Enzymatic Synthesis of Oligodeoxynucleotides*

Lucy M. S. Chang and F. J. Bollum

ABSTRACT: The distribution of oligonucleotide products synthesized enzymatically with calf thymus *terminal*-deoxynucleotidyl transferase was examined using various monomer deoxynucleoside triphosphates and several oligodeoxynucleotide initiators. At monomer:initiator ratios of 1 to 3 most combinations of initiator and triphosphate produce distributions containing one to five monomer additions and these products are readily separable. When reaction bias leads preferentially to longer chain-length products, as may be the case in dATP polymerization, the distribution of products can be changed toward oligomeric products by using certain triphosphate derivatives, for example, *N*-acetyl-dATP. Using the general statistical approach described it is possible to obtain sets of block oligonucleotides having the general structure $(X)_m(Y)_n$, where X is any base of an oligonucleotide initiator of chain length m (m is 3 or greater) and Y is any deoxynucleotide residue with n, the number of residues added, usually being 1 to 5. The extension of statistical synthesis to oligomers having more than two bases should produce sequences of greater complexity and interest.

The enzymatic polymerization of deoxynucleoside triphosphates, catalyzed by a deoxynucleotidyl transferase from calf thymus gland, has been described in earlier publications from this laboratory (*cf.* Kato *et al.*, 1967). The enzyme is fairly nonspecific with regard to base (Lefler and Bollum, 1969; Zmudzka *et al.*, 1969a,b; Hayes *et al.*, 1968; Hansbury *et al.*, 1970), but is quite stringent in its utilization of deoxyribose derivatives. The polymers derived from the action of this enzyme exhibit a narrow molecular weight distribution, and the mean molecular weight is controlled by regulating the ratio of monomer deoxynucleoside triphosphate to initiator oligodeoxynucleotide (Bollum *et al.*, 1964), at least at the high ratios used for polymer synthesis.

The experiments described here were designed to examine the distribution of products formed at low monomer:initiator ratios. The initial experiments with the terminal deoxynucleotidyl transferase activity demonstrated that oligodeoxynucleotide synthesis (Bollum, 1962) occurred by repeated addition to the initiator. The present investigation demonstrates some generalities that extend the use of this enzyme for oligomer synthesis. These results are important for three reasons. First, they demonstrate that essentially all initiator molecules in solution can participate in the chain-extension reaction. This means that enzyme–initiator interaction is (almost) statistical rather than stoichiometric and explains why the degree of polymerization is determined by the monomer:initiator ratio at low as well as high ratios. Secondly, oligomers of intermediate chain lengths, say up to 12 nucleotides and with uniform or mixed sequence, can be readily obtained by enzymic synthesis. Enzymic synthesis has advantages over enzymic degradation of polymers where it is difficult to control the extent of degradation, and over chemical synthesis where intermediate-chain-length products are obtained in poor yield. Finally, it is possible to use synthesis and isolation for preparation of certain defined oligomer sequences. The particular limitation of this synthetic approach is the fact that yield of product is distributed over several products rather than discrete, but it has the advantage that no oligonucleotide material is destroyed in the synthesis. All material is isolated either as useful product or unchanged initiator. This is the primary practical finding of this work.

Materials and Methods

Terminal transferase was prepared as previously described (Yoneda and Bollum, 1965), and had a specific activity of about 15,000 dATP[1] units/mg (about 45,000 dCTP units/mg). It was generally used at about 750 units/ml in the final reaction. Electrophoretically purified DNase I was purchased from Worthington Biochemicals.

Deoxynucleoside triphosphates were synthesized by the morpholidate procedure described by Moffat and Khorana (1961) and were purified by chromatography on Dowex 1-Cl at pH 4.7 and DEAE-HCO_3 at pH 8. They were used as potassium salts. N,Oac_2-dATP[2] was prepared from N,Oac_2-dAMP by the imidazolate procedure of Hoard and Ott (1965). It is not possible to obtain completely selective O deacylation by alkaline hydrolysis and so the final product, containing Nac-dATP and dATP, was purified by preparative chromatography on Whatman No. 3MM sheets developed with ethanol–1 M ammonium acetate (1:1). Radioactive dNTPs were purchased from Schwarz BioResearch. They were used at specific activities of 1000–6000 cpm/nmole.

Oligodeoxythymidylates were synthesized as described by Khorana *et al.* (1962). Oligo(dA), oligo(dC), and oligo(NacdG) were prepared by separation of DNase I digests of the appropriate polymer on DEAE (Cl) columns. [^{3}H]-Oligodeoxythymidylates and [^{3}H]oligodeoxyadenylates were similarly prepared by degradation and separation of digests of radioactive polymers. The [^{3}H]dA oligomers had a specific

* From the Department of Biochemistry, University of Kentucky Medical School, Lexington, Kentucky 40506. *Received August 19, 1970.* This research was supported by Public Health Service Research Grant No. CA-08487 from the National Cancer Institute. The present work is paper VI in a series entitled Deoxynucleotide-polymerizing Enzymes of Calf Thymus Gland. The previous paper in the series is Deoxynucleotide-polymerizing Enzymes of Calf Thymus Gland. V. Homogeneous Terminal Deoxynucleotidyl Transferase.

[1] The abbreviations used are those approved by OBN-IUPAC and Biochemistry. In addition Nac is used for *N*-acetyl substitution and Nan is used for *N*-anisoyl substitution on certain nucleotide derivatives. The ratio *M/I* is used to designate the molar ratio of monomer deoxynucleoside triphosphate to initiator oligodeoxynucleotide.

[2] We are indebted to Mrs. Constance F. Lefler for this preparation.

TABLE I: Initiator Utilization.[a]

Initiator	Concn (μM)	M/I	μM Oligomer in Acid-Insoluble Product				Fraction of Oligomer Utilized at 160 min
			10 min	30 min	60 min	160 min	
A [^{3}H]d(pA)$_3$	13.6	206	5.03	10.06	12.1	13.6	1.00
	40.7	69	8.55	22.0	28.5	39.1	0.96
	67.8	41	9.49	29.2	41.4	58.3	0.86
	95.0	29	8.55	31.35	46.6	72.2	0.76
	122.1	23	8.56	34.2	58.7	81.9	0.67
B [^{3}H]d(pA)$_4$	12.6	222	8.82	11.2	11.72	12.6	1.00
	37.9	74	14.78	30.32	35.2	37.9	1.00
	63.2	44	13.3	40.5	50.6	62.6	0.99
C [^{3}H]d(pA)$_5$	13.4	209	8.31	10.32	10.74	13.0	0.97
	22.4	125	13.2	15.7	17.92	21.3	0.95

[a] Initiator incorporation for oligodeoxyadenylates was carried out by incubating various concentrations of radioactive initiator in 0.2 M potassium cacodylate (pH 7.2), 4 mM $MgCl_2$, 1 mM mercaptoethanol, 59 units of enzyme, and 2.8 mM dATP in a total volume of 0.125 ml at 35°. 25-μl aliquots were sampled at various times and worked up on filter paper disks (Bollum, 1959).

activity of 8000 cpm/nmole of nucleotide and the [^{3}H]dT oligomers had 5000 cpm/nmole.

The DEAE-cellulose used for column chromatography was the microcrystalline form designated Whatman DE-52. The DEAE-cellulose used in sheet form is Whatman DE-81. These products were purchased from H. Reeve Angel, Co., Clifton, N. J. DE-52 columns were used at either pH 4.7 or 8.1 and developed with linear NaCl gradients. Strips of DE-81 1.5 in. wide were attached to glass frames and developed for 16 hr with 0.5 M NH_4HCO_3. Fractions from columns were read at 260 nm and aliquots counted in 0.4% 2,5-bis-2-(5-*t*-butylbenzoxazolyl)thiophene in toluene after drying on paper disks. DE-81 strips were dried and cut into 0.25-in. segments before counting as above. All samples were counted in 4-dram vials (Kimble 60965 with plastic caps, containing 3 ml of scintillation fluid) that can be inserted into regular size plastic vials normally used with the Packard scintillation counter.

General Procedure

The first set of experiments was to find out how many initiator molecules could participate in chain-extension reactions. This was carried out using a fixed amount of enzyme and monomer while varying the concentration of initiator. If the initiator–enzyme interaction is stoichiometric then only a fixed number of initiator molecules should undergo reaction. If the initiator–enzyme interaction is statistical rather than stoichiometric then all initiator molecules should undergo chain extension, provided of course that reaction conditions are such that each enzyme–initiator complex has a number of chances to react with monomer. This condition specifies a rather large excess of monomer, and so the experiments on initiator utilization were run at monomer:initiator ratios greater than 20 (Table I). Initiator utilization is conveniently measured as incorporation of radioactive initiator into acid-insoluble material. At lower ratios chromatographic procedures were used to detect products.

After the statistical nature of the enzyme–initiator interaction is established then it is possible to use fixed amounts of enzyme, a fairly large number of initiator molecules, relative to enzyme molecules, and a fixed, but low, monomer:initiator ratio to produce distributions of products susceptible to precise analysis by separation into discrete homogeneous fractions. That is, the reaction conditions are set for oligomer synthesis and the product distribution can be analyzed precisely (in most cases) by chromatography.

All of the experiments on product distribution are rather similar, but the details of analysis differ somewhat as a result of availability of labeled reaction components. Reactions were set up with a monomer:initiator ratio in the range of 1 to 3. The reactions were allowed to proceed to completion by overnight incubation and the distribution of oligonucleotide products was examined. With readily available initiators, pT_m, pA_m, and pC_m, 2.5-ml reactions were used and the products were separated on 1 × 15 cm DE-52 columns. The molar ratio of [^{14}C]nucleotide to total nucleotide in each peak is an automatic measure of chain length. Both number-average and weight-average distributions can be computed from this kind of analysis. When only small amounts of initiator were available 0.1-ml reactions were run with [^{14}C]dNTPs and the distributions were analyzed by counting products separated on DE-81 strips. This analysis gives only weight distributions in products, weighting factors must be used for comparison to number distributions, and the class of molecules with zero additions does not score in the analysis. When radioactive dXTPs were not available, radioactive initiators were used and the distribution of initiator in products was analyzed on DE-81 strips as above. This analysis gives a number distribution of products directly, but chain length must be assigned by running appropriate markers.

Results

Initiator Utilization. DINUCLEOTIDES. Earlier work from this laboratory failed to demonstrate dinucleotide initiation

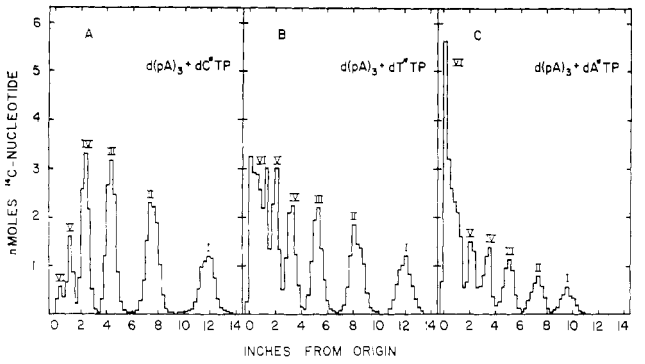

FIGURE 1: DE-81 chromatograms of $d(pA)_3$-initiated polymerizations. (A) [^{14}C]dCTP to $d(pA)_3$ ratio is equal to 1.43. I, A_3C^*; II, $A_3C_2^*$; III, $A_3C_3^*$; IV, $A_3C_4^*$; V, $A_3C_5^*$; and VI, $A_3C_{>5}^*$. (B) [^{14}C]dTTP to $d(pA)_3$ ratio is equal to 1.48. I, A_3T^*; II, $A_3T_2^*$; III, $A_3T_3^*$; IV, $A_3T_4^*$; V, $A_3T_5^*$; and VI, $A_3T_{>5}^*$. (C) [^{14}C]dATP to $d(pA)_3$ ratio is equal to 1.48. I, A_3A^*; II, $A_3A_2^*$; III, $A_3A_3^*$; IV, $A_3A_4^*$; V, $A_3A_5^*$; and VI, $A_3A_{>5}^*$. Since the initiator used is not radioactive, it was not detected on the chromatogram.

(Bollum, 1962). A publication from the Los Alamos group (Hayes *et al.*, 1966) claims utilization of [^{3}H]$d(pT)_2$ and [^{3}H]$d(pA)_2$, and suggests special conditions (0.12 M phosphate buffer, 15°) to improve utilization of these materials for chain initiation. Dinucleotide initiation would be a useful adjunct in specific sequence synthesis and we would be eager to confirm this observation. We have retested [^{3}H]$d(pA)_2$, [^{3}H]$d(pT)_2$, and $d(pNacG)_2$ using our usual conditions and those suggested by Hayes *et al.* (1966). At the concentrations of initiator used in our experiments even a weakly bound initiator should react, but we find no radioactive initiator in polymer and observe no effect on the distribution of products by adding nonradioactive initiator. Since no 5′–3′-exonuclease is present in terminal transferase and our results are on oligonucleotides that are analytically pure, we feel our present and earlier findings are correct.

HIGHER OLIGONUCLEOTIDES. The rate of utilization of homooligonucleotides is related to initiator concentration, initiator chain length and monomer concentration, but the fraction of initiator utilized when reaction is carried to completion depends only on the monomer:initiator ratio. Table I lists initiator incorporation data as a function of time and monomer:initiator ratio for oligodeoxyadenylates of chain length 3, 4, and 5 in the presence of 1.5×10^{-7} M enzyme. The difference in reaction rate of oligomers with different chain length at comparable concentration is interpreted to be the result of binding properties of the enzyme; that is, the affinity of the enzyme for tetranucleotide is much greater than trinucleotide but not much different from pentanucleotide. This interpretation is consistent with the asymptotic K_m data previously published (Kato *et al.*, 1967). With all these oligomers the time course of utilization shows rapid and complete utilization of oligomer when oligomer concentration is relatively low and monomer:initiator ratios are high. With increased concentration of oligomer, shown in Table I, row A with the trideoxyadenylate, more oligomers are incorporated, and eventually all oligomer would probably be incorporated if sufficient monomer is present. At high concentration of oligomer but low monomer:initiator ratios, shown in Table III, row B, not all of the added initiator undergoes reaction, but increased initiator utilization is observed with slight increase in monomer:initiator ratio. The product distribution studies are to be detailed later.

Less extensive data on oligodeoxythymidylate (Kato *et al.*, 1967) and oligodeoxycytidylate (Hayes *et al.*, 1966) incorporation have been published from other investigations. In those experiments complete incorporation of pyrimidine trinucleotide initiators is not achieved even at high monomer:initiator ratios (greater than 100) and long incubation periods. This is presumably a consequence of the lower affinity of the pyrimidine initiator as compared to the products of reaction. Thus a product PyPyPyA, PyPyPyAA, or PyPyPyAAA must have an overwhelming competitive advantage over the PyPyPy initiator. This is not always true, especially in purine-initiated pyrimidine triphosphate polymerizations to be presented later. From the specific activity of the pure terminal transferase (Chang and Bollum, 1971) we estimate that all of the reaction mixtures contain about 1.5×10^{-7} M enzyme and $13.6\text{–}122 \times 10^{-6}$ M initiator. The initiator incorporation studies demonstrate that in every case tested, for each molecule of enzyme 70 or more initiator molecules are converted into acid-insoluble products after 30-min reaction. The usual ratio of initiator to enzyme for polymer synthesis is about 40, with an M/I ratio of 100 to 1000 and reaction time usually more than 12 hr. Under these conditions we can predict that polymer chain length will almost always reflect the M/I ratio, primarily because complete utilization of initiator does occur at early stages of the synthesis. There will be some exceptions with poorly bound initiators (*e.g.*, $d(pT)_3$ or $d(pC)_3$) and a tightly bound product (*e.g.*, poly(dA)). These cases have been described before in our work (Kato *et al.*, 1967) and by Hayes *et al.* (1966).

Distribution of Products. DINUCLEOTIDE INITIATORS. As stated in the previous section we are unable to obtain evidence for the participation of dinucleotides as initiators by initiator incorporation. Nevertheless we have continued to examine this question by looking at the distribution of products at high concentrations of dinucleotides. Chromatographic results obtained in the presence of $d(pA)_2$ and $d(pNacG)_2$ and [^{14}C]dATP, [^{14}C]dCTP, and [^{14}C]dTTP demonstrated only high molecular weight products. Since dinucleotides are not utilized and do not affect the uninitiated synthesis of polymer (K. Kato and F. J. Bollum, unpublished, 1967), we believe that all of the product formed in the presence of dinucleotides is the result of self-initiated synthesis. Since we do not yet know the mechanism of self-initiation, a discussion of this result is not attempted here.

TRINUCLEOTIDE INITIATORS. Purine trinucleotides, $d(pNacG)_3$ and $d(pA)_3$ were examined for product distribution with various dNTPs. The results are presented in Table II, rows A and B and Figure 1. The numbers in the body of the Table II are the fractions of [^{14}C]dNTP in oligomer products, analyzed on DE-81 chromatogram strips. Typical chromatograms with $d(pA)_3$ initiator and several dNTPs are shown in Figure 1. From this set of chromatograms it is quite clear that the final distribution of products is determined by the monomer component. The data presented for $d(pNacG)_3$ cannot be interpreted on relative K_m since these have not been determined for NacG-oligonucleotides. We note, however, that in the case of $d(pA)_3$ more than 50% of the product is in oligomers with less than 4 additions of pyrimidine triphosphates, and greater than 50% is in products with greater than 4 additions of dATP. That is, the formation of a product that binds less avidly than the initiator results in bias toward shorter chain lengths. The opposite is true if the product binds more strongly than initiator.

TETRANUCLEOTIDE INITIATORS. The most intensive study in the present work concerns tetranucleotide initiators.

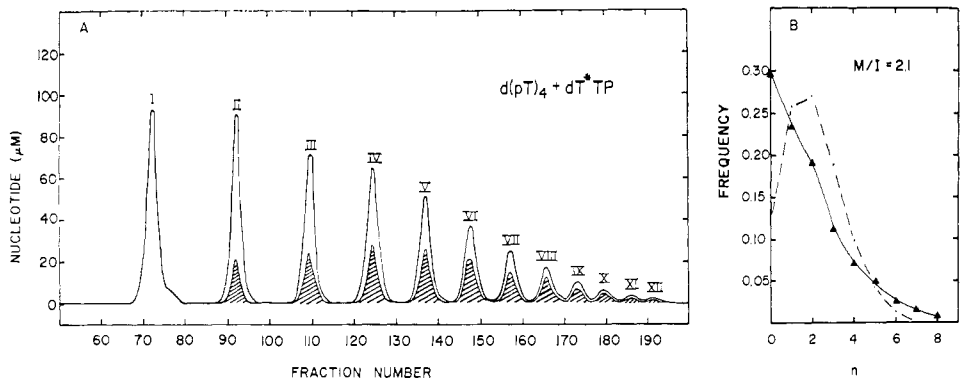

FIGURE 2: Enzymatic synthesis of oligodeoxythymidylates. A reaction mixture containing 0.2 M potassium cacodylate (pH 7.2), 1 mM $CoCl_2$, 1 mM mercaptoethanol, 0.5 mM $d(pT)_4$, 1.04 mM [^{14}C]dTTP (5.1 × 10^5/cpm μmole), and 1875 units of terminal transferase in a total volume of 2.5 ml was incubated at 35° overnight. (A) A 1 × 15 cm DE-52 column was used to separate the products of the $d(pT)_4$-initiated [^{14}C]-dTTP polymerization. A linear 1.2-l. gradient of 0.05–0.5 M NaCl in 0.02 M sodium acetate (pH 4.7) was used to separate the oligonucleotide mixture. Total concentration of nucleotide in each fraction is determined using the extinction coefficients for oligothymidylates of various chain lengths according to Cassani and Bollum (1969), and the concentration of radioactive nucleotide was determined by drying a 0.1-ml aliquot of each fraction on a paper disk and counting in the scintillation counter. The solid line is total nucleotide concentration and the hatched area is radioactive nucleotide concentration. Chain length of each peak is determined by the molar ratio of [^{14}C]nucleotide to total nucleotide. I, T_4; II, T_4T^*; III, $T_4T_2^*$; IV, $T_4T_3^*$; V, $T_4T_4^*$; VI, $T_4T_5^*$; VII, $T_4T_6^*$; VIII, $T_4T_7^*$; IX, $T_4T_8^*$; X, $T_4T_9^*$; XI, $T_4T_{10}^*$; and XII, $T_4T_{11}^*$. (B) A comparison of the number distribution of products of oligodeoxythymidylate synthesis with the Poisson distribution. Monomer to initiator ratio is equal to 2.1. (---) Theory; (▲) found.

TABLE II: Product Distributions.[a]

Initiator	Monomer	*M/I*	Weight Fraction of Product with				
			$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n > 4$
A $d(pNacG)_3$	dTTP	0.83	0.061	0.169	0.127	0.145	0.498
	dCTP	0.72	0.132	0.363	0.270	0.235	0.000
	dATP	0.65	0.077	0.224	0.218	0.150	0.338
B $d(pA)_3$	dTTP	1.48	0.107	0.152	0.145	0.136	0.460
	dCTP	1.43	0.168	0.262	0.217	0.238	0.115
	dATP	1.48	0.061	0.083	0.109	0.113	0.571
C $d(pT)_4$	dTTP	1.35	0.131	0.200	0.207	0.134	0.328
	dCTP	1.30	0.145	0.273	0.330	0.183	0.065
	dATP	1.34	0.072	0.122	0.151	0.118	0.537
D $d(pC)_4$	dTTP	1.26	0.132	0.071	0.034	0.038	0.725
	dCTP	1.21	0.187	0.232	0.194	0.152	0.235
	dATP	1.25	0.056	0.049	0.028	0.060	0.806

[a] The terminal transferase reactions were carried out in a total volume of 0.1 ml. [^{14}C]Monomer and [^{12}C]initiator were used in these experiments. For purine triphosphate polymerization, the reactions contain 0.2 M potassium cacodylate (pH 7.2), 4 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 75 units of terminal transferase. For pyrimidine triphosphate polymerizations, the reactions contained 0.2 M potassium cacodylate (pH 7.2), 1 mM $CoCl_2$, 1 mM 2-mercaptoethanol, and 75 units of terminal transferase. After incubation at 35° overnight, reactions were terminated by addition of 10 μl of 0.2 M Versene. Products of terminal transferase reactions were analyzed by chromatography on DE-81 as described in text. The summation of radioactive monomer incorporated in each peak is divided by the total monomer in product to give the weight fraction of the monomer in each product. *n* equals the number of monomers added onto the initiator.

Two generalizations can be made from these experiments and they apply to trinucleotides as well. First, in homopolymerizations the distribution is biased toward longer chain lengths than predicted by random statistics. This is probably a result of chain-length enhancement of initiator binding. This result is demonstrated in Figure 2A where the column chromatogram and distribution of products (Figure 2B) in the polymerization of [^{14}C]dTTP on $d(pT)_4$ is illustrated. Similar results are obtained with $d(pC)_4$ + dCTP and $d(pA)_4$ + dATP, and the analyses of these polymerizations are shown in Table II, row C and D, and Table III, row C.

In mixed polymerizations different biases moderate the distribution of products. Purine initiators have lower K_m's than pyrimidine initiators. If a purine triphosphate is polymerized on a pyrimidine initiator, the product formed probably results in a shift of binding constant to that characteristic

TABLE III: Initiator Distributions.[a]

Initiator	Monomer	M/I	Number Fraction of Initiator with $n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n > 5$
A [^{3}H]d(pA)$_2$	dATP	15	1.00[b]						
B [^{3}H]d(pA)$_3$	dATP	2.2	0.529	0.091	0.063	0.057	0.044	0.056	0.161
	dATP	2.8	0.486	0.092	0.061	0.056	0.047	0.048	0.211
	dNacATP	2.8	0.353	0.123	0.128	0.136	0.098	0.063	0.105
	dATP	10.8	0.278	0.048	0.031	0.029	0.027	0.027	0.560
C [^{3}H]d(pA)$_4$	dATP	3.3	0.092	0.114	0.127	0.151	0.130	0.386[c]	
	dNacATP	3.3	0.060	0.147	0.287	0.267	0.169	0.079[c]	

[a] Refer to legends in Table II and text for detail. [b] No initiator was found in product. [c] Oligomers with total chain length greater than 9 ($n \geq 5$) did not migrate from the origin.

of purine initiators. The distribution then moves toward longer chain lengths since the products are bound more tightly than the remaining initiator. Since the distribution is affected by base composition at the 3′-hydroxyl end and chain length, utilization of initiator is not complete, and the number of additions is considerably greater than predicted by statistics. Figure 3 (and Table II, row D) demonstrates the distribution resulting from the polymerization of dATP on d(pC)$_4$ at $M/I = 1.25$. In this polymerization more than 80% of the monomer polymerized appears in oligomers with chain length greater than 9. When a pyrimidine triphosphate is polymerized on a purine initiator the products formed change to binding constants characteristic of pyrimidine initiators. Binding of these products are less than the purine initiator. The result is nearly complete utilization of initiator and a distribution showing a different statistical bias. C and T initiators have different affinities and Figure 4 (also Table II, row C) shows the polymerization of dCTP on d(pT)$_4$,

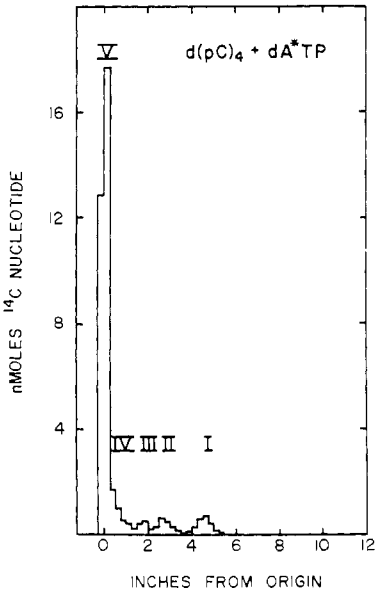

FIGURE 3: DE-81 chromatogram of [^{14}C]dATP polymerization on d(pC)$_4$. Monomer to initiator ratio is equal to 1.25. I, C_4A^*; II, $C_4A_2^*$; III, $C_4A_3^*$; IV, $C_4A_4^*$; and V, $C_4A_{>4}^*$.

indicating the narrow range of products formed in this polymerization at $M/I = 2$.

BLOCKED TRIPHOSPHATES. The results presented thus far indicate that the narrowest distribution of products is obtained when the products are bound less tightly than initiators. This binding is affected by the nature of the heterocyclic base and can also be changed by blocking the amino groups on certain purines. The experiment presented in Figure 5 illustrates the product distributions obtained when dATP and dNacATP are substrates in [^{3}H]d(pA)$_4$ initiated polymerizations. The analysis of this reaction and similar [^{3}H]-d(pA)$_3$ initiated reactions are presented in Table III, rows B and C. It is clear that blocking the amino group on the monomer leads to growing chains with less affinity for enzyme than unblocked A. This results in greater initiator utilization and the products formed are of lower average chain length. This kind of experiment could probably be generalized to the dGTP, dNacGTP pair, but the pronounced tendency of oligodeoxyguanylates to aggregate complicates the chromatographic separation.

Discussion

In an ideal linear condensation polymerization with a fixed number of initiator molecules and equal reactivity of all monomer and initiator molecules the total products should fit a Poisson distribution (Flory, 1953). This specifies that at any stage of the reaction the average chain length of the product is a function of the fraction of monomer polymerized and the monomer:initiator ratio. If monomer and initiator utilization is complete, then at the end of the reaction the average added chain length is just M/I. In our present nomenclature where m = initiator chain length and n = number of residues added, $\bar{n} = M/I$ and

$$\bar{n} = \sum_{n=0}^{n=i} x_i n_i$$

where x_i is the mole fraction of oligomer with n_i additions.

Although it is quite obvious from the outset that interposition of a protein catalyst might have considerable influence on the statistics of the condensation polymerization being considered, we feel that a description of the product distribution in statistical terms is of greater practical value than detailed speculation about the chemistry involved. It should

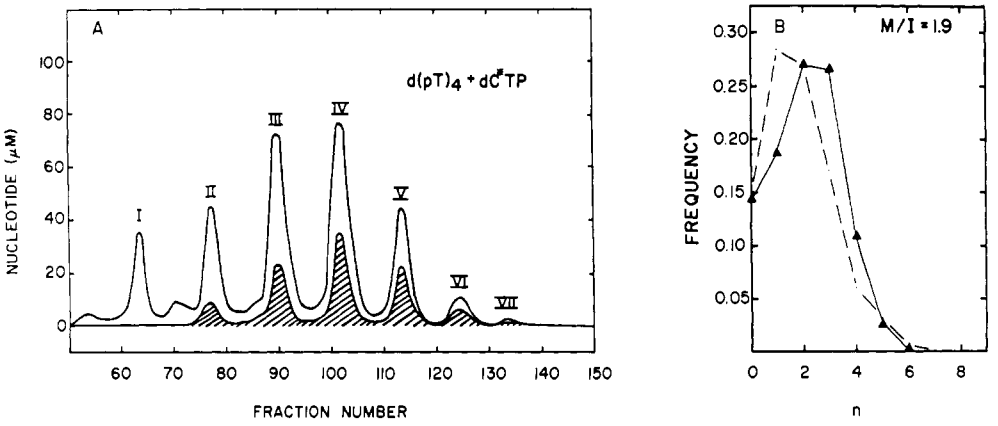

FIGURE 4: Enzymatic polymerization of [^{14}C]dCTP on d(pT)$_4$. The synthesis was carried out by incubation at 35° overnight in 0.2 M potassium cacodylate (pH 7.2), 1 mM $CoCl_2$, 1 mM mercaptoethanol, 0.4 mM d(pT)$_4$, 0.96 mM [^{14}C]dCTP (5.62 × 10^6 cpm/μmole), and 1875 units of terminal transferase. (A) Separation of products was carried out by chromatography on DE-52 at pH 8.0 in 0.05 M Tris-Cl with a linear 1.2-l. gradient of 0.05–0.5 M NaCl. I, T_4; II, T_4C^*; III, $T_4C_2^*$; IV, $T_4C_3^*$; V, $T_4C_4^*$; VI, $T_4C_5^*$; and VII, $T_4C_6^*$. (B) A comparison of the distribution of products of [^{14}C]dCTP polymerization on d(pT)$_4$ with the Poisson Distribution. Monomer to initiator ratio is equal to 1.98. (---) Theory; (▲) found. Refer to Figure 2 and text for further detail.

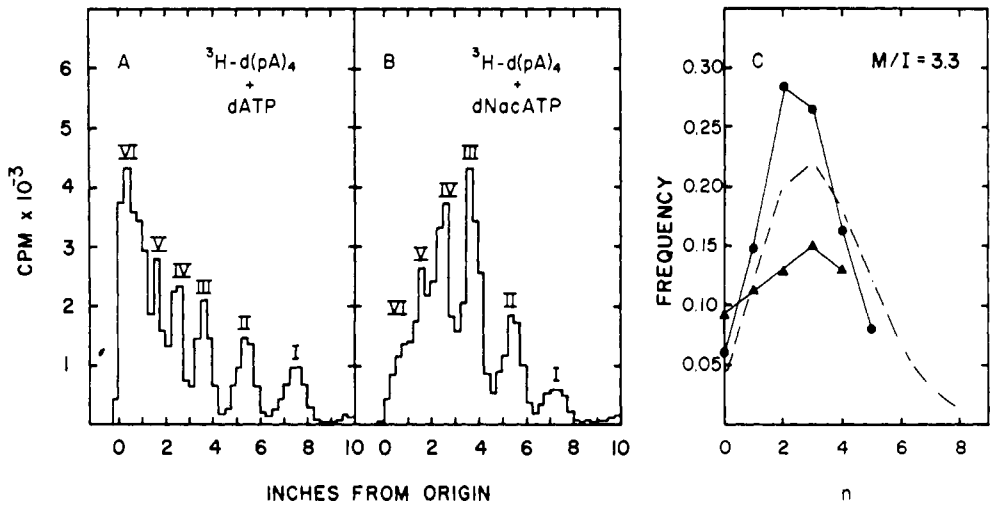

FIGURE 5: Change of product distribution with triphosphate derivative. Terminal transferase reactions were carried out in 0.1 ml under conditions described in the text using [^{3}H]d(pA)$_4$ and dATP or dNacATP. After overnight incubation at 35°, reactions were terminated by addition of 10 μl of 0.2 M Versene, the *N*-acetyl blocking group was removed from the products by addition of an equal volume of concentrated NH_4OH and allowing to stand at room temperature for 6 hr. The NH_4OH was then removed by evaporating the reaction mixture to dryness. The products were redissolved in H_2O before being applied to DE-81 paper. (A) dATP polymerization on [^{3}H]d(pA)$_4$, M/I = 3.3. (B) dNacATP polymerization on [^{3}H]d(pA)$_4$, M/I = 3.3. I, A_4; II, A_5; III, A_6; IV, A_7; V, A_8; and VI A_9. (C) Comparison of product distribution in parts A and B with the Poisson distribution, (---) theory; (▲) dATP plus [^{3}H]d(pA)$_4$; (●) dNacATP plus [^{3}H]d(pA)$_4$.

be clear, however, that with the systems presented here a large amount of chemical data could be generated and analyzed for reaction mechanism.

There are two aspects of the present work that bear some discussion. The first concerns the mechanism of the reaction and the second is the practical application of mechanism. It is clear from all of the results presented that the reaction does not follow random statistics precisely, but the description of mechanism in statistical terms seems appropriate because changing populations of molecules are involved. Under these circumstances we attempt to describe the biases in the reaction that influence the final distribution of products. The reaction begins with a population of initiator molecules homogeneous in base composition and chain length and the reactions occurring influence the nature of the distribution at the end of the reaction. The interaction of enzyme and initiator is determined by intrinsic characteristics of the base and chain length (Kato *et al.*, 1967). We suggest that two factors influence the reactivity of the growing chains: chain length and base composition near the growing end. Thus the "principle of equal reactivity" of growing chains becomes invalid very early during the polymerization, and certain biases develop. The first bias is examined in homopolymerizations, say A_m + dATP; and the second in what may be called copolymerizations, for example, T_m + dCTP. The rules that develop are, first, that in homopolymerizations the chain length distribution will be broad with short-chain initiators (increased binding of growing chains) and possibly Poisson with long-chain initiators (equal binding of all chains). Secondly, that in copolymerizations the distribution of products will be governed by the monomer being polymerized. If, by virtue of the monomer added, the ends of the growing chain bind to the enzyme more avidly than the ends of the initiator chain the distribution will be broadened.

If the ends of the growing chains contain bases that are less tightly bound than initiator chains the distribution will be quite sharp. It may be noted that all of these phenomena occur without a change in the mean chain length of the total distribution of material (that is, unreacted initiator is averaged in), but will affect the average product chain length. We suggest that the practical solution for oligomer synthesis, which means building as narrow a distribution as possible, can often be realized by using monomers that reduce the affinity of the growing chains for the enzyme. In the present experiments this is clearly demonstrated in only one case, by substituting dNacATP for dATP. Since dNacGTP, dNacCTP, or dNanCTP will also substitute for normal substrates, we assume this solution will be generally applicable. As chains are extended the presence of blocking groups will also prevent some aggregation problems. The experiments presented therefore provide a practical description of the rules for statistical synthesis of certain classes of oligodeoxynucleotide.

References

Bollum, F. J. (1959), *J. Biol. Chem. 234*, 2733.

Bollum, F. J. (1962), *J. Biol. Chem. 237*, 1945.

Bollum, F. J., Groeniger, E., and Yoneda, M. (1964), *Proc. Nat. Acad. Sci. U. S. 51*, 853.

Cassani, G., and Bollum, F. J. (1969), *Biochemistry 8*, 3928.

Chang, L. M. S., and Bollum, F. J. (1971), *J. Biol. Chem. 246* (in press).

Flory, P. J. (1953), Principles of Polymer Chemistry, Ithaca, N. Y., Cornell Univ. Press, Chapter VIII, pp 336–339.

Hansbury, E., Kerr, V. N., Mitchell, V. E., Ratliff, R. L., Smith, D. A., Williams, D. L., and Hayes, F. N. (1970), *Biochim. Biophys. Acta 199*, 322.

Hayes, F. N., Hansbury, E., Mitchell, V. E., Ratliff, R. L., and Williams, D. L. (1968), *Eur. J. Biochem. 6*, 485.

Hayes, F. N., Mitchell, V. E., Ratliff, R. L., Schwartz, A. W., and Williams, D. L. (1966), *Biochemistry 5*, 3625.

Hoard, D. E., and Ott, D. G. (1965), *J. Amer. Chem. Soc. 87*, 1785.

Kato, K., Gonçalves, J. M., Houts, G. E., and Bollum, F. J. (1967), *J. Biol. Chem. 242*, 2780.

Khorana, H. G., Vizsolyi, J. P., and Ralph, R. K. (1962), *J. Amer. Chem. Soc. 84*, 414.

Lefler, C. F., and Bollum, F. J. (1969), *J. Biol. Chem. 244*, 594.

Moffat, J. G., and Khorana, H. G. (1961), *J. Amer. Chem. Soc. 83*, 649.

Yoneda, M., and Bollum, F. J. (1965), *J. Biol. Chem. 240*, 3385.

Zmudzka, B., Bollum, F. J., and Shugar, D. (1969a), *Biochemistry 8*, 3049.

Zmudzka, B., Bollum, F. J., and Shugar, D. (1969b), *J. Mol. Biol. 49*, 169.

Acid-Catalyzed Exchange of Hydrogen in Thiamine and Related Compounds*

D. W. Hutchinson

ABSTRACT: The protons of the C(2′)-methyl group of thiamine, oxythiamine, and imidazolethiamine undergo acid-catalyzed exchange in deuterium oxide. The position of exchange has been established by nuclear magnetic resonance studies on the sulfonic acids and 5-(2-hydroxyethyl)-4-methylthiazole produced by bisulfite cleavage of deuterated thiamine and oxythiamine. Confirmation of the position of exchange in thiamine and oxythiamine has been provided by mass spectrometry. A convenient method for the tritiation of thiamine has been developed.

The base-catalyzed exchange of the C(2′) proton of the thiazolium ring of thiamine (Breslow, 1958) has been the subject of a recent study when it was observed that the rate of exchange decreased rapidly with increasing acidity of the reaction medium (Mieyal *et al.*, 1967). In the current investigation, the exchange of the protons of the C(2′)-methyl group of the pyrimidine residues of thiamine chloride hydrochloride (I), oxythiamine chloride hydrochloride (II), and imidazolethiamine (III) has been observed in strongly acid solution.

Materials and Methods

Thiamine chloride hydrochloride (BDH) and oxythiamine chloride hydrochloride (Sigma) were commercially available. Imidazolethiamine (Masuda, 1961) was a gift from Dr. A. J. Knell. Nuclear magnetic resonance spectra were recorded at 60 MHz on a Perkin-Elmer R12 spectrometer, using sodium 2,2-dimethyl-2-silapentane-5-sulfonate as internal standard. The tritium content of samples was measured with a Packard 4000 scintillation counter using a dioxan-based scintillation medium which contained naphthalene (60 g/l.), 2,5-diphenyloxazole (4 g/l.), and 1,4-bis[2-(5-phenyloxazolyl)]benzene (1 g/l.).

Deuterium Exchange of Thiamine (I). The nuclear magnetic resonance spectrum of a freshly prepared solution of thiamine chloride hydrochloride (100 mg) in 6 N deuterium chloride in deuterium oxide (1 ml) was recorded when signals at τ 0.25 (1 H, s), 1.9 (1 H, s), 4.38 (2 H, s), 6.08 (2 H, t, J = 6 Hz), 6.78 (2 H, t, J = 6 Hz), 7.33 (3 H, s), and 7.42 (3 H, s) were observed. The spectrum was rerecorded at intervals over a period of 1 week, during which time the peak at τ 7.33 gradually disappeared ($t_{1/2}$ = 20 hr at room tempera-

* From the School of Molecular Sciences, University of Warwick, Coventry, England. *Received September 9, 1970.*