VOLUME 120, NUMBER 34
SEPTEMBER 2, 1998

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Effects of Monomer and Template Concentration on the Kinetics of Nonenzymatic Template-Directed Oligoguanylate Synthesis

Anastassia Kanavarioti,* Claude F. Bernasconi, and Eldon E. Baird

Contribution from the Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064

Received March 4, 1998

Abstract: To identify key parameters which influence the efficiency of nonenzymatic template-directed oligonucleotide reactions, a kinetic study of oligoguanylate synthesis on a polycytidylate (poly(C)) template has been performed. This oligomerization is satisfactorily described by three kinetic processes: (i) dimerization to form the first primer (k_2), (ii) elongation of a preformed primer (k_i, $i \geq 3$), and (iii) hydrolysis of the monomer to form deactivated material (k_h), with $k_h < k_2 < k_i$. This is the first reported study that includes rate determinations of k_i as a function of the concentration of both poly(C) template and the activated monomer, guanosine 5′-monophosphate-2-methylimidazolide (2-MeImpG), in the range 2 mM $\leq$ [poly(C)] $\leq$ 50 mM and 5 mM $\leq$ [2-MeImpG] $\leq$ 50 mM. k_i values determined under conditions where the template is fully saturated with monomer are practically independent of both monomer and polymer concentration and thus strongly support a template-directed elongation model. Values of k_i determined with a partially occupied template support a mechanism wherein the reaction of the oligonucleotide primer with a template-bound monomer is assisted by the presence of two additional downstream template-bound 2-MeImpG molecules. Comparison between the kinetic parameters obtained here and the ones determined in the montmorillonite-catalyzed oligoadenylate polymerization allows the proposition that the ratio of the rate constants k_i/k_h determines efficiency and the ratio k_i/k_2 determines the degree or extent of a polymerization. These conclusions provide new design principles for the optimization of nonenzymatic polymerizations.

The nonenzymatic template-directed polymerization of activated mononucleotides[1−3] provides a basis for the design of self-replicating systems[4] that could mimic at least one of the necessary chemical reactions for the origin of life.[5] Although polymerizations on templates of mixed sequence have been reported,[6] poly(C) directed oligoguanylate synthesis with guanosine 5′-monophosphate-2-methylimidazolide (2-MeImpG) as the activated monomer provides the highest yield of long (>40 bases) 3′-5′ linked oligomers.[1] The detailed mechanism of nonenzymatic oligonucleotide synthesis is not well understood, and the kinetics are largely unexplored. Given the sequence-dependent limitations[6e] of the template-directed syntheses, we set out to determine kinetic parameters for oligoguanylate synthesis, to determine key mechanistic requirements for efficient polymerizations.

The kinetics of the template-directed reaction of 2-MeImpG at 23 °C in the presence of 0.05 M poly(C) have been described.[7a] A model based on cooperative association of the reactive monomer on the template[8] was developed. In this model oligomerization is defined as a stepwise process with

(1) (a) Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1982**, *162*, 201−218. (b) Inoue, T.; Orgel, L. E. *Science* **1983**, *219*, 859−862. (c) Wu, T. F.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 317−322.

(2) (a) Joyce, G. F. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 41−51. (b) Joyce, G. F.; Orgel, L. E. In *The RNA World*; Gesteland, R. F., Atkins, J. F. Eds.; Cold Spring Harbor Lab. Press: Cold Spring Harbor, 1993; p 1−25.

S0002-7863(98)00723-9 CCC: $15.00
Published on Web 08/15/1998

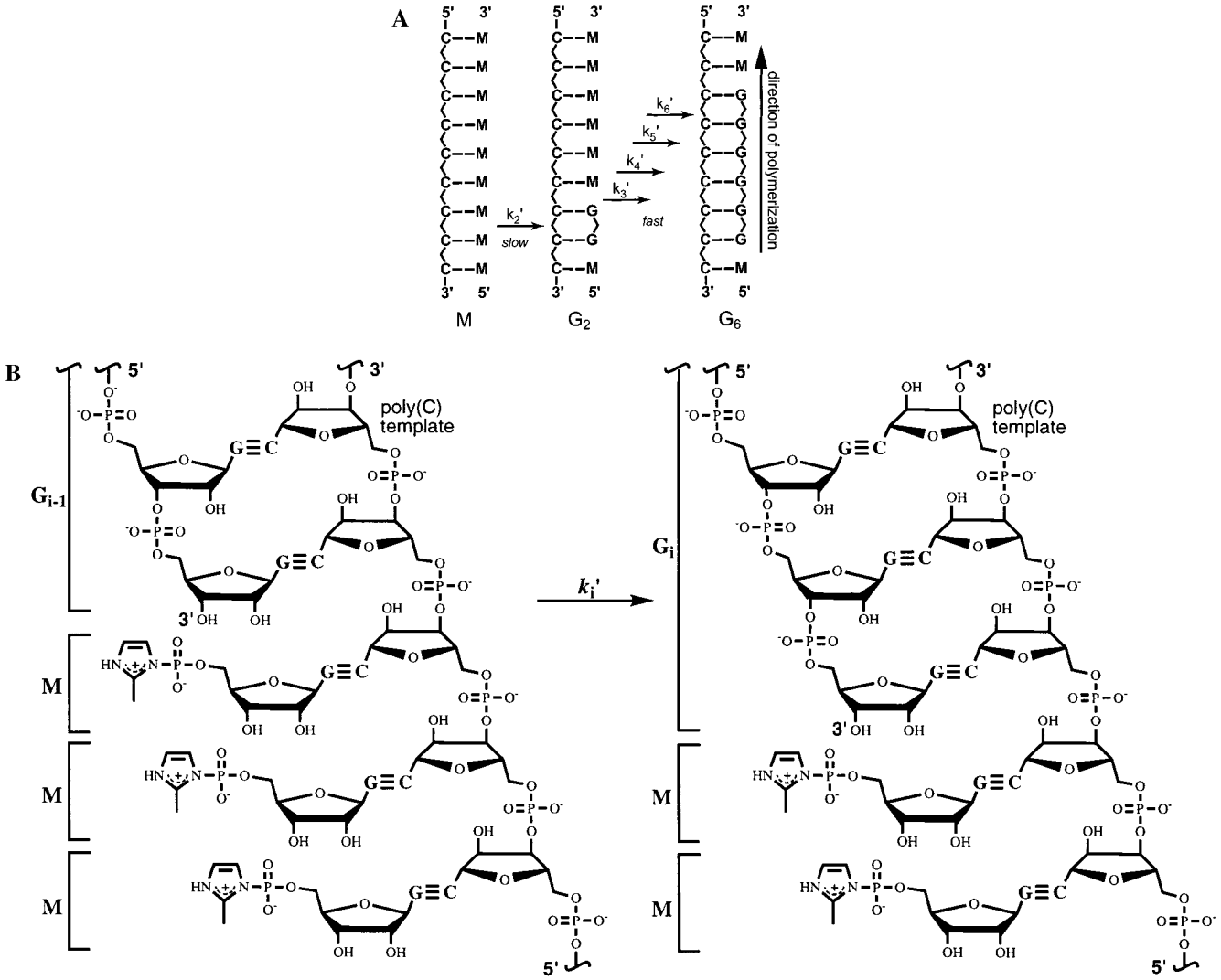


Figure 1. (A) Schematic representation of the stepwise oligoguanylate polymerization on a poly(C) template. Poly(C) is shown as a string of C's. M stands for the reactive guanosine monomer. (B) Elongation of a template bound oligoguanylate G_{i-1} to form the one elongated by one unit, G_i, by nucleophilic attack of the 3′-OH group at the P−N bond of 2-MeImpG, M, and displacement of 2-methylimidazole. Two additional molecules are shown template-bound downstream of the reacting 2-MeImpG monomer to illustrate that this reaction, $G_{i-1} + M \rightarrow G_i$, is facilitated within a complex that consists of the template, the oligomer, and, at least, three template-bound 2-MeImpG molecules.

each consecutive step representing the formation of an oligomer of length i by reaction of 2-MeImpG with an oligomer of length $i - 1$, (eq 1, Figure 1, parts A and B).[7a] In eq 1, M stands for the monomer 2-MeImpG, while G_2, G_3, ..., G_{i+1} are oligomers of length 2, 3,..., $i + 1$, with no distinction made for possible isomers. It is implied that the reacting monomers are template bound and associated at the 3′-end of a growing oligomer. Reaction in this system was shown to occur in the 5′ to 3′ direction only.[1a] k_2', k_3', ..., k_i' are the formal pseudo-first-order rate constants, dependent on M concentration and defined by eqs 2 and 3.

$$M \xrightarrow{k_2'} G_2 \xrightarrow{k_3'} G_3 \xrightarrow{k_4'} G_4 \xrightarrow{k_5'} G_5 \xrightarrow{k_6'} G_6 \longrightarrow\longrightarrow G_{i-1} \xrightarrow{k_i'} G_i \xrightarrow{k_{i+1}'} G_{i+1} \quad (1)$$

$$d[G_2]/dt = k_2'[M] - k_3'[G_2] \quad (2)$$

$$d[G_i]/dt = k_i'[G_{i-1}] - k_{i+1}'[G_i] \quad (3)$$

It was demonstrated that despite the complexity of the system, the oligomerization can be satisfactorily described by just three kinetic parameters, k_2', k_3', and k_i' ($i \geq 4$), with k_i' independent of i in the range $4 \leq i \leq 14$.[7a] It was also suggested that the elongation of a preformed oligomer, i.e., dimer or longer, is assisted by the presence of at least two additional next-neighbor monomer units (Figure 1B).[7a] It has subsequently been reported that a higher yield of longer oligoguanylates can be obtained by *decreasing* the concentration of the poly(C) template.[9] This

(3) (a) Schwartz, A. W.; Orgel, L. E. *Science* **1985**, *228*, 585−587. (b) Sawai, H.; Higa, K.; Kuroda, K. *J. Chem. Soc., Perkin Trans. 1* **1992**, 505−508. (c) Prabahar, K. J.; Cole, T. D.; Ferris, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 10914−10920. (d) Bolli, M.; Micura, R.; Eschenmoser, A. *Chem. Biol.* **1997**, *4*, 309−320.

(4) (a) Zielinski, W. S.; Orgel, L. E. *Nature* **1987**, *327*, 346−347. (b) Kanavarioti, A. *J. Theor. Biol.* **1992**, *158*, 207−219. (c) Li, T.; Nicolaou, K. C. *Nature* **1994**, *369*, 218−221. (d) Sievers, D.; von Kiedrowski, G. *Nature* **1994**, *369*, 221−224. (e) Rembold, H.; Robins, R. K.; Seela, F.; Orgel, L. E. *J. Mol. Evol.* **1994,** *38*, 211−214. (f) Bag, B. G.; von Kiedrowski, G. *Pure Appl. Chem.* **1996**, *68*, 2145−2152.

(5) (a) Orgel, L. E. *Sci. Am.* **1994**, *271*, 53−61. (b) Kanavarioti, A. *Origins Life Evol. Biosph.* **1994**, *24*, 479−495. (c) Bohler, C.; Nelsen, P. E.; Orgel, L. E. *Nature* **1995**, *376*, 578−581. (d) Ferris, J. P.; Hill, A. R.; Liu, R. H.; Orgel, L. E. *Nature* **1996**, *381*, 59−61. (e) Eschenmoser, A.; Kisakurek, M. V. *Helv. Chim. Acta* **1996**, *79*, 1249−1259. (f) Lazcano, A.; Miller, S. L. *Cell* **1996**, *85*, 793−798. (g) Schwartz, A. W. *Chem. Biol.* **1996**, *3*, 515−518. (h) Miller, S. L. *Nat. Struct. Biol.* **1997**, *4*, 167−169. (i) Deamer, D. W. *Microb. Mol. Biol. Rev.* **1997,** *61*, 239−261.

surprising observation provided impetus to extend the earlier kinetic measurements to include not only a wide range of monomer concentrations but also a wide range of template concentrations.

We report here the kinetics of the oligomerization in the range of 0.005 M ≤ [2-MeImpG] ≤ 0.05 M and 0.002 ≤ [poly(C)] ≤ 0.05 M; the concentration of poly(C) is expressed in cytidine equivalents. The present study confirms that k_i' is independent of i in the range $4 \leq i \leq 16$. More importantly, the large range of poly(C) concentrations investigated provides a test for the validity of the proposed mechanism. It turns out that the observed k_3' and k_i' rate constants for oligoguanylate elongation fit the earlier template-directed model perfectly and strengthen the conclusion of next-neighbor assistance. In contrast, the k_2' values for dimer formation are not consistent with a template-directed mechanism of dimerization and will be reported elsewhere.

Experimental Section

Materials and Methods. The materials used and the procedures employed here for obtaining and analyzing the data have been described in detail.[7a] 2-MeImpG was synthesized in our laboratory and it was better than 96% pure, as determined by analysis with high performance liquid chromatography (HPLC). *N*-(2-Hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid (HEPES) and (ethylene-dinitrilo)tetraacetic acid disodium salt (EDTA) were purchased from Aldrich; the potassium salt of poly(C), about 100 to 300 units long, and Pancreatic ribonuclease A (RNase A) were from Sigma. Reactions were run at 23 °C and at pH 8.0 in the presence of 0.5 M HEPES buffer, 1.2 M NaCl, and 0.2 M $Mg(Cl)_2$. Samples were prepared and quenched at regular intervals by dilution and with addition of EDTA to chelate Mg^{2+}. The analysis of the aliquots was done by HPLC on a 1090 Hewlett-Packard liquid chromatograph with use of an RPC5 column with a $NaClO_4$ gradient at pH 12 (Figure 2). Product identification was simplified by hydrolyzing the imidazolide activated monomer and oligomers at pH 3 and by degrading poly(C) to monomer by RNase A. The identification of the oligomers was based on the analysis done by Inoue and Orgel.[1a] The concentration of each oligomer produced was calculated from the known length of the oligomer, the corresponding HPLC area (see Supporting Information), the initial monomer concentration, the dilution, and the conversion factor of 3.08 pmol of 5′GMP per HPLC unit at pH 12.

Results

General Features. 2-MeImpG is soluble in water at concentrations of ≤0.1 M. It hydrolyzes readily to form guanosine 5′-monophosphate and 2-methylimidazole.[10] At relatively high concentrations of substrate the products include a small percentage of dimers formed by nucleophilic attack of either one of the ribose-hydroxyls, 3′-OH and 2′-OH, or the phosphoryl-oxyanion on the P−N bond of another 2-MeImpG molecule with displacement of its 2-methylimidazole moiety.[11] In the presence of poly(C), under appropriate conditions, 2-MeImpG yields almost quantitatively oligoguanylates up to 40-bases long that are mostly 3′-5′ linked.[1a] Favorable conditions for oligomerization are pH 8.0, 1.2 M NaCl, and 0.2 M $Mg(Cl)_2$, which were used in this investigation as well as previous kinetic studies.

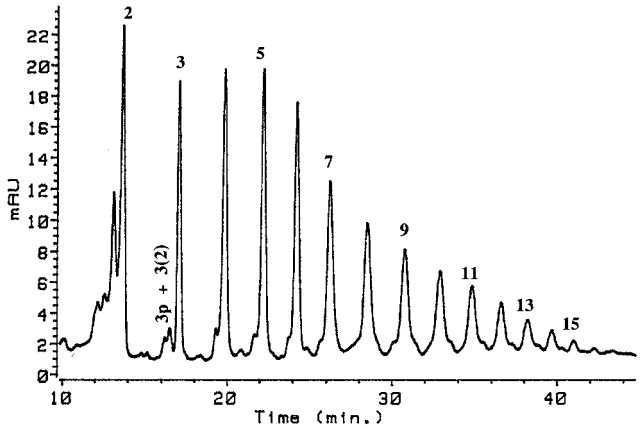


Figure 2. Representative HPLC profile of the oligomerization of 2-MeImpG: Shown reaction of 45 mM 2-MeImpG with 5 mM poly(C) after 2 h of incubation at 23 °C (2-MeIm moieties have been hydrolyzed and poly(C) has been enzymatically degraded; see Experimental Section). Peaks represent oligoguanylates of increasing length. mAU stands for milliabsorbance units at 254 nm.

The kinetic analysis was based on eqs 2 and 3 which can be approximated by eq 4 and, for the last detectable oligomer of length n, by eq 5. In eq 5 the rate term for the conversion of the last detectable oligomer to the next oligomer is omitted. Equation 5 allows the determination of an approximate value for k_n', which then can be used in eq 4 to allow determination of k_{n-1}', and all the preceding k_i' values all the way back to k_2' (see Supporting Information). In eqs 4 and 5 $\Delta t = t_2 - t_1$ is the time interval between two quenches and $\Delta[G_i]$ the observed concentration change of oligomer G_i in the respective time interval. Values for $[G_i]$ were obtained by averaging the observed concentrations at t_1 and t_2. More than six aliquots were used for rate determinations under each condition. These calculations were performed with Microsoft Excel on a Macintosh IIci computer. The obtained rate constants were also verified by computer simulation with KINSIM.[12] The determined rate constants, k_3' and k_i', are reported in Table 1.

$$d[G_i]/dt = k_i'[G_{i-1}] - k_{i+1}'[G_i] \approx \Delta[G_i]/\Delta t \qquad (4)$$

$$d[G_n]/dt = k_n'[G_{n-1}] \approx \Delta[G_n]/\Delta t \qquad (5)$$

Apparent second-order rate constants k_2, k_3, and k_i (not reported) can be easily calculated from the pseudo-first-order rate constants k_2', k_3', and k_i', respectively, and the known monomer concentration based on the equations $k_2' = k_2[M]$, $k_3' = k_3[M]$, and $k_i' = k_i[M]$, where [M] is the average value of the monomer concentration between the initial and the endpoint of the interval Δt. We prefer to discuss the mechanism in terms of pseudo-first-order rather than the second-order rate constants because the chemical processes for which rate constants are measured occur within a preformed complex.

(6) (a) Joyce, G. F.; Orgel, L. E. *J. Mol. Biol.* **1986**, *188*, 433−441. (b) Joyce, G. F.; Orgel, L. E. *J. Mol. Biol.* **1988**, *202*, 677−681. (c) Ng, K. M. E., Orgel, L. E. *J. Mol. Evol.* **1989**, *29*, 101−107. (d) Wu, T., Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 5496−5501. (e) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 7963−7969.

(7) (a) Kanavarioti, A.; Bernasconi, C. F.; Alberas, D. J.; Baird, E. E. *J. Am. Chem. Soc.* **1993**, *115*, 8537−8546. (b) Oligoguanylates are tightly bound to poly(C) at room temperature (Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256). The possibility that the excess of monomer present in solution decreases the binding to such an extent that the shorter oligomers are distributed partly on the template and partly in solution can be discounted on the basis of the observation that the observed k_i' ($i \geq 4$) is independent of length. However, the observation that $k_3 < k_i$ could be attributed to a partially template-bound dimer.

(8) Kanavarioti, A.; Hurley, T. B.; Baird, E. E. *J. Mol. Evol.* **1995**, *41*, 161−168.

(9) Kanavarioti, A.; Baird, E. E. *J. Mol. Evol.* **1995**, *41*, 169−173.

(10) (a) Kanavarioti, A. *Origins Life Evol. Biosph.* **1986**, *17*, 85−103. (b) Kanavarioti, A.; Bernasconi, C. F.; Doodokyan, D. L.; Alberas, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 7247−7257. (c) Kanavarioti, A.; Rosenbach, M. T. *J. Org. Chem.* **1991**, *56*, 1513−1521.

(11) Kanavarioti, A. *Origins Life Evol. Biosph.* **1997**, *27*, 357−376.

(12) Barshop, B. A.; Wrenn, R. F.; Frieden, C. *Anal. Biochem.* **1983**, *130*, 134.

Table 1. Rate Data for Elongation[a]

M/T,[b] mM	r[c]	θ[d]	$[M]_f$,[e] mM	k_3',[f] h^{-1}	k_i',[g] h^{-1}	ref[h]	F_E[i] O·M	O·M_2	O·M_3	O·M_4
5/50	0.88	0.036	3.2	0.09	0.22	7a, 9	0.367	0.160	0.074	0.036
8/50	0.88	0.082	3.9	0.17	0.31	7a	0.483	0.267	0.156	0.094
10/50	0.91	0.117	4.1	0.26	0.59	7a	0.533	0.320	0.203	0.132
15/50	0.92	0.209	4.5	0.52	0.83	7a	0.617	0.419	0.296	0.215
20/50	0.91	0.305	4.8	0.54	1.09	7a, 9	0.673	0.490	0.370	0.285
30/50	0.90	0.500	5.0	0.69	1.34	7a	0.753	0.601	0.492	0.409
40/50	0.87	0.696	5.2	0.87	1.53	7a, 9	0.820	0.700	0.608	0.534
45/50	0.88	0.794	5.3	1.18	1.63	7a	0.856	0.756	0.678	0.612
10/30	0.92	0.185	4.5	0.30	0.60		0.599	0.397	0.274	0.195
15/30	0.91	0.341	4.8	0.53	0.90		0.690	0.513	0.395	0.309
20/30	0.92	0.499	5.0	0.91	1.17		0.752	0.600	0.491	0.408
30/30	0.92	0.820	5.3	0.84	*j*		0.867	0.774	0.700	0.637
8/25	0.87	0.148	4.3	0.38	0.41		0.567	0.359	0.238	0.162
20/35	0.91	0.431	5.0	0.57	1.15		0.727	0.565	0.451	0.366
5/20	0.88	0.065	3.7	*j*	0.32	9	0.449	0.233	0.129	0.074
20/20	0.92	0.737	5.3	0.73	1.53	9	0.834	0.722	0.635	0.565
5/10	0.89	0.098	4.0	*j*	0.42	9	0.508	0.293	0.178	0.112
40/30	0.90	1.000	10.0	1.13	1.66					
20/10	0.93	1.000	10.0	0.80	1.59					
40/10	0.92	1.000	30.0	1.40	1.90					
50/10	0.90	1.000	40.0	1.13	2.15					
20/2	0.93	1.000	18.0	*j*	2.38	9				
45/5	0.93	1.000	40.0	1.13	2.54					

[a] Rate data determined as described under Results. For a listing of the HPLC areas of product peaks as a function of incubation time and the rate constants obtained from the HPLC data see Supporting Information. [b] Initial monomer/template concentration in mM; template concentration in cytidine equivalents. [c] r is the ratio of $[M]/[M]_O$ where [M] is the average value for the monomer concentration between the initial and the end point of the interval Δt and $[M]_O$ is the formal concentration of the substrate. In actuality r corrects for both the purity of the substrate as well as the fact that the activated monomer is consumed during incubation. Reactions were monitored only for a comparatively short Δt, which is why r is close to unity. [d] Occupancy, θ, calculated as described in Results. [e] The concentration of free monomer in solution. [f] Pseudo-first-order rate constant for the reaction of a dimer to form a trimer within a preformed complex; k_3' values accurate to ±30%. [g] Pseudo-first-order rate constant for the reaction of an oligomer of length n ($n \geq 3$) to form the oligomer of length $n + 1$ within a preformed complex; k_i' values accurate to ±20%. [h] This work, unless otherwise noted (see ref 7a or 9). [i] Parameter F_E calculated as described in the Discussion for mechanisms O·M, O·M_2, O·M_3, and O·M_4, respectively. F_E values for the first 8 entries differ slightly from the ones reported in ref 7a, because they have been recalculated here based on the refined q and Q values.[8] [j] Could not be determined with good accuracy.

Discussion

Distribution of Monomer Stacks on the Template. Several lines of evidence strongly suggest that the observed dramatic enhancement of the synthesis of oligoguanylates in the presence of poly(C) is the result of a template-directed oligomerization: (i) there is precedent that stacks of guanosine monomers form 1:1 complexes with poly(C) at pH >7, mimicking the poly(C)·poly(G) double helix;[13] (ii) oligoguanylate yields reach up to 0.95 equiv of the template present in the solution,[1,7a] the concentration of poly(C) being expressed in cytidine equivalents; and (iii) under certain conditions, there is an inverse dependence between template concentration and degree of oligomerization which can be attributed to the fact that it is not a higher template concentration, but a more highly occupied template that leads to faster elongation and results in longer oligomers products.[9]

Stacking and binding of 2-MeImpG at the cytidine sites of poly(C) was evidenced by the hypochromicity of the mixture compared to its components.[8] Analysis of the hypochromicity data, under conditions identical to those of the oligomerization reactions, allowed determination of the fraction of double helical polycytidylate/G complex from which the template occupancy (θ) was calculated. The free monomer concentration, $[M]_f$, was then calculated on the basis of eqs 6 and 7 where $[M]_T$ is the concentration of bound monomer and $[M]_{tot}$ the total monomer concentration. A plot of θ as a function of $[M]_f$, so-called binding isotherm, is S shaped, indicating cooperative binding.[8]

$$\theta = [M]_T/[\text{poly(C)}] \quad (6)$$

$$[M]_{tot} = [M]_T + [M]_f \quad (7)$$

The simplest model for cooperative binding requires two association constants, one (q) for complexation at an isolated site and one (Q) for complexation adjacent to an occupied site, with Q presumed to be independent of the length of the stack. This model can be mathematically described by eq 8 where α_H, the Hill constant,[14] is a measure of the cooperativity of the association where $\alpha_H = (Q/q)^{1/2}$. Fitting the experimental isotherm to eq 8 provided a concentration of 5.55 ± 0.15 mM guanosine monomer at half occupancy of all template sites and association constants $q = 2.22\ M^{-1}$ and $Q = 180\ M^{-1}$.[8] These q and Q values were used to calculate θ and $[M]_T$ for any given set of experimental conditions M/T (see Table 1).[15] An occupancy of 1.0 was assigned to M/T combinations where total monomer concentration exceeded template concentration by more than 10 mM (see last six entries in Table 1).

$$\theta = (Q[M])^{\alpha_H}/\{1 + (Q[M])^{\alpha_H}\} \quad (8)$$

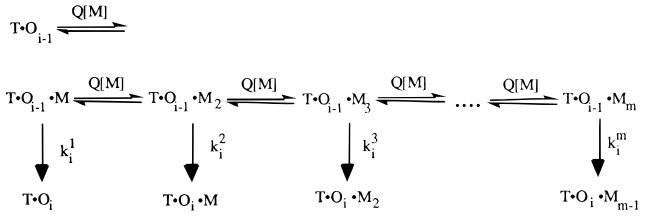
Mechanism of Elongation. If, as we believe, elongation occurs on the template, it is reasonable to assume that elongation of a preformed oligomer will be at its optimal on a fully occupied template. This is because each oligomer will have a

(13) Zimmerman, S. B. *J. Mol. Biol.* **1976**, *106*, 663−672.

(14) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry*; Freeman and Co.: San Francisco, CA., 1980; Part III, p 864 and references in ref 8.

(15) This was done as a spreadsheet calculation with Microsoft Excel on a MacIIci by varying $[M]_f$ by a few percent of a mM at a time and calculating θ, $[M]_T$, and $[M]_{tot}$ for a given [poly(C)], using eqs 6 and 7 and the equilibria implied by Scheme 1. The experimental values of $[M]_{tot}$ were then matched to calculated ones and the corresponding θ and $[M]_T$ for a given [poly(C)] was read out from the spreadsheet. θ and $[M]_f$ are listed in Table 1.

Scheme 1



monomer bound adjacent to its reactive 3′-end. On a less than fully occupied template, some fraction of the oligomers will not have a monomer adjacent to their 3′-end and this will reduce the elongation rate. The data in Table 1 which show an increase in k_3' and k_i' with increasing occupancy (θ) and a leveling off at full occupancy are consistent with this model. The last six entries in Table 1 which refer to situations of fully occupied templates also show that it is in fact θ that determines k_3' and k_i' and not the particular initial concentrations of monomer or template. Taking the average of the last six entries one obtains $k_3' = 1.12 \pm 0.21\ h^{-1}$ and $k_i' = 2.04 \pm 0.39\ h^{-1}$.

As will be shown next, data collected under conditions where the templates are not fully occupied can be analyzed in such a way as to yield limiting values of k_3' and k_i' that correspond to those obtained directly for fully occupied templates. This analysis not only shows good agreement with the results obtained at full occupancy, but gives insights into the detailed mechanism of elongation. This mechanism is shown in Scheme 1. This scheme describes the elongation of a template-bound oligomer,[7b] O_{i-1} of length $i-1$, to form oligomer O_i of length i by addition of one monomer.[7a] The assumption is made that the association constant of a monomer, M, at a site adjacent to the oligomer is equal to the association of a monomer adjacent to a stack of monomers (Q). The first equilibrium step in Scheme 1 exemplifies this association. It is followed by a rate step (vertical) with a rate constant k_i^1 for the actual chemical bond formation process. Depending on [M], more than one monomer, i.e., a stack of two or more monomers, will be template bound at a site adjacent to the oligomer. This is illustrated by the additional Q[M] equilibria in Scheme 1. It is assumed that covalent bond formation occurs with the rate constants k_i^2, k_i^3, ..., k_i^m depending on how many monomers are stacked up at the 3′-end of the oligomer.

O·M Mechanism. The analysis presented below is a summary of what has been originally described in detail.[7a] Assigning different rate constants for each complex allows various mechanistic possibilities to be explored. The simplest assumption is that there is an inherent rate constant for bond formation (k_i*) that is independent of the number of monomers (one or more) stacked up at the 3′-end of the oligomer. We call this the O·M mechanism. However, statistical corrections are required to take into account that monomers can be stacked both at the 5′-end and at the 3′-end with the same Q. This is because only monomers associated at the 3′-end lead to the 3′-5′ elongation, so that complexes which have monomers associated at the 5′-end are considered unreactive. The corrections described in more detail elsewhere[7a] are given by eq 9.

$$k_i^1 = \frac{1}{2}k_i^*; \quad k_i^2 = \frac{2}{3}k_i^*; \quad k_i^3 = \frac{3}{4}k_i^*; \quad ...; \quad k_i^m = \frac{m}{m+1}k_i^* \tag{9}$$

Henceforth O_{i-1} is for simplicity abbreviated O. The rate of elongation is then given by eq 10 and the mass balance for oligomer O by eq 11. In eqs 10 and 11 [T·O·M], [T·O·M_2], and [T·O·M_m] are the actual concentrations of the complexes with one, two, or m monomers.

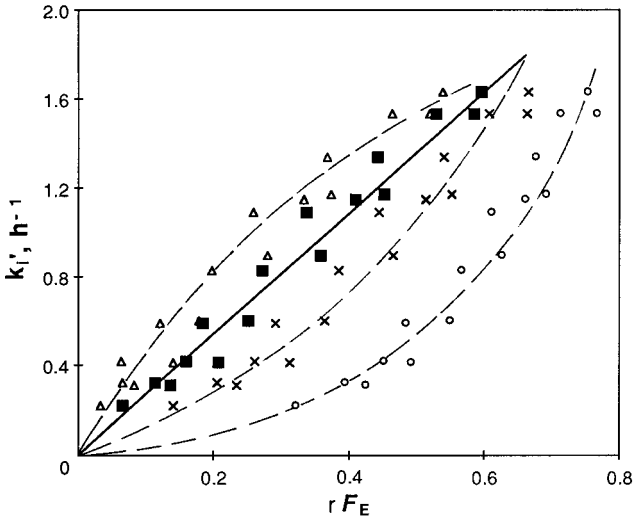


Figure 3. Plots of k_i' ($i \geq 4$) according to eq 12 for the elongation process of a trimer or longer oligomer: circles, O·M mechanism; crosses, O·M_2 mechanism; filled squares, O·M_3 mechanism (the proposed one); triangles, O·M_4 mechanism.

$$-\frac{d[O]}{dt} = k_i^*\left\{\frac{1}{2}[T\cdot O\cdot M] + \frac{2}{3}[T\cdot O\cdot M_2] + \frac{3}{4}[T\cdot O\cdot M_3] + ... + \frac{m}{m+1}[T\cdot O\cdot M_m]\right\} \tag{10}$$

$$[O] = \sum_{j=0}^{m}[T\cdot O\cdot M_j] \tag{11}$$

Dividing both sides of eq 10 by [O] affords eq 12

$$-\frac{d[O]}{[O]\,dt} = k_i' = k_i^* r F_E \tag{12}$$

with $-d[O]/[O]\,dt$ being equal to the observed k_i', r being a measure of the purity of substrate (see caption of Table 1), and F_E being given by eq 13

$$F_E = \frac{1}{2}f_{T\cdot O\cdot M} + \frac{2}{3}f_{T\cdot O\cdot M_2} + \frac{3}{4}f_{T\cdot O\cdot M_3} + ... + \frac{m}{m+1}f_{T\cdot O\cdot M_m} \tag{13}$$

where the $f_{T\cdot O\cdot M_m}$ terms represent the fraction of the oligomer present in the T·O·M_m stack. These fractions were calculated by two methods described previously,[7a] with equal results. The resulting F_E values are reported in Table 1 under the heading O·M.

Equation 12 calls for a linear relationship between the observed pseudo-first-order rate constants and the parameter F_E (or rF_E more precisely) which goes through 0 and provides a slope equal to k_i*. However, the plot of k_i' vs rF_E shown in Figure 3 is curved upward and is thus inconsistent with eq 12. We conclude that the O·M mechanism is inadequate.

O·M_2, O·M_3, and O·M_4 Mechanisms. The next mechanism to which we tried to fit the data is one where complexes with only one monomer at the 3′-end are much less reactive than the ones that have two or more monomers: Specifically we assume $k_i^1 \approx 0$ while all other rate constants in Scheme 1 are the same except for statistical corrections; this is the O·M_2 mechanism. The statistical corrections are given by eq 14 while the F_E term (eq 12) is given by eq 15. The calculated F_E values

are reported in Table 1 and a plot of k_i' vs rF_E for the O·M_2 mechanism is included in Figure 3. This plot shows better adherence to eq 12 than the corresponding plot for the O·M mechanism but there is still room for improvement. The best fit with the data is obtained with the O·M_3 mechanism (Figure 3), which is based on the assumption that efficient covalent bond formation only occurs when at least three monomers are stacked up at the 3′-end of the oligomer. This is expressed in eq 16 with the F_E term given by eq 17 (Table 1). Other mechanisms that were tried, such as the O·M_4 mechanism (eqs 18 and 19) and the O·M_5 mechanism, gave poorer fits with concave up curved plots (shown in Figure 3 for the O·M_4 mechanism).

O·M_2 Mechanism:

$$k_i^1 \approx 0; \quad k_i^2 = \frac{1}{3}k_i^*; \quad k_i^3 = \frac{2}{4}k_i^*; \quad ...; \quad k_i^m = \frac{m-1}{m+1}k_i^* \tag{14}$$

$$F_E = \frac{1}{3}f_{T\cdot O\cdot M_2} + \frac{2}{4}f_{T\cdot O\cdot M_3} + ... + \frac{m-1}{m+1}f_{T\cdot O\cdot M_m} \tag{15}$$

O·M_3 Mechanism:

$$k_i^1 = k_i^2 \approx 0; \quad k_i^3 = \frac{1}{4}k_i^*; \quad k_i^4 = \frac{2}{5}k_i^*; \quad ...; \quad k_i^m = \frac{m-2}{m+1}k_i^* \tag{16}$$

$$F_E = \frac{1}{4}f_{T\cdot O\cdot M_3} + \frac{2}{5}f_{T\cdot O\cdot M_4} + ... + \frac{m-2}{m+1}f_{T\cdot O\cdot M_m} \tag{17}$$

O·M_4 Mechanism:

$$k_i^1 = k_i^2 = k_i^3 \approx 0; \quad k_i^4 = \frac{1}{5}k_i^*; \quad k_i^5 = \frac{2}{6}k_i^*; \quad ...; \quad k_i^m = \frac{m-3}{m+1}k_i^* \tag{18}$$

$$F_E = \frac{1}{5}f_{T\cdot O\cdot M_4} + \frac{2}{6}f_{T\cdot O\cdot M_5} + ... + \frac{m-3}{m+1}f_{T\cdot O\cdot M_m} \tag{19}$$

The slope of the plot of k_i' according to eq 12 for the O·M_3 mechanism yields $k_i^* = 2.49 \pm 0.54$ h^{-1}. This value is in satisfactory agreement with the average $k_i' = 2.04 \pm 0.39$ h^{-1} obtained from the last six entries in Table 1, i.e., conditions where the template is fully occupied.

The value of $k_i^* = 2.49$ h^{-1} obtained at 23 °C compares satisfactorily with the value of $k_i' = 1.5$ h^{-1} obtained for the incorporation of 2-MeImpG in the growing RNA strand on a DNA template at 10 °C.[16] The slower elongation rate in the latter system is partially due to the lower temperature but also partially due to the nature of the template strand (DNA instead of RNA). This is because RNA synthesis is somewhat faster on RNA templates than on DNA templates. For example, at 0 °C the elongation of an RNA strand with 2-MeImpG exhibits a $k_i' = 0.23$ h^{-1} on a DNA template[1c] and $k_i' = 0.83$ h^{-1} on an RNA template.[17]

The elongation of dimers to form trimers (k_3') follows a similar pattern as that observed for the k_i' ($i \geq 4$) steps. Plots of k_3' vs rF_E are shown in Figure 4 for the O·M and the O·M_3 Mechanisms. The k_3' data exhibit more scatter than the k_i' data, nevertheless the curvature shown by the O·M mechanism is clear. In contrast, the O·M_3 Mechanism exhibits a good linear correlation and provides a $k_3^* = 1.43 \pm 0.31$ h^{-1}, in good agreement with the average value of $k_3' = 1.12 \pm 0.21$ h^{-1} obtained from the last six entries (Table 1, fully occupied template). A plot of the k_3' data as a function of rF_E for the O·M_2 mechanism shows slight upward curvature and a plot of the k_3' data as a function of rF_E for the O·M_4 mechanism shows slight downward curvature (not shown). Therefore, we conclude that the O·M_3 mechanism is the preferred mechanism for all elongation steps, including the elongation of the dimer. This mechanism implies that reaction is facilitated by the presence of two additional 2-MeImpG molecules associated downstream of the reacting monomer. Observations in the template-directed incorporation of all four bases as 2-MeIm derivatives with hairpin oligonucleotides are also consistent with the concept of next-neighbor assistance.[6d,e] However, whether the O·M_3 mechanism is preferred with all four nucleobases and whether this assistance is due to the stabilization provided by an incipient double helical complex remains to be seen. Within this mechanism the reacting monomer is "sandwiched" in the midst of a double helical complex (see Figure 1B) with the consequence that in the case of a mismatch by the presence of a non-Watson−Crick base pair the stability of this complex would diminish substantially. In the context of the origin of life and chemical evolution, the stacking requirement provides a mechanism based fidelity which makes it less likely that a mismatched monomer will be incorporated into the growing strand.

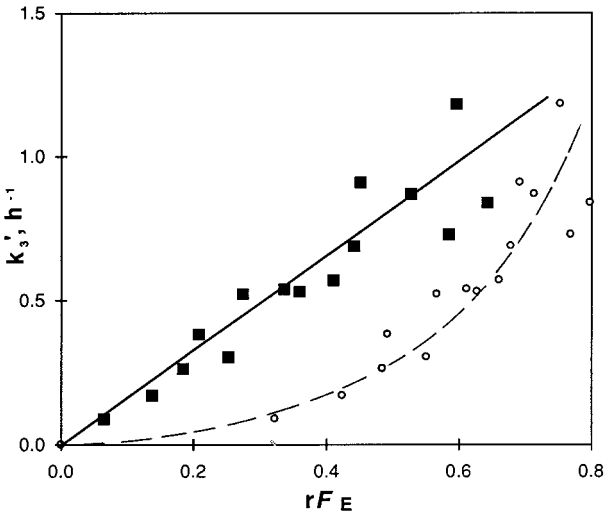


Figure 4. Plots of k_3' according to eq 12 for the elongation process of a dimer: circles, O·M mechanism; filled squares, O·M_3 mechanism (the proposed one).

Implications for the Design of Efficient Oligomerizations. The values of $k_3^* = 1.43$ h^{-1} and $k_i^* = 2.49$ h^{-1}($i \geq 4$) obtained in this study for a range of 0.002 M ≤ [poly(C)] ≤ 0.05 M are in good agreement with the values of $k_3^* = 1.7$ h^{-1} and $k_i^* = 2.9$ h^{-1} ($i \geq 4$) obtained at a constant poly(C).[7a] This agreement demonstrates internal consistency. A comparison of the kinetic parameters of nucleotide oligomerizations may offer design principles for the optimization of such systems. We are only aware of two complete kinetic studies done with nonenzymatic oligomerizations: our own poly(C)/2-MeImpG (G-system) and Ferris,[18] oligomerization of adenosine 5′-monophosphate imidazolide, ImpA, on Na^+-montmorillonite (A system). The first is template directed, whereas the latter one is mineral catalyzed. Both the efficiency and the degree of oligomerization are higher with the G system. The efficiency is defined as the fraction of monomers incorporated into oligomers. The degree or extent of oligomerization is determined by the length of the last

(16) Kurz, M.; Göbel, K.; Hartel, C.; Göbel, M. W. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 842−845.

(17) Prakash, T. P.; Roberts, C.; Switzer, C. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1522−1523.

(18) Kawamura, K.; Ferris, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 7564−7572.

detectable oligomer.[19] Indeed, the efficiency of the G system is almost quantitative,[1a] whereas in the A system only 50% of the monomer is incorporated into oligomers, the other 50% is hydrolyzed to form 5′AMP.[18] In addition, the G system forms oligomers up to 40-units long, whereas the A system yields oligomers only up to 11-units long. Mechanistically, these two systems are similar because they are characterized by the same three processes of (i) dimerization to form the first primer (k_2), (ii) elongation of a preformed primer (k_i, $i \geq 3$), and (iii) hydrolysis of the monomer to form deactivated material (k_h), with $k_h < k_2 < k_i$. Comparison of the corresponding rate constants (A system:[18] $k_h = 7.33 \times 10^{-2}$ h^{-1}, $k_2 = 13.6$ M^{-1} h^{-1}, $k_i = 160$ M^{-1} h^{-1}, and G system:[7a] $k_h = 6.4 \times 10^{-3}$ h^{-1}, $k_2 = 0.33$ M^{-1} h^{-1}, $k_i = 44$ M^{-1} h^{-1})[19a,b] shows that all processes are faster with the A system. Specifically, in the A-system hydrolysis, k_h, is 11 times faster, dimerization is 41 times faster, and elongation is 3.7 times faster than in the G system. However, the ratios are $k_i/k_h = 6875$ and $k_i/k_2 = 133$ for the G system and only 2182 and 12, respectively, for the A system. It is the ratios k_i/k_h and k_i/k_2 which determine efficiency and degree of polymerization, respectively, and not the actual rate constants of the processes involved. This can be understood as follows: In a polymerization the activated monomer is primarily consumed by the elongation process. Therefore it is the elongation process that is in competition with the hydrolysis and this is why k_i/k_h determines efficiency. Furthermore, elongation competes with dimerization, such that the magnitude of the k_i/k_2 ratio determines the degree of polymerization with a larger k_i/k_2 ratio providing for longer products. Hence the fact that a larger k_i/k_h yields better efficiency and a larger k_i/k_2 leads to a higher degree of polymerization may be exploited in the design of second-generation oligomerizing systems.

(19) (a) Detectability is similar in both systems. Conditions are not the same for the two oligomerizing systems, but are optimal for each. Formation of oligoadenylates was done with 0.015 M ImpA in the presence of 50 mg of Na^+-Vol montmorillonite with 0.2 M NaCl, 0.075 M $MgCl_2$, and 0.1 M HEPES (pH 8.0) at 25 °C. (b) For comparison rate constants are used from the reaction of 0.04 M 2-MeImpG in the presence of 0.05 M poly(C).[7a] Conditions are as described in the Experimental Section. Second-order rate constant k_i can also be calculated from the 7th entry in Table 1 based on the equation $k_i = k_i'/([M]*r)$, where [M] is the initial monomer concentration.

Conclusions

This work represents a thorough kinetic investigation of the poly(C)-directed oligoguanylate oligomerization of 2-MeImpG. A comprehensive set of data was obtained where both the monomer and the polymer concentrations were varied. The new data set is consistent with a template-directed mechanism proposed earlier based on much more limited data. A crucial feature of this mechanism is that the elongation is more efficient when the polymer/primer/monomer complex contains at least two additional monomers as immediate neighbors on the template. The catalytic activity of neighboring monomers may point to stacking interactions playing a role in orienting the reactive residue correctly[20] and also indicate that mononucleotides may function as a primitive "polymerase" by regulating the efficiency of primer elongation. A comparison between kinetic parameters for the guanosine polymerization with the montmorillonite-catalyzed adenosine polymerization leads to insights about how the ratio of the rate constants k_i/k_h determines efficiency and the ratio k_i/k_2 determines the degree of a polymerization. A detailed understanding of this chemistry provides an underpinning for the design of second generation systems that exhibit high efficiency and high degrees of polymerization.

Acknowledgment. We thank NASA's Exobiology program for support of this work through a cooperative agreement with Dr. Sherwood Chang of NASA/Ames Research Center (Grant No. NCC 2-534) and Dr. L. E. Orgel of the Salk Institute for providing us with the RPC-5 material.

Supporting Information Available: Tables S1−S15: HPLC areas of product peaks as a function of incubation time; Tables S1R−S15R: calculated rate constants from the HPLC data (31 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA9807237

(20) Suggested by a reviewer.