Sequence Specific Nonenzymatic Ligation of Single- and Double-Stranded DNA by Triple Helix Formation

Thesis by

Kevin J. Luebke

In Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1992

(Submitted April 17, 1992)

Acknowledgements

I would like to thank my advisor, Professor Peter Dervan, for his guidance and support in this work. I have learned many scientific lessons from him which will certainly continue to guide me throughout my research career.

I could not have asked for better labmates than Kim Liu, Warren Wade, Jumi Shin, and Milan Mrksich. Their camaraderie and support are deeply appreciated. I am grateful to Sam Gellman for providing insight and encouragement when I was getting started in the lab. I would also like to thank Sheila David, Kim Waldron, and Cathy Murphy (and all of the Dervan and Barton group members at the Beckman Institute) for making me feel welcome in their lab and keeping me company while I wrote this thesis. Scott Singleton and Karen Draths have helped in a number of ways with the preparation of this thesis, including providing critical review of much of the manuscript.

David Mack and Michelle Unterman deserve much credit for excising me from the lab at times when that task was both difficult and necessary. I would like to thank Dane Wittrup for his friendship; for three years, Dane and I shared an apartment and observations about science and life. I enjoyed that time very much. I have often relied on Mark Distefano to lend a helping hand, a valuable insight, or a sympathetic ear. Laura Kiessling is a great person with whom to laugh, cry, talk science, and generally hang around. I dearly value the friendship of Martha Oakley, a caring and thoughtful person who has shared much with me and from whom I have learned a great deal. Kathlynn Brown is a wonderful friend

who never fails to bolster my outlook, refresh my perspective, and keep me having fun.

Finally, I would like to thank Mom, Dad, Ken, and Suzie. Their love and support are a big part of me without which this thesis would not have been possible.

Abstract

Phosphodiesters link the units of chemical information within nucleic acids. The formation of a phosphodiester linkage by condensation of phosphate and hydroxyl termini of DNA requires activation of the phosphate toward nucleophilic displacement and proper placement of the hydroxyl for nucleophilic attack on the activated phosphate in competition with water. Sequence information is transferred when a nucleic acid template promotes this condensation reaction by the sequence specific formation of a complex in which condensing functionalities are juxtaposed. This thesis describes investigations of the sequence specific formation of phosphodiester linkages by the assembly of triple-helical complexes. In the first part of Chapter 1, sequence specific recognition of double-stranded DNA by triple helix formation is reviewed. Structural features of nucleic acid triple helices, the sequence specificity of their formation, and functions associated with them are considered in this review. In the second part of Chapter 1, literature regarding template-directed formation of phosphodiesters in aqueous solution is reviewed.

Chapter 2 describes investigations of the sequence specific formation of a phosphodiester linkage between pyrimidine oligodeoxyribonucleotides using a double-stranded DNA template to juxtapose their termini in a triple helix. Several approaches to activation of the condensing phosphate were explored. The most effective of these was activation *in situ* with the condensing agent *N*-cyanoimidazole with which condensation yields greater than 80% could be obtained. The reaction was directed by the double-

stranded template and could be shown to form a 3',5' phosphodiester linkage between the two oligodeoxyribonucleotides. A single mismatch in one of the condensing oligodeoxyribonucleotides at the condensing terminus resulted in at least a 25 to 60 fold decrease in the rate of the reaction. The relevance of this reaction to the detection of sequences, nucleic acid catalysts, and the evolution of template-directed information transfer is discussed.

Chapter 3 describes the nonenzymatic sequence specific ligation of blunt-ended duplex DNA by triple helix formation. Using a pyrimidine oligodeoxyribonucleotide as a template and *N*-cyanoimidazole as a condensing agent, a double-stranded plasmid with homopurine tracts at one 3' terminus and one 5' terminus could be covalently circularized in yields exceeding 50%. Ligation on both strands was demonstrated in some of the circularized product. Ligation of duplexes with homopurine tracts at their 3' termini was directed by a pyrimidine oligodeoxyribonucleotide of two segments joined 3' to 3' through an abasic linker, creating a duplex of the sequence type 5'-$(\text{purine})_m(\text{pyrimidine})_n$-3'. The linkages formed in the ligation reaction were demonstrated to be substrates for a restriction endonuclease, identifying them as phosphodiesters. The sequence specificity of these reactions is not accessible by enzymatic ligation of double-stranded DNA.

Table of Contents

List of Figures and Tables

Chapter 1

Chapter 2

Figures

Tables

Chapter 3

Chapter 1
Introduction

Sequence Specific Recognition of Double-Stranded DNA by Triple Helix Formation

The translation of chemical information into a structural and behavioral phenotype is a defining aspect of living systems.[1,2] The three-dimensional structures and catalytic activities of biological macromolecules are encoded in their monomer sequences, which are decoded ultimately from the monomer sequences of nucleic acids. Molecular recognition provides the basis of the codes through which chemical information is translated. For example, sequence information can be transcribed from a polydeoxyribonucleotide to a complementary polyribonucleotide by virtue of the specific, noncovalent interactions comprised by Watson-Crick base pairing (Figure 1.1).

The same code, relating adenine with uracil and guanine with cytosine, governs another information transfer process in which the monomer sequence of a polyribonucleotide is translated into a sequence of amino acids. Molecular recognition of amino acids and their cognate transfer RNA molecules by aminoacyl tRNA synthetases associates the amino acids with trinucleotide sequences. The information within a polyribonucleotide message is translated

into a polypeptide through the Watson-Crick base pairing of these sequences with their complements in the message.[3]

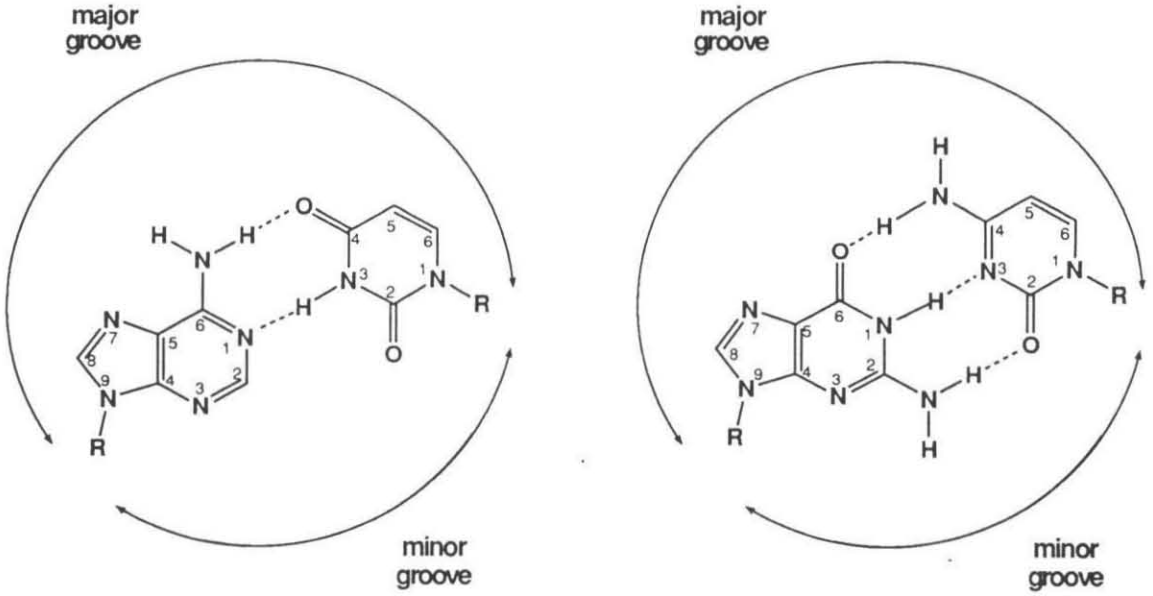


Figure 1.1. The Watson-Crick base pairs of adenine with uracil (left) and guanine with cytosine (right).

In addition to coding for the sequences of polypeptides, DNA carries information that regulates the production of polypeptides. In many cases, this information is translated by sequence specific recognition of double-stranded DNA. The regulation of protein synthesis by transcription factors and transcriptional repressors that bind sequence specifically to duplex DNA is an important example of this type of information transfer.[4,5]

The sequence-specific recognition of double-stranded DNA by polynucleotides.

The sequence specific recognition of double-stranded DNA by polynucleotides was suggested by Jacob and Monod in 1961 in their original articulation of the regulator gene hypothesis.[6] They speculated that a

polynucleotide possessing a base sequence complementary to an operator site could behave as a transcriptional repressor by binding to that site. After the demonstration of transcriptional repression by proteins in bacterial systems,[7-10] Miller and Sobell considered an amended form of the original suggestion by Jacob and Monod. Miller and Sobell imagined the repressor to be a ribonucleoprotein complex that bound specifically to operator sites of the DNA duplex. They developed this hypothesis further by proposing a model in which the polynucleotide component of the repressor was accomodated in the major groove of the double helix, forming a local triple helix.[11]

The model that Miller and Sobell offered was based on observations of three-stranded complexes of homopurine and homopyrimidine nucleic acids.[12-14] The first of these was reported by Felsenfeld, Davies, and Rich, who found, while studying mixtures of poly(U) and poly(A), that mixing curves and sedimentation coefficients indicated the formation of a stable 2:1 complex, 2poly(U)•poly(A), in the presence of 10 mM $MgCl_2$.[15] The structural details of the complex were unknown; however, it was proposed that base triplets can be formed in which one strand of poly(U) associates with poly(A) through Watson-Crick hydrogen bonding and another strand of poly(U) associates with poly(A) through an alternate base pairing arrangement. In the arrangement proposed, O4 and N3 of uracil are hydrogen bonded to N6 and N7 of adenine, respectively. This base-pairing scheme has been termed Hoogsteen hydrogen bonding, because it was first observed directly by Hoogsteen in the crystal structure of a complex of 1-methyl thymidine and 9-methyl adenine (Figure 1.2).[16-19]

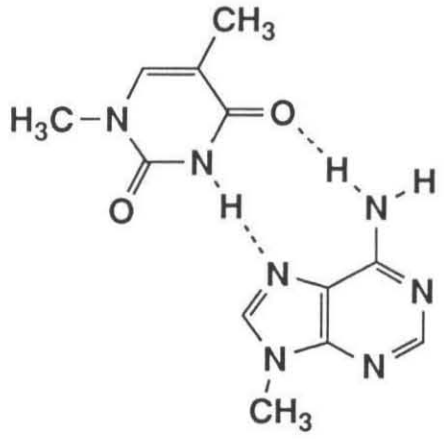
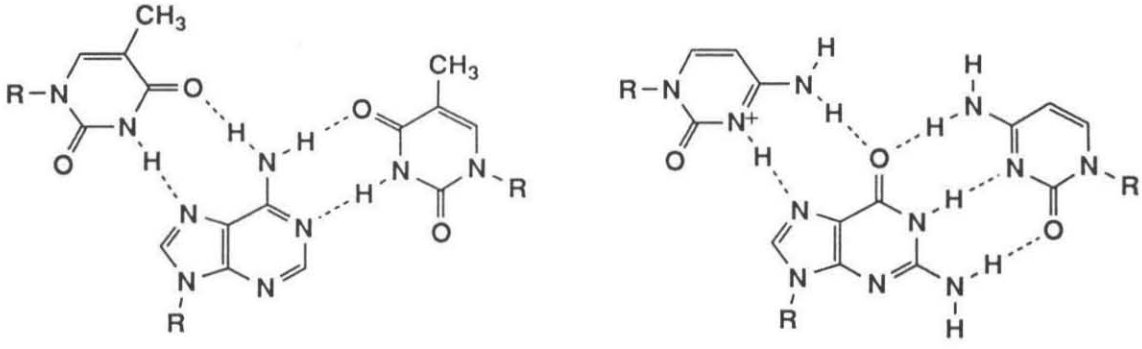


Figure 1.2. The crystal structure reported by Hoogsteen of a complex of 1-methyl thymidine and 9-methyl adenine.[16]

Three-stranded complexes of the type 2poly(C)•oligo(G) had also been reported by Lipsett and coworkers.[20,21] These complexes were favored by low pH, and data from infrared spectroscopy suggested protonation at N3 on the cytosines of one strand.[22] Miller and Sobell noted that a cytosine•guanine•cytosine triplet could be formulated in which a protonated cytosine is hydrogen bonded to guanine with the same geometry as the Hoogsteen configuration of adenine and uracil. Thus, isomorphous base triplets UAU (or TAT) and C+GC could be envisioned (Figure 1.3). The existence of these isomorphous triplets would allow the formation of a three-stranded complex between a homopurine•homopyrimidine duplex of any sequence and a corresponding homopyrimidine strand.

In the model for gene repression that Miller and Sobell proposed, operator sequences would contain stretches of purines that could accomodate pyrimidine sequences from the hypothetical ribonucleoprotein repressor by Hoogsteen hydrogen bonding, forming a local triple helix. Affinity and specificity of the repressor for the operator site would be determined by the

specific hydrogen bonding between operator and repressor sequences, T recognizing AT base pairs and protonated C recognizing GC base pairs. The plausibility of the recognition model that Miller and Sobell proposed rested on the prediction that triple-stranded complexes could be formed containing both G and A. That prediction was confirmed by Morgan and coworkers when they reported that the synthetic DNA duplex poly(dG-dA)•poly(T-dC) associates with poly(U-C) or poly(T-dC) in acidic solution to form a triple-stranded complex.[23-26]



TAT base triplet C^+GC base triplet

Figure 1.3. Isomorphous base triplets formed by association of a pyrimidine strand in the major groove of double helical DNA by Hoogsteen hydrogen bonding.

The Structure of Nucleic Acid Complexes Containing Two Pyrimidine Strands and One Purine Strand. Though the earliest reports of three-stranded nucleic acid complexes included structural proposals, a detailed, experimentally verified structural description of nucleic acid triplexes is still emerging. Ultimately, this structural description would include information about backbone conformation and base stacking as well as hydrogen bonding and

strand orientation. A high-resolution x-ray crystal structure of a triple helical DNA or RNA has not yet been reported.

Assuming an *anti* conformation of the bases with respect to the sugars, Watson-Crick and Hoogsteen base paired polypyrimidine strands would be oriented antiparallel to each other in a three-stranded complex. The first experimental confirmation of this prediction came from affinity cleaving studies by Moser and Dervan.[27] In those experiments, an oligodeoxyribonucleotide, $(dT)_{15}$, was chemically synthesized in which one thymidine (T*) was modified at C5 by covalent attachment of ethylenediamine tetraacetic acid (EDTA). In the presence of oxygen and a reducing agent, the iron complex of EDTA is known to generate a diffusible species that oxidatively cleaves the DNA backbone.[28-31]

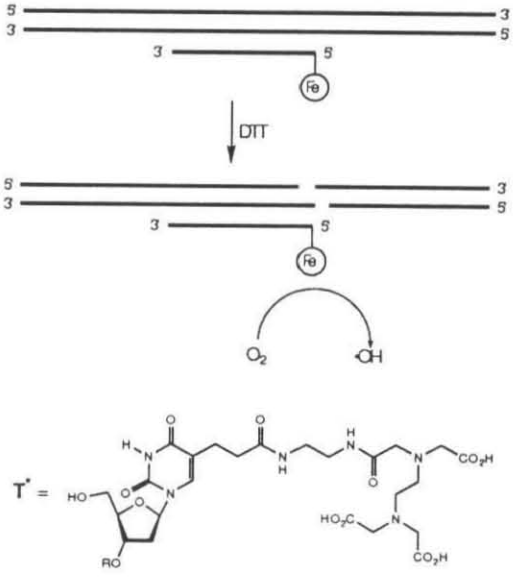


Figure 1.4. A schematic representation of the triple helix affinity cleaving experiment. An oligonucleotide containing T* associates with a DNA duplex by triple helix formation. In the presence of Fe(II), O_2, and dithiothreitol (DTT), the DNA is cleaved proximal to the T*.

The oligodeoxyribonucleotide in which the 5' terminal thymidine was derivatized with EDTA was found to cleave duplex DNA containing a stretch of fifteen contiguous AT base pairs in the presence of Fe(II), O_2, and dithiothreitol. The pattern of cleavage was centered at the 5' end of the stretch of 15 adenines and the 3' end of the stretch of thymidines. When the thymidine derivatized with EDTA was located centrally in $(dT)_{15}$, cleavage of the same duplex was located centrally within the fifteen AT base pairs. Therefore, the single-stranded oligodeoxyribonucleotide binds to the duplex with its 5' terminus proximal to the 5' side of the stretch of adenines, that is, parallel to the purine strand and antiparallel to the pyrimidine strand of the Watson-Crick duplex. In addition, analysis of the cleavage pattern confirmed that the diffusible species responsible for strand oxidation was generated in the major groove of the Watson-Crick duplex, consistent with Hoogsteen hydrogen bonding of the oligodeoxyribonucleotide.[27]

Recent nuclear magnetic resonance studies have further substantiated a picture of the DNA triple helix in which a pyrimidine strand is associated with a Watson-Crick duplex in the major groove by Hoogsteen hydrogen bonding and demonstrated the presence of cytosines protonated at N3 in the Hoogsteen base paired strand.[32-34] The pyrimidine imino proton resonances were observed and assigned in a 2 to 1 complex of $d(TC)_4$ with $d(GA)_4$ at a pH below 6.1. Protonation of cytosines at N3 in one of the strands was thereby confirmed. Furthermore, a through-space coupling (NOE) was observed between the T and C imino protons in that strand and H8 of the associated purine, confirming Hoogsteen hydrogen bonding.[33]

Although a number of studies have addressed the issue of the backbone conformations in triple helices, a general, definitive conclusion has not yet

been provided. X-ray fiber diffraction studies of poly(U)•poly(A)•poly(U)[35] and poly(T)•poly(dA)•poly(T)[36] performed by Arnott and coworkers revealed an axial rise per residue of 3.26 Å and a helical twist of 30° (12 base triplets per turn) in both triplexes. Comparison with calculated diffraction patterns led these investigators to conclude that the experimental diffraction patterns for DNA and RNA were most consistent with an A' helix having base tilts of 7-9° and C3'-endo sugar puckers in all three strands. The structure of poly(U)•poly(A)•poly(U) has also been studied by Raman spectroscopy with the conclusion that the backbone conformations of the two pyrimidine strands in the triple helix differ from each other. The C3'-endo conformation was assigned to the sugars of the Watson-Crick pyrimidine strand, and it was suggested that the sugars of the other strand of poly(U) might be in the C2'-endo conformation.[37]

One- and two-dimensional NMR studies of DNA triple helices have led to other conclusions. Patel and coworkers studied a DNA triplex of 11-mers containing both G and A in the purine strand.[32] From the chemical shifts of the adenine H8 signals, these investigators concluded that the stacking of the purines was consistent with an A-form helix. The A-form is defined by C3'-endo sugar puckers, though Patel and coworkers did not interpret their data in terms of the sugar conformations in the individual strands. Meanwhile, in their studies of triplexes formed from $d(TC)_4$ and $d(GA)_4$, Rajagopal and Feigon found evidence in NOE intensities that the two pyrimidine strands had C3'-endo sugar puckers; whereas, the purine strand did not.[33] Later, Macaya, Schultze, and Feigon performed an analysis of the through-bond coupling constants in the sugar residues of an intramolecular triple helix.[38] This analysis indicated that the sugars of all of the purines and all of the

thymidines except the terminal one are predominantly S-type (near C2'-endo) and that the sugars of only some of the cytidines are largely of the N-type (near C3'-endo) conformation. Most of these N-type cytidines were found to be in the Hoogsteen base paired segment.

As with double-helical nucleic acids, structural parameters such as sugar pucker and base tilt of triple-helical nucleic acids are likely to depend on the presence or absence of the 2'-hydroxyl, the local base sequence, environmental factors (e. g., ionic strength, pH, temperature), and the structural context of the triple helix (e. g. intrastrand triplex[34,39], local triplex within a larger duplex, etc.).[40] Discrepancies between different studies in the assigned backbone conformations could be due in part to the roles of these elements in the various experiments.

Sequence Specificity of Triple Helix Formation. The applicability of local triple helix formation to the sequence specific recognition of double-stranded DNA depends on 1) the existence of a code relating the sequence of an oligonucleotide to a sequence of double-stranded DNA, 2) the fidelity of the code, and 3) the generalizability of the code to a variety of duplex sequences. Hoogsteen hydrogen bonding of pyrimidines to purines provides the basis of one code in which T recognizes an AT base pair and C recognizes a GC base pair. Thus, a local triple helix can be formed within duplex DNA by association of a pyrimidine oligonucleotide having a corresponding sequence parallel to a duplex purine sequence. Recognition of a single homopurine site within a larger DNA duplex was first demonstrated by the method of affinity cleaving. Moser and Dervan showed that both strands of a 4.06 kilobasepair plasmid could be cleaved at a single site having the sequence 5'-d[$A_5(GA)_5$]-3'

by an oligodeoxyribonucleotide with the sequence 5'-d[T*$T_4$$(CT)_5$]-3', where T* is thymidine modified at the C5 position by EDTA.[27]

One indication of the specificity of local triple helix formation with pyrimidine oligodeoxyribonucleotides is the site specific affinity cleavage of genomic DNA. In one example, cleavage on both strands of the 48,502 base pair genome of bacteriophage lambda was exacted at an eighteen base pair sequence of purines, 5'-d[$A_4GA_6GA_4GA$]-3', with the corresponding pyrimidine oligodeoxyribonuclotide•EDTA.[41] Moreover, the 340 kilobase pair chromosome III of *Saccharomyces cerevisiae* has been cleaved at a 20 base pair purine target site by a pyrimidine oligodeoxyribonucleotide with EDTA•Fe at its 5' and 3' ends.[42] Clearly, a pyrimidine oligodeoxyribonucleotide can recognize its target sequence within the context of a large number of other sequences.

A more precise description of the specificity of local triple helix formation in terms of relative affinities of oligonucleotides for binding sites[43] can be expressed as differences in free energy of association or, equivalently, ratios of equilibrium binding constants.[44,45] Quantitative affinity cleavage titration has provided binding constants for the association of matched and mismatched pyrimidine oligodeoxyribonucleotides with a fifteen base pair homopurine site in a 339 base pair DNA duplex. At 24 °C, a single internal mismatch was found to decrease the binding constant by a factor of 60 to 150. Thus, oligodeoxyribonucleotides can discriminate between binding sites on duplex DNA that differ by only a single base pair by two orders of magnitude.[46]

Cooperative association of two oligonucleotides with their target sites to form local triple helical structures might enhance the sequence specificity

of triple helix formation. Two pyrimidine oligodeoxyribonucleotides will associate cooperatively at adjacent sites on a DNA duplex by triple helix formation. The cooperativity might be due to favorable base stacking interactions between adjacent oligodeoxyribonucleotides, hydrogen bonding between 5' and 3' terminal hydroxyls, and/or induced conformational changes in the Watson-Crick strands propagated to adjacent sites.[47] Cooperativity between oligonucleotides in the formation of triple helices can also be effected by the interaction of discrete dimerization domains. Oligodeoxyribonucleotide segments that dimerize in a Watson-Crick double helix can be appended to oligodeoxyribonucleotides that bind to adjacent sites on a DNA duplex to confer cooperativity to triple helix formation.[48]

Several strategies have been applied to the recognition of sequences containing all four base pairs. In addition to the TAT and C+GC base triplets, a GTA triplet can be formed in some sequence contexts.[49,50] The design of non-natural bases for the recognition of pyrimidines within stretches of purines is another approach to generalizing the pyrimidine code for triple helix formation.[51] A third approach, extending the range of sequences that can be recognized to those of the general type 5'-$(purine)_m(pyrimidine)_n$-3', is the association of a pyrimidine oligodeoxyribonucleotide with purines on alternate strands of the double helix (Figure 1.5). A pyrimidine oligodeoxyribonucleotide composed of two segments coupled 3' to 3' through a 1,2-dideoxy-D-ribose linker satisfies the structural requirements for binding alternate strands of duplex DNA. The abasic linker connects the two segments across the major groove, bridging two base pairs in a nonspecific manner at the purine-pyrimidine junction of the target site.[52]

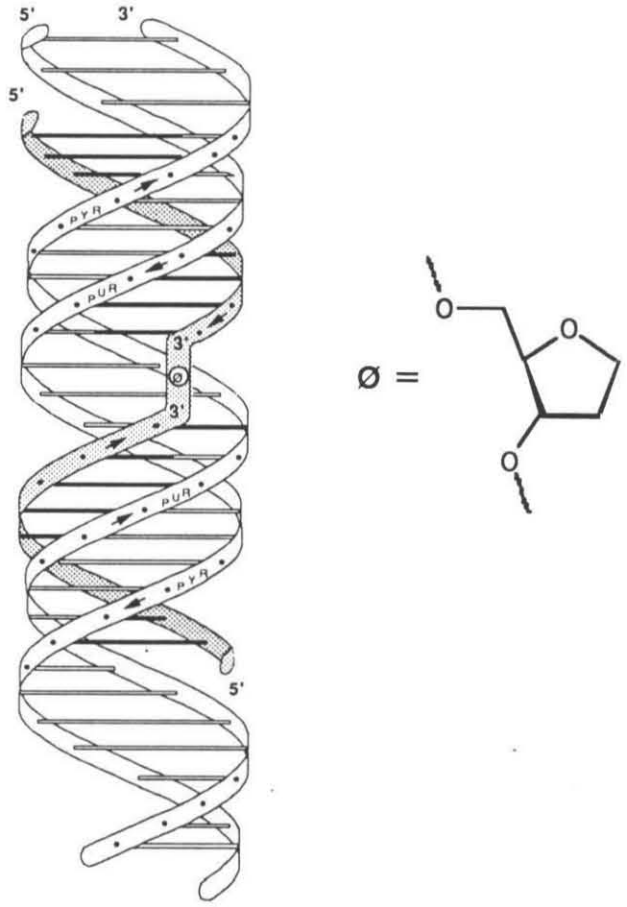


Figure 1.5. Ribbon diagram of an alternate-strand triple helix. A pyrimidine oligodeoxyribonucleotide (shaded) with segments coupled 3'-3' through a 1,2-dideoxy-D-ribose linker (ϕ) associates with purines on alternate strands of the double helix.

In addition to triple helices of the general type pyrimidine•purine•pyrimidine, another class of three-stranded structures has been found, comprising triplexes of the type purine•purine•pyrimidine.[53-58] Specificity in the formation of triplexes of this type is based on the recognition of GC base pairs by G and the recognition of AT base pairs by A or T (Figure 1.6.). Intermolecular[53,57] and intramolecular[56,58,59] triplexes of this class have been studied. Whereas

triplexes of the type pyrimidine•purine•pyrimidine are most stable within AT rich sequences, triplexes of the type purine•purine•pyrimidine are most stable within GC rich sequences.

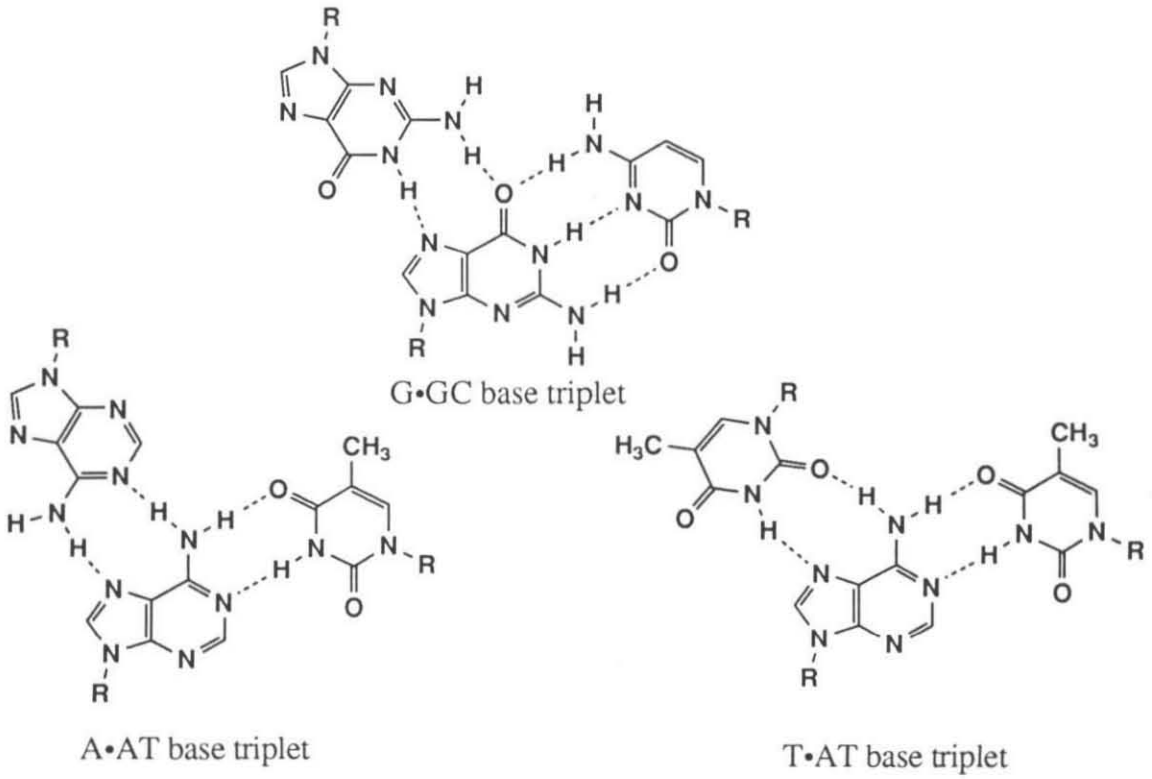


Figure 1.6. Models proposed by Beal and Dervan for G•GC, A•AT,and T•AT base triplets within a purine•purine•pyrimidine triple helix in which the purine strands are antiparallel and the bases are in the *anti* conformation.

Functional Roles for Nucleic Acid Triple Helix Formation. The ribonucleoprotein transcriptional repressor postulated by Miller and Sobell has not been discovered. The naturally occurring transcriptional repressors characterized to date contain no polynucleotide or oligonucleotide component. Nonetheless, there are hints of a natural, biological role for triple helix formation.[60-65] Oligopurine-oligopyrimidine mirror repeats, which have been

shown to form intramolecular triplexes in underwound DNA, are found in or close to transcriptional units and recombination hot spots in eukaryotes.[58,66-76] Also, stretches of homopurine•homopyrimidine sequences have been found to be 3.5-4 times more abundant than statistically predicted in higher eukaryotes and eukaryotic viruses.[60,61] A monoclonal antibody to triplex DNA binds to eukaryotic chromosomes,[77] and recently, a protein was isolated from human cells with a modest specificity for TAT triple helical DNA.[78]Furthermore, a local three-stranded structure, containing two identical and parallel strands, is likely to be an intermediate in the process of homologous recombination.[79]

While its function in natural systems remains to be delineated, the nucleic acid triple helix provides a structural basis for the design of agents for the sequence specific modification and cleavage of double-stranded DNA and the sequence specific inhibition of DNA binding proteins. These functions have application in molecular biology as tools for gene isolation, recombinant DNA manipulations, DNA sequencing, and genomic mapping. Furthermore, they might find utility in the treatment of genetic, neoplastic, and viral diseases.

Local triple helix formation can interfere in a sequence specific manner with the actions of DNA binding and modifying agents, including DNA binding proteins. DNA cleavage by methidiumpropylEDTA•Fe(II)[80], alkylation by dimethyl sulfate,[80,81] and photoinduced pyrimidine dimerization[82,83] are inhibited at sites of triple helix formation within a DNA duplex. The endonuclease DNAseI is prevented from cleaving duplex DNA in triple helical regions,[80] and triple helix formation inhibits restriction endonucleases with recognition sequences overlapped by the bound

oligodeoxyribonucleotide.[84-86] Inhibition of restriction methylases by triple helix formation is the basis for the "Achilles' heel" approach to enzymatic cleavage of DNA at sites defined by triple helix formation.[87-89] In this technique, triple helix formation protects an overlapping methylase site from modification while unprotected methylase sites are enzymatically rendered refractory to endonuclease cleavage. After triple helix disruption, restriction enzyme digestion results in highly efficient cleavage at the unmethylated site. Triple helix formation has been found to block binding of the eukaryotic transcription factor SP1[86] and to repress transcription *in vitro*.[57]

Sequence specific cleavage of double-stranded DNA can be effected by formation of a local triple helix with an oligonucleotide to which a DNA cleaving agent has been appended.[27,80,90-93]One example of this function of the local triple helix is the affinity cleavage experiment described above.[27] In this experiment, the cleaving agent, EDTA•Fe, is sequence neutral, producing strand cleavage wherever the triple helix forms.[94,95] Another sequence neutral DNA cleaving moiety that has been used in conjunction with triple helix formation is the *p*-azidophenacyl group. Attached to the terminal phosphate of an oligodeoxyribonucleotide the *p*-azidophenacyl group can be photocrosslinked to a bound DNA duplex, and crosslinks can be converted to strand cleavage under alkaline conditions.[92,93]

The sequence specificity of triple helix formation can be coupled with additional specificity by functionalizing an oligodeoxyribonucleotide with a base-specific DNA modifying agent. For example, an oligodeoxyribonucleotide derivatized at its 5' terminus with a bromoacetyl moiety and bound in the major groove of duplex DNA places the electrophilic center proximal to a guanine two base pairs to the 5' side of the binding site. Attack of the

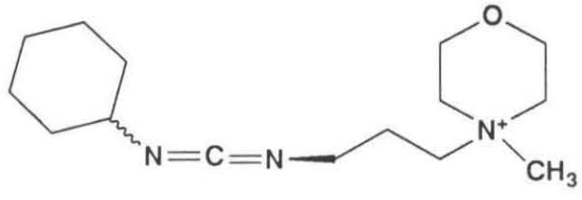
electrophilic carbon at N7 of that guanine results in covalent attachment of the oligodeoxyribonucleotide to the duplex. Treatment with base results in depurination and cleavage of the DNA backbone at the position of alkylation.[80,96]

Sequence specific modification and cleavage of DNA exemplify a more general function of local triple helix formation: assembly of the triple helical complex can place reactive groups in proximity, promoting their reaction. The reaction chemically identifies the location of sequence information that has been recognized by association of an oligonucleotide with a DNA duplex. Sequence information could be transferred between double- and single-stranded nucleic acids if formation of a triple helix promoted a chemical reaction that coupled units of sequence information. Condensation of phosphate and hydroxyl strand termini would create phosphate diester linkages, resulting in products with continuous phosphodiester backbones.

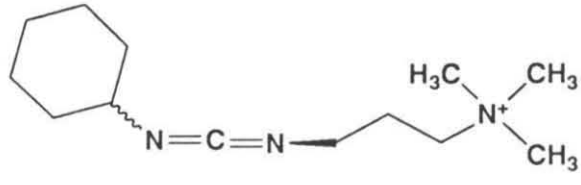
Formation of a phosphorus-oxygen bond between phosphate and hydroxyl termini of DNA in aqueous solution requires chemical activation of the phosphate for nucleophilic substitution and positioning of the hydroxyl for attack on the activated phosphate in competition with water. This esterification reaction is accomplished enzymatically by DNA ligases, which utilize energy from an ATP or NAD cofactor to activate the phosphates.[97,98] Because of the requirement for proximity of the condensing termini, the assembly of a triple helical complex offers an approach to nonenzymatic sequence specific ligation of DNA. A review of some of the previously pursued approaches to formation of phosphodiesters in aqueous solution follows.

Template-Directed Formation of Phosphodiesters in Aqueous Solution

Activation of Phosphates with Water Soluble Carbodiimides. Investigations of the formation of phosphodiesters in aqueous solution were first reported by Naylor and Gilham. They found that water soluble carbodiimides **I** and **II** (Figure 1.7) could be used to activate phosphate monoesters in aqueous solution, promoting the quantitative conversion of nucleotide 2'(3') phosphates to the corresponding nucleotide 2',3' cyclic phosphates in less than one hour at pH 6, 25 °C.[99] A probable mechanism for the carbodiimide-promoted formation of phosphodiesters is shown in the scheme in Figure 1.8.[100,101]



I



II

Figure 1.7. Water soluble carbodiimides used in the studies of Naylor and Gilham.[99]

Naylor and Gilham reasoned that internucleotide phosphodiesters might be formed if hydroxyls were positioned for attack on activated

phosphates within base paired complexes of polynucleotides. They found that thymidine hexanucleotide would react in the presence of polyadenylic acid, carbodiimide **II**, and 1 M sodium chloride at -3 °C to produce thymidine dodecanucleotide in 5% yield in 4 days. They observed no production of dodecanucleotide in the absence of polyadenylic acid or carbodiimide and concluded that the polyadenylic acid behaved as a template to align the termini of the oligothymidylates and promote their condensation.[99]

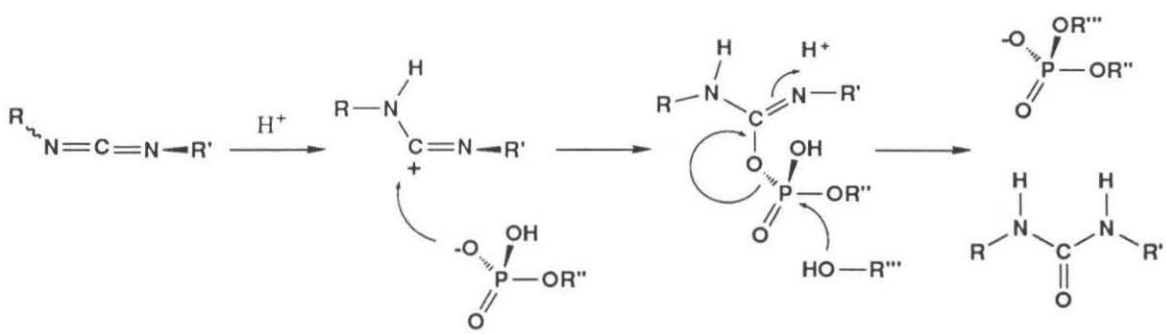
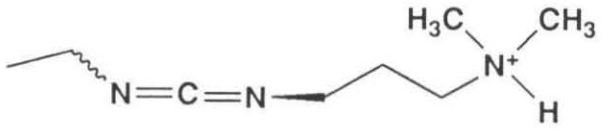


Figure 1.8. A probable mechanism for the carbodiimide-promoted formation of phosphodiesters.

Shabarova and coworkers have used a tetradecadeoxyribonucleotide template of mixed sequence to direct the condensation of two oligodeoxyribonucleotides that form a continuous Watson-Crick double helix with the template.[102] The water soluble carbodiimide 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (Figure 1.9) was used as a condensing agent. A phosphodiester linkage could be formed by condensation of a 3' hydroxyl with a 5' phosphate or by condensation of a 3' phophate with a 5' hydroxyl, though the latter reaction was somewhat more efficient. The reaction was sensitive to mismatches in the double helix, a mismatch at the condensing 3' terminus resulting in a significant decrease in yield under otherwise equal conditions.



EDC

Figure 1.9. The structure of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).

Orgel and coworkers have studied the template-directed condensation of mononucleotides as a nonenzymatic model of RNA polymerase action and prebiotic sequence information transfer.[103] They found that polyuridylic acid promotes the dimerization and trimerization of adenosine-5'-phosphate mononucleotide under conditions which lead to the formation of polyU•A•polyU triple helices (0.15 M $MgCl_2$, 0.2 M NaCl, pH 6.5, 0 °C).[104] Similarly, polycytidylic acid was found to promote the oligomerization of guanylic acid under conditions that lead to the formation of polyC•G double helices.[105] The polynucleotide templates and EDC were required for the reaction. Furthermore, the template directed reactions were base specific, polyU and polyC having no effect on the condensation of their complementary mononucleotide phosphates with cytidine or uridine. Analysis of the oligomerization products revealed that the predominant linkages formed were 5',5' phosphoanhydrides and 2',5' phosphodiesters. Little 3',5' phosphodiester was observed among the products.

Preformation of Phosphorimidazolides. As an alternative to activation of phosphates with water soluble carbodiimides, Orgel and coworkers have utilized nucleotide 5'-phosphorimidazolides as preformed activated nucleotides (Figure 1.10).[106-120] In the presence of a polyuridylic acid

template, $MgCl_2$ (75 mM), and imidazole, adenosine-5'-phosphorimidazolide reacts at 0 °C to form adenosine dimers and trimers linked almost exclusively (approximately 95%) by 2',5' phosphodiesters.[107] Total yields as high as 50% can be obtained. Investigation of the inefficient condensation of uridine-5'-phosphorimidazolide with nucleotide-5'-methylphosphates in the absence of polynucleotide template led Lohrmann and Orgel to conclude that the preponderance of 2',5' linkages over 3',5' linkages in the template-directed oligomerization is due to the combination of a higher intrinsic reactivity of the 2' hydroxyl over the 3' hydroxyl (a factor of 6-9) and association of the reactant nucleotides on the template in an orientation that favors reaction of the 2' hydroxyl (a factor of approximately 3).[115]

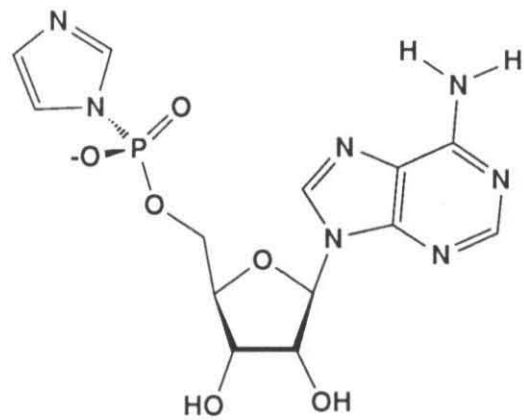


Figure 1.10. Adenosine-5'-phosphorimidazolide.

Metal ions have a significant effect on the efficiency and regiochemistry of the condensation reactions of nucleotide-5'-phosphorimidazolides. A number of divalent metal ions promote the oligomerization of adenosine 5'-phosphorimidazolide in the absence of polynucleotide template.[112,121] For example, in the presence of 25 mM Pb^{2+},

a 50% yield of oligoadenylates can be obtained in the absence of polynucleotide template, under conditions that afford a 2% yield of oligoadenylates in the absence of metal ions. The order of effectiveness of metal ions in the catalysis of adenosine phosphorimidazolide condensation is $Pb^{2+}>Co^{2+}>Zn^{2+}\geq Mn^{2+}>Ni^{2+}>Cd^{2+}>Ca^{2+}\geq Mg^{2+}\approx$no metal$\approx Cu^{2+}\geq Hg^{2+}$. In all cases where the polynucleotide template is absent, the predominant linkage formed (>80%) is the 2',5' phosphodiester. The mechanism by which metal ions catalyze the condensation reactions of nucleotide phosphorimidazolides is not known; however, the metal center is likely to behave as a template for the reactions by orienting phosphate and hydroxyl groups for reaction in a coordination complex. Lewis acid catalysis is also a likely role for the metal ion.[121]

In the condensations of nucleotide phosphorimidazolides on polynucleotide templates, Pb^{2+} and Zn^{2+} not only enhance the efficiency of the oligomerization but also modulate the regioselectivity of phosphodiester formation. In the polycytidylic acid-directed polymerization of guanosine-5'-phosphorimidazolide, Zn^{2+} and Pb^{2+} increase the total yield of oligomer as well as the mean oligomer length.[108,114] Oligomers of 30 to 40 units in length can be produced in the presence of Zn^{2+} and Pb^{2+}; whereas, little polymerization proceeds beyond 10 units under the same conditions in the absence of those metal ions. As in the absence of a metal ion catalyst, the nucleotide residues are predominantly 2',5' linked when Pb^{2+} is the catalyst.[113] However, this regioselectivity is reversed when Zn^{2+} is the catalyst, the 3',5' linkage predominating.[120]

In contrast, the polyuridylic acid-directed oligomerization of adenosine-5'-phosphorimidazolide is not catalyzed by Zn^{2+}, and catalysis by Pb^{2+}

increases the proportion of the 3',5' linkage in the products of that reaction.[111] As well as behaving as templates for orientation of the phosphate and hydroxyl reactants and as Lewis acids, metal ions might effect the efficiency and regiochemistry of polynucleotide directed condensations by binding to sites on the nucleic acid remote from the reactive groups, altering the structure of the helical complex in which condensation occurs and the orientation of the condensing units on the template.[120]

In the absence of metal ion catalysis, the efficiency and regioselectivity of polycytidylate-directed condensation of guanosine-5'-phosphorimidazolides can be altered by substitution on the imidazole moiety.[116,117] In particular, the guanosine-5'-phosphorimidazolide of 2-methylimidazole condenses more efficiently than the guanosine-5'-phosphorimidazolides of imidazole, 2-ethylimidazole, or 4(5)-methylimidazole. Furthermore, condensation of the 2-methyl derivative yields oligomers in which the linkages are almost exclusively 3',5' phosphodiesters. Substitution on the imidazole ring apparently alters the orientation of the phosphorimidazolide with respect to the adjacent 2' and 3' hydroxyls in the base paired complex, thereby altering its reactivity.[117]

The nature of the base pairing between the activated monomer and the polynucleotide template is an additional factor that can effect the efficiency and regioselectivity of phosphorimidazolide condensation. In a one to one complex, 3-isoadenosine (Figure 1.11) is likely to interact with a polyuridylic acid template through Hoogsteen base pairing rather than Watson-Crick base pairing,[122] and the oligomerization of its phosphorimidazolide in the presence of one equivalent (in nucleotides) of polyuridylic acid was found to be more efficient than the corresponding reaction of adenosine-5'-

phosphorimidazolide.[118] Whereas the reaction of adenosine-5'-phosphorimidazolide yields almost exclusively 2',5' linkages, condensation of 3-isoadenosine-5'-phosphorimidazolide produces 3',5' linked oligomers. Thus, the template directed condensation of mononucleotide phosphorimidazolides is sensitive to intrinsic structural features of the monomer-template complex such as the nature of the base pairing interaction.

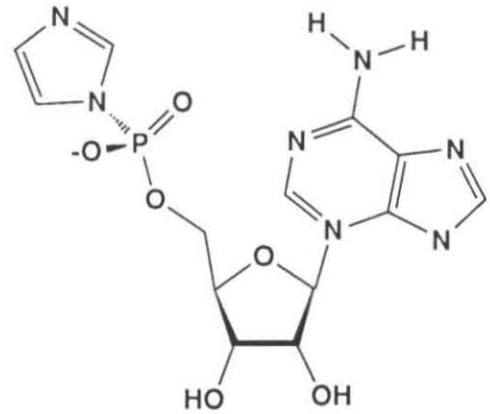
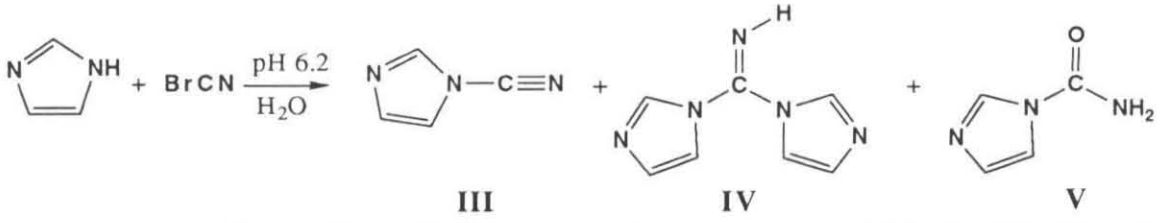


Figure 1.11. 3-Isoadenosine-5'-phosphorimidazolide.

Activation of Phosphates with Condensing Agents Derived from Cyanide. In addition to activation of phosphates as phosphorimidazolides, other alternatives to water soluble carbodiimides have been sought for the formation of phosphodiesters in aqueous solution. Condensing agents derived from cyanide have been studied extensively in this regard because of their possible relevance to prebiotic condensation reactions.[123-127] Diiminosuccinonitrile is produced by oxidation of the hydrogen cyanide tetramer diaminomaleonitrile and has been found to promote the condensation of oligoadenylates in aqueous solution.[125] Cyanamide has been found to bring about the condensation of mononucleotides in the presence of montmorillonite clay.[127]

Cyanogen bromide (BrCN) is a particularly efficient condensing agent.[125,128-130] Using BrCN in the presence of imidazole and divalent metal ions, oligoadenylates can be condensed on a polyuridylate template with an overall yield of 68% in 20 hours at room temperature. As with the template-directed condensations of phosphorimidazolides, metal ions influence the regiochemistry of condensation. The 2',5'- and 3',5'-phosphodiester linkages predominate in the presence of Co^{2+}, Zn^{2+}, and Ni^{2+}, whereas, the 5',5'-pyrophosphate linkage predominates in the presence of Mn^{2+}. The highest yield of the 3',5'-phosphodiester is obtained in the presence of Ni^{2+}.[129]

BrCN reacts with imidazole in aqueous solution to produce *N*-cyanoimidazole, *N,N'*-iminodiimidazole, and *N*-carboxamidoimidazole (Figure 1.12). Purified N-cyanoimidazole and N,N'-iminodiimidazole are water soluble condensing agents with activity similar to that of BrCN in the presence of imidazole.



III **IV** **V**

Figure 1.12. Formation of *N*-cyanoimidazole (**III**), *N,N'*-iminodiimidazole (**IV**), and *N*-carboxamidoimidazole (**V**) in the reaction of BrCN with imidazole in aqueous solution.

Two mechanisms have been proposed for the condensation reactions. By analogy to the reaction of carbonyl diimidazole with nucleotides,[131] a phosphorimidazolide intermediate has been suggested (Figure 1.13).

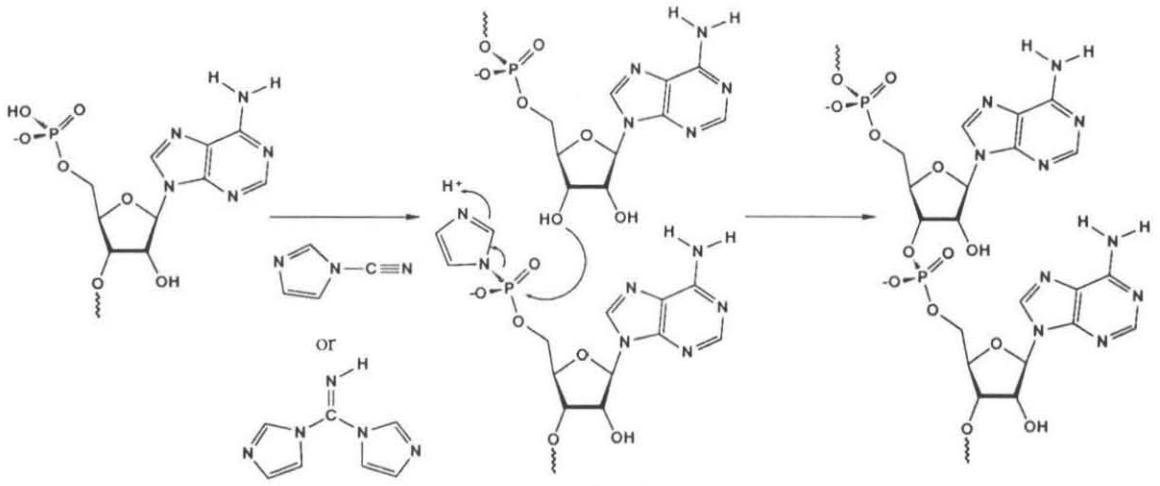


Figure 1.13. Mechanism proposed for the condensation promoted by *N*-cyanoimidazole and *N,N'*-iminodiimidazole, proceeding through a phosphorimidazolide intermediate.

Alternatively, direct attack of hydroxyl on an adduct of the phosphate and condensing agent might occur (Figure 1.14). Reaction of adenosine monophosphate with BrCN and imidazole at pH 6.0 does produce a detectable amount of phosphorimidazolide; however, neither mechanism has been ruled out.[129,132]

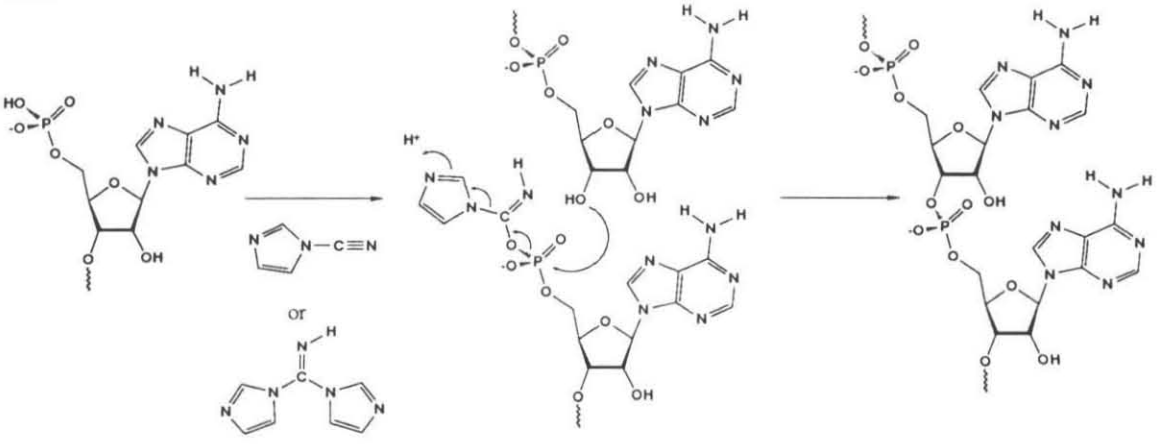


Figure 1.14. Mechanism proposed for the condensation promoted by *N*-cyanoimidazole and *N,N'*-iminodiimidazole, proceeding by direct attack of hydroxyl on an adduct of the phosphate and condensing agent.

Work Described in this Thesis. With the exception of special self-replicating systems,[133] the formation of phosphodiester bonds between strands of DNA creates new base sequences from existing base sequences. Therefore, DNA ligation within a triple helical complex can, in principle, translate chemical information between single- and double-stranded DNA employing a code that relates the sequence of an oligodeoxyribonucleotide to a sequence of double-stranded DNA. Although some of the template-directed condensations described above occurred in three-stranded complexes, information transfer in these cases did not rely on such a code.

Chapter 2 of this thesis describes the sequence specific ligation of two pyrimidine oligodeoxyribonucleotides directed by a double-stranded template through triple helix formation.[134] In this reaction, sequence information is transferred from the double-stranded template to the product. Conditions are found for efficient phosphodiester formation, and the sequence specificity of the reaction is investigated.

Chapter 3 describes the nonenzymatic sequence specific ligation of blunt-ended duplex DNA by triple helix formation.[135] Using a pyrimidine oligodeoxyribonucleotide as a template, two homopurine•homopyrimidine duplexes are ligated, creating a longer homopurine•homopyrimidine duplex. Also, a homopurine•homopyrimidine duplex is ligated to a homopyrimidine•homopurine duplex using a pyrimidine oligodeoxyribonucleotide of two segments joined 3' to 3' through an abasic linker. A duplex of the sequence type 5'-$(purine)_m(pyrimidine)_n$-3' is created. Sequence information is transferred from the single-stranded template to the double-stranded product in both of these reactions. In each case, product is obtained in which both strands are ligated. The sequence specificity of this

nonenzymatic approach to duplex ligation is not currently accessible in enzymatic ligations.

Chapter 2

Nonenzymatic Ligation of Oligodeoxyribonucleotides on a Duplex DNA Template by Triple Helix Formation

By juxtaposing the termini of oligodeoxyribonucleotides head-to-tail in a triple-helical complex, double-stranded DNA could act as a template to direct the sequence specific formation of a phosphodiester linkage (Figure 2.1). Sequence specificity in the formation of the complex would be derived from Hoogsteen hydrogen bonding, and the sequence information inherent in the double-stranded template would be manifested in the single-stranded product.

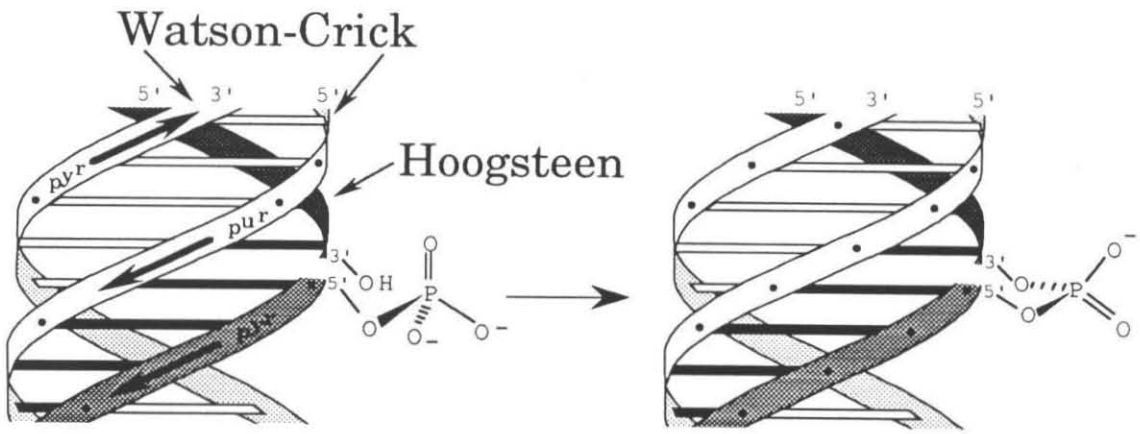


Figure 2.1. Ligation of a 3'-hydroxyl and an activated 5'-phosphate of two oligonucleotides is directed by a double-stranded template through formation of adjacent triple helical complexes.

To investigate this reaction, a double-stranded template, 37 base pairs in length, and two pyrimidine oligodeoxyribonucleotides, 15mers **A** and **B**,

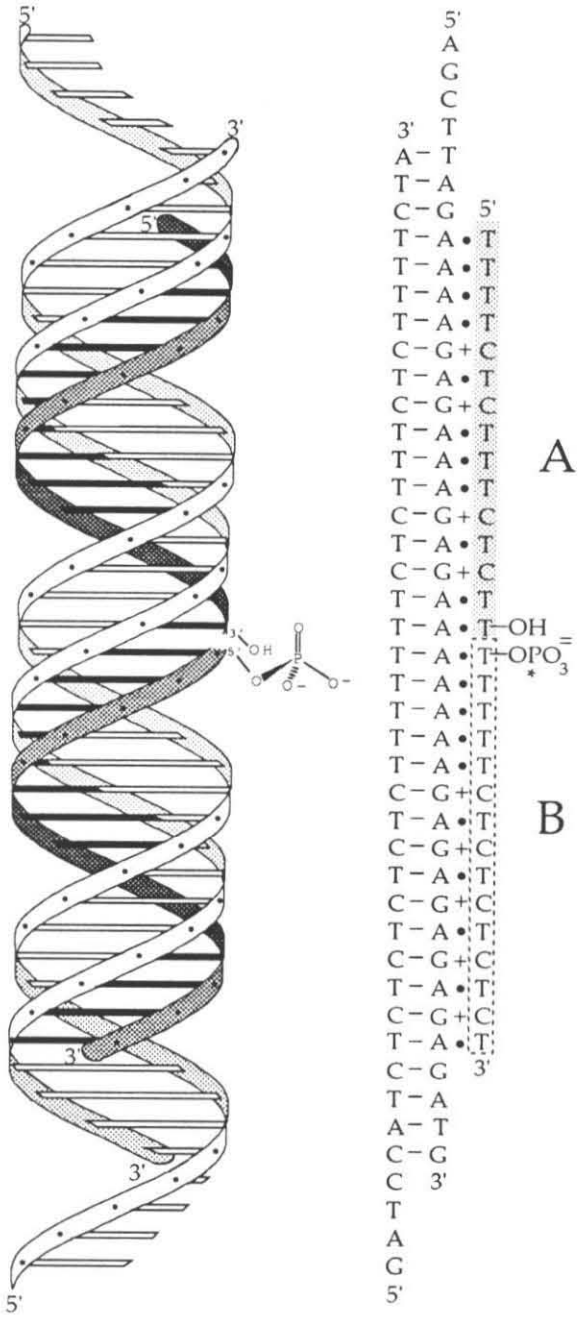
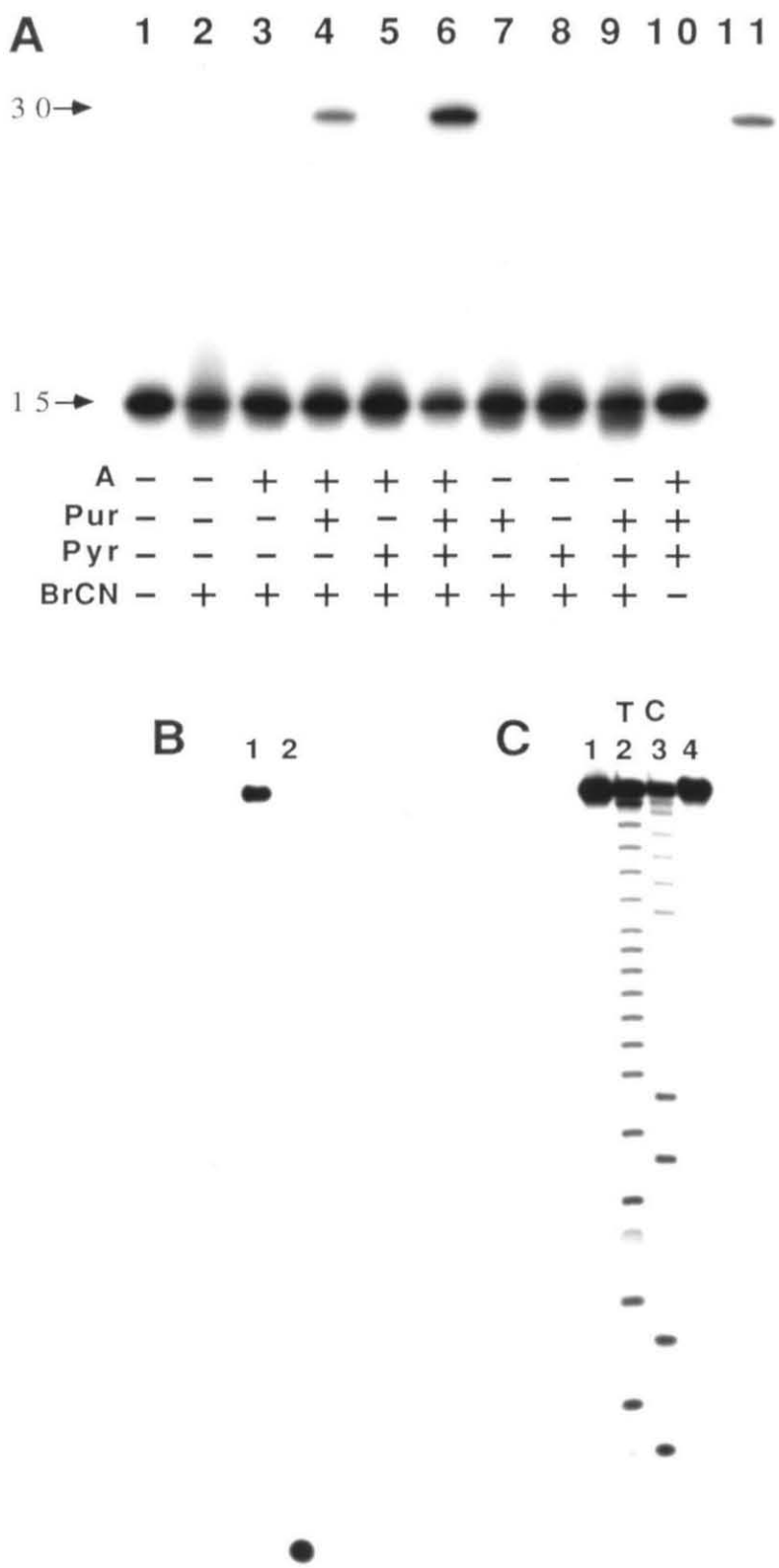


Figure 2.2. In the triple-stranded complex, the pyrimidine Hoogsteen strands (**A** and **B**) are the condensing oligodeoxyribonucleotides (shaded). The 3' terminus of oligodeoxyribonucleotide **A** is proximal to the 5' terminus of oligodeoxyribonucleotide **B**. Both are parallel to the purine strand of the template.

were synthesized by chemical methods and purified by gel electrophoresis. Oligodeoxyribonucleotides **A** and **B** are complementary in a Hoogsteen sense to adjacent unique 15 base pair sequences on the double-stranded template. In a triple helical complex, the 3' end of **A** would be proximal to the 5' end of **B**. Enzymatic phosphorylation of oligodeoxyribonucleotide **B** at its 5' terminus with [γ-^{32}P]ATP and polynucleotide kinase provides a terminal phosphomonoester for condensation with the 3' hydroxyl of **A** and a radioactive label with which to observe the condensation reaction. Condensation of **A** and **B** would yield a 30 nucleotide product containing the label. This product can be separated electrophoretically from the 15 nucleotide starting material. Four approaches to activating the terminal phosphate to nucleophilic attack were investigated: 1. activation *in situ* with cyanogen bromide and imidazole or one of their addition products, *N*-cyanoimidazole, 2. formation of a phosphorimidazolide, 3. activation *in situ* with the water soluble carbodiimide 1-ethyl-3,3-dimethylaminopropyl carbodiimide (EDC), and 4. enzymatic activation with T4 DNA ligase or T4 RNA ligase.

Activation in situ with Cyanogen Bromide and Imidazole or N-Cyanoimidazole. (^{32}P)**B** was incubated in the presence of **A** (0.7 μM) and the duplex DNA template (0.15 μM). The 5' terminal phosphate of (^{32}P)**B** was activated *in situ* with BrCN, imidazole, and $NiCl_2$. After 9 hours (20 °C), the reaction products were separated by denaturing polyacrylamide gel electrophoresis and visualized by autoradiography (Figure 2.3A). Condensation of 15mers **A** and (^{32}P)**B** in the presence of duplex template

Figure 2.3. Condensation of oligodeoxyribonucleotides **A** and (^{32}P)**B** and product analysis; autoradiograms of high resolution, denaturing 20% polyacrylamide gels: (A) Production of a radioactive 30mer by template-directed condensation of 15mer **A** and (^{32}P)**B**. Lanes 1-10: All reaction mixtures contain 5'-^{32}P-phosphorylated **B**. Activation of the 5'-phosphate of **B** was initiated by addition of a freshly prepared solution of BrCN (0.4 M) to the reaction mixtures. Reaction mixtures initially contained 40 mM BrCN, 20 mM imidazole•HCl (pH 7.0), 20 mM $NiCl_2$, and 100 mM NaCl in a total volume of 30 μL. In control experiments, BrCN, oligodeoxyribonucleotide **A**, and one or both strands of the template were omitted from the reaction. After a 9 h. reaction time (20 °C), the DNA was precipitated with ethanol, and the precipitates were dissolved in buffer and loaded on the gel. Lane 11: authentic synthetic 30mer with the sequence 5'-$T_4(CT_2)T_2(CT_2)T_6(CT_5)$-3' phosphorylated at its 5' end with polynucleotide kinase and [γ-^{32}P]ATP. (B) Enzymatic hydrolysis of the condensation product by treatment with calf spleen phosphodiesterase. Lane 1: condensation product. Lane 2: condensation product after treatment with 4 μg of calf spleen phosphodiesterase at 37 °C for 8 h. (C) Sequencing analysis of the 5'-phosphorylated condensation product. Lane 1: purified, ^{32}P 5' end-labeled condensation product. Lane 2: T-selective chemical cleavage reaction ($KMnO_4$).[137] Lane 3: C-selective cleavage reaction (hydrazine, NaCl).[136] Lane 4: 5' end-labeled condensation product treated with 4 μg of calf spleen phosphodiesterase at 37 °C for 8 h.


A
1 2 3 4 5 6 7 8 9 10 11
30→
15→
A – – + + + + – – – +
Pur – – – + – + + – + +
Pyr – – – – + + – + + +
BrCN – + + + + + + + + –
B
1 2
C
T C
1 2 3 4

was apparent on the autoradiogram by the production of a radioactive 30mer in 40% yield. The product of the reaction can be identified as a 30mer by comparison of its electrophoretic mobility with that of an authentic 5'-phosphorylated 30mer having the sequence expected for the head-to-tail **A-B** condensation product. Within experimental limits, no condensation product (<1%) is apparent in the absence of **A** or in the absence of template under the same conditions. No product (<1%) is detected after incubation of A, (^{32}P)**B**, and the pyrimidine strand of the template alone; however, some product (15% yield) is found after incubation of **A**, (^{32}P)**B**, and the purine strand of the template. Similar results are obtained when BrCN and imidazole are replaced by 1 mM *N*-cyanoimidazole.

The requirement for **A** in the coupling reaction to produce 30mer is consistent with condensation of (^{32}P)**B** exclusively with **A**. No self-condensation of (^{32}P)**B** is evident. Optimal formation of the 30mer occurs when both strands of the DNA duplex template are present. Remarkably, condensation of (^{32}P)**B** and **A** is observed in the presence of the purine strand of the template alone, even though the condensing oligodeoxyribonucleotides are not complementary to the purine template in an anti-parallel Watson-Crick fashion. Condensation could result from accomodation of a minimum of four base mismatches necessary for Watson-Crick hybridization or the formation of a double-stranded DNA complex in which the purine and pyrimidine strands are parallel to each other (*vide infra*).

If the double-stranded template is directing condensation of oligodeoxyribonucleotides **A** and (^{32}P)**B**, the 30mer produced will have a 5'-hydroxyl, a 3',5'-phosphodiester incorporating the radiolabeled phosphate,

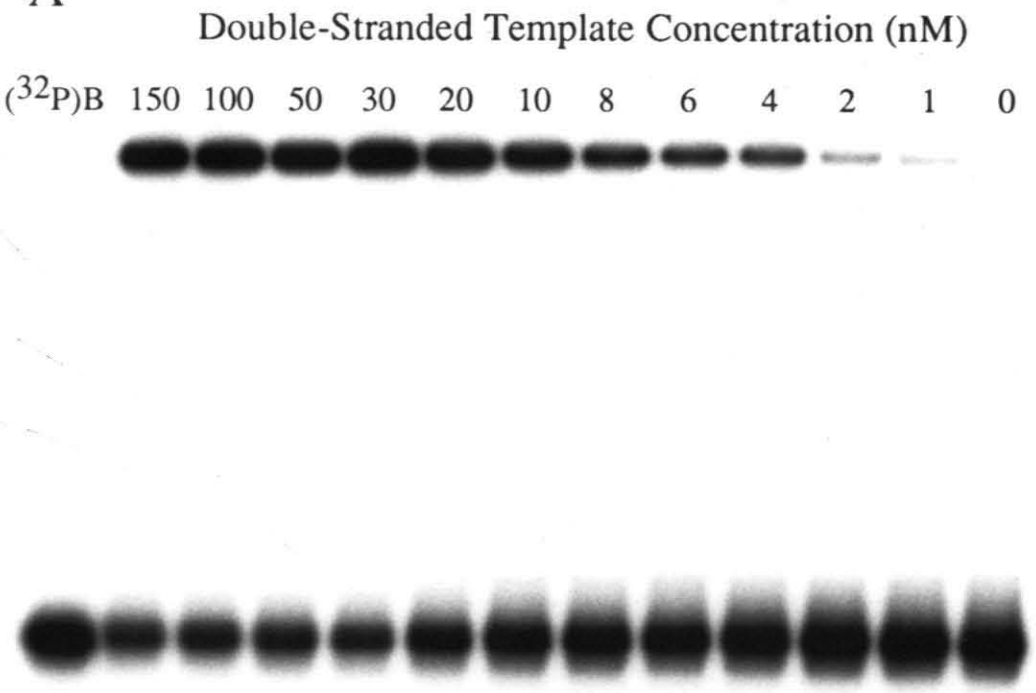
and a nucleotide sequence corresponding to a head-to-tail alignment of **A** and (^{32}P)**B**. To analyze the product composition, the 30mer produced in the reaction using the double-stranded template was purified by gel electrophoresis and digested with calf spleen phosphodiesterase, a 5'-exonuclease that degrades DNA terminated by 5'-hydroxyl (Figure 2.3B). After treatment of the condensation product with this phosphodiesterase, only a trace of the 30mer is detectable, and most of the radioactive label is found in mononucleotides. The susceptibility of the reaction product to degradation by the exonuclease indicates that it possesses a 5'-hydroxyl. The release of mononucleotides containing most of the radioactive label and the absence of a radioactive 16mer corresponding to enzymatic termination at the coupling site indicate that the predominant linkage formed in the condensation reaction is a 3', 5'-phosphodiester bond.

For sequencing analysis, the condensation reaction was carried out with oligodeoxyribonucleotide **B** phosphorylated by using unlabeled ATP. The **A-B** product was then purified by gel electrophoresis and 5' labeled with polynucleotide kinase and [γ-^{32}P]ATP. The 5' end-labeled product was purified by gel electrophoresis and subjected to chemical sequencing reactions specific for C^{136} and T^{137} (Figure 2.3C). The sequence of the 30mer verifies the condensation of the 3' terminus of **A** with the 5' terminus of **B**. Finally, the 5'-phosphorylated product was treated with calf spleen phosphodiesterase and shown to be resistant to degradation, confirming that the exonuclease had little endonuclease or 3'-exonuclease activity.

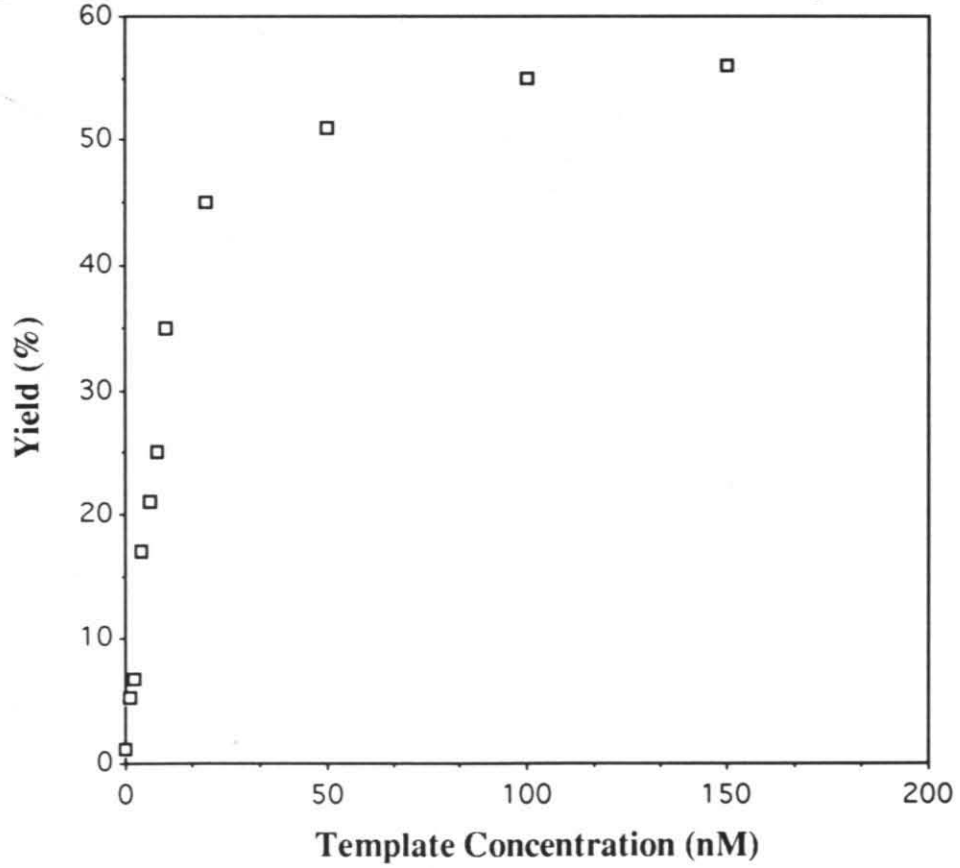
The dependence of the yield of the condensation of (^{32}P)**B** and **A** (1.3 μM) on the concentration of the double-stranded template was determined 10 hours after initiation with 40 mM BrCN (20 mM imidazole•HCl, pH 7.0,

Figure 2.4. Dependence of condensation reaction on the concentration of double-stranded template. (A) Autoradiogram of high-resolution, denaturing 20% polyacrylamide gel. All lanes contain (^{32}P)**B.** The far left lane contains untreated (^{32}P)**B**. In all other lanes, (^{32}P)**B** was treated for 10 hours at 20 °C in a reaction mixture containing oligodeoxyribonucleotide **A** (1.3 μM), 20 mM imidazole•HCl, pH 7.0, 20 mM $NiCl_2$, 40 mM BrCN and double-stranded template at 0, 1, 2, 4, 6, 8, 10, 20, 30, 50 , 100, and 150 nM, as indicated above each lane. Reactions were terminated by ethanol precipitation, and the precipitated DNA was dissolved in formamide buffer and loaded on the gel. (B) Plot of the yield of the condensation reaction *vs.* the concentration of double-stranded template. Yields were measured by excising radioactive bands and quantitating them by scintillation counting. The yield reported is the percentage of the total radioactivity in a lane within the band corresponding to the 30mer product.

A


Double-Stranded Template Concentration (nM)
(^{32}P)B 150 100 50 30 20 10 8 6 4 2 1 0

B


Yield (%)
0
10
20
30
40
50
60
0
50
100
150
200
Template Concentration (nM)

20 mM $NiCl_2$, 20 °C) (Figure 2.4). The yield (percent of total radioactivity incorporated into 30mer) is plotted versus the template concentration in Figure 2.4B. Little increase in yield is achieved by increasing the concentration of the template from 0.10 to 0.15 μM. This plateau is interpreted to represent the state in which association of (^{32}P)**B** with template has been driven nearly to completion.

In their studies of the use of BrCN and imidazole as a condensing agent for the formation of phosphodiesters, Kanaya and Yanagawa[129] and Ferris, Huang, and Hagan[132] found that the efficiency and regiochemistry of condensation were dependent on the presence of divalent metal ions. The metal ion dependence of condensation of **A** and (^{32}P)**B** activated by BrCN and imidazole and directed by the double-stranded template has been investigated. The yield of the reaction (12 hours, 21 °C) in the presence of various polyvalent cations was measured by electrophoretic analysis, and the results are listed on Table 2.1.

Of the cations tested, the most efficient condensation was in the presence of first row transition metals and Cd(II). Maximum yields were obtained in the presence of Ni(II), Cu(II), and Zn(II), and these yields (approximately 50%) were within experimental error of each other. A trace of 30mer (approximately 1% yield) was produced in the presence of alkaline earth metals (20 mM), spermine (1 mM), Pb(II) (10 mM), and $Co(NH_3)_6Cl3$ (20 mM). No condensation (<1%) was evident in the absence of metal or in the presence of 20 mM Hg(II) or Au(III). The specificity of the reaction for first row transition metals is suggestive of a role for these metal ions as Lewis acid catalysts, perhaps by coordination to imidazole.

Table 2.1. Effects of various polyvalent cations on the yield of condensation of (^{32}P)**B** and **A** (0.7 μM) directed by double-stranded template (0.15 μM) in the presence of 100 mM NaCl, 20 mM imidazole•HCl, pH 7.0, and 40 mM BrCN

Polyvalent cation	Condensation yield (%)
none added	<1
$Co(NH_3)_6Cl_3$ (20 mM)	~1
spermine•4HCl (1 mM)	~1
$MgCl_2$ (20 mM)	~1
$CaCl_2$ (20 mM)	~1
$Sr(NO_3)_2$ (20 mM)	~1
$BaCl_2$ (20 mM)	~1
$PbCl_2$ (10 mM)	~1
$MnCl_2$ (20 mM)	24
$CoCl_2$ (20 mM)	40
$NiCl_2$ (20 mM)	52
$CuCl_2$ (20 mM)	56
$ZnCl_2$ (20 mM)	47
$CdCl_2$ (20 mM)	40
$HgCl_2$ (20 mM)	<1
$AuCl_3$ (20 mM)	<1

The dependence of condensation yield on the concentration of $NiCl_2$ was determined using *N*-cyanoimidazole (1 mM) as condensing agent. In Figure 2.5, the yield of the reaction, determined by phosphorimager analysis, is plotted *vs.* the $NiCl_2$ concentration. Little improvement in yield is gained by increasing the $NiCl_2$ concentration above 10 mM under the conditions of the reaction, with optimum yield after 7 hours being approximately 90%.

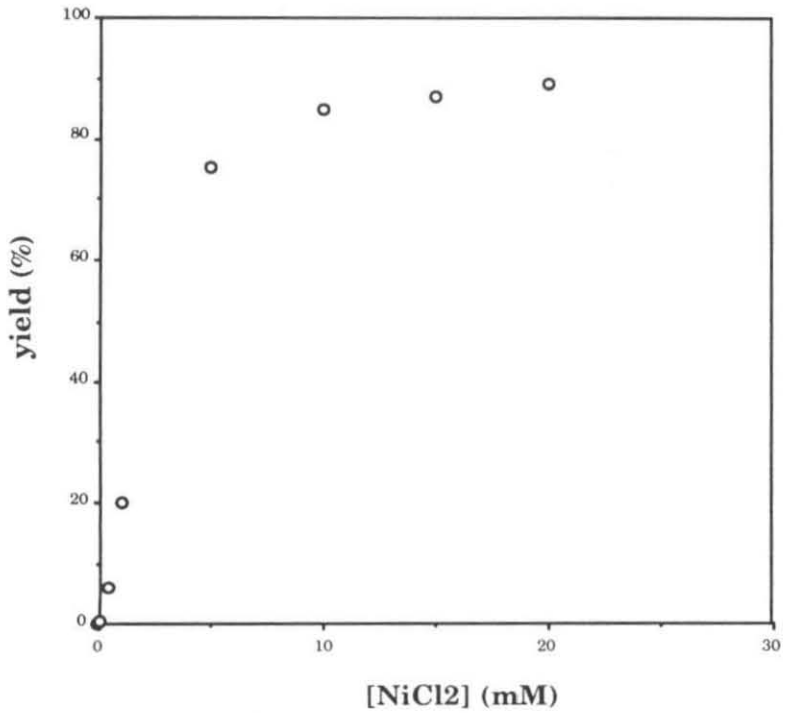


Figure 2.5. Dependence of condensation yield on $NiCl_2$ concentration. Reaction mixtures contained (^{32}P)**B**, **A** (2 μM), double-stranded template (0.15 μm), *N*-cyanoimidazole (1 mM), and $NiCl_2$ at various concentrations. The reactions were allowed to proceed for 7 hours at 20 °C before being stopped by ethanol precipitation and electrophoretic analysis.

Solutions of BrCN (40mM), imidazole (20 mM, pH 7.0), and $NiCl_2$ (20 mM) are acidic, having a pH of 6.1 upon mixing. The pH of these solutions

decreases with time due to hydrolysis of BrCN, falling to pH 5.6 in 6 hours. Similarly, solutions containing 20 mM $NiCl_2$ and 1 mM *N*-cyanoimidazole have a pH of 5.3 upon mixing and become slightly more acidic with time, the pH dropping to 5.2 in one hour. The effect of added buffers on the yield of the condensation reaction when BrCN and imidazole are used to activate the phosphate of (^{32}P)**B** was investigated. Highest yields were obtained in the absence of added buffers (52%), in 230 mM Tris•HCl, pH 7.0 (43%), and in 230 mM MES, pH 5.5 and 6.0 (66%). Low yields (<5%) were obtained in 230 mM HEPES, pH 6.8, 7.0, and 7.2. Low yield (<20%) was also obtained in the presence of 280 mM imidazole•HCl, pH 7.0. The effect of buffers on the reaction efficiency appears to result from a combination of factors. One factor might be the pH dependence of the stability of the triple helix, complexation being favored at low pH by protonation of cytosine.[27,32,33,44] However, the yield of the reaction in Tris•HCl, pH 7.0, is comparable to the yield in the acidic unbuffered solution. Complexation of the requisite divalent metal ions by buffer, reaction of buffer with the condensing agent, and general acid or general base catalysis of hydrolytic breakdown of condensing agent by buffer might contribute to the effect of buffers on the reaction.

To control the pH of the reaction and circumvent the interference of added buffers, the $NiCl_2$ was used as a buffer, employing the $Ni(H_2O)/Ni(OH)$ equilibrium to maintain pH. *N*-cyanoimidazole (1 mM) was used as the condensing agent in reaction mixtures containing (^{32}P)**B**, **A** (1.0, 2.0 or 12.7 μM), double stranded template (0.15 μM), and $NiCl_2$ (10 mM, pH 5.6, 5.8, 6.1, 6.7, 6.9, or 7.1). The result is plotted in Figure 2.6.

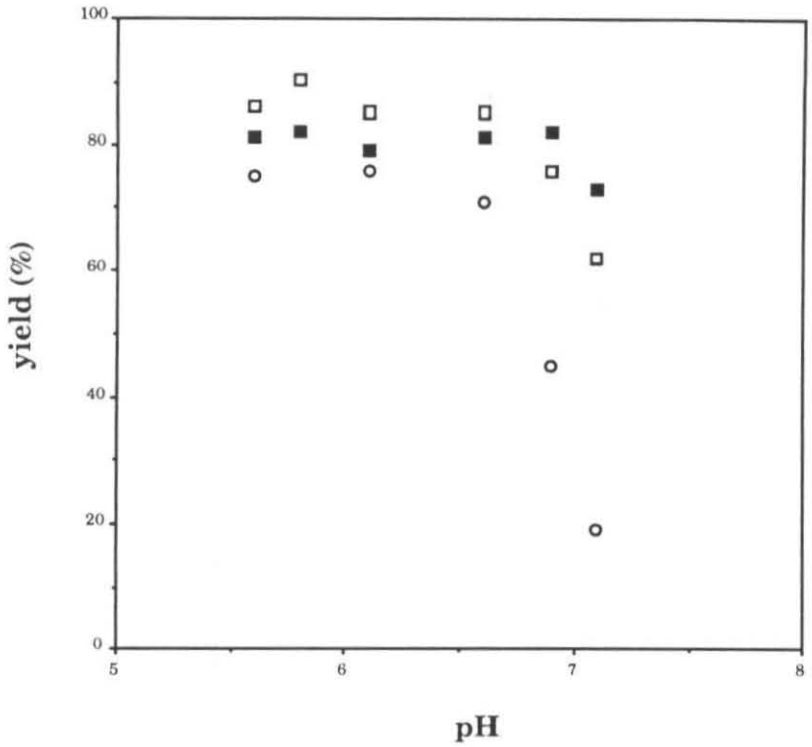


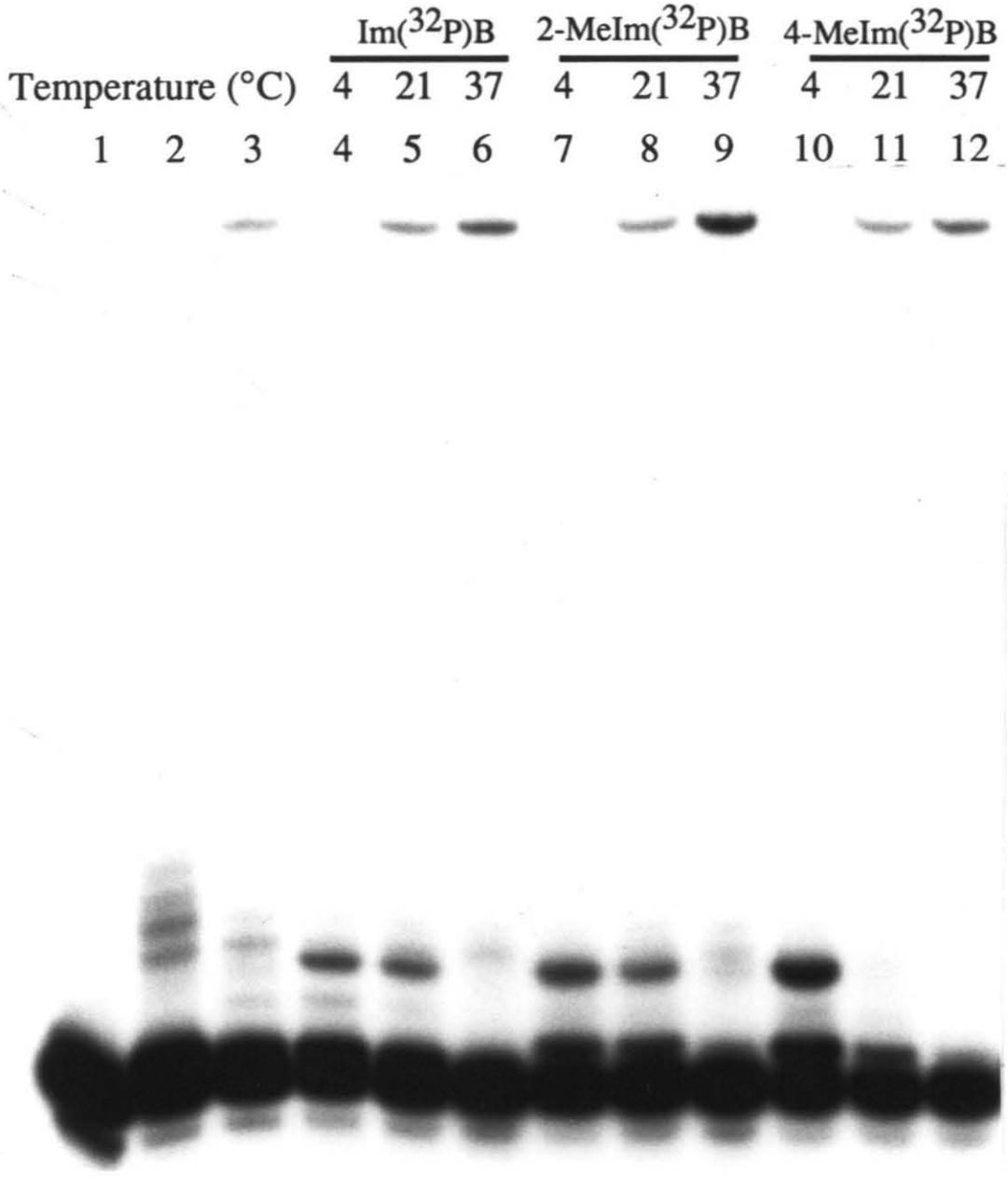
Figure 2.6. Effect of pH on yield of condensation of (^{32}P)**B**, **A** (12.7 μM, ■ ; 2 μM ,□; or 1 μM,○), in the presence of double-stranded template (0.15 μM), 1 mM *N*-cyanoimidazole, and 10 mM $NiCl_2$, pH 5.6, 5.8, 6.1, 6.6, 6.9, or 7.1, adjusted with concentrated NaOH. Reaction was for 7 h. at 20 °C.

At the concentrations of **A** that were studied, the yield of the reaction after 7 hours showed little dependence on pH below 6.7. Above pH 6.7, however, the condensation yield decreased with increasing pH when the concentration of **A** was 1.0 or 2.0 μM. There might also be a slight decrease in yield with 12.7 μM **A** at pH 7.1, but this apparent decrease could be due to uncertainty in the measurement of the reaction yield. The decrease in pH dependence with increasing concentration of **A** suggests that the decrease in yield with increasing pH is due primarily to destabilization of the triple helix. This destabilization can be compensated by increasing the concentration of **A**.

Activation by Formation of Phosphorimidazolide. A modification of the method of Chu and Orgel[138] was employed to form the phosphorimidazolide of oligodeoxyribonucleotide **B**. Because Inoue and Orgel had observed differences in the regiochemistry and efficiency of template-directed condensations using the 4-methyl-, 2-methyl-, and unmethylated phosphorimidazolides of mononucleotides,[117] the 4- and 2-methyl as well as unmethylated phosphorimidazolides of (^{32}P)**B** were investigated. Reaction of (^{32}P)**B** with EDC in a buffer of 200 mM imidazole•HCl, 2-methylimidazole•HCl, or 4-methylimidazole•HCl, pH 6.1, yielded, after ethanol precipitation, a product that was slightly less electrophoretically mobile in a 20% polyacrylamide gel than (^{32}P)**B**. Optimal yield of this product (60%) was obtained in 3 hours at room temperature. It was tentatively identified as the phosphorimidazolide of (^{32}P)**B** (Im(^{32}P)**B**, 2-MeIm(^{32}P)**B**, or 4-MeIm(^{32}P)**B**). The product mixture could be enriched in the putative phosphorimidazolide (80%) by electrophoretic purification; however, purification had no significant effect on the yield of condensation reactions.

These putative phosphorimidazolides were incubated in the presence of 1 μM duplex template and 1.3 micromolar **A** in a solution containing 20 mM imidazole•HCl, 2-methylimidazole•HCl, or 4-methylimidazole•HCl, respectively, pH 7.0, 100 mM NaCl, 20% tetrahydrofuran, and 20 mM $ZnCl_2$, 20 mM $Co(NH_3)_6Cl_3$, or 40 mM $CaCl_2$. After 20 hours at 4 °C, 21 °C, or 37 °C, the reaction products were separated by denaturing polyacrylamide gel electrophoresis and visualized by autoradiography. The results of reactions in the presence of $ZnCl_2$ are shown in Figure 2.7.

Figure 2.7. Reactions of putative phosphorimidazolides of (^{32}P)**B**; autoradiogram of high resolution, denaturing 20% polyacrylamide gel. Lane 1: (^{32}P)B, untreated. Lane 2: Products of treatment of (^{32}P)**B** with EDC and imidazole•HCl, pH 6.1, at 21 °C, 5 h. Lane 3: The products in lane 2 were treated with **purine**(Watson-Crick) (1 μM) and **A**(Watson-Crick) (1 μM), in 20 mM imidazole•HCl, pH 7.0, 100 mM NaCl, 20% tetrahydrofuran, and 40 mM $CaCl_2$. Lanes 4-12: Reactions of the putative phosphorimidazolides of (^{32}P)**B** and imidazole (Im(^{32}P)**B**), 2-methylimidazole (2-MeIm(^{32}P)**B**), or 4-methylimidazole (4-MeIm(^{32}P)**B**) at 4 °C, 21 °C, or 37 °C in the presence of **A** (1.3 μM), the purine strand of the template (1 μM), $ZnCl_2$ (20 mM), 100 mM NaCl, and 20% tetrahydrofuran.Reactions were for 20 hours.


Im(32P)B
2-MeIm(32P)B
4-MeIm(32P)B
Temperature (°C) 4 21 37 4 21 37 4 21 37
1 2 3 4 5 6 7 8 9 10 11 12

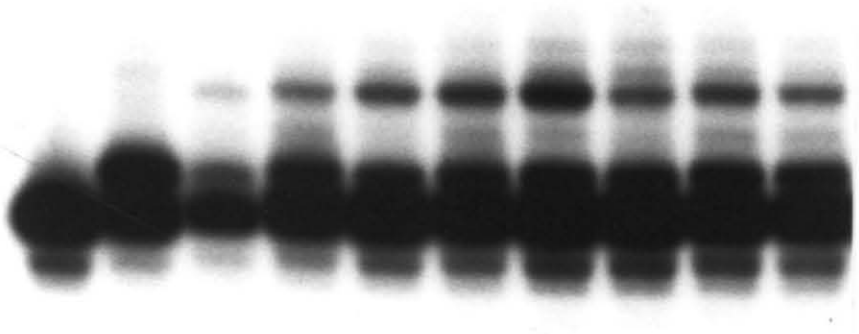
Several labeled products are observed in all cases. The least mobile of these products migrated as would be expected for a 30mer in this type of gel, approximately with the marker dye xylene cyanol. It is present in each of lanes 4-12. A product with the same mobility is also present in lane 3, where Im(^{32}P)**B** was incubated at 21°C, with 40 mM $CaCl_2$, in the presence of **A**(Watson-Crick), [5'-$(TC)_5T_5$-3'], and an oligodeoxyribonucleotide, **Purine**(Watson-Crick), [5'-AGCTTA$(AG)_6A_{10}(GA)_6$TG-3'], with which (^{32}P)**B** and **A**(Watson-Crick) can form a continuous Watson-Crick duplex, juxtaposing the 5' terminus of (^{32}P)**B** with the 3' terminus of **A**(Watson-Crick). Much smaller, but detectable, amounts (<0.5% of total radioactivity) of a species with the same mobility were also produced in the presence of $Co(NH_3)_6Cl_3$ and $CaCl_2$. At a given temperature, the yields are comparable for each of the phosphorimidazolides and increase with increasing temperature from 4 °C to 37 °C. The optimum yield of this product, obtained with 2-MeIm(^{32}P)**B** at 37 °C, is approximately 2-5%.

A series of 4-6 products with electrophoretic mobilities slower than (^{32}P)**B** is apparent in all lanes with phosphorimidazolides, even lane 2, where Im(^{32}P)**B** had not been treated following formation with EDC and imidazole. A plausible explanation of these products is that they are the result of reaction of EDC with nucleophilic sites on the bases of (^{32}P)**B** during formation of the phosphorimidazolides. Another product, similar in electrophoretic mobility to this series of products, is present when the phosphorimidazolides have been treated with zinc at 4 °C and 21 °C. It is not produced by treatment with $Co(NH_3)_6Cl_3$ or $CaCl_2$, and it is not present after treatment at 37 °C in the presence of zinc. Noteworthy is the absence of this product after treatment of 4-MeIm(^{32}P)**B** in the presence of zinc at

21°C. Bands corresponding in mobility to the putative phosphorimidazolides are clearly resolved from the (^{32}P)**B** band in several lanes, suggesting that the lifetime of the activated species under the conditions of the reactions is comparable to the reaction time, 20 hours.

Experiments were performed to determine the dependence of formation of the low mobility product on the double-stranded template and oligodeoxyribonucleotide **A**. The reaction was run as above with 2-MeIm(^{32}P)**B** at 21 °C for 22 hours. In different reactions, oligodeoxyribonucleotide **A**, the purine strand of the template, the pyrimidine strand of the template, two of the above, or all three of the above were omitted. The result of analysis of the reactions by polyacrylamide gel electrophoresis and autoradiography is shown in Figure 2.8. Whereas the low mobility product is formed in a yield of approximately 5% when all of the components of the reaction mixture are present, none is detectable in the absence of **A** or the purine strand of the template. The lower limit of detection is estimated to be <0.5% of the total radioactivity in a lane. As with the condensation with BrCN and imidazole, some product(<1%) is observed in the presence of **A** and the purine strand of the template alone. The requirement of 15mer **A** in the reaction encourages the identification of this product as an adduct of **A** and (^{32}P)**B**. Furthermore, the dependence of the reaction for both strands of the duplex template supports a model in which formation of this adduct is directed by juxtaposition of **A** and (^{32}P)**B** in a triple helical complex such as that depicted in Figure 2.2.

Figure 2.8. Template-directed condensation of 2-MeIm(^{32}P)**B** with **A**; autoradiogram of high resolution, denaturing 20% polyacrylamide gel. Lane 7: Reaction of 2-MeIm(^{32}P)**B** in the presence of **A** (1.3 μM) and double-stranded template (1 μM). Lane 1: (^{32}P)**B**, untreated. Lane 2: 2-MeIm(^{32}P)**B**, untreated. In control reactions, **A**, the purine strand of the template, and/or the pyrimidine strand of the template were omitted. All reaction mixtures contained 20 mM 2-methylimidazole•HCl, pH 7.0, 100 mM NaCl, 20% tetrahydrofuran, and 20 mM $ZnCl_2$.



1	2	3	4	5	6	7	8	9	10	
		−	+	+	+	+	−	−	−	A
		−	−	+	−	+	+	−	+	Purine
		−	−	−	+	+	−	+	+	Pyrimidine

In addition to $ZnCl_2$, $Co(NH_3)_6Cl_3$, and $CaCl_2$, the template-directed reaction of 2-MeIm(^{32}P)**B** was attempted in the presence of 20 mM $MgCl_2$, $MnCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $CdCl_2$, $Hg(OAc)_2$, or K_3IrCl_6, 10 mM $PbCl_2$, or 1 mM spermine. Under conditions that produced the 30mer in 4% yield in the presence of 20 mM $ZnCl_2$ (28 hours, 37 °C), the product appeared in approximately 1% yield in the presence of 20 mM $MnCl_2$. None of the other metal salts investigated or spermine detectably promoted the reaction under the conditions tested; however, a new product with electrophoretic mobility intermediate between those of (^{32}P)**B** and the low mobility product resulted from the reaction in the presence of 1 mM spermine. The yield of this product was 38% and is likely to be a (^{32}P)**B**-spermine adduct.

Several roles can be imagined for the zinc ion in this reaction: electrostatic stabilization of the locally high negative charge density required for triple helix formation, modulation of the structure of the triple-helical complex by coordination to bases and phosphates to provide optimal geometric disposition of the reactive groups, as a template to approximate the reactive functionalities (presumably the 5'-phosphorimidazolide of **B** and the 3'-hydroxyl of **A**) by coordinating them directly, and as a Lewis acid catalyst. Though electrostatic stabilization of the triple helix is a likely role for zinc ions in these experiments, that role is not sufficient to account for the specificity of the reaction for zinc. $Co(NH_3)_6Cl_3$ and the other polyvalent cations tested could also serve this role. Similarly, a role for the zinc as a Lewis acid catalyst of the reaction, possibly labilizing the phosphorimidazolide by coordination to the lone pair of the imidazole nitrogen, does not account for the inadequacy of metal ions such as Cu(II) and Ni(II) which also have a high affinity for imidazole.

The specific requirement for Zn(II) in the reaction suggests that it is involved in orienting the reactive functionalities with respect to each other within the triple helix, either by altering the overall structure of the triple helix or providing a template for approximation of the reacting groups (or both). The ability of a metal ion to behave as a template for the reaction would be expected to be sensitive to the ligand geometries accessible to the metal center. The nature of the deformation of a triple helical complex upon binding of a metal would depend upon the distribution of the metal among the possible base and backbone binding sites. Each of these effects could contribute to the specificity of the reaction on Zn(II). Orgel and co-workers have reported effects of metal ions on the efficiency and regiochemistry of condensations of mononucleotide phosphorimidazolides in the presence[108,111,113,114,120] and absence[112,121] of single stranded templates. Their results are also consistent with metal ions altering the structure of the helical template-substrate complex and behaving as templates themselves to orient reactive groups.

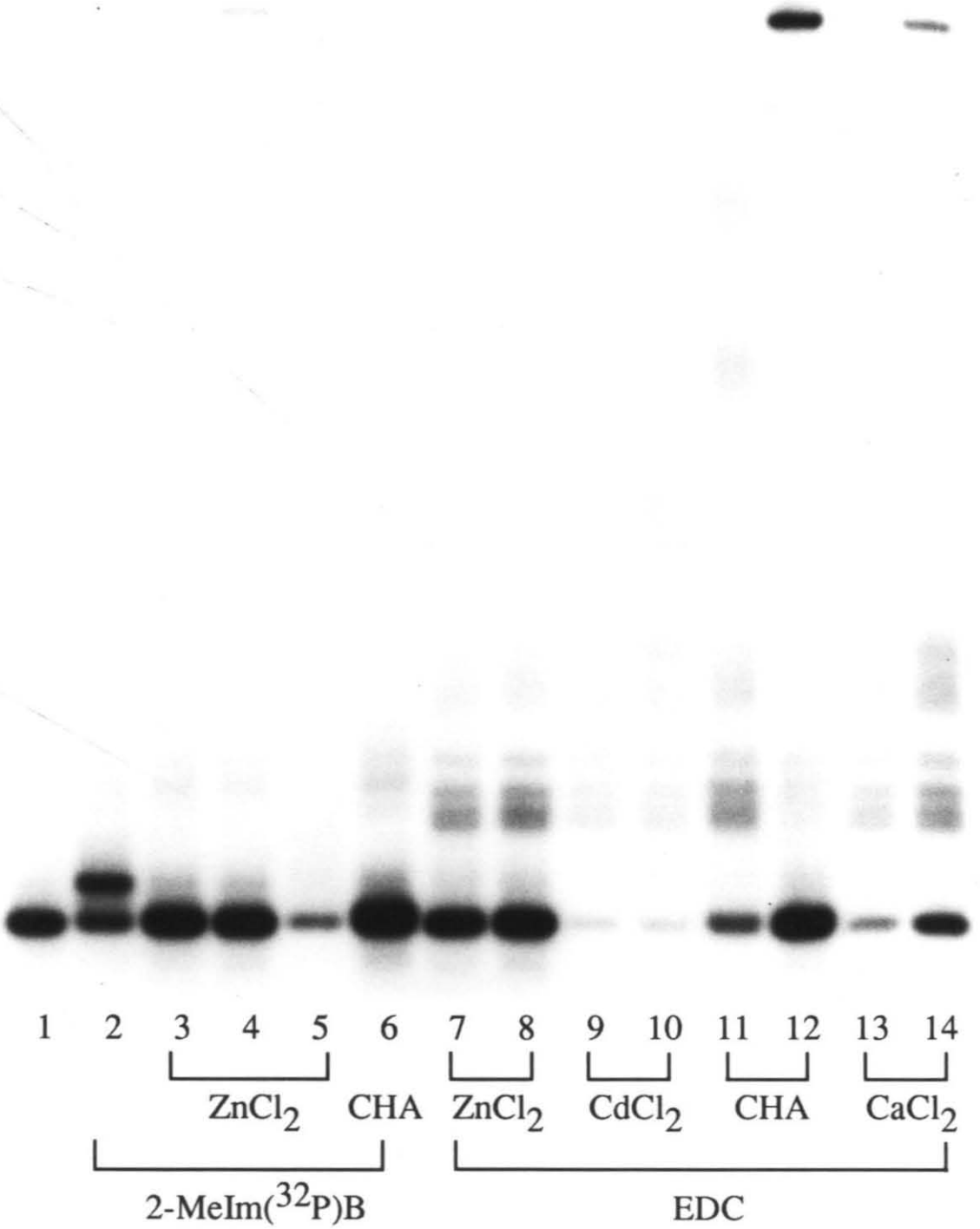
The difference in the efficiencies of condensation of (^{32}P)**B**-5'-phosphorimidazolide and (^{32}P)**B** activated by reaction with *N*-cyanoimidazole is relevant to the mechanism of phosphate activation by the latter. In alternative mechanistic hypotheses, a) condensation occurs by direct attack of the nucleophile on an adduct of *N*-cyanoimidazole with the phosphate or b) an intermediate phosphorimidazolide is formed, which is attacked by the nucleophile.[129,132] The relative inefficiency of condensation of the phosphorimidazolides, despite their apparent stability to the reaction conditions, suggests that the phosphorimidazolide is not the most reactive

intermediate in the much more efficient *N*-cyanoimidazole-promoted condensation.

Activation in situ with the Water Soluble Carbodiimide 1-ethyl-3, 3-dimethylaminopropyl Carbodiimide (EDC). EDC at a final concentration of 60 mM was added to solutions containing (^{32}P)**B**, 1.3 μM **A**, 1.0 μM double stranded template, 100 mM HEPES, pH 6.8, 100 mM NaCl, 30% THF, and 20 mM $ZnCl_2$, 20 mM $CdCl_2$, 20 mM $Co(NH_3)_6Cl_3$, or 40 mM $CaCl_2$. Also, for comparison, 2-MeIm(^{32}P)**B** was used in reaction mixtures containing 1.3 μM **A**, 1.0 μM double stranded template, 100 mM NaCl, 30% THF, 100 mM HEPES, pH 6.8, or 20 mM 2-methylimidazole•HCl, pH 7.0, and 20 mM $ZnCl_2$ or $Co(NH_3)_6Cl_3$. In control experiments, the template was omitted. An autoradiogram of the electrophoretic analysis after 20 hours reaction at 37 °C is shown in Figure 2.9.

As seen in previous experiments, reaction of 2-MeIm(^{32}P)**B** in the presence of duplex template, oligodeoxyribonucleotide **A**, 2-methylimidazole buffer, and $ZnCl_2$ resulted in the formation of a labeled product with electrophoretic mobility as expected for a 30mer. In addition, a product with the same mobility resulted from the reaction with EDC containing template and $Co(NH_3)_6Cl_3$ (10% yield) or $CaCl_2$ (3% yield). This product was not yielded in the reaction with EDC in the presence of $ZnCl_2$ or in the reaction of 2-MeIm(^{32}P)**B** in HEPES buffer in the presence of $Co(NH_3)_6Cl_3$. This difference in polyvalent cation requirement confirms that the activated species in the reactions of 2-MeIm(^{32}P)**B** is not simply a stable adduct of EDC with (^{32}P)**B** but a material that is chemically distinct from the active species in the reaction with EDC.

Figure 2.9. Activation of (^{32}P)**B** with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC); autoradiogram of high resolution, denaturing 20% polyacrylamide gel. Lanes 7-14: (^{32}P)**B** after reaction with EDC (60 mM) for 20 hours at 37 °C in the presence of 1.3 μM **A**, 1.0 μM double-stranded template (lanes 8, 10, 12, and 14) or no template (lanes 7, 9, 11, and 13), 100 mM HEPES, pH 6.8, 100 mM NaCl, 30% THF, and 20 mM $ZnCl_2$, 20 mM $CdCl_2$, 20 mM $Co(NH_3)_6Cl_3$ (CHA), or 40 mM $CaCl_2$. For comparison, Lane 1: Untreated (^{32}P)**B**, Lane 2: Untreated 2-MeIm(^{32}P)**B.** Lane 3: 2-MeIm(^{32}P)**B** treated with 1.3 μM **A**, 100 mM NaCl, 30% THF, 20 mM 2-methylimidazole•HCl, pH 7.0, and 20 mM $ZnCl_2$ in the absence of duplex template. Lane 4: 2-MeIm(^{32}P)**B** treated with 1.3 μM **A**, 100 mM NaCl, 30% THF, 20 mM 2-methylimidazole•HCl, pH 7.0, and 20 mM $ZnCl_2$ in the presence of 1 μM duplex template. Lane 5: 2-MeIm(^{32}P)**B** treated with 1.3 μM **A**, 100 mM NaCl, 30% THF, 100 mM HEPES, pH 6.8, and 20 mM $ZnCl_2$ in the presence of 1 μM duplex template. Lane 6: 2-MeIm(^{32}P)**B** treated with 1.3 μM **A**, 100 mM NaCl, 30% THF, 100 mM HEPES, pH 6.8, and 20 mM $Co(NH_3)_6Cl_3$ (CHA) in the presence of 1 μM duplex template.


1 2 3 4 5 6 7 8 9 10 11 12 13 14
ZnCl2 CHA ZnCl2 CdCl2 CHA CaCl2
2-MeIm(32P)B
EDC

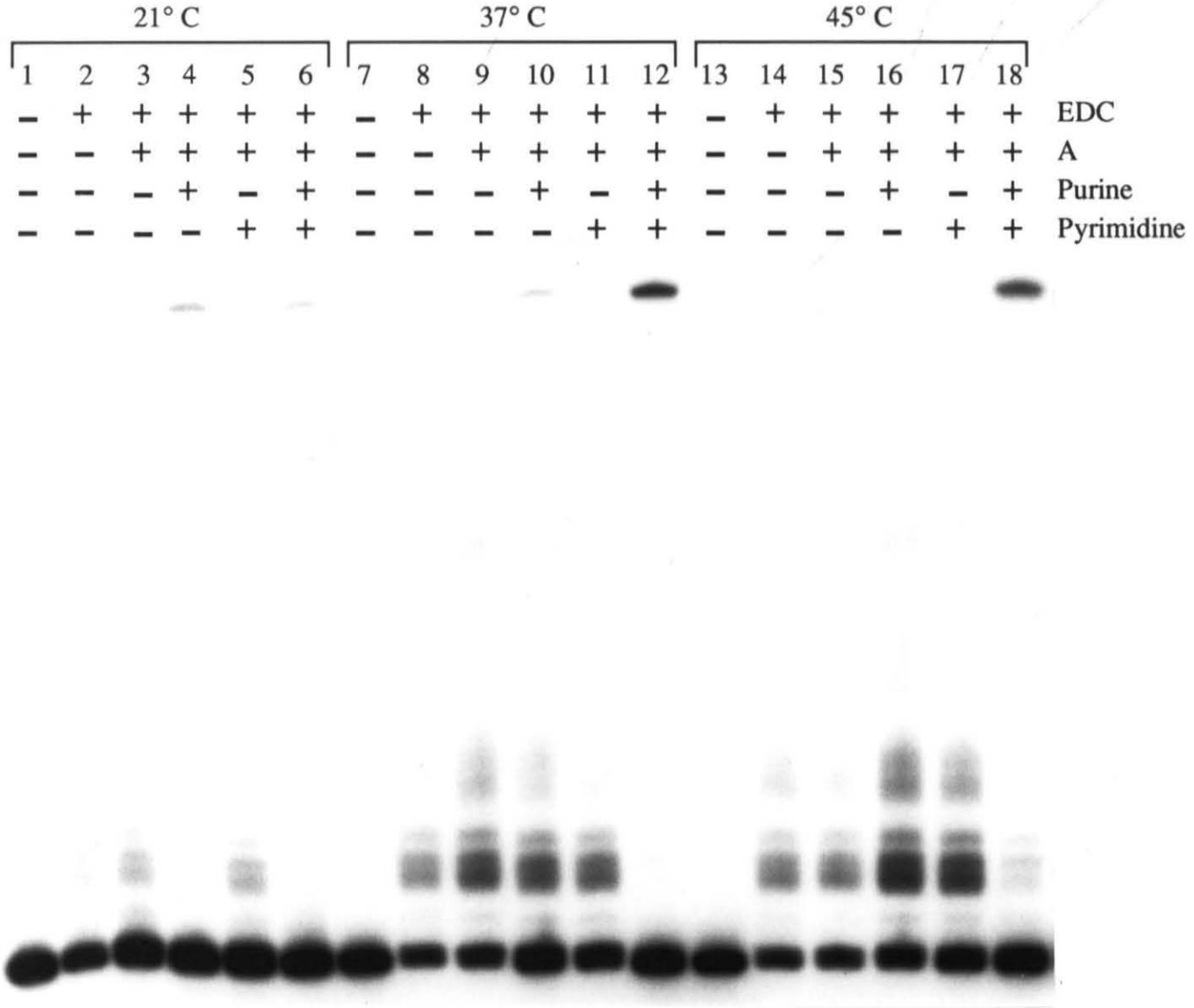
After treatment of (^{32}P)**B** with EDC, much of the total radioactivity was incorporated into products with a broad range of electrophoretic mobilities. Reaction of EDC with nucleophilic sites on the bases of (^{32}P)**B** is likely to account for these products. Their occurrence is dependent on both the polyvalent cation present and the presence or absence of the template. Within this experiment, their yield is lowest in the presence of template and $Co(NH_3)_6Cl_3$.

A set of control experiments was performed at 21 °C, 37 °C, and 45 °C, in the presence of 100 mM HEPES, pH 6.8, 100 mM NaCl, 30% THF, and 20 mM $Co(NH_3)_6Cl_3$, omitting template strands and oligodeoxyribonucleotide **A** (Figure 2.10). In the presence of the double-stranded template and **A**, the 30mer was produced in 10% yield at 37 °C and 45 °C and in 1.5% yield at 21 °C. At 21 °C and 37 °C, a product with the same mobility was produced in 2% yield in the presence of the purine strand of the template alone.

Of the reaction temperatures investigated in this experiment, 37 °C is optimal with regard to yield of the 30mer product relative to the yield of side products. The yield of side products generally increased with increasing temperature. At 21 °C, the yield of these products was clearly diminished in the presence of the purine strand of the template as well as the double-stranded template. At the higher temperatures, the yield of these products remained diminished in the presence of the double-stranded template but was comparable in the presence of the purine strand of the template, the pyrimidine strand of the template, and no template.

Although oligodeoxyribonucleotides **A** and **B** are not complementary to the purine strand of the template in an anti-parallel Watson-Crick

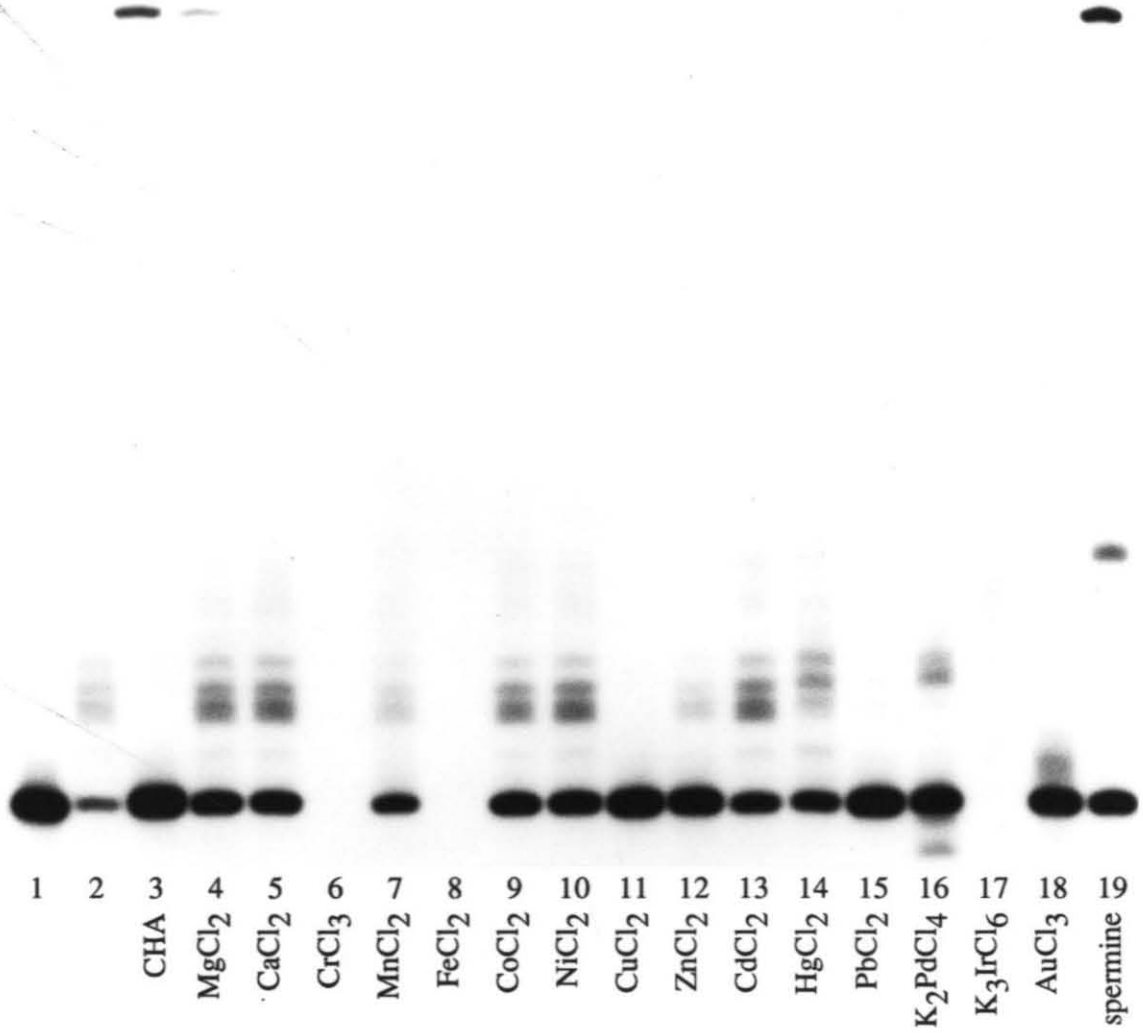
Figure 2.10. Template-directed condensation of (^{32}P)**B** with **A** using EDC as an activating agent: template requirements and temperature dependence; autoradiogram of high resolution, denaturing 20% polyacrylamide gel. All lanes contain (^{32}P)**B** after treatment for 20 hours at the indicated temperature in solutions containing 100 mM HEPES, pH 6.8, 100 mM NaCl, 30% THF, and 20 mM $Co(NH_3)_6Cl_3$. Lane 6, 12, and 18: Treatment with 60 mM EDC in the presence of **A** (1.3 μM), and double-stranded template (1 μM). Lanes 5, 11, and 17: The purine strand of the template was omitted. Lanes 4, 10, and 16: The pyrimidine strand of the template was omitted. Lanes 3, 9, and 15: Both strands of the template were omitted. Lanes 2, 8, and 14: **A** and template were omitted from reaction with 60 mM EDC. Lanes 1, 7, and 13: EDC, **A**, and template were omitted.


21° C
37° C
45° C
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
– + + + + + – + + + + + – + + + + + EDC
– – + + + + – – + + + + – – + + + + A
– – – + – + – – – + – + – – – + – + Purine
– – – – + + – – – – + + – – – – + + Pyrimidine

fashion, protection of (^{32}P)**B** from reaction with EDC at 21 °C in the presence of the purine strand of the template is indicated by the absence of side products in that reaction, suggesting the formation of a base paired complex. The absence of detectable amounts of the 30mer after reaction at 45 °C in the presence of the purine strand alone indicates the instability of the complex at that temperature under the conditions of the reaction. EDC is evidently as effective a condensing agent at 45 °C as at 37 °C, the yields of 30mer in the presence of double stranded template being equal at those temperatures.

In addition to $ZnCl_2$, $Co(NH_3)_6Cl_3$, $CdCl_2$, and $CaCl_2$, the template-directed reaction of (^{32}P)**B** with **A**, using EDC as the condensing agent, was attempted at 37 °C in the presence of 20 mM $MgCl_2$, $CrCl_3$, $FeCl_2$, $MnCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $HgCl_2$, K_2PdCl_4, $AuCl_3$, or K_3IrCl_6, 10 mM $PbCl_2$, or 1 mM spermine (Figure 2.11). The 30mer was produced in yields greater than 1% in the presence of $MgCl_2$ (2%) and spermine (20%). Traces of the product (<1%) were obtained in the absence of polyvalent cation and in the presence of $MnCl_2$. Cobalt hexammine and spermine were the only polyvalent cations investigated which both stimulated formation of the 30mer and resulted in diminished yield of side products. Reaction in the presence of 1 mM spermine, however, resulted in a distinct side product in 10% yield. As in the reaction of 2-MeIm(^{32}P)**B** in the presence of spermine, this species is likely to be the result of condensation of the activated phosphate of **B** with an amine of spermine, producing a spermine(^{32}P)**B** phosphoramidate. The absence of radioactive label after treatment with K_3IrCl_6, $CrCl_3$, and $FeCl_2$ is presumed to be due to

Figure 2.11. Polyvalent cation dependence of the template-directed reaction of (^{32}P)**B** with **A** using EDC (60 mM) as a condensing agent; autoradiogram of high resolution denaturing 20% polyacrylamide gel. All lanes contain (^{32}P)**B**. Lane 1: Untreated (^{32}P)**B**. Lanes 2-19: (^{32}P)**B** treated for 20 hours at 37 °C with 60 mM EDC in the presence of **A** (1.3 μM), double-stranded template (1 μM), 100 mM HEPES, pH 6.8, 100 mM NaCl, and 30% THF. In lane 2, no polyvalent cation was added to the reaction mixture. In lanes 3-19, the the polyvalent cation salt indicated was added. $PbCl_2$ (lane 15) was 10 mM. Spermine tetrahydrochloride (lane 19) was 1 mM. All other polyvalent cations were present at 20 mM. CHA (lane 3) is $Co(NH_3)_6Cl_3$.

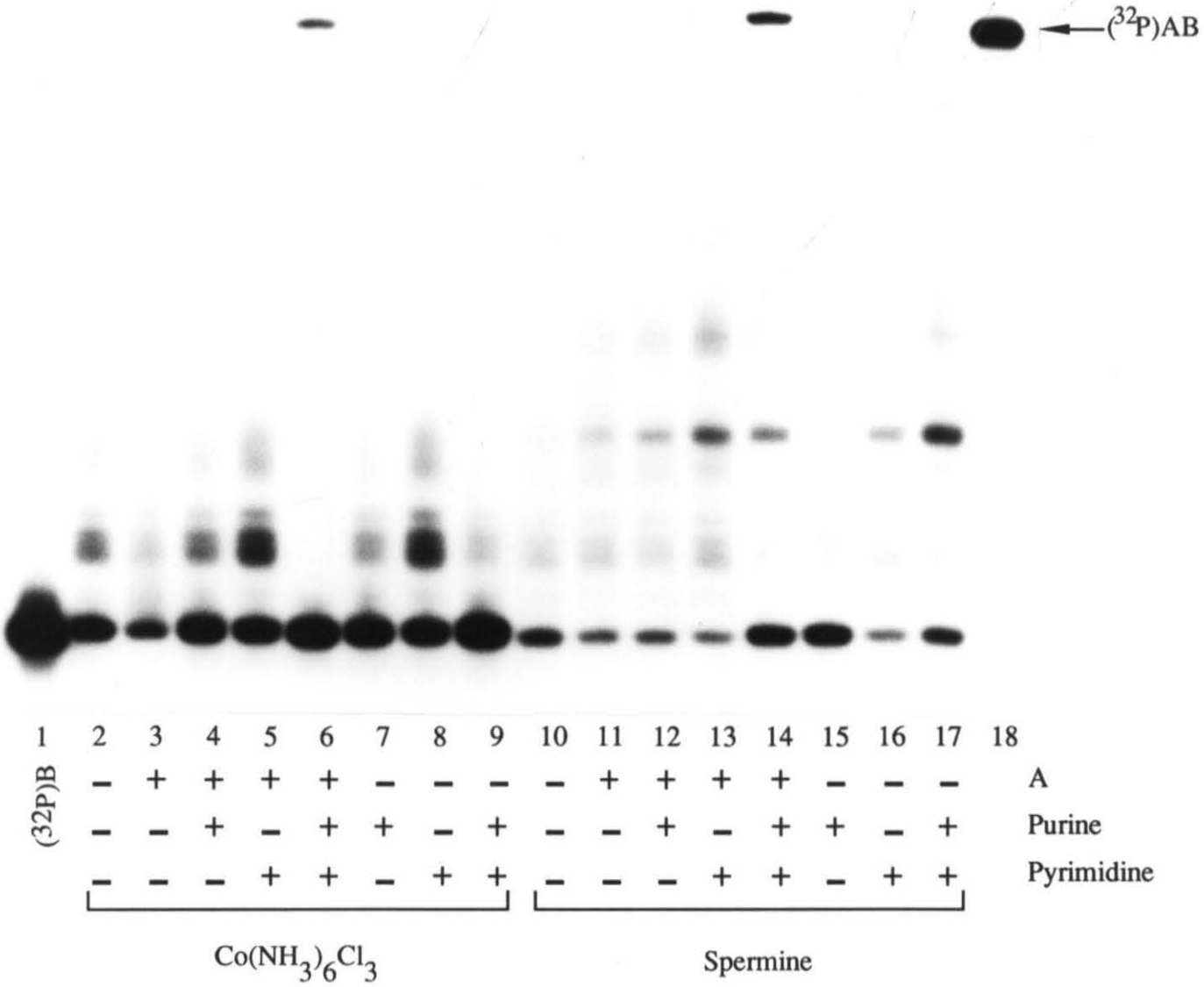

1
2
3 CHA
4 $MgCl_2$
5 $CaCl_2$
6 $CrCl_3$
7 $MnCl_2$
8 $FeCl_2$
9 $CoCl_2$
10 $NiCl_2$
11 $CuCl_2$
12 $ZnCl_2$
13 $CdCl_2$
14 $HgCl_2$
15 $PbCl_2$
16 K_2PdCl_4
17 K_3IrCl_6
18 $AuCl_3$
19 spermine

degradation or precipitation of the DNA by those metals.

Comparison of the reactions of (^{32}P)**B** at 37 °C, in the presence of 1 mM spermine or 20 mM $Co(NH_3)_6Cl_3$ and 60 mm EDC is shown in the autoradiogram of Figure 2.12. The yield of the 30mer in the presence of **A** and the double-stranded template was 20% in the presence of spermine and 10% in the presence of $Co(NH_3)_6Cl_3$. The electrophoretic mobility of the product is compared with that of an authentic 5' phosphorylated 30mer having the sequence expected for the head-to-tail **A-B** condensation product. The authentic 30mer, synthesized chemically and phosphorylated with [γ–^{32}P]ATP and polynucleotide kinase, is slightly more mobile than the product of the condensation reaction, a difference attributable to the effect of the negative charge added to the authentic sample upon 5' phosphorylation.

Enzymatic Ligation. In studies of the action of T4 DNA ligase on triple-stranded nucleic acids, Raae and Kleppe found that ligation of decathymidylate, dT_{10}, was catalyzed in the presence of equimolar mixtures of dA_n and dT_n and dA_n and U_n.[139] The authors cited this result as support for the conclusion that polynucleotide ligase is able to catalyze ligation of dT_{10} present in triple-stranded DNA and further concluded that the enzyme is able to act on the Hoogsteen base paired strand of a triple helix. T4 DNA ligase and T4 RNA ligase were tested for their ability to catalyze the ligation of **A** and (^{32}P)**B** in the presence of ATP (1 mM). No ligation was observed when (^{32}P)**B** was treated with T4 RNA ligase (10 units) in the presence or absence of **A** or template. The results of

Figure 2.12. Template-directed condensation of (^{32}P)**B** with **A** using EDC (60 mM) as a condensing agent: comparison of reactions in the presence of $Co(NH_3)_6Cl_3$ and spermine; autoradiogram of high resolution, denaturing 20% polyacrylamide gel. Lane 1: Untreated (^{32}P)**B**. Lane 18: Authentic synthetic 30mer with the sequence 5'-$T_4(CT_2)T_2(CT_2)T_6(CT_5)$-3' phosphorylated at its 5' end with polynucleotide kinase and [γ-^{32}P]ATP. Lanes 2-17: (^{32}P)**B** treated for 20 hours at 37 °C with 60 mM EDC in the presence of 100 mM HEPES, pH 6.8, 100 mM NaCl, and 30% THF. In lanes 2-9, $Co(NH_3)_6Cl_3$ was present in the reaction mixtures at 20 mM. In lanes 10-17, spermine tetrahydrochloride was present at 1 mM. **A** was included at 1.3 μM, and template was included at 1.0 μM.

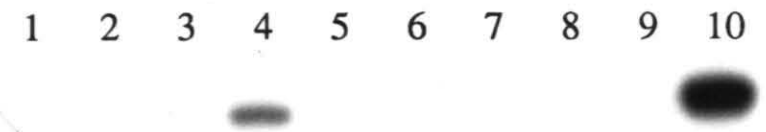

(^{32}P)AB
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
(^{32}P)B
A
Purine
Pyrimidine
$Co(NH_3)_6Cl_3$
Spermine

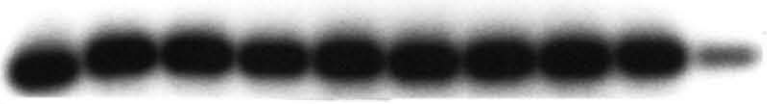
electrophoretic analysis of reactions with T4 DNA ligase are shown in Figure 2.13.

Treatment of (^{32}P)**B** with T4 DNA ligase (20 units) in the presence of double stranded template (0.15 μM) and **A** (2.0 μM) resulted in the formation of product in trace amounts (<1.0%) (lane 6). As a size standard and a control for the activity of the enzyme under the experimental conditions, (^{32}P)**B** was treated with T4 DNA ligase in the presence of **Purine**(Watson-Crick) (0.18 μM) and **A**(Watson-Crick) (2.0 μM) (lane 10). A ligation product was afforded in 85% yield. This product was slightly less mobile in the gel than that yielded in the presence of **A** and the double stranded template, consistent with both being 30 nucleotide ligation products of different sequences. In the DNA triple helix studied here, T4 DNA ligase was unable to catalyze ligation of Hoogsteen base paired oligodeoxyribonucleotides with an efficiency comparable to ligation of Watson-Crick base paired oligodeoxyribonucleotides. With regard to the conclusions of Raae and Kleppe, it is possible that the ability of T4 DNA ligase to catalyze ligation of Hoogsteen base paired oligodeoxyribonucleotides is dependent upon sequence or that the dT_{10} ligated in the experiments of those authors was not the Hoogsteen base paired strand of a triple helix. The dT_{10} might not have been the Hoogsteen strand of a triple helix if, for example, the equimolar mixture of dA_n and dT_n was present as a mixture of $dT_n \bullet dA_n \bullet dT_n$ and dA_n.

No ligation was detected (<0.5%) in the absence of template or in the presence of the pyrimidine strand of the template alone (lanes 2, 3, 5, and 8). Traces (<1.0% yield) of a ligation product were observed when **A** was excluded from mixtures containing the double-stranded template or the

Figure 2.13. Attempted enzymatic ligation of (^{32}P)**B** and **A** by triple helix formation; autoradiogram of high resolution, denaturing 20% polyacrylamide gel. All lanes contained (^{32}P)**B**. Lane 1: Untreated (^{32}P)**B**. Lanes 2-10: (^{32}P)**B** treated for 20 hours at 21 °C with 20 units T4 DNA ligase in solutions containing 50 mM Tris•HCl, pH 7.0, 20 mM $MgCl_2$, and 2 mM adenosine triphosphate. When **A** or **A**(Watson-Crick) were included in the reaction mixture, their concentrations were 2 μM. The purine strand of the template, pyrimidine strand of the template, or double-stranded template were included at 0.15 μM. **Purine**(Watson-Crick) was 0.18 μM when present.





–	+	+	+	+	–	–	–	–	A
–	–	+	–	+	+	–	+	–	purine
–	–	–	+	+	–	+	+	–	pyrimidine
–	–	–	–	–	–	–	–	+	A (Watson-Crick)
–	–	–	–	–	–	–	–	+	purine (Watson-Crick)

purine strand of the template alone (lanes 7 and 9). This product apparently results from ligation of two molecules of (^{32}P)**B**. **A** and (^{32}P)**B** are ligated with a 15% yield in the presence of the purine strand of the template alone (lane 4).

*The Complex of (^{32}P)**B** and **A** with the Purine Strand of the Template.* The nature of the complex that leads to condensation of **A** and (^{32}P)**B** in the presence of the purine strand of the template alone is unknown. At least 80% of the isolated condensation product (BrCN/imidazole as condensing agent) is a substrate for calf spleen phosphodiesterase, and at least 50% of the radioactivity in this product can be hydrolyzed to the labeled mononucleotide by the enzyme. These observations indicate that the predominant product of the reaction is the 3',5' phosphodiester-linked **A-B** condensation product. More complete digestion might be afforded by the use of more enzyme or a longer reaction time. Thus, in the predominant reactive complex, the 3' hydroxyl of **A** is proximal to the 5' phosphate of (^{32}P)**B**.

The protection of the bases of (^{32}P)**B** from reaction with EDC at 21 °C in the presence of the purine strand of the template alone (Figure 2.10) indicates that the complex is base paired, though a perfectly matched Watson-Crick duplex is not possible. The loss of this protection (and condensation reaction with **A**) at higher temperatures suggests that the complex is unstable at 45 °C. The enzymatic ligation of **A** and (^{32}P)**B** in the presence of the purine strand of the template alone (Figure 2.13) is further evidence of a base paired complex that aligns 5' phosphate and 3' hydroxyl termini.

One possible complex is a Watson-Crick duplex in which mismatches are accomodated. The minimum number of base mismatches in a Watson-Crick duplex of **A**, (^{32}P)**B**, and the purine template that apposes the 3' hydroxyl of **A** and the 5' phosphate of (^{32}P)**B** is four. The pairing in this duplex is shown in Figure 2.14. An alternative is a double-stranded complex in which G is paired with C and T with A, and both strands proceed from 5' to 3' in the same direction, a parallel stranded duplex.

Figure 2.14. Alignment of **A** and **B** on the purine strand of the template in a mismatched Watson-Crick duplex. Mismatches are identified with an x.

Parallel stranded duplex DNA is precedented in which the base pairing is reverse Watson-Crick; however, the duplexes of this type reported so far have included only AT base pairs.[140-149] Reverse Watson-Crick base pairing of C and G would require formation of the enol-imine tautomer of C or a displacement of the C with respect to the G to align hydrogen bond donors with hydrogen bond acceptors, allowing the formation of two hydrogen bonds in a base pair that is not isomorphous with the reverse Watson-Crick AT base pair.[140]

One parallel stranded alternative to the reverse Watson-Crick duplex is a Watson-Crick duplex in which the purine residues have adopted the *syn* conformation about the glycosidic bonds. Involvement of the *syn* conformation about the glycosidic bond in guanosine in DNA double

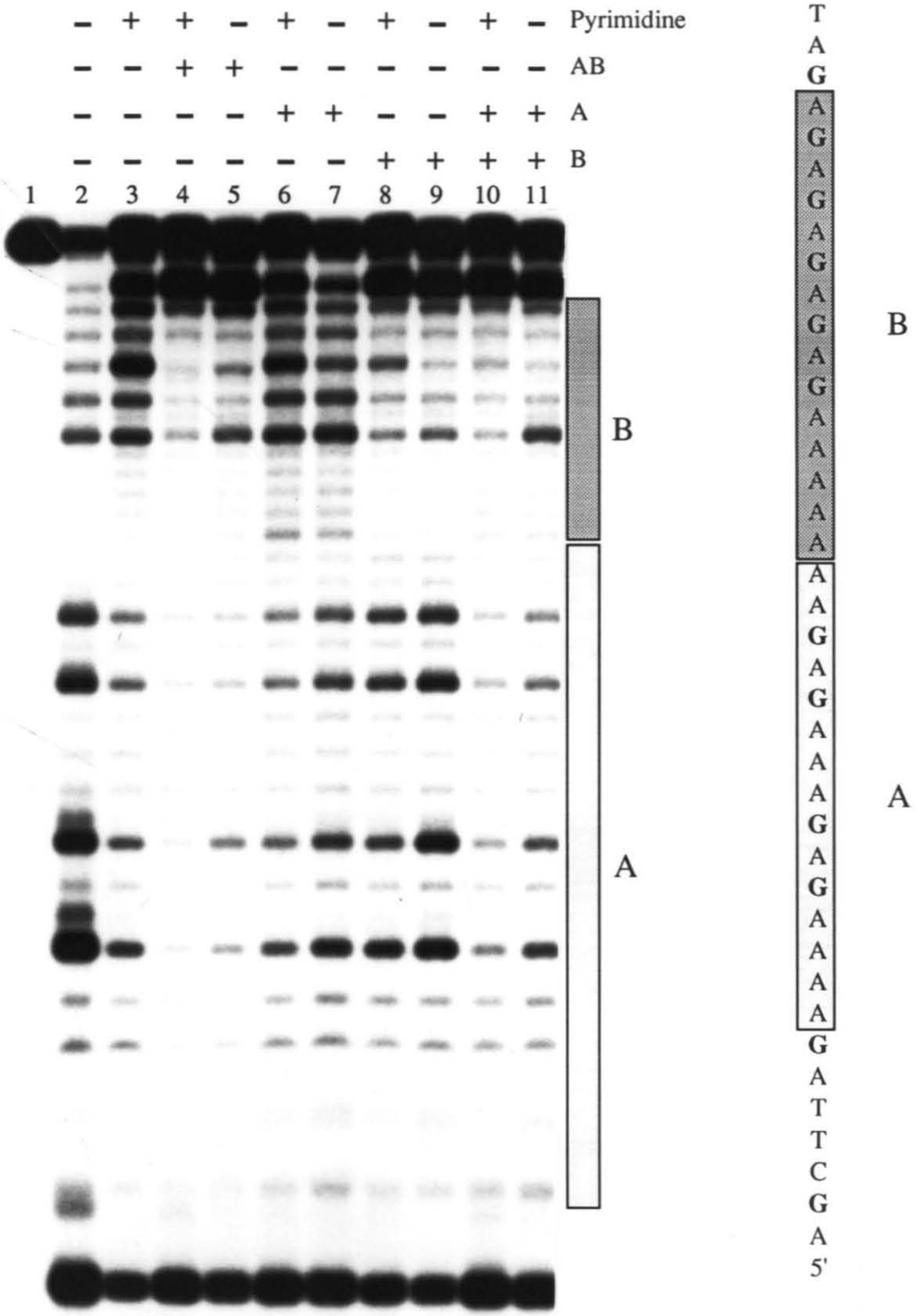
helices is well established in the Z form of DNA. However, in Z form DNA, guanosines in the *syn* conformation alternate on a single strand with cytosines in the *anti* conformation, resulting in a left-handed double helix with anti-parallel strands.[40]

Another conceivable parallel stranded complex is a Hoogsteen base paired duplex. Poly(A) in which the amino group at the 2 position of adenine has been blocked by alkylation has been found to form a double helix with poly(U), and spectroscopic and fiber diffraction data have supported a model for the structure of this complex in which the strands are parallel and the base pairing is of the Hoogsteen type.[150] Parallel stranded duplexes with unmodified AT and CG Hoogsteen base pairs have not been previously reported.

Methylation at N7 of guanosine by dimethyl sulfate (DMS) was employed as a probe for the structure of the complex of oligodeoxyribonucleotides **A** and **B** with the purine template. The purine template was labeled at its 5' terminus by phosphorylation with [γ-^{32}P]ATP and polynucleotide kinase. It was reacted at room temperature with DMS (0.2%) in 100 mM HEPES, pH 6.8, and 20 mM $MgCl_2$ in the presence and absence of pyrimidine oligodeoxyribonucleotides. The products were treated with 10% piperidine at 95 °C and analyzed by electrophoresis in a denaturing high resolution 20% polyacrylamide gel. An autoradiogram of the gel is shown in Figure 2.15A. A qualitative summary of the results is shown in Figure 2.15B.

In the absence of any pyrimidine oligodeoxyribonucleotide (lane 2), a band is present corresponding to alkylation and subsequent cleavage at each guanosine in the sequence of the purine template. In the presence of

Figure 2.15A. Dimethyl sulfate protection study of the complexes formed by association of **A**, **B**, **A-B**, and the pyrimidine strand of the template with the purine strand of the template; autoradiogram of high resolution, denaturing 20% polyacrylamide gel. The purine strand of the template, labeled with ^{32}P at its 5' terminus by phosphorylation with polynucleotide kinase and [γ-^{32}P]ATP, is in each lane. Lane 1 contains the untreated, labeled oligodeoxyribonucleotide. Lanes 2-11 contain the labeled purine strand of the template after treatment at 20 °C for 20 minutes with 0.2% dimethyl sulfate in 100 mM HEPES, pH 6.8, and 20 mM $MgCl_2$, followed by treatment with 10% (by volume) piperidine in water at 95 °C for 30 minutes. The pyrimidine strand of the template was 0.3 μM when included. **A** was 1.0 μM when included. **B** was 1.7 μM when included. **A-B** was 0.5 μM when included.


– + + – + – + – + – Pyrimidine
– – + + – – – – – – AB
– – – – + + – – + + A
– – – – – – + + + + B
1 2 3 4 5 6 7 8 9 10 11
B
A
3'
G
T
A
G
AGAGAGAGAGAAAAA
AAGAGAAAAGAGAAAAA
G
A
T
T
C
G
A
5'
B
A

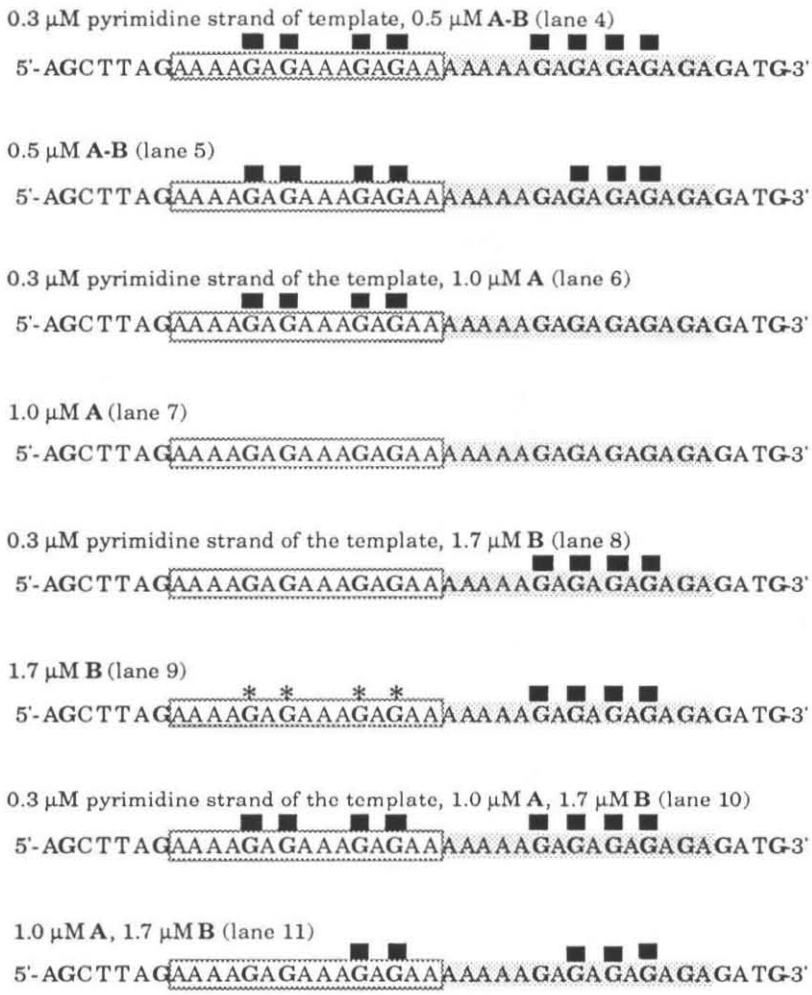


Figure 2.15B. Summary of results of the dimethyl sulfate protection study of the complexes formed by association of **A**, **B**, **A-B**, and the pyrimidine strand of the template with the purine strand of the template. Reactivity with dimethyl sulfate at each guanosine in the purine strand of the template, as judged by extent of cleavage, was compared to the reactivity in the Watson-Crick duplex formed by association of the purine and pyrimidine strands of the template (Figure 2.15A, lane 3). Protection from reaction (with respect to the Watson-Crick duplex) is indicated by ■. Greater reactivity with respect to the Watson-Crick duplex is indicated by *.

the pyrimidine strand of the template (0.3 μM, lane 3), there is generally less cleavage, with bands due to cleavage at the guanosines farthest from the 5' label being more equally represented. When both the pyrimidine strand of the template and the **A-B** 30mer (0.5 μM) are present (lane 4), additional protection from alkylation is seen at guanosines within the homopurine segment complementary in the Hoogsteen sense to **A-B**. This protection is consistent with Hoogsteen hydrogen bonding of the **A-B** strand with the purine strand of the duplex and has been previously observed for DNA triple helices.[80]

When the pyrimidine strand of the template is omitted and only the **A-B** 30mer is present (lane 5), guanosines within the homopurine segment for triple helix formation remain protected to an extent comparable to that in the triple helical complex. This protection of guanosines might be due to formation of a duplex in which **A-B** is base paired with its complementary sequence on the purine strand, imposing a parallel orientation on the strands. The extent of protection is consistent with Hoogsteen hydrogen bonding. Alternatively, **A-B**, which is in approximately 100-fold excess over the labeled purine strand, might form a 2 to 1 complex with the purine strand, a triple helix in which the Watson-Crick base pairing contains four mismatches (Figure 2.14).

In the presence of the pyrimidine strand of the template and oligodeoxyribonucleotide **B** (1.7 μM, lane 8), the five guanosines in the site for binding of **B** are protected from alkylation relative to the Watson-Crick duplex, and the four guanosines in the site for binding of **A** are at least as reactive as in the Watson-Crick duplex. In the presence of oligodeoxyribonucleotide **B** only (lane 9), the guanosines in the binding site

for **B** remain protected and those in the binding site for **A** are more reactive than in the Watson-Crick duplex. This specificity of **B** for its site of Hoogsteen complementarity is suggestive of a parallel stranded duplex. In order to form a triple helix involving the five protected guanosines with the purine strand and **B** only, accomodation of at least three mismatches would be required in the Watson-Crick pairing. Though accomodation of three mismatches might be possible under the conditions of the experiment, the reactivity of the four guanosines in the binding site for **A**, where **B** can form a Watson-Crick duplex with only one mismatch, suggests that a Watson-Crick duplex with B accomodating four mismatches is unlikely. This argument is not conclusive, however, because cooperativity between the binding of **B** in Watson-Crick and Hoogsteen modes could stabilize the complex that contains four mismatches.

In the presence of the pyrimidine strand of the template and oligodeoxyribonucleotide **A** (1.0 μM, lane 6), there is slight protection of guanosines in the binding site for **A** relative to alkylation of the Watson-Crick duplex (lane 2). However, in the presence of oligodeoxyribonucleotide **A** only (lane 7), alkylation at all guanosines is approximately equal to that in the presence of the pyrimidine strand of the template alone. Though **A** does associate with the purine strand, protecting it generally from reaction with DMS, there is not a clear specificity of **A** for its Hoogsteen complementary binding site. A mixture of several complexes might be formed.

In the presence of the pyrimidine strand of the template, **A**, and **B** (lane 10), a pattern of protection similar to that in the presence of the

pyrimidine strand of the template and **A-B** (lane 4) is seen.When the pyrimidine strand of the template is omitted, and **A** and **B** are present, there is a slight decrease in the extent of protection at guanosines in the site complementary (Hoogsteen) to **A** and at a single guanosine in the site complementary (Hoogsteen) to **B**. This result is qualitatively similar to that obtained when the labeled purine strand is treated with DMS in the presence of **A-B**. As in that case, protection could be due to formation of a parallel-stranded duplex or to formation of a triplex in which mismatches are accomodated in the Watson-Crick base pairing.

The results of these chemical protection experiments suggest that several different base paired complexes might be formed by association of **A** or **B** with the purine strand of the template. The specificity of **B** for its target site is strongly suggestive of a parallel-stranded duplex. Furthermore, the effective protection of N7 in this complex is the expected outcome of Hoogsteen hydrogen bonding, though reverse Watson-Crick hydrogen bonding is not ruled out. The apparent lack of specificity of **A** for its binding site is evidence for the likely formation of other complexes in addition to a parallel-stranded duplex, a mismatched Watson-Crick duplex, for example. In the presence of a several complexes, little can be concluded about the nature of the complex that promotes condensation of **A** and (^{32}P)**B**. Several complexes formed concurrently might promote condensation with varying efficiencies.

The Effect of Mismatches on the Ligation of Pyrimidine Oligodeoxyribonucleotides on a Double-Stranded Template. The sequence specificity of the template-directed condensation reaction can be considered

in terms of relative rates of condensation of different oligodeoxyribonucleotides on the double stranded template. The initial rate of the reaction is expected to equal the product of a pseudo-first-order rate constant and the concentration of the reactive complex in which both oligodeoxyribonucleotides are juxtaposed head-to-tail by association with the template. The sequence of the condensing oligodeoxyribonucleotides can effect this rate by effecting the rate constant for reaction of the complex and by effecting the concentration of the reactive complex.

Due to the specificity of triple helix formation, mismatches between an oligodeoxyribonucleotide and its binding site on the template will decrease the affinity of that oligodeoxyribonucleotide for its site on the template. At oligodeoxyribonucleotide concentrations below those at which the template is saturated, this decrease in affinity will decrease the concentration of the reactive complex, thereby decreasing the rate at which products are formed. This aspect of sequence specificity is observed in the affinity cleavage experiment, where the rate of cleavage of the target duplex depends on the extent of occupancy of the target site by the EDTA-bearing oligonucleotide.[27,46]

Because the condensation reaction depends upon the alignment of the terminal phosphate and hydroxyl groups in the triple helix, the rate of the reaction will depend upon the structure and conformational rigidity of the fully formed complex. The accomodation of mismatches between a condensing oligodeoxyribonucleotide and the duplex template will perturb the structure and dynamics of the complex. This perturbation is expected to alter the rate of formation of product in the reaction by altering the rate constant.

In order to study the sequence specificity of the Hoogsteen strand ligation, a set of nine oligodeoxyribonucleotides, **A'(1)-A'(9),** was prepared with the sequence of oligonucleotide **A** varied at one position. Thymidine was replaced by cytidine, adenosine, or guanosine at a) the 3' terminus of **A**, b) one nucleotide from the 3' terminus of **A**, and c) 6 nucleotides from the 3' terminus of **A**. Thus, in forming a triple helix at the binding site for **A**, each of the new oligodeoxyribonucleotides must accomodate a single mismatch of A, G, or C against an AT base pair (Figure 2.16).

A'(1): 5'-TTTTCTCT**C**TCTCTT-3'
A'(2): 5'-TTTTCTCT**A**TCTCTT-3'
A'(3): 5'-TTTTCTCT**G**TCTCTT-3'
A'(4): 5'-TTTTCTCTTTCTC**C**T-3'
A'(5): 5'-TTTTCTCTTTCTC**A**T-3'
A'(6): 5'-TTTTCTCTTTCTC**G**T-3'
A'(7): 5'-TTTTCTCTTTCTCT**C**-3'
A'(8): 5'-TTTTCTCTTTCTCT**A**-3'
A'(9): 5'-TTTTCTCTTTCTCT**G**-3'

To study the effect of mismatches on the extent of formation of the reactive complex, the yield of the ligation reaction after 6-8 hours at 24 °C was measured using varying concentrations (0.01-30.0 μM) of oligodeoxyribonucleotide **A** and each of the mismatched oligodeoxyribonucleotides and constant concentrations of duplex template (0.15 μM), condensing agent (1 mM *N*-cyanoimidazole), and $NiCl_2$ (10 mM, pH 6.9). Reactions were initiated by addition of *N*-cyanoimidazole and stopped by ethanol precipitation. All reactions with a given

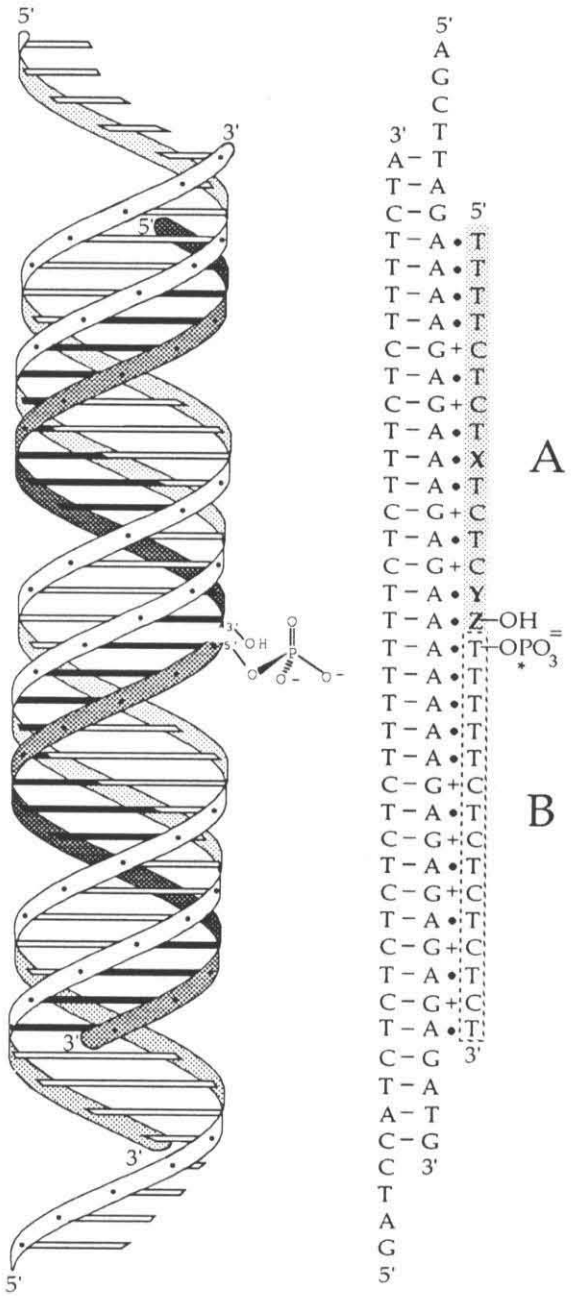


Figure 2.16. In the mismatched triple-stranded complexes, cytidine, guanosine, and adenosine are located at position X, Y, or Z of **A**, placing a mismatch against an AT base pair 6 nucleotides from the condensing 3' terminus of **A**, 1 nucleotide from the condensing 3' terminus of **A**, or at the condensing 3' terminus of **A**.

oligodeoxyribonucleotide **A** or **A'** were stopped within a period of five minutes of each other. The products were analyzed electrophoretically, and yields were quantitated by phosphor storage analysis as the percent of total (^{32}P)**B** incorporated into 30mer product.

The mean yield of two trials of the condensation reaction with **A** (no mismatch) is plotted *vs.* the logarithm of the micromolar concentration of **A** in Figure 2.17. The yield of the reaction in the absence of template at the highest concentration of **A**, 30 μM, was found in a separate experiment to

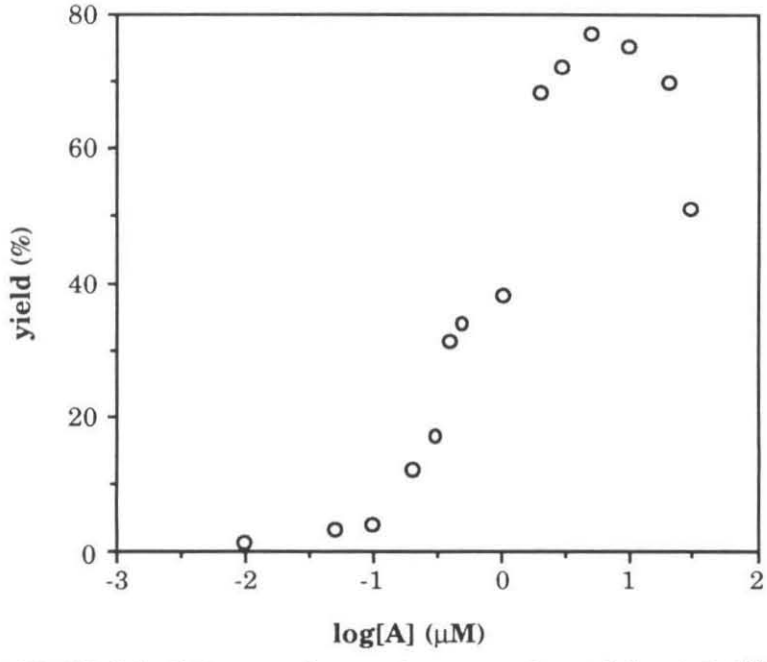


Figure 2.17. Yield of the condensation reaction of **A** with (^{32}P)**B** directed by the double-stranded template as a function of the concentration of **A**.

be 5%, confirming that at least 90% of the reaction observed in the presence of template is template-directed. The yield of the reaction at a given time before completion was expected to reflect the relative occupancy of the template by **A**. In the absence of other effects, the maximum yield for

a given reaction time would be approached with increasing concentration of **A** as saturation of the double-stranded template by **A** at its proper binding site was approached. However, the yield of the reaction decreases with increasing concentration above 5.0 μM **A** . A plausible explanation for this decrease in yield could be the association of **A** with the template at mismatched sites overlapping the binding site of (^{32}P)**B**, displacement of (^{32}P)**B** being driven by the high concentration of **A**. Another cause of the decrease might be catalysis of the hydrolytic breakdown of the condensing agent by **A**.

The yield of the condensation reactions with **A'(1)** and **A** are plotted together *vs.* the logarithm of the micromolar concentration of those oligodeoxyribonucleotides in Figure 2.17. Based on the quantitative affinity cleavage titration analysis of Singleton and Dervan,[46] **A'(1)** was expected to bind to the site for **A** on the template with an equilibrium association constant approximately 50 times lower than the equilibrium association constant for the binding of **A** at that site, due to the internal mismatch of a C against an AT base pair. This expectation led to the prediction that attainment of some fraction of the maximum yield in the condensation reaction would require a higher concentration of **A'(1)** than **A**, because a higher concentration of **A'(1)** would be required in order to form an equal concentration of the reactive complex. However, nearly equal yields of product were obtained at equal concentrations of **A** and **A'(1)**. Either the expected effect of a mismatch on binding constant was not present under the conditions of the experiment, or the yield of the reaction in the experiment described does not directly reflect the equilibrium concentration of the complex of template and **A** or **A'(1)**.

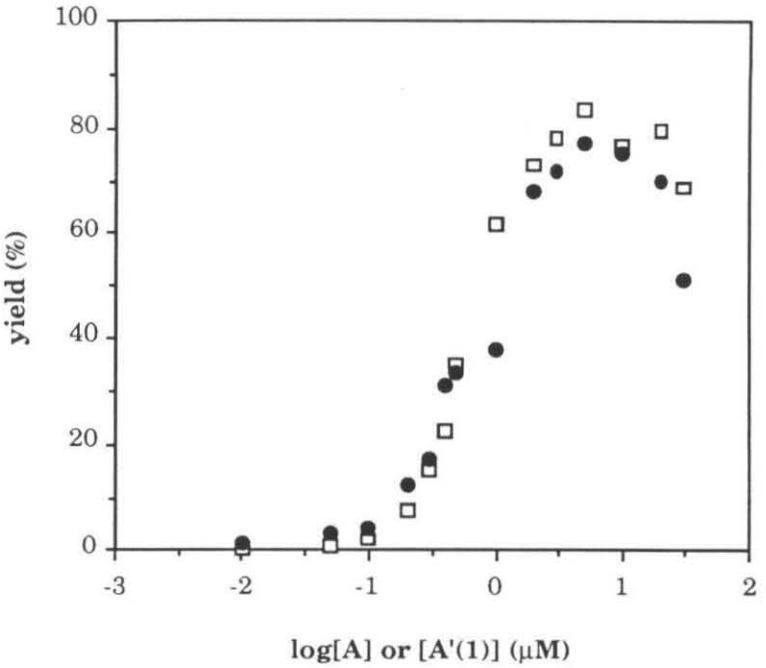


Figure 2.18. Yield of the condensation reaction of **A'(1)** or **A** with (^{32}P)**B** directed by the double-stranded template as a function of the concentration of **A'(1)** (□) or **A** (●).

The effect of a single internal base mismatch on the equilibrium association constant of a pyrimidine oligodeoxyribonucleotide with a DNA duplex was determined by Singleton and Dervan at 24 °C, in 100 mM Na^+, 1 mM spermine tetrahydrochloride, and 50 mM Tris•acetate, pH 7.0.[46] The extent to which the difference between these conditions and the conditions under which the condensation reactions were performed (1 mM *N*-cyanoimidazole, 10 mM $NiCl_2$, pH 6.9, 24 °C) effects the energetics of accomodating a mismatch in the triple helix are unknown. It is unlikely, however, that the difference would entirely negate the effect of a mismatch on the equilibrium binding constant. It is most likely that the yield of the condensation reaction as it was performed in this experiment does not

directly reflect the equilibrium concentration of the complex of template and **A** or **A'(1)**.

A plausible cause of the apparent lack of correspondence between the condensation yield and the equilibrium concentration of triple helical complex is that the reaction was initiated before equilibrium was reached. Upon mixing, oligodeoxyribonucleotides **A** or **A'(1)**, (^{32}P)**B**, and the double-stranded template associate at a rate that depends upon their concentrations.[151] At concentrations of **A** (or **A'(1)**) where its association with the template is not rapid with respect to the reaction time, the concentration of the complex is lower than the equilibrium concentration for a significant fraction of the total reaction time if condensing agent is added immediately after mixing the oligodeoxyribonucleotides, as was done in this experiment. Thus, the dependence of the yield of the reaction on the concentration of **A** reflects the dependence of the rate at which equilibrium is reached on the concentration of A as well as the equilibrium concentration of the complex.

If the rate of association of **A** with the template is on the order of the reaction time at concentrations of **A** where the template is near saturation at equilibrium, then the dependence of yield on the concentration of **A** primarily reflects the rate of formation of the triple helical complex. At concentrations of **A** and **A'(1)** where either would saturate the template at equilibrium, equivalent rates of association would result in similar functions of yield in terms of oligodeoxyribonucleotide concentration. The rates of association of **A** and **A'(1)** with the duplex template could be nearly equal if the difference in equilibrium binding constant were reflected

primarily in the dissociation rates of the matched and mismatched complexes.

An alternative explanation for the apparent lack of correspondence between the condensation yield and the equilibrium concentration of triple helical complex postulates that the oligodeoxyribonucleotides **A** and **A'(1)** catalyze the template-directed phosphodiester formation. Catalysis of the reaction by **A** or **A'(1)** would cause an increase in the yield of the reaction with increasing concentration above that at which the template is saturated. Both this and the previous explanation require that the double-stranded template be saturated with matched and mismatched oligodeoxyribonucleotide at concentrations well below 5.0 μM. Again, because of the differences between the reaction conditions in this and other studies, it is difficult to judge the plausibility of this assumption.

The yields of the condensation reactions with **A'(2)**-**A'(9)** are plotted *vs.* the logarithm of the micromolar concentrations of those oligodeoxyribonucleotides in **Appendix A**. For comparison, the mean yield of two trials of the condensation reaction with **A** (no mismatch) plotted *vs.* the logarithm of the micromolar concentration of **A** is also shown in each plot. These reactions differ in their maximum yields under the conditions of the reaction and the extent to which their yields decrease with increasing concentration above 5.0 μM. However, the shapes of the curves are qualitatively similar, and within experimental error, each reaches its maximum at a concentration of approximately 5.0 μM **A'**. Thus, for each of the oligodeoxyribonucleotides targeted to the binding site for **A** on the double-stranded template, the maximum rate of the condensation reaction

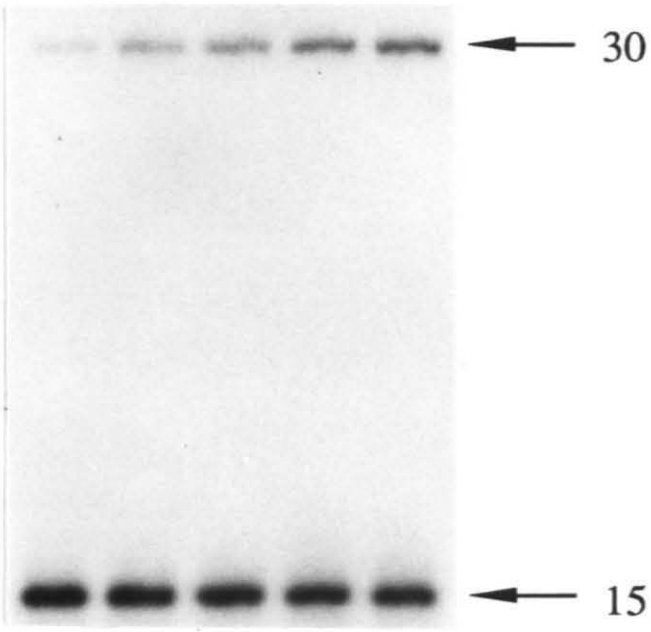
with (^{32}P)**B** can be obtained with a 5.0 μM concentration of those oligodeoxyribonucleotides.

A pseudo-first-order rate constant was measured for the template-directed condensation of (^{32}P)**B** with each of the oligodeoxyribonucleotides targeted to the binding site for **A** on the double-stranded template. Reactions were performed at 24 °C with (^{32}P)**B**, 5.0 μM **A** or **A'**, 0.15 μM double-stranded template, and 10 mM $NiCl_2$, pH 6.9. The reaction mixtures were preincubated for 3 hours at 24 °C prior to initiation of the reaction by addition of *N*-cyanoimidazole to a final concentration of 1.0 mM. After initiation of the reaction, aliquots were removed from the reaction mixtures at ten minute intervals and quenched by addition of sodium acetate, pH 5.2 and ethanol, followed by freezing in dry ice. After 50 minutes, the last aliquot was quenched, and the precipitates were analyzed by electrophoresis. An autoradiogram of a gel on which the time course of the reaction with **A** was analyzed is shown in Figure 2.19A. The radioactivity in the bands corresponding to 15mer starting material and 30mer product was quantitated by storage phosphor analysis.

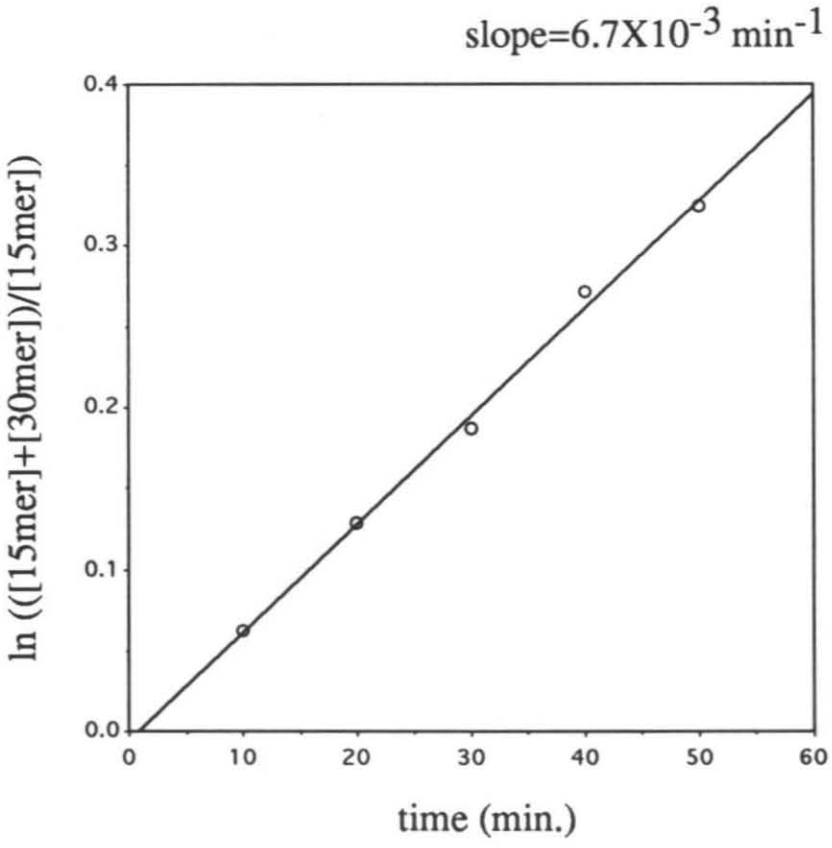
The pseudo-first-order rate constant was determined for each reaction as the slope of the line defined by plotting *vs.* time the natural logarithm of the ratio of the total radioactivity and the radioactivity remaining as starting material. This analysis is shown in Figure 2.19B for the reaction of Figure 2.19A. This determination of the rate constant relies on several assumptions: 1) the amounts of radioactivity precipitated and loaded on the gel reflect the relative amounts of 15mer and 30mer in solution; 2) the fraction of (^{32}P)**B** of which the terminal phosphate is activated rapidly

Figure 2.19. Determination of the pseudo-first-order rate constant from the initial rate of template-directed condensation of **A** with (^{32}P)**B**. (A) Autoradiogram of high resolution, denaturing 20% polyacrylamide gel showing the time course of the production of 30mer from the 15mer (^{32}P)**B** by condensation with 15mer **A**. (B) Graphical analysis of the gel shown in (A). The pseudo-first-order rate constant of the reaction is taken to be the slope of the line fit to a plot of ln(([15mer]+[30mer])/[15mer]) *vs.* reaction time. Concentrations of (^{32}P)**B** and **A**(^{32}P)**B** in solution are assumed to be proportional to the amount of radioactivity observed on the gel for those species.

A


30
15
10 20 30 40 50
reaction time (min.)

B


slope=6.7X10-3 min-1
ln (([15mer]+[30mer])/[15mer])
0.0
0.1
0.2
0.3
0.4
0 10 20 30 40 50 60
time (min.)

reaches a steady state and does not change during the 50 minute course of the reaction; and 3) the concentration of the fully formed triple helical complex of the template with **A** and (^{32}P)**B** does not change during the course of the reaction. The quality of the fits of the first order rate law to the data supports the validity of these assumptions.

The rate constants measured for the condensation reactions of **A** and **A'(1)-A'(9)** are listed on Table 2.2. Each determination is the mean of three

Oligonucleotide	rate constant ($\times 10^{-3}$ min^{-1})	number of determinations
A	9.7 ± 2.5	5
A'(1)	3.4 ± 1.6	4
A'(2)	5.4 ± 2.5	5
A'(3)	2.8 ± 1.5	4
A'(4)	1.2 ± 0.4	6
A'(5)	0.9 ± 0.3	5
A'(6)	0.4 ± 0.1	6
A'(7)	0.2 ± 0.05	5
A'(8)	0.4 ± 0.1	3
A'(9)	0.2 ± 0.05	3

Table 2.2. Rate constants measured for the template-directed condensation of (^{32}P)**B** with **A** and **A'(1)-A'(9)** at 24 °C with 1 mM *N*-cyanoimidazole, 10 mM $NiCl_2$, pH 6.9, 0.15 μM double-stranded template, and 5.0 μM **A** or **A'**.

to six trials, and uncertainties are expressed as standard deviations in the means. These rate constants are represented graphically in Figure 2.20. A mismatch at the 3' terminus of **A** decreases the rate of the reaction by a

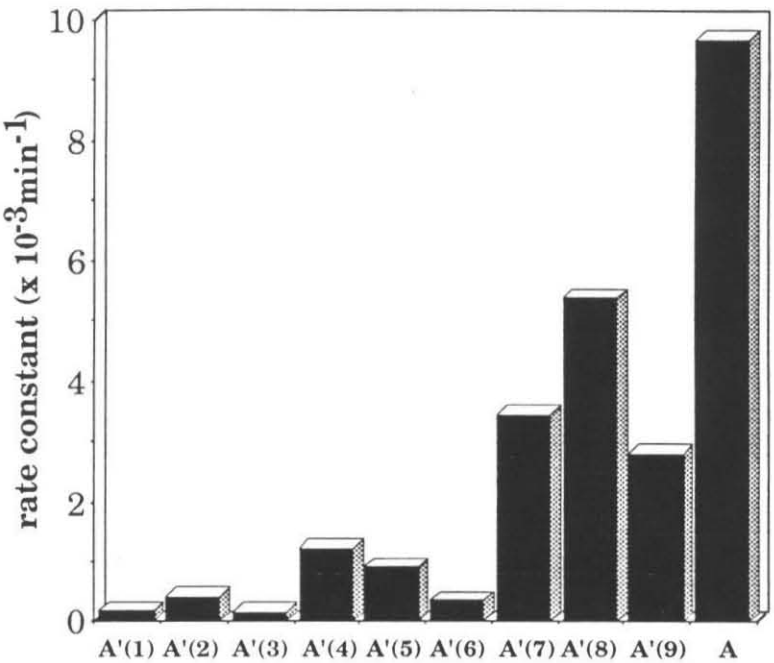


Figure 2.20. Graphical representation of rate constants listed on Table 2.2

factor of 25 to 60. This decrease is comparable to the factor by which affinity of a pyrimidine oligodeoxyribonucleotide for a target site on a DNA duplex is decreased by a single internal mismatch. Mismatches at the base position one removed from the 3' terminus of **A** have a generally smaller, though still significant effect on the rate of condensation, decreasing the rates by factors of 8 to 27. Surprisingly, a mismatch six nucleotides removed from the condensing terminus apparently has an effect on the rate of the reaction. Application of Student's t test[152,153] to the values obtained for **A** and **A'(8)** (A in third strand) lends at least 95% confidence to the conclusion that the difference in measured rates is not entirely due to experimental uncertainty. The smaller, more precise values for the reaction rates of **A'(7)** and **A'(9)** are even more likely to differ from the reaction rate of **A** by a statistically significant amount.

If the assumption is made that the double-stranded template is saturated with oligodeoxyribonucleotide targeted to the binding site for **A** in each of these rate determinations, then the rates measured represent the rates of reaction of the fully formed template-substrate complex. There is some doubt of the validity of this assumption due to the ambiguity in interpretation of the experiments discussed above relating reaction yield to oligodeoxyribonucleotide concentration. Thus, the relative rates might reflect the combined effects of the the reactivity of the assembled complex and the extent to which the complex is assembled. In either case, the decreasing rate as the mismatch is placed closer to the site of reaction is likely to reflect the effect on reactivity of the assembled complex.

A mismatch at the ligating terminus can be imagined to effect the rate of condensation in two ways: 1. the interaction of the mismatched base with its corresponding base pair on the template is less stable than the matched Hoogsteen hydrogen bonding interaction, allowing the terminus greater mobility and 2. noncanonical interactions between the mismatched base and its corresponding base pair on the template place the equilibrium position of the condensing hydroxyl in an unfavorable orientation for attack on the activated phophate. Both of these effects lower the effective concentration of the condensing hydroxyl in the correct position for attack on the activated phosphate.

Similar effects might result from a mismatch one base removed from the reacting terminus. Structural distortion due to the mismatch might be propagated to the neighboring position, unfavorably orienting the reacting hydroxyl. However, NMR data have suggested that mismatches in the Hoogsteen base paired strand of an intramolecular triple helix at the site of

a GC base pair can be accomodated with little distortion of the overall structure of the complex.[154] Alternatively, the mismatch one base from the terminus might destabilize the interaction of the terminal base with the template, allowing the terminal hydroxyl greater mobility.

It is highly unlikely that structural distortion resulting from a mismatch six bases from the ligating terminus would be propagated through the complex to alter the orientation of the terminal hydroxyl. It is possible that the mismatch makes the entire structure somewhat more dynamic, increasing the mobility of the apposed termini with respect to each other. Alternatively, the decrease in measured rate upon placing a mismatch six bases from the terminus might indicate the extent to which the mismatch effects the occupancy of the oligodeoxyribonucleotide at its target site on the template.

Enzymatic ligation of oligodeoxyribonucleotides on a single-stranded DNA template has been developed as a means of detecting single-base mutations.[155] No enzymatic ligation is observed if either of the bases flanking the site of ligation is mismatched. A highly sensitive gene amplification and detection technique, termed the ligase chain reaction, employs a thermostable DNA ligase to combine thermal cycles of annealing and denaturation with the oligonucleotide ligation assay.[156] In principle, the nonenzymatic ligation of single-stranded pyrimidine oligodeoxyribonucleotides on a double-stranded template could similarly be used as a means of detecting single-base mutations without the requirement of a single-stranded template. However, maintenance of a chain reaction would be complicated by the simultaneous operation of both Hoogsteen and Watson-Crick base pairing.

Conclusions. Double-stranded DNA can serve as a template to align reactive termini of oligonucleotides and promote their condensation. In this work, condensation of a 5' phosphate with a 3' hydroxyl was effected, producing a phosphodiester linkage. The reaction was found to be particularly efficient when the phosphate was activated in the presence of transition metal ions such as Ni^{+2} with cyanogen bromide and imidazole or *N*-cyanoimidazole. Condensation could also be observed when the phosphate was activated as a phosphorimidazolide or as an *O*-phosphoryl urea with a water soluble carbodiimide. A 25-60-fold decrease in rate is observed when the base at the position of the condensing hydroxyl terminus is mismatched. Smaller decreases in rate result from the placement of mismatches further from the ligating terminus.

In a formal sense, this reaction allows detection of a sequence of double-stranded DNA within a larger duplex, ligation being observed only if the targeted template is present. Greater sequence specificity might be afforded the nonenzymatic reaction if its sensitivity to a single mismatch could be coupled with the full sensitivity of the affinity of the condensing strands to single mismatches. This combination of effects might be observed at lower concentrations of the mismatched oligodeoxyribonucleotides. Unlike the affinity cleavage experiment, information is not gained about the location of the target within the duplex, only about its presence or absence. A practical limitation to the detection of sequences by the condensation reaction studied here is the requirement for an amount of target duplex comparable to or in excess of the labeled oligonucleotide. With large DNA molecules (e.g., the 48.5 kbp genome of

bacteriophage lambda), this amount of DNA can be experimentally cumbersome.

The significant effects of mismatches on the rates of oligodeoxyribonucleotide ligation directed by a double-stranded template suggest that the chemical ligation can serve as a probe of the stucture and dynamics of three-stranded complexes. Shabarova and coworkers have suggested that comparison of chemical ligation efficiencies in double- and triple-stranded complexes can serve as an indicator of backbone conformation.[157] Ligation efficiencies might also serve as guides for the optimization of structures designed to recognize double-stranded DNA by triple helix formation.

Though the double-stranded template does not behave as a catalyst under the conditions studied, the condensation reaction has relevance to catalysis by nucleic acids. It demonstrates the potential of a double helix to align substrates for a reaction by providing adjacent binding sites in its major groove. Thus, the grooves of a double helix can, in principle, behave as enzymatic active sites. It is noteworthy that a number of the RNA catalysts discovered so far catalyze phosphoryl transfer reactions such as the transesterifications catalyzed by the Group I introns.[158,159] Attempts to observe transesterification directed by the double stranded template are described in **Appendix B**.

In the system studied here, chemical information is transferred from the double-stranded template to the product. By sequence-specific condensation of two oligonucleotides, an oligonucleotide of a new sequence is created. The new sequence is neither identical to nor the Watson-Crick complement of either strand of the template; yet, it contains the sequence

information of the double-stranded template translated through the recognition code of triple helix formation. The existence of multiple codes by which chemical information can be recognized and translated (e.g., Watson-Crick hydrogen bonding, Hoogsteen hydrogen bonding, reverse-Watson-Crick hydrogen bonding, etc.) has implications for the evolution of a template-directed information transfer system. Such systems have been postulated as precursors to biological self-replication.[1,2,103] On one hand, multiple codes allow the system to explore a larger sequence space, providing the potential for variability without mutation. On the other hand, they provide pathways of information transfer that can compete with self replication, making stable maintenance of information content less likely.

Materials and Methods

Ultraviolet absorption spectra were recorded with a Hewlett Packard 8452A diode array spectrophotometer or a Perkin-Elmer Lambda 4C spectrophotometer. Radioactivity was quantitated with a Beckman LS3801 scintillation counter or by phosphor storage analysis with a Molecular Dynamics Phosphorimager. Autoradiograms were made with Kodak XAR 5 X-ray film and developed with a Kodak M35A X-omat film processor. The melting point of *N*-cyanoimidazole was measured with a Thomas-Hoover

apparatus. Measurements of pH were made with an Orion Research pH meter. Lyophilization was performed with a Savant Speed-Vac.

Reagents for oligonucleotide synthesis were obtained from Applied Biosystems, Inc. and phosphoramidites were obtained from Cruachem. Imidazole and metal salts were obtained from Aldrich. Cobalt hexammine trichloride was obtained from Eastman Kodak as were cyanogen bromide and hydrazine. Tris, MES, and HEPES buffers, 1-ethyl-3-(3-dimethylamino-propyl)carbodiimide (EDC), and spermine were obtained from Sigma. Sephadex NAP-5 columns were from Pharmacia.

Reactions with oligonucleotides were performed in polypropylene conical tubes in solutions prepared with water purified with a Milli-Q purification system. Oligonucleotides were precipitated by the addition of 3-5 volumes punctilious ethanol, 0-4 °C, and 0.1 volume of 3 M sodium acetate, pH 5.2, or 1 volume of saturated sodium chloride. The resulting mixtures were centrifuged at 14,000 RPM, 0-4 °C, for 30 minutes in an Eppendorf microcentrifuge.

Polyacrylamide Gel Electrophoresis. Denaturing polyacrylamide gels contained 45% (by weight) urea (ICN) and 15% or 20% polyacrylamide (Boehringer Mannheim) crosslinked 1:20 with *N,N'*-methylene-bis-acrylamide (USB) in a buffer containing 89 mM Tris•HCl, pH 8.3, 89 mM boric acid, and 2 mM ethylenediamine tetraacetic acid (TBE). Urea was omitted from nondenaturing gels. Polymerization was initiated by the addition of ammonium persulfate and *N,N,N'N'*-tetramethylethylene-diamine. Gels for the purification of oligonucleotides were 2 mm thick and 38 cm long. Gels for the purification of radiolabeled oligonucleotides were 2

mm thick and 16 cm long. High resolution gels were 0.4 mm thick and 40 cm long. Gels were run in the Tris•borate•EDTA (TBE) buffer described above. Gels containing radiolabeled material were visualized by autoradiography or phosphorimager analysis. Quantitation of relative amounts of radioactivity by phosphorimager analysis and by excising the gel slices containing radioactivity and quantitating the radioactivity by scintillation counting gave equivalent results within the limits of certainty of the measurements in the assays reported here. Yields are expressed as the percentage of total radioactivity present in a lane on a gel as product. Control experiments were performed to substantiate the assumption that 15mer starting material and 30mer product of condensation reaction were precipitated with ethanol with approximately equal efficiencies.

Preparation of oligodeoxyribonucleotides. Oligodeoxyribonucleotides were obtained deprotected from the Caltech Microchemical Facility or synthesized with a Beckman System 1 Plus or Applied Biosystems 380B automated DNA synthesizer by standard phosphite triester methodology using β-cyanoethylphosphoramidites.[160] Those synthesized with the Beckman DNA synthesizer were removed from the support and deprotected concurrently by warming in concentrated ammonium hydroxide solution at 55 °C for 24 hours. Those synthesized with the ABI DNA synthesizer were removed from the support with concentrated ammonium hydroxide at room temperature on the machine using an automated deprotection cycle. After removal from the support, approximately one milliliter of fresh concentrated ammonium hydroxide was added and the solution was warmed at 55 °C for 24 hours. After

deprotection, ammonia was removed by bubbling argon through the solution for several hours or by letting the solution set open in a fume hood for twelve to sixteen hours. The solutions were lyophilized.

All oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis. Template oligonucleotides and oligonucleotide A-B were purified on gels of 15% polyacrylamide. All other oligonucleotides were purified on 20% polyacrylamide gels. After deprotection and drying, the oligonucleotides were suspended in a solution of 85% (by volume) formamide in 75 mM Tris•HCl, pH 8.3, 75 mM H_3BO_3, and 1.5 mM EDTA. In some cases, xylene cyanol and bromophenol blue were included as marker dyes. The resulting mixtures were heated to 95 °C for several minutes, then deposited in the wells of a gel. Approximately half of a 1 μmole synthesis was loaded on one gel. The gel was run until the full length oligonucleotide was estimated to have migrated at least two thirds of the length of the gel.

The DNA was visualized by ultraviolet shadowing over an activated TLC plate, and the upper two thirds of the band were cut from the gel, crushed, and soaked in water or 0.2 M NaCl at 37 °C for 12 to 24 hours. The supernatant was removed from the gel by centrifugation or filtration (.45 μm filter) and desalted by dialysis against water. Template oligodeoxyribonucleotides were ethanol precipitated and redissolved in water instead of dialysis. The concentrations of all oligodeoxyribonucleotides were determined by their ultraviolet absorbance at 260 nm. Extinction coefficients were calculated accounting for nearest neighbor hypochromicity.[161,162]

Table 2.3. Extinction Coefficients of Oligodeoxyribonucleotides

Oligodeoxyribonucleotide	Extinction coefficient at 260 nm
Purine strand of template	$4.4 \times 10^5\ M^{-1}cm^{-1}$
Pyrimidine strand of template	$3.4 \times 10^5\ M^{-1}cm^{-1}$
Oligodeoxyribonucleotide **A**	$1.2 \times 10^5\ M^{-1}cm^{-1}$
Oligodeoxyribonucleotide **B**	$1.2 \times 10^5\ M^{-1}cm^{-1}$
Oligodeoxyribonucleotide **A-B**	$2.4 \times 10^5\ M^{-1}cm^{-1}$
Purine(Watson-Crick)	$4.7 \times 10^5\ M^{-1}cm^{-1}$
A(Watson-Crick)	$1.2 \times 10^5\ M^{-1}cm^{-1}$

Annealing and Purification of Double-Stranded Template. Equimolar amounts of the purine strand of the template and the pyrimidine strand of the template were combined and dried under vacuum (speed-vac). The residues were redissolved in water, heated at 90 °C for 15 minutes, allowed to cool to room temperature, then chilled to 6 °C. The solution was mixed with 8 ml of 12.5% Ficoll containing xylene cyanol and bromophenol blue and loaded on a nondenaturing 15% polyacrylamide gel. Individual template strands were loaded in adjacent lanes as size markers. The gel was run at 4 °C until the bromophenol blue reached the bottom. The gel was stained in a solution of ethidium bromide (Sigma, 2 μg/ml in TBE) and visualized with ultraviolet light. The gel slice containing the duplex was excised, crushed, and soaked in 0.2 M NaCl at 37 °C for 20 hours. The suspension was filtered (0.45 μm), concentrated with n-butanol, and ethanol precipitated. The precipitate was dried under vacuum and redissolved in water. The concentration of the duplex template was

determined by the absorbance of the solution at 260 nm. The extinction coefficient of the duplex was calculated to be 5.8×10^5 $M^{-1}cm^{-1}$ by summing the extinction coefficient of double stranded DNA with the extinction coefficients calculated for the single stranded terminal regions assuming nearest neighbor hypochromicities.

5' Labelling of Oligodeoxyribonucleotides. 70 pmoles of oligodeoxyribonucleotide were phosphorylated in a reaction mixture containing kinase buffer, adenosine 5'-[γ-^{32}P]triphosphate (Amersham), and polynucleotide kinase (50-100 units, New England BioLabs), in a total volume of 50 µl. After reaction at 37 °C for one to three hours, the reaction mixture was concentrated to approximately 5 µL, mixed with 10 µL of formamide loading buffer, heated to 95 °C for several minutes, and purified by electrophoresis in a denaturing 20% polyacrylamide gel. Following electrophoresis, the labeled oligonucleotide was visualized by autoradiography. The upper half of the band was excised, crushed, and soaked in water or 0.2 M NaCl at 37 °C for 16 to 24 hours. The supernatant was removed by filtration or centrifugation and the DNA was precipitated with ethanol. The precipitate was dried under vacuum and dissolved in water, typically to a specific activity of 50,000-100,000 CPM/ µL.

Activation of Terminal Phosphate with Cyanogen Bromide and Imidazole. Reactions were performed in 30 µL solutions initially containing (^{32}P)**B**, **A** at the concentration indicated, purine strand of the template, pyrimidine strand of the template, or duplex template at the concentration indicated, 20 mM imidazole•HCl, pH 7.0, polyvalent counterion as indicated in the

discussion of results, and 100 mM NaCl. Reaction was initiated by addition of cyanogen bromide to a concentration of 40 mM. The reaction was allowed to proceed 9 hours at room temperature followed by ethanol precipitation of the DNA. The reactions were analyzed by high resolution polyacrylamide gel electrophoresis.

Analysis of Product Formed in Condensation Reaction Directed by Double Stranded Template Using Cyanogen Bromide and Imidazole as Condensing Agent. The product of the condensation reaction of (^{32}P)**B** was isolated for analysis by digestion with calf spleen phosphodiesterase. The condensation reaction was run as above on a 28-fold scale, using 20 mM $NiCl_2$, 1.3 μM **A**, and 0.15 μM duplex template. The product was purified by electrophoresis in a denaturing 20% sequencing gel and extracted from the gel by crushing the gel and soaking it in 0.2 M NaCl. The supernatant was removed by filtration and concentrated under vacuum. The internally labeled reaction product was precipitated with ethanol, dried, and dissolved in 80 μL of water. 2 μL of this solution were incubated with 4 μg of calf spleen phosphodiesterase (Boehringer Mannheim) in a total volume of 20 μL at 37 °C for 8 hours. The water was then evaporated to dryness, and the residue was suspended in formamide electrophoresis buffer and analyzed by electrophoresis in a denaturing 20% polyacrylamide sequencing gel.

For sequencing analysis, the condensation reaction was carried out with oligodeoxyribonucleotide **B** phosphorylated with unlabeled (^{31}P) ATP (Calbiochem). Phosphorylation and template-directed condensation were carried out as above, locating the products on gels from the distance

migrated by labeled products run on the same gel. After isolation from the gel, the unlabeled condensation product was dissolved in 40 μL of water and phosphorylated with polynucleotide kinase and γ-^{32}P ATP in a 60 μL reaction volume. The 5'-labeled material was purified by gel electrophoresis and dissolved in water to a specific activity of approximately 15,000 CPM/μL.

Cytosine specific cleavage was effected by reaction with hydrazine in NaCl solution followed by piperidine treatment.[136] 2 μL of the product solution were mixed with 30 μL of hydrazine and 18 μL of 5 M NaCl. The reaction was allowed to proceed for 45 minutes at room temperature and stopped by addition of 200 μL of a solution containing 0.6 M sodium acetate, 0.1 mM EDTA, and 25 μg/ml tRNA. The DNA was precipitated by addition of 750 μL of ethanol, and the precipitate was washed once with 70% ethanol and dried under vacuum. It was treated with 10% (vol) piperidine in water at 90 °C for one hour and dried under vacuum overnight. To remove lingering traces of piperidine, the dried residue was two times wetted with 10 μL of water and dried again under vacuum. The final dried residue was suspended in formamide loading buffer for analyis by polyacrylamide gel electrophoresis.

Thymidine specific cleavage was effected by reaction with potassium permanganate followed by piperidine treatment.[137] Two μL of the 5' labeled product solution were treated with potassium permanganate at a concentration of 46 μg/ml in a total of 25 μL aqueous solution. After 2 minutes at room temperature, the reaction was stopped by addition of 30 μL of a solution containing 16% β-mercaptoethanol, 0.5 mg/mL tRNA, and 5% (by weight) sodium acetate, pH 5.2. The DNA was precipitated with 180

μL of ethanol and the precipitate was washed with 70% ethanol and dried under vacuum. The dried precipitate was treated with piperidine and analyzed electrophoretically as described for the cytosine specific reaction. 2 μL of the solution of 5' labeled condensation product were also incubated with 4 μg of calf spleen phosphodiesterase in a total volume of 20 μL at 37 °C for 8 hours. The water was then evaporated and the residue was suspended in formamide electrophoresis buffer and analyzed by electrophoresis in a denaturing 20% polyacrylamide sequencing gel.

Preparation of *N*-Cyanoimidazole was by the method of Giesemann.[163] A solution of 5.5 g (0.05 moles) of BrCN in 25 ml of benzene was added dropwise to a solution of 3.2 g (0.05 moles) imidazole in 50 ml of benzene warmed to 50 °C. After addition, the mixture, in which a pale yellow, fine precipitate was forming, was warmed for an additional ten minutes, then cooled to 4 °C. After 24 hours at 4 °C, the bright yellow precipitate was removed by filtration, and the clear, colorless filtrate was concentrated under vacuum. The white, crystalline residue was purified by sublimation, yielding 0.53 g (6 mmoles, 12% yield) of product, mp=59-60°C.

Template-Directed Condensation Reaction of A and (^{32}P)B Using *N*-Cyanoimidazole as Condensing Agent. Reaction was initiated in mixtures containing **A**, (^{32}P)**B**, purine strand of the template, pyrimidine strand of the template, or duplex template, and divalent metal ion at the concentrations indicated in the discussion of results by addition of a freshly prepared solution of *N*-cyanoimidazole in water to a final volume of 30 μL.

After the indicated reaction time, the DNA was precipitated with ethanol, and the reaction products were analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Condensation Reactions of (^{32}P)B Activated with Imidazole, 2-Methylimidazole, or 4-Methylimidazole and Water Soluble Carbodiimide. 1 μL of (^{32}P)**B** was activated for each reaction by mixing with 100 μL of a 2.2 mg/100 μL solution of EDC in 200 mM imidazole•HCl, 2-methylimidazole•HCl, or 4-methylimidazole•HCl, pH 6.1. Activation was allowed to proceed for 1 to 5 hours at room temperature after which the DNA was ethanol precipitated (using NaCl). The precipitate was dried briefly under vacuum. The precipitate was used directly in 30 μL condensation reaction mixtures containing 20 mM imidazole•HCl, pH 7.0, 100 mM NaCl, polycation salt at the concentration indicated in the discussion, 20% THF, 1 μM purine strand of the template, pyrimidine strand of the template, or duplex template, and 1.3 μM **A**. Reactions were performed at the temperatures indicated for 20-24 hours, then ethanol precipitated and analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Electrophoretic Purification of Putative 5' Phosphorimidazolide of Oligodeoxyribonucleotide B. 2X40 μL of (^{32}P)**B** were mixed with 300 μL of 2.3 mg EDC/100 μL in 200 mM 2-methylimidazole•HCl, pH 6.1. The reaction was allowed to proceed for four hours, after which the DNA was ethanol precipitated (40 μL of saturated NaCl and 1.2 mL ethanol). The

precipitate was dried under vacuum, suspended to the extent possible in formamide loading buffer, and run on a denaturing 20% polyacrylamide gel. The starting material and product were visualized on the gel by autoradiography and the product band was excised. The excised gel slice was crushed and soaked in 200 mM 2-methylimidazole, pH 10, at 37 °C for 4 hours. The supernatant was removed (centrifugation) and the DNA was precipitated (as above with NaCl). The precipitate was dissolved in 80 μL of water. A second analysis by denaturing 20% polyacrylamide gel electrophoresis indicated that approximately 80% of the purified material was the putative phosphorimidazolide.

Activation of (^{32}P)B in Condensation Reaction with EDC. Condensation reactions using EDC directly to activate the phosphate were performed in a total volume of 30 μL in HEPES buffer, pH 6.8, containing polyvalent cation salt as indicated in the discussion of results, 100 mM NaCl, 30% THF and 60 mM EDC. Reactions were performed at the indicated temperature for 20 hours, followed by ethanol precipitation of the DNA and analysis of the products by denaturing 20% polyacrylamide gel electrophoresis.

Ligation of Oligodeoxyribonucleotides on Single- and Double-Stranded Templates Using T4 DNA Ligase and T4 RNA Ligase. Reaction mixtures contained 50 mM Tris•HCl, pH 7.0, 20 mM $MgCl_2$, 0.15 μM purine strand of the template, pyrimidine strand of the template, or duplex template, or 0.18 μM **purine**(Watson-Crick) (or no template), 2 μM **A** or **A**(Watson-Crick), or no **A**, 1 μL of (^{32}P)**B**, 2 mM ATP, and 20 units New England BioLabs T4 DNA ligase, 1 unit Boehringer Mannheim T4 DNA ligase, or 10

units New England BioLabs T4 RNA ligase in a total volume of 30 μL at 21 °C. After 20 hours, the DNA was precipitated with sodium acetate and ethanol and analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Analysis of the Product of Condensation of A and (^{32}P)B in the Presence of the Purine Strand of the Template. The condensation reaction was performed at room temperature in a reaction mixture containing 28 μL of (32P)**B**, 20 mM imidazole•HCl, pH 7.0, 20 mM $NiCl_2$, 100 mM NaCl, 0.7 μM **A**, 0.15 μM purine strand of the template, and 40 mM cyanogen bromide in a total volume of 840 μL. After 10.5 hours, the reaction was stopped by ethanol precipitation, and the product was purified by polyacrylamide gel electrophoresis and extracted from the gel by methodology already described. The purified product was dissolved in 18 μL of water. 2 μg of calf spleen phosphodiesterase was added, and the mixture was warmed at 37 °C for 5 hours, after which another 2 μg of calf spleen phosphodiesterase was added followed by an additional 3 hours of reaction at 37 °C. Finally, the mixture was lyophilized and analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Protection of the Purine Strand of the Duplex Template from Reaction with Dimethyl Sulfate by Pyrimidine Oligodeoxyribonucleotides. The purine strand of the template (45 pmoles) was 5' labeled as described above, and the purified product was dissolved in 1 ml of water. 1 μL of this solution was treated at 20 °C with 0.2% dimethyl sulfate (2 μL of a freshly prepared solution in water added to a total reaction volume of 20 μL) in a solution

containing 100 mM HEPES, pH 6.8, and 20 mM $MgCl_2$. Mixtures also contained 0.3 μM pyrimidine strand of the template, 0.5 μM **A-B**, 1.7 μM **B**, 1.0 μM **A** or a combination of these oligodeoxyribonucleotides at the listed concentrations. Reactions were initiated by addition of the dimethylsulfate solution and were allowed to proceed for 20 minutes. Reactions were stopped by addition of 10 μL of a solution containing 0.1 mg/mL tRNA, 7% 2-mercaptoethanol, and 12% sodium acetate, pH 5.2. The DNA was precipitated by addition of 150 μL of ethanol. The precipitates were dried and treated with 100 μL of 10% piperidine at 95 °C for 30 minutes. Piperidine was removed under vacuum, and the reactions were analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Effect of Mismatches on Hoogsteen Ligation. In order to standardize their purity, oligodeoxyribonucleotides **A** and **A'(1)-A'(9)** were purified by reverse phase FPLC following purification by electrophoresis. A pro-RPC column was used. Flow rate 4.0 mL/min. Buffer A, 100 mM TEAA, pH 7.0; buffer B 40% CH3CN, 100 mM TEAA, pH 7.0. Inject at 100% A for 9 minutes, ramp to 100% B over 40 minutes, maintain 100% B for 10 minutes, ramp down to 0% B in 5 minutes, wash the column with 100% A for 10 minutes. Each of the oligonucleotides eluted as a single peak at approximately 40% B and was collected manually in 2 to 4 mL. The collected fractions were lyophilized, desalted with a Waters Sep-Pak column, lyophilized again, and dissolved in 1.0 mL of water. The concentration of this solution was determined by UV absorbance at 260 nm. 850 μL of the solution were concentrated to 500 μL and further desalted

with a Sephadex NAP-5 column. The eluate was concentrated to a final oligodeoxyribonucleotide concentration of 100 μM.

Condensation reactions to determine the dependence of ligation yield on the concentration of oligodeoxyribonucleotide **A** (or mismatched oligodeoxyribonucleotide **A'**) were performed in a total volume of 30 μL containing 10 mM $NiCl_2$, 0.15 μM duplex template, 2 μL of (^{32}P)**B** (final concentration of approximately 10 nM), and **A** (or mismatched A') at concentrations varying from 0.01 μM to 30 μM. $NiCl_2$ was added as a 100 mM solution with its pH adjusted to 6.9 with concentrated NaOH. Reaction was initiated by addition of a freshly prepared solution of *N*-cyanoimidazole in water to a concentration of 1 mM. Reactions were allowed to proceed for 6 to 8 hours at room temperature followed by ethanol precipitation with sodium acetate and analysis by denaturing 20% polyacrylamide gel electrophoresis. The yields of the reaction were quantitated by phosphorimager analysis.

In order to measure the initial rates of condensation, reaction mixtures were prepared and equilibrated for 3 hours prior to addition of *N*-cyanoimidazole. Reactions were then initiated by addition of a freshly prepared solution of 10 mM *N*-cyanoimidazole to a concentration of 1 mM. In addition to *N*-cyanoimidazole, reaction mixtures initially contained 10 mM $NiCl_2$, pH 6.9, 0.15 μM duplex template, 5 μM **A** or mismatched **A'**, and 2 μL of (^{32}P)**B**. After initiating the reactions, the mixtures were divided into aliquots, one for each time point. At each time point, reaction was stopped in an aliquot by quenching with 0.3 volumes 20% sodium acetate, pH 5.2, and 25 volumes of ethanol and freeezing in dry ice. Analysis by

polyacrylamide gel electrophoresis and phosphorimager analysis afforded the percent yield of the reaction as a function of time.

The Results of the Following Experiments are Described in Appendix B: Attempts to Observe Template-Directed Transesterification with Oligonucleotides Containing Only Deoxyribonucleotides. Oligodeoxyribonucleotides **A**, $\mathbf{A_{t1}}$, and $\mathbf{A_{t2}}$ were 5' labeled as described above. Reaction mixtures contained, in 30 μL total, 1 μL of the labeled solution of **A**, $\mathbf{A_{t1}}$, or $\mathbf{A_{t2}}$, 20 mM imidazole•HCl, pH 7.0; 0.15 μM duplex template, template purine strand, template pyrimidine strand, or no component of the template; 2.6 μM $\mathbf{B_{t1}}$, 2.0 μM $\mathbf{B_{t2}}$, 4.0 μM $\mathbf{B_{t3}}$, 6.6 μM **B**, 3.3 μM **B**, or no **B**; and 1 mM spermine, 20 mM $MgCl_2$, $CaCl_2$, $BaCl_2$, $Sr(NO_3)_2$, $MnCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $ZnCl_2$, $CdCl_2$, $Hg(ClO_4)_2$, K_2PdCl_4, $AuCl_3$, $HfOCl_2$, $La(OAc)_3$, $Ce(NO_3)_3$, $Nd(NO_3)_3$, $Eu(OAc)_3$, $Ho(OAc)_3$, or $Lu(OAc)_3$, or 4 mM $PbCl_2$. The reaction mixtures were warmed to 37 °C for 5 to 9 days after which the DNA was precipitated with ethanol and analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Analysis of the Product of Attempted Transesterification Reaction by Electrophoretic Comparison with Authentic Samples of Possible Products and Reactivity in Acid and Base. The product of the attempted transesterification was purified by electrophoresis in a denaturing 20% polyacrylamide sequencing gel. The band was visualized by autoradiography, excised, crushed, and soaked in 0.2 M NaCl. The isolated product, ethanol precipitated from the 0.2 M NaCl supernatant, was compared electrophoretically to oligodeoxyribonucleotides **CC** and **CT**

chemically synthesized and 5' labeled as described above. The reactivity to acid and base of the linkage formed in the reaction was investigated by dissolving the isolated product (or labeled **CC** or **CT**) in 100 μL of 100 mM HCl/NaCl, pH 2, 2% NaOAc, pH 5.2, water, 10% piperidine, or concentrated NH_4OH and warming the solution to 90 °C for two hours. Solutions containing NH_4OH and piperidine were lyophylized, and the DNA was precipitated from the other solutions with ethanol. The products were then analyzed by denaturing 20% polyacrylamide gel electrophoresis.

Preparation of Oligonucleotides Containing Ribonucleotides for Attempts to Observe Template-Directed Transesterification. Oligonucleotide **ArU** was prepared using a CPG support derivatized with 5'-trityl,2'(3')-acetyl uridine (Glenn Research) on a Beckman System 1 Plus DNA Synthesizer using standard coupling cycles for a 1 μmole synthesis. The oligonucleotide was cleaved from the support, deprotected, and purified by the general procedure outlined above for oligodeoxyribonucleotides. The extinction coefficients for U and TpU[162] were used in the calculation of the extinction coefficient of **ArU**.

Oligonucleotide $\mathbf{rB_{t2}}$ was synthesized coupling 5'-trityl-2'-t-butyldimethylsilyl cytidine 3'-β-cyanoethylphosphoramidite (Penninsula) at the designated position.[163] Synthesis was performed on a 1 μmole scale on an Applied Biosystems 380B Automated DNA Synthesizer. RNA Cycles provided by Applied Biosystems were used for the coupling, capping, and 5' deprotection of the ribocytidine. The oligonucleotide was not cleaved from the support on the machine, but approximately 25% of the oligonucleotide and support was treated for 24 hours at room temperature with methanol

saturated with ammonia. The supernatant was lyophilized and the residue was purified electrophoretically by the method described above, excising only the top half of the least mobile band, eluting into water and desalting with Sephadex NAP-5 columns instead of dialysis. The resulting solution was divided into two portions and lyophilized. The residue from one of the portions was treated at room temperature with 200 microliters of 1 M tetrabutylammonium fluoride (TBAF) in tetrahydrofuran (Aldrich). Most of the white residue dissolved or became finely suspended, though a small amount remained undispersed. After 6 hours, the mixture was centrifuged, and the supernatant was removed and quenched with 300 μL of water.

The resulting solution was desalted using a NAP-5 column. The untreated residue was also dissolved in 500 microliters of water and further desalted with a Sephadex NAP-5 column. Removal of the t-butyldimethylsilyl group by TBAF treatment was confirmed by 5' labeling approximately 50 pmoles of the material which had and had not been treated and comparing their electrophoretic mobilities (denaturing high resolution 20% polyacrylamide gel).

Attempt to Observe Template-Directed Transesterification with Oligonucleotides Containing Ribonucleotides. Oligonucleotide **ArU** (100 pmoles) was 5' labeled with polynucleotide kinase and γ-^{32}P ATP as described above. Reaction mixtures contained, in a total volume of 30 μL, one μL of (^{32}P)**ArU**, 60 mM imidazole•HCl, pH 7.0, 1.9 μM rB_{t2}, 0.15 μM duplex template or 0.28 μM **apR**, and 20 mM $MgCl_2$, $CaCl_2$, $BaCl_2$, $Sr(NO_3)_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $ZnCl_2$, $HfOCl_2$, $Ce(NO_3)_3$, $Nd(NO_3)_3$,

$Eu(OAc)_3$, $Ho(OAc)_3$, or $Lu(OAc)_3$ or 1 mM spermine, or 4 mM $PbCl_2$. The mixtures were warmed at 37 °C in darkness for 18 days before ethanol precipitation and electrophoretic analysis.

Chapter 3

Nonenzymatic Sequence Specific Ligation of Double-Helical DNA

Formation of phosphodiester linkages between the termini of DNA duplexes is a requisite operation in the procedures of molecular cloning. It is typically performed enzymatically using a DNA ligase.[3,164] A nonenzymatic approach to the ligation of double-stranded DNA can be imagined in which a single-stranded template aligns the termini of two DNA duplexes in a continuous triple helix. Juxtaposition of the DNA termini by a guide sequence in a triple-helical complex, accompanied by chemical activation of the terminal phosphates, would be expected to promote ligation of the double-helical DNA (Figure 3.1). Sequence specificity would be imparted on the reaction by the sequence specificity of triple helix formation.

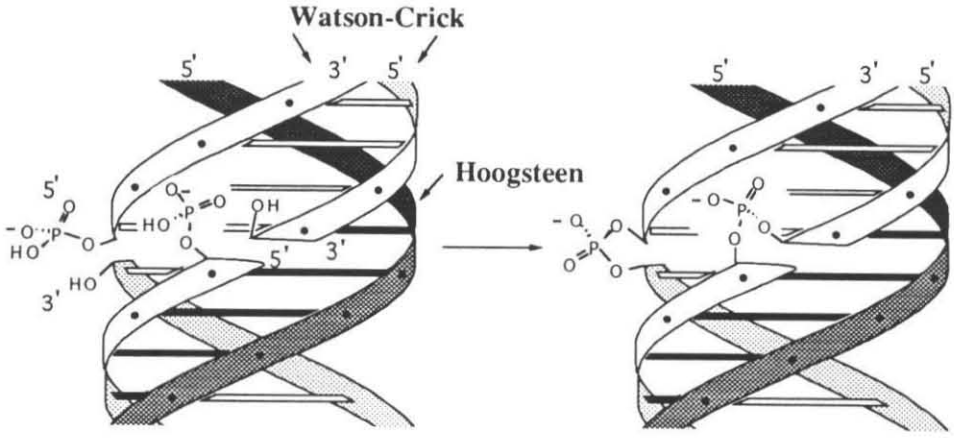


Figure 3.1. Alignment of the 5'-phosphate and 3'-hydroxyl termini of two blunt-ended DNA duplexes by association of an oligonucleotide template in a triple-helical complex.

A plasmid, pDSL, was constructed which, upon cleavage with the restriction endonucleases *Dra* I and *Stu* I, yields a 3.7 kilobase pair (kbp) blunt-ended linear duplex possessing a 15 base pair purine tract for triplex formation at each end (Figure 3.2). 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.

A mixture of double-helical DNA (1.7 nM), 30mer template (17 nM), and $ZnCl_2$ (20 mM) was allowed to react with the condensing agent *N*-cyanoimidazole (1.0 mM). After 7 hours (20°C, pH 4.9), the reaction products resulting from single- and double-stranded ligation reactions were separated by agarose gel electrophoresis in the presence of ethidium bromide (Figure 3.3). 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 DNA 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 in the gel 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 3.3, lane 7), with gel electrophoretic mobilities identical to those produced by treatment of the linearized plasmid with T4 DNA ligase

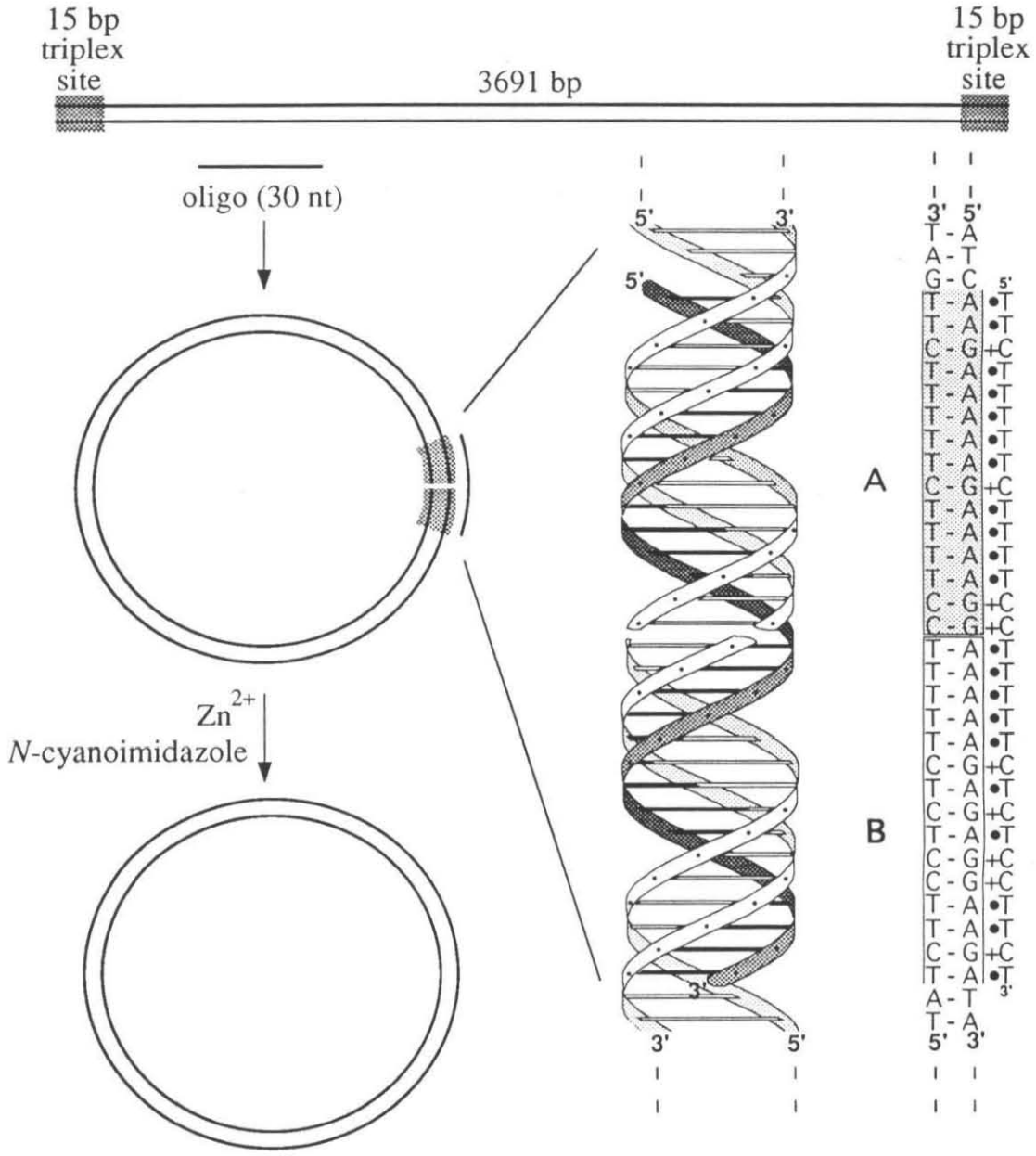


Figure 3.2. Covalent circularization of the linear DNA molecule by association of the template strand with both ends of the double-helical DNA.Upon chemical activation, the proximal 3'-hydroxyl and 5'-phosphate termini would be susceptible to covalent ligation on one or both strands.

and ATP (lane 8). The similarity of products of the non-enzymatic and enzymatic reactions indicates that they are the results of end-to-end ligation. In a separate experiment, the effect of the single-stranded template on the efficiency of enzymatic ligation was investigated. The 30-mer was not found to inhibit or enhance enzymatic ligation.

In controls, no reaction is observed in the absence of the oligodeoxyribonucleotide template (lane 3), in the presence of an oligodeoxyribonucleotide (30mer) complementary in the Watson-Crick sense to the purine tract of the plasmid ends (lane 4), or in the case of the plasmid containing a triple helix forming site at only one terminus (lanes 5 and 6). These requirements for triple helix forming sites at both termini and for the oligodeoxyribonucleotide complementary in the Hoogsteen sense support a model in which the oligodeoxyribonucleotide 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). 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 might 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 work-up.

Approximately 15% of the 3.7 kbp DNA is converted into products with slower gel mobilities than form II DNA (lane 7). In Figure 3.4, the

Figure 3.3. 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 pDSL (0.8 μg) was linearized with *Dra* I and *Stu* I, deproteinized with phenol, desalted by elution through Sephadex G-25, and treated in a total volume of 200 μL with 1 mM *N*-cyanoimidazole, 20 mM $ZnCl_2$, , and 17 nM single-stranded template for 7 hours 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, *N*-cyanoimidazole 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 enzyme digest; 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 hours at 20 °C with 2000 units of T4 DNA ligase 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).

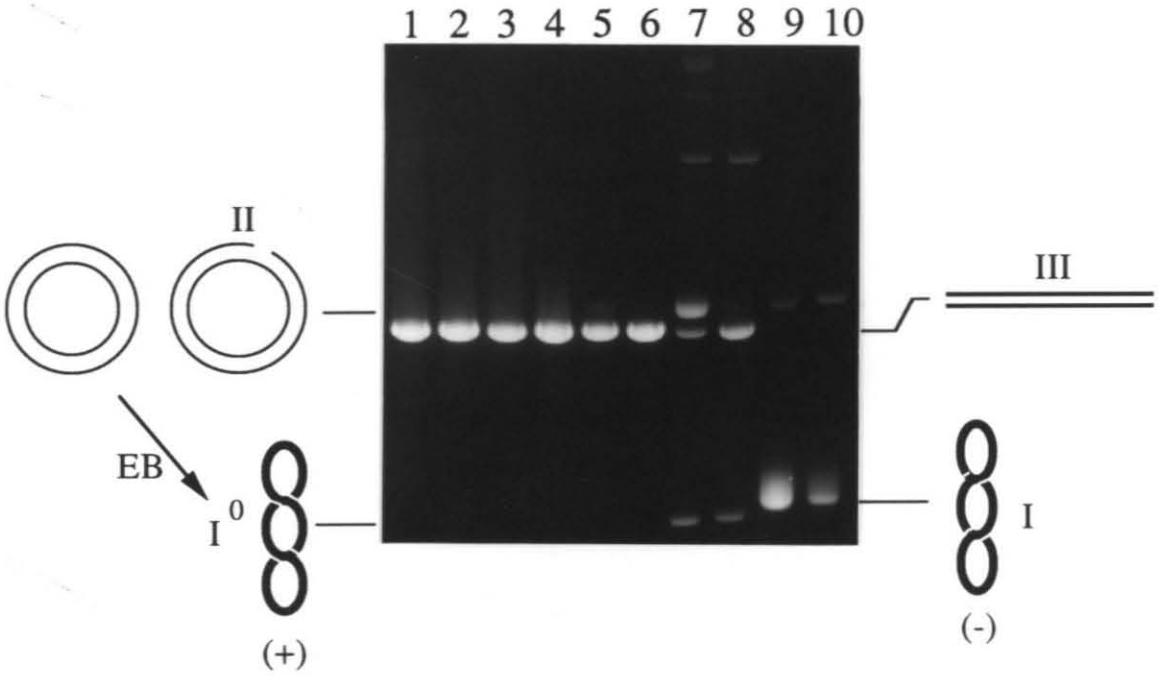
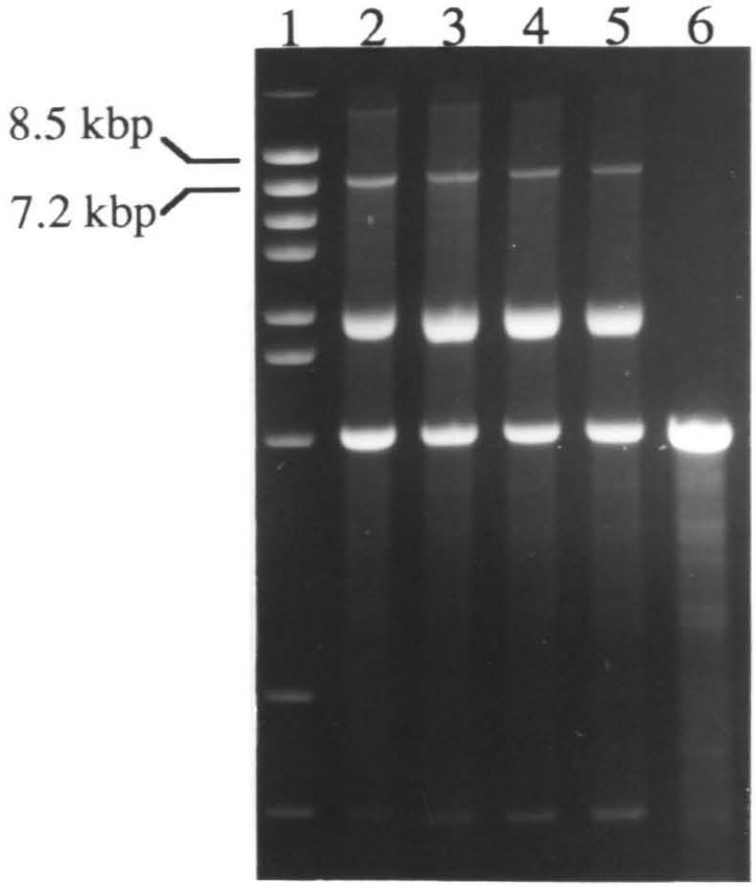

1 2 3 4 5 6 7 8 9 10
II
III
EB
I 0
I
(+)
(-)

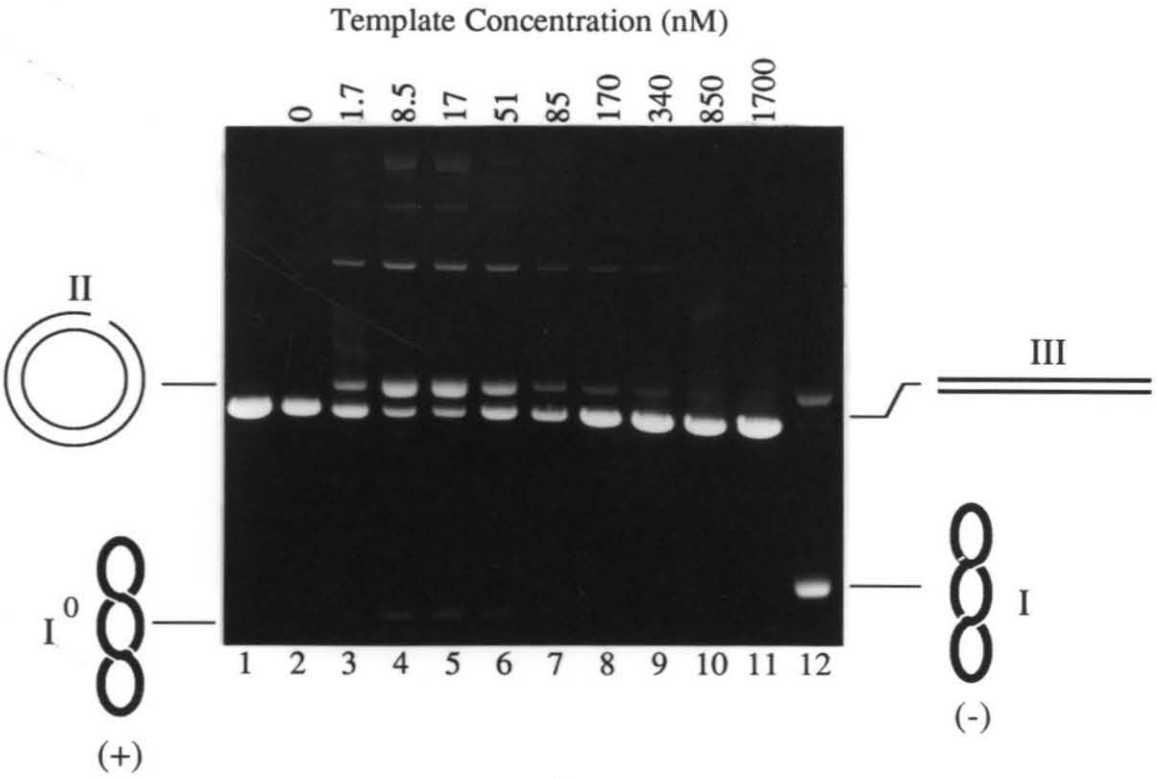
Figure 3.4. Effect of plasmid concentration on relative yields of products; 1.2% agarose gel containing 0.4 mg/L ethidium bromide. Ligation reactions were performed at 24 °C with 17 nM template, 20 mM $ZnCl_2$, 1 mM N-cyanoimidazole, and 0.8 μg pDSL linearized with *Dra* I and *Stu* I in a total reaction volume making the plasmid concentration 8.5 nM (lane 2), 4.3 nM (lane 3), 2.8 nM (lane 4), or 1.7 nM (lane 5). A *Bst*E II digest of lambda DNA was loaded in lane 1, and 0.8 μg of plasmid pDSL, untreated after linearization with *Dra* I and *Stu* I, was loaded in lane 6.


1
2
3
4
5
6
8.5 kbp
7.2 kbp

effect of varying the plasmid concentration in the reaction mixture on the relative yields of the products is shown. The least mobile of the condensation products was not visible on this gel. However, there is a clear increase in the yields of the other two low mobility products as the plasmid concentration is increased from 1.7 nM to 8.5 nM, suggesting that the low mobility products are the result of intermolecular ligations. There is a concomitant decrease in the yield of the doubly ligated circular monomer. By comparison with the mobilities of DNA size standards, the most abundant of these can be identified as a 7.4 kbp linear dimer of the starting material. The product of second greatest abundance on the gel in Figure 3.3 does not have the second slowest migration. It migrates more slowly than the 14.1 kbp size standard and is therefore not an 11.1 kbp trimer of the starting material. This band is assigned to the circular dimer covalently closed on one strand. The remaining product does have a mobility consistent with an 11.1 kbp trimer of the starting material but might also be the circular dimer covalently closed on both strands.

The dependence of the ligation reaction on the template concentration is shown in Figure 3.5. The yield of all products reaches a maximum when the template concentration is 8.5 nM to 17 nM. Above 17 nM, the yield of all products decreases with increasing template concentration. This dependence is seen quantitatively on the plot of total yield of ligation products *vs.* template concentration in Figure 3.6. The decrease in yield at template concentrations above 17 nM can be explained as the result of favoring the association of two template molecules with each plasmid molecule at high concentrations of template. A duplex molecule in which both template binding sites are occupied by association with different template molecules will not be circularized. The optimal template

Figure 3.5. Dependence of nonenzymatic ligation of pDSL on concentration of 30 nucleotide template; 1% agarose gel containing 0.4 mg/L ethidium bromide. Lanes 2-11: plasmid pDSL, linearized with *Dra* I and *Stu* I, was treated (7 hours, 20 °C) at a concentration of 3.4 nM with 20 mM $ZnCl_2$, 20 mM imidazole•HCl, pH 7.0, 40 mM BrCN, and template at the concentrations indicated. Lane 1: Plasmid pDSL, 0.8 μg, untreated after linearization with *Dra* I and *Stu* I. Lane 12: 0.8 μg supercoiled plasmid pDSL (form I).


Template Concentration (nM)
0
1.7
8.5
17
51
85
170
340
850
1700
II
III
I 0
I
(+)
(-)
1 2 3 4 5 6 7 8 9 10 11 12

concentration would be predicted from this analysis to be that at which one half of the template binding sites are occupied by template. Assuming approximately equal affinities of the template for each of the two triple helix forming sites on the plasmid, this concentration approximates the dissociation constant for a complex of the template with one of the triple helix forming sites.

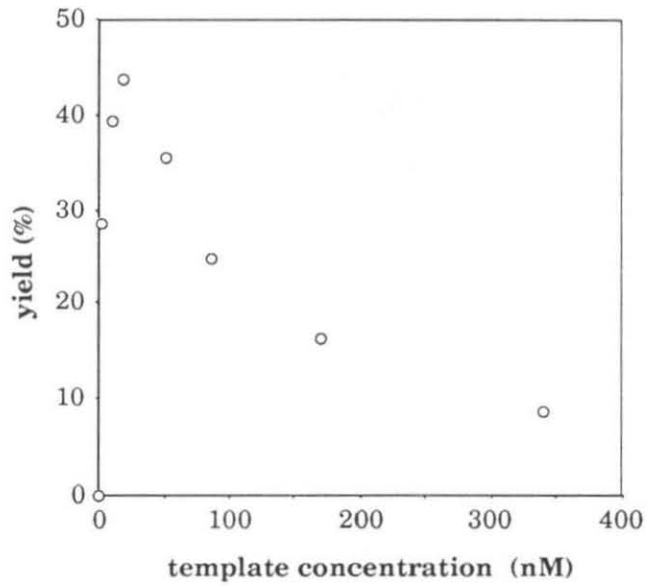


Figure 3.6. Graphical presentation of the total yield of all DNA products as a function of template concentration in the reaction of pDSL with BrCN and imidazole. Yields were determined by densitometric analysis of the gel in Figure 3.5.

No reaction was observed if the template-directed ligation with *N*-cyanoimidazole was performed in the absence of divalent metal ion or with $ZnCl_2$ replaced by 20 mM $MgCl_2$, $CaCl_2$, $BaCl_2$, or $SrCl_2$. In the presence of 20 mM $MnCl_2$ or 20 mM $CoCl_2$, a 30% yield of form II DNA and traces of form I DNA and dimer were produced. When the ligation reaction was performed in the presence of 20 mM $NiCl_2$ or 20 mM $CuCl_2$, the DNA was

precipitated. If the reaction mixtures containing $CuCl_2$ or $NiCl_2$ were treated with EDTA (25 mM final concentration) prior to desalting, the DNA could be analyzed by electrophoresis, but no reaction was evident.

The circular product of the ligation reaction (20 mM $ZnCl_2$, 17 nM template, 1.0 mM *N*-cyanoimidazole, 6.5 hour reaction at 20 °C) could be propagated in *Escherichia coli*. After extensive desalting of the reaction mixture, it was used to transform competent cells by standard methodology, selecting for transformants by their resistance to tetracycline.[164] No transformants were obtained from plasmid that had been cleaved with *Dra* I and *Stu* I and added directly to the transformation mixture without ligation, confirming that unligated plasmid could not be taken up by the bacteria and ligated intracellularly. Ligated plasmid DNA grown in *E. coli* was sequenced, confirming that the product was the expected ligation product. The ability of *E. coli* to propagate the ligation product indicates that at least some of the product contains circular plasmid in which a phosphodiester was formed in the ligation reaction. It does not, however, provide evidence that a phosphodiester was the predominant linkage formed.

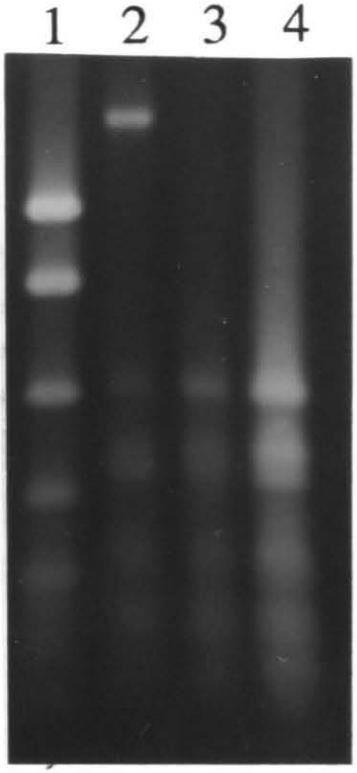
Phosphodiester bond formation in the chemical ligation reaction would afford a cleavage site for the restriction endonuclease *Mnl* I. *Mnl* I hydrolyzes the phosphodiesters on both strands of double-helical DNA seven base pairs to the 3' side of the sequence 5'-CCTC-3'. Therefore, digestion of the products of the nonenzymatic ligation reaction with *Mnl* I will produce a set of restriction fragments consistent with enzymatic cleavage at the ligation site if phosphodiesters are the predominant linkage formed in the reaction.

The pattern of fragments resulting from *Mnl* I digestion of the products of the ligation reaction (60% total yield) is shown in lane 3 of the gel shown in Figure 3.7. As a control, the plasmid prepared by growing the ligation product in *E. coli* was also digested with *Mnl* I, and the cleavage products were analyzed by electrophoresis in lane 4 of the gel shown in Figure 3.7. The pattern of fragments is apparently identical.

If the linkage formed in the ligation reaction were refractory to cleavage by *Mnl* I, a 700 base pair restriction fragment, larger than any of the other fragments of the digestion, would be expected in the *Mnl* I cleavage products. To demonstrate that this fragment could be observed in this assay, a plasmid was constructed that was identical to the plasmid for the initial ligation reaction except for replacement of the *Mnl* I recognition site seven base pairs from the ligation site by the sequence 5'-CTTC-3'. The nonenzymatic ligation reaction was performed with this plasmid, using the corresponding pyrimidine oligodeoxyribonucleotide as a template. A 60% yield of ligation products was obtained. The fragments from *Mnl* I digestion of these products are in lane 2 of Figure 3.7. A fragment with molecular weight larger than the largest fragment in the *Msp* I digest of plasmid pBR 322 (lane 1, 621 base pairs) is clearly visible, demonstrating that the linkages formed in the ligation reaction are stable to the conditions of the *Mnl* I digestion. This fragment is absent from lanes 3 and 4. *Mnl* I certainly cleaved the products of the nonenzymatic ligation reaction at the ligation site. Cleavage by a restriction endonuclease supports the identification of the bonds formed in the reaction as phosphodiesters.

This analysis does not rigorously prove the identity of the predominant linkages formed in the ligation reaction. It is possible that the enzyme *Mnl* I is not absolutely specific for the phosphodiester seven base

Figure 3.7. Analysis of products of nonenzymatic ligation reaction of plasmid pDSL by restriction endonuclease digestion; 1.8% agarose gel containing 0.4 mg/L ethidium bromide. Reactions and digestions were performed as described inthe Materials and Methods section. Lane 1: *Msp* I digest of pBR 322. Lane 2: *Mnl* I digestion of products of nonenzymatic ligation of plasmid pDS2, lacking an *Mnl* I recognition site proximal to the ligation site. Lane 3: *Mnl* I digestion of products of nonenzymatic ligation of plasmid pDSL. Lane 4: *Mnl* I digestion of authentic circular plasmid with the sequence expected for the product of ligation of pDSC.


1
2
3
4

pairs from its recognition sequence. The precise cleavage positions of *Mnl* I and other restriction endonucleases that cleave DNA at sites remote from their recognition sequences have been reported to vary somewhat.[164] Cleavage of the DNA at phosphodiesters near the linkages formed in the ligation reaction would be indistinguishable from cleavage of the new linkages in the low resolution assay employed here. A more convincing analysis of the linkages formed in the ligation reaction could be performed if ligation formed the site of recognition and cleavage of a restriction endonuclease. However, restriction endonucleases which cleave at their recognition sites generally recognize inverted repeat sequences (for example 5'-AGGCCT-3') which cannot be constructed by ligation of homopurine duplexes to form longer homopurine duplexes. Thus, the more general problem of recognizing and ligating duplexes that contain both purines and pyrimidines is encountered.

Nonenzymatic Ligation of Double-Stranded DNA by Alternate-Strand Triple Helix Formation. Formation of a local triple helix within duplex DNA can be extended to sequences of the general type 5'-$(purine)_m(pyrimidine)_n$-3' by association of a pyrimidine oligodeoxyribonucleotide with purines on alternate strands of the double helix. A pyrimidine oligodeoxyribonucleotide constituted of two segments coupled 3' to 3' through a 1,2-dideoxy-D-ribose linker (ϕ) satisfies the structural requirements for binding alternate strands of duplex DNA. The abasic linker connects the two pyrimidine segments across the major groove, bridging two base pairs in a nonspecific manner at the purine-pyrimidine junction of the target site (Figure 1.5).[52]

Application of alternate-strand triple helix formation to template-directed ligation of duplex DNA includes the particular case in which the crossover is effected at the ligation site (Figure 3.8). By alternating strands across the ligation junction, the template aligns two duplexes with purine tracts at their 3' termini, enabling formation of sequences of the type 5'-$(purine)_m(pyrimidine)_n$-3'. A noteworthy feature of a nonenzymatic ligation of this type is that it could form recognition sites for restriction endonucleases.

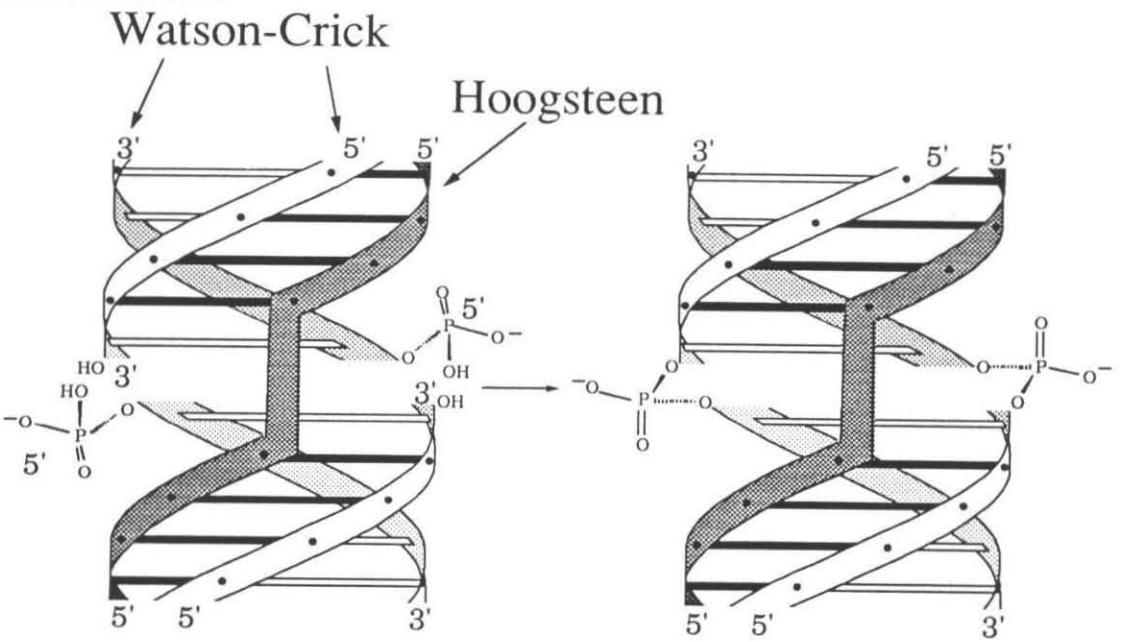


Figure 3.8. Juxtaposition and ligation of blunt-ended DNA duplexes by their concurrent association with a template composed of two oligodeoxyribonucleotide segments coupled 3' to 3' through a 1,2-dideoxy-D-ribose linker. In the triple helix formed, the abasic linker bridges alternate Watson-Crick strands across the site of ligation.

The two base pairs flanking the ligation site are not directly hydrogen bonded to the template. Prior to covalent ligation, the triple helical complex might have greater conformational flexibility around the condensing termini when alternate strands are bound than when a template that

hydrogen bonds continuously to the duplex is employed. Alignment of the duplex termini depends on general alignment of the helical axes of the ligating duplexes. Furthermore, nonspecific interaction of the linker with the ligating duplexes might impart some rigidity to the complex. Stacking of the base pairs at the ligation junction might also decrease the conformational flexibility of the complex in a manner that orients the condensing termini favorably.

To investigate nonenzymatic ligation of double-helical DNA by alternate-strand triple helix formation, a 2.7 kilobase pair (kbp) plasmid, pASL, was constructed which could be cleaved with the restriction endonuclease *Stu* I to afford a blunt-ended linear duplex with a sixteen base pair tract of purines at each 3' terminus. A template molecule was synthesized which contains two fifteen nucleotide segments that are complementary in the Hoogsteen sense to the fifteen base pair purine tract one base pair removed from each terminus of the linear duplex. These segments are coupled 3' to 3' through phosphodiesters to the abasic 1,2-dideoxy-D-ribose linker. Association of this template molecule with both ends of the linear duplex by formation of an alternate-strand triple-helix would circularize the double-stranded DNA. Upon chemical activation, the 3'-hydroxyl and 5'-phosphate termini would be susceptible to covalent ligation on one or both strands if properly oriented with respect to each other in the complex (Figure 3.9).

A mixture of linearized plasmid (2.5 nM), 3'-3' linked template (36 nM), and $ZnCl_2$ (100 mM) was allowed to react with the condensing agent *N*-cyanoimidazole (1 mM). After 25 hours (20 °C, pH 4.9), the reaction products were separated by agarose gel electrophoresis in the presence of

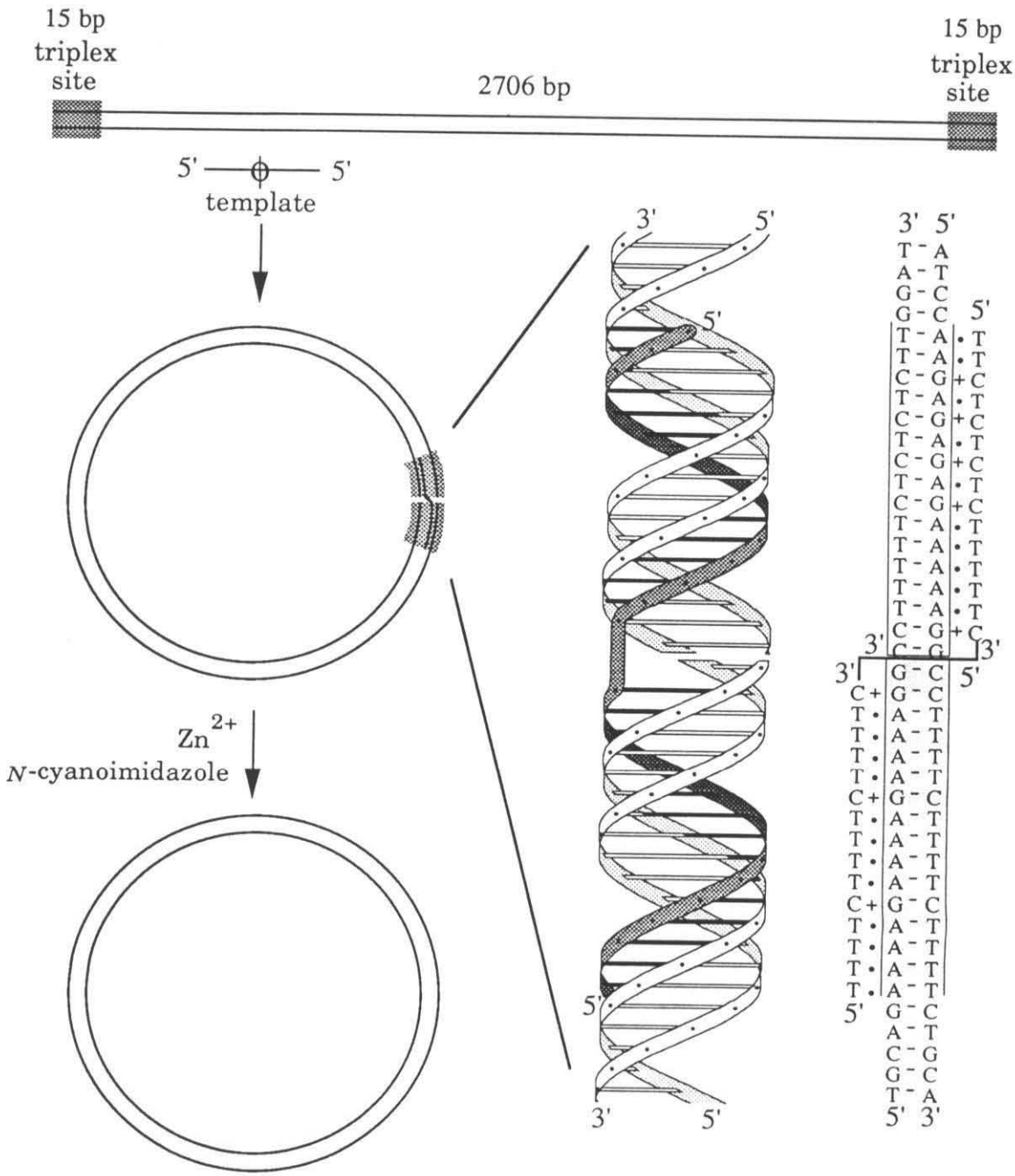


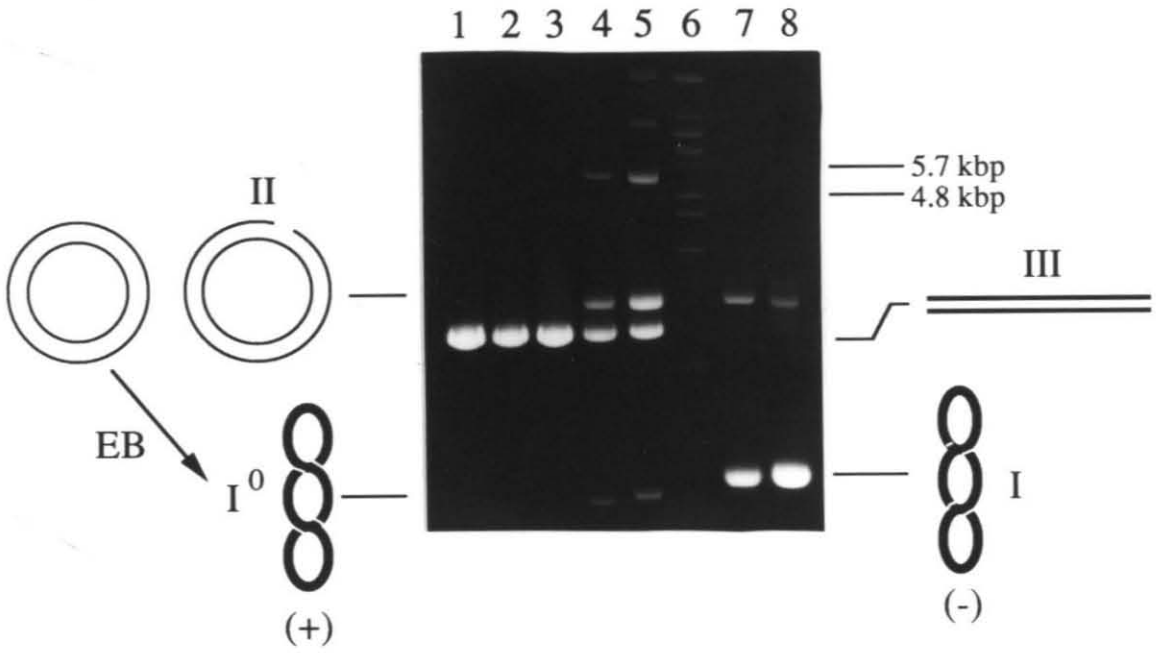
Figure 3.9. Covalent circularization of linear double-helical DNA with purine tracts at its 3' termini. The reaction is directed by a single-stranded template constructed to bind to purine tracts on alternate strands of a Watson-Crick duplex. *N*-cyanoimidazole in the presence of Zn^{2+} promotes condensation.

ethidium bromide (Figure 3.10). Several DNA products from the chemical ligation reaction are observed (Figure 3.10, lane 4), and a set of products with identical mobilities is observed when the linearized plasmid is treated with T4 DNA ligase and ATP (Figure 3.10, lane 5). The similarity of products of 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 template or in the absence of condensing agent, *N*-cyanoimidazole. These requirements support a model in which the template molecule promotes covalent ligation of the duplex termini by apposing them in a three-stranded complex.

A reaction product with electrophoretic mobility corresponding to that of form I°, migrating slightly faster in the gel than the negatively supercoiled plasmid (form I), is present in 7% yield. This product is circularized plasmid in which both strands have been ligated. The predominant reaction product (35% yield) is identified as form II DNA. Form II DNA might arise from two pathways: (i.) circularization followed by ligation of only one of the two strands and (ii.) nicking of the doubly ligated product during the course of the reaction and workup. In a control experiment, approximately 20% of the plasmid is nicked after treatment under the nonenzymatic ligation conditions (Figure 3.10, lane 7). This amount of nicking is insufficient to account for the predominance of form II DNA among the reaction products, and it is concluded that ligation of the two strands of the plasmid DNA is sequential, where the rate of ligation of the second strand is comparable to the rate of ligation of the first strand.

The time course of the reaction reveals that the yield of form II DNA increased to its maximum in the first five hours and subsequently decreased, while the yield of form I° DNA increased continuously within

Figure 3.10. Ligation of linear 2.7 kbp plasmid pASL; analysis by electrophoresis in a 1.2% agarose gel containing 0.4 mg/L ethidium bromide (EB). Lane 4: Plasmid pASL linearized with *Stu* I and treated for 25 hours at 24°C with 36 nM 3'-3' linked template, 1 mM *N*-cyanoimidazole, and 100 mM $ZnCl_2$. In controls, lane 1: pASL linearized with *Stu* I and analyzed without further treatment; lane 2: pASL linearized with *Stu* I and treated as in lane 4, omitting *N*-cyanoimidazole; lane 3: pASL linearized with *Stu* I and treated as in lane 4, omitting 3'-3' linked template; lane 5: pASL linearized with *Stu* I and treated with T4 DNA ligase and ATP; lane 6: *Bst*E II digest of lambda DNA as molecular weight marker; lane 7: un-cut pASL treated as in lane 4; lane 8: un-treated pASL.


1 2 3 4 5 6 7 8
5.7 kbp
4.8 kbp
II
III
EB
I 0
I
(+)
(-)

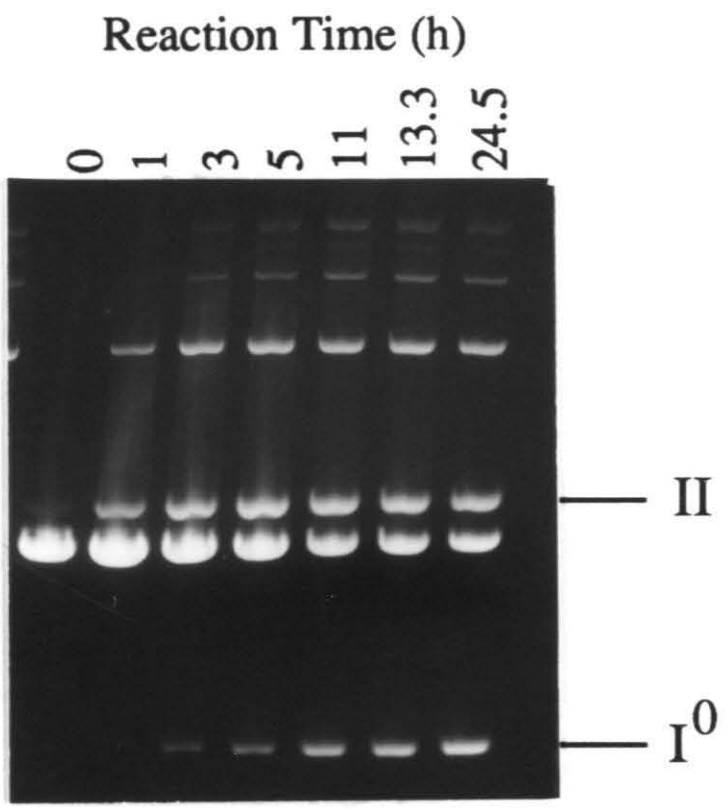
the 24.5 hour course of the reaction (Figure 3.11). Due to the symmetry of the three-stranded complex formed, the two strands of the duplex are geometrically similar prior to ligation, each being directly hydrogen bonded to the template to the 5' side of the ligation junction. The apparent similarity of the rates of ligation of the first and second strands suggests that the structure of the complex is not greatly perturbed by ligation of the first strand.

Products with slower electrophoretic mobilities than the form II DNA are apparent in the nonenzymatic and enzymatic ligation reactions. By comparison with the mobilities of DNA size standards, the most abundant of these can be identified as a 5.4 kbp linear dimer of the starting material. The other products are also assigned as multimers of the starting material formed by intermolecular ligation. Intermolecular ligation is expected to be favored by increasing the concentration of plasmid DNA in the reaction mixture. As anticipated, the extent of intermolecular ligation increases as the concentration of plasmid in the reaction mixture is increased from 2.5 nM to 12.5 nM (Figure 3.12).

The dependence of the reaction on the concentration of the template was examined (Figure 3.13). The optimal template concentration is between 9 nM and 36 nM, affording a total yield of all products of approximately 75% when corrected for the 6% circular plasmid in the starting material. As with the ligation promoted by a template of continuous directionality, the yield of products decreases as the template concentration is raised above an optimal value, interpreted to be that at which one half of the template binding sites are occupied by template. It is noteworthy that this optimal concentration, which approximates the dissociation constant for a complex of the template with one of the triple helix forming sites, is roughly

Figure 3.11. Time course of nonenzymatic ligation reaction at 24°C of pASL, directed by the 3'-3' linked template. (A) Analysis by electrophoresis in a 1.2% agarose gel containing 0.4 mg/L ethidium bromide of the nonenzymatic ligation reaction performed as described in the Materials and Methods section. The reaction was stopped at the times indicated above the lanes on the gel. (B) Yields of form I° (▲) and form II (□) plasmid plotted as a function of reaction time.

A


Reaction Time (h)
0
1
3
5
11
13.3
24.5
II
I0

B

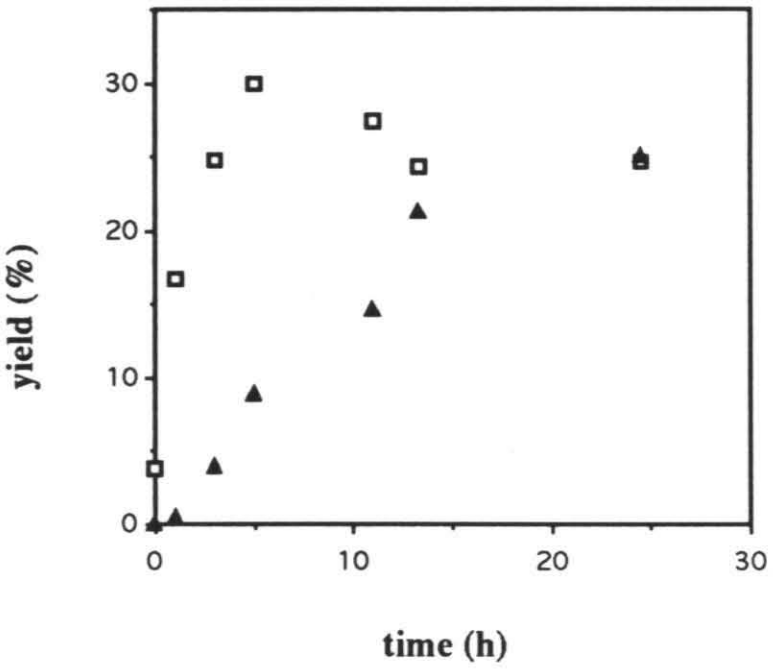

yield (%)
0
10
20
30
0
10
20
30
time (h)

Figure 3.12. Effect of plasmid concentration on relative yields of products; 1.2% agarose gel containing 0.4 mg/L ethidium bromide. Ligation reactions were performed at 24 °C with 36 nM 3'-3' linked template, 100 mM $ZnCl_2$, 1 mM *N*-cyanoimidazole, and 0.9 μg pASL linearized with *Stu* I in a total reaction volume making the plasmid concentration 12.5 nM (lane 2), 6.3 nM (lane 3), 4.2 nM (lane 4), or 2.5 nM (lane 5). A *Bst*E II digest of lambda DNA was loaded in lane 1, and 0.9 μg of plasmid pASL untreated after linearization with *Stu* I was loaded in lane 6.

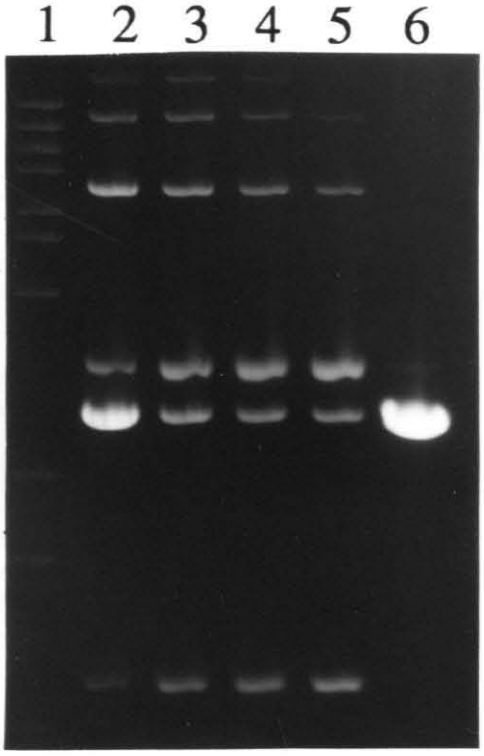
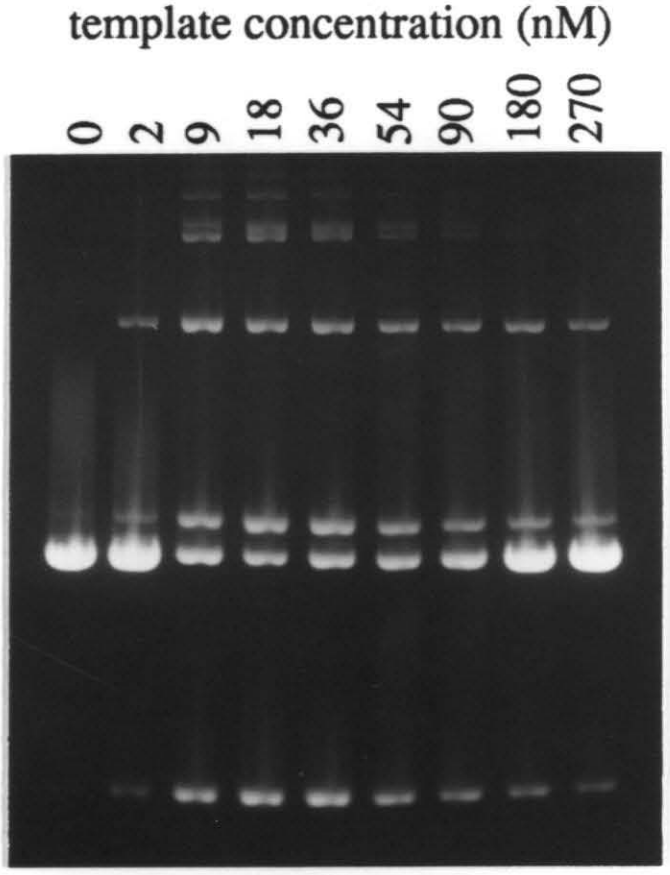
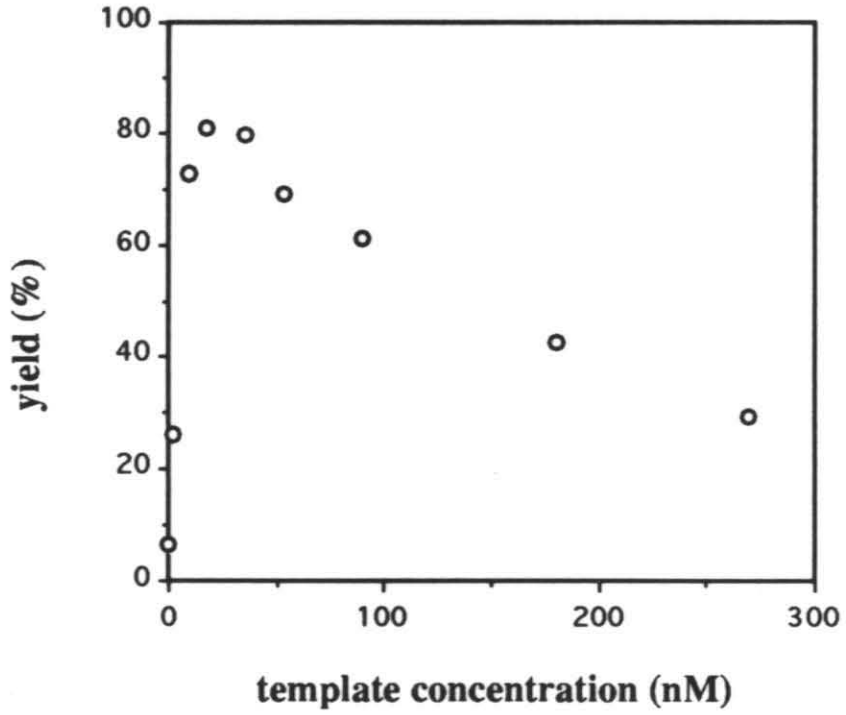

1 2 3 4 5 6

Figure 3.13. Dependence of nonenzymatic ligation reaction on concentration of 3'-3' linked template. Plasmid pASL, linearized with *Stu* I, was treated (24 hours, 24 °C) at a concentration of 2.5 nM with 100 mM $ZnCl_2$, 1 mM *N*-cyanoimidazole, and template at the concentrations indicated. (A) Analysis by electrophoresis in a 1.2% agarose gel containing 0.4 mg/L ethidium bromide. (B) Graphical analysis of the total yield of all DNA products as a function of template concentration.Control experiments indicate that the circular plasmid present (6%) in the absence of template is due to incomplete enzymatic linearization of the plasmid rather than ligation in the absence of template.

A


template concentration (nM)
0
2
9
18
36
54
90
180
270

B


100
80
60
40
20
0
yield (%)
0
100
200
300
template concentration (nM)

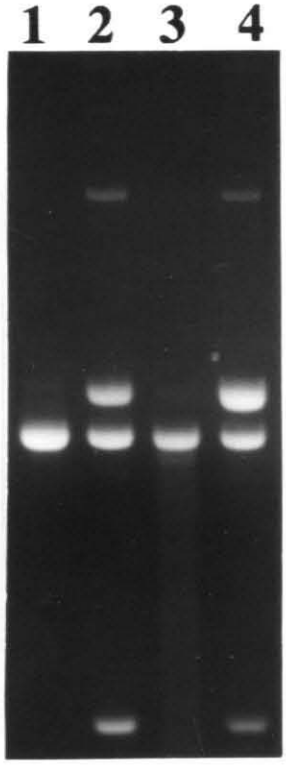
equivalent for the 3'-3' linked template and the template of continuous directionality.

Ligation of the linear plasmid regenerates a *Stu* I site. The products of the chemical ligation reaction (Figure 3.14, lane 2) are quantitatively cleaved to the linear monomer by treatment with *Stu* I (Figure 3.14, lane 3). In a control experiment, the linkages formed in the reaction are stable to the conditions of the enzyme digestion in the absence of *Stu* I (Figure 3.14, lane 4). The susceptibility of the chemical ligation products to cleavage by a restriction endonuclease supports the identification of the bonds formed in the nonenzymatic ligation reaction as phosphodiesters.

Blunt-ended duplexes were chosen as substrates for nonenzymatic ligation in order to avoid complication of the reaction analysis by ligation due to the Watson-Crick base pairing of cohesive ends. If the base pairing at cohesive ends were sufficiently stable to promote condensation of duplex termini, the sequence specificity of the template-directed reaction would be lost. If, however, this base pairing were not sufficient to promote condensation, nonenzymatic ligation by alternate-strand triple helix formation could be applied with sequence specificity to the joining of ends created by any restriction endonuclease having a recognition sequence of the general type 5'-PuPuNNPyPy-3'.

In order to determine the efficacy of base pairing in overhangs of four bases at promoting ligation upon activation with *N*-cyanoimidazole, the 2.7 kbp plasmid pASL was linearized by cleavage at unique sites with *Hin*d III or *Eco*R I and treated with *N*-cyanoimidazole (1.0 mM) and $ZnCl_2$ (100 mM) at 20 °C for 24 hours (Figure 3.15). The plasmid was also linearized with *Stu* I and treated with *N*-cyanoimidazole (1.0 mM) and $ZnCl_2$ (100 mM) at 24 °C

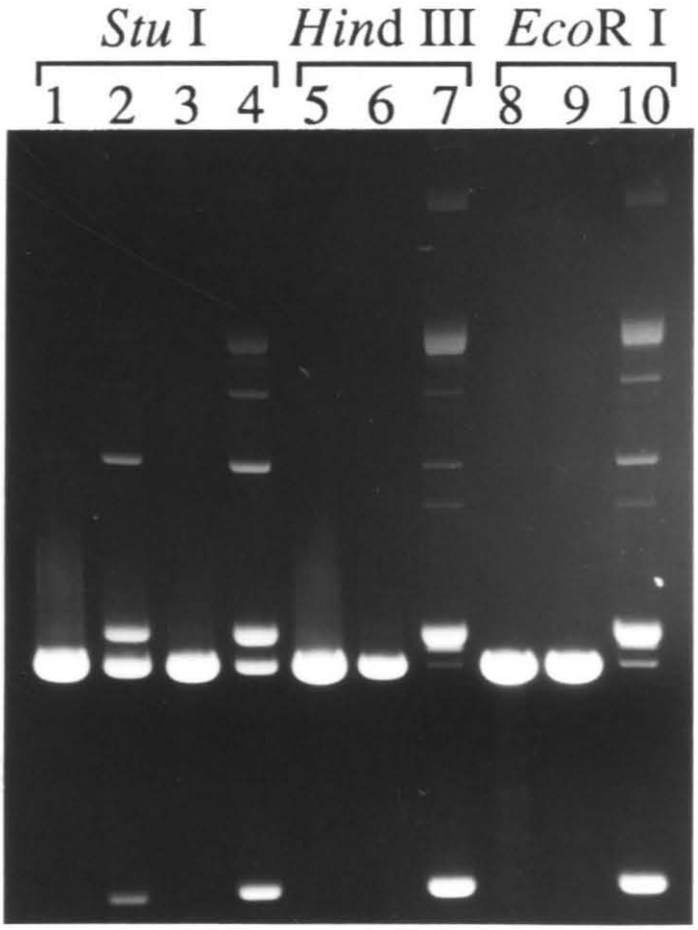
Figure 3.14. Analysis of products of nonenzymatic ligation by restriction endonuclease digestion; 1.2% agarose gel containing 0.4 mg/L ethidium bromide. Reactions and digestions were performed as described in the Materials and Methods section. Lane 1: Plasmid pASL linearized with *Stu* I; lane 2: products of nonenzymatic ligation reaction; lane 3: products of nonenzymatic ligation reaction digested with *Stu* I; lane 4: products of nonenzymatic ligation reaction treated as in lane 3 with the omission of *Stu* I.


1 2 3 4

for 24 hours in the presence and absence of 36 nM 3'-3' linked template. No products were formed when plasmid cleaved with *Stu* I, *Hin*d III, or *Eco*R I was treated with *N*-cyanoimidazole and $ZnCl_2$ in the absence of template, though ligation products were observed as expected when plasmid cleaved with *Stu* I was treated with *N*-cyanoimidazole, $ZnCl_2$, and 3'-3' linked template. Evidently, four-base overhangs of the sequence 5'-AATT-3' (*Eco*R I, recognition sequence 5'-GAATTC-3') and 5'-AGCT-3' (*Hin*d III, recognition sequence 5'-AAGCTT-3') are insufficient to align duplex termini for condensation promoted by *N*-cyanoimidazole. Sequence information from these overhangs would not be expected to interfere with the sequence specificity of nonenzymatic ligation of duplex DNA by triple helix formation; however, duplexes possessing compatible cohesive termini are likely to be acceptable as substrates for nonenzymatic sequence-specific ligation by alternate-strand triple helix formation. It remains to be seen if incompatible overhangs could be induced to ligate, either enzymatically or nonenzymatically, by alternate-strand triple helix formation.

Conclusions. Blunt-ended DNA can be ligated on one and both strands by chemical methods. Using a pyrimidine oligodeoxyribonucleotide as a template, DNA duplexes terminated by purine sequences can be aligned in a continuous triple helix with sequence specificity conferred by Hoogsteen hydrogen bonding. Upon activation with the condensing agent *N*-cyanoimidazole, the juxtaposed phosphate and hydroxyl termini will condense to form phosphodiester linkages. A 3.7 kbp linear plasmid can be covalently circularized by this approach and the circular product can be grown in bacteria.

Figure 3.15. Investigation of the efficacy of complementary four-base overhangs at promoting ligation in the presence of *N*-cyanoimidazole and $ZnCl_2$; analysis by electrophoresis in a 1.2% agarose gel containing 0.4 mg/L ethidium bromide. Lanes 1, 5, and 8: Plasmid pASL (0.9 μg) linearized with *Stu* I, *Hin*d III, and *Eco*R I, respectively, and analyzed without further treatment. Lane 2: Plasmid pASL linearized with *Stu* I and treated with *N*-cyanoimidazole (1.0 mM) and $ZnCl_2$ (100 mM) at 24 °C for 24 hours in the presence of 36 nM 3'-3' linked template. Lanes 3, 6, and 9: Plasmid pASL (0.9 μg) linearized with *Stu* I, *Hin*d III, and *Eco*R I, respectively, followed by treatment for 24 hours at 24 °C with *N*-cyanoimidazole (1.0 mM) and $ZnCl_2$ (100 mM). Lanes 4, 7, and 10: Plasmid pASL (0.9 μg) linearized with *Stu* I, *Hin*d III, and *Eco*R I, respectively, followed by treatment for 24 hours at 24 °C with 60 units of T4 DNA ligase and 2 mM ATP in 50 mM Tris•HCl, pH 7.0, and 20 mM $MgCl_2$.


Stu I
*Hin*d III
*Eco*R I
1 2 3 4 5 6 7 8 9 10

In the three-stranded complex that orients the condensing termini, the two strands that are ligating are structurally distinguishable. One, composed of purines, is directly hydrogen bonded to the continuous template, while the strand composed of pyrimidines is not directly hydrogen bonded to the template. Thus, these two strands might ligate at different rates. The relative rates of ligation of the two strands in this complex might be studied most easily in the template-directed ligation of two DNA stem-loops.

The stem-loops shown in Figure 3.16 and the products of their single- and double-strand ligation could be separated by denaturing polyacrylamide gel electrophoresis. The 5' terminus of stem-loop **A** will be within the condensing pyrimidine strand; whereas, the 5' terminus of stem-loop **B** will be within the condensing purine strand. Thus, the two strands can be specifically radiolabeled by phosphorylation of one of the two hairpins with [γ-^{32}P]ATP and polynucleotide kinase, allowing the ligation of the two strands to be monitored separately. Because the products and reactants of this ligation are small (relative to a 3.7 kbp plasmid), an analysis of the products by high resolution gel electrophoresis could confirm the identity of the linkages formed more rigorously than has been possible in the plasmid.

Alternate-strand triple helix formation can be applied to the template-directed ligation of double-helical DNA when the strand crossover is effected at the site of ligation. The apposition of the duplex termini in the alternate-strand triple helix is sufficient to promote their condensation, despite the absence of direct hydrogen bonding to the template by the two base pairs flanking the ligation site. The application of alternate-strand triple helix formation to template-directed ligation of duplex DNA extends

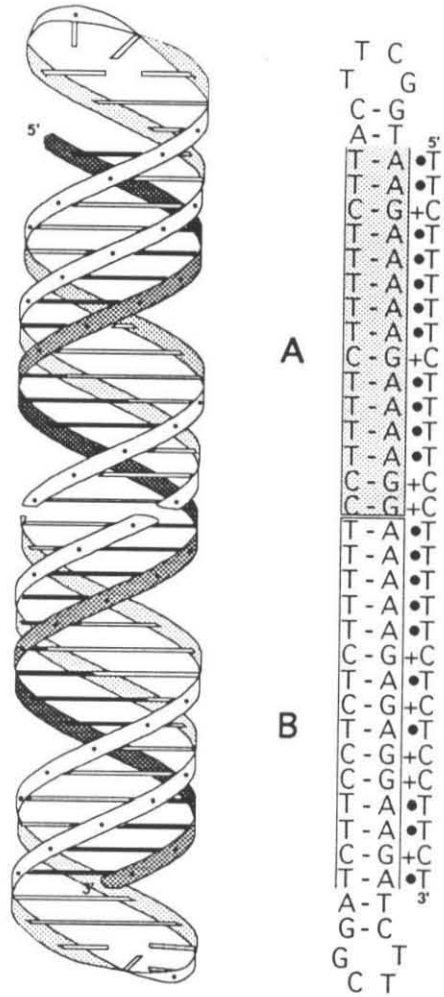


Figure 3.16. Proposed approach to a mechanistic investigation of template-directed nonenzymatic ligation of double-stranded DNA employing a template of continuous directionality. Two unique DNA stem-loops, **A** and **B**, are aligned in a continuous triple-helical complex by association with the 30 nucleotide template strand. Condensation of phosphate and hydroxyl termini would yield linear (single ligation) and circular (double ligation) products which would be electrophoretically separable. Specific radiolabeling of **A** or **B** would render ligation of the purine and pyrimidine strands distinguishable.

the range of sequences that can be created from blunt-ended duplexes by the nonenzymatic approach. Included among the sequences accessible is a set of inverted repeat sequences such as restriction endonuclease recognition sites of the type 5'-PuPuNNPyPy-3'.

Applications of the sequence-specific ligation of double-stranded DNA to problems in cloning or sequencing have yet to emerge. In principle, a double-stranded fragment of DNA could be specifically isolated or manipulated within a diverse mixture of fragments with knowledge of its sequence at one terminus by sequence-specifically ligating it to a "tag" duplex such as a cloning vector. Many of the conceivable applications of this process are already made possible by the polymerase chain reaction. Another application of nonenzymatic template-directed ligation might be the sequence-specific ligation of double-stranded RNA. In that context, the regioselectivity of phosphodiester formation would be of interest, 2' and 3' hydroxyls being available for condensation.

The nonenzymatic ligation of double-stranded DNA by triple helix formation creates sequences of duplex DNA using information from a single-stranded oligonucleotide. Because this template-directed approach involves the transfer of sequence information to the double-helical substrate, the nonenzymatic reaction supecedes in specificity the related enzymatic reaction.

Materials and Methods

Oligodeoxyribonucleotides and *N*-cyanoimidazole were prepared as described in Chapter 2 of this thesis. Sephadex G-50 was obtained from Aldrich and equilibrated with water. Spun columns were prepared by centrifugation of this suspension in the barrels of 1 ml syringes plugged with cotton. Phenol was obtained from BRL and was molecular biology grade. It was equilibrated with water and stored frozen at -20 °C. Protein extractions were performed with reagent grade chloroform and isoamyl alcohol. Restriction endonucleases and T4 DNA ligase were obtained from New England BioLabs and Boehringer Mannheim and used with the buffers provided. T4 polynucleotide kinase was obtained from New England BioLabs. Ethidium bromide was from Sigma and cesium chloride for density gradients was from Fisher Biotech. Density gradient centrifugation was performed using standard methodology[164] with a Beckman L8-70 Ultracentrifuge. L-broth and L-plates for cell culture were prepared by standard methodology.[164]Plasmids were sequenced by the method of Sanger[165] using a kit supplied by Pharmacia.

Agarose gels were run in 40 mM Tris•acetate, pH 8.0, 1 mM EDTA. Nucleic acid grade agarose was obtained from BRL and low melting point agarose was obtained from FMC. For the analysis of ligation reactions, 200 mL of agarose containing 0.4 mg/L ethidium bromide was used in a horizontal slab gel that was 20 cm long and approximately 0.5 cm thick. Agarose gels were visualized by illumination with UV light and photographed with Polaroid type 55 film. Reaction yields were estimated by densitometric analysis of the negative of the photograph using a Bromma LKB Ultroscan-XL densitometer and integrating the peaks of the

densitometer trace by cutting them out and weighing them. 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 the linear and nicked circular DNA.

Construction of Plasmid pDSL and Plasmid Lacking an *Mnl* I Recognition Site at the Ligation Junction, pDS2. 4 nmoles of each oligodeoxyribonucleotide R55 and Y55 or R'55 and Y'55 (no *Mnl* I site) were 5'-phosphorylated with ATP and polynucleotide kinase and annealed by mixing, heating to 90 °C, and subsequent cooling to 4 °C.

R55:
5'-GATCAAGAAAAAGAAAAGGCCTAGAGATCTCTTTTAAAAAGAGAGGAAGATACGT-3'

Y55:
5'-ACGTATCTTCCTCTCTTTTTAAAAGAGATCTCTAGGCCTTTTCTTTTTCTTGATC-3'

R'55:
5'-GATCAAGAAAAAGAAAAGGCCTAGAGATCTCTTTTAAAAAGAGAAGAAGATACGT-3'

Y'55:
5'-ACGTATCTTCTTCTCTTTTTAAAAGAGATCTCTAGGCCTTTTCTTTTTCTTGATC-3'

Plasmid pBR 322 (18 μg) was digested with *Dra* I A total of 120 units of enzyme was used, and the digestion was monitored by electrophoresis in a 1% agarose gel containing ethidium bromide. The 3652 base pair fragment was separated from the other fragments by electrophoresis in a 1% low melting point agarose gel containing ethidium bromide. The band containing this fragment was excised, and the DNA isolated by melting the gel slice at 70 °C in 0.8 mL of 50 mM Tris•HCl, pH 8.0, 1 mM EDTA, extracting the gel with water-saturated phenol, 25:24:1

phenol:chloroform:isoamyl alcohol, and 24:1 chloroform:isoamyl alcohol. The annealed oligodeoxyribonucleotide duplex and the 3652 base pair *Dra* I fragment of pBR322 were ligated using T4 DNA ligase and ATP. Ligation was carried out at room temperature. The insert was in approximately 1000-fold excess over the plasmid fragment. Following ligation, the ligase was removed by phenol extraction, and the DNA was precipitated with ethanol. The precipitate was digested with *Bgl* II, and the larger of the resulting fragments was purified by electrophoresis in 1% low melting point agarose. The purified DNA was treated with ligase and ATP at 4 °C.

Cells of the HB101 strain of *E. coli* were transformed with the resulting construct. The cells were made competent by soaking in 0.1 M $CaCl_2$ at 4 °C overnight. They were transformed by mixing with the ligation products at 4 °C for 15 minutes, warming to 37 °C for 5 minutes, allowing to cool to room temperature and culturing in L-broth. The resulting culture was streaked on L-plates containing 15 μg/L tetracycline. The plates were incubated at 37°C and colonies were visible within 24 hours. Six of the colonies were inoculated into L-broth containing 15 μg/L tetracycline, grown at 37 °C, and sequenced using a primer with the sequence of pBR322 nucleotides 3136 to 3150. Of these, a clone having the desired sequence inserted once was prepared on large scale by density gradient centrifugation for use in double strand ligation experiments.

Template Concentration Dependence of Continuous Directionality Double-Strand Ligation. Plasmid pDSL (4.4 μg) was digested with 480 units *Dra* I and 192 units of *Stu* I in a total volume of 500 μL of the *Dra* I buffer. Digestion was for 4 hours at 37 °C and was followed by extraction twice with 500 μL of water-saturated phenol and desalting with a NAP-5 column. The

1 mL of solution containing the eluted plasmid was concentrated to 120 μL. Ligation reaction mixtures initially contained, in 100 μL total volume, 10 μL of plasmid solution, 20 mM $ZnCl_2$, 20 mM imidazole•HCl, pH 7.0, template oligodeoxyribonucleotide at the concentration indicated in the discussion of results, and 40 mM BrCN. Before addition of BrCN, the other components of the reaction mixtures were equilibrated together at room temperature for 1.5 hours. Reaction was initiated by addition of the BrCN (10 μL of a freshly prepared 0.4 M solution in water) and was allowed to proceed for seven hours at room temperature. After seven hours, the reaction mixtures were diluted to 200 μL and desalted with Sephadex G-50 spun columns. The eluates were frozen and concentrated to 5-10 μL, mixed with 5 μL of 12.5% Ficoll, and analyzed by electrophoresis in a 1% agarose gel containing 0.4 mg/L ethidium bromide.

Continuous Directionality Double-Strand Ligation with *N*-Cyanoimidazole. Plasmid pDSL (8 μg) was digested with *Stu* I (192 units) and *Dra* I (480 units) in the *Dra* I reaction buffer. Digestion was for 4.5 hours at 37 °C and was followed by extraction twice with 500 μL of water saturated phenol and desalting with a NAP-5 column, yielding an 8 μg/mL solution of linearized plasmid. 100 μL of this solution were treated in a total volume of 200 μL with 1 mM *N*-cyanoimidazole, 20 mM $ZnCl_2$, and 17 nM single-strand template for 7 hours at 20 °C. The pH of an equivalently constituted solution was 4.9. The reaction mixture was desalted with a Sephadex G-50 spun column, and the eluate was frozen and concentrated to 5-10 μL, mixed with 5 μL of 12.5% Ficoll, and analyzed by electrophoresis in a 1% agarose gel containing 0.4 mg/L ethidium bromide.

In controls, single-stranded template or *N*-cyanoimidazole was omitted, or the single stranded template was replaced by the 30 nucleotide Watson-Crick complement of the plasmid purine strand. Controls were also performed in which *Dra* I, *Stu* I, or both were omitted from the enzyme digest. As an electrophoretic standard for the products of end-to-end ligation, plasmid cut with *Dra* I and *Stu* I was treated for 7 hours at 20 °C with 2000 units of T4 DNA ligase in 200 μL of a solution containing 20 mM $MgCl_2$, 50 mM Tris•HCl, pH 7.0, and 15 mM ATP.

Enzymatic Ligation of Linearized pDSL in the Presence of Single-Strand Template. Plasmid pDSL (0.4 micrograms), digested with *Dra* I and *Stu* I, was treated with T4 DNA ligase (400 units, New England BioLabs) in a 200 μL reaction mixture containing 20 mM $MgCl_2$, 50 mM Tris•HCl, pH 7.0, and 2.0 mM ATP. In different reactions, 17 nM single-strand template was included or omitted. Work-up and electrophoretic analysis were by the standard procedures after 7.5 hours at room temperature.

Transformation of Competent Cells with the Products of Template-Directed Double-Strand Ligation; Preparation of Plasmid pDSC. Plasmid pDSL (0.4 μg), digested with *Dra* I and *Stu* I, was treated with 40 mM BrCN, 20 mM $ZnCl_2$, 20 mM imidazole•HCl, pH7.0, and 8.5 nM single strand template in a total volume of 200 μL for 7 hours at room temperature. In controls, template or BrCN were omitted. In another control, 0.4 μg of linearized plasmid were ligated enzymatically with 800 units of T4 DNA ligase in a 200 μL reaction containing 20 mM $MgCl_2$, 50 mM Tris•HCl, pH 7.0, and 2 mM ATP. Enzymatic ligation was performed at room temperature for 6.5 hours then extracted with phenol, chloroform and n-butanol. All reaction

mixtures were desalted with Sephadex G-50 spun columns and concentrated to 5 μL under vacuum. HB101 cells made, competent by $CaCl_2$ treatment, were treated with the concentrates in the standard transformation protocol, and 150 μL of transformation suspension were streaked on L-plates containing 15 μg/L tetracycline. Colonies were cultured and sequenced according to standard methods. The chemical ligation product, pDSC, was grown and prepared by standard methods.

Double-Strand Ligation Product Analysis by Enzymatic Digestion. Plasmids pDSL and pDS2 (8 μg of each) were digested with *Dra* I and *Stu* I. The linearized plasmids were treated with 1 mM *N*-cyanoimidazole, 20 mM $ZnCl_2$, and 17 nM template in a total volume of 2 mL for 7 hours at room temperature. The single strand template used for each plasmid was of the appropriate Hoogsteen complementary sequence. The reaction mixtures were desalted with Sephadex G-50 spun columns in 200 μL aliquots, and the aliquots for each reaction were pooled. 180 μL of each were concentrated under vacuum and analyzed by agarose gel electrophoresis to confirm that ligation had occurred to a similar extent with each plasmid. The remainder of each solution was centrifugally concentrated (Centricon-30) to 40 μL. The concentrates were twice diluted with 1.5 mL 300 mM Tris•HCl, pH 9.0, and concentrated. The resulting solutions were diluted several times with 40-200 microliters 10X *Mnl* I buffer and concentrated again. Finally, these solutions were twice diluted with 2 ml water and concentrated to 30 μL.

The concentrated reaction products and 8 μg of pDSC were digested with *Mnl* I. Digestions were at 37 °C in *Mnl* I buffer with 0.1 mg/mL BSA. They were started with 7.5 units of enzyme and 7.5 additional

units of enzyme were added after 24 and 48 hours. After a total digestion time of 88 hours (68 hours for pDSC), 75 μL of each solution were desalted with a spun column, concentrated under vacuum, and analyzed on a 1.8% agarose gel. 2.5 micrograms of an *Msp* I digest of pBR322 were also included on the gel as a molecular weight standard. After electrophoresis, the gel was stained for 0.5 hours with a solution of 0.8 μg/mL ethidium bromide in 1X TAE buffer.

Effect of Plasmid Concentration on Product Distribution in Continuous Strand Duplex Ligation. Plasmid pDSL (0.8 μg), cleaved with *Dra* I and *Stu* I, was treated in the indicated total reaction volume with 1 mM *N*-cyanoimidazole, 20 mM $ZnCl_2$, and 17 nM single-strand template for 10 hours at 20 °C. The reaction mixtures were finally diluted to 200 μL, desalted with Sephadex spun columns, concentrated, and analyzed by electrophoresis in a 1.2% agarose gel.

Construction of Plasmid pASL for Alternate-Strand Duplex Ligation Experiments. Plasmid pUC18 (1.86 μg, BRL) was linearized with restriction endonuclease *Pst* I (40 units in a total volume of 20 μL, 5.5 hours, 37 °C) and further cut with *Bam*H I (140 units in a total volume of 100 μL containing 0.1 mg/mL BSA, 10 hours, 37 °C). The reaction mixture was extracted with phenol, and the DNA was precipitated with ethanol.

Oligodeoxyribonucleotides with the sequences shown below were synthesized with an ABI 380B DNA synthesizer, deprotected, and purified by electrophoresis in a denaturing 15% polyacrylamide gel.

5'-GATCCAAGAGAGAGAAAAAGGCCTTTTCTTTTCTTTTCTGCA-3'

5'-GAAAAGAAAAGAAAAGGCCTTTTTCTCTCTCTTG-3'

130 pmoles of each were mixed, and the mixture was dried under vacuum, dissolved in 10 μL of water, heated to 90 °C for 10 minutes, and allowed to cool to room temperature. The annealed insert and the *Pst* I/*Bam*H I cut pUC18 were ligated in a reaction mixture containing 50 mM Tris•HCl, pH 7.0, 20 mM $MgCl_2$, 3.0 mM ATP, and 200 units of T4 DNA ligase in a total volume of 50 μL. Ligation was at 4 °C for 7 hours. Subcloning efficiency competent DH5α cells (BRL) were transformed by treating them with half of this mixture in a standard transformation protocol[164]. The resulting mixture was streaked on L-plates covered with 50 μL of a solution of 20 mg/mL Bluo-Gal (BRL) in DMF and containing 50 μg/L ampicillin. 10 white colonies were cultured and sequenced, and all contained the desired insert. One of these cultures was used for large scale preparation of pASL.

Large scale preparation was by density gradient centrifugation or with a Qiagen maxi prep kit using the manufacturer's recommended procedure. The effect on ligation efficiency of additional purification by FPLC after plasmid preparation using the Qiagen kit was investigated. A MonoQ 5/5 column was employed. Buffer A contained 0.75 M NaCl and buffer B contained 0.85 M NaCl, both in 10 mM Tris•HCl, pH 8.0, 1 mM EDTA. Approximately 40 μg of DNA were injected in each separation. Best separation of nicked and supercoiled plasmid was achieved with a flow rate of 0.5 mL/min, eluting for 10 minutes with 100% buffer A, linearly increasing the concentration of buffer B to 50% over 5 minutes, linearly increasing the concentration of buffer B to 88% over the next 21 minutes,

maintaining 88% buffer B for 9 minutes, then increasing the concentration of buffer B to 100% over 10 minutes. The concentration of buffer B was maintained at 100% for 10 minutes before returning to 100% buffer A over 10 minutes. The collected peak eluted between 88% and 100% B with this protocol. The DNA in the collected fractions was precipitated with ethanol, and the ethanol precipitates were dissolved in 500 microliters of water and further desalted with a Sephadex NAP-5 column. Ligation yields were significantly lower when reactions were performed with plasmid prepared by methods other than density gradient centrifugation.

Preparation of Alternate-Strand Binding Template. The template for the alternate strand ligation experiment was prepared by methods previously described.[52] 3'-trityl-5'-B-cyanoethyl phosphoramidites of thymidine and cytosine were synthesized by the method of van de Sande.[148] 5'-trityl protected abasic linker[166,167] was provided by Thomas Povsic and the β-cyanoethyl phosphoramidite was prepared by standard procedures.[160] 3'-Trityl thymidine coupled through the 5' position to controlled pore glass was provided by Dr. David Horne. The synthesis was performed on an ABI 380B automated DNA synthesizer, coupling the first 15 nucleotides in the 5' to 3' direction. A 3',3' linkage was formed to the abasic linker, and the remaining 15 nucleotides were coupled in the conventional 3' to 5' direction. From spectrophotometric quantitation of trityl responses, the average coupling yield for the coupling steps in the 5' to 3' direction and the coupling of the abasic linker was 86%.

The template was deprotected and purified by electrophoresis in a denaturing 20% polyacrylamide gel as described earlier. The product migrated near the xylene cyanol as was expected for the desired

oligodeoxyribonucleotide. It was purified by the standard procedure, desalting the solution by gel filtration on NAP-5 columns rather than dialysis, and the concentration of the solution was determined by its absorbance at 260 nm using an extinction coefficient for the template of 2.5×10^5 $M^{-1}cm^{-1}$.

Ligation Reactions with Alternate-Strand Binding Template. Plasmid pASL (5 pmoles) was linearized with restriction endonuclease *Stu* I obtained from New England Biolabs using the provided buffer. The linearized plasmid was deproteinized by extraction three times with water-saturated phenol and three times with 25:24:1 phenol:chloroform:isoamyl alcohol, then desalted by gel filtration (NAP-5 column). Reactions were performed at 24 °C, pH 4.9 with the template, plasmid, and $ZnCl_2$ concentrations indicated in the discussion of results. They were initiated by addition of a freshly prepared 10 mM solution of *N*-cyanoimidazole to a final concentration of 1 mM. Reactions were stopped by centrifugation through approximately 0.7 mL of Sephadex G-50 and freezing.

Prior to analysis by agarose gel electrophoresis, the solutions were concentrated under vacuum to approximately 10 μL, and 5 μL of a 12.5% solution of Ficoll containing bromophenol blue and xylene cyanol were added. Reactions were analyzed by electrophoresis in 1.2% agarose gels containing 0.4 mg/L ethidium bromide, 40 mM Tris•acetate, pH 8.0, and 1 mM EDTA. Approximately 0.9 μg of DNA were loaded in each lane. 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 the linear and nicked circular DNA. The *Bst*E II digest of lambda DNA,

used as an electrophoretic size marker, was obtained from New England BioLabs.

To generate enzymatic ligation products from pASL linearized with *Stu* I, 0.9 μg of plasmid cut with *Stu* I was treated for 25 hours at 24 °C with 60 units of T4 DNA ligase (Boehringer Mannheim) in 200 μL of a solution containing 20 mM $MgCl_2$, 50 mM Tris•HCl, pH 7.0, and 2.0 mM ATP. The reaction mixture was treated in the same manner as the nonenzymatic ligation reactions.

The time course of the nonenzymatic ligation reaction was studied by performing the reaction with 7.2 μg of pASL cut with *Stu* I in a total volume of 1.6 mL containing 100 mM $ZnCl_2$, 36 nM template, and 1 mM *N*-cyanoimidazole. At the times indicated after addition of *N*-cyanoimidazole, 200 μL aliquots were withdrawn, desalted by centrifugal gel filtration through Sephadex G-50, and frozen in dry ice. The frozen solutions were concentrated and analyzed by agarose gel electrophoresis as described above.

To confirm that the linkages formed in the nonenzymatic ligation reaction are susceptible to cleavage by *Stu* I, 9.0 μg of pASL linearized with *Stu* I were treated with 36 nM template, 100 mM $ZnCl_2$, and 1 mM *N*-cyanoimidazole in a total volume of 2 mL for 24 hours at 24 °C. Gel filtration through Sephadex G-25 (Pharmacia) was employed to remove $ZnCl_2$ from the solution. One third of the product mixture was treated with 320 units of *Stu* I in a total volume of 500 μL for 24 hours at 37°C. In a control, one third of the product mixture was incubated in 500 μL of *Stu* I reaction buffer for 24 hours at 37 °C in the absence of enzyme. The digest and control were desalted by gel filtration (Sephadex G-25), concentrated under vacuum, and

analyzed by agarose gel electrophoresis with the undigested reaction products and the linear starting material as described above.

References

(1) Orgel, L. E. In *Cold Spring Harbor Symposia on Quantitative Biology;* Cold Spring Harbor: 1987; Vol. LII; pp. 9-16.

(2) Joyce, G. F. In *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harbor, 1987; Vol. LII; pp. 41-51.

(3) Watson, J. D.; Hopkins, N. H.; Roberts, J. W.; Steitz, J. A.; Weiner, A. M. *Molecular Biology of the Gene;* 4th ed.; Benjamin/Cummings: Menlo Park, 1987.

(4) Struhl, K. *Trends Biochem. Res.* **1989**, *14*, 137-140.

(5) Brennan, R. G.; Matthews, B. W. *Trends Biochem. Res.* **1989**, *14*, 286-290.

(6) Jacob, F.; Monod, J. *J. Mol. Biol.* **1961**, *3*, 318-356.

(7) Bourgeois, S.; Cohn, M.; Orgel, L. E. *J. Mol. Biol.* **1965**, *14*, 300.

(8) Gallant, J.; Stapleton, R. *J. Mol. Biol.* **1964**, *8*, 442.

(9) Garen, A.; Otsuji, N. *J. Mol. Biol.* **1964**, *8*, 841.

(10) Sadler, J. R.; Novick, A. *J. Mol. Biol.* **1965**, *12*, 305.

(11) Miller, J. H.; Sobell, H. M. *Proc. Natl. Acad Sci. U. S. A.* **1966**, *55*, 1201-1205.

(12) Felsenfeld, G.; Miles, H. T. *Annu. Rev. Biochem.* **1967**, *36*, 407-448.

(13) Hattori, M.; Frazier, J.; Miles, H. T. *Biopolymers* **1976**, *15*, 523.

(14) Michelson, A. M.; Massoulie, J.; Guschlbauer, W. *Prog. Nucl. Acids Res. Mol. Biol.* **1967**, *6*, 83-141.

(15) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.

(16) Hoogsteen, K. *Acta Cryst.* **1963**, *16*, 907.

(17) Hoogsteen, K. *Acta Cryst.* **1959**, *12*, 822-823.

(18) Mathews, F. S.; Rich, A. *J. Mol. Biol.* **1959**, *8*, 89.

(19) Mathews, F. S.; Rich, A. *J. Mol. Biol.* **1964**, *8*, 89-95.

(20) Lipsett, M. N. *Biochem. Biophys. Res. Comm.* **1963**, *11*, 224-228.

(21) Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256-1260.

(22) Howard, R. B.; Frazier, J.; Lipsett, M. N.; Miles, H. T. *Biochem. Biophys. Res. Commun.* **1964**, *17*, 93.

(23) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.

(24) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073-3091.

(25) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073-3091.

(26) Lee, J. S.; Woodsworth, M. L.; Latimer, L. J. P.; Morgan, A. R. *Nucleic Acids Res.* **1984**, *12*, 6603-6614.

(27) Moser, H. E.; Dervan, P. B. *Science (Washington, D. C.)* **1987**, *238*, 645-650.

(28) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968-972.

(29) Hertzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934.

(30) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457.

(31) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565.

(32) de los Santos, C.; Rosen, M.; Patel, D. *Biochemistry* **1989**, *28*, 7282.

(33) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859.

(34) Sklenar, V.; Feigon, J. *Nature* **1990**, *345*, 836.

(35) Arnott, S.; Bond, P. J. *Nature, New Biol.* **1973**, *244*, 99-101.

(36) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*, 509-521.

(37) O'Connor, T.; Bina, M. *J. Biomol. Struct. Dyn.* **1984**, *2*, 615-625.

(38) Macaya, F. R.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1992**, *114*, 781-783.

(39) Haner, R.; Dervan, P. B. *Biochemistry* **1990**, *29*, 9761.

(40) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

(41) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927-7929.

(42) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73.

(43) Letai, A. G.; Palladino, M. A.; Fromm, E.; Rizzo, V.; Fresco, J. R. *Biochemistry* **1988**, *27*, 9108-9112.

(44) Plum, G. E.; Park, Y. W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9436-9440.

(45) Mergny, J.-L.; J.-S., S.; Rougee, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791-9798.

(46) Singleton, S. F.; Dervan, P. B., *In Press.*

(47) Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 7286.

(48) Distefano, M. D.; Shin, J. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 5901-5902.

(49) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-971.

(50) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *Biochemistry* **1992**, *31*, 2830-2834.

(51) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *In press*

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

(53) Beal, P. A.; Dervan, P. B. *Science* **1991**, 1360-1363.

(54) Boles, T. C.; Hogan, M. E. *Biochemistry* **1987**, *26*, 367-376.

(55) Broitman, S. L.; Im, D. D.; Fresco, J. R. *Proc. Natl. Acad. Sci. U.S.A* **1987**, *84*, 5120-5124.

(56) Chen, F.-M. *Biochemistry* **1991**, *30*, 4472-4479.

(57) Cooney, M.; Czernuszewica, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456-459.

(58) Kohwi, Y.; Kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3781-3785.

(59) Jayasena, S. D.; Johnston, B. H. *Biochemistry* **1992**, *31*, 320-327.

(60) Beasty, A. M.; Behe, M. J. *Nucleic Acids Res.* **1988**, *16*, 1517-1528.

(61) Behe, M. J. *Biochemistry* **1987**, *26*, 7870-7875.

(62) Hopkins, R. C. *Comments Mol. Cell. Biophys.* **1984**, *2*, 153.

(63) Minton, K. W. *J. Exp. Path.* **1985**, *2*, 135.

(64) Morgan, A. R. *Trends Biochem. Sci.* **1979**, *4*, N244.

(65) Young, L. S.; Dunstan, H. M.; Witte, P. R.; Smith, T. P.; Ottonello, S.; Sprague, K. U. *Science* **1991**, *252*, 542-546.

(66) Voloshin, O. N.; Mirkin, S. M.; Lyamichev, V. I.; Belotserkovskii, B. P.; Frank-Kamenetskii, M. D. *Nature* **1988**, *333*, 475-476.

(67) Wells, R. D.; Collier, D. A.; Hanvey, J. C.; Shimizu, M.; Wohlrab, F. *FASEB J.* **1988**, *2*, 2939-2949.

(68) Nickol, J. M.; Felsenfeld, G. *Cell* **1982**, *35*, 467-477.

(69) Mirkin, S. M.; Lyamichev, V. I.; Drushlyak, K. N.; Dobrynin, V. N.; Filippov, S. A.; Frank-Kamenetskii, M. D. *Nature* **1987**, *330*, 495-497.

(70) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomol. Struct. Dyn.* **1986**, *3*, 667-670.

(71) Johnston, B. H. *Science* **1988**, *241*, 1800-1804.

(72) Htun, H.; Dahlberg, J. E. *Science* **1988**, *241*, 1791-1796.

(73) Hentchel, C. C. *Nature* **1982**, *295*, 714-716.

(74) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6292-6296.

(75) Glover, J. N. M.; Pulleyblank, D. E. *J. Mol. Biol.* **1990**, *215*, 653-663.

(76) Christophe, D.; Cabrer, B.; Bacolla, A.; Targovnik, H.; Pohl, V.; Vassart, G. *Nucleic Acids Res.* **1985**, *13*, 5127.

(77) Lee, J. S.; Burkholder, G. D.; Latimer, L. J. P.; Haug, B. L.; Braun, R. P. *Nucleic Acids Res.* **1987**, *15*, 1047.

(78) Kiyama, R.; Camerini-Otero, R. D. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 10450-10454.

(79) Hsieh, P.; Camerini-Otero, C. S.; Camerini-Otero, R. D. *Genes Dev.* **1990**, *4*, 1951-1963.

(80) Povsic, T. J. Thesis, California Institute of Technology, 1992.

(81) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Nucleic Acids Res.* **1989**, *18*, 157-161.

(82) Lyamichev, V. I.; Frank-Kamenetskii, M. D.; Soyfer, V. N. *Nature* **1990**, *344*, 568-570.

(83) Lyamichev, V. I.; Voloshin, O. N.; Frank-Kamenetskii, M. D.; Soyfer, V. N. *Nucleic Acids Res.* **1991**, *19*, 1633-1638.

(84) Francois, J.-C.; Saison-Behmoaras, T.; Thuong, N. T.; Helene, C. *Biochemistry* **1989**, *28*, 9617-9619.

(85) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Nucleic Acids Res.* **1990**, *18*, 157-161.

(86) Maher, L. J.; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-730.

(87) Koob, M.; Szybalski, W. *Science* **1990**, *250*, 271.

(88) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172.

(89) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639-1642.

(90) Francois, J.-C.; Saison-Behmoaras, T.; Chassignol, M.; Thuong, N. T.; Helene, C. *J. Biol. Chem.* **1989**, *264*, 5891-5898.

(91) Pei, D.; Corey, D. R.; Schultz, P. G. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9858-9862.

(92) Praseuth, D.; Perroault, L.; Doan, T. L.; Chassignol, M.; Thuong, N.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349-1353.

(93) Perrouault, L.; Asseline, U.; Rivalle, C.; Thuong, N. T.; Bisagni, E.; Giovannangeli, C.; Le Doan, T.; Helene, C. *Nature* **1990**, *344*, 358-360.

(94) Dervan, P. B. In *Nucleic Acids and Molecular Biology*; F. Eckstein and D. M. J. Lilley, Ed.; Springer-Verlag: Heidelberg, 1988; Vol. 2; pp 49-64.

(95) Dervan, P. B. *Science* **1986**, *232*, 464-471.

(96) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9428-9430.

(97) Fareed, G. C.; Richardson, C. C. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 665.

(98) Little, J. W.; Zimmerman, S. B.; Oshinsky, C. K.; Gellert, M. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 2004.

(99) Naylor, R.; Gilham, P. T. *Biochemistry* **1966**, *5*, 2722-2728.

(100) Khorana, H. G. *Chem. Rev.* **1953**, *53*, 145-166.

(101) Williams, A.; Ibrahim, I. T. *Chem. Rev.* **1981**, *81*, 589-636.

(102) Dolinnaya, N. G.; Sokolova, N. I.; Gryaznova, O. I.; Shabarova, Z. A. *Nucleic Acids Res.* **1988**, *16*, 3721-3738.

(103) Orgel, L. E.; Lohrmann, R. *Acc. Chem. Res.* **1974**, *7*, 368-377.

(104) Sulston, J.; Lohrmann, R.; Orgel, L. E.; Miles, H. T. *Proc. Natl. Acad. Sci. USA* **1968**, *59*, 726-733.

(105) Sulston, J.; Lohrmann, R.; Orgel, L. E.; Schneider-Bernloehr, H.; Weimann, B. J.; Miles, H. T. *J. Mol. Biol.* **1969**, *40*, 227.

(106) Tohidi, M.; Zielinski, W. S.; Chen, C. B.; Orgel, L. E. *J. Mol. Evol.* **1987**, *25*, 97-99.

(107) Weimann, B. J.; Lohrmann, R.; Orgel, L. E.; Schneider-Benloehr, H.; Sulston, J. E. *Science* **1968**, *161*, 387.

(108) van Roode, J. H. G.; Orgel, L. E. *J. Mol. Biol.* **1980**, *144*, 579-585.

(109) Zielinski, W. S.; Orgel, L. E. *Nucleic Acids Res.* **1985**, *13*, 2469-2484.

(110) Zielinski, W. S.; Orgel, L. E. *Nucleic Acids Res.* **1987**, *4*, 1699-1715.

(111) Sleeper, H. L.; Lohrmann, R.; Orgel, L. E. *J. Mol. Evol.* **1979**, *13*, 203.

(112) Sawai, H.; Orgel, L. E. *J. Am. Chem. Soc.* **1975**, *97*, 3532-3533.

(113) Lohrmann, R.; Orgel, L. E. *J. Mol. Biol.* **1980**, *142*, 555-567.

(114) Lohrmann, R.; Bridson, P. K.; Orgel, L. E. *Science* **1980**, *208*, 1464-1465.

(115) Lohrmann, R.; Orgel, L. E. *Tetrahedron* **1978**, *34*, 853-855.

(116) Inoue, T.; Orgel, L. E. *Science* **1983**, *219*, 859-862.

(117) Inoue, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1981**, *103*, 7666-7667.

(118) Hill, A. R., Jr.; Kumar, S.; Leonard, N. J.; Orgel, L. E. *J. Mol. Evol.* **1988**, *27*, 91-95.

(119) Chen, C. B.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1985**, *181*, 271-279.

(120) Bridson, P. K.; Orgel, L. E. *J. Mol. Biol.* **1980**, *144*, 567-577.

(121) Sawai, H. *J. Am. Chem. Soc.* **1976**, *98*, 7037-7039.

(122) Leonard, N. J.; Cruickshank, K. A.; Groziak, M. P.; Clauson, G. L.; Devadas, B. *Ann. N.Y. Acad. Sci.* **1986**, *471*, 255-265.

(123) Ferris, J. P.; Edelson, E. H.; Mount, N. M.; Sullivan, A. E. *J. Mol. Evol.* **1979**, *13*, 317-330.

(124) Ferris, J. P.; Hagan, W. J., Jr.; Alwis, K. W.; McCrea, J. *J. Mol. Evol.* **1982**, *18*, 304-309.

(125) Ferris, J. P.; Yanagawa, H.; Dudgeon, P. A.; Hagan, W. J., Jr.; Mallare, T. R. *Origins Life* **1984**, *15*, 29.

(126) Ferris, J. P.; Yanagawa, H. *J. Org. Chem.* **1984**, *49*, 2121-2125.

(127) Ibanez, J. D.; Kimball, A. P.; Oro, J. *Science* **1971**, *173*, 444-446.

(128) Shabarova, Z. A.; Merenkova, I. N.; Oretskaya, T. S.; Sokolova, N. I.; Skripkin, E. A.; Alexeyeva, E. V.; Balakin, A. G.; Bogdanov, A. A. *Nucleic Acids Res.* **1991**, *19*, 4247-4251.

(129) Kanaya, E.; Yanagawa, H. *Biochemistry* **1986**, *25*, 7423-7430.

(130) Sokolova, N. I.; Ashirbekova, D. T.; Dolinnaya, N. G.; Shabarova, Z. A. *FEBS Lett.* **1988**, *232*, 153-155.

(131) Hoard, D. E.; Ott, D. G. *J. Am. Chem. Soc.* **1965**, *87*, 1785.

(132) Ferris, J. P.; Huang, C.-H.; Hagan, W. J., Jr. *Nucleosides Nucleotides* **1989**, *8*, 407-414.

(133) Zielinski, W. S.; Orgel, L. E. *Nature* **1987**, *327*, 346-347.

(134) Luebke, K. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 8733-8735.

(135) Luebke, K. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 7447-7448.

(136) Maxam, A. M.; Gilbert, W. *Methods Enzymology* **1980**, *65*, 497-559.

(137) Kochetkov, N. K.; Budovskii, E. I. In *Organic Chemistry of Nucleic Acids*Plenum: New York, 1972; Vol. Part B; pp 412-416.

(138) Chu, C. F.; Orgel, L. E. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 963.

(139) Raae, A. J.; Kleppe, K. *J. Am. Chem. Soc.* **1978**, *17*, 2939-2942.

(140) Pattabiraman, N. *Biopolymers* **1986**, *25*, 1603-1606.

(141) Il'ychova, I. A.; Lysov, Y. P.; Chernyi, A. A.; Shchyolkina, A. K.; Gottikh, B. P.; Florentiev, V. A. *J. Biomol. Struct. Dyn.* **1990**, *7*, 879-897.

(142) Klysik, J.; Rippe, K.; Jovin, T. M. *Biochemistry* **1990**, *29*, 9831-9839.

(143) Rich, A.; Davies, D. R.; Crick, F. H. C.; Watson, J. D. *J. Mol. Biol.* **1961**, *3*, 71-86.

(144) Ramsing, N. B.; Rippe, K.; Jovin, T. M. *Biochemistry* **1989**, *28*, 9528-9535.

(145) Rippe, K.; Jovin, T. M. *Biochemistry* **1989**, *28*, 9542-9549.

(146) Rippe, K.; Ramsing, N. B.; Jovin, T. M. *Biochemistry* **1989**, *28*, 9536-9541.

(147) Shchyolkina, A. K.; Lysov, Y. P.; Il'ichova, I. A.; Chernyi, A. A.; Golova, Y. B.; Chernov, B. K.; Gottikh, B. P.; Florentiev, V. L. *FEBS Lett.* **1989**, *244*, 39-42.

(148) van de Sande, J. H.; Ramsing, N. B.; Germann, M. W.; Elhorst, W.; Kalisch, B. W.; Kitzing, v. E.; Pon, R. T.; Clegg, R. C.; Jovin, T. M. *Science (Washington D. C.)* **1988**, *241*, 551.

(149) Otto, C.; Thomas, G. A.; Rippe, K.; Jovin, T. M.; Peticolas, W. L. *Biochemistry* **1991**, *30*, 3062-3069.

(150) Ishikawa, F.; Frazier, J.; Howard, F. B.; Miles, H. T. *J. Mol. Biol.* **1972**, *70*, 475-490.

(151) Maher, L. J.; Dervan, P. B.; Wold, B. J. *Biochemistry* **1990**, *29*, 8820-8826.

(152) Student *Biometrika* **1908**, *6*, 1.

(153) Day, R. A., Jr.; Underwood, A. L. *Quantitative Analysis;* 5th ed.; Prentice-Hall: Englewood Cliffs, 1986.

(154) Macaya, R. F.; Gilbert, D. E.; Malek, S.; Sinsheimer, J. S.; Feigon, J. *Science* **1991**, *254*, 270-274.

(155) Landegren, U.; Kaiser, R.; Sanders, J.; Hood, L. *Science* **1988**, *241*, 1077-1080.

(156) Barany, F. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 189-193.

(157) Dolinnaya, N. G.; Pyatrauskene, O. V.; Shabarova, Z. A. *FEBS Lett.* **1991**, *284*, 232-234.

(158) Cech, T. R.; Bass, B. L. *Ann. Rev. Biochem.* **1986**, *55*, 599-629.

(159) Cech, T. R. *Science* **1987**, *236*, 1532-1539.

(160) Atkinson, T.; Smith, M. In *Oligonucleotide Synthesis; a Practical aApproach*; M. J. Gait, Ed.; IRL Press: Oxford, 1984; pp 35-81.

(161) Warshaw, D.; Tinoco, I. *J. Mol. Biol.* **1966**, *20*, 29.

(162) Cantor, C. R.; Warshaw, D. *Biopolymers* **1970**, *9*, 1059.

(163) Giesemann, H. *J. Prakt. Chem.* **1955**, *1*, 345-348.

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

(165) Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5463-5467.

(166) Eritja, R.; Walker, P. A.; Randall, S. K.; Goodman, M. F.; Kaplan, B. E. *Nucleosides Nucleotides* **1987**, *6*, 803-814.

(167) Takeshita, M.; Chang, C.-N.; Johnson, F.; Will, S.; Grollman, A. P. *J. Biol. Chem.* **1987**, *262*, 10171.

(168) Cech, T. R. *Cell* **1986**, *44*, 207.

(169) Liu, L. *CRC Crit. Rev. Biochem.* **1983**, *15*, 1-9.

(170) Westheimer, F. H. *Science* **1987**, *235*, 1173-1178.

(171) Usher, D. A.; Richardson, D. I.; Oakenfull, D. G. *J. Am. Chem. Soc.* **1970**, *92*, 4699-4710.

(172) Dugas, H.; Penney, C. *Bioorganic Chemistry*; Springer-Verlag: New York, 1981.

(173) Ogilvie, K.; Usman, N. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5764-5768.

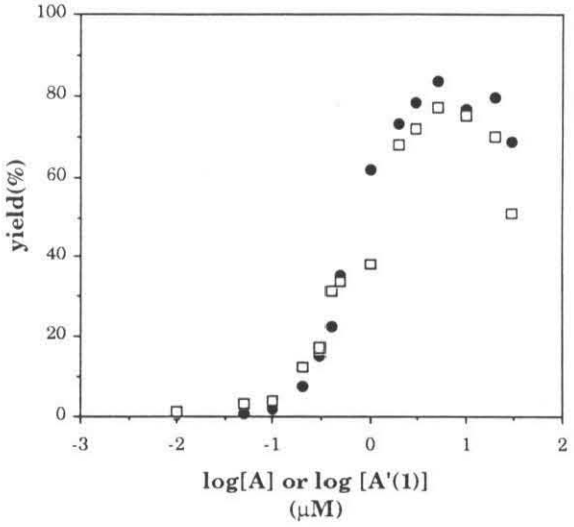
Appendix A

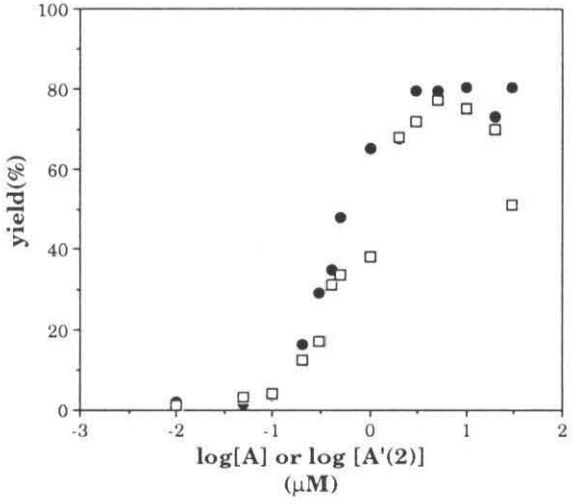
The following are plots of the yield of the template-directed condensation of (^{32}P)**B** with **A** (□) and **A'(1)-A'(9)** (●) vs. the logarithm of the micromolar concentration of **A** or **A'(1)-A'(9)**. The sequences of the double-stranded template and oligodeoxyribonucleotides **B**, **A**, and **A'(1)-A'(9)** are as described in Chapter 2 of this thesis. For ease of reference, the sequences of **A'(1)-A'(9)** are shown below with the mismatched base shown in bold:

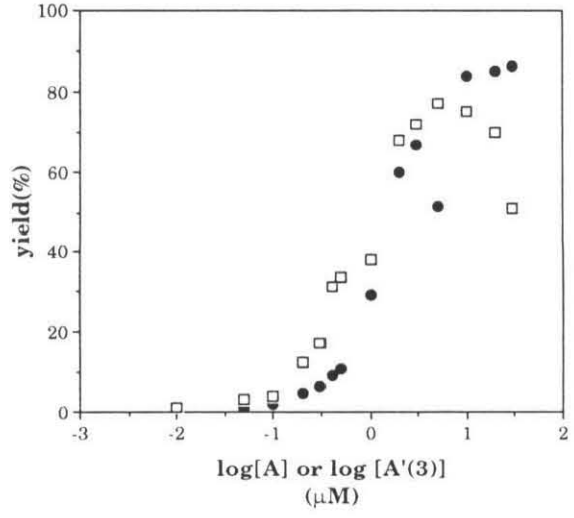
A'(1): 5'-TT TTC TC**T**C T CTC TT-3'
A'(2): 5'-TT TTC TC**TA**T CTC TT-3'
A'(3): 5'-TT TTC TC**TG**T CTC TT-3'
A'(4): 5'-TT TTC TC TTT CTC **C**T-3'
A'(5): 5'-TT TTC TC TTT CTC **A**T-3'
A'(6): 5'-TT TTC TC TTT CTC **G**T-3'
A'(7): 5'-TT TTC TC TTT CTC T**C**-3'
A'(8): 5'-TT TTC TC TTT CTC T**A**-3'
A'(9): 5'-TT TTC TC TTT CTC T**G**-3'

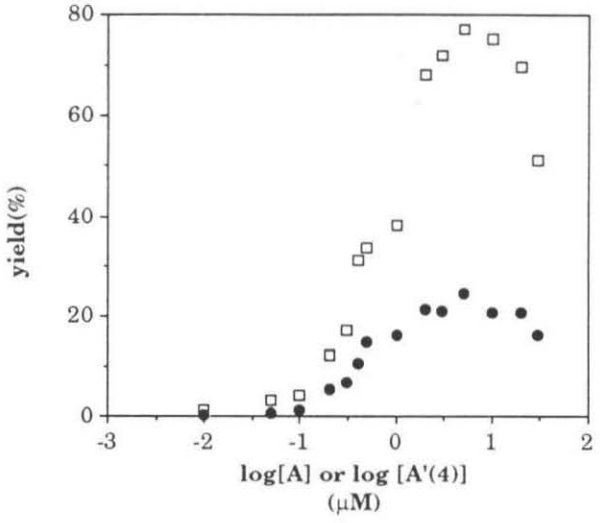
The yield of the ligation reaction after 6-8 hours at 24 °C was measured using varying concentrations (0.01-30.0 μM) of oligodeoxyribonucleotide **A** and each of the mismatched oligodeoxyribonucleotides and constant concentrations of duplex template (0.15 μM), condensing agent (1 mM *N*-cyanoimidazole), and $NiCl_2$ (10 mM, pH 6.9). Reactions were initiated by addition of *N*-cyanoimidazole and stopped by ethanol precipitation. All reactions with a given oligodeoxyribonucleotide **A** or **A'** were stopped within a period of five minutes of each other. The products were analyzed electrophoretically, and yields were quantitated by phosphor storage analysis as the percent of total (^{32}P)**B** incorporated into 30mer product.

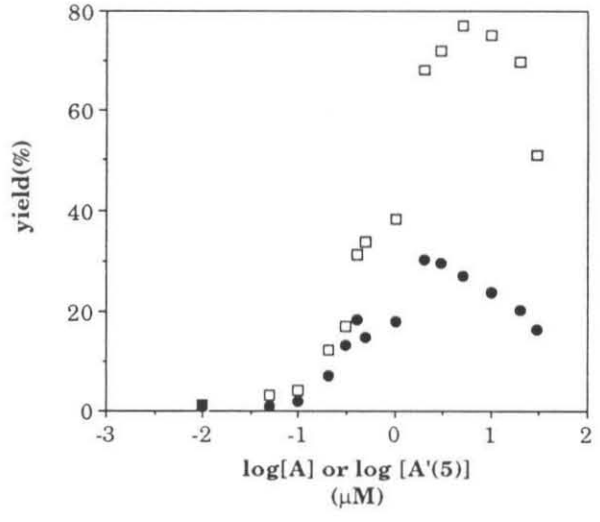
For comparison, the mean yield of two trials of the condensation reaction with **A** (no mismatch) is plotted *vs.* the logarithm of the micromolar concentration of **A** in each of the following plots.

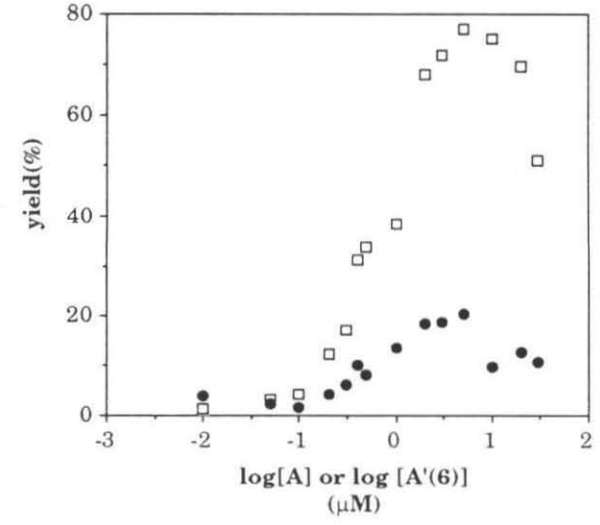

100
80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(1)]
(μM)

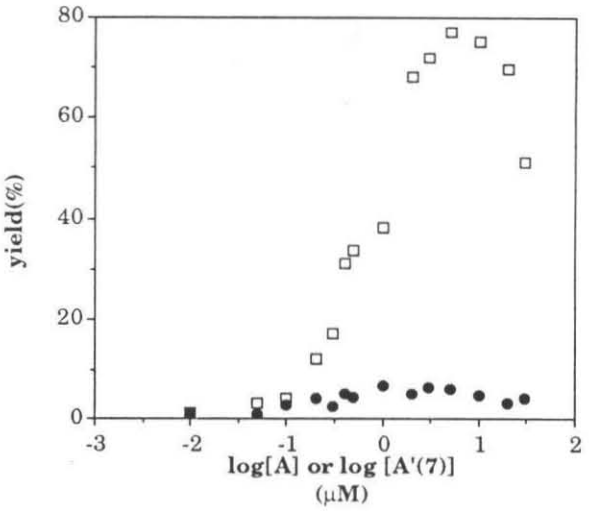

100
80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(2)]
(μM)

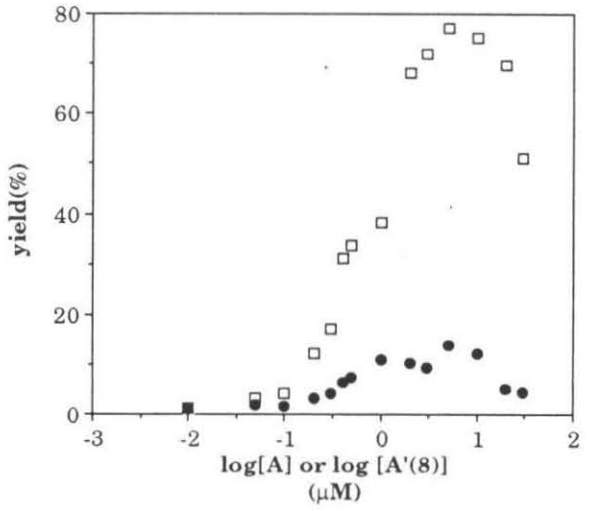

100
80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(3)]
(μM)

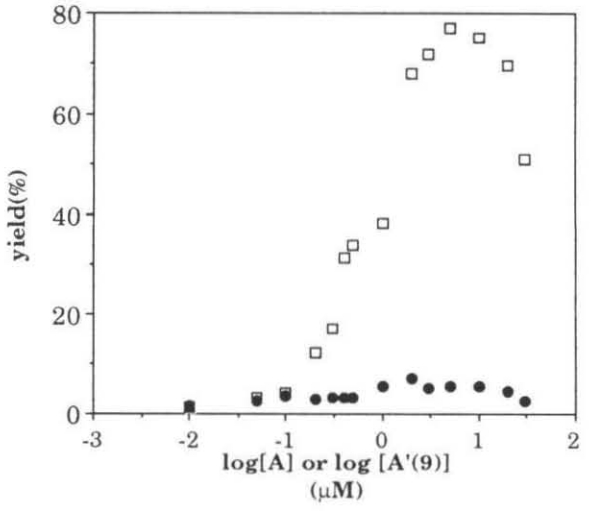

80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(4)]
(μM)


80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(5)]
(μM)


80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(6)]
(μM)


80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(7)]
(μM)


80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(8)]
(μM)


80
60
40
20
0
yield(%)
-3
-2
-1
0
1
2
log[A] or log [A'(9)]
(μM)

Appendix B

Attempts to Observe Transesterification of Phosphodiesters Directed by Triple Helix Formation.

In the experiments described in Chapters 2 and 3, phosphodiester formation was accomplished by nucleophilic displacement of imidazole, *O*-phosphoryl urea, or *O*-phosphoryl *N*-carboxamidoimidazole, highly labile ligands. A terminal hydroxyl of an oligonucleotide could also be imagined to react with a proximal phosphodiester in a triple helical complex to displace another oligodeoxyribonucleotide, effecting a transesterification. Transesterification reactions of phosphodiesters occur in a number of biological strand transfer processes such as RNA splicing[158,168] and DNA topoisomerization.[169] Phosphodiesters are notoriously resistant to nucleophilic displacement,[170] and the duplex-directed transesterification reaction imagined was expected to be slow. Nonetheless, two considerations allowed hope that such a reaction might be observable.

First, electrophoretic analysis of radiolabeled oligonucleotides provided a highly sensitive assay. Because transesterification could produce labeled oligonucleotides longer than the starting material, interference due to oxidative degradation of labeled starting material could be avoided. Transesterification products would migrate more slowly in a polyacrylamide gel than starting material; whereas, degradation products would all migrate more rapidly than starting material. Thus, long reaction times could be employed, and background interference with detection would be low. A yield of 0.1% was considered to be observable.

Second, the relative lability of RNA to hydrolysis due to anchimeric assistance from the adjacent 2'-hydroxyl[171,172] suggested that proper orientation of a hydroxyl with respect to a phosphodiester could promote an observable reaction. Comparison of the efficiency observed by Naylor and Gilham for the cabodiimide-promoted cyclization of nucleotide 2'(3') phosphates (quantitative conversion in less than one hour at 25 °C)[99] with the efficiency of EDC-promoted ligation directed by the duplex template described here (20% yield in 20 hours at 37°C in the presence of spermine), suggested that the neighboring hydroxyl and phosphate termini in the triplex are not as well disposed for reaction as the 2' hydroxyl and 3' phosphate of a ribonucleotide. However, the efficiency of template-directed reactions is known to be sensitive to a number of factors that alter the orientations of the reacting groups, and it was conceivable that a particular combination of oligodeoxyribonucleotide leaving group and metal ion could provide the requisite orientation for observable reaction. Furthermore, it was thought that catalysis by metal ions could assist the reaction.

The oligodeoxyribonucleotides shown in Figure B.1 were synthesized (See Materials and Methods, Chapter 2). Aligned by triple helix formation as shown, a terminal 3' or 5' hydroxyl would be proximal to a phosphodiester linking a di- or tetranucleotide to the sequence of **B** or **A**. Oligonucleotides targeted to the binding site for **A** were labeled by phosphorylation at their 5' termini with polynucleotide kinase and [γ-^{32}P]ATP. The labeled oligodeoxyribonucleotides were mixed with the corresponding oligodeoxyribonucleotide (1-3 μM) targeted to the binding site for **B**, the double-stranded template described in Chapter 2 (0.15 μM), 20 mM imidazole•HCl, pH 7.0, and spermine (1 mM), $PbCl_2$ (4 mM), $MgCl_2$,

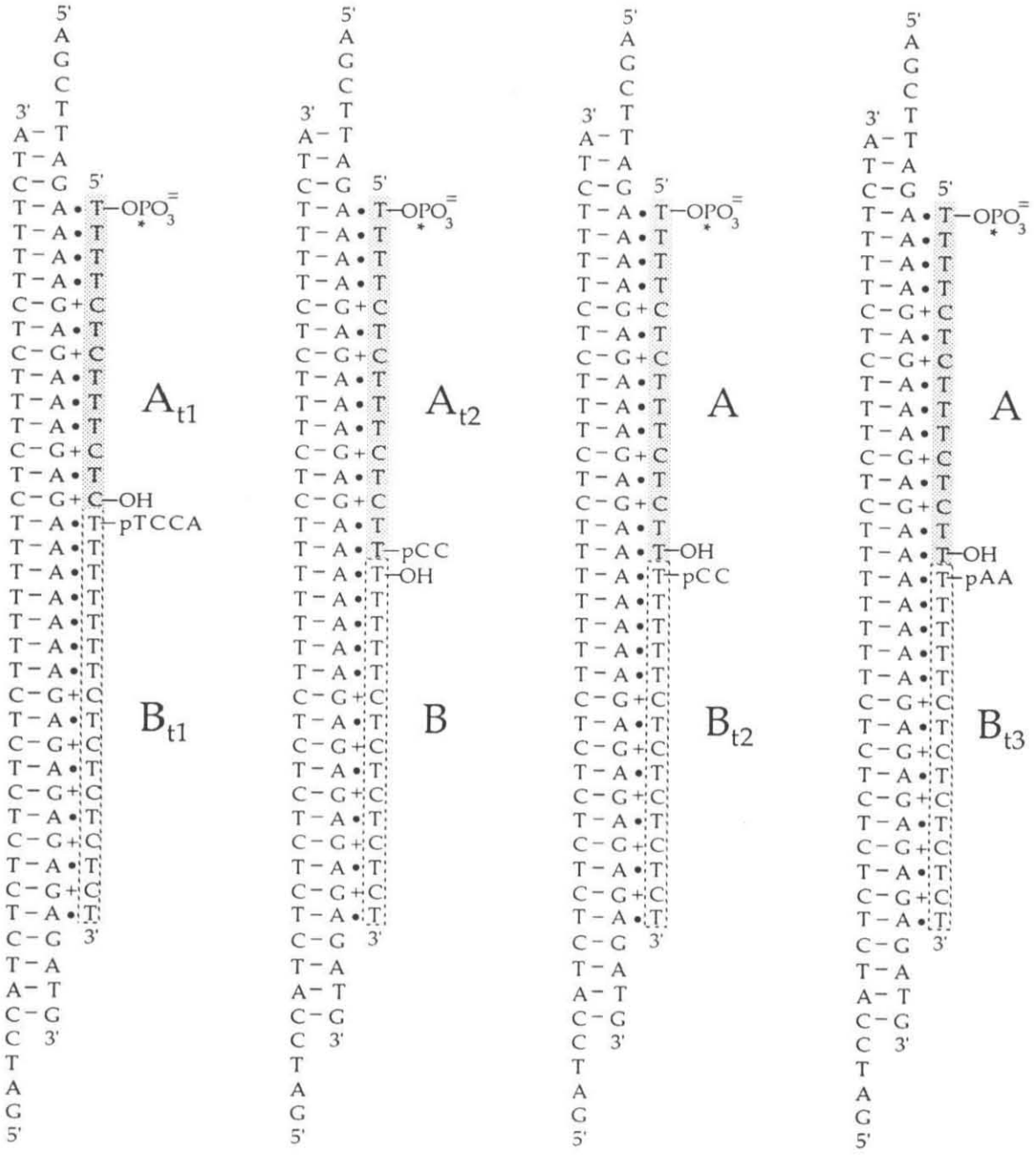


Figure B.1. Oligodeoxyribonucleotides studied in an attempt to observe template-directed transesterification. Triple-helical complexes formed by the association of **A**, $\mathbf{A_{t1}}$, or $\mathbf{A_{t2}}$ and **B**, $\mathbf{B_{t1}}$, $\mathbf{B_{t2}}$, or $\mathbf{B_{t3}}$ with the double-stranded template were expected to place a terminal 3' or 5' hydroxyl proximal to a 3',5' phosphodiester. The targeted phosphodiester links a di- or tetranucleotide to the Hoogsteen segment of an oligodeoxyribonucleotide.

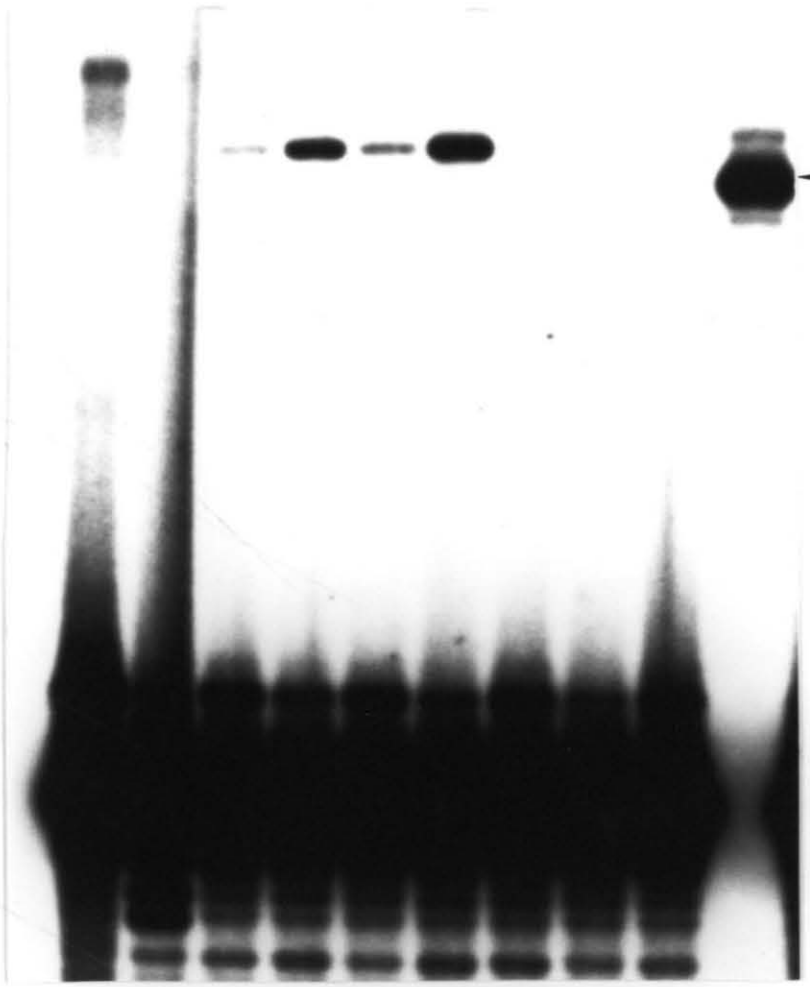
$CaCl_2$, $BaCl_2$, $Sr(NO_3)_2$, $MnCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $ZnCl_2$, $CdCl_2$, $Hg(ClO_4)_2$, K_2PdCl_4, $AuCl_3$, $HfOCl_2$, $La(OAc)_3$, $Ce(NO_3)_3$, $Nd(NO_3)_3$, $Eu(OAc)_3$, $Ho(OAc_3)$,or $Lu(OAc)_3$ (all 20 mM). Metal salts and spermine were omitted from one mixture of each set. The mixtures were warmed to 37 °C for 5-7 days, after which the DNA was precipitated with ethanol and analyzed by electrophoresis in a denaturing polyacrylamide gel.

Reaction of **B** and $\mathbf{A_{t2}}$ yielded a trace (<1% of total radioactivity) of a radioactively labeled product having an electrophoretic mobility approximately like that of a 30mer. This material was produced in the presence of $MgCl_2$, $CaCl_2$, $BaCl_2$, $Sr(NO_3)_2$, $MnCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $ZnCl_2$, $CdCl_2$, $Hg(ClO_4)_2$, and $PbCl_2$. None (<0.1%) was observed after treatment in the presence of the other metal salts tested or spermine, or in the absence of polyvalent cation. No product (<0.1%) of approximately correct mobility was observed in the mixtures containing the other combinations of oligodeoxyribonucleotides.

The reaction with **B** and $\mathbf{A_{t2}}$ was performed again in the presence of 20 mM $MgCl_2$, omitting **B**, the purine strand of the template, or the pyrimidine strand of the template in control experiments. The reaction was carried out for nine days prior to analysis. The result of electrophoretic analysis is shown on the autoradiogram in Figure B.2A. The product of the reaction is slightly less mobile in the gel than the expected product (^{32}P)**AB** (lane 10). The reaction has an absolute requirement for oligodeoxyribonucleotide **B** and is optimal in the presence of the double-stranded template, identifying it as a template-directed reaction between the termini of $\mathbf{A_{t2}}$ and **B**. On the other hand, requirement for the double-stranded template or any template at all is not absolute.

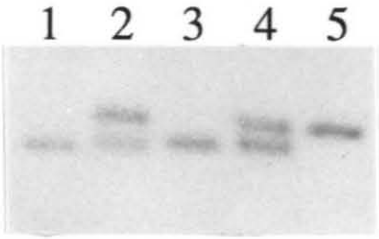
Figure B.2. Characterization of the reaction of $\mathbf{A_{t2}}$ and **B**; autoradiograms of high resolution, denaturing 20% polyacrylamide gels.(A) Dependence of the reaction on double-stranded template and **B**. Each of lanes 1-9 contains $\mathbf{A_{t2}}$, labeled at its 5' terminus with polynucleotide kinase and [γ-^{32}P]ATP. A low mobility product is observed (lanes 3-6) when $\mathbf{A_{t2}}$ has been allowed to react with **B** (3 μM) in the presence of 20 mM imidazole•HCl, pH 7.0, and 20 mM $MgCl_2$ for 9 days at 37 °C. Optimal yield of the product is obtained in the presence of double-stranded template (0.15 μM) (lane 6), though it is also obtained inthe presence of the purine strand of the template alone (lane 4), the pyrimidine strand of the template alone (lane 5), and in the absence of template (lane 3). No product is obtained in the absence of **B** (lanes 2 and 7-9). Lane 10: Authentic sample of the oligodeoxyribonucleotide (^{32}P)**AB** expected from transesterification by nucleophilic attack of the 5'-hydroxyl of **B** on $\mathbf{A_{t2}}$ to displace the 3'-terminal dinucleotide. Lane 1: Untreated (^{32}P)$\mathbf{A_{t2}}$. (B) Electrophoretic comparison of (^{32}P)**CC** and (^{32}P)**CT** with the product of reaction of (^{32})P$\mathbf{A_{t2}}$ and B. Lane 1: (^{32}P)**CC**. Lane 2: (^{32}P)**CC** and reaction product. Lane 3: (^{32}P)**CT**. Lane 4: (^{32}P)**CT** and reaction product. Lane 5: Reaction product.

A



1	2	3	4	5	6	7	8	9	10	
	−	+	+	+	+	−	−	−		B
	−	−	+	−	+	+	−	+		purine
	−	−	−	+	+	−	+	+		pyrimidine

B



The product of the reaction is not the expected product, being less electrophoretically mobile. If the reaction were a transesterification, it could proceed by displacement of a mononucleotide rather than a dinucleotide, resulting in a 31mer, less electrophoretically mobile than the expected 30mer. Such a reaction could involve attack by the 3'-hydroxyl of $\mathbf{A_{t2}}$ to liberate thymidine or by the 5'-hydroxyl of **B** to liberate cytidine. The 31mers produced would contain a central 5'-CC-3' or a central 5'-CT-3', respectively.

The 31mers that would be produced by these reactions, identified here as **CC** or **CT**, respectively, were synthesized independently, labeled at their 5' termini with polynucleotide kinase and [γ-^{32}P]ATP, and compared electrophoretically with the purified product of the template-directed reaction. The result is shown on the autoradiogram in Figure B.2B. The product of the template-directed reaction is less electrophoretically mobile than either of the possible transesterification products, ruling out transesterification as the observed reaction. Due to the inefficiency of the reaction, further determination of the nature of the product was not pursued.

Additional experiments were performed to explore the posibility of transesterification in oligonucleotides containing ribonucleotides as a 3' terminal nucleophile and a 5' terminal leaving group. Oligonucleotides **ArU** and **CrC-B**, shown below where U denotes uridine and rC denotes ribocytidine, were prepared. Oligonucleotide **ArU** was synthesized by standard methodology,[160] starting with 5'-dimethoxytrityl-2'(3')-*O*-acetyluridine attached to the CPG support through the remaining ribose

hydroxyl. Normal deprotection in concentrated ammonia removed the terminal acetyl group.

ArU: 5'-TTTTCTCTTTCTCTU-3'

CrC-B: 5'-CrCTTTTTCTCTCTCTCT-3'

Oligonucleotide **CrC-B** was prepared incorporating ribocytidine at the indicated position as the 2'-t-butyldimethylsilyl protected phosphoramidite.[173] The product was removed from the support and deprotected in ammonia-saturated methanol at room temperature for 20 hours. The silyl protecting group was removed by suspension of the product in a 1 M solution of tetrabutylammonium fluoride (TBAF) in tetrahydrofuran for 6 hours followed by quenching with water and desalting by gel filtration (Sephadex G-25). A portion of the product was not treated with TBAF. This material and the fully deprotected product were both 5'-radiolabeled by phosphorylation with [γ-32P]ATP and polynucleotide kinase and compared electrophoretically (denaturing 20% polyacrylamide gel). The silyl group had a significant effect on the electrophoretic mobility of the product, so its removal by TBAF treatment could be confirmed. The TBAF treated sample contained two well resolved bands. The most mobile of which contained 84% of the radioactivity (band excision and scintillation counting). The least mobile band, containing 16% of the radioactivity, co-migrated with the labeled material that had not been treated with TBAF.

ArU, labeled by phosphorylation at its 5' terminus, was mixed with **CrC-B** (2 μM, unlabeled) and the duplex template from Chapter 2 (0.15 μM) or an oligodeoxyribonucleotide with the same sequence in the opposite direction 5'-3' as the purine strand of the duplex (0.3 μM) in a solution containing 60 mM imidazole•HCl, pH 7.0, and the polyvalent cation salts

used in the attempted transesterification experiments described above. The mixtures were warmed at 37 °C for 17 days, then ethanol precipitated and analyzed by electrophoresis. None of the mixtures yielded a radioactive product with electrophoretic mobility similar to a 30mer. A product of much lower mobility was afforded in the presence of the alkaline earth metals and the single-stranded template. Identification of this product was not pursued.