DE-AC02-76CH00016 at BNL and by National Institutes of Health Grants GM34468 and 36520 at MSU.

Supplementary Material Available: Crystallographic details, a stereoscopic packing diagram, deviations from planarity, and final positional and thermal parameters for the non-hydrogen atoms of **3** (4 pages); observed and calculated structure factors for **3** (16 pages). Ordering information is given on any current masthead page.

Nonenzymatic Sequence-Specific Ligation of Double-Helical DNA

Kevin J. Luebke and Peter B. Dervan*

Arnold and Mabel Beckman Laboratories of Chemical Synthesis
California Institute of Technology
Pasadena, California 91125

Received May 31, 1991

Formation of a phosphorus–oxygen bond between phosphate and hydroxyl termini of DNA in aqueous solution requires chemical activation of the phosphate for nucleophilic substitution and positioning of the hydroxyl for attack on the activated phosphate in competition with water.[1-4] This esterification reaction is accomplished enzymatically by DNA ligases, which utilize energy from an ATP or NAD cofactor to activate the phosphates.[5] We report a *nonenzymatic* approach to ligation of double-helical DNA employing a single-stranded template to align two duplex strand termini in a local triple helix (Figure 1). A triple-stranded complex is formed by association of a pyrimidine oligodeoxyribonucleotide in the major groove of the Watson–Crick duplex with sequence specificity derived from Hoogsteen hydrogen bonding.[6-8] Juxtaposition of the two DNA termini by a guide sequence in a triple helix, accompanied by chemical activation of the terminal phosphates, promotes ligation of the double-helical DNA.

A 3.7 kilobase pair (kbp) blunt-ended linear DNA duplex possessing a 15 base pair purine tract for triplex formation at each end was constructed.[9] A 30 nucleotide template strand, complementary in a Hoogsteen sense to the continuous 30 base pair triplex site formed by apposing the termini of the double-helical DNA, was also synthesized. In a triple-stranded complex formed by association of the template strand with both ends of the double-helical DNA, the linear DNA molecule would be circularized. Upon chemical activation, the proximal 3′-hydroxyl and 5′-phosphate termini would be susceptible to covalent ligation on one or both strands (Figure 2).

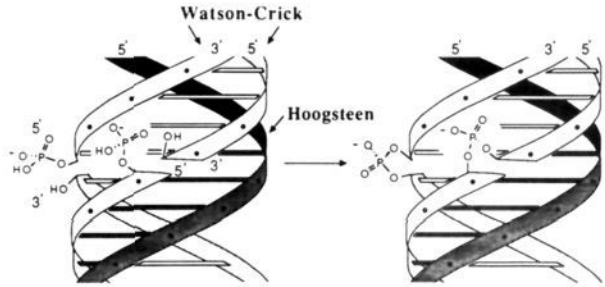

Figure 1. The 5′-phosphate and 3′-hydroxyl termini of two blunt-ended DNA duplexes can be aligned for condensation by association of an oligonucleotide template in a triple-helical complex.

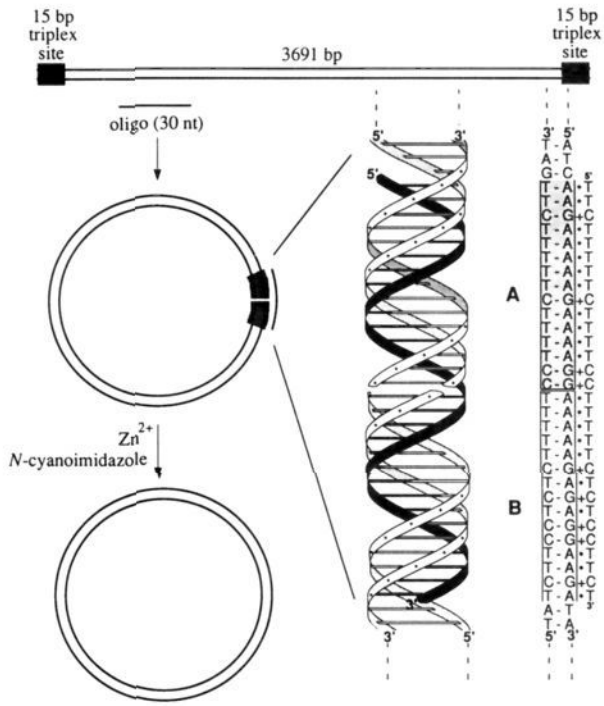

Figure 2. The substrate linear double-helical DNA can be covalently circularized in a reaction that requires single-stranded template, purine tracts on both ends of the double-helical DNA termini, and activation by *N*-cyanoimidazole (NCI) in the presence of Zn^{2+}.

A mixture of double-helical DNA (1.7 nM), 30-mer template (17 nM), and $ZnCl_2$ (20 mM) was allowed to react with the condensing agent *N*-cyanoimidazole (0.1 mM).[10,11] After 7 h (20 °C, pH 4.9), the reaction products resulting from single- and double-strand ligations were separated by agarose gel electrophoresis in the presence of ethidium bromide (Figure 3). Covalent closure of one strand of the linear DNA produces a circular molecule (form II) that migrates more slowly in the gel than the linear starting material (form III). If both strands of the plasmid are covalently closed, the circular DNA is positively supercoiled by intercalation of ethidium bromide (EB) contained in the gel. The positively supercoiled DNA (form I°) migrates more rapidly than the linear starting material (form III) and the negatively supercoiled DNA isolated from bacteria (form I).

Several DNA products from the chemical ligation reaction are observed (Figure 3, lane 7), with gel electrophoretic mobilities identical with those produced by treatment of the linearized

(1) Naylor, R.; Gilham, P. T. *Biochemistry* **1966**, *5*, 2722–2728.
(2) (a) Orgel, L. E.; Lohrmann, R. *Acc. Chem. Res.* **1974**, *7*, 368–377. (b) Inoue, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1981**, *103*, 7666–7667. (c) Hill, A. R., Jr.; Kumar, S.; Leonard, N. J.; Orgel, L. E. *J. Mol. Evol.* **1988**, *208*, 91–95. (d) Lohrmann, R.; Bridson, P. K.; Orgel, L. E. *Science (Washington, D.C.)* **1980**, *208*, 1464–1465. (e) Bridson, P. K.; Orgel, L. E. *J. Mol. Biol.* **1980**, *144*, 567–577. (f) Lohrmann, R.; Orgel, L. E. *J. Mol. Biol.* **1980**, *142*, 555–567.
(3) (a) Dolinnaya, N. G.; Sokolova, N. I.; Gryaznova, O. I.; Shabarova, Z. A. *Nucleic Acids Res.* **1988**, *16*, 3721–3738. (b) Sokolova, N. I.; Ashirbekova, D. T.; Dolinnaya, N. G.; Shabarova, Z. A. *FEBS Lett.* **1988**, *232*, 153–155.
(4) Luebke, K. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 8733–8735.
(5) (a) Fareed, G. C.; Richardson, C. C. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 665. (b) Little, J. W.; Zimmerman, S. B.; Oshinsky, C. K.; Gellert, M. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 2004.
(6) Moser, H. E.; Dervan, P. B. *Science (Washington, D.C.)* **1987**, *238*, 645 and references cited therein.
(7) Praseuth, D.; Perroault, L.; Doan, T. L.; Chassignol, M.; Thuong, N.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *58*, 1349.
(8) (a) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859. (b) de los Santos, C.; Rosen, M.; Patel, D. *Biochemistry* **1989**, *28*, 7282. (c) Sklenar, V.; Feigon, J. *Nature* **1990**, *345*, 836.
(9) The plasmid was constructed by the insertion of the sequence 5′-ACGTATCTTCCTCTCTT*TTTAAA*GAGATCTCT*AGGCCTT*$_3$CT$_5$CT-TGATC-3′ (*DraI* and *StuI* sites are 5′-TTTAAA-3′ and 5′-AGGCCT-3′, respectively) into the 3652-bp *DraI* fragment of pBR322. Cleavage with *DraI* and *StuI* yields a linear 3691-bp DNA substrate with one 15-bp purine site for triple-helix formation at each blunt end. The duplex termini resulting from endonuclease cleavage are 3′-hydroxyl and 5′-phosphate.
(10) (a) Kanaya, E.; Yanagawa, H. *Biochemistry* **1986**, *25*, 7423–7430. (b) Ferris, J. P.; Huang, C.-H.; Hagan, W. J., Jr. *Nucleosides Nucleotides* **1989**, *8*, 407–414.
(11) *N*-Cyanoimidazole was prepared as in the following: Giesemann, H. *J. Prakt Chem.* **1955**, *1*, 345–348.

0002-7863/91/1513-7447$02.50/0

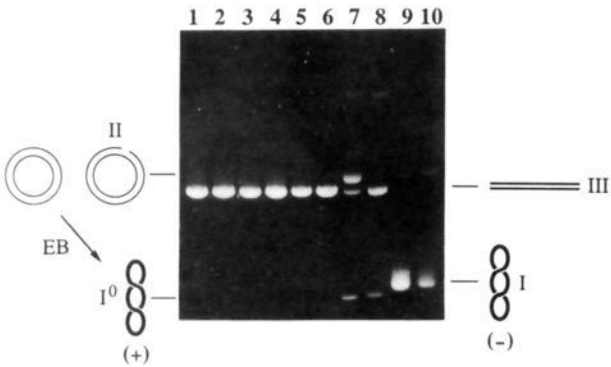

Figure 3. Ligation of linear 3.7-kbp DNA; analysis by electrophoresis in a 1% agarose gel containing 0.4 mg/L ethidium bromide (EB). Lane 7: Plasmid DNA (0.8 μg) was linearized with *Dra*I and *Stu*I,[9] deproteinized with phenol, desalted by elution through Sephadex G-25, and treated in a total volume of 200 μL with 0.1 mM NCI, 20 mM $ZnCl_2$, and 17 nM single-strand template for 7 h at 20 °C (pH 4.9). The reaction mixture was desalted with Sephadex G-50 and concentrated under vacuum before loading on the gel. Control experiments: lane 1, linearized DNA; lane 2, NCI was omitted; lane 3, single-stranded template was omitted; lane 4, the single-stranded template was replaced by the 30-nucleotide Watson–Crick complement of the plasmid purine strand; lanes 5 and 6, *Stu*I and *Dra*I, respectively, were omitted from the enzyme digest;[9] lane 10, *Dra*I and *Stu*I were both omitted from the enzyme digest; lane 8: Plasmid cut with *Dra*I and *Stu*I was treated for 7 h at 20 °C with 2000 units of T4 DNA ligase (New England Biolabs) in 200 μL of a solution containing 20 mM $MgCl_2$, 50 mM Tris·HCl, pH 7.0, and 15 mM ATP. Lane 9: 0.8 μg of supercoiled plasmid DNA (form I).

plasmid with T4 DNA ligase and ATP (lane 8). The similarity of products of the nonenzymatic and enzymatic reactions indicates that they are the results of end-to-end ligation. In controls, no reaction is observed in the absence of the oligodeoxyribonucleotide template (lane 3); in the presence of an oligodeoxyribonucleotide (30-mer) complementary in the Watson–Crick sense to the purine tract of the plasmid ends (lane 4); and in the case of the plasmid containing only one triple helix forming site at the termini (lanes 5 and 6).[12] These requirements for triple-helix sites at both termini and for the oligonucleotide complementary in the Hoogsteen sense support a model in which the oligonucleotide behaves as a template for ligation by combining the duplex termini in a local triple helix. Requirement of the condensing agent *N*-cyanoimidazole suggests that the circularization is covalent (lane 2).

One reaction product (15% yield) migrates slightly faster in the gel than negatively supercoiled plasmid (form I).[13] We assign this band to positively supercoiled plasmid (form I°). This product is circularized plasmid in which *both strands have been ligated*. The predominant reaction product (45% yield) is identified as form II DNA. This product may arise from circularization and ligation of only one of the two strands or from nicking of the doubly ligated product during the course of the reaction and workup.[14]

Approximately 15% of the 3.7-kbp DNA is converted into products with slower gel mobilities than form II DNA (lane 7). By comparison with the mobility of a DNA size standard (data not shown), the most abundant of these can be identified as a 7.4-kbp linear dimer of the starting material. The other two bands can be assigned to the circularized dimer, covalently closed on one and both strands. Consistent with these tentative assignments, the yields of the three products decrease with decreasing DNA concentration in the reaction mixture.

Phosphodiester bond formation in the chemical ligation reaction would afford a cleavage site for the restriction endonuclease *Mnl*I. *Mnl*I hydrolyzes the phosphodiesters on both strands of double-helical DNA, seven base pairs to the 3′ side of the sequence 5′-CCTC-3′. Digestion of the products of the nonenzymatic ligation reaction with *Mnl*I reveals a set of restriction fragments consistent with enzymatic cleavage at the ligation site.[15] Cleavage by a restriction endonuclease supports the identification of the bonds formed in the ligation reaction as phosphodiesters.[16]

In summary, we find that blunt-ended double-helical DNA can be ligated on one and both strands by chemical methods. Because this template-directed approach involves sequence information from the double-helical substrate in the ligation reaction, the nonenzymatic reaction supercedes in specificity the related enzymatic process.

Acknowledgment. This work was supported in part by the National Institutes of Health (GM-35724) and the Caltech Consortium in Chemistry and Chemical Engineering (founding members: E. I. du Pont de Nemours and Company, Inc., Eastman Kodak Company, Minnesota Mining and Manufacturing Company, and Shell Oil Company Foundation). We are grateful for a National Research Service Award to K.J.L. from the National Institute of General Medical Science.

(12) A small amount of reaction (~1%) was observed when the plasmid was linearized only with *Dra*I, which could be due to a weak interaction of the template with the DNA to the side of the restriction cut removed by *Stu*I.

(13) Yields were estimated by densitometric analysis of a negative photograph of the agarose gel. This analysis did not account for differential staining of the different forms of DNA by ethidium bromide and probably underestimates the yield of form I° DNA, which is expected to have a lower affinity for ethidium bromide than linear and nicked circular DNA.[14]

(14) Approximately 15% of the plasmid is nicked after treatment under the nonenzymatic ligation conditions (lane 10).

(15) In a control experiment, the ligation reaction was performed on an otherwise identical plasmid with the sequence 5′-CTTC-3′ in the template binding site replacing the *Mnl*I site. Digestion of the products of this reaction with *Mnl*I and electrophoretic analysis confirmed that the linkages formed in the ligation reaction are stable to the enzyme digestion conditions.

(16) This analysis does not rigorously prove the identity of the linkages formed in the ligation reaction. It is conceivable that *Mnl*I can catalyze the cleavage of linkages other than phosphodiesters or that the enzyme is not absolutely specific for the phosphodiester 7 base pairs from its recognition sequence.[17] However, the nonenzymatic ligation products can be grown in *Escherichia coli*, confirming that at least some of the linkages formed are phosphodiesters. We find that use of an alternate-strand binding template[18] enables template-directed double-strand ligation that re-creates a restriction enzyme cleavage site. The linkages formed can be enzymatically cleaved, lending further support to the conclusion that phosphodiester linkages are formed in the nonenzymatic reaction reported here.

(17) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Press: Cold Spring Harbor, 1989; Vol. 1, p 5.8.

(18) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435–2437.