Ozenberger, B. A., Nahlik, M. S., & McIntosh, M. A. (1987) *J. Bacteriol. 169*, 3638–3646.
Ozenberger, B. A., Brickman, T. J., & McIntosh, M. A. (1989) *J. Bacteriol. 171*, 775–783.
Pettis, G. S., & McIntosh, M. A. (1987) *J. Bacteriol. 169*, 4154–4162.
Pettis, G. S., Brickman, T. J., & McIntosh, M. A. (1988) *J. Biol. Chem. 263*, 18857–18863.
Pierce, J. R., & Earhart, C. F. (1986) *J. Bacteriol. 166*, 930–936.
Pierce, J. R., Pickett, C. L., & Earhart, C. F. (1983) *J. Bacteriol. 155*, 330–336.
Rock, C. O. (1982) *J. Bacteriol. 152*, 1298–1300.
Rose, R. E., DeJesus, C. E., Moylan, S. L., Ridge, N. P., Scherer, D. E., & Knauf, V. C. (1987) *Nucleic Acids Res. 15*, 7197.
Roskoski, R., Gevers, W., Kleinkauf, H., & Lipmann, F. (1970a) *Biochemistry 9*, 4839–4845.
Roskoski, R., Kleinkauf, H., Gevers, W., & Lipmann, F. (1970b) *Biochemistry 9*, 4846–4851.
Rusnak, F., Faraci, W. S., & Walsh, C. T. (1989) *Biochemistry 28*, 6827–6835.
Rusnak, F., Liu, J., Quinn, N., Berchtold, G. A., & Walsh, C. T. (1990) *Biochemistry 29*, 1425–1435.
Safford, R., Windust, J. H. C., Lucas, C., de Silva, J., James, C. M., Hellyer, A., Smith, C. G., Slabas, A. R., & Hughes, S. G. (1988) *Eur. J. Biochem. 174*, 287–295.
Sakaitani, M., Rusnak, F., Quinn, N., Frigo, T. B., Tu, C., Berchtold, G. A., & Walsh, C. T. (1990) *Biochemistry 29*, 6789–6798.
Sanger, F., Nicklens, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Scherer, D. E., & Knauf, V. C. (1987) *Plant Mol. Biol. 9*, 127–134.
Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M., Bibb, M. J., & Hopwood, D. A. (1989) *EMBO J. 8*, 2717–2725.
Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342–1346.
Staab, J. F., & Earhart, C. F. (1990) *J. Bacteriol. 11*, 6403–6410.
Staab, J. F., Elkins, M. F., & Earhart, C. F. (1989) *FEMS Microbiol. Lett. 59*, 15–20.
Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.
Vogel, H. J., & Bonner, D. M. (1956) *J. Biol. Chem. 218*, 97.
Walker, J. E., Saraste, M., Runswick, M., & Gay, N. J. (1982) *EMBO J. 1*, 945–951.
Weckermann, R., Fürbass, & Marahiel, M. A. (1988) *Nucleic Acids Res. 16*, 11841.
Wierenga, R. K., Terpstra, P., & Hol, W. G. J. (1986) *J. Mol. Biol. 187*, 101–107.
Winkler, F. K., D'Arcy, A., & Hunziker, W. (1990) *Nature 343*, 771–774.
Wion, K. L., Kirchgessner, T. G., Lusis, A. J., Schotz, M. C., & Lawn, R. M. (1987) *Science 235*, 1638–1641.
Woodrow, G. C., Young, I. G., & Gibson, F. (1975a) *J. Bacteriol. 124*, 1–6.
Woodrow, G. C., Young, I. G., & Gibson, F. (1975b) *Biochim. Biophys. Acta 582*, 145–153.
Young, I. G., Langman, L., Luke, R. K. J., & Gibson, F. (1971) *J. Bacteriol. 106*, 51–57.
Zocher, R., Keller, U., & Kleinkauf, H. (1982) *Biochemistry 21*, 43–48.

Chemical Synthesis of Oligodeoxynucleotide Dumbbells

Gary W. Ashley* and Diana M. Kushlan
Department of Chemistry, Northwestern University, Evanston, Illinois 60208
Received September 4, 1990; Revised Manuscript Received November 5, 1990

ABSTRACT: The chemical synthesis of DNA dumbbells is investigated by using two sequences, *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) and *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGAA). This method readily and inexpensively yields multimicromole quantities of circular DNA, allowing detailed structural and physical studies to be carried out. Linear oligomers of sequence d(GCG-T_4-CGCCGC-T_4-GCG) having either 3′- or 5′-phosphates were cyclized in 40% and 67% isolated yield, respectively, by using 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide. Formation of the circular product is confirmed by a 28 °C increase in the optical melting temperature, anomalously rapid electrophoretic migration, sequential nuclear Overhauser enhancements between protons of G1 and G20, and observed nuclear coupling between the ligated phosphorus and protons of both G1 and G20. *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGCC) was synthesized from the corresponding linear 3′-phosphate in 80% yield by the same procedure. Chemical ligation is most effective for 3′-phosphorylated nick sites flanked by two purine bases.

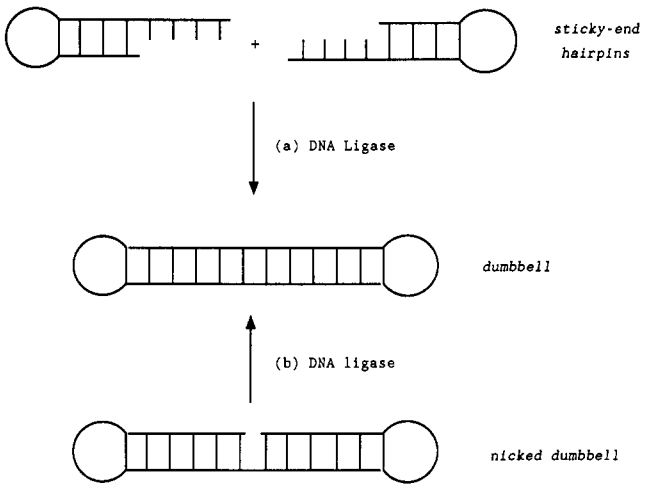
DNA structures containing unpaired bases play significant roles in biological processes. Such noncanonical structures occur as cruciform junctions in supercoiled DNA (Lilley, 1989), near control regions in the chromosome (Muller & Fitch, 1982), and in bulged regions during replication (Yager & von Hipple, 1987). A major interest so far has been the thermodynamics of such conformations (Senior et al., 1988), although their unusual drug binding properties are coming under intense investigation (Williams & Goldberg, 1988; Guo et al., 1990).

Proper understanding of the link between noncanonical DNA structure and biological function requires structural

* To whom correspondence should be addressed.

0006-2960/91/0430-2927$02.50/0

Scheme I

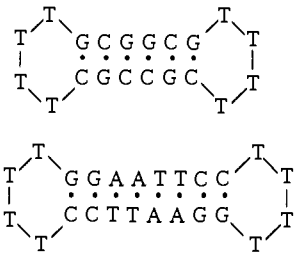


information at atomic resolution, yet such data have been extremely difficult to obtain due to the high conformational mobility exhibited by such structures. There is at present only one X-ray crystal structure of a hairpin loop, and this structure has a Z-form stem as well as clear crystal packing artifacts in the loop region (Chattopadhyaya et al., 1990). Most information concerning hairpin structures has emerged from high-resolution NMR studies, particularly 2D NMR experiments, which have the advantage of studying the oligomer in solution under conditions that are arguably closer to being physiologically relevant (Hare & Reid, 1986; Blommers et al., 1989). Both X-ray and NMR studies of noncanonical structures suffer from conformational heterogeneity due to monomer–dimer equilibria and end-fraying effects.

Closed circular single-strand DNA loops promise to alleviate difficulties due to conformational interconversion. If made with the proper sequence, such circles exist as "dumbbells", a double-helical stem closed off by two hairpin loops. DNA dumbbells have to date been synthesized enzymatically by one of two strategies (Scheme I), yet both procedures are limited in scope. Ligation of two sticky-ended hairpins (Wemmer & Benight, 1985) cannot be used to generate small dumbbells, given the requirement for a number of bases to provide sufficient base pairing in the hairpins combined with the need for several bases for the sticky ends. The alternate method relies on enzymic ligation of a nicked dumbbell and has been used to prepare a small dumbbell containing two TTTTT ("T_5") loop segments, *cyclo*-d(TTCC-T_5-GGAATTCC-T_5-GGAA) (Erie et al., 1989). The corresponding sequence containing two T_4 loop segments failed (Erie et al., 1989), suggesting that the specificity of the ligase may also severely constrain the size dumbbell which may be prepared. The quantity of circular DNA produced by these enzymatic procedures may also be limited by the expense of DNA ligase, making enzymatic procedures undesirable for large-scale (multimicromole) production.

We have found that DNA dumbbells may be synthesized in high yield by using purely chemical techniques. This chemical ligation conveniently and inexpensively produces circular DNA in multimicromole quantities, without many of the limitations inherent to enzymatic procedures. We report here the details of the chemical syntheses of two DNA dumbbells, *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) and *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGAA) (Chart I). These syntheses outline both the merits and difficulties associated with the chemical ligation process. We also provide the first detailed characterization of the products of chemical ligation.

Chart I



Materials and Methods

Oligomer Synthesis. DNA oligomers were synthesized by standard solid-phase cyanoethyl phosphoramidite chemistry (Beaucage & Caruthers, 1981). Phosphorylation used 2-[[2-[(4,4′-dimethoxytrityl)oxy]ethyl]sulfonyl]ethyl 2-cyanoethyl *N,N*-diisopropylphosphoramidite (Horn & Urdea, 1986). For synthesis of 5′-phosphates this reagent was used in the last coupling step. For synthesis of 3′-phosphates, the phosphorylation reagent was used in the first coupling step, with the supported nucleoside being lost upon deprotection. Oligomers were purified by anion-exchange chromatography (Sepharose Q-Fast Flow, 10 mM NaOH, 0.2–1.0 M NaCl). After neutralization (1 N HCl) and concentration in vacuo, the DNA was desalted on HPLC (absorption onto C_{18} silica in H_2O followed by elution with a gradient up to 50% CH_3OH in H_2O). Oligomers were dried in vacuo and then analyzed by NMR and electrophoresis as described below. DNA concentrations were determined from UV absorption, using extinction coefficients derived by nearest-neighbor analysis (Puglisi & Tinoco, 1989).

Chemical Ligation Reactions. The purified, desalted oligomer (1 mM strands) in 50 mM MES[1]–NaOH, pH 6.0, and 20 mM $MgCl_2$ is cooled on ice and treated with 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (final concentration of 0.2 M). The reaction is kept at 4 °C, with aliquots being removed every 24 h and analyzed by HPLC as described below. Upon completion (2–4 days), the mixture is evaporated in vacuo. The residue is dissolved in 1.0 mL of 0.1 M triethylammonium acetate, pH 6.5 (TEAA), and purified by reversed-phase HPLC on C_{18} silica. The column (10 × 250 mm) is equilibrated in 10% methanol–90% TEAA at a flow rate of 3 mL/min. The sample is loaded and the column is washed 5 min with 10:90 CH_3OH/TEAA prior to elution with a linear gradient to 50:50 CH_3OH/TEAA over 60 min. The circular material elutes first as a sharp peak, followed by a broad band of smaller peaks centered around the starting material. To remove TEAA, the product is brought to pH 12 by addition of NaOH and subjected to anion-exchange chromatography as described above. The DNA peak is collected, neutralized with 1 N HCl, concentrated, and desalted by HPLC.

Gel Electrophoresis. Oligomers were dissolved in formamide loading buffer and analyzed on 20% acrylamide gels (1:20 cross-linking) containing 8.3 M urea and 90 mM Tris–borate–EDTA, pH 8.5. The gel temperature was held at 50 °C with a circulating bath to ensure denaturation. DNA bands were visualized by staining with methylene blue.

[1] Abbreviations: COLOC, correlation spectroscopy by long-range couplings; DQF-COSY, double-quantum-filtered correlation spectroscopy; EDC-HCl, 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride; MES, 2-(*N*-morpholino)ethanesulfonic acid; NOESY, nuclear Overhauser and exchange spectroscopy; TEAA, 0.1 M triethylammonium acetate, pH 6.5.

Table I: Isolated Yields of Cyclic DNA Oligomers Obtained from Chemical Ligation[a]

precursor sequence	reaction time (days)	isolated yield (%)
pGCG-T_4-CGCCGC-T_4GCG	4	40
GCG-T_4-CGCCGC-T_4-GCGp	2	67
TTCC-T_4-GGAATTCC-T_4-GGAAp	4	30
ATTCC-T_4-GGAATTCC-T_4-GGAp	2	80

[a] DNA oligomers were cyclized with 0.2 M EDC-HCl in 50 mM MES, pH 6.0, and 20 mM $MgCl_2$. Yields refer to the recovery of DNA OD_{260} units after purification by reversed-phase HPLC, removal of Et_3N–HOAc by anion-exchange chromatography, and desalting.

NMR Spectroscopy. NMR spectroscopy was performed at 30 °C on a Varian XLA-400 spectrometer. ^{31}P NMR (161 MHz) was referenced to H_3PO_4 in an external cylindrical capillary, set to −0.73 ppm to correct for solvent diamagnetic susceptibility (Tebby, 1987). ^1H NMR (400 MHz) was referenced to internal sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$. DNA samples were exchanged into D_2O by repeated evaporation from 99.8 atom % D_2O followed by dissolution in 0.5 mL of 99.98 atom % D_2O. The residual HOD resonance was suppressed by using presaturation during the relaxation delay. Phase-sensitive 2D NMR experiments were acquired by using the hypercomplex method (States et al., 1982) and were processed by using Varian VXR software. The NOESY and DQF-COSY experiments consisted of 512 increments in t_1 and 1024 points in t_2. The NOESY experiment used 64 transients per increment with a relaxation delay of 2.5 s, while the DQF-COSY experiment used 32 transients per increment with a relaxation delay of 1.4 s. Both data sets were zero-filled to 2K × 2K and weighted by using a pseudo-echo envelope prior to transformation. The ^1H–^{31}P COLOC experiment consisted of 120 increments in t_1, 208 transients per increment, a relaxation delay of 2.5 s, and 512 points in t_2. Spectral widths of 1205 and 496 Hz were used for ^1H and ^{31}P dimensions, respectively. Fixed delays of $D3 = 52$ ms and $D4 = 31$ ms were used. The data set was zero-filled to 1K × 1K and weighted with pseudo-echo envelopes in both dimensions prior to Fourier transformation.

Optical Melting Profiles. DNA samples (3 OD_{260}) dissolved in 3.0 mL of 10 mM Tris-HCl and 10 mM NaCl, pH 7.5, were placed in a sealed cuvette fitted with a magnetic stir bar and thermometer. The cuvette was kept in a stirred, thermostated cell holder, and the absorbance at 260 nm was monitored by using a Hewlett-Packard 8420A diode-array spectrophotometer set for a 2.5-s signal integration time. The sample temperature was increased by 1 °C/min, and readings were taken every 2 °C. The melting temperature, T_m, was estimated as the midpoint of the transition curve.

Results

Chemical Ligation of 5′-Phosphorylated DNA. d(pGCG-T_4-CGCCGC-T_4-GCG) was found to cyclize upon treatment with 0.2 M EDC-HCl (Table I), although the yield of circular product was at most 50% despite extensive variations in reaction conditions. After 4 days at 4 °C, electrophoresis of the reaction mixture indicated a roughly 1:1 mixture of product (rapid migration) and a material migrating similar to starting material (slow migration). Increased reaction times led to slow destruction of the cyclic product and accumulation of currently unidentified byproducts. The suspected linear and circular materials from a 4-day reaction were separated by HPLC and analyzed by their optical melting profiles, which suggested that they were in fact circular and linear oligomers. Analysis of the linear oligomer by ^{31}P NMR indicated loss of the phosphate monoester resonance, and resubmission of the linear DNA to the reaction conditions did not result in further cyclization. These results suggest that the recovered linear DNA has undergone chemical modification.

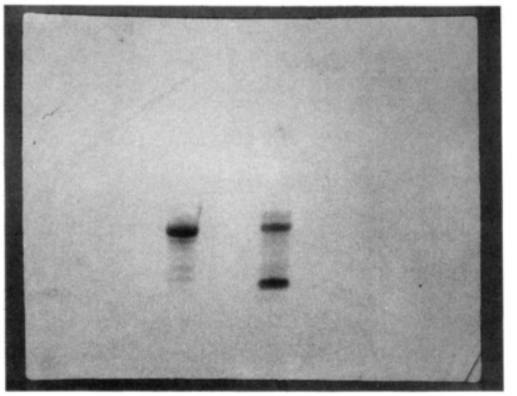


FIGURE 1: Electrophoretic analysis of d(ATTCC-T_4-GGAATTCC-T_4-GGAp) cyclization reaction. Left lane: Starting nicked oligomer. Right lane: Cyclization reaction after 24 h at 4 °C. The reaction aliquot was removed and electrophoresed according to the procedure given under Materials and Methods.

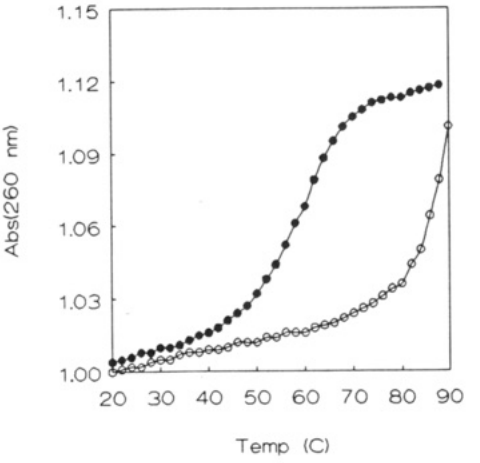


FIGURE 2: Optical melting profiles for nicked and ligated DNA dumbbells. (●) Nicked, d(GCG-T_4-CGCCGC-T_4-GCGp); (○) ligated, *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG).

Chemical Ligation of 3′-Phosphorylated DNA. The 20-mer d(GCG-T_4-CGCCGC-T_4-GCGp) was submitted to cyclization by using 0.2 M EDC-HCl under the same conditions as for the 5′-phosphorylated oligomer. Cyclization of the 3′-phosphate was essentially complete after 2 days at 4 °C as evidenced by gel electrophoresis and HPLC (Table I).

The 24-mer d(TTCC-T_4-GGAATTCC-T_4-GGAAp) was treated in the same manner. Cyclization of this sequence was less efficient, with the reaction being only 25% complete after 2 days at 4 °C. An analogous 24-mer having the nick site shifted by one base, d(ATTCC-T_4-GGAATTCC-T_4-GGAp), cyclized rapidly to give a 75:25 mixture of cyclic and linear materials, respectively, after 24 h at 4 °C (Figure 1). The linear DNA was essentially consumed after 48 h.

Characterization of Cyclic Products. The thermal stability of base-stacked *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) was compared with that of nicked d(pGCG-T_4-CGCCGC-T_4-GCG) (Figure 2). The linear oligomer shows $T_m = 58$ °C, whereas the circular oligomer gives an estimated $T_m = 86$ °C, giving an increase of 28 °C. Exact determination of the T_m for the cyclic product was not possible as an upper plateau region was not observed before the limit of temperature was reached. The estimated value is made by assuming equivalent hyperchromicity for the linear and circular oligomers. A

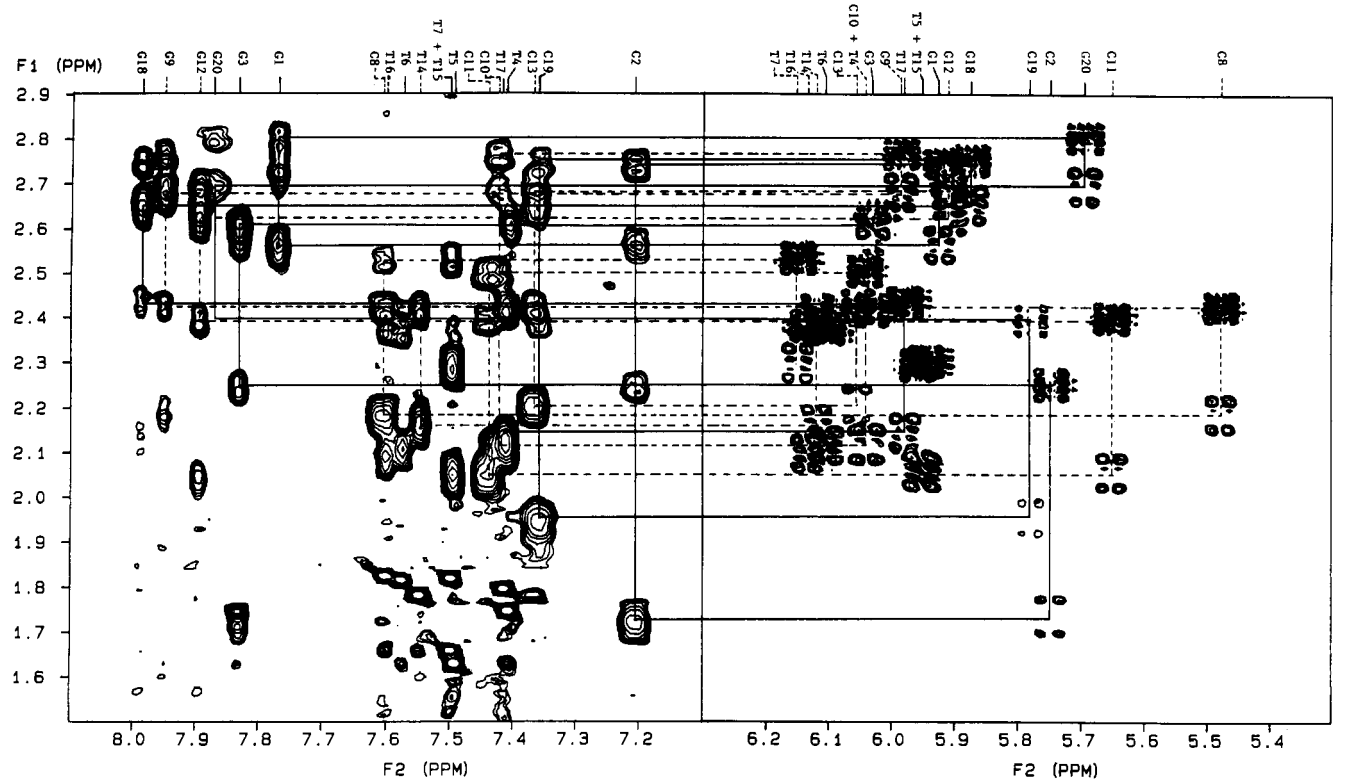


FIGURE 3: Combined NOESY–DQF-COSY connectivity map for *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG). The NMR experiments used 3.5 μmol of DNA dissolved in 0.5 mL of D_2O. Base-to-H2′/H2″ NOESY (300-ms mix) cross-peaks are plotted at the left, and H1′–H2′/H2″ DQF-COSY cross-peaks are plotted on the right. Solid lines connect sequential assignments on the GCGGCG side, and dashed lines, those on the CGCCGC side of the double-helical stem region. Connections within the two T_4 loop segments are not shown for clarity. Residues are numbered starting from the 5′-end of the nicked precursor, d(GCG-T_4-CGCCGC-T_4-GCGp). The expected cross-peak between G20 H8 and C19 H2″ is not observed.

Table II: Resonance Assignments for *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG)

residue	H5	H6	H8	H1′	H2′	H2″	H3′	H4′	P
G1			7.773	5.925	2.566	2.746	5.000	4.409	−1.074
C2	5.342	7.212		5.749	1.725	2.244	4.801	4.110	−1.072
G3			7.834	6.030	2.610	2.610	4.970	4.370	−1.095
T4	1.751	7.410		6.033	2.114	2.417	4.810	4.220	−0.870
T5	1.637	7.496		5.943	2.051	2.301	4.650	4.190	−1.366
T6	1.822	7.579		6.105	2.111	2.375	4.720	4.185	−1.150
T7	1.825	7.504		6.151	2.297	2.530	4.708	4.250	−1.366
C8	5.859	7.610		5.482	2.179	2.417	4.815	4.160	−0.566
G9			7.952	5.987	2.681	2.769	5.015	4.400	−1.130
C10	5.391	7.430		6.043	2.114	2.506	4.850	4.245	−1.070
C11	5.622	7.450		5.655	2.044	2.391	4.859	4.128	−0.919
G12			7.892	5.908	2.623	2.736	4.986	4.368	−0.890
C13	5.384	7.373		6.057	2.204	2.430	4.755	4.236	−0.995
T14	1.787	7.552		6.120	2.156	2.414	4.815	4.195	−0.760
T15	1.665	7.502		5.961	2.058	2.294	4.720	4.070	−1.080
T16	1.830	7.605		6.140	2.090	2.383	4.855	4.060	−1.399
T17	1.799	7.418		5.980	2.140	2.432	4.680	4.200	−1.283
G18			7.990	5.876	2.653	2.751	4.949	4.346	−1.123
C19	5.384	7.373		5.781	1.940	2.387	4.855	4.193	−0.919
C20			7.872	5.698	2.686	2.793	5.015	4.355	−1.006

similar increase in thermal stability of approximately 41 °C was observed for *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGAA) when compared with the precursor d(TTCC-T_4-GGAATTCC-T_4-GGAAp) (data not shown).

Confirmation of the cyclic structure postulated for *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) was obtained by NMR spectroscopy. Proton resonance assignments were obtained by using a combination of NOESY and double-quantum-filtered COSY experiments as shown in Figure 3 (Chazin et al., 1986). Residue numbers are assigned from the 5′-end of the nicked precursor, with ligation occurring between G1 and G20. The tracing obtained was verified by the alternate NOESY-only method (Hare et al., 1983). While an unbroken chain tracing was not obtained for the entire circle due to breaks in base stacking in the T_4 loop regions, complete assignments of all double-helical residue proton resonances except for H5′/H5″ were obtained. All thymidine spin systems have been identified, although there are still ambiguities in the assignment of some sugar proton resonances in these residues due to overlap. The proton resonance assignments for *cyclo*-d(GCG-T_4-CGCGC-T_4-GCG) are given in Table II.

NOESY cross-peaks were observed between the protons of G1 and G20, indicating that these residues are proximal. The correlation between H1′ of G20 and H8 of G1 expected for a double-helical structure is observed, as is that between H2″ of G20 and H8 of G1. For conclusive evidence of ligation,

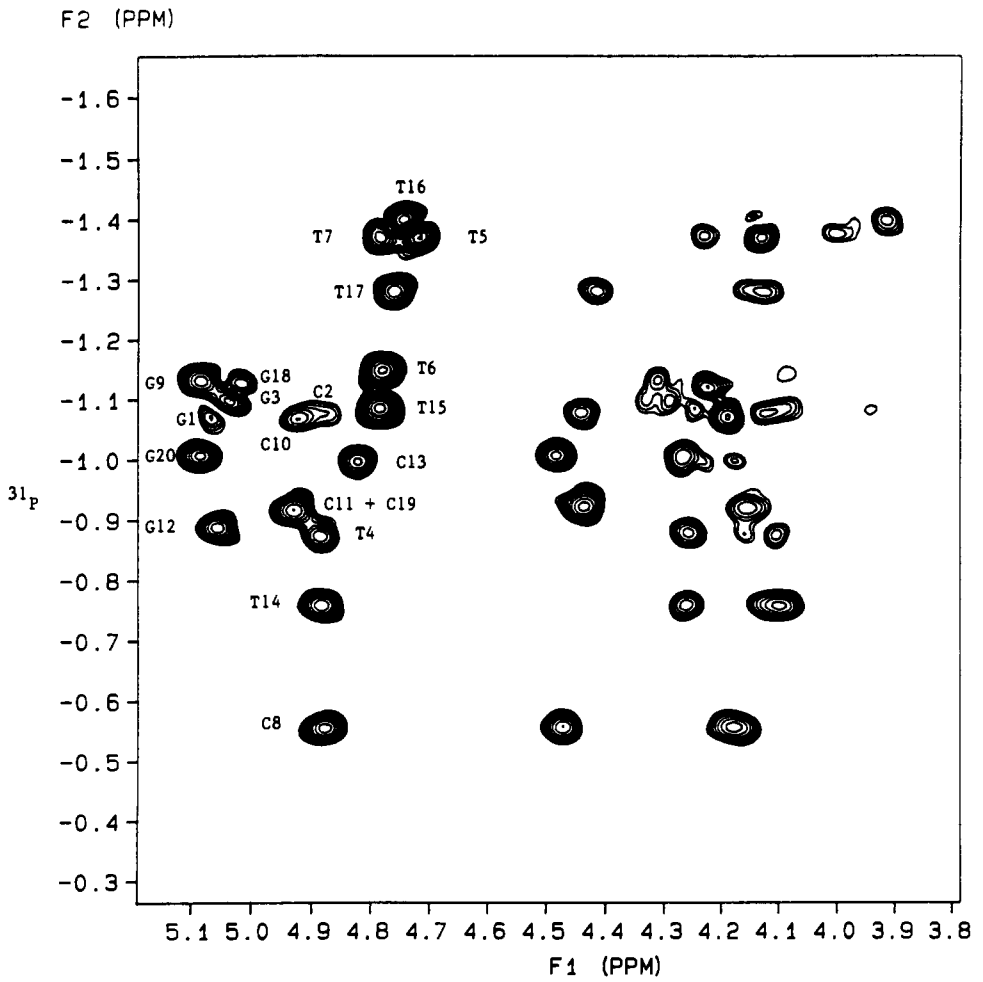


FIGURE 4: ^{1}H–^{31}P COLOC spectrum of *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG). Assignments are given next to the cross-peaks representing coupling to each residue H3′, with G1 being the correlation between G1 H3′ and the phosphorus on the 3′-side of G1.

coupling between protons on G1 and G20 and the phosphorus (P20) at the ligation site was observed using COLOC (Figure 4). The NMR results also demonstrate the absence of EDC-modified bases on the DNA product.

Discussion

DNA dumbbells have achieved substantial popularity as models for the properties of large DNA segments. Dumbbells have been primarily used to study the thermodynamics of DNA melting (Wemmer & Benight, 1985; Erie et al., 1987, 1989; Benight et al., 1988), although they have also found use in structural studies of DNA nick sites (Snowden-Ifft & Wemmer, 1990). To date, DNA dumbbells have been prepared by T4 DNA ligase mediated closure of nick sites. This procedure is limited in scope, as demonstrated by the inability DNA ligase to prepare *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGAA) (Erie et al., 1987). DNA ligase is also a relatively expensive reagent, with approximately $25 of enzyme being used to synthesize 0.1 μmol of DNA dumbbell (Wemmer & Benight, 1985). As many physical studies, such as detailed NMR solution structure studies, require DNA syntheses on 10-μmol scales, the cost of DNA ligase can be high.

We have found that the chemical ligation method of Shabarova (1988) is advantageous for the preparation of DNA dumbbells. In the present case nicked dumbbells are ligated, although the method could in principle also be used for ligation of two sticky-ended hairpins. The present scheme has the advantage of being able to prepare significantly smaller dumbbell sequences than the hairpin-fusion method and should proceed in higher yield as it involves only a single ligation reaction. The hairpin fusion method may be attractive for the preparation of larger dumbbells, where the yields from chemical synthesis of the long precursors may be a consideration. We have demonstrated that the method can readily supply multimicromole quantities of DNA dumbbells and that it succeeds on at least one sequence known to fail with T4 DNA ligase. It may therefore be generally more useful for the preparation of unusual DNA conformations required for detailed structural, thermodynamic, and biological investigations. There are limitations to the chemical ligation technique, however, which have also been investigated.

Both 5′- and 3′-phosphorylated oligomers cyclize in the presence of the water-soluble carbodiimide EDC, yet the 5′-phosphate appears to be more susceptible to side reactions and the yield of DNA dumbbell is relatively low. Cyclizations of the 3′-phosphates proceed in good-to-excellent yield and without significant side product formation, giving a 65–80% yield of dumbbell after purification. This is in accord with the observations of Shabarova, who noted higher yields in the chemical ligations of linear DNA segments when 3′-phosphates were used. Our results suggest that the major limitation in using 5′-phosphates in chemical modification is their susceptibility toward dephosphorylation, although the mechanism for this process is currently unknown. The increased nucleophilicity of the primary 5′-hydroxyl group over the secondary 3′-hydroxyl group is apparently sufficient to overcome competing reactions and significantly improve the ligation yield.

Comparison of the optical melting profile of the products with the nicked starting materials reveals large increases in T_m. This is consistent with previous studies of dumbbell structures, which show a 25–50 °C increase in thermal stability of the duplex form upon cyclization (Erie et al., 1987).

Two-dimensional NMR studies provided information on the structure of *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) and demonstrated that normal ligation had occurred. It should be noted that this is the first detailed characterization of an oligonucleotide prepared according to the chemical ligation technique. The observation of nuclear Overhauser enhancements between protons of the nick-site residues G1 and G20 confirmed the close spatial arrangement of these two residues. Observation of NOESY cross-peaks between G1 and G20 does not provide clear evidence of proper ligation, however, as the precursors themselves would be expected to show similar nuclear Overhauser enhancements on the basis of their relatively high T_m values. More conclusive evidence for ligation was obtained by examination of scalar coupling between protons and phosphorus using the COLOC experiment (Kessler et al., 1984). The phosphorus of each dinucleotide unit 5′-XpY-3′ is coupled to the H3′ proton of nucleoside X and to the H4′/H5′/H5″ system of nucleoside Y, although correlations to H5′/H5″ pairs are typically weak due to the strong geminal coupling between the two protons. Since the H3′ and H4′ protons were assigned from DQF-COSY and NOESY experiments, the COLOC experiment allowed assignment of the phosphate groups. The ligated phosphorus (P20) was found to be coupled to the H3′ proton of G20 and to the H4′/H5′/H5″ system of G1, conclusively demonstrating proper ligation of the precursor to form *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG).

The NMR data on *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) indicate intriguing conformational properties. In particular, the H8 base resonance of G20 is quite broad, suggesting that this residue may be in exchange between two conformations. It has been suggested that GG steps in DNA tend to prefer an A-like conformation (Heinemann et al., 1990), although calculations indicate that GG steps are among the least stable pairs in A-form (Haran et al., 1987). Also, the sugar resonances of C8 are quite sensitive to temperature, with H1′ showing a 0.1 ppm dowfield shift in going from 20 and 30 °C, and another 0.1 ppm downfield shift in going from 30 and 40 °C. As C8 lies at a helix–loop junction, such shifts indicate conformational mobility in this region, even though the sample is 50 °C below the helix melting temperature. This dumbbell may therefore offer a unique opportunity to investigate conformational interconversions in a small DNA model. Detailed solution structure information on this dumbbell is currently being pursued.

One potential difficulty with the chemical ligation of DNA oligomers is the known propensity of EDC to covalently modify unpaired bases, particularly G and T (Gilham, 1962), a technique used to map single-stranded regions DNA (Ganguly & Prockop, 1990). We find no evidence by ^{1}H NMR for attachment of EDC to the unpaired T residues in either DNA dumbbell produced. We do note the formation of unidentified byproducts at extended reaction times, however. These products begin to appear after 3 days at 4 °C and become a significant fraction of the reaction mixture after 7 days. We suspect on the basis of their electrophoretic mobilities that these materials represent slow covalent modification of the DNA by EDC. For this reason, the extended reaction times advocated by Shabarova for chemical ligation should be avoided when dealing with oligomers containing unpaired bases. We find little improvement in product yield after 2–3 days at 4 °C.

The most stringent test of chemical ligation for the production of DNA dumbbells is the successful production of *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGAA) from the corresponding 3′-phosphorylated 24-mer. This sequence is resistant to cyclization using T4 DNA ligase (Erie et al., 1987). Since the corresponding sequence containing T_5 loop segments rather than the smaller T_4 loop segments can be cyclized enzymatically, it has been proposed that the smaller T_4 loop segment distorts the structure of the nicked dumbbell and prevents proper alignment of the chain ends. Given the success of chemical ligation at cyclizing the T_4-loop sequence, however, it is apparent that the previously noted difficulty in synthesis of this dumbbell lies in the substrate specificity of T4 DNA ligase rather than in any conformational property of the oligomer and thus represents an inherent limitation of the enzymatic procedure.

Limitations on the DNA sequence required for successful cyclization have been noted. In contrast to the results of Shabarova with linear DNA ligations (Shabarova, 1988), we have found that purine–purine steps ligate more efficiently than purine–pyrimidine steps. Thus, d(ATTCC-T_4-GGAATTCC-T_4-GGAp), having two purines flanking the nick site, is more efficiently ligated than d(TTCC-T_4-GGAATTCC-T_4-GGAAp), having a purine and a pyrimidine at the nick site. This is thought to be due to the requirement for stable base stacking at the nick site.

We are currently investigating the chemical syntheses of more complex circular DNA structures, including models for cruciform junctions and internal bubbles, which are unlikely to be successfully prepared by using enzymatic ligation.

Acknowledgments

We thank Dr. Zoe Shabarova for information concerning the chemical ligation technique, Dr. Robert Letsinger for helpful discussions on DNA synthesis, and Dr. Ka-Leung Ngai for assistance in oligomer synthesis.

References

Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859.

Benight, A. S., Schurr, J. M., Flynn, P. F., Reid, B. R., & Wemmer, D. E. (1988) *J. Mol. Biol. 200*, 377.

Blommers, M. J. J., Walters, J. A. L. I., Haasnoot, C. A. G., Aelen, J. M. A., van der Marel, G. A., van Boom, J. H., & Hilbers, C. W. (1989) *Biochemistry 28*, 7491.

Chattopadhyaya, R., Grzeskowiak, K., & Dickerson, R. E. (1990) *J. Mol. Biol. 211*, 189.

Chazin, W. J., Wuthrich, K., Hyberts, S., Rance, M., Denny, W. A., & Leupin, W. (1986) *J. Mol. Biol. 190*, 439.

Erie, D., Sinha, N., Olson, W., Jones, R., & Breslauer, K. (1987) *Biochemistry 26*, 7150.

Erie, D. A., Jones, R. A., Olson, W. K., Sinha, N. K., & Breslauer, K. J. (1989) *Biochemistry 28*, 268.

Ganguly, A., & Prockop, D. J. (1990) *Nucleic Acids Res. 18*, 3933.

Gilham, P. T. (1962) *J. Am. Chem. Soc. 84*, 687.

Guo, Q., Lu, M., Seeman, N. C., & Kallenbach, N. R. (1990) *Biochemistry 29*, 570.

Haran, T. E., Shakked, Z., Wang, A. H.-J., & Rich, A. (1987) *J. Biomol. Struct. Dyn. 5*, 199.

Hare, D. R., & Reid, B. R. (1986) *Biochemistry 25*, 5341.

Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319.

Heinemann, U., Alings, C., & Lauble, H. (1990), in *Structure & Methods, Volume 3: DNA & DNA* (Sarma, R. H., & Sarma, M. H., Eds.) pp 39–53, Adenine Press, Guilderland, NY.
Horn, T., & Urdea, M. (1986) *Tetrahedron Lett. 27*, 4705.
Kessler, H., Griesinger, C., Zarbock, J., & Loosli, H. R. (1984) *J. Magn. Reson. 57*, 331.
Lilley, D. M. J. (1989) *Chem. Soc. Rev. 18*, 53.
Muller, U. R., & Fitch, W. M. (1982) *Nature (London) 298*, 582.
Puglisi, J. D., & Tinoco, I., Jr. (1989) *Methods Enzymol. 180*, 304.
Senior, M. M., Jones, R. A., & Breslauer, K. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6242.
Shabarova, Z. A. (1988) *Biochemie 70*, 1323.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.
Tebby, J. C. (1987) in *Phosphorus-31 NMR Spectroscopy in Stereochemical Analysis* (Verkade, J. G., & Quin, L. D., Eds.) pp 1–60, VCH, Deerfield Beach, FL.
Wemmer, D. E., & Benight, A. S. (1985) *Nucleic Acids Res. 13*, 8611.
Williams, L. D., & Goldberg, I. H. (1988) *Biochemistry 27*, 3004.
Yager, T. D., & von Hipple, P. H. (1987) in *Cellular and Molecular Biology* (Neidhart, F., et al., Eds.) pp 1241–1275, American Society of Microbiology, Washington, DC.

Kinetic Mechanism of the *Eco*RI DNA Methyltransferase†

Norbert O. Reich* and Neda Mashhoon
Chemistry Department, University of California, Santa Barbara, California 93106
Received July 3, 1990; Revised Manuscript Received December 14, 1990

ABSTRACT: We present a kinetic analysis of the *Eco*RI DNA N^6-adenosine methyltransferase (Mtase). The enzyme catalyzes the *S*-adenosylmethionine (AdoMet)-dependent methylation of a short, synthetic 14 base pair DNA substrate and plasmid pBR322 DNA substrate with k_{cat}/K_m values of 0.51×10^8 and 4.1×10^8 s^{-1} M^{-1}, respectively. The Mtase is thus one of the most efficient biocatalysts known. Our data are consistent with an ordered bi-bi steady-state mechanism in which AdoMet binds first, followed by DNA addition. One of the reaction products, *S*-adenosylhomocysteine (AdoHcy), is an uncompetitive inhibitor with respect to DNA and a competitive inhibitor with respect to AdoMet. Thus, initial DNA binding followed by AdoHcy binding leads to formation of a ternary dead-end complex (Mtase–DNA–AdoHcy). We suggest that the product inhibition patterns and apparent order of substrate binding can be reconciled by a mechanism in which the Mtase binds AdoMet and noncanonical DNA randomly but that recognition of the canonical site requires AdoMet to be bound. Pre-steady-state and isotope partition analyses starting with the binary Mtase–AdoMet complex confirm its catalytic competence. Moreover, the methyl transfer step is at least 10 times faster than catalytic turnover.

The selective recognition and subsequent modification of DNA by enzymes are critical to all known forms of life. A nearly ubiquitous example of such DNA modification is the methylation of cytosine and adenine bases. This expansion of the information content of DNA has biological consequences ranging from protection of host DNA against restriction endonucleases to regulation of gene expression in eukaryotes (Razin et al., 1984; Modrich & Roberts, 1982). Moreover, sequence-specific modification of DNA represents a dramatic achievement in macromolecular recognition. DNA methylation is catalyzed by a diverse group of enzymes which are uniformly dependent on *S*-adenosylmethionine as a methyl donor. Prokaryotic type II DNA Mtases, with hundreds of sequence specificities, form a large subset of DNA-modifying enzymes; their use in molecular biology is widespread (Modrich & Roberts, 1982; Koob et al., 1988). An understanding of the recognition and catalytic mechanisms of these enzymes may facilitate the design of biocatalysts with novel specificities and chemical mechanisms.

We are investigating the *Eco*RI[1] DNA Mtase to understand the details of DNA sequence discrimination, cofactor interactions, and the catalytic mechanism. The *Eco*RI Mtase is functional as a monomer (MW 38050) and is part of a type II bacterial restriction–modification system. The enzyme methylates the second adenine in the canonical site 5′GAATTC3′ to form N^6-methyladenine and *S*-adenosylhomocysteine (AdoHcy), thereby making the site resistant to cleavage by the corresponding *Eco*RI endonuclease (Rubin & Modrich, 1977). The availability of the Mtase gene (Greene et al., 1981), large amounts of highly purified protein, and relatively simple catalytic requirements make this an ideal system to investigate DNA methylation. Moreover, the X-ray structures of the DNA substrate (Drew et al., 1981), methylated DNA product (Frederick et al., 1988), and *Eco*RI endonuclease–DNA complex (McClarin et al., 1986) provide detailed structural information.

One approach to understanding DNA–enzyme interactions is by functional analysis of DNA substrates which lack structural features proposed to be contacted by the enzyme. Interpretation of such detailed specificity analyses is aided by an understanding of the kinetic mechanism; thus, the order

† This work was supported by National Science Foundation Grant DMB 89-11404 to N.O.R.
* To whom correspondence should be addressed.

[1] Abbreviations: *Eco*RI, plasmid-specified DNA restriction–modification system of *Escherichia coli*; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; AdoMet, *S*-adenosylmethionine; AdoHcy, *S*-adenosylhomocysteine; DE81, diethylaminoethyl ion-exchange filters.

0006-2960/91/0430-2933$02.50/0