a short sequence of nucleotides within a polynucleotide such as a messenger RNA. There are several approaches available at the present time.

Streisinger's group has demonstrated the occurrence of pseudo-wild lysozyme phenotypes in acridine-induced double frameshift mutants of bacteriophage T4 (78-82). By comparing the codons (82, 83) for the amino acids within the double frameshift region of the mutant and the corresponding region of the wild type proteins, one is able to obtain an almost unambiguous sequence up to 10 nucleotides long.

Missense mutations have also been used to derive the sequence of nucleotides within a short tract. For example, Sherman et al. (84) have studied the relationship of gene structure and protein structure of iso-1-cytochrome C from yeast. From a number of amino acid replacement mutants a short sequence of the messenger at the 5' end of the molecule can be derived. Also, the codons for methionine and tryptophan residues may be deduced unambiguously (85). Consequently, if any two of these occur consecutively in a protein, a sequence
in the corresponding mRNA 8 nucleotides long, can be determined.

It may be possible to derive (not entirely unambiguously) the codons for a short sequence of amino acids within a protein by deduction from ribosomal binding studies (82) of fractionated tRNA molecules. However, this method can only answer questions concerning degeneracy in the first letter of codewords.

Finally, recent reports from several laboratories have demonstrated that it is possible to determine short sequences within viral RNA molecules, for example ribosomal binding sites, by virtue of the protection by the ribosome of a short oligonucleotide from nuclease digestion (86, 87). It is conceivably possible that this method could be adapted to determination of a short sequence within any messenger RNA. For example, the translation of a messenger could be stopped at a specific point by omission of a rare amino acid from the synthesizing preparation. The specific polysome class could be isolated by sedimentation or precipitation (67). A sequence of oligonucleotides may then be obtained by mild nuclease digestion of the polysomes followed by isolation of the "protected" sequence. The structure of the sequence could then be determined by existing methods (10).

In summary, the methods presently available for the isolation of nucleic acids are inadequate for the preparation
of homogeneous populations of both deoxyribonucleic acids and
messenger ribonucleic acids. Possibly the main difficulty in
the isolation of these molecules rests in their heterogeneity
(both size and sequence). Because of the high degree of
specificity of base interactions of complementary nucleic
acid sequences, it is likely that a method which took advantage
of this property, would be of general use in the isolation of
specific polynucleotides. In this Introduction, it is proposed
that polynucleotides such as messenger ribonucleic acids may
be isolated by hybridization to a chemically synthesized oligo-
deoxyribonucleotide which is complementary to a short tract
within the polynucleotide. The experiments reported here,
carried out in order to determine the practicability of such a
method, involved a study with chemically synthesized model
compounds.

This thesis is divided into two parts. Part I describes
the chemical synthesis of oligodeoxyribonucleotides of defined
length and sequence. The methods used were similar to those
developed by Khorana and his associates (75). The preparation
of some of the sequences has not been described previously.
Accordingly, these oligonucleotides have been thoroughly
characterized. For the studies described in Part II, micromole
amounts of oligonucleotides were required, considerably more
than the nanomole amounts of the intermediate oligomers which
were required in the chemical synthesis of the yeast alanine tRNA gene (88).

Part II is divided into two sections. Section A is a study of the thermal stability of interactions between complementary oligonucleotides in solution. In Section B, the method used to covalently link oligonucleotides to cellulose and the preparation of oligonucleotide-celluloses in a form suitable for column chromatography, is described. Optimal conditions for the thermal elution of complementary oligonucleotides have been determined.

The oligonucleotide-celluloses were characterized with regard to their ability to bind complementary deoxyribo- and ribo-oligonucleotides. The capacity, resolving power, and reproducibility of these columns have been studied. The ability of an oligonucleotide-cellulose column to selectively retain a complementary (synthetic) oligomer in a mixture with cellular ribonucleic acids has been examined.

The results of these experiments provide a very encouraging basis for an attempted extension of the use of oligonucleotide-celluloses to the isolation of specific, naturally occurring polynucleotides.
PART I

CHEMICAL SYNTHESIS OF OLIGODEOXYRIBONUCLEOTIDES
OF DEFINED LENGTH AND SEQUENCE
INTRODUCTION

Methods for the synthesis of oligodeoxyribonucleotides of defined sequence have been developed in the laboratory of Dr. H.G. Khorana, over the past 15 years (for general review, see 75). These methods, involving reaction in anhydrous solvent, use either a carbodiimide (e.g. dicyclohexylcarbodiimide, DCC) or a sulfonyl chloride (e.g. 2-mesitylenesulfonyl chloride or 2,4,6 triisopropylbenzenesulfonyl chloride) as condensing agent. These procedures are lengthy, requiring much time both to make the suitably protected reactants, as well as to separate the desired products in pure form by anion exchange chromatography. However, they have been used to great advantage, resulting in the synthesis of all sixty-four ribonucleotide codons (89) and oligodeoxyribonucleotides with repeating di- (90), tri- (91) and tetranucleotide (92-94) sequences. The oligodeoxyribonucleotides were used as template primers for *Escherichia coli* DNA polymerase (95). Oligodeoxyribonucleotides and derived polydeoxyribo-nucleotides have been transcribed with *E. coli* RNA polymerase into high molecular weight ribopolymers used in cell-free protein

---

*The preferred method to obtain high molecular weight ribopolymers is to use the short oligodeoxyribonucleotides with *E. coli* DNA polymerase to prepare DNA-like polymers. These are then transcribed into RNA-like polymers (95, 75).*
synthesizing systems to determine and confirm codon assignments (see review, 95). Recently, the goal of these synthetic methods, the complete synthesis of the gene for yeast alanine tRNA, has been achieved (88). The method used involved the chemical synthesis of oligodeoxynucleotides (up to an icosanucleotide in length) of overlapping complementary segments of the derived structure for this tRNA gene. The segments were joined using the enzyme polynucleotide ligase (96, 97).

The chemical synthesis of oligodeoxynucleotides requires two general types of protection in intermediates:

1) protection of the heterocyclic amino groups of cytosine, adenine and guanine, and

2) protection of either the 5' phosphoryl group or the 3' hydroxyl group of the nucleotides.

The heterocyclic amino groups require protecting groups in order to prevent side reactions. (One advantage to using N-acyl protecting groups is that the protected nucleotides are more soluble in anhydrous pyridine, the referred reaction solvent.)

In early experiments, reaction between d-pT-OAc and 5'-O-trityl d-pC in the presence of DCC resulted in phosphorylation of the 6-amino group of d-pC by the activated components, forming the phosphoramide (98). A similar problem where the amino group of deoxyadenosine was phosphorylated was encountered in the attempted synthesis of d-TpApC by reacting d-TpA and d-pC^An-OAc
in the presence of DCC (99).

The chemical characteristics for groups to protect heterocyclic amines are that they must be specific, they must be stable to the conditions of the phosphodiester synthesis and the work up, and they must produce no adverse side effects (e.g., see discussion of d-pG<sub>Bz</sub> below). Also, the protecting groups must be readily removed under conditions which do not degrade phosphodiester or glycosidic linkages.

It has been determined that d-pA may be adequately protected with the benzoyl group (100), while d-pG may be protected with the p-methoxybenzoyl (anisoyl) group (101) (see Figure 1). Deoxyguanosine 5' phosphate was originally protected with the acetyl group (102), however this was not quite stable enough (103) (loss of about 10% during normal work up of a condensation reaction); thus the benzoyl derivative was preferred. Subsequently, the N-benzoyl group was found to interfere with the anion exchange chromatography purification step by lowering the pK of the hydroxyl at position 6, resulting in partial ionization at pH 8.5 (104). The consequence is poor resolution of the reactants and products. The preferred protecting group for d-pG now is the isobutyryl group (75) (see Figure 1). In the chemical synthesis work described in this thesis, d-pG residues were protected with the benzoyl group.
Figure 1. N-acyl protecting groups for nucleoside 5’monophosphates.
All three N-acyl protecting groups are removed by concentrated NH₄OH (105). They are all stable to strong alkali (pH 12) for short periods at low temperature, by virtue of the fact the acyl groups enolize above pH 12, making them resistant to attack by hydroxyl ion (100).

The second type of protecting groups involve protection of the 5'phosphoryl group and the 3'hydroxyl group. It has been found that the cyanoethyl group is suitable for protecting the 5'phosphoryl group, while the acetyl group is satisfactory for the 3'hydroxyl (75). Both of these groups may be removed in stepwise synthesis, with strong alkali (pH 12, 20 min, 0°C), conditions in which the N-acyl groups are stable (see above). However, one disadvantage is that both groups are alkali labile; thus in stepwise synthesis, one cannot be selectively removed.

Two other groups which have been proposed for the 5'phosphoryl group are the trichloroethyl ester (106) and the S-ethylphosphoro-thioates (107). Both of these groups are stable to alkali. The trichloroethyl group is readily removed by reduction (106), while the thioethyl group is removed by mild oxidation (107).

Recently, Ohtsuka et al. (108) have reported the use of the anilidate group to protect a 5'terminal phosphate in stepwise synthesis. This group is stable to the alkaline conditions required to remove the 3'-O-acetyl groups, and is removed readily by brief treatment with isoamyl nitrite.
In the chemical synthesis of the yeast alanine tRNA gene, block polymers were prepared with a 5' terminal nucleoside (protected with an acid labile, alkali stable, trityl group). In the final joining stages the 5' phosphoryl group was introduced by the enzyme polynucleotide kinase (109).

Phosphate-terminated oligonucleotides are required for the synthesis of phosphodiester-linked oligonucleotide-celluloses. In the present work, it was decided to synthesize the 5'-phosphorylated oligonucleotides directly.

The general method for stepwise synthesis of oligodeoxyribonucleotides bearing a 5' phosphoryl group involves:

1) condensation of two suitably protected nucleotides, alkaline treatment, and work up to give the protected dinucleotide, and

2) re-protection of the 5' phosphoryl group of the dinucleotide with the cyanoethyl group and condensation of this with another suitably protected mononucleotide (see Figure 2).

In the present series of experiments, complementary oligonucleotides of repeating sequence of d(pApApG)_n, d(pTpTpC)_n and d(pCpTpT)_n (n = 2, 3, 4) have been prepared, as well as the homodeoxyribooligomers, d(pT)_n and d(pA)_n. The homo-oligodeoxyribo nuleotides were made by polymerization of the suitably protected monomer units (d-pT, for the oligodeoxythymidylates;
Figure 2. General method for the synthesis of protected deoxyribotrinucleotides.
and d-pA^Bz, for the oligodeoxyadenylates). Homodeoxythymidylates
of length 3, 6, 9 and 12 were also prepared by stepwise
synthesis of d-pTpTpT, and chemical polymerization of this
trinucleotide.

The oligonucleotides of repeating sequence were synthesized
by the chemical polymerization of suitably protected tri-
The first two protected trinucleotides were prepared by stepwise
synthesis, according to the scheme outlined in Figure 2.
The trinucleotide d-pC^An pTpT was prepared by a slightly different
stepwise procedure.

In the Experimental and Results section, Part I, procedures
which follow directly from published reports have not been
repeated, but rather the references have been listed and the
yields and chromatographic and spectral data on the compounds
synthesized, are discussed.
MATERIALS AND METHODS

Nucleotides and Chemicals.

The oligodeoxyribonucleotides synthesized were all prepared from the four deoxyribonucleoside 5' monophosphates. These four nucleotides, d-pA, d-pG, d-pC and d-pT, were purchased from P-L Biochemicals, Inc., Milwaukee, Wisconsin, and Raylo Chemicals, Ltd., Edmonton, Alberta. The purity of these nucleotides was checked (at least 2 μmoles of each nucleotide) by paper chromatography in systems A, B and C (see page 34), and found to be 98% or better. The nucleotides were used without further purification.

The following list of chemicals and other materials, and their purification methods, were used. For the purification of chemicals by distillation, the chemicals were distilled through a 50 cm fractionating column containing glass tubing chips, except where noted otherwise. Distillations at reduced pressure were carried out using an oil vacuum pump which gave a pressure of about 0.1 mm.

Pyridine (B & A, purified grade) was distilled (bp 114°) at standard pressure from toluenesulfonyl chloride (Eastman Organics, Rochester, N.Y.) (25 g/litre). The distillate was stored at least two weeks over calcium hydride chips (Alpha
Benzylo chloride (Eastman Organics, Rochester N.Y.) was distilled under reduced pressure and stored in a dark bottle.

p-Anisoyl chloride (p-methoxybenzoyl chloride) (Aldrich Chemical Co., Inc., Milwaukee, Wisconsin) was distilled under reduced pressure (bp 87°). The clear distillate was stored in a dark bottle.

Acetic anhydride (B & A) was purified by distillation at reduced pressure, and stored in a dark bottle.

Hydracrylonitrile (β-cyanoethanol) (Eastman Organics, Rochester, N.Y.) was used as supplied.

Dicyclohexycarbodimide (DCC) (Aldrich Chemical Co., Inc., Milwaukee, Wisconsin) was used as supplied.

2-Mesitylenesulfonyl chloride (2,4,6-trimethylbenzenesulfonyl chloride) (Aldrich Chemical Co., Inc., Milwaukee, Wisconsin) was purified by charcoal filtration (2 to 3X) and recrystallization from warm cyclohexane (3X). The white crystals were dried over P₂O₅, in vacuo, and stored, in vacuo, over P₂O₅.

Triethylamine (Eastman Organics, Rochester, N.Y.) was distilled at standard pressure (bp 87°) from toluenesulfonyl chloride (25 g/litre).

Tributylamine (Eastman Organics, Rochester, N.Y.) was distilled under reduced pressure (bp 45°).

Trihexylamine (K & K Laboratories, Inc., Plainview, N.Y.) was
distilled under reduced pressure without the use of a fractionating column.

**Phosphorus pentoxide** (Granusic®) was the product of J.T. Baker Chemical Co., Phillipsburg, N.J.

**Bio-Rad AG 50W-X2** (Bio-Rad Laboratories, Richmond, California) 200-400 mesh, capacity 0.7 meq/ml resin bed, was regenerated by washing with at least five volumes 2 N NaOH, 2 volumes 1 N NaOH (50% ethanol), water, 5 volumes 2 N HCl, and water to neutrality.

**O-(diethylaminoethyl)cellulose**, standard grade (Schleicher & Schuell, Inc., Keene, New Hampshire), was precycled according to the method in the Whatman Technical Bulletin (116). Various preparations ranged from 0.9 to 0.98 meq/g dry cellulose.

All other chemicals used were standard laboratory chemicals of reagent grade.

**Methods**

In the synthesis of the oligonucleotides described in the Experimental and Results section, a number of procedures are similar for the different nucleotides. For brevity, these have been summarized below.

1. **Conversion of nucleotides to pyridinium salts.**

   The salt or free acid form of the nucleotide was dissolved in water (10 ml/5 mmoles) and passed through a Bio-Rad AG 50W-X2
column, 2.5 x 10 cm (pyridinium form, freshly prepared). The nucleotide was eluted in 100 to 200 ml dilute aqueous pyridine (5%) and concentrated either by rotary evaporation or freeze-drying.

2. Anhydrous conditions.

Nucleotides were rendered anhydrous by repeated evaporation (at least 3X) of dry pyridine on an oil vacuum pump equipped with a dry ice, methanol cold trap. The nucleotides then were suspended in the desired volume of anhydrous pyridine for the reaction. All synthetic steps, whether it be a protection step, or one involving the synthesis of an inter-deoxyribonucleotide bond were carried out in dry pyridine, in the dark, at room temperature, and all reaction flasks were shaken vigorously on a mechanical wrist action shaker.


When sulfonyl chloride condensation reactions are stopped by the addition of water, heat is produced. The reaction mixture was therefore first cooled in an ice-water bath prior to the addition of water, so that the flask did not become warm. This was done to minimize loss of protecting groups.


a. DCC Reactions. For working up DCC reactions, an equal volume of water was added, and the dicyclohexyl urea filtered off and washed with dilute aqueous pyridine. The
filtrate was extracted 3X with an equal volume of either cyclohexane or diethyl ether to remove unreacted carbodiimide, and the nucleotide solution left overnight at room temperature.

b. Sulfonyl chloride reactions. An equal volume of water and of trialkylamine was added to the reaction mixture which was left for 4 to 16 hours at room temperature.

c. Removal of terminal protecting groups. The general procedure for the removal of the cyanoethyl group from the 5'-phosphomonoester and the acetyl group from the 3'-hydroxyl of the nucleotides without removing the acyl protecting group on the heterocyclic amino residues was to cool the aqueous pyridine mixture to 0°, add an equal volume of 2 N NaOH (0°), stir for 20 minutes at 0°, and then rapidly neutralize with Bio-Rad AG 50W resin (hydrogen form). The nucleotides were recovered by filtering off the resin and washing it extensively with dilute aqueous pyridine.

5. Anion exchange chromatography of protected nucleotides.

The products of a condensation reaction were purified by chromatography on DEAE-cellulose columns. The protected nucleotides were chromatographed using either TEAB (pH 8.5) or TEAA (pH 5.5) buffers, normally in a cold room at 8° in an effort to minimize loss of protecting groups. Ninety-five percent ethanol (10 to 35% by volume) was included in the elution buffers.
6. Concentration of protected nucleotides following anion exchange chromatography.

It was found most convenient to concentrate the protected nucleotides by first removing the triethylammonium ions by adding Bio-Rad AG 50W resin (hydrogen form) until the pH dropped to below 5. The resin was then filtered off and the nucleotides recovered by extensively washing the resin with dilute aqueous pyridine. The nucleotides were then concentrated on a Büchi Rotavapor R evaporator using the vacuum of an oil pump. The condensor was cooled to -2° (Haake model KT41 circulating bath) such that a reasonable temperature gradient between condensor and evaporation flask could be maintained while keeping the evaporation flask at less than 20°. The nucleotides were rendered anhydrous, dissolved in a small volume of dry pyridine and added dropwise to anhydrous ether. The white precipitates were collected by centrifugation.

7. Removal of all protecting groups.

When all protecting groups (CE, OAc, N-Bz, N-An) were to be removed, the nucleotides, in a small volume of aqueous pyridine, were incubated with 3 volumes of concentrated NH₄OH for either a) 48 hours at room temperature (91); or b) 3 hours at 60° (105). The NH₄OH was removed by rotary evaporation at 35°.

8. Characterization of protected di- and trinucleotides.

The nucleotides were characterized by paper chromatography
prior to and after removal of protecting groups. The deprotected nucleotides were then treated with an aliquot of *Escherichia coli* alkaline phosphatase (Worthington or Miles) and again checked for homogeneity on paper. The dephosphorylated nucleotide was then degraded with an aliquot of snake venom phosphodiesterase (Worthington) and the products separated on paper (normally system A or D) (see 10 below). These were eluted (110) and the absorbance of each accurately measured in order to obtain molar ratios as a confirmation of structure.

9. The purification and characterization of longer oligonucleotides from the polymerization of protected trinucleotides.

After the polymerization reaction, the nucleotides were deprotected (see 7 above) and separated by anion exchange chromatography (DEAE-cellulose, CO$_3^-$ or Cl$^-$ form), in the presence of 7 M urea, followed by extended paper chromatography of appropriate peaks in system C (see 10 below). The major bands were eluted with 0.1% NH$_4$OH, and the purity of these longer oligonucleotides checked by alkaline phosphatase digestion of a sample of the nucleotide and chromatography in system C. The assumption is that major contaminants will be pyrophosphates and cyclic oligonucleotides, resistant to alkaline phosphatase (ref. 101, footnote 23). Thus, if all the alkaline phosphatase treated oligonucleotide moves ahead of non-treated nucleotide, then the oligomer contains a terminal 5'-phosphate group and is the desired, linear molecule.

All paper chromatography was done on Whatman #40 (double acid washed) paper, using the descending technique. Four systems have been used:

A 95% ethanol:M ammonium acetate, pH 7.5 (7:3)
B isobutyric acid:M ammonium hydroxide (5:3)
C n-propanol:concentrated ammonia:water (55:10:35)
D isopropanol:concentrated ammonia:water (7:1:2)

In checking the purity of nucleotides at least 2 μmoles (20 absorbance units of non-protected nucleotide, and as much as 40 absorbance units of protected nucleotide) were used. Nucleotides were detected under short ultraviolet light using a Chromatovue from Ultraviolet Products Incorporated.
EXPERIMENTAL AND RESULTS

1. Synthesis of Oligodeoxythymidylate Polymers Terminated with a 5'-Phosphate and 3'-Hydroxyl \( (d(pT)_n \ n = 2 \text{ to } 12) \).

The synthesis of oligodeoxythymidylate polymers was achieved by two separate procedures:

a) the polymerization of the mononucleotide, \( d-pT \), and

b) the polymerization of the trinucleotide, \( d-pTpTpT \).

a) Polymerization of \( d-pT \). The polymerization of \( d-pT \) was carried out according to Khorana and Conners (111), with the exception that no \( d-pT-OAc \) was used. The resultant nucleotides were chromatographed on a DEAE-cellulose column (CO\(_3\) form) and eluted with a linear gradient of \( NH_4HCO_3 \). Nucleotides up to the decamer, \( d(pT)_{10} \), were obtained plus a peak of higher oligonucleotides eluted with \( M NH_4HCO_3 \). (Rechromatography of the higher oligomers on a smaller DEAE-cellulose column resulted in resolution of oligonucleotides as long as \( d(pT)_{16} \).) The peaks corresponding to \( d(pT)_2 \), \( d(pT)_4 \), \( d(pT)_5 \) and \( d(pT)_7 \) were identified by alkaline phosphatase digestion followed by venom phosphodiesterase digestion to obtain the molar ratio of dT to d-pT (see Table I). Table III lists the yields of the various oligomers as well as the yields reported in the literature (111). Comparison with the reported yields indicates that this polymerization did not go as well, however, adequate yields of the
Table II. Characterization of dephosphorylated oligonucleotides by venom phosphodiesterase digestion and estimation of molar ratios.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1(^a) d-TpT</td>
<td>15</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.07</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:1</td>
</tr>
<tr>
<td>2(^a) d-TpTpTpT</td>
<td>16</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.04</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:3</td>
</tr>
<tr>
<td>3(^a) d-TpTpTpTpT</td>
<td>13</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.98</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:4</td>
</tr>
<tr>
<td>4(^a) d-TpTpTpTpTpTpT</td>
<td>7</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.70</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:6</td>
</tr>
<tr>
<td>5 d-TpT</td>
<td>16</td>
<td>1.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:1</td>
</tr>
<tr>
<td>6 d-TpTpT</td>
<td>7</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:2</td>
</tr>
<tr>
<td>7 d-TpTpC</td>
<td>5</td>
<td>1.03</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>1:1:1</td>
</tr>
<tr>
<td>8 d-CpTpT</td>
<td>12</td>
<td>-</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>1.88</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1:2</td>
</tr>
<tr>
<td>9 d-ApA</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>1.04</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>1:1</td>
</tr>
<tr>
<td>10 d-ApApG</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>0.93</td>
<td>-</td>
<td>-</td>
<td>1.16</td>
<td>1.00</td>
<td>-</td>
<td>1:1:1</td>
</tr>
</tbody>
</table>

\(^a\)Oligonucleotides obtained by polymerization of d-pT. All others were from stepwise condensation reactions.

\(^b\)Theoretical molar ratios of the expected nucleosides and nucleotides reading in order from left to right across the table.
Table III. Yield of oligonucleotides $d(pT)_n$ from the polymerization of $d$-pT.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>A.U. (267 nm)</th>
<th>Yield (%)</th>
<th>Literature yield %$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>di</td>
<td>9900</td>
<td>17.0</td>
<td>12.82</td>
</tr>
<tr>
<td>tri</td>
<td>9600</td>
<td>16.5</td>
<td>14.77</td>
</tr>
<tr>
<td>tetra</td>
<td>6500</td>
<td>11.2</td>
<td>12.19</td>
</tr>
<tr>
<td>penta</td>
<td>4090</td>
<td>7.0</td>
<td>7.29</td>
</tr>
<tr>
<td>hexa</td>
<td>2215</td>
<td>3.8</td>
<td>7.29</td>
</tr>
<tr>
<td>hepta</td>
<td>1200</td>
<td>2.1</td>
<td>4.39</td>
</tr>
<tr>
<td>octa</td>
<td>680</td>
<td>1.2</td>
<td>2.7</td>
</tr>
<tr>
<td>nona</td>
<td>410</td>
<td>0.7</td>
<td>1.37</td>
</tr>
<tr>
<td>deca and higher</td>
<td>643</td>
<td>1.1</td>
<td>2.63</td>
</tr>
</tbody>
</table>

$^a$Yield reported by Khorana and Conners (111).
oligomers up to the nonanucleotide were obtained, and as judged by chromatography in system C, the penta- through nonanucleotides were free from contaminants and suitable for use in experiments without further purification.

b) **Polymerization of the trinucleotide, d-pTpTpT.**

(i) **Stepwise synthesis of d(pT)$_3$.** The large scale synthesis of d(pT)$_3$ was carried out by the stepwise condensation of suitably protected deoxythymidylate units according to the following scheme.

\[
\text{CEd-pT} + \text{d-pT-OAc} \\
\quad \quad 1) \text{DCC, pyridine;} \\
\quad \quad 2) \text{N NaOH, 20 min, 0°.} \\
\text{d-pTpT} \downarrow \\
\text{DCC, hydrazonitrile, pyridine} \\
\quad \quad \text{CEd-pTpT} + \text{d-pT-OAc} \\
\quad \quad 1) \text{MsSO$_2$Cl, pyridine, THA} \\
\quad \quad 2) \text{N NaOH, 20 min, 0°.} \\
\text{d-pTpTpT}
\]

The procedure for the synthesis of the $\beta$-cyanoethyl derivative of d-pT has been described (112, 113). In a similar procedure, starting with 4 mmoles of pyridinium d-pT, a yield of 3.56 mmoles (89%) of the monocyanoethylated product was obtained.
The product was homogeneous in system A with $R_f = 0.75$, $R_{d-pT} = 1.61$ (Table IV).

The 3'-O-acetyl derivative of d-pT was prepared according to Khorana and Vizsolyi (101). The product, obtained in quantitative yield, was homogeneous in systems A and B, with $R_f$ in system A of 0.49, $R_{d-pT} = 1.21$, and $R_f$ in system B of 0.73, $R_{d-pT} = 1.50$. The spectra of both d-pT-OAc and CEd-pT were identical to d-pT.

The synthesis of the dinucleotide d-pTpT was carried out according to the general procedure reported by Narang et al. (103). The pyridinium salts of CEd-pT and d-pT-OAc (0.8 mmole each) were combined, rendered anhydrous, and dissolved in 3 ml dry pyridine. DCC (1.6 gm, 8 mmoles) was added and the reaction sealed and shaken for 4½ days. After the normal work up for a DCC condensation reaction (see Methods), the products were chromatographed on a DEAE-cellulose column (Fig. 3). Tubes 95 to 150 contained the desired product. The yield was 0.42 mmole (52.5%).

The dinucleotide was concentrated by rotary evaporation and characterized by paper chromatography and enzymatic degradation. The d-pTpT was homogeneous in system B, with $R_f = 0.42$. Enzymatic removal of the 5'-phosphate gave d-TpT, $R_f$ (system B) = 0.59. The molar ratio of dT to d-pT after venom phosphodiesterase digestion was 1.09 to 1.00 (Table II).
Table IV. Summary of $R_f$ values of non-protected and protected nucleotides.

<table>
<thead>
<tr>
<th>CPD.</th>
<th>Colour</th>
<th>Solvent System A</th>
<th>Solvent System B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$R_f$</td>
<td>$R_{d-pT}$</td>
</tr>
<tr>
<td>CEd-pT</td>
<td>P</td>
<td>0.75</td>
<td>1.61</td>
</tr>
<tr>
<td>d-pT-0Ac</td>
<td>P</td>
<td>0.49</td>
<td>1.21</td>
</tr>
<tr>
<td>d-pTpT</td>
<td>P</td>
<td>0.36</td>
<td>-</td>
</tr>
<tr>
<td>d-TpT</td>
<td>P</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CEd-pTpT</td>
<td>P</td>
<td>0.63</td>
<td>-</td>
</tr>
<tr>
<td>d-pTpTpT</td>
<td>P</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>d-TpTpTpT</td>
<td>P</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>d-pCAn</td>
<td>Bf</td>
<td>0.47</td>
<td>1.27</td>
</tr>
<tr>
<td>d-pCAn-0Ac</td>
<td>Bf</td>
<td>0.60</td>
<td>1.62</td>
</tr>
<tr>
<td>d-pTpTpCAn</td>
<td>Bf</td>
<td>0.27</td>
<td>0.68</td>
</tr>
<tr>
<td>d-pTpTpC</td>
<td>P</td>
<td>0.15</td>
<td>0.38</td>
</tr>
<tr>
<td>d-TpTpC</td>
<td>P</td>
<td>0.47</td>
<td>1.10</td>
</tr>
<tr>
<td>d-pTpT-0Ac</td>
<td>P</td>
<td>0.40</td>
<td>0.99</td>
</tr>
<tr>
<td>CEd-pCAn</td>
<td>Bf</td>
<td>0.60</td>
<td>2.34</td>
</tr>
<tr>
<td>dIEEd-pCAn</td>
<td>Bf</td>
<td>0.63</td>
<td>4.60</td>
</tr>
<tr>
<td>d-pCAnTpT</td>
<td>Bf</td>
<td>0.24</td>
<td>0.63</td>
</tr>
<tr>
<td>d-pCpTpT</td>
<td>P</td>
<td>0.12</td>
<td>0.28</td>
</tr>
<tr>
<td>d-CpTpT</td>
<td>P</td>
<td>0.38</td>
<td>1.04</td>
</tr>
<tr>
<td>d-pABz</td>
<td>PBf</td>
<td>0.50</td>
<td>1.30</td>
</tr>
<tr>
<td>*d-Ap-pABz</td>
<td>PBf</td>
<td>0.70</td>
<td>1.80</td>
</tr>
<tr>
<td>*d-Ap-pA</td>
<td>P</td>
<td>0.29</td>
<td>0.75</td>
</tr>
<tr>
<td>CEd-pABz</td>
<td>PBf</td>
<td>0.74</td>
<td>1.88</td>
</tr>
<tr>
<td>d-pABz-0Ac</td>
<td>PBf</td>
<td>0.64</td>
<td>1.60</td>
</tr>
<tr>
<td>d-pABz-pA</td>
<td>PBf</td>
<td>0.44</td>
<td>1.26</td>
</tr>
</tbody>
</table>
Table IV. Summary of \( R_f \) values of non-protected and protected nucleotides - Continued.

<table>
<thead>
<tr>
<th>CPD</th>
<th>Colour</th>
<th>Solvent System A</th>
<th>Solvent System B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( R_f )</td>
<td>( R_{d-pT} )</td>
</tr>
<tr>
<td>d-pApA</td>
<td>P</td>
<td>0.14</td>
<td>0.37</td>
</tr>
<tr>
<td>d-ApA</td>
<td>P</td>
<td>0.45</td>
<td>1.21</td>
</tr>
<tr>
<td>CEd-pApBz</td>
<td>PB&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.85</td>
<td>1.73</td>
</tr>
<tr>
<td>d-pG&lt;sup&gt;Bz&lt;/sup&gt;</td>
<td>Y&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.41</td>
<td>-</td>
</tr>
<tr>
<td>*d-G&lt;sup&gt;Bz&lt;/sup&gt;-pG&lt;sup&gt;Bz&lt;/sup&gt;(?)</td>
<td>Y&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.78</td>
<td>-</td>
</tr>
<tr>
<td>d-pG&lt;sup&gt;Bz&lt;/sup&gt;-OAc</td>
<td>Y&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.41</td>
<td>-</td>
</tr>
<tr>
<td>d-pApA&lt;sup&gt;Bz&lt;/sup&gt;pG&lt;sup&gt;Bz&lt;/sup&gt;</td>
<td>Y&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.43</td>
<td>1.00</td>
</tr>
<tr>
<td>d-pApApG</td>
<td>P</td>
<td>0.12</td>
<td>0.23</td>
</tr>
<tr>
<td>d-ApApG</td>
<td>P</td>
<td>0.32</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Abbreviations:

- \( R_{d-pT} \) - \( R_f \) relative to dTMP
- \( P \) - purple
- \( PB<sup>f</sup> \) - purple-blue fluorescent
- \( B<sup>f</sup> \) - blue fluorescent
- \( Y<sup>f</sup> \) - yellow fluorescent.

*These structures are the 5' linked pyrophosphates, d-N5'p-p5'. They are also abbreviated as

\[ \langle pA<sup>Bz</sup> \rangle, \langle pA<sup>Bz</sup> \rangle, \text{ and } \langle pG<sup>Bz</sup> \rangle. \]
Figure 3. Anion exchange chromatography d-pTptT. The sample (500 ml) was applied to a DEAE-cellulose column in the carbonate form (2.5 x 40 cm), pre-equilibrated with 0.05 M NH$_4$HCO$_3$, pH 8.0. After washing to remove the pyridine, the nucleotides were eluted with a linear gradient of NH$_4$HCO$_3$ (0.05 M to 0.15 M, 4 l total), at a flow rate of 2.5 ml per min.
The dinucleotide d-pTpT was converted to the cyanoethylated derivative using DCC and hydracrylonitrile \(^{103}\). Starting with 1.9 mmoles of the pyridinium salt of d-pTpT, a yield of 1.3 mmoles (68\%) of the desired product was obtained. The CEd-pTpT had a \(R_f\) of 0.63, and \(R_{d-pTpT}\) of 1.75 in system A.

The trinucleotide d-pTpTpT was prepared by condensing CEd-pTpT with a five-fold excess of the protected mononucleotide, d-pT-OAc. The trihexylammonium salts of CEd-pTpT (1.1 mmoles) and d-pT-OAc (5.1 mmoles) were rendered anhydrous and dissolved in 10 ml dry pyridine. \(\text{MsSO}_2\text{Cl} (3.24 g, 14.6 mmoles) was added and the reaction allowed to proceed for 4 hour. After the usual workup for a sulfonyl chloride condensation reaction (see Methods), the products were chromatographed on a DEAE-cellulose column (carbonate form (Fig. 4). The desired trinucleotide (39\% yield) was located in tubes 325 to 440. The product was pure as judged by chromatography in system B (\(R_f = 0.28\)).

Alkaline phosphatase digestion of a sample of the trinucleotide resulted in complete conversion to d-TpTpT (\(R_f = 0.46\) in system B). A molar ratio of dT to d-pT of 1.00 to 1.80 was obtained after venom phosphodiesterase digestion (Table II).

(ii) Polymerization of the trinucleotide d-pTpTpT (91).
The trinucleotide d-pTpTpT was polymerized using the condensing agent, \(\text{MsSO}_2\text{Cl}\). 0.16 mmmole of trinucleotide (4600 A.U. 266 nm) were converted to the pyridinium salt on a Bio-Rad AG 50W column,
Figure 4. Anion exchange chromatography of d-pTpTpT. The sample (2 l) was chromatographed on a DEAE-cellulose (CO₃²⁻) column (5 x 30 cm), pre-equilibrated with 0.05 M NH₄HCO₃. The nucleotides were eluted with a linear gradient of NH₄HCO₃ (0.05 M to 0.275 M, 14 l total) at a flow rate of 5 ml per minute.
and approximately 25% of this was acetylated in 1 ml dry pyridine and 0.3 ml acetic anhydride (overnight). After a workup similar to that for the synthesis of d-pT-OAc, the product, d-pTpTpT-OAc, was rendered anhydrous and precipitated with diethyl ether. This acetylated derivative was combined with the remainder of the trinucleotide and dissolved in 2 ml dry pyridine. Sufficient triethylamine (160 μl, 0.46 mmole) was included to solubilize the nucleotide. MsSO₂Cl (425 mg, 1.91 mmoles) was added and the reaction mixture immediately concentrated to a gum. After 3 hour, the reaction was stopped by the addition of 2 ml each of pyridine, water and triethylamine, and the flask was stored overnight (16 hour). The acetate group was removed by ammonolysis (see Methods), and the nucleotides were concentrated for chromatography on DEAE-cellulose (Fig. 5).

The peaks containing the various oligomers, d(pT)₃, d(pT)₆, d(pT)₉ and d(pT)₁₂ (indicated in Fig. 5), were pooled and concentrated to remove NH₄HCO₃. These samples were purified by extended chromatography on Whatman #40 paper in system C. The major bands were eluted with 0.1% NH₄OH, and the purity of the compound checked by alkaline phosphatase digestion of an aliquot (10 A.U.) and rechromatography.

Although the yield of the oligomers (Table V) was relatively high judged on this final polymerization step, the overall yields

---

⁹The overall yield is that yield based on the amount of starting protected mononucleotide. The overall yield for the hexanucleotide, d(pT)₆, is the product of the % yields in the syntheses of d-pTpT (52.5%), CEd-pTpT (68%), d-pTpTpT (39%), and the hexanucleotide, d(pT)₆ (18.5%, Table V).
Figure 5. Anion exchange chromatography separation of products from the polymerization of d-pTpTpT. The sample (100 ml) was applied to a DEAE-cellulose (CO\textsuperscript{3}) column (2.5 cm x 45 cm), pre-equilibrated with 0.075 M NH\textsubscript{4}HCO\textsubscript{3}. The nucleotides were eluted with a linear gradient of NH\textsubscript{4}HCO\textsubscript{3} (0.075 M to 1.0 M, 6 l total) at a flow rate of 3 ml per min (note that the absorbance 270 nm scale changes at tube number 143).
Table V. Yield of oligonucleotide obtained by polymerization of suitably protected trinucleotides.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Length of chromatography (hours)</th>
<th>Distance moved (cm)</th>
<th>A.U. Streaked on paper</th>
<th>A.U. recovered</th>
<th>% Yield&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Overall yield (%)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Yield in polymerization of mononucleotide (%)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>d(pT)&lt;sub&gt;6&lt;/sub&gt;</td>
<td>24</td>
<td>17</td>
<td>1,008</td>
<td>848</td>
<td>18.5</td>
<td>2.6</td>
<td>7.29</td>
</tr>
<tr>
<td>d(pT)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>48</td>
<td>15</td>
<td>620</td>
<td>527</td>
<td>11.5</td>
<td>1.6</td>
<td>1.37</td>
</tr>
<tr>
<td>d(pT)&lt;sub&gt;12&lt;/sub&gt;</td>
<td>48</td>
<td>8</td>
<td>300</td>
<td>118</td>
<td>2.6</td>
<td>0.36</td>
<td>not given; probably 0.4</td>
</tr>
<tr>
<td>d(pTpTpC)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>48</td>
<td>28.5</td>
<td>1,750</td>
<td>1,130</td>
<td>9.75</td>
<td>1.55</td>
<td>-</td>
</tr>
<tr>
<td>d(pTpTpC)&lt;sub&gt;3&lt;/sub&gt;</td>
<td>48</td>
<td>16.8</td>
<td>378</td>
<td>208</td>
<td>1.8</td>
<td>0.29</td>
<td>-</td>
</tr>
<tr>
<td>d(pTpTpC)&lt;sub&gt;4&lt;/sub&gt;</td>
<td>48</td>
<td>6.5</td>
<td>355</td>
<td>184</td>
<td>1.58</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td>d(pCpTpT)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>24</td>
<td>15</td>
<td>921</td>
<td>745</td>
<td>16.9</td>
<td>4.5</td>
<td>-</td>
</tr>
<tr>
<td>d(pCpTpT)&lt;sub&gt;3&lt;/sub&gt;</td>
<td>48</td>
<td>12</td>
<td>1,755</td>
<td>1,264</td>
<td>6.1</td>
<td>1.56</td>
<td>-</td>
</tr>
<tr>
<td>d(pApApG)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>72</td>
<td>19.0</td>
<td>461</td>
<td>127</td>
<td>4.5</td>
<td>0.11</td>
<td>-</td>
</tr>
<tr>
<td>d(pApApG)&lt;sub&gt;3&lt;/sub&gt;</td>
<td>120</td>
<td>8.8</td>
<td>307</td>
<td>56</td>
<td>1.97</td>
<td>0.048</td>
<td>-</td>
</tr>
<tr>
<td>d(pApApG)&lt;sub&gt;4&lt;/sub&gt;</td>
<td>168</td>
<td>5.3</td>
<td>140</td>
<td>39</td>
<td>1.37</td>
<td>0.034</td>
<td>-</td>
</tr>
</tbody>
</table>

<sup>a</sup>- % of total nucleotide polymerized;

<sup>b</sup>- overall yield is based on the yields of all intermediates from the protected mononucleotides times the final yield at the polymerization step. For d(pT)<sub>n</sub> this is the product of yields of d-pTpT (52.5%), CEd-pTpT (68%), and d-pTpTpT (39%) times column F. For d(pTpTpC)<sub>n</sub> this is the product of yields of d-pTpT (52.5%), CEd-pTpT (68%), and d-pTpTpC<sub>A</sub> (44.5%), times column F. For d(pCpTpT)<sub>n</sub> this is the product of yields of d-pTpT (52.5%), d-pTpT-OAc (100%), and d-pCpTpT (46%), times column F. For d(pApApG)<sub>n</sub> this is the product of yields of d-pAbzApAbz (45.5%), CEd-pAbzApAbz (69%), and d-pAbzApAbz<sub>A</sub> (7.8%).

<sup>c</sup>- yield in polymerization of mononucleotide is the yield of the corresponding oligonucleotide obtained from the polymerization of d-pT (111).
were not high. Considering those yields obtainable from a direct polymerization of \( d-pT \) (Table V), the amount of extra work involved in the stepwise synthesis of a trinucleotide and polymerization of this, as well as the lack of intermediate oligomers \((d(pT)_{n}, n = 4, 5, 7, 8, 10, 11)\) that could be useful in further block synthesis, it is obviously preferable to prepare the homo-oligodeoxyribonucleotides by a polymerization of suitably protected mononucleotides.

2. Synthesis of oligonucleotides of repeating sequence \( d(pTpC)_{n} \) \((n = 1 \text{ to } 4)\).

The synthesis of the oligodeoxyribonucleotides of repeating sequence \( d-pTpC \) was achieved by the polymerization of the protected trinucleotide \( d-pTpC^{An} \).

a) Synthesis of the protected trinucleotide, \( d-pTpC^{An} \).

The trinucleotide \( d-pTpC^{An} \) was prepared by stepwise synthesis of appropriately protected mono- and dinucleotides according to the general procedure reported by Narang et al. (103). The procedure is outlined below.

\[
\begin{align*}
CEd-pT + d-pT-OAc & \\
\text{1) DCC, pyridine;} & \text{2) } N \text{ NaOH, } 20 \text{ min, } 0^\circ.
\end{align*}
\]

\[
CEd-pTpT + d-pC^{An}-OAc
\]

\[
\begin{align*}
\text{1) MsSO}_2\text{Cl, pyridine, THA;} & \\
\text{2) } N \text{ NaOH, } 20 \text{ min, } 0^\circ.
\end{align*}
\]
The steps up to and including the synthesis of CEd-pTpT have been described above. The procedure used to prepare the N-anisoylated derivative of d-pC has been described in the literature (114). The work up used was that described with the exception that the d-pCAn was eluted from the Bio-Rad AG 50W column with 25% aqueous pyridine rather than the reported 5%. (This higher amount of pyridine was used because of the low solubility of the protected derivative in dilute aqueous pyridine.) The yield of the d-pCAn was quantitative. The product (blue fluorescent \( B^f \) under short UV) had an \( R_f \) of 0.47 and \( R_{d-pT} \) of 1.27 in system A. All preparations contained two minor contaminants. One was present in very small amounts (<1%), was blue fluorescent and had an \( R_f \) of 0.79. The other had an \( R_f \) of 0.83 and was identified by its spectrum as anisic acid. The preparation was used as such. The spectral data on the major spot (\( R_f 0.47 \)) agreed with that reported for d-pCAn (114); \( \lambda_{max} 302; \lambda_{min} 235, \) inflection 255.

The acetylation of d-pCAn was done as reported by Schaller and Khorana (115). The product obtained after lyophilization was a fluffy white powder which was insoluble in dry pyridine. The yield was quantitative and the spectral characteristics the same as for d-pCAn. The product had an \( R_f = 0.60 \) in system A, with a trace contaminant (\( B^f \)) at \( R_f = 0.73 \).

The synthesis of the protected trinucleotide, d-pTpTpCAn was carried out by the general procedure of Narang et al. (103).
CEd-pTpT (1 mmole) and d-pCAn-0Ac (3 mmoles) were combined (with 175 µl, 0.51 mmole THA) and rendered anhydrous. The sample was suspended in 10 ml dry pyridine and MsSO₂Cl (2 g, 9 mmoles) was added. The nucleotide suspension became homogeneous within 1 to 2 min. The reaction was left 6 hour and then worked up according to the general procedure for a sulfonyl chloride condensation reaction (see Methods). The products were chromatographed on a DEAE-cellulose column (acetate form) (Fig. 6). The desired trinucleotide was obtained from tubes 500 to 700. The yield was 0.445 mmole (44.5%) based on a calculated 

\[ E_{270} = 34,200 \]

for d-pTpTpCAn. The protected trinucleotide was homogeneous in system A with \( R_f = 0.27, R_{d-pT} = 0.68 \). The spectral properties of d-pTpTpCAn were \( \lambda_{max_1} = 186, \lambda_{max_2} = 303, \lambda_{min_1} = 237 \) and \( \lambda_{min_2} = 296 \) at pH 7.0.

In the characterization of d-pTpTpCAn, ammonolysis gave d-pTpTpC (\( R_f = 0.15, R_{d-pT} = 0.38 \), system A). Dephosphorylation of d-pTpTpC gave d-TpTpC (\( R_f = 0.47, R_{d-pT} = 1.1 \), system A). The molar ratio of dT:d-pT:d-pC obtained after venom phosphodiesterase digestion was 1.03:1.00:1.00 (Table II).

---

10 It was found that the monotrihexylammonium salt of the protected cytidine nucleotide was very insoluble in dry pyridine, and at concentrations greater that 1 mmole/10 ml consistently came out of solution. However, when the mixture was made anhydrous and the nucleotides suspended in the reaction volume of dry pyridine, fortunately, on the addition of the condensing agent, MsSO₂Cl, the protected d-pCAn goes into solution for the duration of the reaction.
Figure 6. Anion exchange chromatography of products from the synthesis of d-pTpTpC\textsuperscript{An}.

The sample (2 l) was loaded onto a DEAE-cellulose (\textsuperscript{\textregistered}0Ac) column (5 x 35 cm), pre-equilibrated with 0.1 M TEAA, pH 6.5 (35% ethanol). The sample was washed in with starting buffer and eluted with a linear gradient of TEAA (16 l), pH 6.5, 35% ethanol (0.1 M to 0.2 M) at a flow rate of 3 ml per min. The homogeneity of the trinucleotide peak was examined by looking at the ratio of absorbance at 303/270 nm.
b) Polymerization of the trinucleotide, d-pTpTpC<sup>An</sup> (91).

The pyridinium salt of d-pTpTpC<sup>An</sup> (0.445 mmole, 14,500 A.U. @ 270 nm) was polymerized by a procedure similar to that described for the polymerization of d-pTpTpT. The trinucleotide (25% as d-pTpTpC<sup>An</sup>-OAc) was rendered anhydrous and dissolved in 2 ml dry pyridine and 200 µl (0.58 mmole) trihexylamine. MsSO<sub>2</sub>Cl (2.7 mmoles, 0.6 g) was added and the mixture concentrated to a gum. The reaction was left for 2 hour and then worked up as described for the polymerization of d-pTpTpT. The products were chromatographed on a DEAE-cellulose column (Fig. 7). The tri-, hexa-, nona- and dodecanucleotide peaks were concentrated to remove ammonium bicarbonate and chromatographed in system C. The major bands corresponding to the hexa-, nona- and dodecanucleotides were eluted with 0.1% NH<sub>4</sub>OH and checked for purity by alkaline phosphatase treatment of a sample (10 A.U.) and rechromatography in system C. The yields of the oligonucleotides are given in Table V.

3. Synthesis of oligonucleotides of the sequence d(pCpTpT)<sub>n</sub> (n = 1 to 4).

Oligonucleotides of the sequence d(pCpTpT)<sub>n</sub> were prepared by the polymerization of the protected trinucleotide d-pC<sup>An</sup>pTpT.

a) Synthesis of d-pC<sup>An</sup>pTpT. The synthesis of the protected trinucleotide d-pC<sup>An</sup>pTpT was according to the procedure outlined.
Figure 7. Anion exchange chromatography of products from the polymerization of d-pTpTpC_{An}. The sample (210 ml) was loaded onto a DEAE-cellulose ($\text{CO}_3^{2-}$) column (2.5 x 36 cm), pre-equilibrated with 0.075 M $\text{NH}_4\text{HCO}_3$. The nucleotides were eluted with a linear gradient of $\text{NH}_4\text{HCO}_3$ (0.075 M to 0.65 M, 4 l total), at a flow rate of 2.5 ml per min. (Note that the absorbance 270 nm scale changes at tube number 165.)
The preparation of the dinucleotide d-pTpT and the protected mononucleotide d-pC^An has been described earlier. The dinucleotide was acetylated by reacting 1.75 mmoles of the pyridinium salt of d-pTpT in dry pyridine (20 ml) with 8.2 ml (87.5 mmoles) of acetic anhydride. The reaction was left for 5½ hour and then worked up in a similar procedure for the acetylation of d-pT (101). The yield of d-pTpT-OAc was quantitative, and the product had an R_f in system A of 0.40 (R_d-pT = 0.99).

The cyanoethylated derivative of d-pC^An (5.5 mmoles) was prepared by reaction in 25 ml dry pyridine with 25 ml hydracynitride (280 mmoles) and 4.5 g (21.6 mmoles) DCC, for 24 hr. The work up was similar to that for the preparation of CEd-pT (101) with the exception that the product was incubated for 4 hour at room temperature at a pH of 8.5 (adjusted with M TEAB, rather
than NH$_4$OH) in order to convert any dicyanoethylated derivative
to the mono-substituted product. The product obtained after
precipitation from dry pyridine by ether was a white powder.
The yield was 4.15 mmoles (75%). The spectrum was identical to
d-pC$^{An}$. The $R_f$ in system A of the major spot was 0.60
($R_{d-pT} = 2.34$). Minor spots at $R_f = 0.49$ (1% of total A.U. put
on paper) and $R_f = 0.63$ (6% of total A.U. put on paper) were
also evident. (This latter, major contaminant was probably
residuel dicyanoethyl derivative, as in a subsequent preparation,
incubation at pH 8.5 for 8 hour eliminated this spot.) The
preparation was used as such.

The trinucleotide d-pC$^{An}$pTpT was prepared using the condensing
agent MsSO$_2$Cl. The two reactions, CEd-pC$^{An}$ (4.15 mmoles) and
d-pTpT-OAc (1.75 mmoles) were combined and suspended in 30 ml
pyridine with 200 µl (0.58 mmole) trihexylamine. The nucleotides
were rendered anhydrous and suspended in 15 ml dry pyridine.$^{11}$
MsSO$_2$Cl (2 g, 9 mmoles), was added and the reaction allowed to
proceed for 5½ hour. The work up was that normally followed for
a sulfonyl chloride condensation (see Methods). The products
were chromatographed on a DEAE-cellulose column (Fig. 8).

$^{11}$The protected cytidine mononucleotide was insoluble in dry pyridine.
However, as soon as the sulfonyl chloride was added, the reaction
mixture was homogeneous within 1 to 2 min. This is similar to the
observation made during the synthesis of d-pTpTpC$^{An}$. 
Figure 8. Anion exchange chromatography purification of d-pC\textsuperscript{An}pTpT.
The sample was loaded onto a DEAE-cellulose (acetate form) column (5 x 40 cm),
pre-equilibrated with 0.12 M TEAA, pH 5.5. The nucleotides were eluted with a
linear gradient 0.12 M to 0.22 M TEAA, pH 5.5, 25% ethanol (14 l total), at a
flow rate of 3 ml per min.
Tubes 440 to 660 contained the desired product, d-pC\textsuperscript{An}pTpT. The yield was 27,300 A.U.\textsubscript{272} nm (46%) (a calculated extinction coefficient $E_{270} = 34,200$ was used). This protected trinucleotide was pure as judged by chromatography in system A ($R_f = 0.24$, $R_d$-pT = 0.63).

After ammonolysis, the d-pCpTpT had an $R_f$ 0.12 in system A ($R_d$-pT = 0.28). Removal of the 5'-phosphate with alkaline phosphatase resulted in complete conversion to the trinucleotide d-CpTpT ($R_f = 0.38$, system A). The molar ratio of dC to d-pT after venom phosphodiesterase digestion was 1.00 to 1.88 (Table II) confirming the structure.

b) Polymerization of d-pC\textsuperscript{An}pTpT (91). The protected trinucleotide d-pC\textsuperscript{An}pTpT (0.77 mmole, 26,400 A.U.\textsubscript{272} nm) was polymerized in a similar manner to that described for d(pT)\textsubscript{3} and d-pTpTpC\textsuperscript{An}. Thus, 20% of the trinucleotide was converted to d-pC\textsuperscript{An}pTpT-OAc and combined with the remaining nucleotide. The sample was dissolved in pyridine and a minimum amount of trihexylamine added to solubilize the nucleotides (300 µl, 0.87 mmole). The sample was rendered anhydrous, dissolved in 5 ml dry pyridine and MsSO\textsubscript{2}Cl (2.04 g, 9.2 mmoles) was added. The solution was immediately concentrated to a gum and left 2 hour. The reaction was worked up according to that described for the polymerization of d(pT)\textsubscript{3}. The products were separated on a DEAE-cellulose (carbonate form) column (Fig. 9). The peaks
Figure 9. Anion exchange chromatography of products from the polymerization of d-pC^AnpTpT. The sample was applied to a DEAE-cellulose (CO\textsuperscript{3}H\textsuperscript{+}) column (2.5 x 50 cm), pre-equilibrated with 0.075 M NH\textsubscript{4}HCO\textsubscript{3}. The nucleotides were eluted with a linear gradient of NH\textsubscript{4}HCO\textsubscript{3} (7 l total), 0.015 M to 0.8 M. The flow rate was 20 ml per 9 min. (Note that the A\textsubscript{270 nm} scale changes at tube number 200.)
containing the hexa- and nonanucleotides were concentrated and purified by extended chromatography in system C. The purity of these nucleotides was checked by digestion with alkaline phosphatase, and rechromatography in system C. Table V lists the yields of the oligonucleotides at the final polymerization step and the overall yields.

4. **Synthesis of oligodeoxyadenylate polymers, d(pA)<sub>n</sub> (n = 2 to 12).**

The oligodeoxyadenylates used in this study were prepared by Dr. Michael Smith. The method used in their preparation was the polymerization of d-pA<sup>Bz</sup> with DCC in anhydrous pyridine (100).

5. **Synthesis of oligodeoxyribonucleotides of repeating sequence, d(pApApG)<sub>n</sub> (n = 1 to 4).**

The synthesis of deoxyribo-oligomers of repeating sequence d-(pApApG)<sub>n</sub> was achieved by the polymerization of the protected trinucleotide, d-pA<sup>Bz</sup>pA<sup>Bz</sup>pG<sup>Bz</sup>.

a) **Synthesis of the protected trinucleotide, d-pA<sup>Bz</sup>pA<sup>Bz</sup>pG<sup>Bz</sup>.**

The trinucleotide d-pA<sup>Bz</sup>pA<sup>Bz</sup>pG<sup>Bz</sup> was prepared by a stepwise condensation of suitably protected mono- and dinucleotides according to the following scheme.
The synthesis of d-\text{pA}^\text{Bz}\ was carried out according to the procedure of Ralph and Khorana (100). It was found that, by increasing slightly the volumes of pyridine and benzoyl chloride relative to the nucleotide (30 ml pyridine, 3 ml benzoyl chloride per mmole nucleotide instead of 20 ml pyridine, 2.5 ml benzoyl chloride (100)) and by keeping the reaction time to a minimum, the desired product could be obtained in quantitative yield with a minimum of impurities. The product obtained by lyophilization was a fluffy white powder. It has an $R_f$ in system A of 0.55 ($R_{d-pT} = 1.36$). However, in all preparations there was a second,
faster migrating spot \( R_f = 0.7, R_{d-pT} = 1.91 \). The second spot is probably the dinucleoside pyrophosphate \( \text{pA}^{\text{Bz}} \) and, as judged by the absorbance at 282 nm, was present, consistently, in several preparations, at 10 to 11% of the total nucleotide. It was found that this major contaminant could be removed by chromatography on a DEAE-cellulose \((\text{CO}_{3}^\text{m})\) column, eluting with a linear gradient of TEAB to 0.12 M.

Benzoylated adenosine residues are difficult to detect under UV light, and in order to see the impurity, at least 25 to 30 A.U. at 280 nm must be spotted on the paper. To distinguish these benzoylated adenosine residues from anisoylated cytosine residues, I have described the former as purple blue fluorescent \((\text{PB}^f)\) and the latter as blue fluorescent \((\text{BF}^f)\), under short UV.

The faster moving spot was characterized by chromatography prior to and after removal of the benzoyl group. Aside from the \( R_f \) values in systems A and B (Table IV), the compound had an \( R_f = 0.54 \) (system C) and 0.17 (system D). In all of these systems this (debenzoylated) compound was not completely resolved from \( \text{d-pA} \). However, in a fifth chromatography system (0.1 M phosphate buffer, pH 6.8 (100 ml: \((\text{NH}_4\)_2\text{SO}_4 \ (60 g): \text{n-propanol (2 ml))}, the two are resolved. The compound was stable to alkaline phosphatase, while digestion with venom phosphodiesterase converted it quantitatively to the mononucleotide \( \text{d-pA} \). These data are consistent with the structure of the contaminant being the \( \text{N,N}^\prime \)-dibenzoyl dinucleoside pyrophosphate, \( \text{pA}^{\text{Bz}} \).

The spectrum of this compound was slightly different from that of \( \text{d-pA}^{\text{Bz}} \) (\( \lambda_{\text{max}} \text{d-pA}^{\text{Bz}} 282 \text{ nm}, \lambda_{\text{max}} \text{pA}^{\text{Bz}} 283 \text{ nm} \)).
(25% in ethanol) at 8°. Under these conditions the contaminant was eluted in a trailing shoulder of the main peak of UV absorbing material. Later it was found by Dr. M. Smith that chromatography of the nucleotides on a similar column in the absence of ethanol in the elution gradient gave an improved separation. Under these conditions, the pyrophosphate elutes as a small peak not completely resolved from the main peak.

CED-pA_Bz was prepared according to Ohtsuka et al. (113). Starting with 7 mmoles of d-pA_Bz (purified) the desired product 6.8 mmoles (97%) was obtained. The product had an R_f in system A of 0.74, R_d-pT = 1.88.

The 3'-O-acetyl protected derivative of d-pA_Bz was prepared according to a method described for the preparation of d-pA_Ac-OAc (100). Starting with 1.4 mmoles of d-pA_Bz, a yield of 0.98 m mole (70%) of d-pA_Bz-OAc was obtained. (In other preparations, the yields were normally higher; 90 to 100%.) The product obtained by precipitation from pyridine with anhydrous ether was a white powder. It was homogeneous in system A (R_f = 0.64, R_d-pT = 1.60).

The spectra of the CE- and 3'-O-acetyl derivatives of d-pA_Bz were similar to that of d-pA_Bz.

The dinucleotide d-pA_Bz-pA_Bz was prepared by condensation of equimolar amounts of CEd-pA_Bz and d-pA_Bz-OAc (103). 2.9 mmoles of CEd-pA_Bz and 3.1 mmoles of d-pA_Bz-OAc were condensed
in 15 ml dry pyridine with 4.2 mmoles DCC and 1.2 g Bio-Rad AG 50W resin in the pyridinium form. After 6 days the reaction was worked up as for a DCC condensation (see Methods). The products were chromatographed on a DEAE-cellulose column (Fig. 10). Tubes 260 to 360 contained the dinucleotide, which was concentrated by the procedure described in Methods. The product was obtained as a white powder by precipitation from dry pyridine with ether. The yield was 48,000 A.U. 282 nm (1.32 mmoles, 45.5% based on a calculated $E_{282}^\infty$ of 36,400). The protected dinucleotide was homogeneous in system A ($R_f = 0.44$, $R_{d-pT} = 1.26$).

Characterization of the product, after ammonolysis, showed that d-pApA had an $R_f = 0.14$ in system A. Alkaline phosphatase digestion resulted in complete conversion to d-ApA ($R_f$ in system A = 0.45). The molar ratio of dA to d-pA after venom phosphodiesterase digestion was 1.00 to 1.04, confirming the structure d-pApA.

$\text{CEd-pA}^\text{Bz} \cdot \text{pA}^\text{Bz}$ was prepared according to the general procedure reported by Narang et al. (103). A yield of 0.9 mmole (69%) of the $\text{CEd-pA}^\text{Bz} \cdot \text{pA}^\text{Bz}$ was obtained. The product was homogeneous in system A with $R_f = 0.85$ ($R_{d-pT} = 1.73$).

The protected nucleotide d-pG$^\text{Bz}$ was prepared by reacting the pyridinium salt of d-pG with benzoyl chloride in dry pyridine (102). The volumes of pyridine and benzoyl chloride were increased from that reported by Ralph et al. (102), to
Figure 10. Anion exchange chromatography purification of d-pA$^{Bz}pA^{Bz}$.
The sample (1800 ml) was loaded onto a DEAE-cellulose ($CO_3^-$) column (5 x 25 cm),
pre-equilibrated with 0.05 M TEAB (25% ethanol by volume) at 8°. The nucleotides
were eluted with an 8 l gradient in TEAB (25% ethanol), from 0.05 M to 0.175 M,
at a flow rate of 5 ml per min.
30 ml pyridine, 3 ml benzoyl chloride per mmole mononucleotide, and the reaction was allowed to proceed for 2 hour. After a work up similar to that described in the literature (102), the product was chromatographed to check its purity. The major spot (yellow fluorescent, Y) had an Rf = 0.48, R_{d-pT} = 1.87 in system A. A second, minor spot (also Y) had an Rf = 0.78. (This was probably the pyrophosphate, pG\text{Bz}^- . The product was usually obtained as a light brown powder and was routinely purified by chromatography on a DEAE-cellulose (acetate) column at room temperature, eluting with a linear gradient of TEAA pH 5.5, to 0.3 M (25% in ethanol). The major peak at 290 nm, eluted between 0.15 M and 0.175 M TEAA, was recovered as a fluffy white powder after lyophilization. The d-pG\text{Bz}^- was homogeneous in system A

14An early attempt to purify d-pG\text{Bz}^- on a DEAE-cellulose (CO\text{3-}) eluting with a gradient of TEAB pH 8.5 (25% ethanol) resulted in very poor resolution of the UV absorbing material. The nucleotide eluted over a range of 0.05 M to 0.2 M TEAB. A possible explanation for this poor resolution may be that, as mentioned in the Introduction, benzoylation of the heterocyclic amino group of d-pG lowers the pK of the ring hydroxyl. If there was a change in pH of the gradient (due to loss of CO2 over the 48 hour chromatography period, or due to using older, higher pH, TEAB buffer to make up the 0.3 M reservoir), then as the concentration of eluting ion increased, so would the net charge on the protected nucleotide. Aggregation of the nucleotides may also be contributing to the poor resolution observed.(116).

15Although this column removed only 5% of the UV absorbing material (probably the pyrophosphate, pG\text{Bz}^- ), it also removed all the visible (brown) colour from the sample. At the end of the gradient at least 4 distinct coloured bands are visible on the column.
with \( R_f = 0.48 \). A second UV absorbing peak (5% of the A.U. at 290 nm) eluted between 0.225 M to 0.275 M TEAA. This peak corresponded to the yellow fluorescent spot at \( R_f 0.78 \) and was probably the pyrophosphate \( \text{d-pG}^{\text{Bz}} \). The spectrum of d-pG^{\text{Bz}} was identical to that reported by Ralph et al. (102); \( \lambda_{\text{max}} 291, 262, \text{shoulder} 241, \lambda_{\text{min}} 276 \) and 223. The yields ranged from 50 to 80%.

The acetylation of d-pG^{\text{Bz}} was carried out by the method used for the acetylation of d-pA^{\text{Bz}}. The product, obtained by precipitation of the nucleotide from pyridine with anhydrous ether, was a white powder. The yield was quantitative. The d-pG^{\text{Bz}}-0Ac had an \( R_f = 0.53, R_{\text{d-pG}} = 2.5 \), and the spectral properties were similar to those of d-pG^{\text{Bz}}.

The trinucleotide d-pA^{\text{Bz}}pA^{\text{Bz}}pG^{\text{Bz}}-0Ac was prepared by condensation of CEd-pA^{\text{Bz}}pA^{\text{Bz}} (0.87 mmole) and d-pG^{\text{Bz}}-0Ac (4.35 mmoles) in 10 ml dry pyridine (103). The nucleotides were solubilized by the addition of 200 µl (0.58 mmole) trihexylamine, and the condensing agent MsSO_2Cl (2.45 g, 11 mmoles) added. The reaction was left 6 hour, and then worked up in the usual way for a sulfonyl chloride condensation (see Methods). The nucleotides were chromatographed on a DEAE-cellulose (C0^3) column (Fig. 1). Trinucleotide was eluted between 0.185 M to 0.225 M TEAB. However, the product was impure,\(^\text{16}\) and was rechromatographed on a second DEAE-cellulose (C0^3)

\(^{16}\)The major contaminant was not clearly resolved from d-pA^{\text{Bz}}pA^{\text{Bz}}pG^{\text{Bz}} by paper chromatography in system A; however, removal of the benzoyl group by ammonolysis and rechromatography gave two nucleotide spots (\( R_f 0.12, \lambda_{\text{max}} 256 \), the trinucleotide d-pApApG, and \( R_f 0.20, \lambda_{\text{max}} 252 \) with the spectrum of a G compound). The minor spot was stable to alkaline phosphatase digestion and attempts to digest this nucleotide with venom phosphodiesterase were unsuccessful, resulting in trace amounts of a product with an \( R_f \) similar to dG. The remainder of the nucleotide was undegraded.
Figure 11. Anion exchange chromatography of d-pA^Bz_pA^Bz_pG^Bz at pH 8.5.
The sample (2 l) was applied to a DEAE-cellulose (CO_3^2-) column, 5 x 35 cm, pre-equilibrated with 0.125 M TEAB, pH 8.5. The column was washed with 0.175 M TEAB pH 8.5, and then eluted with a gradient of TEAB, 0.175 M to 0.275 M (16 l total) at a flow rate of 3 ml per min. All solutions were 25% ethanol, by volume.
column, eluting with TEAB, pH 8.5. The product after this second chromatography at pH 8.5 was still contaminated. Since the benzoylated derivative, d-pG<sub>Bz</sub> was eluted as a sharp peak from a DEAE-cellulose (acetate) column at pH 5.5, it was decided to try to purify the trinucleotide d-pA<sub>Bz</sub>pA<sub>Bz</sub>pG<sub>Bz</sub> under similar conditions (Fig. 12). The UV absorbing material eluted between 0.15 M and 0.2 M TEAA had a "benzoylated G" spectrum, while tubes numbers 50 to 82 contained the trinucleotide d-pA<sub>Bz</sub>pA<sub>Bz</sub>pG<sub>Bz</sub>. The yield at this point was 3,980 A.U. (Based on a calculated \( E_{280} \) for the protected trinucleotide of 50,500 the yield was 7.8%). The product was homogeneous in system A with \( R_f = 0.42 \).

After ammonolysis, the trinucleotide had an \( R_f = 0.12 \) and \( R_{d-pT} = 0.23 \) in system A. The trinucleotide was dephosphorylated by alkaline phosphatase to d-ApApG (\( R_f = 0.32 \) in system A), and the molar ratio of dA to d-pA to d-pG after venom phosphodiesterase digestion was 0.93:1.16:1.00, consistent with the structure d-ApApG.

b) Polymerization of the protected trinucleotide, d-pA<sub>Bz</sub>pA<sub>Bz</sub>pG<sub>Bz</sub>. The trinucleotide d-pA<sub>Bz</sub>pA<sub>Bz</sub>pG<sub>Bz</sub> was polymerized according to the procedure described by Narang et al. (91). Approximately 20% of the 4400 A.U. (0.087 mmole) was acetylated and combined with the remaining trinucleotide. The nucleotides were suspended in 3 ml dry pyridine and 50 µl trihexylamine (0.15 mmole). The condensation was carried out by adding 230 mg (1.35 mmoles) of MsSO<sub>2</sub>C1 and concentrating the mixture to a gum. After 2 hour, the reaction was worked up in a similar way to that described
Figure 12. Anion exchange chromatography of d-pA<sup>Bz</sup> gA<sup>Bz</sup> pG<sup>Bz</sup> at pH 5.5. The sample (2.1, 7400 A.U. <sub>280</sub>) was applied to a DEAE-cellulose (acetate) column (1.2 x 40 cm), pre-equilibrated with 0.175 M TEAA, pH 5.5 (30% ethanol, by volume). The nucleotides were eluted with a linear gradient of TEAA (2.5 1), pH 5.5 (0.175 M to 0.4 M, 30% ethanol) at a flow rate of 2 ml per min.
earlier for the polymerization of d-pTpTpT. The nucleotides were chromatographed on a DEAE-cellulose (chloride) column (Fig. 13). The tubes containing the hexa-, nona- and dodeca-nucleotides were pooled, diluted five-fold with distilled water and desalted on a small DEAE-cellulose (CO$_3^-$) column, 1.2 x 15 cm. The nucleotides were eluted with M NH$_4$HCO$_3$ and concentrated by rotary evaporation prior to chromatography on Whatman #40 paper in system C (see Table V). The overall yields of the oligonucleotides are very low; however, this is not primarily due to a low yield at the final polymerization step, but rather, due to the low yield of the trinucleotide.
Figure 13. Anion exchange chromatography of the products from the polymerization of d-pA<sup>3</sup>Bz<sup>2</sup>pA<sup>3</sup>Bz<sup>2</sup>pG<sup>3</sup>Bz<sup>2</sup>. The sample (100 ml) was applied to a DEAE-cellulose (Cl<sup>-</sup>) column (1.2 x 50 cm), pre-equilibrated with 0.1 M sodium acetate, pH 5.5, 0.02 M in NaCl, 7 M in urea. The nucleotides were eluted with this buffer using an increasing gradient of NaCl.
DISCUSSION

Oligodeoxyribonucleotides of the series $d(pT)_n$ and $d(pA)_n$ have been prepared, as well as oligonucleotides containing the repeating sequences, $d(pT)pTpC)_n$, $d(pC)pTpT)_n$, and $d(pA)pApG)_n$ ($n = 1$ to 4).

The homo-oligodeoxyribonucleotides were prepared because of their relative ease of synthesis (particularly oligodeoxythymidylates), compared with the synthesis of oligonucleotides containing specific, mixed nucleotide sequences. These oligonucleotides were suitable for the study of a number of the properties of interactions between complementary oligonucleotides, both in solution and on oligonucleotide-cellulose.

The three complementary repeating sequences were prepared in order that one might be able to evaluate the effect of mixed base sequences (including GC base pairs) on the stability of oligonucleotide interactions. The decision to synthesize repeating trinucleotide sequences was governed by the availability of general methods for the synthesis of protected deoxyribotrinitucleotides (103). These could then be polymerized, in one step, to defined, repeating sequences, 6, 9, and 12 nucleotides long (91).

The yields of dinucleotides (45 to 55%) and trinucleotides (40 to 45%) were somewhat lower than those reported by Narang
et al. (103) for the synthesis of similar compounds (60 to 70% and 50 to 60%). However, all intermediates were homogeneous and demonstrated to contain the appropriate nucleoside and nucleotide(s). Yields of from 2 to 10 μmoles of oligonucleotides (prepared by the polymerization of suitably protected trinucleotides) were practicable, and these amounts are more than adequate for the synthesis of oligonucleotide-celluloses.

Further details on the synthesis of particular sequences are given below.

Jacob and Khorana (117) have reported the synthesis of the oligomers d(TpTpC)ₙ by stepwise condensation of suitably protected mononucleotides to the 3' hydroxyl end of a growing chain. The 5' nucleotide was 5'~0~-trityl dT. In the present experiments, the protected deoxyribotrinucleotide, d-pTpTpCAn, was prepared, and this was polymerized to obtain the hexa-, nona-, and dodecanucleotides.

The synthesis of the oligonucleotides d(pCpTpT)ₙ (n = 1 to 4) has not been reported. These were prepared by a modification of the general procedure described by Narang et al. (103, 91) (see Figure 2). The protected trinucleotide d-pCAnTpT was prepared by condensation of the mononucleotide CEd-pCAn with the di-nucleotide, d-pTpT-0Ac. The product, d-pCAnTpT, was then polymerized to give the hexa-, nona-, and dodecanucleotides.

The synthesis of oligomers of repeating sequence, d(pApApG)ₙ
has been reported in the literature (91). In the present experiments the protected trinucleotide, d-\(\text{pA}^{Bz}\text{pA}^{Bz}\text{pG}^{Bz}\) was prepared rather than d-\(\text{pA}^{Bz}\text{pA}^{Bz}\text{pG}^{Ac}\), as reported by Narang et al. (103). This trinucleotide was then polymerized to obtain higher oligomers.

The synthesis of the three protected deoxyribotrinucleotides, d-\(\text{pTpC}^{An}\), d-\(\text{pC}^{An}\text{pTpT}\), and d-\(\text{pA}^{Bz}\text{pA}^{Bz}\text{pG}^{Bz}\), has, so far, not been reported in the literature. These trinucleotides were therefore carefully characterized by chromatography, prior to and after the removal of the N-acyl protecting groups and also after treatment with alkaline phosphatase (Table IV). The resultant trinucleoside diphosphates were degraded with venom phosphodiesterase and the molar ratios of the nucleoside and nucleotides determined as a confirmation of the structure (Table II).

Studies with the isomeric repeating sequences, d(\(\text{pTpTpC}\))\(_n\) and d(\(\text{pCpTpT}\))\(_n\) with d(\(\text{pApApG}\))\(_n\) sequences resulted in a number of interesting observations, due to the slightly different hybrid structures that can form as a result of a "shift in the pairing frame" (see below) in d(\(\text{pTpTpC}\))\(_n\), d(\(\text{pApApG}\))\(_n\) interactions.

\[
\begin{align*}
\text{3'} & \text{pTpTpCpTpTpC} \\
\end{align*}
\]

"shifted pairing frame"
PART II

THERMAL ELUTION OF OLIGONUCLEOTIDES ON
CELLULOSE COLUMNS CONTAINING OLIGONUCLEOTIDES
OF DEFINED LENGTH AND SEQUENCE
INTRODUCTION

It was decided to examine the interaction of oligonucleotides in solution prior to studying the elution of complementary oligomers from oligonucleotide-celluloses. A study of the thermal dissociation curves of complementary oligomers would provide a guide to the properties of the various oligonucleotide-cellulose: oligonucleotide interactions.

A number of studies on homopolymer:oligonucleotide interactions have been reported (see 118-123). In these experiments, the hybrid formation or destruction has been followed by one or more of the following: the decrease in absorbance observed on mixing a complementary oligomer and polymer, thermal denaturation profiles, sedimentation studies in the analytical ultracentrifuge, gel filtration on Sephadex columns, equilibrium dialysis, or the change in specific rotation of the mixtures. In these studies, the structures involved either 2-stranded or 3-stranded complexes (depending on the ionic strength and temperature of the mixture), and in most cases, the interactions are greatly stabilized by cooperative binding of the complementary oligonucleotide.\(^{17}\)

In studies applicable to the problems discussed in this

\(^{17}\)Cooperative binding refers to the reduction of free energy or stabilization that results due to vertical stacking of the consecutive bases of adjacent oligomers.
thesis, it was felt that oligonucleotide:oligonucleotide inter-
actions would provide a more useful study for the extrapolation
of data to experiments concerning the isolation of nucleic acids
(i.e. oligonucleotide:random sequence polymer interactions). The
reason for this is that, as mentioned above, there is a considerable
stabilization associated with cooperative binding of small
oligomers to a complementary homopolymer. For example, Pitha and
Ts' o (118) reported that adenosine binds to poly U in the presence
of the adenosine heptanucleoside hexaphosphate (A(pA)_6). These
workers feel the interaction of adenosine is stabilized by
cooporative binding of this nucleoside between two A(pA)_6 residues.
Lipsett et al. (121) were able to observe an interaction between
poly U and oligomers as small as the trinucleotide, ApApA. In
other work, Lipsett (122) observed a Tm of 25.5° for the inter-
action of GpG with poly C. However, for the interaction of a
short sequence with its complement within a random sequence polymer
(see illustration below), end to end binding (i.e. vertical stacking

```
-A
-G
-U·A
-U·A
-U·A
-U·A
-G
-C
-G
-A
-C
```
of the oligonucleotide units) is not possible.

Naylor and Gilham (124) have reported the interaction between oligodeoxyribonucleotides of the series, \( d(pT)_n \) and \( d(pA)_n \). In their investigation, the interactions were studied by measuring the % hypochromicity observed on mixing the two components and reducing the temperature to 0°.

One criticism of the oligonucleotide:oligonucleotide interaction model system is that in oligomer:random sequence polymer interactions, the effect of the long "tails" on the polymers may be to greatly reduce the stability of the interaction. Studies concerned with a non-complementary sequence (or base), attached to a complementary sequence have been reported by Lipsett and coworkers (121) and Bautz and Bautz (125). For example, the stability of poly U.ApApApApU and poly U.ApApApU complexes have been studied. It was found that the terminal uridine residue reduces the stability of the interaction, although the stoichiometry of the reaction indicated the "tails" were bent out from the helix, to permit the next oligonucleotide to bind in a cooperative way. These extra uridine residues also prevented the formation of triple-stranded structures (121), however three-stranded structures with ApApApU.poly U do form (125). It is difficult to extrapolate these observations to oligomer:random sequence interactions, as again, end to end binding of the oligonucleotide is not the case, and the reduction in stability of the hybrid in the
experiments reported may be due to partial interference of vertical stacking. Experiments of Szybalski and his coworkers (126-130), in which differential hybridization of the two strands of a DNA duplex with poly G, poly (I, G) or poly (U, G) has been used to facilitate the separation of the two strands by ultracentrifugation in cesium chloride, suggest that it should be possible to isolate a short hybrid region with long unbound "tails". However, these workers feel that the region of hybrid structure extends for 15 to 40 nucleotides (129).\(^{18}\) The increased length of the hybrid region, plus the fact that the interactions are all GC interactions, would contribute greater stability to these types of structures compared to that anticipated in the isolation of nucleic acids by oligomer:random sequence polymer interactions.

Niyogi and Thomas (131, 132) have reported the formation of specific, stable oligoribonucleotide:DNA hybrids with oligomers as short as the decanucleotide for both T4 and T7 viral DNA: oligomer interactions. McConaughy and McCarthy (133) have reported stable oligonucleotide:DNA complexes with oligodeoxyribonucleotide chain lengths approximately 11-14 residues. Clearly, the minimum stable length of an oligonucleotide:DNA complex must be a function of solvent, temperature, and base composition.

Gillespie and Spiegelman (134) reported that the minimum chain

\(^{18}\)Recent evidence would suggest that the interaction of consecutive cytidylate tracts with guanylate residues may be much shorter (14, 15).
length oligoribonucleotide (from rRNA) that hybridized with homologous DNA was 50 at 67°, 32 at 55°, 17 at 44° and greater than 6 at 23° and 37° (in 2 x SSC).\textsuperscript{19} In these experiments of Thomas, McCarthy and Spiegelman, and their associates, the oligonucleotide fractions were obtained by limited degradation of the nucleic acid (with a nuclease, by partial depurination or alkali), and the resultant oligomers separated according to chain length. The observations by these workers that stable specific oligonucleotide:random sequence polymer hybrids can be formed provide encouragement for the success of techniques designed to isolate a particular nucleic acid by hybridization with a small, complementary oligonucleotide of defined sequence. The possibility of success is further encouraged by a report by Jovin and Kornberg (135), in which a deoxythymidine polynucleotide-cellulose\textsuperscript{20} was shown to retain, specifically, single-stranded DNA (molecular weight $2.5 \times 10^6$ daltons) when the DNA was extended by terminal nucleotidyl transferase (136) in the presence of dATP. In another experiment, the deoxythymidine polynucleotide-cellulose was extended with dCTP. This mixed polynucleotide-cellulose then retained single-stranded DNA which had been extended with dITP.

\textsuperscript{19}The buffer used by Gillespie and Spiegelman was not reported in their abstract (134). However, it is reported to be 2 x SSC, as a personal communication in (131).

\textsuperscript{20}See footnote 22.
The two complexes are illustrated, schematically, below. In I, the maximum interaction

\[
\text{cell-} \overline{\text{p}} \text{T}_p \text{T}_p \text{T}_p \text{T}_p \text{T} \quad \overline{\text{A}_p \text{A}_p \text{A}_p \text{A}_p} \quad \text{-DNA}
\]

\[
\text{cell-} \overline{\text{p}} \text{T}_p \text{T}_p \text{T}_p \text{T}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}_p \text{C}
\quad \overline{\text{p}_p \text{p}_p \text{p}_p \text{p}_p \text{p}_p} \quad \text{-DNA}
\]

is probably \(d(pT)_{12}d(pA)_{12}\), although the majority of the oligothymidylates would be shorter. In II, the length of cytidylic acid residues may be considerably greater (up to 200 residues (135)).

A short, remarkably stable hybrid region (6 GC and 1 AU base pairs) has been isolated from the stem portion of \textit{E. coli} tyrosine tRNA (137). There is also a short hybrid region in the anticodon loop of \textit{E. coli} methionine tRNA\(f\) (4 GC, 1 AU base pairs) (138) as well as a region in yeast aspartic tRNA (5 GC base pairs) (139).

By suitable choice of temperature and Mg\(^{+2}\) concentration, it has been possible to form short hybrid regions between overlapping complementary synthetic oligodeoxynucleotides. These oligomers were subsequently joined, enzymatically, during the synthesis of the yeast alanine tRNA gene (88).
Possibly the best model system to use in studying oligomer: random sequence polymer interactions would be polymers which contained a short sequence complementary to a synthetic oligonucleotide.\(^{21}\)

In the studies reported in this thesis, oligonucleotide: oligonucleotide interactions between the series \(d(pA)_n\) and \(d(pT)_n\) were investigated by studying the thermal dissociation curves for a number of mixtures. Some interactions between the complementary oligonucleotides of mixed, repeating base sequence have also been studied.

The synthesis of polynucleotide-celluloses\(^{22}\) was first reported

---

\(^{21}\) The synthesis of such polymers could be done enzymatically. A sample of synthetic oligonucleotide could act as a primer for the terminal deoxynucleotidyl transferase, described by Bollum (136). The oligonucleotide, extended in the 3'-hydroxyl direction (50 to 100 nucleotide units), could then be copied specifically by \(E. coli\) DNA polymerase, using such conditions as those described by Wu and Kaiser (16) and Wu (17). (A small primer may be needed to initiate the repair reaction.) The complementary polymer could then be extended further in the 3' direction, using terminal deoxynucleotidyl transferase to provide a second "tail". The strand separation would be greatly facilitated by starting with the chemically synthesized oligonucleotide attached to an insoluble support such as cellulose (135).

\(^{22}\) The term "polynucleotide-cellulose" refers to cellulose containing random length oligodeoxyribonucleotides. This term was first used by Gilham (68), and is retained here as a means of easily distinguishing between a cellulose containing oligonucleotides of random length, and one containing oligonucleotides of defined length (and sequence). These latter cellulosics are referred to as "oligonucleotide-celluloses".
by Gilham (68). Random length homo-oligodeoxyribonucleotides of T, A and C were prepared by polymerization of the corresponding mononucleotide in anhydrous pyridine using the condensing agent, DCC. After the polymerization was complete, dry cellulose powder and more DCC was added to the reaction mixture to link the linear oligomers to the cellulose (68, 69). The resulting polynucleotide-celluloses were demonstrated to retain complementary oligonucleotides. Thus, deoxythymidine polynucleotide-cellulose retained a mixture of deox adenylate oligomers (tri- through heptanucleotides). The various oligomers could be separated by stepwise thermal elution. Similarly, deoxyadenosine polynucleotide-cellulose was able to retain and resolve a series of uridine oligonucleotides (69).

Gilham and Robinson (70) have reported the use of deoxythymidine polynucleotide-cellulose in fractionating an enzymic digest of viral RNA. Bromegrass virus RNA was digested with pancreatic ribonuclease and the resultant oligonucleotides separated according to length by anion-exchange chromatography. Oligonucleotides of the same chain length were then chromatographed on a deoxythymidine polynucleotide-cellulose column using stepwise temperature elution. The columns were able to separate oligomers based on their content of consecutive adenosine sequences. However, the columns apparently did not resolve such oligomers as those containing tetra-adenosine sequences.
(i.e. interrupted tetra-adenosine sequences). This observation
may be explained by experiments reported by Bautz and Bautz (125).
These workers studied the differences in stacking energy between
different dinucleotides and concluded that guanine interacts
(stacks) relatively strongly with an adjacent adenine, and
within a tract of adenylic acid residues does not greatly
destabilize the overall interaction. However, in an attempt to
quantitate the effect of mismatching of base pairs, a study by
Bautz and Bautz (125) of poly U. poly (A, G interactions with
varying amounts of guanylate demonstrated a linear relationship
in which 1% mismatching of bases reduced the Tm by about 0.7°C.
This observation is in agreement with that of Kotaka and Baldwin
(140) and Laird et al. (29). According to Laird et al., the
relationship is valid for hybrids in which there is <30% mismatching of bases and the polymers are larger than 50 nucleotides long. If this relationship can be extrapolated to much shorter nucleotides, it is difficult to understand why the interrupted
tetra-adenosine sequence containing oligonucleotides were retained
by the deoxthymidine polynucleotide-cellulose column.

Edmonds and Abrams (71) have reported the use of deoxy-
thymidine polynucleotide-cellulose to isolate a polyadenylate
associated with an adenosine triphosphate polymerase in calf
thymus nuclei, while Edmonds and Caramela (72) have used the
same technique to recover from Ehrlich ascites nuclei the 1% fraction of RNA which is rich in adenosine.

Recently, Gilham (144) has reported an alternate method for the covalent attachment of oligonucleotides to cellulose. This involves incorporation of nucleotides onto cellulose paper strips by specific activation in aqueous solution of a terminal phosphate group using the water-soluble carbodiimide, N-cyclohexyl-N'-β(4-methylmorpholinium)ethylcarbodiimide p-toluenesulfonate. At pH 6.0, it has been shown that this reagent is a specific phosphate activating agent that can effect the efficient formation of phosphodiester bonds, in the presence of a high concentration of hydroxyl group (124). One novel aspect of this reaction is that, in solution, in the presence of cellulose, there is no incorporation of nucleotide onto the cellulose, presumably due to a low concentration of available hydroxyl group. However, when the reaction mixture is streaked on paper, and slowly air dried, the reaction is concentrated on the cellulose fibres. Under these conditions, 71% incorporation of d-pG, 67% incorporation of d-pC, 64% incorporation of UDP, etc. have been reported (141).

One problem associated with the water-soluble carbodiimide method is that the yield of the nucleotide incorporated decreased with increasing length of the oligonucleotide (141). (For example, d(pT)$_4$ was incorporated at a level of 47% (141), while d(pT)$_{12}$ was
incorporated at a level of only 11.8% and 16.7% (two experiments) (142).

In the experiments reported in this thesis, oligonucleotide-celluloses containing oligomers of defined length and sequence were prepared using the water-soluble carbodiimide method.\(^{23}\) The method, as reported by Gilham (141), was altered by using more carbodiimide (in two applications), and by leaving the reaction for a much longer time. Other minor alterations in buffer concentration are described in Methods. The reason for using more carbodiimide, and streaking it on the reaction paper in two application steps was based on data concerning the half-life of this reagent. Metz and Brown (144) studied the stability of the water-soluble carbodiimide by infrared spectroscopy, following the decrease in the 2100 cm\(^{-1}\) peak due to the N=C=N stretching mode. In water (pH not specified), these workers observed a half-life of 240 hr. However, in 0.1 M sodium cacodylate buffer, pH 6.0, the t\(_\frac{1}{2}\) was only 1 hour. Using a different assay, Naylor and Gilham (124) reported a t\(_\frac{1}{2}\) of 6 hour,

\(^{23}\)Alternate methods were tried to link the oligonucleotides to cellulose. For example, the imidazolate of d-pT (143) was reacted with dried cellulose powder or cellulose paper strips under anhydrous conditions. The imidazolate of d-pT was also reacted with dry phosphocellulose powder. However, none of the conditions used resulted in incorporation of nucleotide, and because it was discovered by altering the conditions for the water-soluble carbodiimide reaction, the yields of nucleotide incorporated were consistently greater than 60%, further efforts to find another method for incorporation of oligonucleotide were abandoned.
in water, at pH 6.0. From these data it seemed possible that the stability of the carbodiimide may be affecting the incorporation of oligonucleotide, which proceeds much slower than the incorporation of monomer units.

Since a considerably greater amount of carbodiimide was used to effect the linkage of the oligomer to the cellulose, the possibility of side reactions occurring was investigated, in that the integrity of the functional groups on the bases is essential for hybrid formation to occur.

Using the modified water-soluble carbodiimide procedure, oligonucleotide-cellulose containing the oligodeoxyribonucleotides, $d(pT)_6$, $d(pT)_9$, $d(pT)_9$, and $d(pT)_{12}$ were prepared. These oligonucleotide-celluloses, in the form of small columns, were examined for their ability to retain complementary oligomers. The retained oligonucleotides could be eluted conveniently with a linear temperature gradient.

The capacity of the oligonucleotide-cellulose, the resolution that may be obtained between consecutive deoxyribo-oligomers, and the elution of ribo-oligomers were studied. Several different preparations of the same oligonucleotide-cellulose ($cell-d(pT)_9$) were examined in order to learn something of the reproducibility of retention properties of oligonucleotide-celluloses. Celluloses of mixed, repeating base sequences ($cell-d(pTpTpC)_{2,3}$ and $cell-d(pCpTpT)_{2,3}$) were tested for their
ability to retain the complementary oligonucleotides $d(pApA_pG)_2,3$.

Preliminary experiments, concerning the ability of an oligonucleotide-cellulose column to select a complementary sequence from a mixture of nucleic acids, were also carried out.
MATERIALS AND METHODS

The complementary oligodeoxyribonucleotides $d(pT)_n$, $d(pA)_n$, $d(pTpT)_{2,3,4}$, $d(pCpT)_{2,3,4}$ and $d(pApApG)_{2,3,4}$ were prepared by chemical synthetic methods (see Part I of this thesis). The dephosphorylated oligonucleotides $d(A)_n$ were prepared by alkaline phosphatase digestion of the series $d(pA)_n$. The oligoribonucleotides $r(A)_9$ and $r(A)_6$ were purchased from Miles Laboratories (Elkhart, Indiana). Polyadenylic acid ($S_{20} 7.0 - 11.5$) was also the product of Miles Laboratories.

A. Interaction of Complementary Oligonucleotides in Solution.

1) Estimation of average extinction coefficient per base at $25^\circ$ for oligonucleotides of mixed base sequence. In order to prepare equimolar mixtures (based on mononucleotide concentration) of complementary oligonucleotides, it is necessary to be able to estimate the concentration of these nucleotides. One convenient method is to assay for total phosphate and obtain an $E_{p}^{24}$ value. Alternatively, an estimate of the extinction coefficient may be obtained by degrading the oligomer with a nuclease to the level of mononucleotides (124), thus eliminating the residual hypochromism. By following the increase in absorbance at the wavelength of maximum absorbance, one can

$E_{p}^{24}$ is the molar extinction coefficient per phosphate residue.
estimate the residual hypochromicity and subtract this value from the average theoretical extinction coefficient ($E_{\text{theor}}$). This calculated value is defined as the average extinction coefficient per base ($E_{\text{calc}}$). In the experiments described in this thesis, the concentration of mixed base sequence oligonucleotides was determined spectrophotometrically from $E_{\text{calc}}$ values for the oligomers determined by degradation of a sample of each oligonucleotide with venom phosphodiesterase.

A sample of each oligomer (about 0.7 A.U. for purine oligonucleotides and about 2 A.U. for pyrimidine oligonucleotides) was pre-incubated at 35° in 1 ml of 0.1 M tris acetate, pH 8.0, in a 1 ml cuvette placed in a water-jacketed sample holder of a recording spectrophotometer. An aliquot of venom phosphodiesterase (Worthington, 7.5 mg/ml) was added to the sample cuvette and the change in absorbance at the maximum wavelength recorded. The amount of enzyme added was determined experimentally such that degradation of the oligomer was complete within 1 to 2 hour. The % hypochromicity was calculated using the formula:

\[
\% \text{ Hypochromicity} = \frac{\Delta \text{A.U. } \lambda_{\text{max}}}{\text{initial A.U. } \lambda_{\text{max}} + \Delta \text{A.U. } \lambda_{\text{max}}} \times 100\%
\]

$E_{\text{theor}}$ is the average molar extinction coefficient per base calculated from the molar extinction coefficients of the individual nucleotides at the $\lambda_{\text{max}}$ of the oligonucleotide.
The $E_{\text{calc}}$ was calculated as described above. It should be noted that this value is actually $E_{\text{calc}}^{35^\circ}$ and because most often the extinction coefficient at $25^\circ$ is more useful, the hypochromicity resulting from a change in temperature from $35^\circ$ to $25^\circ$ was estimated and this value used to correct the $E_{\text{calc}}^{35^\circ}$ to the $E_{\text{calc}}^{25^\circ}$.

2) Thermal denaturation curves for complementary interactions of the type $d(pT)_n \cdot d(pA)_n$ and $d(pT)_n \cdot \text{poly} A$.

The interaction between oligonucleotides was studied by following the increase in absorbance at 260 nm of a mixture of two complementary sequences. The method used was similar to that described by Gilham (68) and Naylor and Gilham (124).

Approximately 0.025 mmole of oligonucleotide (based on the mononucleotide unit) were mixed with 0.025 mmole of a complementary oligonucleotide or polynucleotide (total concentration of nucleotide = 0.05 mM). The sample was made up to 1 ml with molar buffered saline, MBS (0.01 M $\text{NaH}_2\text{PO}_4$, pH 7.0, 1 M in NaCl) and stored at $4^\circ$ for 16 to 18 hour. The sample was then transferred to a pre-chilled ($0^\circ$) 1 ml, 1 cm light path, quartz cuvette in the sample holder of a double beam recording spectrophotometer (Unicam Sp.800 with Sp.825 programme controller Sp.850
The temperature in the cuvette was reduced further to approximately \(-7^\circ\) and the sample incubated at this temperature for at least 1 hour. The thermal denaturation profile was then observed by slowly raising the temperature of the sample. This was accomplished by connecting the magnet on the thermoregulator of the Haake circulating bath to the slow shaft of a variable speed motor (Gerald K. Heller, model 2T60-1110 variable speed motor with S-10 motor controller). The motor was set to turn the magnet at a rate which caused a rise in temperature of no greater than \(1.5^\circ\) per 3 min. The spectrophotometer was programmed to record the absorbance of the sample for a five second duration every 3 min, and the record was plotted out on the recorder. The SP.21 recorder was set at a full scale expansion of 0.20 A.U. The temperature increase was continued until an upper "break" in the thermal profile was observed.

26The jacketed cuvette holder of the spectrophotometer was pre-chilled to \(0^\circ\) using a circulating refrigerated bath (Haake Model KT41) filled with either 95% ethanol or 50% ethylene glycol. The blank cuvette contained 1 ml MBS. A temperature probe connected to a tele-thermometer (YSI tele-thermometer model 42 SC with range \(-40^\circ\) to \(150^\circ\)) was placed in a third cuvette filled with MBS. This "temperature cell" was placed in sample slot 2, adjacent to the sample cuvette. (A check showed that the temperature of the sample cuvette was identical to that recorded in the "temperature cell".) Condensation of water vapour on the optical surfaces at less than \(10^\circ\) was a problem. However, it was found that passing a stream of dry nitrogen into the cuvette holder area eliminated this problem, provided the cell compartment door was not opened below \(0^\circ\).
Because of the stacking of linearly arranged purine bases, single stranded polynucleotides and oligonucleotides containing tracts of purines (particularly adenylic acid residues) exhibit a thermal denaturation profile (145). Consequently, the purine containing oligodeoxyribonucleotides as well as poly A were denatured individually. The denaturation curve for a .025 mM solution of purine oligonucleotide was then subtracted from a thermal denaturation curve for a mixture of complementary nucleotides (.050 mM nucleotide) to obtain the true thermal denaturation profile (see Fig. 14A).

3) \( T_m \) and % hypochromicity.

The \( T_m \) is defined as the temperature at the midpoint of the increase in absorbance (at 260 nm) (145) (see Fig. 14B). The % hypochromicity is defined as the per cent increase in absorbance relative to the sum of the absorbance of the individual oligonucleotides at 25\(^\circ\). For interactions between shorter oligonucleotides, especially oligonucleotides of mixed repeating base sequence, which are incomplete at -7\(^\circ\), the % hypochromicity is calculated on the maximum difference in absorbance observed. In these cases the \( T_m \) was estimated, based on an expected maximum hypochromicity of 15% for \( d(pA)_n \cdot d(pT)_n \) interactions and 9% for \( d(pApApG)_n \cdot d(pCpTpT)_n \) (or \( d(pTpTpC)_2 \)) interactions. This \( T_m \) value is recorded with an *, indicating it is an estimated value.
Figure 14. Definition of terms related to thermal denaturation curves.

A. Theoretical thermal denaturation curves for interaction of complementary oligonucleotides. Curve a is the thermal denaturation curve observed spectrophotometrically for the interaction of two complementary oligonucleotides. Curve $b_1$ is the denaturation profile of the pyrimidine oligonucleotide and $b_2$, the purine oligonucleotide. Curve c, the true denaturation profile, is obtained by subtracting curves $b_1$ and $b_2$ from a. Since the increase in absorbance of a pyrimidine oligonucleotide over the temperature range $-5^\circ\text{C}$ to $70^\circ\text{C}$ is negligible, curve c is routinely obtained by subtracting curve $b_2$ from a.

B. Determination of Tm and % hypochromicity (145). The thermal denaturation profile is equivalent to curve c in Figure 14A. The Tm is the temperature at one half the increase in absorbance. The % hypochromicity is calculated from the formula:

$$\frac{A_{260}(M) - A_{260}(B)}{A_{260\text{ individual oligomers at } 25^\circ\text{C}}} \times 100\%$$
B. Synthesis and Properties of Oligonucleotide-Celluloses.

1) Synthesis of oligonucleotide-celluloses. Reaction of oligonucleotides on cellulose paper with a water-soluble carbodiimide.

The procedure used to link oligonucleotides to cellulose paper was slightly modified from that described by Gilham (124). Approximately 50 to 150 A.U. of oligonucleotide (NH$_4^+$ salt converted to Na$^+$ salt on a small Bio-Rad AG 50W column, 0.9 cm x 5 cm) was dried overnight, in vacuo, over P$_2$O$_5$. Four tubes were set up according to the outline below. The contents of each tube were streaked onto a 5 x 12 cm area of pre-washed Whatman 3 MM paper strip.

<table>
<thead>
<tr>
<th>Tube</th>
<th>Oligonucleotide (0.2 M Na$^+$ Mes$^-$)</th>
<th>Buffer (pH 6.0)</th>
<th>Water</th>
<th>Carbodiimide</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50-150 A.U.</td>
<td>0.1 ml</td>
<td>0.4 ml</td>
<td>50 mg</td>
</tr>
<tr>
<td>B</td>
<td>50-150 A.U.</td>
<td>0.1 ml</td>
<td>0.4 ml</td>
<td>--</td>
</tr>
<tr>
<td>C</td>
<td>--</td>
<td>0.1 ml</td>
<td>0.4 ml</td>
<td>50 mg</td>
</tr>
<tr>
<td>D</td>
<td>--</td>
<td>0.1 ml</td>
<td>0.4 ml</td>
<td>--</td>
</tr>
</tbody>
</table>

$^{27}$ The Whatman 3 MM was washed in 0.001 M EDTA (pH 7.0), distilled water, and air-dried. The paper was then soaked in methanol:HCl (99:1 v/v) for 3 days. The paper was washed in excess distilled water until the pH was neutral, and soaked briefly in 0.05 M Na$^+$Mes$^-$, pH 6.0. The paper was again washed in distilled water and air-dried. Initially, the paper was methylated because it is known that cellulose contains a number of free carboxyl groups which can react with a carbodiimide (146). This can rearrange to give a stable acylurea, which in the case of a water-soluble carbodiimide with a quaternary nitrogen group would introduce anion-exchanger properties to the cellulose, and may therefore affect the elution of anionic oligonucleotides (142). Although at high salt concentrations (1 M NaCl) the ionic interference of oligonucleotide elution is unlikely the methylation step was retained because it was felt that the proposed intermediate of carbodiimide and phosphate on the oligonucleotide could react with free carboxyl groups producing linkages of the type

$$\text{cell-O-P-O-CH}_2$$

Oligonucleotides linked to cellulose via a mixed anhydride would be less stable than those linked via a phosphodiester linkage,
The paper strips were hung up to dry at room temperature for 24 to 48 hour. (The papers were not exposed periodically to an atmosphere of water vapor, as described by Gilham (124).) In some cases, a second 50 mg of N-cyclohexyl-N'-β(4-methyl-morpholinium)ethylcarbodiimide p-toluenesulfonate was applied to papers A and C. This was done by dissolving the reagent in 0.1 ml buffer and 0.4 ml water. Papers B and C were streaked with the buffer and water mixture minus the carbodiimide. If two applications of the carbodiimide were done, the second was routinely streaked on 48 hour after the first, and the papers left another 48 hour (i.e. 96 hour total). In some experiments, the amount of buffer was increased to 0.2 ml (0.2 M Na⁺Mes⁻) and the amount of water reduced to 0.3 ml. Finally, in some experiments, because of the limited amount of oligonucleotide available, tube B was omitted. The omission of tube B alters the calculation of the % nucleotide incorporation (see below), and owing to this change, the estimate of nucleotide incorporation is probably up to 10% higher than the true value.

After the reaction had proceeded (for 24 to 96 hour) the paper strips were suspended in a trough in a small chromatography tank (Kontes, thin-layer chromatography tank) and eluted by the descending method with about 10 ml of 0.05 M NaH₂PO₄, pH 7.0. The eluate was collected and made up to 25 ml with the phosphate buffer. The spectra of the four samples were run and the amount
of oligonucleotide incorporated onto the paper calculated using one of the following two formulae.

**Calculation of nucleotide incorporation when tube B is included.**

(A, B, C and D are the absorbance of the 25 ml solutions at the wavelength of maximum absorption of the oligonucleotide.)

\[
\text{A.U. Oligonucleotide incorporated} = \left[ (B-D)_{\text{A.U./ml}} + (C-D)_{\text{A.U./ml}} - (A-D)_{\text{A.U./ml}} \right]_{25 \text{ ml}}
\]

**Calculation of nucleotide incorporation when tube B is omitted.**

(A, C and D are the absorbance of the 25 ml solutions at the wavelength of maximum absorption of the oligonucleotide. B is the absorbance units of oligonucleotide that was placed in tube A, initially.)

\[
\text{A.U. Oligonucleotide incorporated} = B_{\text{A.U.}} + \left[ (C-D)_{\text{A.U./ml}} - (A-D)_{\text{A.U./ml}} \right]_{25 \text{ ml}}
\]

2) **Possible side reactions in the reaction of oligonucleotides with the carbodiimide.**

In experiments to check that the oligonucleotides were not involved in side reactions with the carbodiimide under the conditions of the incorporation experiments, the model compounds d-pT, d-pC and d-pA (approximately 20 μmoles of the Na⁺ salts), were reacted in aqueous solution and on paper, in the presence of Na⁺Mes⁻ buffer.
(pH 6.0) with varying amounts of carbodiimide reagent and for varying lengths of time. (The details of the experimental conditions are given in the Results.) At the end of the incubation, the reaction mixture was diluted to 10 ml with water and the pH adjusted to 7.1 with a dilute solution of triethylamine. The samples were loaded directly onto DEAE-cellulose (acetate) columns (1.2 cm x 40 cm), pre-equilibrated with 0.01 M TEAA, pH 7.1, and eluted with a linear gradient of TEAA, pH 7.1, 0.01 M to 0.15 M. In order to more closely approximate reaction conditions for the nucleotides with the carbodiimide during an incorporation experiment, one experiment was done in which d-pT and carbodiimide were streaked onto Whatman 3 MM in the usual manner (see 1 above). After 24 hour, the papers were eluted with 10 ml phosphate buffer (0.01 M, pH 7.0) and the eluate loaded directly onto a DEAE-cellulose (acetate) column as described above.

3) Preparation of column matrix and procedure for packing column.

The oligonucleotide-cellulose paper strips were prepared for column chromatography by cutting the paper into 1 mm squares, suspending these in 20 ml MBS and stirring vigorously with a magnetic stirrer for 1 to 2 hour. After this time most of the paper was in finely divided fibres with a very few squares of paper remaining.\textsuperscript{28}

\textsuperscript{28}Attempts were made to obtain a more uniform particle size of the cellulose. This was done by suspending the paper in 100 ml MBS and mixing in a Waring Blender. However, when the paper just reached the stage where all squares of paper had been fragmented, the resulting slurry, when packed into a column gave a prohibitively slow flow rate, particularly at temperatures less than 5\textdegree. Therefore, the method described above was used.
The jacketed columns used for the column chromatography were prepared by Vancouver Scientific Glassblowing. These columns (dimensions in Figure 15) were prepared to accommodate the fittings illustrated in the figure. (These fittings were purchased from Hoeffer Scientific Instruments, California.) One advantage of this column is that the temperature jacket extends well beyond both ends of the column matrix and therefore there is no chance of there being a temperature gradient within the column. Also, the dead space is negligible (<200 μl, including the lead to the fraction collector).

The cellulose slurry was packed into the column using a pressure bulb. The matrix was further packed by applying pressure with a glass rod which just fit the inside of the column. Although an estimate of the amount of pressure used to pack the column is difficult to give, it should be said that it was considerably greater than what one ordinarily uses to pack such columns as DEAE-cellulose columns (2 to 5 psi) and that the cellulose from a 5 cm x 12 cm piece of Whatman 3 MM could be packed into a column 0.9 cm x 5 cm. Although the column was packed very tightly, flow rates in excess of 0.5 ml/min could be obtained, even at temperatures as low as -5°.

4) Procedure for running oligonucleotide-cellulose columns.

a) Thermal elution of oligonucleotides. The oligonucleotide-cellulose columns were routinely prepared (and
Figure 15. Dimensions (in centimeters) of jacketed column used for temperature-controlled elution of oligonucleotides from columns of oligonucleotide-cellulose. The fittings illustrated were purchased from Hoeffer Scientific Instruments.
regenerated) by washing them rapidly with at least 3 column volumes of 0.1 N NaOH. The cellulose was then equilibrated with MBS at the temperature that the nucleotides were to be loaded (usually -4° to -5°). The oligonucleotide-celluloses were stable to repeated washing with the dilute alkali, and could be reused many times, with the results for a particular column invariant after several washings and also after storage (as a slurry in MBS at -20°) for several months.

The columns were loaded with an oligonucleotide sample (usually about 5 A.U. in 1 ml MBS) under a slight pressure such that a 1 ml sample was loaded in approximately 15 min. The sample was washed in with 1 ml MBS and the column washed with this buffer at -4° to -5°, at a flow rate of 1.7 ml/min (controlled by using a Buchler peristaltic pump). This initial washing was continued until the absorbance of the effluent dropped to zero. The bound nucleotide was then eluted with a stepwise temperature gradient or linear temperature gradient.\(^{29}\)

\(^{29}\)The column temperature was controlled by using a Haake model KT41 circulating bath. The linear increase in temperature was obtained by attaching the magnet of the thermo-regulator to the slow shaft of a variable speed motor (G.K. Heller, model 2T60-1110). Although the specifications of the motor state that speeds of 0-4000 RPM are obtainable, the motor was unreliable at the necessary slow speed, and frequently stalled. The problem was overcome by connecting the slow shaft of the motor to a gear system (ratio 4.6 to 1) supported in a gear box which was attached to the motor by a bracket. The magnet was then attached to the large gear. In this way, the motor could be set at a speed such that the magnet on the thermo-regulator turned at a slow, reliable, constant rate (1 revolution per 14 min), causing an increase in temperature of 1° every 22 min.
The latter was set up to give an increase of $1^\circ$ per 22 min (i.e. $0.5^\circ$ per 1.7 ml fraction). The choice of this flow rate and temperature gradient was somewhat arbitrary, although it was felt that since the bed volume of the column was approximately 3 ml, and the fraction size 1.7 ml, then every two fractions are equivalent to at least one column volume, and this also corresponds to a change of $1^\circ$ (i.e. this flow rate and temperature gradient result in one bed volume of eluant through the column per increase of $1^\circ$).

b) **Other methods of elution of oligonucleotides.** A number of methods other than thermal elution were used to effect the elution of oligonucleotides from oligonucleotide-cellulose columns. These include reverse salt gradients (high salt to low salt), organic solvents (formamide, dimethylformamide, dimethylsulfoxide), as well as a gradient of sodium perchlorate. The details of these experiments are given in the Results.

5. **Preparation of oligonucleotides of the series**

$d(A)_n$ and $r(A)_n$.

The oligonucleotide series $d(A)_n$ ($n = 6, 7, 8, 9$) was prepared by alkaline phosphatase digestion of $d(pA)_n$. Approximately 1 μmole of nucleotide was incubated in 100 μl of 0.1 M $NH_4HCO_3$, pH 8.0 with 10 μl alkaline phosphatase (Miles, 10 mg/ml, 33U/mg) at $45^\circ$ for 3 hour. The nucleotide was separated from the enzyme by elution on either a small DEAE-cellulose
(CO$_3^{m-}$) column, 1.2 cm x 5 cm (d(A)$_9$ and d(A)$_8$ preparations), with a linear gradient of NH$_4$HCO$_3$, or on a Sephadex G50 column, 0.8 cm x 55 cm (d(A)$_7$ and d(A)$_6$ preparations), with 0.01 M NH$_4$HCO$_3$. In both cases the enzyme activity was assayed using the substrate p-nitrophenylphosphate, and the enzyme was found to be separated from the oligonucleotide.$^{30}$

The oligomer r(A)$_g$ was prepared by periodate oxidation of r(A)$_g$ (147) followed by alkaline phosphatase digestion. The r(A)$_g$ was separated from the enzyme on a DEAE-cellulose (CO$_3^{m-}$) column as described above. This preparation was subjected to another round of periodate oxidation followed by phosphatase, and the resultant r(A)$_7$ was purified on a Sephadex G50 column as described above.

6. Isolation of Salmon Liver RNA.

The procedure for the isolation of RNA from the liver of a Chinook salmon (Oncorhynchus tschawytscha) was based on a method using diethyl pyrocarbonate as an inhibitor of nuclease activity (148). The liver (4.8 g, stored frozen at -20$^\circ$) was broken into small pieces and homogenized at 0$^\circ$ in 12 ml homogenizing medium (0.25 M sucrose, 1.5 mM MgCl$_2$, 2.5 mM tris HCl, pH 7.6; TMS) in a glass-terflon homogenizer. The homogenate was centrifuged at 2200 RPM (SS-34 rotor) at 4$^\circ$ for 10 min.

$^{30}$It was not really necessary to first remove the enzyme from the oligonucleotide prior to elution of the oligomer from an oligonucleotide-cellulose column.
The nuclear pellet was washed with 4 ml TMS and recentrifuged. The pooled supernatants were centrifuged again, at 12,500 RPM (SS-34) (4°) for 15 min to pellet the mitochondria (149). The pellet was washed (4 ml TMS) and the mitochondria re-pelleted. The reddish supernatants were pooled to give a total of 20 ml. Concentrated SDS buffer (2 ml, 11% SDS w/v in 0.5 M tris HCl, pH 7.6, 50 mM MgCl₂) was added to the 20 ml of extract, and the two were mixed thoroughly. Diethylpyrocarbonate (Baycvin) (0.66 ml) was added to bring the NaCl concentration to 1.7 M. The solution was left a further 15 min at 37°, and the heavy precipitate removed by centrifugation (10,000 x g, SS-34 rotor, 4°). Some flocculent precipitate remained and was removed by pipetting the solution through a small pad of glass wool. The filtrate was added to 60 ml of 95% ethanol and stored overnight at -20°. The precipitate was dissolved in 8 ml 50 mM Na₂HPO₄, pH 7.0, 0.4 M NaCl and dialyzed³¹ overnight in the cold room against this buffer (2 x 2 litres). Sufficient NaCl was added to bring the solution to 1 molar, and the RNA was again precipitated with 95% ethanol. The precipitate was redissolved in 8 ml 0.01 M tris HCl, pH 7.4, 5 mM MgCl₂. Deoxyribonuclease (Worthington, ribonuclease free, 0.5 mg/ml) (200 µl) was added and the solution incubated at 35° for 30 min.

³¹The dialysis tubing was heating at 80° in 0.001 M EDTA, pH 7.0, for 1 hour, then washed with distilled water and stored in EDTA, pH 7.0, at 4° (150).
SDS (0.33 ml of a 25% solution w/v), diethylpyrocarbonate (0.27 ml) and solid NaCl (0.86 g) were added and the solution incubated at 37° for 5 min. The precipitate was removed by centrifugation and the RNA precipitated with 95% ethanol (2 volumes). The RNA was redissolved in 10 ml of 0.025 M tris HCl (pH 8.0), 0.01 M MgCl₂ and dialyzed at 4° against this buffer for 36 hour. The yield of RNA was 230 A₂₆₀ units (A₂₈₀/A₂₆₀ ratio = 0.48). High voltage acrylamide gel electrophoresis according to a procedure described by Peacock and Dingman (151) showed this preparation contained material migrating in regions consistent with 4S (and 5S) RNA, the two rRNAs, as well as two larger molecular weight RNAs.

Recently, diethylpyrocarbonate has been shown to react with at least one component of nucleic acids, adenine, forming the 5(4)-N-carbethoxyaminimidazole-4(5)-N'-carbethoxycarboxamidine. This finding may explain the extreme sensitivity of the infectivity of tobacco mosaic virus RNA to this reagent (152). Because of this finding, diethyl pyrocarbonate is no longer the one of choice in the isolation of nucleic acids (at least when biological activity of the molecules is to be retained).
RESULTS AND DISCUSSION

A. Interaction of Complementary Oligonucleotides in Solution.

1) Estimation of the average extinction coefficient per base at 25\(^\circ\) for oligonucleotides of mixed base sequence.

The average extinction coefficient per base at 25\(^\circ\) for the oligonucleotides of mixed base sequence was determined by degrading the oligonucleotides with venom phosphodiesterase as described in Methods. The procedure was checked by degrading a sample of d(pT)\(_6\) (1.91 A.U. at 267 nm). The % hypochromicity observed was 6.0%, from which an \(E_{calc}^{25}\) value of 9,024 was calculated. Naylor and Gilham (124) obtained a value of 9,000.

The data for sample degradations of d(pApApG)\(_2\) (0.72 A.U. at 257 nm) and d(pApApG)\(_3\) (0.725 A.U. at 257 nm) are shown in Figure 16. The amount of enzyme added in each case, 2 \(\mu\)l (7.5 \(\mu\)g), resulted in complete degradation within 30 min for the hexanucleotide and within 45 min for the nonanucleotide.

The data for all the oligonucleotides of mixed base sequence is summarized in Table VI. Because most often the \(E_{calc}^{25}\) is a more useful value, the hypochromicity for the purine containing oligonucleotides (\(E_{25}^{25}\) and \(E_{35}^{25}\) for pyrimidine containing oligomers are the same) was calculated by estimating the hypochromicity due to a 10\(^\circ\) decrease in temperature (35\(^\circ\) to 25\(^\circ\)). This was
Figure 16. Venom phosphodiesterase digestion of $d(pApG)_2$ and $d(pApG)_3$. Details in text.
Table VI. Average molar extinction coefficient per base at 35°

for oligonucleotides of mixed base sequence.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>A.U. (λmax)</th>
<th>Δ A.U. (λmax)</th>
<th>% hypochromicity</th>
<th>E_{35}^{\text{theor}}</th>
<th>E_{25}^{\text{calc}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>d(pT)_6</td>
<td>1.91 (267)</td>
<td>0.122 (267)</td>
<td>6.0</td>
<td>9,600</td>
<td>9,024</td>
</tr>
<tr>
<td>d(pTpTpC)_2</td>
<td>1.42 (267)</td>
<td>0.105 (267)</td>
<td>6.9</td>
<td>9,317</td>
<td>8,670</td>
</tr>
<tr>
<td>d(pTpTpC)_3</td>
<td>1.58 (267)</td>
<td>0.111 (267)</td>
<td>6.5</td>
<td>9,317</td>
<td>8,712</td>
</tr>
<tr>
<td>d(pCpTpT)_2</td>
<td>1.73 (267)</td>
<td>0.125 (267)</td>
<td>6.7</td>
<td>9,317</td>
<td>8,697</td>
</tr>
<tr>
<td>d(pCpTpT)_3</td>
<td>1.435 (267)</td>
<td>0.100 (267)</td>
<td>6.5</td>
<td>9,317</td>
<td>8,712</td>
</tr>
<tr>
<td>d(pCpTpT)_3^a</td>
<td>1.745 (267)</td>
<td>0.120 (267)</td>
<td>6.5</td>
<td>9,317</td>
<td>8,712</td>
</tr>
<tr>
<td>d(pApApG)_2</td>
<td>1.84 (257)</td>
<td>0.485 (257)</td>
<td>20.8</td>
<td>14,500</td>
<td>11,480</td>
</tr>
<tr>
<td>d(pApApG)_2^a</td>
<td>0.72 (257)</td>
<td>0.212 (257)</td>
<td>22.8</td>
<td>14,500</td>
<td>11,200</td>
</tr>
<tr>
<td>d(pApApG)_3</td>
<td>0.725 (257)</td>
<td>0.231 (257)</td>
<td>23.0</td>
<td>14,500</td>
<td>11,160</td>
</tr>
<tr>
<td>d(pApApG)_3^a</td>
<td>0.78 (257)</td>
<td>0.213 (257)</td>
<td>21.2</td>
<td>14,500</td>
<td>11,420</td>
</tr>
</tbody>
</table>

^aThis sample is a repeat of the one immediately above it in the table.
approximately 2.7%.

A value of 8,700 for the average molar extinction coefficient (per base) for the oligonucleotides, d(pTpTpC)_n and d(pCpTpT)_n (n = 2, 3), has been used. For d(pApApG)_n oligomers the average molar extinction coefficient (per base) obtained at 35° was 11,315. The corresponding value at 25° is 11,020. For spectrophotometric determination of nucleotide concentration, a value of 11,000 was used.

The $E_{25^\circ}^{\text{calc}}$ values for oligodeoxythymidylates (9,000) and for oligoriboadenylates (11,000) have been previously determined (124, 153). The value for oligodeoxyadenylates was assumed to be 11,000.

2) Thermal denaturation curves for complementary oligonucleotide interactions and Poly A:d(pT)_n interactions.

a) Typical thermal denaturation curves. Typical thermal denaturation curves for an oligonucleotide:oligonucleotide interaction (d(pT)_8·d(pA)_8) is shown in Figure 17B. For comparison, Figure 17A shows the oligonucleotide:polynucleotide interaction (d(pT)_8·polyA). In each case the purine containing component has been melted separately (curve b) and subtracted from the observed profile (curve a) to obtain curve c, the true thermal denaturation profile. For the interaction d(pT)_8·d(pA)_8 in MBS, the Tm is 15.0° with a % hypochromicity of 14%. For the interaction d(pT)_8·polyA in MBS, the Tm is 24.0°, and the calculated
Figure 17. Thermal denaturation profiles for the interactions $d(pT)_8\cdot polyA$ (A) and $d(pT)_8\cdot d(pA)_8$ (B).

- • - observed thermal dissociation profile for the mixtures.

- o - observed thermal dissociation profile for the purine component.

- - - true thermal dissociation profile for the mixture.
hypochromicity, 22%.

b) Width of the thermal transition. The oligonucleotide interactions (as well as the oligonucleotide:polynucleotide interactions) all melt over a wide temperature range (approximately 30°). This is much wider than for naturally occurring DNA molecules (range of melting approximately 10° in SSC (154)). According to Szybalski (145), the width of the thermal transition in these cases is a result of the "inhomogeneities of the linear sequence of base pairs or clusters of alternating base pairs" (i.e., A,T rich regions melt out at lower temperatures, the basis behind denaturation mapping (155, 156). The breadth of the denaturation curve is also a function of the ionic strength and solvent. For interactions of the type d(pT)\textsubscript{n} 'oligo d(pA)\textsubscript{n}, as well as d(pT)\textsubscript{n} 'polyA interactions, there can be no "inhomogeneity in linear sequence", but in these cases the breadth of the curves must be related to the length of the oligonucleotide. Naylor and Gilham (124) observed approximately a 30° range of melting for the interaction d(pT)\textsubscript{12} d(pA)\textsubscript{6}, as well as for d(pT)\textsubscript{6} 'polyA in MBS. Cassani and Bollum (123) also observed a relatively wide range of melting\textsuperscript{33} for oligonucleotide:polynucleotide interactions. (The wide range of melting was for duplex interactions only. The triplex interactions for oligo dA:2poly dT were much sharper.

\textsuperscript{33}The buffers used were 1) SSC; 2) 40 mM potassium phosphate, pH 7.0, and 3) 40 mM potassium phosphate, pH 7.0, containing 8 mM MgCl\textsubscript{2}.}
with a transition range of approximately 5° to 10°.)

c) % Hypochromicity. The % hypochromicity of each interaction was estimated as described in Methods. The % hypochromicity for \( d(pT) \_n \cdot d(pA) \_n \) interaction which were complete within the temperature range studied was 13 to 16% (see Table VII). These values agree with those reported by Naylor and Gilham (124). The only minor discrepancies were for \( d(pT) \_g \cdot d(pA) \_4 \) (14%) and \( d(pT) \_9 \cdot d(pA) \_4 \) (13%), indicating complete interaction. Naylor and Gilham (124) reported an incomplete interaction, 3% and 5% respectively. However, in their procedure, they were studying the % hypochromicity on reducing the temperature to 0°, while in these experiments, the temperature has been routinely decreased to about -7°, permitting more complete interactions for the less stable hybrids.

d) Summary of data on Tm curves for various interactions of complementary oligonucleotides and polynucleotides. In Table VII a summary of the data for the various homo-oligonucleotide interactions studied, is given. The Tm and % hypochromicities were obtained as described in Methods.

e) Thermal denaturation curves for longer oligonucleotides. In interactions between the series \( d(pT) \_n \) and \( d(pA) \_n \), as the length of the oligonucleotides increased, the slope of the transition noticeably decreased, which is inconsistent with what one would expect for longer, completely homogeneous interactions. Also,
Table VII. Tm values (°) and percent hypochromicity for
\( d(pT)_n:d(pA)_n \) and \( d(pT)_n:polyA \) interactions. The
bracketed number is the % hypochromicity.

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>0.0 (14)</td>
<td>3.5 (13)</td>
<td>7.0 (12)</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>4.5 (16)</td>
<td>8.5 (14)</td>
<td>16.0 (14)</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>4.0 (15)</td>
<td>8.5 (16)</td>
<td>11.5 (15)</td>
<td>19.0 (14)</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>10.5 (13)</td>
<td>15.0 (14)</td>
<td>17.0 (14)</td>
<td>21.5 (16)</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>13.5 (13)</td>
<td>18.0 (15)</td>
<td>24.5 (15)</td>
<td>24.0 (15)</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>18.5 (15)</td>
<td>24.0 (14)</td>
<td>29.0 (14)</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>19.5 (14)</td>
<td>26.5 (15)</td>
<td>31.0 (15)</td>
</tr>
</tbody>
</table>

Poly A  6.5 (18) | 16.0 (21) | 24.0 (22) | 28.5 (20) | -
in some of these interactions, a two-step melting curve was observed. Both these observations suggested that the longer oligonucleotides were less pure than one hoped (although the two-step melting curve could suggest the presence of other than double-stranded structures (123)).

If these longer oligonucleotides are less pure, in order to purify them, it is necessary to consider the nature of the contaminants. In chemical polymerization reactions, the major impurities are cyclic phosphates (157) (up to the cyclic penta-nucleotide, at which point these components become negligible) and pyrophosphates (114). At the end of a DCC polymerization reaction, the sample is rendered anhydrous and treated with excess acetic anhydride in pyridine to convert the pyrophosphates to mixed anhydrides which will break down in aqueous solution. If pyrophosphates are contaminating the longer oligonucleotides then a purification step which would take advantage of the abrupt change in polarity of the molecule at the pyrophosphate linkage would be the best purification procedure. Anion exchange chromatography has already not resolved these impurities. The contaminating pyrophosphate in a linear oligonucleotide peak is most likely the next highest oligomer in length, and, therefore, would have the same net charge (e.g. d(pA)$_{11}$ and pA$_3$(pA)$_7$ both have a net charge of minus 12 at neutral pH). At the decanucleotide level and longer, extended paper chromatography in system C (91)
would also probably give very poor resolution between the linear oligonucleotides and the pyrophosphates.

Consequently, it was decided to try to purify a sample of d(pA)$^{11}$ on an oligonucleotide-cellulose column containing covalently attached d(pT)$^{9}$ (see Part II, B). This column would be expected, by careful choice of temperature elution steps, to resolve all pyrophosphates except possibly the 9:3, 10:2 and 11:1 structures. The nucleotide retained at 29° but eluted at 50° from a cell-d(pT)$^{9}$ column (see inset Figure 18B) was mixed with an equimolar amount of d(pT)$^{9}$, and d(pT)$^{12}$, and the thermal denaturation profiles of these mixtures studied (Fig. 18A to D). Figure 18A and C are the Tm curves for the two interactions, d(pA)$^{11}$·d(pT)$^{9}$ and d(pA)$^{11}$·d(pT)$^{12}$ prior to the purification of the d(pA)$^{11}$. Figure 18B and D are the same interactions after the d(pA)$^{11}$ had been purified on a cell-d(pT)$^{9}$ column. The increased slope of the thermal transition, complete absence of the two-step melting curve, and increase in Tm, all suggest that the original sample of d(pA)$^{11}$ was impure prior to its chromatography on the oligonucleotide-cellulose column.

Considering the basis for the purification of the d(pA)$^{11}$ sample,  

\[ \text{<pA(pA)$^8$; the 10:2 pyrophosphate is:} \]

\[ \text{<pA(pA)$^{10}$,} \]

\[ \text{<pA(pA)$^9$; the 11:1 pyrophosphate is} \]

\[ \text{<pA(pA)_{10}} \]
Figure 18. Thermal denaturation profiles for the interactions d(pA)$_{11}$·d(pT)$_9$ and d(pA)$_{11}$·d(pT)$_{12}$.

A. d(pA)$_{11}$·d(pT)$_9$ prior to purification of d(pA)$_{11}$ on a cell-d(pT)$_9$ column.

B. d(pA)$_{11}$·d(pT)$_9$ after purification of d(pA)$_{11}$ on a cell-d(pT)$_9$ column.

C. dpA$_{11}$·d(pT)$_{12}$ prior to purification of d(pA)$_{11}$ on a cell-d(pT)$_9$ column.

D. d(pA)$_{11}$·d(pT)$_{12}$ after purification of d(pA)$_{11}$ on a cell-d(pT)$_9$ column.

The inset in Figure 18B is the elution profile for the stepwise thermal elution of d(pA)$_{11}$ on a cell-d(pT)$_9$ column.

- - observed thermal dissociation profile of the mixture.

- - - observed thermal dissociation profile of the purine oligomer.

- - - - true thermal dissociation profile for the interaction.
it seems quite likely that the majority of the impurities were pyrophosphates.

f) **Thermal denaturation curves for oligonucleotides of mixed, repeating base sequence (Table VIII).**

In a study of the interactions between some of the complementary oligodeoxyribonucleotides d(pApApG)\textsubscript{n}, and d(pTpTpC)\textsubscript{n} and d(pCpTpT)\textsubscript{n} (n = 2, 3, 4) the following two generalizations can be made.

(i) The percent hypochromicity is significantly lower (7 to 9% for a complete interaction) than for the homo-oligonucleotide interactions (13 to 16%) at the same nucleotide concentrations.

(ii) Although these mixed deoxyoligomers are capable of forming GC base pairs, which are known to be more stable than AT base pairs (145), for an analogous length of oligo dA'oligo dT interaction, the Tm values are not a great deal higher (e.g. d(pCpTpT)\textsubscript{3}:d(pApApG)\textsubscript{3} had a Tm of 31°, \textsuperscript{35} while d(pT)\textsubscript{9}:d(pA)\textsubscript{9} had a Tm of 24.5°), and in some cases apparently lower (e.g. d(pCpTpT)\textsubscript{3}:d(pApApG)\textsubscript{2} had a Tm of between -5° to 0°, while d(pT)\textsubscript{9}:d(pA)\textsubscript{6} melted at 8.5°).

\textsuperscript{35}The value 31° does not agree with the value (25°) in Table VIII. The explanation is (analogous to the situation with the longer oligodeoxyadenylate) the d(pApApG)\textsubscript{3} sample was contaminated, probably with the pyrophosphate, and possibly by as much as 30%. When the d(pApApG)\textsubscript{3} was purified on an oligonucleotide-cellulose column, and this interaction restudied, the higher Tm value of 31° was obtained (see Figure 33D).
Table VIII. Thermal stability of complementary oligonucleotides of mixed base sequence.

All interactions were studied in MBS, unless otherwise noted.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>d(pTpTpC)$_2$</td>
<td></td>
<td>-no interaction</td>
<td>-incomplete</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-&lt;2% hypochromicity</td>
<td>-Tm* &lt; -5°</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-in-7° to 20° range</td>
<td>-hypochromicity</td>
<td>4.7%</td>
</tr>
<tr>
<td>d(pTpTpC)$_3$</td>
<td></td>
<td>-incomplete</td>
<td>-Tm 20.5°</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Tm* 0° to 5°</td>
<td>-hypochromicity</td>
<td>7.4%</td>
</tr>
<tr>
<td>d(pTpTpC)$_4$</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-no Tm curve, but 14% hypochromicity over the range -5° to 70°</td>
</tr>
<tr>
<td>d(pTpTpC)$_4$ in 1/10 x MBS</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-normal Tm curve -Tm 14.5° -hypochromicity 10.2%</td>
</tr>
<tr>
<td>d(pCpTpT)$_2$</td>
<td></td>
<td>-</td>
<td>-incomplete</td>
<td>-incomplete</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Tm* &lt; -5°</td>
<td>-Tm ~5°</td>
<td>-hypochromicity 4.5% - hypochromicity 6.3%</td>
</tr>
<tr>
<td>d(pCpTpT)$_3$</td>
<td></td>
<td>-incomplete</td>
<td>-Tm 25.0°</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Tm* -5° to 0°</td>
<td>-hypochromicity</td>
<td>8.7%</td>
</tr>
</tbody>
</table>

* Indicates that the Tm value has been estimated.
Beyond expecting GC containing complementary deoxyoligomers to be more stable than homo-oligodeoxyribonucleotides of A and T, it is difficult to predict how much more stable these base substitutions would make the interactions. For high molecular weight polydeoxyribonucleotides, there is a 1° increase in Tm per 2.5% increase in GC content (158). If this relationship is valid for the interactions of the deoxyoligomers studied here, a 33 1/3% increase in GC content should result in a 13° increase in the Tm. For a longer interaction (9 base pairs), d(pA)⁹·d(pT)⁹ had a Tm of 24.5°, while d(pApApG)₃·d(pCpTpT)₃ had a Tm of 31° (a difference of 6.5°). This is a few degrees less than expected.

In Table VIII it is noted that d(pTpTpC)₂·d(pApApG)₂ showed no interaction. However, d(pTpTpC)₂·d(pApApG)₃ resulted in a partial interaction within the temperature range studied. The dodecanucleotides, d(pTpTpC)₄ and d(pApApG)₄, did not give a normal thermal denaturation profile (in MBS) over the range -5° to 70°, although 14% hypochromicity was observed. In 1/10 x MBS, a normal curve was observed, with Tm 14.5°. For the two interactions, d(pTpTpC)₃·d(pApApG)₃ and d(pCpTpT)₃·d(pApApG)₃, the Tm values were 20.5° and 25.0°, respectively. However, as shown in Figure 33C and D, when the purine oligonucleotide was purified on a cell-d(pCpTpT)₃ column, and these interactions restudied, the Tm of both interactions increased by 6°.
g) Summary of data from the thermal denaturation studies of complementary oligonucleotides.

For interactions between oligodeoxythymidylates and oligodeoxyadenylates, the data in Table VII indicate that the thermal stability of the interactions is dependent on the length of both oligomers. Within a given series (for example, d(pT)$_{12}$·d(pA)$_n$), there is an increase of approximately 3° to 4° for each additional base pair. However, due to the presence of impurities, particularly in the longer deoxyadenylates, which result in Tm values at least 1° to 2° lower than the true value (see section 2e and Figure 18, above), it does not seem wise to derive a relationship between oligomer length and thermal denaturation temperature, until these longer oligomers have been purified on complementary oligonucleotide-celluloses (see inset, Figure 18) and the particular interactions re-examined.

The width of the thermal transitions (about 30°) indicates that there is considerable overlap of the melting ranges for successive members of a series. This would suggest that oligonucleotide-celluloses may not be capable of resolving consecutive oligomers. However, this was not the case (see Part B, section 2, d, below).

The % hypochromicity for oligodeoxythymidylate:oligodeoxyadenylate interactions is about 15%. For polyA:oligodeoxythymidylate interactions it is about 22%. Both of these values
are in agreement with the data of Naylor and Gilham (124).

From the data presented here on the interaction of deoxy-oligomers of mixed base sequence, it is not possible to draw a strong conclusion as to the effect of GC base pairs on thermal stability. In the case of the hexanucleotide:nonanucleotide interactions (Table VIII), for reasons that are not clear, these interactions were less stable than the corresponding homooligodeoxyribonucleotide interactions (Table VII). However, \(d(pCpTpT)_3 \cdot d(pApApG)_3\) \(T_m = 31.0^\circ\), see Figure 33) melts significantly higher than \(d(pT)_9 \cdot d(pA)_9\) \(T_m = 24.5^\circ\). Also, even when one overlooks the possible destabilization effect of an unpaired terminal nucleotide, \(d(pTpTpC)_3 \cdot d(pApApG)_3\) \(T_m = 26.5^\circ\), see Figure 33) melts considerably higher than \(d(pT)_8 \cdot d(pA)_8\) \(T_m = 15.0^\circ\).

B. Synthesis and Properties of Oligonucleotide-Celluloses.

1) Synthesis of oligonucleotide-celluloses.

   a) Incorporation of oligonucleotide by the water-soluble carbodiimide method. The incorporation of oligonucleotides onto cellulose paper was accomplished using the water-soluble carbodiimide \(N\)-cyclohexyl-\(N^\prime\)-β(4-methylmorpholinium)-ethyl carbodiimide p-toluenesulfonate (CMC-OTS). The products of this reaction are considered to consist of oligonucleotides connected to the cellulose by ester linkages between the terminal
phosphate group of the nucleotide chains and the sterically favoured hydroxyl groups of the cellulose (141).

The results for the incorporation of a number of oligonucleotides and the conditions for the particular reaction are given in Table IX. It should be noted that by increasing the amount of carbodiimide (in two applications) and the time of the reaction, incorporation of nucleotides was consistently greater than 60% \(^\text{36}\), even for the longer oligonucleotides (cf. Gilham (141) reported a decreasing yield of incorporation as the length of the oligonucleotide increased.) A procedure which gives high levels of incorporation is extremely valuable in the synthesis of oligonucleotide-celluloses, particularly for the longer, more complex oligomers, because these are considerably more difficult to prepare, and the yields obtained in chemical syntheses are low.

b) Examination of the CMC-OTs reaction. Because in the incorporation of oligonucleotides onto cellulose paper, much higher levels of carbodiimide and longer times have been used than those reported by Gilham (141), it was considered important to re-examine this reaction for possible byproducts.

At pH 8.0, CMC-OTs has been reported to substitute uracil,

\(^{36}\)When the tube B is omitted, the \% incorporation, marked with a *, is probably an over-estimate by possibly as much as 10% (see Methods).
<table>
<thead>
<tr>
<th>Prep #</th>
<th>Oligonucleotide</th>
<th>Carbodiimide amount&lt;sup&gt;a&lt;/sup&gt; (mg)</th>
<th>Total incorp. time&lt;sup&gt;b&lt;/sup&gt; (hr)</th>
<th>Calc. incorp. A.U.&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Incorp. %&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>d(pT)&lt;sub&gt;6&lt;/sub&gt;</td>
<td>(56.5) 266 nm</td>
<td>50</td>
<td>26</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>d(pT)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>(47) 266 nm</td>
<td>50</td>
<td>26</td>
<td>6.0</td>
</tr>
<tr>
<td>3</td>
<td>d(pT)&lt;sub&gt;8&lt;/sub&gt;</td>
<td>(48) 266 nm</td>
<td>50</td>
<td>48</td>
<td>27.0</td>
</tr>
<tr>
<td>4</td>
<td>d(pT)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>(78) 266 nm</td>
<td>50</td>
<td>24</td>
<td>39</td>
</tr>
<tr>
<td>5</td>
<td>d(pTpTpC)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(90) 267 nm</td>
<td>50</td>
<td>64</td>
<td>24.5</td>
</tr>
<tr>
<td>6</td>
<td>d(pTpTpC)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(124) 267 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>92.5*</td>
</tr>
<tr>
<td>7</td>
<td>d(pTpTpC)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(124) 267 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>96.5*</td>
</tr>
<tr>
<td>8</td>
<td>d(pTpTpC)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(124) 267 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>92.5*</td>
</tr>
<tr>
<td>9</td>
<td>d(pT)&lt;sub&gt;12&lt;/sub&gt;</td>
<td>(47.5) 266 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>30 *</td>
</tr>
<tr>
<td>10</td>
<td>d(pT)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>(49.6) 266 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>48.4*</td>
</tr>
<tr>
<td>11</td>
<td>d(pTpTpC)&lt;sub&gt;3&lt;/sub&gt;</td>
<td>(104) 267 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>77.0*</td>
</tr>
<tr>
<td>12</td>
<td>d(pCpTpT)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(133) 267 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>96.4*</td>
</tr>
<tr>
<td>13</td>
<td>d(pCpTpT)&lt;sub&gt;2&lt;/sub&gt;</td>
<td>(132) 267 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>102.8*</td>
</tr>
<tr>
<td>14</td>
<td>d(pT)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>(72) 266 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>69.3*</td>
</tr>
<tr>
<td>15</td>
<td>d(pT)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>(48) 266 nm</td>
<td>2 x 50</td>
<td>96</td>
<td>47.0*</td>
</tr>
</tbody>
</table>

<sup>a</sup>Indicates experiments in which tube B was omitted.

<sup>b</sup>Amount of carbodiimide - amount of CMC-OTs added/tube. 2 x 50 mg means 50 mg was added at start, and at 48 hours, a further 50 mg was streaked onto the paper according to the procedure described in Methods.

<sup>c</sup>Total time of incorporation - time from first streaking of paper till time of elution of non-incorporated material.

<sup>d</sup>Calculated incorporation - see formulae I and II in Methods.

<sup>e</sup>Incorporated - A.U. incorporated expressed as % of oligonucleotide streaked onto paper.
thymine and guanine nucleotides (159, 160), as well as pseudo-uridine (161). The structure of the addition product for d-pT, CMCd-pT, is shown below. An analogous structure has been suggested for the uracil addition product, while the purine ring of the guanine nucleotide is thought to be substituted at the N-1 position (160). These nucleotides all have pK values of about 9, and based on the observation that the rate of substitution goes up as the pH increases, and as the pK decreases, it has been suggested that the reacting species is the anion of the pyrimidine or purine base (160).
However, below pH 7.0, this reagent has been reported not to participate in a base addition reaction (Ho and Gilham (160)), but only to activate the free phosphate group. The reagent has been reported to effect the synthesis of the four ribonucleoside 2',3' cyclic phosphates in water at pH 6.0, starting with the 2'(3') monophosphates (124). Naylor and Gilham (124) have also reported the use of this reagent under aqueous conditions at pH 6.0 to link together two deoxythymidylate hexanucleotides (hybridized to poly A to provide the correct orientation of reacting groups). This procedure may well have been of considerable importance in the synthesis of large deoxyribonucleotide oligomers, had not the enzyme polynucleotide ligase been discovered (162-166).

If with higher levels of CMC-OTs, the reagent did substitute deoxythymidylate residues at the N-3 position, this would interfere with the ability of the thymidylate residue to hydrogen bond with an adenylate residue (167). Consequently, the reaction of CMC-OTs with nucleotides has been carefully studied. In Figure 19A, B and C, the results of the reaction of d-pT with CMC-OTs, in solution, at pH 8.0 (borate buffer) for 7 hr (A), pH 6.0 (Na⁺Mes⁻ buffer) for 7 hr (B), and at pH 6.0 (Na⁺Mes⁻ buffer) for 24 hr (C) are given. In Figure 19B and C, a small peak of UV-absorbing material (about 1.5%) elutes at the position of CMCd-pT (peak b in these graphs). This peak from a 20-fold, large scale preparation was pooled, concentrated and characterized.
Figure 19. Reaction of d-pT with CMC-OTs.

A. 20 μmoles of d-pT in 2 ml 0.2 M sodium borate, pH 8.0, 7 hr (68 mg CMC-OTs).
B. 20 μmoles of d-pT in 2 ml 0.1 M Na^+Mes^-, pH 6.0, 7 hr (68 mg CMC-OTs).
C. 20 μmoles of d-pT in 2 ml 0.1 M Na^+Mes^-, pH 6.0, 24 hr (68 mg CMC-OTs).
D. Eluate from reaction paper A.
E. Eluate from reaction paper B.
F. Eluate from reaction paper C.

In this figure, the position of elution of the urea and any unreacted carbodiimide is indicated by a, CMC-d-pT by b, p-toluenesulfonate by c, and unreacted d-pT by d. Shaded area is A_{254nm} X 10.
Chromatographic, high voltage paper electrophoresis, and spectral data are in agreement with this peak being CMCd-pT (see Table X and Figure 20).

In Figure 19D, E and F, the eluates from a typical paper incorporation experiment (tubes/papers A, B and C) have been chromatographed. This was done to see if some side reaction occurred unique to the paper reaction, which did not otherwise occur in solution, at pH 6.0. It is clear that no significant side reaction has occurred unique to the paper reaction.\(^{37}\) \(^{38}\)

The reaction of d-pT (Fig. 21A), d-pC (Fig. 21B) and d-pA (Fig. 21C) with CMC-OTs, at pH 6.0, was studied using a higher level of carbodiimide (100 mg) and incubating for a longer period (96 hr). Under these conditions, CMCd-pT (peak b) was produced (2\% of total nucleotide) while no corresponding derivatives for d-pC or d-pA were observed. However, what was most noticeable was the appearance of a second side product (peak x). For the

\(^{37}\)The d-pT (peak d in Figure 19D and E) appears to contain two components. However, by careful checking it was found that this split peak was the result of loading the sample in a mixture of Na\(^+\)Mes\(^-\) and phosphate buffers onto the DEAE-cellulose column equilibrated with dilute TEAA buffer (i.e. some unexplained \"salt effect\").

\(^{38}\)Although no byproducts have been produced, unique to the paper reaction, the significance of a 2\% level of substitution of d-pT residues is discussed more fully in section 2f, below.
Table X. Characterization of byproducts produced on reaction of d-pT with CMC-OTs at pH 6.0.

A. CMC d-pT (peak b)

<table>
<thead>
<tr>
<th>Chromatography system</th>
<th>$R_f$ (b)</th>
<th>$R_f$ (standard CMCd-pT)</th>
<th>$R_f$ (d-pT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.71</td>
<td>0.72</td>
<td>0.47</td>
</tr>
<tr>
<td>B</td>
<td>0.90</td>
<td>0.90</td>
<td>0.47</td>
</tr>
</tbody>
</table>

| High voltage electrophoresis$^a$ | 3.9 cm | 4.2 cm | 16.3 cm |

B. Compound x

<table>
<thead>
<tr>
<th>Chromatography system</th>
<th>$R_f$ (x)</th>
<th>$R_f$ (x) after alkaline phosphatase</th>
<th>$R_f$ (d-pT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.82</td>
<td>0.81</td>
<td>0.44</td>
</tr>
<tr>
<td>B</td>
<td>0.88</td>
<td>-</td>
<td>0.50</td>
</tr>
<tr>
<td>C (two spots)</td>
<td>0.30 and 0.86</td>
<td>-</td>
<td>0.52</td>
</tr>
<tr>
<td>D (two spots)</td>
<td>0.06 and 0.68</td>
<td>-</td>
<td>0.12</td>
</tr>
</tbody>
</table>

| High voltage electrophoresis | 11.6 cm | 11.7 cm | 20.8 cm |

$^a$High voltage electrophoresis was carried out in 0.05 M NH$_4$HCO$_3$, pH 8.0, using Whatman 40 paper. The samples were run for 40 to 60 min at 3000 volts (120 volts/cm), 49 milliamps.
Figure 20. Spectrum of peak b (CMCd-pT).
Figure 21. Reaction of d-pT, d-pC and d-pA with 100 mg (total) of CMC-OTs, at pH 6.0.

20 μmoles of nucleotide in 2 ml 0.1 M Na⁺Mes⁻, pH 6.0 (plus 1 drop of butanol) was incubated at 30° in the presence of 50 mg CMC-OTs. After 48 hr, a further 50 mg of reagent (CMC-OTs) was added, and the incubation continued for another 48 hr. The nucleotides were chromatographed on DEAE-cellulose as described in Methods. Shaded area is A 254 nm X 10^7V.

A, d-pT; B, d-pC; C, d-pA.
In the d-pT reaction, this compound was completely resolved from the p-toluenesulfonate (peak c) (Fig. 21A). In the case of d-pC, x is a very small peak (0.5%) (Fig. 21B), while for the d-pA reaction, it is probably at least 2% of the total nucleotide.

Peak x from the d-pT reaction was concentrated and characterized by chromatographic and electrophoretic data (Table X). The spectrum of this compound in water at pH 7.0, and in 0.1 N HCl and 0.1 N NaOH was similar to d-pT. Although identity of compound x is unknown, it is known that the compound is not degraded by alkaline phosphatase digestion (Table X), suggesting substitution of the 5'-phosphate. (The compound was not cyclic deoxythymidine 3'-5'-monophosphate.) Naylor and Gilham (124) have reported that reaction of 25 μmoles of d-pA with 0.43 μmoles ethanol in 300 μl water (pH 6.0, 25°, 24 hr) gave 20% 5'-ethyl phosphate. It is possible that compound x is the analogous, 5'-butyl phosphate (a drop of butanol was added to each tube to prevent bacterial growth during the 96 hr incubation at 30°, see Figure 21). This is supported by the observation that x is converted, quantitatively, to a product with $R_f = 0.49$ (in system A) after digestion with an aliquot of venom phosphodiesterase.

c) Summary of the synthesis of oligonucleotide-celluloses. In the experiments described in

---

40 I have assumed that the peaks "x" in Figure 21A to C are analogous compounds.
section 1a and b above, it is apparent that oligonucleotides bearing a terminal phosphate group may be linked, conveniently, and efficiently, to cellulose paper using a procedure modified slightly from that reported by Gilham (141). When the amount of condensing agent (water-soluble carbodiimide) is doubled, and streaked onto the paper in two applications, at time zero and at 48 hours, as described in section 1a above, the incorporation of oligonucleotide is consistently greater than 60%. Also, as far as can be detected, only one minor byproduct, CMCd-pT, is produced under the conditions of the reaction. The byproduct appears at a level of 2% of the total nucleotide, which means that approximately 17% (168) of d(pT)$_9$ oligomers, linked to cellulose by this method, will be substituted with the carbodiimide. However, this is not a serious problem as these derivatives are readily broken down at a pH greater than 10.5 (160).

2) Oligonucleotide-cellulose columns.

a) Thermal elution of complementary oligonucleotides.

(i) Stepwise thermal elution. Gilham (69) has reported the elution of oligodeoxyadenylates (tri- through heptanucleotides) from a thymidine polynucleotide-cellulose column,\(^4\) using stepwise

\(^4\)This cellulose is prepared by the polymerization of d-pT in anhydrous pyridine, using DCC as condensing agent. After the polymerization is complete, dry cellulose powder and more DCC is added, and the oligonucleotides are covalently attached to the cellulose. The cellulose therefore contains oligodeoxythymidylates of random length. The deoxythymidine polynucleotide-celluloses, prepared by Gilham, contain about 1.2 mmoles of total nucleotide (68, 69). If one estimates that in a normal polymerization reaction, 7% of the nucleotide is recovered as the octamer and higher oligomers (111), then this cellulose contains about 80 mmoles of oligonucleotides.
thermal elution. According to this report, the successive oligomers are eluted at the following temperatures: tri (5°); tetra (15°); penta (25°); hexa (30°); and hepta (35°) (i.e. the various oligonucleotides elute within a 5° to 10° temperature range.

When an oligonucleotide-cellulose, cell-d(pT)$_8$ (27 A.U. d(pT)$_8$, 0.37 μmoles of oligonucleotide) was loaded at 0° with the complementary oligodeoxyadenylate, d(pA)$_8$, and eluted with a stepwise temperature gradient (Fig. 22A), the d(pA)$_8$ appeared to be sub-fractionated into several components. The d(pA)$_8$ sample was re-checked for purity by chromatography in system C, and found to be >95% pure. Therefore, it would seem that this oligonucleotide-cellulose column releases d(pA)$_8$ over a wide temperature range (30°), which, from the Tm data, is not entirely unexpected. It is also possible that the column was overloaded, and that some of the oligodeoxyadenylate elutes at each temperature step because, as the temperature is increased, the restrictions on the hybrid structures become more stringent. However, this possibility was ruled out (see below, Studies on the capacity of oligonucleotide-celluloses).

(ii) Linear temperature gradient elution. When the same oligodeoxyadenylate was eluted from cell-d(pT)$_8$ with a continuous linear-temperature gradient, the bound material was eluted over a 24° temperature range, with a Tm$^C$ of approximately 20° (Fig. 22B).

Tm$^C$ - temperature at the midpoint of the eluted peak.
Figure 22. Elution of $d(pA)_8$ on cellulose-$d(pT)_8$.

$d(pA)_8$ (2.6 A.U. in 1 ml MBS) was applied to cellulose-$d(pT)_8$ column (27 A.U. $d(pT)_8$, 0.9 cm x 5 cm) at $0^\circ$. The column was washed slowly at 1 ml per 22 min. In Figure A, the bound nucleotide was eluted by a stepwise thermal elution as indicated in the figure. In B, the bound nucleotide was eluted by a linear temperature gradient. In both A and B, the fraction size was 1 ml.
The elution of complementary oligomers on an oligonucleotide-cellulose is characterized by a relatively wide range of elution for the oligomer, compared with Gilham's deoxythymidine polynucleotide-cellulose column (69). The major difference between these two columns is that the polynucleotide-cellulose contains at least 200 times more bound oligonucleotide (see footnote 40). Also, the polynucleotide-cellulose may contain oligomers as long as d(pT)$_{12}$ or longer. In experiment described later (section 2, d, below) the range of elution on oligonucleotide-cellulose columns was reduced to approximately 8$^\circ$ to 12$^\circ$, by increasing the amount of oligonucleotide linked to the cellulose.

(iii) Typical elution profile. Because most of the studies on oligonucleotide-cellulose columns reported in this thesis have involved the use of linear temperature gradient elution, it seems worthwhile to describe in some detail a typical elution profile (see Fig. 23).

For all complementary nucleotides loaded onto an oligonucleotide-cellulose column, there was always a certain amount of UV-absorbing material not retained by the column. (For convenience this is referred to as peak A, tubes 2 to 20 in elution profile figures.) Peak A varied in size depending on the length of the oligomers; as the oligomer length decreased, peak A increased. It is possible that peak A represents impurities such as cyclic oligonucleotides or pyrophosphates,
Figure 23. Typical elution profiles for oligonucleotide-celluloses.

The elution of $d(pA)_8$ (A) and $d(pA)_11$ (B) on a cellulose column (39 A.U. $d(pT)_9$ (0.9 cm x 5 cm). The oligonucleotide sample (approximately 5 A.U. in 1 ml MBS) was slowly applied to the column at -4°.

The column was washed at this temperature until the absorbance of the eluate dropped to less than 0.003. The bound nucleotide was then eluted with a linear temperature gradient as indicated. The fraction size was 1.7 ml per 11 min, and the temperature gradient approximately 0.5° per 11 min. The dotted curve in Fig. 23B is the elution profile for a sample of $d(pA)_11$ which has been purified on a cellulose column by stepwise thermal elution (see inset Figure 18B).
the major contaminants in the chemical synthesis (151, 114). However, one would expect that the longer the oligonucleotides are, the less pure they would be, and thus the longer the oligomer, the larger peak A should be.⁴³

It is possible that peak A represents material which is complementary, but insufficient time was allowed during the loading and washing for it to form stable hybrid structures. This explanation is unlikely however, because when the column was loaded very slowly (by gravity, 2 hr), the size of peak A was not reduced. Diluting the sample several fold also did not decrease peak A. From capacity studies described later, it is known that the column was not overloaded.

Another characteristic of the elution profile was the symmetry of the elution of the retained nucleotide. As the length of the oligonucleotide increased, a significant amount eluted in a leading edge associated with the main peak (see Figure 23B). The nature of this material is not known. However, when the d(pA)₁₁ was eluted with a stepwise gradient (inset Figure 18B) and the material released between 29⁰ and 50⁰ reapplied

⁴³The statement "the longer the oligonucleotides are, the less pure they would be", is probably true. However, the impurities are not necessarily eluted in peak A. They may form sufficiently stable interactions and be retained by the column. Then as the temperature increases, they should be eluted at a lower temperature than the major peak.
to a cell-d(pT)$_9$ column it eluted as a symmetrical peak with no suggestion of a leading edge (Fig. 23B, dotted curve). This indicated that the materials in the leading edge were impurities that form less stable hybrid structures than the linear oligodeoxyadenylate. This suggestion is supported by Tm data on the purified deoxyadenylate (see Figure 18).

The volume in which the sample is eluted (60 to 70 ml) and the Tm$_C$ value were independent of the sample size (from 1 to 30 ml). However, there was some (random) variability in the sharpness of the elution curve, and reproducibility of the Tm$_C$ value for a particular oligonucleotide-cellulose and complementary oligonucleotide was $\pm 0.5^\circ$.

One final observation on the general elution profile was that after the elution of the bound nucleotide, in order to ensure that no nucleotide was still retained on the column, and to "regenerate" the cellulose, these columns were routinely washed at room temperature with 0.1 N NaOH. In all cases a small peak of UV-absorbing material was eluted. Spectra of the fractions showed that in the first fraction there was, possibly, a very small amount of nucleotidic material. (This alkali wash peak is not shown in most elution profiles.)

---

$^{44}$The recovery of nucleotide was 85 to 100%.
b) **Alternate methods for the elution of oligonucleotides from oligonucleotide-cellulose columns.**

Although the major portion of the work reported in this thesis has been concerned with thermal elution of oligonucleotides, other elution methods have been investigated, briefly.

If the principle behind these oligonucleotide-cellulose columns involves retention of complementary oligonucleotides by hydrogen bond formation, then any conditions which destabilize hybrid structure should be capable of effecting the elution of bound oligonucleotide.

Szybalski (145) has briefly reviewed the reagents or conditions which can reversibly destabilize hybrid structure. It is known that increasing the ionic strength of the solvent increases the stability of DNA to thermal denaturation (169, 170). This stabilization is based on the electrostatic neutralization of the charges on the phosphate groups by the cations in the solvent (145). Water-miscible organic solvents also affect the stability of hybrid structures, resulting in a decrease in Tm. If the denaturing solvent concentration is sufficiently high, denaturation of the DNA can be achieved at temperatures as low as room temperature (145). The effect of the solvents formamide, N,N'-dimethyl formamide, and dimethylsulfoxide on the stability of DNA have been studied by Levine et al. (171), Marmur and Ts'o (172), Herskovits (173, 174) and others. Certain salts at high
concentrations (e.g. NaClO₄) are also known to have a marked effect on hybrid stability (175). The Tm of sea urchin DNA (37% GC) is reduced approximately 40° in 6 M NaClO₄ buffered with EDTA at pH 7.0 (175).

In the experiments described here, the conditions investigated to effect the elution of bound nucleotides were reverse salt gradients (a linear gradient of decreasing salt concentration) and gradients of increasing organic solvent concentrations (formamide; N,N'-dimethylformamide, DMF; and dimethyl sulfoxide, DMSO). A gradient of increasing sodium perchlorate concentration was also used.

The "model system" employed to study these various conditions was the elution of d(pA)₇ (5 to 6 A.U. in 1 ml MBS, loaded at -4°) on a preparation of cell-d(pT)₉ (39 A.U. d(pT)₉, 0.9 cm x 5 cm). After the sample was loaded, the column was eluted at -4° with MBS until the absorbance at 260 (or 270) nm dropped. The temperature was then raised to 8 or 9° and a second peak of nucleotide eluted. The columns were then eluted as described below.

(i) **Reverse salt gradient.** A linear gradient of NaCl (1.0 M to zero) was run in 0.01 M NaH₂PO₄, pH 7.0 (total 200 ml) (Fig. 24B).

(ii) **Formamide.** A linear gradient of formamide (0 to 30%, v/v) in 0.01 M NaH₂PO₄, pH 7.0 1 M NaCl (total 200 ml) was run (Fig. 24C).
Figure 24. Alternate methods for the elution of oligonucleotides on oligonucleotide-cellulose columns. Cell-d(pT)$_9$ (39 A.U. d(pT)$_9$, 0.9 cm x 5 cm) was loaded with a sample of d(pA)$_7$ at -4°C in MBS. The nucleotides were eluted as described in the text.

A - linear temperature gradient.

B - reverse salt gradient, 1 M to 0.0 M NaCl.

C - linear gradient of formamide, 0 to 30%.

D - linear gradient of DMF, 0 to 30%.

E - exponential gradient of NaClO$_4$, 0 to 4.1 M.

In C and D, the dotted curve is the absorbance at 270 nm corrected for the absorbance due to formamide (C) or DMF (D).
(iii) **N,N'-dimethylformamide.** A linear gradient of DMF (0 to 30%, v/v) in 0.01 M Na$_2$PO$_4$, pH 7.0, 1 M NaCl (total 200 ml) was run (Fig. 24D).

(iv) **Dimethylsulfoxide.** A linear gradient of DMSO (0 to 30%, v/v) in 0.01 M Na$_2$PO$_4$, pH 7.0, 1 M NaCl (total 200 ml) was run. Under these conditions, the oligonucleotide d(pA)$_7$ was not eluted. At the end of the gradient, the column was washed with MBS. The temperature was then raised to 30° and the elution continued. The nucleotide was recovered quantitatively at this point.

(v) **Sodium perchlorate.** An exponentially increasing gradient of sodium perchlorate consisting of 100 ml of MBS in the mixing chamber and 100 ml of 7.2 M NaClO$_4$ (buffered with sodium EDTA, pH 7.0) (175) in a second flask was run (Fig. 24E).

In all these experiments, the flow rate was 1.7 ml per 11 min.

Elution of oligonucleotides with a reverse salt gradient resulted in a very broad peak of nucleotide (Fig. 24B). Similarly, formamide and DMF (Fig. 24C and D) did not give as sharp an elution of bound nucleotide as did the linear temperature gradient (Fig. 24A). From the relationship reported by McConaughy et al. (177) (a 1% increase of formamide concentration results in a drop in the

---

45 This gradient was set up such that the rate of removal of liquid from the mixing chamber was equal to the rate of addition of 7.2 M NaClO$_4$ to the mixing chamber. The gradient that results is an exponential one. The concentration of NaClO$_4$ reaches a maximum (approximately 4.1 M) halfway through the gradient (176).
Tm of 0.72°, in either 1 x or 5 x SSC), the concentration of formamide that should be required to elute d(pA)$_7$, at 8° can be calculated, assuming this relationship is valid for the buffer used in these experiments. The Tm of d(pA)$_7$ on cell-d(pT)$_9$ in MBS was 19°. Therefore, the concentration of formamide required to elute the d(pA)$_7$ at 8° could be approximately 14% DMF, v/v. The nucleotide actually eluted at a lower concentration of formamide (about 10%).

The oligonucleotide d(pA)$_7$ was eluted most sharply by the NaClO$_4$ gradient (the peak of nucleotide elutes, approximately, between 1 to 2 M NaClO$_4$). The sharpness of the peak is not unexpected from the data of Hamaguchi and Geiduschek (175).

c) Capacity of oligonucleotide-cellulose columns.

In early experiments to attach oligonucleotides onto cellulose (Table IX, preparations 1 to 5), incorporation of less than 60% was obtained, giving celluloses with less than 40 A.U. of covalently bound oligonucleotide. Since it was felt that at least 2 A.U. and preferably up to 5 or 6 A.U. of complementary oligonucleotide should be loaded to study the binding and elution of a particular oligomer, it was necessary to examine the capacity of the columns to determine at what point the column was overloaded. A cell-d(pT)$_8$ column (27 A.U. d(pT)$_8$) should be capable of retaining 33 A.U. of d(pA)$_8$, and thus a sample size of 5 A.U. should be well within the theoretical capacity of the column.
To examine the capacity of a cell-d(pT)$_8$ column (27 A.U. d(pT)$_8$, 0.9 cm x 5 cm), the column was loaded with six different amounts of d(pA)$_8$ oligomer (1.6, 3.4, 6.2, 10.6, 16.4 and 25.3 A.U. each in 1 ml MBS). The retained nucleotide was eluted with a linear temperature gradient. Figure 25 gives the elution profiles for three samples of d(pA)$_8$ (3.4, 10.6 and 25.3 A.U.). It can be seen that on increasing the sample load, there is an increase in the amount of material not retained by the column (peak A), and the retained peak changes from a symmetrical one to an asymmetrical one, with a pronounced leading edge. Data related to the % A.U. not retained, and the Tm$^C$ for each sample size are given in Table XI.

In determining the capacity of the column, it is necessary to define capacity. It seems convenient to use the two criteria:

(i) the non-proportionate increase in the nucleotide not retained by the column, relative to the size of the load (Table XI and Figure 26), and

(ii) the asymmetry of the elution of the bound oligonucleotide peak (Fig. 25).

Using the first of these criteria, this column was overloaded between 6 to 10 A.U. (Table XI and Figure 26). Based on the asymmetry of the eluted peak, the column was overloaded with 16.4 A.U. d(pA)$_8$ (the profile was similar to that in Figure 25C), and is probably loaded to its capacity with 10.6 A.U. (When a
Figure 25. The capacity of a cell-\(d(pT)_{8}\) column (27 A.U. \(d(pT)_{8}\), 0.9 cm x 5 cm). Various sample sizes of \(d(pA)_{8}\) were loaded and eluted as described in the text.

A - 3.4 A.U. \(d(pA)_{8}\).

B - 10.6 A.U. \(d(pA)_{8}\).

C - 25.3 A.U. \(d(pA)_{8}\).
Table XI. Capacity of cell-d(pT)$_8$ column (0.9 cm x 5 cm; 27 A.U. d(pT)$_8$). Samples of d(pA)$_8$ were loaded and eluted as described in the text.

<table>
<thead>
<tr>
<th>A.U. loaded</th>
<th>A.U. recovered</th>
<th>A.U. not retained</th>
<th>% A.U. not retained</th>
<th>Tm$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6</td>
<td>1.6</td>
<td>0.15</td>
<td>9.4</td>
<td>26</td>
</tr>
<tr>
<td>3.4</td>
<td>2.8</td>
<td>0.33</td>
<td>11.8</td>
<td>24</td>
</tr>
<tr>
<td>6.2</td>
<td>5.6</td>
<td>0.53</td>
<td>9.5</td>
<td>20</td>
</tr>
<tr>
<td>10.6</td>
<td>9.9</td>
<td>1.5</td>
<td>15.0</td>
<td>20</td>
</tr>
<tr>
<td>16.4</td>
<td>17.2</td>
<td>5.2</td>
<td>30.0</td>
<td>20</td>
</tr>
<tr>
<td>25.3</td>
<td>25.6</td>
<td>10.7</td>
<td>41.9</td>
<td>20</td>
</tr>
</tbody>
</table>
Figure 26. % A.U. \( d(pA)_8 \) not retained by a cell-\( d(pT)_8 \) column vs A.U. \( d(pA)_8 \) loaded.
perpendicular is dropped from the peak tube in Figure 25C, and twice the absorbance to the right of this perpendicular is estimated, a value of 10.4 A.U. is obtained.

Therefore, the capacity of this cell-d(pT)$_8$ column (27 A.U. d(pT)$_8$) is approximately 10 A.U. of d(pA)$_8$, which is 1/3 to 1/4 of the theoretical capacity of 33 A.U.

The leading edge in Figure 25C is presumably due to oligonucleotide which is only partially hydrogen bonded. Raising the temperature results in increasingly more stringent conditions for stable hybrid formation and thus retention of the nucleotide. The oligonucleotides must then rearrange to stabler structures, and the excess oligomers are eluted in this leading edge. This is illustrated below.

```
\[
\begin{align*}
\text{cell-} & \\
& \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \\
& \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \\
& \text{3'} \quad \text{5'}
\end{align*}
\]
```

stable at low temperatures

increase in temperature results in a rearrangement

```
\[
\begin{align*}
\text{cell-} & \\
& \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \text{pT} \\
& \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \\
& \text{3'} \quad \text{5'}
\end{align*}
\]
```

stable at higher temperatures

```
\[
\begin{align*}
& \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \text{A}_p \\
& \text{3'} \quad \text{5'}
\end{align*}
\]
```

eluted early
During this study of capacity it was observed that there was some dependence of the $Tm^C$ on the size of the sample (see Table XI). As the sample size decreased, the $Tm^C$ increased, slightly. This may partially explain the $Tm^C$ discrepancy between the elution of $d(pA)_8$ on cell-$d(pT)_8$ at $20^\circ$ (Table XI) described above, and the observation of Gilham (69) (using deoxythymidine polynucleotide-cellulose) that $d(pA)_7$ is eluted at $35^\circ$.46

d) Resolution of oligodeoxyadenylates on cell-$d(pT)_8$, cell-$d(pT)_9$, and cell-$d(pT)_12$.

The elution of the oligodeoxyadenylate series $d(pA)_n$ ($n = 6$ to 11) was studied on celluloses with covalently attached $d(pT)_8$, $d(pT)_9$ and $d(pT)_12$. The data for these experiments is given in Table XII. A composite graph of the elution profiles for the various adenylate oligomers on these oligodeoxythymidylate-celluloses is given in Figure 27. The data from these experiments may be summarized as follows.

(i) increasing the chain length of the oligodeoxyadenylate resulted in a stabler interaction (higher $Tm^C$ value) on a particular oligonucleotide-cellulose.

46 It should be emphasized that the substituted cellulose that Gilham used contained at least 200 fold more oligomer than the oligonucleotide-cellulose used here. Gilham's cellulose must also have contained oligomers much longer than the octathymidylate (69). The much higher capacity of the polynucleotide-cellulose very likely results in the greater resolution observed.
Table XII. Elution of deoxyadenylates d(pA)$_6$ to d(pA)$_{11}$, on cell-d(pT)$_8$, cell-d(pT)$_9$, and cell-d(pT)$_{12}$.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>A.U. loaded</th>
<th>Tm$^C$ (°C)</th>
<th>Range of elution (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Cell-d(pT)$_8$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(pA)$_6$</td>
<td>7.2</td>
<td>3.5</td>
<td>19</td>
</tr>
<tr>
<td>d(pA)$_7$</td>
<td>6.3</td>
<td>13.5</td>
<td>19</td>
</tr>
<tr>
<td>d(pA)$_8$</td>
<td>6.2</td>
<td>19.0</td>
<td>15</td>
</tr>
<tr>
<td>d(pA)$_9$</td>
<td>6.2</td>
<td>25.5</td>
<td>20</td>
</tr>
<tr>
<td>d(pA)$_{10}$</td>
<td>6.0</td>
<td>28.5</td>
<td>14.5</td>
</tr>
<tr>
<td>B. Cell-d(pT)$_9$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(pA)$_6$</td>
<td>5.4</td>
<td>8.5</td>
<td>19</td>
</tr>
<tr>
<td>d(pA)$_7$</td>
<td>5.5</td>
<td>18.0</td>
<td>16</td>
</tr>
<tr>
<td>d(pA)$_8$</td>
<td>5.5</td>
<td>26.0</td>
<td>19</td>
</tr>
<tr>
<td>d(pA)$_9$</td>
<td>5.6</td>
<td>32.0</td>
<td>16</td>
</tr>
<tr>
<td>d(pA)$_{10}$</td>
<td>5.7</td>
<td>35.0</td>
<td>16</td>
</tr>
<tr>
<td>d(pA)$_{11}$</td>
<td>5.6</td>
<td>37.0</td>
<td>14</td>
</tr>
<tr>
<td>C. Cell-d(pT)$_{12}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(pA)$_6$</td>
<td>4.0</td>
<td>8.5</td>
<td>18</td>
</tr>
<tr>
<td>d(pA)$_7$</td>
<td>4.8</td>
<td>14.5</td>
<td>17.5</td>
</tr>
<tr>
<td>d(pA)$_8$</td>
<td>4.2</td>
<td>20.5</td>
<td>18.5</td>
</tr>
<tr>
<td>d(pA)$_9$</td>
<td>3.0</td>
<td>27.5</td>
<td>11.0</td>
</tr>
<tr>
<td>d(pA)$_{10}$</td>
<td>4.4</td>
<td>34.0</td>
<td>11.0</td>
</tr>
<tr>
<td>d(pA)$_{11}$</td>
<td>4.8</td>
<td>37.5</td>
<td>14.0</td>
</tr>
</tbody>
</table>
Figure 27. Composite graph of elution profiles for elution of d(pA)$_n$ on cell-d(pT)$_8$, A; cell-d(pT)$_9$, B; and cell-d(pT)$_{12}$, C. The samples were all loaded at -4º and eluted with a linear temperature gradient.
(ii) The $\Delta T_m^C$ increment with increasing length of oligodeoxyadenylate is about $6^\circ$ to $8^\circ$ when the length of the dA oligomer is less than or equal to the length of the dT oligomer, linked to the cellulose. Beyond this length, there is a very noticeable decrease in the $\Delta T_m^C$ increment (approximately $2.5^\circ$ to $3.0^\circ$) for each additional nucleotide. These results suggest that all the nucleotides linked to the cellulose are available for hybrid formation with complementary nucleotides.

If one considers the relationship between $T_m^C$ and oligomer length, $n$, for the theoretical case where consecutive oligomers show a $\Delta T_m^C$ of $6^\circ$ when the length of the dA oligomer is the length of the cellulose-bound dT oligomer, and beyond this length the $\Delta T_m^C$ is $3^\circ$, then a plot of $T_m^C$ versus oligodeoxyadenylate length will give two straight lines which intersect where the length of the oligodeoxyadenylate is equal to the length of the oligodeoxythymidylate. When the $T_m^C$ data for the three different oligodeoxythymidylate-cellulose (Table XII) versus oligomer length, $n$, are plotted (Figure 28), at least 2 curves can be drawn in 28A and B. The dotted lines intersect just below where the length of the deoxyadenylate is equal to the length of the cellulose-bound deoxythymidylate, while the solid lines intersect just beyond this point.

It would seem from the data that at least all but one of the nucleotides, and possibly the entire oligonucleotide attached to
the cellulose, is free to interact with a complementary oligonucleotide. The interaction involving the deoxythymidylate residue adjacent to the cellulose may be weaker (due to steric reasons) than the interactions involving the rest of the nucleotides. However, the \( \Delta T_m^C \) increment suggests this nucleotide does interact as strongly. Data presented below in section 2,1 also suggest that all the nucleotides attached to the cellulose are free to interact with complementary nucleotides.

Figure 28 implies that there is a linear relationship between \( T_m^C \) and the length of unbound oligomers. This relationship may not be a linear one.

(iii) increasing the chain length of the oligonucleotide-cellulose increases the \( T_m^C \) of the corresponding complementary oligonucleotide in the case of cell-d(pT)_8 and cell-d(pT)_9. However, for cell-d(pT)_12, although the \( T_m^C \) vs oligomer length gives a linearly increasing relationship from \( n = 6 \) to \( n = 11 \) (Fig. 28C), the \( T_m^C \) values for the various oligodeoxyadenylates are not significantly greater than for cell-d(pT)_9. This may be explained partially by the earlier observation that the \( T_m^C \) value is somewhat dependent on the size of the sample (see capacity discussion above). The cell-d(pT)_12 contained an estimated 30 A.U. of covalently linked d(pT)_12 (Table IX). This column would therefore have as much available "exchanger units" as a cell-d(pT)_9 column with 22 A.U. d(pT)_9. The cell-d(pT)_9
Figure 28. $T_m^C$ of $d(pA)_n$ on celluloses-$d(pT)_{8}$, $d(pT)_{9}$, and $d(pT)_{12}$ vs oligomer length $n$.

A - cell-$d(pT)_{8}$.

B - cell-$d(pT)_{9}$.

C - cell-$d(pT)_{12}$.
column used in these studies contained almost twice that amount of oligothymidylate. The lower elution temperatures for the deoxyadenylate series on cell-d(pT)$_{12}$ may also be partially explained by observations described in part 2f, below.

From Figure 27 and Table XIII, it can be seen that consecutive oligomers are not completely resolved on these small oligonucleotide-cellulose columns (0.9 cm x 5 cm,~35 A.U. oligothymidylate), although oligomers differing in length by two nucleotides are resolved. Therefore, a larger column of cell-d(pT)$_{9}$ of higher capacity (0.9 mm x 15 cm, 150 A.U. d(pT)$_{9}$) was used to study the elution of d(pA)$_n$ ($n = 6$ to 9). The results are given in Table XIII and Figure 29. By increasing the size of the column, it was found that consecutive oligomers were resolved, and the range of elution was reduced from 15°-20° to 8°-12°. One other observation was that the oligonucleotides eluted at 1.5° to 3° higher from the larger column, consistent with the earlier observation that the Tm$^c$ value is dependent to some degree on the capacity of the column.

In summarizing the observations on the resolution of

---

47 In an experiment to see if increasing the flow rate relative to the temperature gradient would resolve consecutive oligomers, it was found that a particular oligomer was eluted over the same temperature range.

48 For this larger column, 4.7 ml fractions were collected per 11 min, and the temperature gradient was 1° per 22 min.
Table XIII. Elution of $d(pA)_n$ on cell-$d(pT)_g$ columns.

A. Cell-$d(pT)_g$ (0.9 cm x 15 cm, 150 A.U. $d(pT)_g$)

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>A.U. loaded</th>
<th>Tm°C (°C)</th>
<th>Range (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d(pA)_6$</td>
<td>8.8</td>
<td>10.0</td>
<td>11</td>
</tr>
<tr>
<td>$d(pA)_7$</td>
<td>7.8</td>
<td>21.0</td>
<td>12</td>
</tr>
<tr>
<td>$d(pA)_8$</td>
<td>6.5</td>
<td>28.0</td>
<td>8.5</td>
</tr>
<tr>
<td>$d(pA)_9$</td>
<td>6.8</td>
<td>33.5</td>
<td>9.5</td>
</tr>
</tbody>
</table>

B. Cell-$d(pT)_g$ (0.9 cm x 15 cm, 39 A.U. $d(pT)_g$)

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>A.U. loaded</th>
<th>Tm°C (°C)</th>
<th>Range (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d(pA)_6$</td>
<td>5.4</td>
<td>8.5</td>
<td>19</td>
</tr>
<tr>
<td>$d(pA)_7$</td>
<td>5.5</td>
<td>18.0</td>
<td>16</td>
</tr>
<tr>
<td>$d(pA)_8$</td>
<td>5.5</td>
<td>26.0</td>
<td>19</td>
</tr>
<tr>
<td>$d(pA)_9$</td>
<td>5.6</td>
<td>32.0</td>
<td>16</td>
</tr>
</tbody>
</table>
Figure 29. Composite elution profile for $d(pA)_n$ ($n = 6$ to 9) on cell-$d(pT)_g$ $(0.9 \text{ cm} \times 15 \text{ cm}, 150 \text{ A.U.} \ d(pT)_g)$. 

\[ A_{260\text{nm}} \] 
\[ VOLUME \text{ ml} \]
\[ TEMPERATURE \] 
\[ 0, 10, 30, 50 \]
consecutive oligomers on small oligonucleotide-cellulose columns, several points are important. The $\Delta T_m^C$ for consecutive oligomers (when the free oligomer is shorter than the oligonucleotide exchanger) is slightly larger than the $\Delta T_m$ for the same series in solution (compare the data in Table XII with those in Table VII). This is true even for a relatively small column with approximately 0.5 μmole of covalently linked exchanger. The range of elution ($15^\circ$ to $20^\circ$) is less than the width of the corresponding thermal dissociation curve (about $30^\circ$). On a column with higher capacity (about 2 μmoles of bound nucleotide (Table XIII), the range of elution of 0.1 μmole of free oligomer is decreased further. Under these conditions consecutive oligomers are just resolved, whereas the thermal denaturation studies, in solution, had suggested that one would get overlap of five or more successive oligomers. Although most of the columns studied contained approximately 0.5 to 1.0 μmole of bound oligomer, preparation of columns with 2 to 5 μmoles of bound oligomer (of defined length and sequence), are practicable, using the existing chemical synthetic developed by Khorana and his associates (75).

e) Elution of $d(pA)_n$, $d(A)_n$, and $r(A)_n$ on cellulose-$d(pT)_9$.

If deoxynucleotide-celluloses are to be used in the isolation of naturally occurring ribonucleic acids, it is important to know how oligoribonucleotides elute from these columns in comparison with oligodeoxyribonucleotides. Chamberlin and his
coworkers (178, 179) have studied a number of polydeoxyribonucleotide and polyribonucleotide interactions. They have found that, depending on the particular base pair, there is a distinct order of thermal stability. For example, the order for increasing Tm values for the following homopolymer interactions was observed: 

dI:rC < dI: dC < rl:dC < rl: rC < and 
dG:dC < dG:rC < rG:dC < rG:rC.

From these experiments, it cannot be predicted whether d(T)n:r(A)n interactions would be more stable or less stable than d(T)n:d(A)n interactions. However, Riley et al. (180) have also studied the homopolymer interactions dA:dT and rA:dT. At 0.1 M sodium ion concentration, the dA:dT interaction had a Tm of 68° while the rA:dT interaction had a Tm of 63°. Therefore, one would predict that oligoriboadenylates would be eluted from oligodeoxythymidylate-celluloses at lower temperatures than the corresponding oligodeoxyadenylates. (It should also be mentioned that in studies with naturally occurring polynucleotides, DNA/RNA hybrids have been found to be less stable than DNA/DNA hybrids (181).)

To examine the elution of oligoribonucleotides, the series r(A)n (n = 6 to 9) was run on a preparation of cell-d(pT)9 (0.9 cm x 5 cm, 39 A.U. d(pT)9). For comparison, the dephosphorylated deoxyribo-series d(A)9 was also run. The results of these experiments are listed in Table XIV (see also
Table XIV. Elution of oligoribo- and oligodeoxyribonucleotides on cell-d(pT)$_9$.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>A.U. loaded</th>
<th>Tm$^c$ ($^\circ$C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>r(A)$_6$</td>
<td>5.4</td>
<td>5.0</td>
</tr>
<tr>
<td>r(A)$_7$</td>
<td>6.5</td>
<td>14.0</td>
</tr>
<tr>
<td>r(A)$_8$</td>
<td>4.3</td>
<td>20.0</td>
</tr>
<tr>
<td>r(A)$_9$</td>
<td>5.1</td>
<td>25.5</td>
</tr>
<tr>
<td>d(A)$_6$</td>
<td>7.0</td>
<td>9.0</td>
</tr>
<tr>
<td>d(A)$_7$</td>
<td>4.0</td>
<td>19.5</td>
</tr>
<tr>
<td>d(A)$_8$</td>
<td>4.2</td>
<td>26.5</td>
</tr>
<tr>
<td>d(A)$_9$</td>
<td>5.2</td>
<td>34.5</td>
</tr>
<tr>
<td>d(pA)$_6$</td>
<td>5.4</td>
<td>8.5</td>
</tr>
<tr>
<td>d(pA)$_7$</td>
<td>5.5</td>
<td>18.0</td>
</tr>
<tr>
<td>d(pA)$_8$</td>
<td>5.5</td>
<td>26.0</td>
</tr>
<tr>
<td>d(pA)$_9$</td>
<td>5.6</td>
<td>32.0</td>
</tr>
</tbody>
</table>
Figure 30). It is clear that the oligoribonucleotides elute at 4° to 9° lower than the corresponding deoxyribo-oligomers. One other interesting observation is that the phosphorylated d(pA)$_n$ elutes at approximately 1° to 2° lower than the dephosphorylated series d(A)$_n$, suggesting that even at a very high salt concentration (M NaCl), the 5'phosphate group on the oligonucleotide is capable of reducing the stability of the hybrid structure.

f) Characteristics of different preparations of oligonucleotide-celluloses.

It was of interest to compare the properties of different preparations of the same oligonucleotide-cellulose in order to see if there was good reproducibility of retention properties. The two preparations that were compared initially were preparation #4, cell-d(pT)$_g$ prepared as described by Gilham (141), and preparation #10, cell-d(pT)$_g$ prepared by a modified procedure (see Methods, Part II and Table IX), using a larger amount of carbodiimide. The oligoadenylates d(pA)$_g$ and d(pA)$_7$ were eluted at a significantly higher temperature from preparation #4 than preparation #10 (see Table XV). The difference in Tm$^c$ values observed is significant in that it is known that the repeat elution of an oligonucleotide from the same column is reproducible to within ± 0.5°.

There are three possible explanations for this variability
Figure 30. $T_m^c \, (^\circ C)$ versus oligomer length for the elution of the series $d(pA)_n$, $d(A)_n$, and $r(A)_n$ on cell-$d(pT)_g$. 
Table XV. Elution of d(pA)₉ and d(pA)₇ on different preparations of cell-d(pT)₉:

<table>
<thead>
<tr>
<th>Cellulose preparation</th>
<th>#</th>
<th>A.U. Incorporated of d(pA)₇</th>
<th>Tm°C (°C) of d(pA)₇</th>
<th>Tm°C (°C) of d(pA)₉</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. cell-d(pT)₉</td>
<td>4</td>
<td>39</td>
<td>18.0 (5.5 A.U.)</td>
<td>32.0 (5.6 A.U.)</td>
</tr>
<tr>
<td>2. cell-d(pT)₉</td>
<td>10</td>
<td>48.5</td>
<td>14.0 (5.7 A.U.)</td>
<td>29.5 (5.3 A.U.)</td>
</tr>
<tr>
<td>3. cell-d(pT)₉</td>
<td>10</td>
<td>48.5</td>
<td>14.0 (5.5 A.U.)</td>
<td>29.5 (5.3 A.U.)</td>
</tr>
<tr>
<td>(treated at pH 10.9 for 20 hr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. cell-d(pT)₉</td>
<td>14</td>
<td>69</td>
<td>19.0 (5.1 A.U.)</td>
<td>-</td>
</tr>
<tr>
<td>(treated at pH 10.9 for 20 hr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. cell-d(pT)₉</td>
<td>15</td>
<td>47</td>
<td>16.5 (5.1 A.U.)</td>
<td>-</td>
</tr>
<tr>
<td>(treated at pH 10.9 for 20 hr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In experiment 3, cell-d(pT)₉ (#10) was incubated at pH 10.9 (Na₂CO₃ buffer) for 20 hr, as described in the text. It was noticed that after this treatment, the flow rate was noticeably reduced, although a rate of 1.7 ml per 11 min could still be maintained. For preparations #14 and #15, the papers were treated at pH 10.9 prior to being cut up to form a column matrix. In both these preparations, the flow rates were normal. Therefore, treatment of cellulose at alkaline pH for extended periods prior to cutting up the paper is preferred.
There are three possible explanations for this variability between different preparations of the same oligonucleotide-cellulose.

(i) The incorporation of d(pT)₉ onto cellulose in preparation #10 used twice as much carbodiimide as was used in preparation #4. It is possible that the reagent has produced some side product, e.g. CMC-dpT derivatives which would interfere with the hydrogen bonding capabilities of the oligodeoxythymidylate. Experiments reported above, showed that under these reaction conditions approximately 2% of d-pT is converted to the CMC-dpT. (If 2% of the thymidylate residues have been substituted in the d(pT)₉ preparation, then 17% of the oligonucleotides contain one or more substituted thymidylate residues (168)).

(ii) These preparations of d(pT)₉ were made by different methods. Preparation #4 was obtained from a polymerization of d-pT, while #10 was from a polymerization of d-pTpTpT (see part I of this thesis).

(iii) There is inherent, unexplained variability in different preparations of the same oligonucleotide-cellulose.

If the cause of the variable properties of the same oligonucleotide-cellulose is due to substitution of deoxythymidylate residues by CMC, then incubation of preparation #10 in 0.2 M Na₂CO₃, pH 10.9, for 20 hr at room temperature should result in complete conversion of the CMC substituted residues to free thymidylates (160). Therefore preparation #10 was incubated at
pH 10.9, and the elution of $d(pA)_7$ and $d(pA)_9$ again studied. The data in Table XV (experiment 3) clearly show that this incubation had no effect whatsoever on the elution temperature of these two oligodeoxyadenylates.

Another possible explanation is that the $d(pT)_9$ used in the preparation of these two celluloses was from different syntheses (Part I). The $d(pT)_9$ used in preparation #4 was identified by its position of elution from a DEAE-cellulose column, as well as molar base ratio data (see part I, Table II). There seems little doubt from these data that this was not $d(pT)_9$. Similarly, the identification of the $d(pT)_9$ obtained from the polymerization of d-pTpTpT seems sound, based on its elution from DEAE-cellulose and co-chromatography in system C with the earlier preparation of $d(pT)_9$, both before and after alkaline phosphatase digestion.

One is therefore left to conclude that the variability between different oligonucleotide-cellulose preparations is the result of unexplained inherent variability. This suggestion is supported by data on two further preparations of cell-d(pT)$_9$ (#14 and #15, Table IX). These preparations were both made with $d(pT)_9$ obtained by the polymerization of d-pTpTpT, and both were made using 100 mg carbodiimide. $d(pA)_7$ elutes from preparation #14 at 19.0°, and from preparation #15 at 16.5° (see Table XV).

It would seem that each preparation of oligonucleotide-
cellulose should be standardized prior to its use in isolating naturally occurring, complementary sequences.

g) Oligonucleotide-cellulose of mixed repeating base sequences and the elution of complementary oligomers.

Oligonucleotides of the general form d(pTpC)n and d(pCpT)n (n = 2,3) were prepared as described in part I, and then covalently linked to cellulose paper using the water-soluble carbodiimide method (Table IX). These celluloses have been tested for their ability to retain complementary oligonucleotides of the type d(pApG)n (n = 2,3). A summary of the TmC data for the AAG oligomers on the different complementary pyrimidine oligonucleotides is listed in Table XVI. The corresponding profiles are shown in Figure 31 and 32.

(i) Celluloses-d(pTpC)2 and -d(pCpT)2. From Table XVI and Figure 31A, it is clear that although cell-d(pTpC)2 does not retain d(pApG)2 at -5.0, it is definitely retarded on this column. (The dotted curve in Figure 31A is the elution profile for the non-complementary oligomer d(pA)9.) The d(pApG)2 peak is displaced approximately 5 to 10 tubes (10 to 20 ml). However, on cell-d(pCpT)2 (where one more GC base pair is possible), the hexanucleotide d(pApG)2 is retained, and elutes at 14.0. (This observation further supports the suggestion that all the nucleotides attached to the cellulose are free to hydrogen bond with a complementary sequence.)
Table XVI. Elution of $d(pApG)_n$ ($n = 2,3$) on oligonucleotide-celluloses of mixed, repeating base sequences.

The columns had a packed dimension of 9 mm × 5 cm.

<table>
<thead>
<tr>
<th>Oligonucleotide-cellulose</th>
<th>$A_{266}$ units incorporated</th>
<th>Sample ($A_{260}$ units)</th>
<th>$Tm^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell-$d(pTtGtT)_2$</td>
<td>92</td>
<td>$d(pApG)_2$ (3.0)</td>
<td>not retained but retarded at $-5^\circ$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d(pApG)_3$ (2.5)</td>
<td>15.5(^\circ)</td>
</tr>
<tr>
<td>Cell-$d(pCtGtT)_2$</td>
<td>97</td>
<td>$d(pApG)_2$ (3.0)</td>
<td>14.0(^\circ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d(pApG)_3$ (2.5)</td>
<td>19.0(^\circ)</td>
</tr>
<tr>
<td>Cell-$d(pTtGtT)_3$</td>
<td>77</td>
<td>$d(pApG)_2$ (3.0)</td>
<td>4.5(^\circ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d(pApG)_3$ (2.5)</td>
<td>33.5(^\circ)</td>
</tr>
<tr>
<td>Cell-$d(pCtGtT)_3$</td>
<td>103</td>
<td>$d(pApG)_2$ (3.0)</td>
<td>16.0(^\circ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d(pApG)_3$ (2.5)</td>
<td>40.5(^\circ)</td>
</tr>
</tbody>
</table>
Figure 31. Elution of \( d(pApG)_2,3 \) on cellulosics \( d(pTpC)_2 \) and \( d(pCpT)_2 \).

A - Cell \( d(pTpC)_2 \) - elution \( d(pApG)_2 \)

B - Cell \( d(pCpT)_2 \) - elution \( d(pApG)_2 \)

C - Cell \( d(pTpC)_2 \) - elution \( d(pApG)_3 \)

D - Cell \( d(pCpT)_2 \) - elution \( d(pApG)_3 \).
Elution profiles for the nonanucleotide \( d(pApApG)_3 \) on both cell\(-d(pTpC)_2 \) and cell\(-d(pCpT)_2 \) are given in Figure 31C and D, respectively. If all the nucleotides on the cellulose can hydrogen bond, then the \( Tm^C \) of \( d(pApApG)_3 \) on both hexanucleotide columns should be similar. In fact, they differ by 4.5\(^\circ\). This difference could possibly be explained by inherent random variation in cellulose preparations as discussed above for different cell\(-d(pT)_9 \) preparations. However, it is also possible that steric interference may destabilize the cell\(-d(pTpC)_2 \)'\( d(pApApG)_3 \) interaction (see scheme below). In the cell\(-d(pTpC)_2 \)'\( d(pApApG)_3 \) interaction, for the stablest structure (4 AT bonds, 2 GC bonds), an extra nucleotide (deoxyguanylate residue) is adjacent to the cellulose, and it is possible that this may destabilize the hybrid structure. Such an effect would be a serious problem in the isolation of naturally occurring nucleic acids by hybridization.

\[
\begin{align*}
\text{cell-} &
\begin{array}{l}
\text{cell-} &
\end{array}
\begin{array}{c}
\begin{array}{c}
5' \quad \tilde{T}pT\tilde{T}pCpT\tilde{T}pC \quad 3'.
\end{array}
\end{array}
\begin{array}{c}
G_pA_pG_pA_pG_pA_pG_pA_pA_p
\end{array}
\text{extra } d-pG
\end{align*}
\]

with a small, internal segment. However, one way around this problem might be to construct an oligonucleotide-cellulose with an "extension block" of nucleotides (e.g. the tetrathymidylate),
and attach the synthetic oligonucleotide of defined sequence to the 3' hydroxyl end of this thymidylate block as illustrated below, where the complementary sequence desired is d-pApC(pT)₅.

\[
\text{cell-T₆T₅T₄T₃A₄C₅T₆T₇T₈T₉T}
\]

Another observation concerning these hexanucleotide-celluloses is that when d(pApApG)₃ is run on cell-d(pTpTpC)₂, the non-retained peak has a trailing edge (Figure 31C) while the retained peak on cell-d(pCpTpT)₂ (Figure 31D) has a distinct leading edge. It seems likely that the d(pApApG)₃ preparation, although purified by chromatography on a DEAE-cellulose column in the presence of 7 M urea, followed by extended paper chromatography (see part 1), still contains a significant amount of contaminants. If the major contaminant is the 6:3 pyrophosphate, it should elute at a temperature similar to the hexanucleotide. Therefore, the 6:3 pyrophosphate should be retarded, but not retained at -5°C, on cell-d(pTpTpC)₂, while on cell-d(pCpTpT)₂, the 6:3 pyrophosphate should elute at approximately 14°C. The correspondence between the predicted elution of this pyrophosphate on both hexanucleotide celluloses, and the observed elution position of this "contaminant"

\[pApApgpApG\]

The 6:3 pyrophosphate is
supports the suggestion of that this material is the pyrophosphate. The stablest interactions for the 6:3 pyrophosphate on the cellulose hexanucleotides are illustrated below:

\[
\begin{align*}
\text{cell-} & \ pT_pT_pC_pT_pT_pC_pA_pG_pA_pA_pG_pA_pA_p & \text{cell-} & \ pC_pT_pT_pC_pT_pT_pC_pA_pG_pA_pA_pA_p \rightarrow \\
& \rightarrow \text{G_p} & \rightarrow \text{G_p} & \rightarrow \\
& \text{A_p} & \text{A_p} & \text{A_p} \\
\end{align*}
\]

4 AT 1 GC

(ii) Cellulose-\(d(pT_pT_pC)\)\(_3\) and \(-d(pC_pT_pT)\)\(_3\).

a) Elution of the hexanucleotide, \(d(pA_pA_pG)\)\(_2\).

Cell-\(d(pT_pT_pC)\)\(_3\) elutes \(d(pA_pA_pG)\)\(_2\) at 4.5°, while the isomeric oligonucleotide-cellulose, cell-\(d(pC_pT_pT)\)\(_3\) elutes this hexanucleotide at 14.5° (Table XVI and Figure 32A and B). Although in both cases the number of interactions (4 AT and 2 GC base pairs) is the same, there is a considerable difference in the Tm\(_c\) values (10°). Again, this difference may (at least partially) be explained by random variability of cellulose preparations. However, other explanations may be responsible. For example, the stablest structure for cell-\(d(pT_pT_pC)\)\(_3\) (based on the maximum number of base pairs) is illustrated below, while for cell-\(d(pC_pT_pT)\)\(_3\)
Figure 32. Elution of $d(pApApG)_{2,3}$ on celluloses-$d(pTpTpC)_{3}$ and -$d(pCpTpT)_{3}$.


The inset in Figure 31D is the profile for the stepwise elution of $d(pApApG)_{3}$ on cell-$d(pCpTpT)_{3}$. The temperature steps are indicated in the figure.
the stablest structure may be either of the following (or a combination of both).

\[
\text{cell-} \quad \begin{array}{cccccccc}
\text{pC} & \text{pT} & \text{p} & \text{pC} & \text{pT} & \text{pC} & \text{pT} & \text{pC} \\
\text{G} & \text{Ap} & \text{Ap} & \text{G} & \text{Ap} & \text{Ap} & \text{G} & \text{Ap} & \text{Ap}
\end{array}
\]

\[
\text{cell-} \quad \begin{array}{cccccccc}
\text{pC} & \text{pT} & \text{p} & \text{pC} & \text{pT} & \text{pC} & \text{pT} & \text{pC} & \text{pT} \\
\text{G} & \text{Ap} & \text{Ap} & \text{G} & \text{Ap} & \text{Ap} & \text{G} & \text{Ap} & \text{Ap}
\end{array}
\]

In the first case, there is a gap of two unpaired bases (pTpT), before hybridization begins, while for the other case, the d(pApApG)\textsubscript{2} can hybridize immediately adjacent to the cellulose. If there is some "stabilizing effect" associated with this latter configuration, then this may explain the difference in Tm values (partially or wholly) for these interactions. A possible "stabilizing effect" might result from hydrogen bonding between nucleotides at the 3' end of the purine oligomer with the cellulose. Secondary binding forces between nucleosides (particularly purinedeoxyribosides) on cellulose have been observed, and the elution of different oligonucleotide isopliths on DEAE-cellulose is greatly improved when the chromatography is carried out in the presence of 7 M urea (182).

b) Elution of the nonanucleotide, d(pApApG)\textsubscript{3}

(Figure 32C and D).

The nonanucleotide d(pApApG)\textsubscript{3} was eluted at 33.5\degree on cell-
d(pTpTpC)\textsubscript{3} and 40.5\degree on cell-d(pCpTpT)\textsubscript{3}. In both cases, a broad
peak of UV-absorbing material (as much as 30% of the UV-absorbance and with a spectrum like d(pApApG)₃) was eluted prior to the major, d(pApApG)₂, peak. The estimated Tm° value for this peak on cell-d(pTpTpC)₃ is 12°, and on cell-d(pCpTpT)₃ is 16°.

If this nonanucleotide is contaminated with the 6:3 pyrophosphate as discussed above, then the Tm° values of the pyrophosphate on these nonanucleotide-celluloses should be related to the Tm° values for the elution of d(pApApG)₂ on the nonanucleotide-celluloses. These values are 4.5° and 14.5°, respectively, for cell-d(pTpTpC)₃ and cell-d(pCpTpT)₃. The "contaminant" elutes at a temperature somewhat higher than expected in both cases.

Since considerable evidence for the presence of a major contaminant in the d(pApApG)₃ preparation was obtained, it was decided to attempt to purify this nonanucleotide on a cell-
d(pCpTpT)₃ column, using stepwise thermal elution (see inset Figure 32D). The material eluting between 32° and 50° was concentrated, dialyzed against distilled water, and lyophilized. The nucleotide was dissolved in MBS and samples for solution thermal denaturation studies prepared as described in Methods. The two interactions, d(pApApG)₃.d(pTpTpC)₃ and d(pApApG)₃.d(pCpTpT)₃ were studied (Figure 33C and D). In both cases, the thermal dissociation profile with the oligonucleotide-cellulose purified d(pApApG)₃ was noticeably sharper, and the Tm value
Figure 33. Thermal denaturation profiles of the interactions d(pApApG)₃ with d(pTpTpC)₃ and d(pCpTpT)₃ prior to and after purification of the d(pApApG)₃ on an oligonucleotide-cellulose column.


— observed thermal dissociation curve for the mixtures.

— observed thermal dissociation curve for the purine oligomer.

— true thermal denaturation profile.
increased by 6.0°. These observations further support the idea that the \(d(pApApG)\) preparation contained a significant amount of impurity.

h) Effect of GC base pairs on the stability of oligonucleotide-cellulose:complementary oligomer interactions.

The data for the elution of the oligomers, \(d(pApApG)_{2,3}\) on complementary, isomeric cellulososes (Table XVI) indicate that the deoxyoligomers containing GC base pairs are significantly more stable than the homo-oligonucleotide-cellulose interactions (Table XII). Note however that for the mixed sequence cellulososes, approximately 0.07 \(\mu\)mole of purine oligomer was chromatographed on a 1 \(\mu\)mole column. For the cell-\(d(pT)\) columns, approximately 0.07 \(\mu\)mole of oligodeoxyadenylate was chromatographed on a cellulose containing 0.5 \(\mu\)mole of \(d(pT)\). This could affect the \(Tm^c\) by 1° to 2° (see section 2d, above).

The oligomer \(d(pA)_6\) is retarded on cell-\(d(pT)_6\) (6 AT base pairs) at -5°, as is \(d(pApApG)_2\) on cell-\(d(pTpTpC)_2\) (4 AT pairs, 1 GC pair).

Elution of oligomers, \(d(pApApG)_{2,3}\) on mixed complementary oligonucleotide-cellulososes in which 4 AT and 2 GC base pairs are possible, resulted in \(Tm^c\) values of approximately 14° to 19°. A possible explanation for the one exception (4.5°) was discussed in section 2g, above. On the other hand, \(d(pA)_6\) eluted from three
different oligothymidylate-celluloses at lower temperatures (Table XII).

In the one possible 6 AT, 2 GC interaction, \( d(pApApG) \), on cell-\( d(pTpC) \), the \( Tm^c \) was 33.5°. The homo-oligodeoxyribonucleotide, \( d(pA)_8 \), eluted at lower temperatures from deoxythymidylate-celluloses in which 8 AT interactions are possible (Table XII).

Similarly, the one possible 6 AT, 3 GC interaction, \( d(pApApG) \), on cell-\( d(pCpTpT) \), had a \( Tm^c \) of 40.5°. Interactions involving 9 AT base pairs are less stable (Table XII).

i) Elution of non-complementary oligonucleotides on oligonucleotide-celluloses.

If the relationship described by Laird et al. (29) (1% mismatching os bases decreases the \( Tm \) by 0.7°) is valid for these short oligomers, substitution of 1/3 of the bases should reduce the \( Tm \) by some 24°.

In experiments to see if non-complementary sequences would be excluded from oligonucleotide-celluloses, oligomers of the series \( d(pA)_n \) were chromatographed on cell-\( d(pCpTpT) \), and cell-\( d(pCpTpT) \). Similarly, \( d(pApApG)_{2,3} \) were chromatographed on cell-\( d(pT)_9 \). In the first case, none of the oligodeoxyadenylate were retained by the \( d(pCpTpT)_{2,3} \) cellulose (for example, see Figure 31A). Less than 1.3% of the \( d(pApApG)_{2} \) sample was retained by a cell-

\( d(pT)_9 \) column at -4°, while <5% of a \( d(pApApG) \) sample was retained
in a similar experiment. It is not known what the thermal stability of the small amount of retained nucleotide was, as the temperature of the column was raised, rapidly, to 30° in order to elute any retained material as a sharp peak. In the latter case, had the nucleotides been perfectly complementary, \( (pApApG)^3 \cdot (pCpTpT)^3 \), the Tm would have been about 40° to 41°, while this very small amount of non-complementary nucleotide that was retained, was eluted at 30° (or lower).

j) Purification of chemically synthesized oligonucleotides on oligonucleotide-cellulose columns.

The observations made in the previous sections concerning the apparent presence of impurities both in the synthetic oligodeoxyadenylates (e.g. \( d(pA)_{11} \)) and in the \( d(pApApG)^3 \) oligomer, suggest that the purification methods used for the separation of these oligonucleotides is not entirely satisfactory (see e.g. 91). In the case of the \( d(pApApG)^3 \) oligomer (prepared by the chemical polymerization of \( d-pA^Bz-pA^Bz-pG^Bz \) in anhydrous pyridine, using the sulphonyl chloride, \( MsSO_2Cl \), there is supposed to be little pyrophosphate produced (183)). The purification steps involved chromatography on a DEAE-cellulose (chloride) column, in the presence of 7 M urea, although from the elution profile (Figure 13), the nonanucleotide peak is not cleanly resolved from other by-products. Also, extended chromatography on Whatman #40 paper in system C (91), still
did not remove a significant amount of impurity, although the major band was clearly separated from a faster moving band and from a fluorescent band which remained at the origin. According to Narang et al. (91) this paper chromatography technique is capable of separating the pyrophosphate, which should move faster than the linear homologue.

However, after these purification steps, the nonanucleotide still contained a major contaminant(s), and based on the retention principle of an oligonucleotide-cellulose column, the suggestion that much of the impurity is the pyrophosphate, is not unreasonable. Therefore, it would seem useful in the chemical synthesis of oligodeoxyribonucleotides to include as a final purification step, chromatography on a complementary oligonucleotide-cellulose column.

k) Elution of RNA from an oligonucleotide-cellulose column, and the selective retention of a complementary oligonucleotide.

As a preliminary experiment related to the ability of an oligonucleotide-cellulose to select a nucleic acid with a complementary base sequence, it was decided to co-chromatograph a mixture of RNA and the oligonucleotide d(pApApG)₃ on cell-d(pCpTpT)₃.

When a sample of RNA⁵⁰ (14 A.U. in 1 ml of MBS) was applied

⁵⁰The RNA used was that isolated from the liver of chinook salmon (Onchorynchus tschawytscha) (see Methods).
to a cell-d(pCpTpT)₃ column at -4⁰, a large peak of RNA was eluted at the front. However, a considerable amount of material was retained, and eluted between 0⁰ and 35⁰ (Figure 34). These conditions of high salt (M NaCl) and low temperature (-4⁰) are known to be ideal for the precipitation of rRNA (for example, 184). If the retention of the RNA by the column is non-specific due to precipitation, as the temperature is slowly increased, the RNA should go back into solution and be eluted. That this RNA was retained due to a precipitation phenomenon is further substantiated by the observation that repeated attempts to chromatograph the RNA resulted in widely varying amounts of RNA being retained, both on a oligonucleotide-cellulose column as well as on a control cellulose column. It was therefore necessary to find conditions for loading and eluting the columns, under which a minimal amount of RNA was precipitated, while the conditions would still permit the binding of d(pApApG)₃. The conditions that were finally chosen were loading the sample in MBS at 10⁰. 5¹

The profiles for elution of d(pApApG)₃ and RNA (loading at 10⁰) are given in Figure 35A and B. A mixture of RNA (two different amounts) and d(pApApG)₃ is shown in Figure 35C and D. (For details of these experiments, see Table XVII.) It is clear

5¹ In the isolation of a naturally occurring RNA such as mRNA, one would include a step that would remove a large portion of the rRNA prior to chromatography on a complementary oligonucleotide-cellulose column.
Figure 34. Elution of RNA on cell-d(pCpTpT)$_3$.

The sample was loaded in MBS at -4°C. The arrow indicates where the column was washed with 0.1 N NaOH at room temperature.
Figure 35. Elution of \(d(pApApG)_3\) and RNA, separately, and as a mixture on \(cell-d(pCpTpT)_3\).

A. \(d(pApApG)_3\).
B. RNA.

The arrow indicates where the column was washed with 0.1 N NaOH at room temperature.
Table XVII. Thermal elution of RNA and $d(pApApG)_3$ on cellulose-$d(pCpTpT)_3$. The samples were loaded onto the column at $10^\circ$, and eluted with a linear temperature gradient.

<table>
<thead>
<tr>
<th>Sample</th>
<th>A.U. loaded</th>
<th>Volume of sample</th>
<th>Figure</th>
<th>$T_m^{c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $d(pApApG)_3$</td>
<td>3.7</td>
<td>1 ml</td>
<td>35A</td>
<td>40.5$^\circ$</td>
</tr>
<tr>
<td>2. RNA</td>
<td>9.3</td>
<td>1 ml</td>
<td>35B</td>
<td>--</td>
</tr>
<tr>
<td>3. RNA + $d(pApApG)_3$</td>
<td>9.3 + 3.7</td>
<td>1 ml</td>
<td>35C</td>
<td>40.5$^\circ$</td>
</tr>
<tr>
<td>4. RNA + $d(pApApG)_3$</td>
<td>84 + 3.7</td>
<td>9 ml</td>
<td>35D</td>
<td>40.5$^\circ$</td>
</tr>
</tbody>
</table>
that the nonanucleotide d(pApApG)_3 is retained when loaded at 10°, even in the presence of a large load of RNA. Also, the d(pApApG)_3 peak is eluted at exactly the same temperature as when no RNA is loaded. The UV-absorbing material eluting between 10° and 35° seems to be contributed to by both the oligonucleotide sample as well as the RNA.
CONCLUSIONS

In the General Introduction a method was proposed for the isolation of specific polynucleotides by hybridization with a synthetic oligonucleotide of defined structure which is complementary to a short sequence within a polynucleotide. To facilitate the separation, it was proposed that the oligonucleotide be covalently linked to an insoluble matrix such as cellulose. The experiments reported here are a study, with chemically synthesized model compounds, of the practicability of such a method.

In part I, the synthesis and purification of complementary homo-oligodeoxyribonucleotides and oligonucleotides of repeating trinucleotide sequences (up to 12 nucleotides long) have been described. The general procedures used in these syntheses were those developed by Khorana and his associates (75). Using these methods it was found that yields of 1 to 10 μmoles of the desired oligonucleotides could readily be obtained. These amounts are suitable for the synthesis of oligonucleotide-celluloses.

In experiments not reported here, the block synthesis of oligonucleotides of defined structure has also been accomplished, and again yields in the range of 1 to 10 μmoles were readily obtainable.

In part II, methods to attach oligonucleotides via a terminal phosphodiester linkage to cellulose, were investigated.
It was found that a modification of the water-soluble carbodiimide method reported by Gilham (141) consistently resulted in high levels of coupling of the oligomer to cellulose paper. Under the conditions of this reaction, one minor byproduct, the carbodiimide substituted deoxythymidylate, was detected, at a maximum level of 2% of the total nucleotide. However, conditions are available for the convenient reversal of the substitution (160).

Complementary oligonucleotides (both homo-oligonucleotides and oligomers of mixed base sequence), from 6 to 12 nucleotides long, form stable complexes with the appropriate oligonucleotide-celluloses, under conditions suitable for hybrid formation (neutral pH, 1 M NaCl, low temperature). The stability of these complexes is related to the total number of base interactions possible. Oligodeoxyriboonucleotides which are capable of forming GC base pairs are more stable than the same length oligomers which contain only adenylate and thymidylate residues.

The capacity of oligonucleotide-celluloses is 1/3 to 1/4 of the theoretical.

Because of the difference in thermal stability of the different oligomers, consecutive oligonucleotides of a particular series, d(pA)_n, may be resolved, partially, or completely, by chromatography on a complementary oligonucleotide-cellulose, eluting with a linear temperature gradient. When the ratio
of bound oligomer (cellulose-bound) to free oligomer is sufficiently high (2 μmoles bound, to 0.1 μmole free), consecutive oligomers are completely resolved.

Evidence is presented which suggests that probably all the nucleotides of a particular oligonucleotide-cellulose are capable of hydrogen bonding with their complementary bases.

The minor variability in retention properties of different preparations of a particular oligonucleotide-cellulose has not been accounted for.

Finally, in all the interactions studied here, complementary strands have consisted of either entirely pyrimidines or entirely purines. Since the vertical stacking of oligopurine tracts is known to contribute additional stabilization to hybrid interactions (145), it would be interesting to investigate the thermal stability of some complementary strands which each contain both pyrimidines and purines. In order to draw useful comparisons with the present data, an obvious complementary set of oligonucleotides that should be prepared are d(pGpTpT)ₙ and d(pApApC)ₙ.

In the application of oligonucleotide-celluloses to the isolation of polynucleotides, a number of experiments are suggested.

First, it is necessary to obtain some firm conclusions as to the steric factors involved in oligonucleotide-cellulose: polynucleotide interactions. It is expected that at least the "tail" on the 3' end of a polynucleotide will destabilize (and
possibly prevent) interaction with the cellulose-bound, complementary oligomer. The design of the oligonucleotide-cellulose may have to be altered to accommodate this.

In a test of the steric interference of the 3' "tail", either a synthetic polynucleotide (prepared enzymatically, see page 80) or a naturally occurring polymer in which short tracts of nucleotides are known, may be used. Pyrimidine tracts, including oligothymidylic acid tracts in two viral DNA molecules (T7 and λ) have been reported, recently (14, 15). The frequency and strand distribution of the oligothymidylic acid tracts in these molecules is discussed in the General Introduction. Short tracts of pyrimidines have also been determined in the minus strand of øX 174(185).

Once the necessary steric easement of the oligonucleotide-cellulose has been defined, the isolation of a natural polynucleotide from a heterogeneous population of molecules could be attempted. An obvious first choice would be the attempted isolation of phage T4 lysozyme messenger ribonucleic acid. T4 lysozyme is a small protein (164 amino acids), which is produced late in the infection cycle. The amino acid sequence has been determined (186). Streisinger and his group have described the isolation of acridine-induced double frameshift mutants of phage T4, which make pseudowild lysozyme molecules. From the altered amino acid sequence within the frameshift region,
short tracts of the messenger nucleotide sequence can be deduced (78-81). For example, a sequence within the wild type lysozyme messenger has been deduced. This sequence would be complementary to a heptanucleotide, d-pApCpTpTpTpT. Similarly, the mutant e JD10eJ42eJ17 messenger should contain a tract of nucleotides complementary to the octanucleotide sequence, d-pApCpApCpTpTpTpT. To date, the synthesis of both the heptanucleotide, d-pApCpTpTpTpT, and the nonanucleotide, d-pApC-pTpTpTpTpTpT, has been accomplished, and both have been linked, efficiently, to cellulose, using the modified water-soluble carbodiimide method. Dr. Michael Doel, of this laboratory, has recently begun the synthesis of the octanucleotide, d-pApCpApCpTpTpT. Preliminary experiments related to the isolation of lysozyme mRNA using these oligomers, are in progress.

One problem associated with the isolation of any group of biological molecules is that of having available a specific assay for the molecules. In the case of phage T4 lysozyme, a convenient, in vitro assay for the messenger (involving the synthesis of active enzyme molecules), has been developed (31).
LITERATURE CITED


139. Gillam, I.C., personal communication.
142. Gilham, P.T., personal communication.


151. Peacock, A.C., and Dingman, C.W., Biochemistry, 7, 668 (1968).


160. Ho, N.W.Y., Gilham, P.T., Biochemistry, 6, 3632 (1967).


