Synthesis and Properties of Oligonucleotide–Cellulose Columns†

Caroline R. Astell‡ and Michael Smith*

ABSTRACT: Deoxyribooligonucleotides of known length and sequence have been linked to cellulose paper in high yield by phosphate ester formation, using the water-soluble carbodiimide, 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate. The resultant deoxyribooligonucleotide–cellulose papers were converted into a form suitable for column chromatography. Oligonucleotide–cellulose columns are capable of retaining complementary deoxyribooligonucleotides under conditions which favor hydrogen-bond formation (neutral pH, high salt concentration, and low temperature). These complementary oligomers are conveniently released from the columns by applying a linearly increasing temperature gradient. The eluted nucleotide peaks are symmetrical and can be characterized by their temperature midpoint (T_m°) and range of elution. The capacity of these columns has been estimated to be 0.25–0.33 of the theoretical maximum. However, for complete resolution between adjacent oligomers of a series, a ratio of matrix-bound oligomer to free oligomer of 20:1 is required. Preliminary evidence suggests that the entire oligonucleotide attached to the cellulose is capable of hydrogen bonding with its complementary sequence. Deoxyribooligonucleotide–celluloses also retain ribooligonucleotides although the stability of these hybrids is less than that of the corresponding deoxyribooligonucleotide. A 5′-terminal phosphate lowers the T_m° of an oligomer by approximately 1–2°.

It is probable that immunochemical precipitation of specific polysomes (Williams and Askonas, 1967; Holme *et al.*, 1971; Uenoyama and Ono, 1972; Palacios *et al.*, 1972; and Palmiter *et al.*, 1972) will provide the basis of a general technique for isolation of mRNA in a translatable form. However, the high degree of specificity which can be achieved by molecular hybridization has also been used for the isolation of several mRNAs (Hayashi *et al.*, 1963; Attardi *et al.*, 1963; Imamoto and Yanofsky, 1967; Stubbs and Hall, 1968). We suggest that it should be possible to isolate a specific mRNA by hybridization involving only a very short length of the polynucleotide chain. The length of a complementary oligonucleotide sequence required to isolate a unique species of mRNA in an organism of given genetic complexity (Table I) is within that to which synthetic methods for deoxyribooligonucleotides can be applied (Khorana, 1968). Consequently, we have initiated studies directed at a general procedure for the isolation of a specific mRNA molecule by hybridization with a matrix-bound synthetic oligonucleotide which is complementary to a partial sequence in the mRNA.

A first requirement for this approach is an *efficient* method for linking one *end* of an oligonucleotide of defined sequence to an insoluble matrix. Second, a complete definition of the interaction of the oligonucleotide–matrix with complementary oligonucleotides is required. The present communication describes efficient synthesis of a variety of deoxyribooligonucleotide–celluloses linked by esterification of an oligonucleotide 5′-phosphate to cellulose. The interactions of these deoxyribooligonucleotide–celluloses with complementary oligonucleotides suggest that the methodology will be useful in the isolation of specific polynucleotides.

Materials and Methods

Nucleotides. The oligonucleotides used were prepared as described previously (Astell and Smith, 1971). The oligomer pd(A-C-T-T-T-T-T) was prepared by standard chemical synthesis (Astell, 1970).

Preparation of Oligonucleotide–Cellulose. The 5′-phosphate-terminated deoxyribooligonucleotides were linked covalently to cellulose paper using a water-soluble carbodiimide (Gilham, 1968, 1971). The method was modified by increasing the concentration of buffer and, most important, the amount of reagent used, resulting in incorporation consistently greater than 60%, even for the longer oligomers. Approximately 500–1500 nmoles (50–150 A.U.[1]) of oligonucleotide (sodium salt) was suspended in 0.5 ml of 0.04 M Mes buffer (pH 6.0). CMC, 120 μmoles (50 mg), was added, and the solution streaked evenly on to a strip of Whatman No. 3MM paper (50 mm × 120 mm, approximately 1 g). [Previously, the Whatman No. 3MM paper was washed in 10^{-3} M EDTA (pH 7.0) and then water and air-dried. The papers were then placed in MeOH–concentrated HCl (99:1) for 3 days, washed with water, 0.05 M Mes buffer (pH 6.0), and water, and air-dried. This treatment was designed to methylate carboxyl groups which might be present in the cellulose (Gilham, 1971).] The paper was air-dried and left at room temperature for 24–48 hr. In most cases, the paper was streaked with another 50 mg of CMC in 0.5 ml of Mes buffer (0.04 M, pH 6.0) and left a further 48 hr. Unincorporated oligomer was eluted from the paper by descending chromatography (0.05 M sodium phos-

† From The Department of Biochemistry, Faculty of Medicine, University of British Columbia, Vancouver 8, British Columbia, Canada. *Received April 28, 1972.* This research was supported by a grant from the Medical Research Council of Canada. A preliminary report of some of this work has appeared (Astell and Smith, 1971).

‡ Recipient of a Killam Predoctoral Fellowship. Present address: the Rockefeller University, New York, N. Y. 10021.

* Author to whom correspondence should be addressed. Medical Research Associate of the Medical Research Council of Canada.

[1] The abbreviations and definitions used are as follows. For nucleotides and oligonucleotides and other reagents, the IUPAC–IUB recommendations are used; cellulose-p(dT)$_8$, an octathymidylate linked to cellulose by a 5′-phosphate; Mes buffer, sodium 2-(*N*-morpholino)-ethanesulfonate; CMC, 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate; A.U., the amount of material with an absorbance of 1.0 when dissolved in 1 ml of solvent in a 1-cm light path (the absorbance is determined at the wavelength of maximum absorbance); T_m°, the temperature at the midpoint of a peak of oligonucleotide eluted from an oligonucleotide-cellulose column; the range of elution is the temperature span over which the oligonucleotide is eluted; MBS, 1 M NaCl in 0.01 M sodium phosphate (pH 7.0).

TABLE I: Minimum Length, n, of Unique Base Sequences in mRNA Derived from Genomes Containing 3×10^3 to 3×10^9 Base Pairs.[a]

Organism (Base Pairs)	4^n	n
	4	1
	16	2
	64	3
	256	4
	1,024	5
R17 (3300);[b] MS-2 (4000);[b] ϕX 174 (5100);[b] SV40 (7000)	4,096	6
	16,384	7
T7 (40,000); λ (50,000)	65,536	8
T2, T4 (2×10^5)	262,144	9
	1,048,576	10
	4,194,304	11
E. coli (4.5×10^6)	16,777,216	12
Yeast (2×10^7)[c]	67,108,864	13
	268,435,456	14
Teleosts (4×10^8)[c]	1,073,741,824	15
Amphibians (1×10^9)[c]		
Reptiles (1.5×10^9)[c]		
Mammals (3.2×10^9)[c]	4,294,967,296	16

[a] The minimum length, n, for a unique oligonucleotide transcribed from a given genome is obtained when 4^n equals the number of base pairs in the genome. Thomas (1966) has presented similar calculations for determining the minimum nonrepeating length in a genome. [b] Base pairs in replicative form of nucleic acids. [c] Base pairs in haploid cells.

TABLE II: Synthesis of Oligonucleotide–Celluloses.

Oligonucleotide[a]	CMC (mg)[b]	Reaction Time (hr)	Amount of Oligonucleotide Incorporated (A.U.)[c]
$p(dT)_6$ (57)	50	26	32 (56)
$p(dT)_9$ (47)	50	26	6 (13)
$p(dT)_9$ (78)	50	24	39 (50)
$p(dT)_8$ (48)	50	48	27 (56)
$p(dT)_9$ (48)	2×50	96	47 (98)
$p(dT)_9$ (72)	2×50	96	69 (96)
$p(dT)_{12}$ (48)	2×50	96	30 (63)
$p(dT\text{-}T\text{-}C)_2$ (90)	50	64	25 (28)
$p(dT\text{-}T\text{-}C)_2$ (124)	2×50	96	97 (78)
$p(dC\text{-}T\text{-}T)_2$ (133)	2×50	96	96 (72)
$p(dT\text{-}T\text{-}C)_3$ (104)	2×50	96	77 (74)
$p(dC\text{-}T\text{-}T)_3$ (132)	2×50	96	103 (78)
$pd[A\text{-}C\text{-}(T)_5]$ (84)	2×50	96	83 (98)

[a] The numbers in parentheses are the amounts (A.U.) of oligonucleotide used in each experiment. [b] When 2×50 mg of CMC was used, 50 mg was added at the start of the reaction and 50 mg after 48 hr as described under Materials and Methods. [c] Numbers in parentheses are the percentage yields.

phate, pH 7.0) and the per cent incorporation was calculated from the amount of nucleotide recovered in the eluate (Gilham, 1968).

Column Chromatography with Deoxyribooligonucleotide–Celluloses. The oligonucleotide–cellulose was cut into 1-mm squares, suspended in MBS, and stirred vigorously to form a pulp. The pulp from 60 cm^2 could be packed under pressure into a jacketed column 9 mm diameter $\times$ 50 mm. The column was washed with 0.1 N NaOH (rapidly) prior to equilibrating with MBS at $-4°$. Samples of complementary oligomers (3–7 A.U.) in 1 ml of MBS were loaded. The column was washed with MBS at $-4°$ until unbound oligonucleotide was eluted. Hydrogen-bonded oligonucleotide was eluted with the same buffer using a temperature gradient, normally a linearly increasing gradient which produced an increase of 0.5°/1.7-ml fraction per 11 min. This flow rate and temperature gradient was used in all experiments except where noted otherwise.

Results and Discussion

Incorporation of Deoxyribooligonucleotides into Cellulose Paper. Gilham (1968) has reported that phosphoryl groups on nucleotides may be linked to cellulose paper sheets using water-soluble carbodiimide, CMC. Although the reaction went reasonably well with small oligomers, there was a decreasing yield of incorporation as the length of the oligonucleotide increased. In two experiments, $p(dT)_{12}$ was incorporated at a level of only 11.8% and 16.7% (P. T. Gilham, personal communication). In the present experiments it was found that by using a higher concentration of Mes buffer (for better pH control) and twice as much carbodiimide (CMC) in two applications, at the start of the reaction and 48 hr later, incorporation of oligomers was consistently greater than 60% (Table II). For example, the per cent incorporation of $p(dT)_9$ using these conditions was 97 and 92, compared with lower and variable yields of 13 and 50% with less carbodiimide (Table II). We feel that the factor most important in limiting the incorporation of oligomers under the improved conditions is the purity of the synthetic oligomers.

Possible Side Reactions Resulting from the Reaction of Nucleotides with CMC. Because the integrity of the groups involved in hydrogen bonding between complementary oligomers is essential, it is important that the coupling reagent does not react with the nucleotide bases. Gilham (1962) and Naylor *et al.* (1965) have reported that CMC reacts specifically with thymine, uracil, and guanine nucleoside residues at pH 8.0 to form the CMC adducts. However, below pH 7.0, the reagent did not participate in a base addition reaction (Ho and Gilham, 1967). Because we used a high molar ratio of CMC to nucleotide (up to 250:1) and extended reaction times (up to 96 hr), the reaction with mononucleotides under our conditions was examined to see if any CMC adducts were formed. We were able to detect a small amount of by-product (1–2%) when CMC was incubated with deoxythymidine 5′-phosphate (Table III). This product was identified as the CMC adduct described by Ho and Gilham (1967) by chromatography, high-voltage electrophoresis, and spectral data (Astell, 1970). In agreement with Naylor *et al.* (1965), deoxyadenosine and deoxycytidine 5′-phosphates did not undergo a base addition reaction.

If 2% of the thymidine residues in a preparation of cellulose–$p(dT)_9$ are substituted with CMC, it can be predicted that 17% of the oligomers will be substituted with the carbo-

TABLE III: Reaction of pdT, pdA, and pdC with CMC.[a]

Nucleotide	Amount (μmoles)	Buffer	Time (hr)	CMC (mg)	Yield of CMC-Substituted Derivative (%)
pdT	20	2 ml of 0.1 M sodium borate, pH 8.0	7	68	99
pdT	20	2 ml of 0.1 M Mes, pH 6.0	7	68	~1–1.5
pdT	20	2 ml of 0.1 M Mes, pH 6.0	24	68	~1–1.5
pdT	20	2 ml of 0.1 M Mes, pH 6.0	96	50 mg at time zero, 50 mg at 48 hr	~2
pdA	20	2 ml of 0.1 M Mes, pH 6.0	96	50 mg at time zero, 50 mg at 48 hr	None detected
pdC	20	2 ml of 0.1 M Mes, pH 6.0	96	50 mg at time zero, 50 mg at 48 hr	None detected
pdT[b]	10	See Methods	24	50 mg	~1

[a] The nucleotides were allowed to react with CMC under the conditions described in the table. At the end of the incubation the samples were diluted to 10 ml with water, loaded onto a DEAE-cellulose column (OAc^-) (12 mm diameter × 400 mm), and eluted with a linear gradient of triethylammonium acetate (0.01–0.15 M, pH 7.1, 1400 ml). [b] 10 μmoles of pdT was incorporated onto cellulose paper in a normal incorporation reaction (see Methods). 53% of the nucleotide was linked covalently to the cellulose. Of the 47% eluted from the paper and run on a DEAE-cellulose column (OAc^-), 1% (47 nmoles) was recovered as the CMC adduct of pdT.

diimide (Steele and Torrie, 1960). However, in agreement with the work of Gilham and his coworkers, we found that at an alkaline pH, the adduct is completely degraded to deoxythymidine 5′-phosphate within 24 hr. Therefore, in some experiments, the oligonucleotide–cellulose papers were incubated in 0.2 M sodium carbonate buffer (pH 10.9) for 20 hr, then washed and converted into a pulp (Table IV).

Properties of Oligonucleotide–Cellulose Columns. As noted earlier the preparations of oligonucleotide–cellulose obtained from oligonucleotide covalently attached to 60 cm^2 (1 g) of cellulose paper yield 9 mm diameter × 50 mm columns. When an aliquot of complementary oligonucleotide (3–7 A.U.) was loaded onto such a column, a small amount of nucleotide was not bound and was eluted by washing the column with MBS at −4°. The bound oligonucleotide could be recovered quantitatively by raising the temperature. This was accomplished either by stepwise increments or more usually by a linear temperature gradient. In one early experiment using thermal elution, $p(dA)_8$ was eluted from a cellulose–$p(dT)_8$ column over a stepwise temperature range of 5–30° (Figure 1a). It appeared that the column was subractionating the sample of $p(dA)_8$. However, the purity of this nucleotide was checked and found to be >95%. When a similar sample was eluted using a linear temperature gradient (Figure 1b), the $p(dA)_8$ was eluted in a broad symmetrical peak with a temperature range of 24°. The T_m° was approximately 20°. This wide temperature range of elution is not entirely unexpected for oligomer–oligomer interactions. Naylor and Gilham (1966) and Astell (1970) have observed that in oligomer–oligomer interactions in solutions, the melting range is of the order of 20–30°. However, Gilham (1964) has reported that with oligonucleotide–celluloses prepared from a random mixture of lengths of homooligonucleotides, a complementary oligonucleotide is eluted within a narrow temperature range (~5°). The capacity of the oligonucleotide–cellulose used by

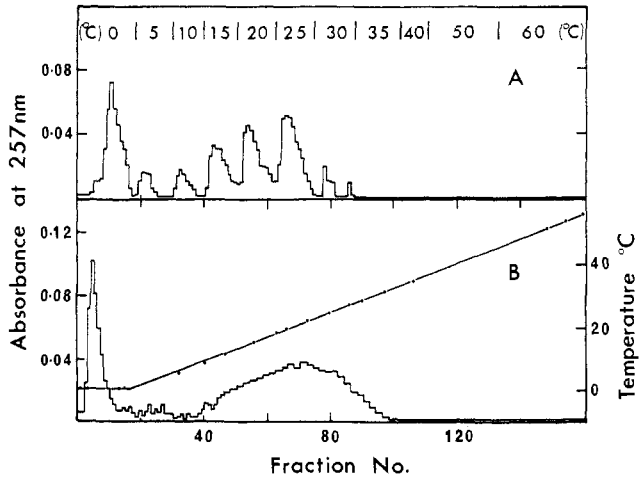

FIGURE 1: Chromatography of $p(dA)_8$ on cellulose–$p(dT)_8$ using (A) a stepwise thermal gradient and (B) a linear thermal gradient. A sample of $p(dA)_8$ (2.6 A.U. in 1 ml of MBS) was applied to a cellulose–$p(dT)_8$ column (27 A. U. $p(dT)_8$; 9 mm diameter × 50 mm) at 0°. The column was washed at a flow rate of 1 ml/22 min. In Figure 1A, the oligonucleotide was eluted by stepwise thermal elution as indicated. In Figure 1B, the oligonucleotide was eluted by a linear temperature gradient. The straight line indicates the temperature of the column. In both A and B, the fraction size was 1 ml.

TABLE IV: Elution of $p(dA)_9$ and $p(dA)_7$ on Different Preparations of Cellulose–$p(dT)_9$.

Cellulose Preparation	A.U. Incorporated	T_m° of $p(dA)_7$	T_m° of $p(dA)_9$
1. Cellulose–$p(dT)_9$	39	18.0 (5.5 A.U.)	32.0 (5.6 A.U.)
2. Cellulose–$p(dT)_9$	48.5	14.0 (5.7 A.U.)	29.5 (5.3 A.U.)
3. Cellulose–$p(dT)_9$ (treated at pH 10.9 for 20 hr)[a]	48.5	14.0 (5.5 A.U.)	29.5 (4.4 A.U.)
4. Cellulose–$p(dT)_9$ (treated at pH 10.9 for 20 hr)	69	19.0 (5.1 A.U.)	
5. Cellulose–$p(dT)_9$ (treated at pH 10.9 for 20 hr)	47	16.5 (5.1 A.U.)	

[a] Cellulose preparation 3 was obtained by incubating preparation 1 at pH 10.9 for 20 hr.

Gilham was at least 200-fold greater than the oligonucleotide–celluloses used here (Astell, 1970) and, as reported below, increasing the capacity of a column reduces the temperature range of elution of an oligonucleotide.

The thermal elution of oligonucleotides over a wide temperature range (both stepwise and linear elution) suggests that there is a heterogeneity of binding sites for oligonucleotides on the column. Several types of heterogeneous binding may occur. For example, the oligonucleotides attached to the cellulose may present a range of available nucleotides free to bind with a complementary nucleotide. However, the amount of sample loaded was well below the capacity of the column (discussed below, and in greater detail elsewhere; Astell, 1970).

At very low temperatures and in high salt, less stable interactions may (and probably do) form. For example, two oligoadenylates may each partially bind to a single oligothymidylate. As the temperature is raised, these structures will rearrange to more stable ones, and excess oligonucleotides will be eluted. However, such an explanation would also predict an asymmetrical elution profile of the bound nucleotide. In Figure 2B the nucleotide is released from the column in a symmetrical peak. A study of capacity (see below) indicates that this type of heterogeneity of binding is not the explanation applicable to the present case.

A third type of heterogeneous binding involves triple-strand structures. Arguments are presented later which suggest that these types of structures do not participate in the interactions described in this report. Several reports in the literature indicate that in polymer–oligomer interactions, the T_m increases as the activity (concentration) of oligonucleotide is increased (Magee *et al.*, 1963; Cassani and Bollum, 1969). It has also been reported that if the concentration of only one of the two components is altered (for example, if the concentration of the polymer is kept constant, and the concentration of free oligomer is increased), not only does the T_m increase, but there is also in increase in the slope of the transition, and thus a decrease in range of melting (see Table II in Cassani and Bollum, 1969). It is possible that these phenomena influence the behavior of oligonucleotide–cellulose columns.

Typical Temperature Gradient Elution. As a basis for further discussion, a typical temperature gradient elution profile of an oligonucleotide will be described (Figure 2A). A sample of $p(dA)_8$ was retained (>90%) by a cellulose–$p(dT)_9$ column. When the temperature was raised, the bound nucleotide was eluted as a symmetrical peak with a T_m^c of 26°. The range of elution was 19°. The reproducibility of the T_m^c was ± 0.5° for a particular preparation of the oligonucleotide–cellulose. (However, different preparations of a particular oligonucleotide–cellulose eluted a complementary oligonucleotide at slightly different temperatures, see Table IV.)

When another oligomer, $p(dA)_{11}$, was chromatographed on the same cellulose–$p(dT)_9$ column (Figure 2b), the elution profile (solid line) indicated that $p(dA)_{11}$ was eluted with a T_m^c of 37°. However, there was a leading edge to the main peak. In order to define this "leaking," a sample of $p(dA)_{11}$ which was retained at 29° but eluted at 50° in a stepwise temperature shift (inset, Figure 2B) was rechromatographed on cellulose-$p(dT)_9$. This sample of $p(dA)_{11}$ was eluted as a sharp, symmetrical peak (T_m^c = 37°), with a complete absence of the leading edge (Figure 2B, dotted curve). It seems likely that the leading edge observed with the initial sample of $p(dA)_{11}$ was due to impurities (probably oligonucleotide pyrophosphates). This suggestion is supported by solution T_m data on the purified $p(dA)_{11}$ with $p(dT)_9$ and $p(dT)_{12}$ (Astell, 1970).

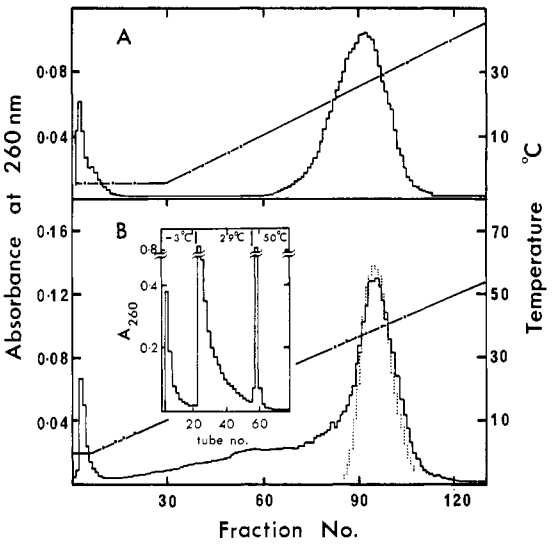

FIGURE 2: Typical elution profiles for oligonucleotide–celluloses. Elution of $p(dA)_8$ (A) and $p(dA)_{11}$ (B) on a cellulose–$p(dT)_9$ column (39 A.U. $p(dT)_9$; 9 mm diameter × 50 mm). The oligonucleotide samples (approximately 5 A.U. in 1 ml of MBS) were slowly applied to the column at −4°. The column was washed at this temperature gradient as described under Materials and Methods. The inset in Figure 2B is the stepwise elution of $p(dA)_{11}$ on cellulose–$p(dT)_9$. The dotted curve in Figure 2B is the elution profile for "oligonucleotide–cellulose purified" $p(dA)_{11}$ (see text).

Alternate Methods for the Elution of Oligonucleotides Interacting with Oligonucleotide–Cellulose Columns. We have tried methods other than thermal elution for the chromatography of complementary oligonucleotides on oligonucleotide–cellulose. If the principle behind these oligonucleotide–cellulose columns involves retention by hydrogen-bond formation, then any conditions which destabilize hybrid structure should be capable of effecting the elution of hydrogen-bonded oligonucleotides. Szybalski (1967) has reviewed the reagents or conditions which can reversibly destabilize hybrid structure. We have examined the use of reverse salt gradients, linear increasing gradients of formamide and *N,N'*-dimethylformamide, as well as an increasing gradient of $NaClO_4$. Although, all of them eluted oligonucleotide, none of the procedures seems to offer an advantage over the thermal elution procedure described above.

Capacity of Oligonucleotide–Cellulose Columns. The capacity of the oligonucleotide–celluloses was examined by chromatographing varying amounts of complementary oligomer, $p(dA)_8$, on a cellulose–$p(dT)_9$ column. Over the range of oligomer chromatographed, (from 5 to 75% of the theoretical amount complementary to the cellulose–$p(dT)_9$ column), it was found that the eluted peak was symmetrical for the smaller amounts, but by the time the column was loaded to 33–50% of its theoretical capacity, an asymmetrical peak was obtained with a leading front edge. The capacity could best be quantitated by comparing the amount of nucleotide not retained by the column against the amount of the theoretical maximum load put on the column (Figure 3). Here it can be seen that the functional capacity of this cellulose–$p(dT)_9$ column is approximately 0.25–0.33 of the theoretical amount.

Resolution of a Series of Oligonucleotides, $p(dA)_n$, on Cellulose–$p(dT)_9$. The resolution of the series $p(dA)_n$ (n = 6–11), was studied on a preparation of cellulose–$p(dT)_9$. A composite elution profile for the series is given in Figure 4A. A tabulated

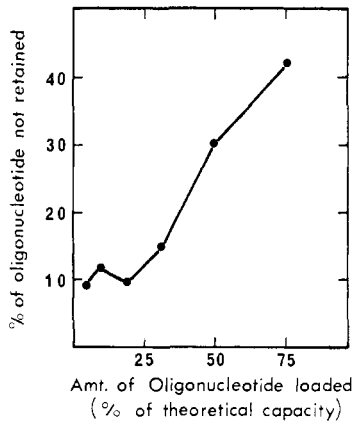

FIGURE 3: Per cent A.U. $p(dA)_8$ not retained by a cellulose-$p(dT)_8$ column. A cellulose-$p(dT)_8$ column (27 A.U. $p(dT)_8$; 9 mm diameter × 50 mm) was loaded with various amounts of $p(dA)_8$ (5–75% of the maximum theoretical capacity of the column. The column was washed and the oligonucleotide eluted as described under Materials and Methods. The amount of material not retained by the column is expressed as a percentage of the oligonucleotide loaded. This is plotted against the load expressed as the percentage of the theoretical capacity of the column.

summary of the T_m^e and range of elution has been presented (see Astell and Smith, 1971). It can be seen that when the $p(dA)_n$ is equal in length to, or shorter than, the cellulose-bound oligonucleotide, oligomers differing in length by 2 nucleotide residues are completely resolved, while oligomers differing in length by only 1 nucleotide are less completely resolved.

In order to determine if the resolution of these columns could be improved, we prepared a column which was approximately three times the usual size [9 mm diameter × 150 mm; 150 A.U., $p(dT)_9$]. The elution of $p(dA)_n$ (n = 6–9) was studied. In this case, successive oligomers of the deoxyadenylate series were completely resolved (Figure 4B).

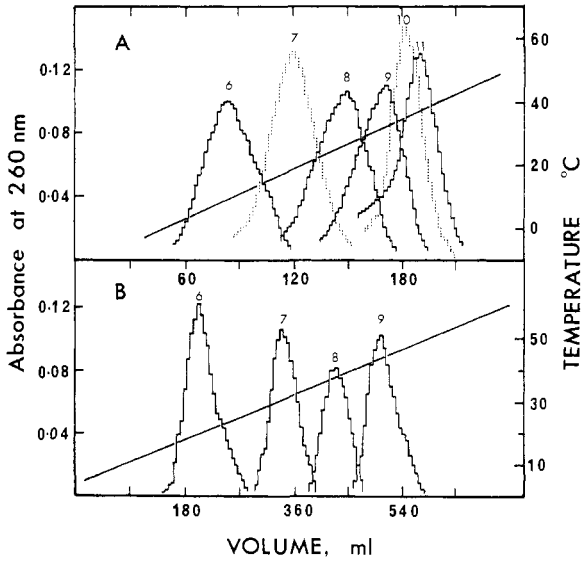

FIGURE 4: Composite elution profile for $p(dA)_n$ on cellulose-$p(dT)_9$. ((A) 9 mm diameter × 50 mm; 39 A.U. $p(dT)_9$ and (B) 9 mm diameter × 150 mm; 150 A.U. $p(dT)_9$). Samples of $p(dA)_n$, n = 6–11, approximately 5–9 A.U., were applied to the cellulose-$p(dT)_9$ columns and eluted with a linear temperature gradient as described under Materials and Methods. For Figure 4B the flow rate was 4.7 ml/11 min per 0.5°. The sample size in A was 1.7 ml and in B 1.6 ml. The number above each peak is the length of the $p(dA)_n$.

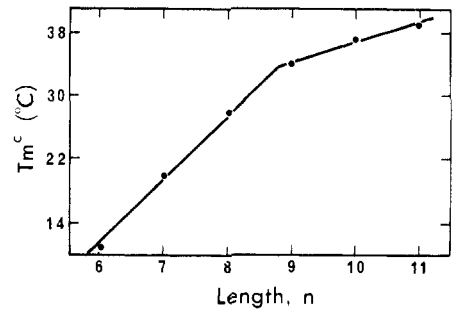

FIGURE 5: T_m^e of $p(dA)_n$ on cellulose-$p(dT)_9$. Samples of $p(dA)_n$ (n = 6 to 11) were loaded onto a cellulose-$p(dT)_9$ column (39 A.U. $p(dT)_9$; 9 mm diameter × 50 mm) and eluted as described under Materials and Methods. The T_m^e values are plotted *vs.* length (n) of the $p(dA)_n$.

In Figure 4B it is interesting to note that even on a larger column (larger both in void volume and capacity), the oligonucleotides are eluted in a relatively smaller volume than on a 9 mm diameter × 50 mm column (Figure 4A). Thus the range of elution is reduced to 8–12°, significantly less than the range of elution on a smaller column (16–19°) (Astell and Smith, 1971). On these larger columns, complete resolution of successive oligomers is achieved when the ratio of bound oligomer to complementary oligomer is approximately 20:1.

Availability for Hydrogen-Bond Formation of Nucleotides in Cellulose-Bound Oligonucleotides. When the length of the free oligomer exceeds that of the bound oligomer, the T_m^e increment between successive oligomers is reduced (Figure 4A). If the T_m^e for each oligomer is plotted *vs.* the length of the oligomer (Figure 5), there appears to be a linear relationship between the two up to the point where the length of the $p(dA)_n$ equals the length of the cellulose-bound $p(dT)_n$. These data indicate that probably the entire oligonucleotide attached to the cellulose is capable of hydrogen bonding with a complementary sequence. More extensive analysis of this observation has been carried out, and the data, supporting the conclusion above, will be published shortly (Astell *et al.*,[2] 1972).

The fact that the T_m^e continues to increase when the length of the free oligonucleotide exceeds the length of the bound, suggests that adjacent bound oligonucleotides participate in the binding of a single oligonucleotide chain. However, these columns have approximately 0.3–1 μmole of bound oligomer/g of cellulose, or fewer than 1 substituted glucose residue in 6000. It is therefore very unlikely that two adjacent oligonucleotides participating in the binding of a single oligodeoxyadenylate, explains the increase in stability of chain lengths 10 or 11 in Figure 5. A possible explanation of this observation might be that there is increased stabilization due to extended stacking of the purine bases, contributing to the overall stability of the interaction.

Interaction of Ribooligonucleotides with Cellulose-$p(dT)_9$. If deoxynucleotide–celluloses are to be used in the isolation of naturally occurring ribonucleic acids, it is important to know how ribooligonucleotides elute from these columns in comparison with deoxyribooligonucleotides. To define the elution of ribooligonucleotides, the series $r(A)_n$ (n = 6–9) was examined on a preparation of cellulose-$p(dT)_9$. For comparison, the dephosphorylated deoxyribo series $d(A)_n$ (n = 6–9) was

[2] Manuscript in preparation.

also studied. The results of these experiments have been tabulated previously (Astell and Smith, 1971; Table I). It is clear from these data that the ribooligonucleotides eluted at 4–9° lower than the corresponding deoxyribooligomers. It is interesting that the phosphorylated $p(dA)_n$ series elutes approximately 1–2° lower than the dephosphorylated series, $d(A)_n$, suggesting that even at a very high salt concentration (1 M NaCl), the 5′-phosphate group on the oligonucleotide is capable of reducing the stability of the hybrid structure. A similar destabilizing effect of a terminal phosphate residue has been observed in oligomer–oligomer interactions (Miles *et al.*[3]) and in oligomer–polymer interactions (Lipsett, 1964).

Characteristics of Different Preparations of Oligonucleotide–Celluloses. We have compared the properties of different preparations of the same oligonucleotide–cellulose (cellulose–$p(dT)_9$), and found that although that $T_m{}^c$ of a particular complementary oligomer is reproducible (±0.5°) with a given batch of oligonucleotide–cellulose, there is variability between different preparations. The $T_m{}^c$ of $p(dA)_7$ and $p(dA)_9$ on five different preparations of cellulose–$p(dT)_9$ is listed in Table IV. A number of possible explanations for this variability have been considered and eliminated (Astell, 1970), and we are unable to offer a suggestion that would explain this result.

Stoichiometry of The Interaction between Complementary Oligomers with Oligonucleotide–Celluloses. We have assumed that the hybrid interactions studied here involve double-stranded structures. Studies with synthetic complementary polymers (adenylate and uridylate and/or thymidylate polymers) have indicated that triple-stranded structures can occur (see Felsenfeld and Miles, 1967, for a review). Similarly, in oligomer–polymer interactions, in certain cases, triple-stranded complexes have been observed (Naylor and Gilham, 1966; Miles *et al.*;[3] Cassani and Bollum, 1969). Cassani and Bollum noted that oligo[d(pT)]:poly(dA) did not form triple strands under their experimental conditions, however, oligo(dA):poly(dT) resulted in three-stranded structures when the length of the oligodeoxyadenylate was less than a critical length which was dependent on the solvent.

Few studies of oligomer–oligomer interactions have been undertaken (Naylor and Gilham, 1966; Miles *et al.*;[3] Astell, 1970). Miles *et al.* reported the presence of triple-stranded structures.

We feel that the formation of triple-stranded structures between oligomers in solution and complementary oligonucleotides attached to the cellulose are, at least for the oligonucleotide–celluloses st ed here, unlikely. In most cases, when triple-stranded s tures form because of steric reasons, the stoichiometry volves two pyrimidine strands and one purine strand (Felsenfeld and Miles, 1967). In the studies reported here, three-stranded structures would need to involve two oligothymidylate chains with one oligodeoxyadenylate. Because fewer than one glucose residues in 6000 is substituted with an oligothymidylate, the chance of two oligopyrimidine chains being sufficiently close to participate in a ternary complex is very unlikely. Therefore, the interactions studied in the present communication are most likely two-stranded structures.

We have prepared cellulose–oligodeoxyadenylates and studied the thermal chromatography of oligothymidylates. In this case, three-stranded structures may participate in the binding of the free oligomers. The results of these experiments will be reported shortly (Astell *et al.*[2]).

[3] H. T. Miles, J. Frazier, and F. M. Rottmann, quoted in Felsenfeld and Miles (1967).

Conclusions

We have suggested that it should be possible to isolate mRNAs by hybridization of a complementary oligonucleotide to a very short length of the polynucleotide chain (Table I). The practical application of this procedure requires that one be able to attach *efficiently* one end of the oligonucleotide to an insoluble matrix (for example, cellulose). Modification of the water-soluble carbodiimide method of Gilham (1968, 1971) for attachment of nucleotides to cellulose through a 5′-phosphate achieved this objective. The reaction was equally facile when the attached nucleotide was deoxyadenylate, deoxycytidylate, or thymidylate. It was possible to attach 1 μmole of oligonucleotide to 1 g of cellulose paper.

It is important to note that the new conditions did not result in reactions involving the heterocyclic bases which could interfere with hydrogen bonding.

Studies of the hydrogen-bonding properties of deoxyribooligonucleotides–celluloses were carried out using complementary deoxyribooligonucleotides. These were used because of their availability through chemical synthesis. However, more limited studies with ribooligonucleotides (Astell and Smith, 1971) showed that the effects are comparable. The lower stability of the ribooligonucleotide interactions confirm those of polynucleotides in solution (Riley *et al.*, 1966).

The experiments reported here demonstrated that the oligonucleotide–celluloses have excellent capacity for complementary oligonucleotides and that the latter can be conveniently eluted from an oligonucleotide–cellulose column using a linearly increasing temperature gradient. Stable interactions were possible with as few as six deoxyadenosine–deoxythymidine or adenosine–deoxythymidine pairs. All of the bases in an oligonucleotide covalently linked to cellulose appear to be available for hydrogen-bond formation. Most importantly, the method is capable of resolving structures differing in length by one base pair. We feel that the system is worth further investigation as a possible tool in mRNA isolation.

Acknowledgments

We are grateful to Dr. P. T. Gilham and Dr. B. D. Hall for encouraging discussions.

References

Astell, C. R. (1970), Ph.D. Thesis, the University of British Columbia, 199 pp.
Astell, C., and Smith, M. (1971), *J. Biol. Chem. 246*, 1944.
Attardi, G., Naono, S., Rouviere, J., Jacob, F., and Gros, F. (1963), *Cold Spring Harbor Symp. Quant. Biol. 28*, 363.
Cassani, G. R., and Bollum, F. J. (1969), *Biochemistry 8*, 3928.
Felsenfeld, G., and Miles, H. T. (1967), *Annu. Rev. Biochem. 36*, 407.
Gilham, P. T. (1962), *J. Amer. Chem. Soc. 84*, 1311.
Gilham, P. T. (1964), *J. Amer. Chem. Soc. 86*, 4982.
Gilham, P. T. (1968), *Biochemistry 7*, 2809.
Gilham, P. T. (1971), *Methods Enzymol. 21*, 191.
Hayashi, M., Spiegelman, S., Franllin, N. C., and Luria, S. E. (1963), *Proc. Nat. Acad. Sci. U. S. 49*, 729.
Ho, N. W. Y., and Gilham, P. T. (1967), *Biochemistry 6*, 3632.
Holme, G., Boyd, S. L., and Sehon, A. H. (1971), *Biochem. Biophys. Res. Commun. 45*, 240.
Imamoto, F., and Yanofsky, C. (1967), *J. Mol. Biol. 28*, 1.
Khorana, H. G. (1968), *Pure Appl. Chem. 17*, 349.

Lipsett, M. (1964), *J. Biol. Chem. 239*, 1257.
Magee, W. S., Jr., Gibbs, J. H., and Zimm, B. H. (1963), *Biopolymers 1*, 133.
Naylor, R., and Gilham, P. T. (1966), *Biochemistry 5*, 2722.
Naylor, R., Ho, N. W. Y., and Gilham, P. T. (1965), *J. Amer. Chem. Soc. 87*, 4209.
Palacios, R., Palmiter, R. D., and Schimke, R. T. (1972), *J. Biol. Chem. 247*, 2316.
Palmiter, R. D., Palacios, R., and Schimke, R. T. (1972), *J. Biol. Chem. 247*, 3296.
Riley, M., Maling, B., and Chamberlin, M. J. (1966), *J. Mol. Biol. 20*, 359.
Steele, R. G. D., and Torrie, J. H. (1960), Principles and Procedures of Statistics, New York, N. Y., McGraw-Hill Book Company, p 353.
Stubbs, J. D., and Hall, B. D. (1968), *J. Mol. Biol. 37*, 289.
Szybalski, W. (1967), *in* Thermobiology, Rose, A. H., Ed., New York, N. Y., Academic Press, p 73.
Thomas, C. A., Jr. (1966), *Progr. Nucl. Acid Res. Mol. Biol. 5*, 315.
Uenoyama, K., and Ono, T. (1972), *J. Mol. Biol. 65*, 75.
Williams, A. R., and Askonas, B. A. (1967), *J. Mol. Biol. 23*, 201.

Determination of the Secondary Structures of Proteins by Circular Dichroism and Optical Rotatory Dispersion†

Yee-Hsiung Chen,‡ Jen Tsi Yang,* and Hugo M. Martinez

ABSTRACT: The rotatory contributions of a protein can be represented by $X = f_H f_H + f_\beta X_\beta + f_R X_R$, where X can be either the ellipticity or the rotation at any wavelength. The The f's are the fractions of the helix (H), β form and unordered form (R); the sum of f's is equal to unity and each f is greater than or equal to zero. With the f values of five proteins obtained by X-ray diffraction studies, the X's of these proteins at any wavelength are fitted by a least-squares method, which defines the X_H, X_β, and X_R. The circular dichroism and optical rotatory dispersion for the helix, β, and unordered form thus determined can be conversely used to estimate the secondary structure of any protein with X's at several wave lengths for the same equation. Results of tests on five additional proteins were satisfactory to good. The calculated spectra were also in satisfactory to good agreement with the experimental results. The β and unordered forms have the same ellipticity and rotation near 222 and 233 nm, respectively, and a very small b_0 of the Moffitt equation. Thus, as a first approximation, the helical content of a protein varies linearly with $[\theta]_{222}$, $[m]_{233}$, or b_0. The present approach is more realistic than previous methods using synthetic polypeptides as reference compounds.

Since the discovery of the optical activity of α helix (Doty and Yang, 1956; Yang and Doty, 1957) and the introduction of the Moffitt equation (Moffitt and Yang, 1956), optical rotatory dispersion (ORD) and, more recently, circular dichroism (CD) have been widely used for studying the conformation and conformational change of proteins and polypeptides in solution. At first, the use of ORD for estimating the helical content of a protein molecule was only suggestive, since X-ray diffraction studies had not yet revealed the three-dimensional structure of any protein (Crick and Kendrew, 1957). The current methods for estimating the helical content of a protein molecule such as the b_0 of the Moffitt equation, the reduced mean residue rotation at 233 nm, and the mean residue ellipticity at 222 nm are based on the use of synthetic polypeptides as model compounds and, furthermore, they completely neglect the optical activity of other structural elements such as the β form (see, for example, Yang, 1969). Recently, Greenfield *et al.* (1967) have attempted to fit the experimental ORD of proteins with a combination of the ORD of pure helix, β form, and random coils as deduced from synthetic poly(L-lysine). Magar (1968) has proposed a method of minimizing the variance between experimental and computed ORD curves for the determination of the three forms. Greenfield and Fasman (1969) have used the same principle for the minimization of the CD spectra. In the method of isodichroic points, Myer (1970) has estimated the fractions of the three forms in proteins from the CD measurements at three wavelengths corresponding to the three isodichroic points. In all cases poly(L-lysine) was used as a model compound for the three forms. All these treatments using synthetic polypeptides, however, lead to estimates that could differ considerably from the X-ray results. For instance, the three-dimensional structure of sperm-whale myoglobin showed 77% α helix and no β form (Kendrew *et al.*, 1960). The ORD of myoglobin, however, could best be fitted with a mixture of 54–55% α helix, 35–36% β form, and 10% coils (Greenfield *et al.*, 1967; Magar, 1968), and the isodichroic points method gave the same estimates as the ORD spectrum (Myer, 1970). The computer fit of the CD spectrum of myoglobin gave about 68% α helix and 5–8% β form, which were closer to the X-ray results (Greenfield and Fasman, 1969).

The use of synthetic polypeptides as model compounds for

† From the Cardiovascular Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94122. *Received May 30, 1972.* This work was aided by U. S. Public Health Service Grants GM-10880 and HL-06285. Presented in part at the 56th Annual Meeting of Federation of American Societies for Experimental Biology, Atlantic City, N. J., April 1972 (Yang and Chen, 1972).

‡ Part of a dissertation to be submitted by Y.-H. C. in partial fulfillment for the requirement of the degree of Doctor of Philosophy, University of California, San Francisco.