Probing Contacts to the DNA Backbone in the *trp* Repressor−Operator Sequence-Specific Protein−Nucleic Acid Complex Using Diastereomeric Methylphosphonate Analogues[†]

Stephen A. Smith and Larry W. McLaughlin*

Department of Chemistry, Merkert Chemistry Center, Boston College, Chestnut Hill, Massachusetts 02167

Received January 13, 1997; Revised Manuscript Received March 13, 1997[⊗]

ABSTRACT: Fourteen analogue DNA sequences containing the *trp* operator sequence and a single diastereomeric methylphosphonate linkage are each prepared from the stereochemically pure nucleoside methylphosphonate dimer building block, prepared as a phosphoramidite. The analogue sequences are shown to be single diastereomers on the basis of HPLC analysis of the digestion mixture; in each case, only a single diastereomeric dimer is present. These analogue sequences can be used effectively to probe for interactions to either of the prochiral phosphate oxygens as illustrated by their use to identify critical interactions in the *trp* repressor−operator complex. In a number of cases, the pairs of diastereomeric analogue sequences exhibit variable binding affinities that can be used to identify one of the prochiral phosphate oxygens as a critical site for complex-stabilizing interactions. Upon the basis of dissociation constants, apparent incremental binding energies are assigned to specific interactions. In all but one example, these identified sites for interactions to the phosphate backbone can be correlated with contacts implicated by the crystal structure analysis of the *trp* repressor−operator complex.

The *trp* repressor−operator complex is a small sequence-specific protein−nucleic acid complex that forms to modulate expression from the *trpEDCBA* operon. The consensus operator sequence is a nearly palindromic 18 base pair sequence composed of two direct sequence repeats. A partial understanding of the sequence-specific nature of this complex, as mediated by the binding of the corepressor L-tryptophan, has been obtained from chemical protection studies (Bass et al., 1987; Carey, 1989), genetic manipulations of both the operator (Kelley & Yanofsky, 1982) and the repressor (Bass et al., 1988; Marmorstein et al., 1991) sequences, and from the detailed pictures provided by two crystallographic analysis of the protein−nucleic acid complex in both a dimeric (Otwinowski et al., 1988) and tetrameric form (Lawson & Carey, 1993). The presence of the corepressor L-tryptophan enhances the affinity of the repressor for the operator by initiating a conformational change to permit the interaction of the sequence-reading helices with two adjacent major grooves of the 18 base pair operator (Zhang et al., 1987). Solution assays including a phosphatase protection assay (Marmorstein et al., 1991) and gel retardation assays (Carey et al., 1991; Haran et al., 1992) both indicate high-affinity binding between the repressor and the operator sequence. However, the nature of the interactions at the interface between the protein and the nucleic acid as implicated by the crystal structures is unusual in that only a single direct amino acid−nucleobase contact appears to be present in each half-site: a bidentate interaction between Arg_{69} and the base residue G_{-9}. Five additional water-mediated interactions are present in each half-site in which three water molecules bridge specific functional groups of the base residues in the operator and those of the amino acid side chains or the peptide backbone in the repressor (Otwinowski et al., 1988). Base analogue (Smith et al., 1994) and sequence mutagenesis studies (Joachimiak et al., 1994) have confirmed that the predicted water-meditated contacts provide significant contributions to the free energy of binding and provide in part an explanation for the formation of the observed high-affinity complex.

In addition to the direct or water-mediated interactions with the base residues, there are 24 contacts to the unesterified oxygens of the internucleotide phosphates (12 in each half-site). The operator sequence is kinked at three sites, and this distortion of the operator sequence may contribute to the recognition of the sequence by the protein. The interactions between the repressor protein and the operator sequence have provided the basis for the proposal of an indirect sequence readout mechanism (Otwinowski et al., 1988), in which the position of the phosphates contributes to sequence selectivity. The nature of the interactions between the internucleotide phosphates and the protein has not been examined to date.

Conventional mutagenesis studies have a limited ability to resolve the nature of specific interactions within protein−DNA complexes and their contribution to high-affinity binding. Base analogues, those in which a single functional group has been excised or otherwise altered, have proven valuable in assessing the contribution of individual interactions to overall complex stability in a number of studies (Aiken & Gumport, 1991; Waters & Connolly, 1994) of sequence-specific protein−nucleic acid complexes including the *trp* repressor−operator complex (Mazzarelli et al., 1992; Smith et al., 1994). Similar studies to probe the nature of specific interactions to the internucleotide phosphate residues are more complex. The internucleotide phosphodiesters contain two unesterified phosphate oxygens that share a negative charge. It appears quite common for a protein to

[†] This work was supported by a grant from the NSF (MCB-9507040).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1997.

[1] Abbreviations: $Np(CH_3)N$ represents 5′ → 3′ nucleoside dimers in which the internucleotide linkage is present as a chiral methylphosphonate derivative. The corresponding dA dimer is represented as $d[Ap(CH_3)A]$.

S0006-2960(97)00078-0 CCC: $14.00

contact only one of the two prochiral nonbridging oxygens. For example, for a protein bound in the major groove of a B-form DNA, the *pro-R* oxygen of the phosphodiester linkage is oriented more toward the major groove, while the *pro-S* oxygen is oriented more toward the minor groove. Preferential, but not exclusive contact with the *pro-R* oxygens could be expected for a major groove binding protein. Studies in which phosphate analogues have been used to probe the presence of interactions to the DNA backbone in protein−nucleic acid complexes have not always considered the diastereomeric nature of such interactions (Botfield & Weiss, 1994; Thorogood et al., 1996). A recent study with T4 endonuclease V (Iwai et al., 1994), and a second with the HIV Tat protein (Pritchard et al., 1994), both involved the preparation of DNA or RNA sequences, respectively, containing racemic methylphosphonates, followed by their resolution into stereochemically pure isomers for analysis of protein binding. Studies with the *Eco*RI restriction endonuclease have described the use of chiral phosphorothioates to probe essential interactions to the phosphates in the complex involving the GAATTC recognition site (Koziolkiewicz & Stec, 1992; Lesser et al., 1992; Kurpiewski et al., 1996). An earlier study with the *lac* repressor employed chiral methylphosphonate analogues (Noble et al., 1984).

In the present report, we describe the use of chiral methylphosphonate derivatives to probe interactions to each of the prochiral oxygens of the internucleotide phosphates in the *trp* repressor−operator complex. Six phosphates in each half-site have been proposed to be involved in interactions with the *trp* repressor. We have probed these six sites using the corresponding 12 diastereomeric R_p and S_p methylphosphonate derivatives.

EXPERIMENTAL PROCEDURES

Materials. Oligodeoxynucleotides were synthesized using nucleoside phosphoramidites on an Applied Biosystems 381A DNA synthesizer. The four fully protected common 2′-deoxynucleoside phosphoramidites containing aryl- or isobutyrylamides were purchased from BioGenex. The *trp* repressor was isolated from an overproducing strain provided as a gift from Prof. P. Sigler (Yale University). Calf intestinal alkaline phosphatase was a product of Boerhinger Mannheim Inc. (Germany).

Methods. Four of the protected methylphosphonate dimers {R_p and S_p d[Ap(CH_3)A]; R_p and S_p d[Tp(CH_3)A]} were prepared with the 3′-hydroxyl protected as the *O*-benzoyl derivative, while the remaining eight dimers used *tert*-butyldiphenylsilyl protection (see text for explanation).

*5′-O-(4,4′-Dimethoxytrityl)thymidylyl-(3′→5′)-N⁶-benzoyl-3′-O-benzoyl-2′-deoxyadenosine 3′-Methylphosphonate (**3**).* The procedure for dimer coupling to prepare the protected derivative of d[Tp(CH_3)A] was adapted from two previously published approaches (Cormier & Pannunzio, 1991; Wozniak et al., 1994). To 2 g (2.6 mmol) of DMT-protected hexafluoroisopropyl methylphosphonate derivative (**1**) of thymidine was added 0.95 g (2.1 mmol) of N^6-benzoyl-3′-*O*-benzoyl-2′-deoxyadenosine corresponding to **2a** followed by 1.2 equiv of DBU (2.5 mmol), and the reaction stirred overnight. TLC analysis (chloroform/methanol, 95/5) indicated that two new products were present. The reaction mixture was evaporated to dryness without any work up, dissolved in a minimum amount of dichloromethane and chromatographed on a column of silica gel (2.5 × 15 cm) using a gradient of methanol (0 to 4%). After collection of the appropriate fractions and evaporation to dryness 880 mg of the R_p diastereomer (39%) and 350 mg of the S_p diastereomer (16%) were obtained. A combined yield of 55% was obtained.

R_f of R_p isomer (chloroform/methanol, 95/5) = 0.25, and R_f of S_p isomer (chloroform/methanol, 95/5) = 0.23. MS FAB calculated for $C_{56}H_{54}N_7O_{13}P$ ($M + H^+$): 1064.3595. Found: 1064.3591.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.8 (s, 1H, N-H of A), 8.4 (s, 1H, H2 of A), 8.0 (s, 1H, H8 of A), 7.6−6.8 (m, 23H, Ar-H), 6.62 (t, 1H, J = 7.0 Hz, H1′ of A), 6.38 (t, 1H, J = 6.0 Hz, H1′ of T), 5.75 (m, 1H, H3′ of A), 5.24 (m, 1H, H3′ of T), 4.48 (m, 1H, H4′ of A), 4.17 (m, 1H, H4′ of T), 3.78 (s, 6H, O-CH_3), 3.6 (m, 2H, H5′ of T), 3.41 (m, 2H, H5′ of A), 3.16 (m, 1H, H2′ of A), 2.83 (m, 1H, H2″ of A), 2.60 (m, 1H, H2′ of T), 2.46 (m, 1H, H2″ of T), 1.5−1.6 (d, 3H, P-CH_3), 1.3 (s, 3H, CH_3 of T).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.8 (s, 1H, N-H of A), 8.35 (s, 1H, H2 of A), 8.0 (s, 1H, H8 of A), 7.6−6.8 (m, 23H, Ar-H), 6.62 (t, 1H, J = 6.8 Hz, H1′ of A), 6.38 (t, 1H, J = 6.8 Hz, H1′ of T), 5.65 (m, 1H, H3′ of A), 5.24 (m, 1H, H3′ of T), 4.40 (m, 1H, H4′ of A), 4.20 (m, 1H, H4′ of T), 3.78 (s, 6H, O-CH_3), 3.6 (m, 2H, H5′ of T), 3.41 (m, 2H, H5′ of A), 3.10 (m, 1H, H2′ of A), 2.85 (m, 1H, H2″ of A), 2.50 (m, 1H, H2′ of T), 2.40 (m, 1H, H2″ of T), 1.5−1.6 (d, 3H, P-CH_3), 1.3 (s, 3H, CH_3 of T).

*5′-O-(4,4′-Dimethoxytrityl)thymidylyl-(3′→5′)-N⁶-benzoyl-2′-deoxyadenosine 3′-Methylphosphonate (**4**).* To 400 mg of either the R_p or the S_p diastereomer of **3** (0.38 mmol, the 3′-*O*-benzoyl derivative of d[Tp(CH_3)A]) in methanol was added 10 mL of a 0.05 M solution (0.50 mmol) of freshly prepared sodium methoxide in methanol, and the reaction was checked every 15 min (TLC) until starting material was largely absent. A cation exchange resin (H^+ form) was added until the pH of the solution was neutral and then the solution was filtered. The filtrate was poured into a solution of 5% sodium bicarbonate and extracted with chloroform. After drying over sodium sulfate and concentrating the solution to dryness, the mixture was dissolved in a minimum amount of chloroform and chromatographed on a column of silica gel column (1 × 10 cm) using a gradient of methanol (0 to 8%). After collection of the appropriate fractions and evaporation to dryness, 175 mg (0.18 mmol) of 48% of the R_p (40% yield for the S_p diastereomer) diastereomer of **4** was obtained.

R_f of R_p isomer (chloroform/methanol, 90/10) = 0.50, and R_f of S_p isomer (chloroform/methanol, 90/10) = 0.50.

MS FAB calculated for $C_{49}H_{51}N_7O_{12}P$ ($M + H^+$): 960.3333. Found: 960.3340.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.8 (s, 1H, N-H of A), 8.4 (s, 1H, H2 of A), 8.0 (s, 1H, H8 of A), 7.6−6.8 (m, 23H, Ar-H), 6.62 (t, 1H, J = 7.0 Hz, H1′ of A), 6.38 (t, 1H, J = 6.0 Hz, H1′ of T), 5.75 (m, 1H, H3′ of A), 5.24 (m, 1H, H3′ of T), 4.6 (m, 1H, H3′ of A), 4.48 (m, 1H, H4′ of A), 4.17 (m, 1H, H4′ of T), 4.05 (d, 1H, 3′OH of A), 3.78 (s, 6H, O-CH_3), 3.6 (m, 2H, H5′ of T), 3.41 (m, 2H, H5′ of A), 3.16 (m, 1H, H2′ of A), 2.83 (m, 1H, H2″ of A), 2.60 (m, 1H, H2′ of T), 2.46 (m, 1H, H2″ of T), 1.5−1.6 (d, 3H, P-CH_3), 1.3 (s, 3H, CH_3 of T).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.8 (s, 1H, N-H of A), 8.35 (s, 1H, H2 of A), 8.0 (s, 1H, H8 of A), 7.6−6.8 (m, 23H, Ar-H), 6.62 (t, 1H, J = 6.82 Hz, H1′ of A), 6.38 (t,

1H, $J = 6.82$ Hz, H1′ of T), 5.24 (m, 1H, H3′ of T), 4.70 (m, 1H, H3′ of A), 4.40 (m, 1H, H4′ of A), 4.20 (m, 1H, H4′ of T), 4.05 (d, 1H, 3′OH of A), 3.78 (s, 6H, O-CH_3), 3.6 (m, 2H, H5′ of T), 3.41 (m, 2H, H5′ of A), 3.10 (m, 1H, H2′ of A), 2.85 (m, 1H, H2″ of A), 2.50 (m, 1H, H2′ of T), 2.40 (m, 1H, H2″ of T), 1.5−1.6 (d, 3H, P-CH_3), 1.3 (s, 3H, CH_3 of T).

*5′-O-(4,4′-Dimethoxytrityl)thymidylyl-(3′→5′)-N⁶-benzoyl-3′-O-(2-cyanoethyl N,N-diisopropylamidophosphoramidite)-2′-deoxyadenosine 3′-Methylphosphonate (**5**).* To 150 mg of the R_p diastereomer of the d[Tp(CH_3)A] derivative **4** (0.16 mmol) dissolved in 2 mL of anhydrous chlorofom was added 3 equiv (88 μL) of diisopropylethylamine (0.51 mmol) and 1.5 equiv (60 μL) of 2-cyanoethyl diisopropylchlorophosphoramidite (0.26 mmol), and the reaction mixture was stirred for 2 h. This mixture was added directly to a silica gel column (1 × 10 cm) without any work up and eluted with chloroform and a gradient of methanol (0 to 5%) containing a trace of triethylamine. After collection of the appropriate fractions, 150 mg (0.12 mmol, 54%) of the R_p (50% yield for the S_p diastereomer) phosphoramidite dimer was obtained.

R_f of R_p isomer (chloroform/methanol/triethylamine, 94/5/1) = 0.35, and R_f of S_p isomer (chloroform/methanol/triethylamine, 94/5/1) = 0.35.

^{31}P NMR ($CDCl_3$) R_p isomer: δ 31.78, 31.84 (methylphosphonate), 149.04, 149.06 (phosphoramidite).

^{31}P NMR ($CDCl_3$) S_p isomer: δ 32.11, 32.27, (methylphosphonate), 149.25, 149.29 (phosphoramidite).

*N⁶-Benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyadenoylyl-(3′→5′)-N⁶-benzoyl-3′-O-benzoyl-2′-deoxyadenosine 3′-Methylphosphonate (**6**).* The corresponding benzoyl-protected dimer of the methylphosphonate of d[Ap(CH_3)A] was prepared by a procedure analagous to that described for **3**. From 1 g (2.2 mmol) of **2**, 284 mg (0.25 mmol) of **6** was obtained (yield of 11% for the R_p diastereomer). An 8% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.22, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.20.

MS FAB calculated for $C_{63}H_{57}N_{10}O_{12}PK$ ($M + K^+$): 1215.3532. Found: 1215.3580.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 9.1−9.2 (s, 2H, N-H of A), 8.1−8.8 (s, 4H, H2 and H8 of A), 7.6−6.8 (m, 31H, Ar-H), 6.60 (t, 1H, $J = 6.8$ Hz, H1′ of A1), 6.45 (t, 1H, $J = 7.2$ Hz, H1′ of A2), 5.75 (m, 1H, H3′ of A2), 5.34 (m, 1H, H3′ of A1), 4.50 (m, 1H, H4′ of A2), 4.35 (m, 1H, H4′ of A1), 3.71 (s, 6H, O-CH_3), 3.45 (m, 2H, H5′ of A2), 3.35 (m, 2H, H5′ of A1), 3.08 (m, 1H, H2′ of A2), 3.06 (m, 1H, H2′ of A1), 2.78 (m, 1H, H2″ of A1), 2.76 (m, 1H, H2″ of A2), 1.5−1.6 (d, 3H, P-CH_3).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 9.2−9.3 (s, 2H, N-H of A), 8.2−8.8 (s, 4H, H2 and H8 of A), 7.6−6.8 (m, 31 H, Ar-H), 6.60 (t, 1H, $J = 6.8$ Hz, H1′ of A1), 6.50 (t, 1H, $J = 7.1$ Hz, H1′ of A2), 5.68 (m, 1H, H3′ of A2), 5.34 (m, 1H, H3′ of A1), 4.42 (m, 1H, H4′ of A2), 4.35 (m, 1H, H4′ of A1), 3.71 (s, 6H, O-CH_3), 3.45 (m, 2H, H5′ of A2), 3.35 (m, 2H, H5′ of A1), 3.08 (m, 1H, H2′ of A2), 3.06 (m, 1H, H2′ of A1), 2.78 (m, 1H, H2″ of A1), 2.76 (m, 1H, H2″ of A2), 1.5−1.6 (d, 3H, P-CH_3).

*N⁶-Benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyadenoylyl-(3′→5′)-N⁶-benzoyl-2′-deoxyadenosine 3′-Methylphosphonate (**7**).* Deprotection of the benzoyl-protected dimer of the methylphosphonate of d[Ap(CH_3)A] was prepared by a procedure analagous to that described for **4**. From 250 mg (0.21 mmol) of **6**, 110 mg (0.10 mmol) of **7** was obtained (yield of 48% for the R_p diastereomer). A 41% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.52, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.52.

MS FAB calculated for $C_{56}H_{53}N_{10}O_{11}PK$ ($M + K^+$): 1111.3270. Found: 1111.3291.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 9.1−9.2 (s, 2H, N-H of A), 8.1−8.8 (s, 4H, H2 and H8 of A), 7.6−6.8 (m, 31H, Ar-H), 6.60 (t, 1H, $J = 6.8$ Hz, H1′ of A1), 6.45 (t, 1H, $J = 7.2$ Hz, H1′ of A2), 5.75 (m, 1H, H3′ of A2), 5.34 (m, 1H, H3′ of A1), 4.50 (m, 1H, H4′ of A2), 4.35 (m, 1H, H4′ of A1), 4.11 (d, 1H, 3′OH of A), 3.71 (s, 6H, O-CH_3), 3.45 (m, 2H, H5′ of A2), 3.35 (m, 2H, H5′ of A1), 3.08 (m, 1H, H2′ of A2), 3.06 (m, 1H, H2′ of A1), 2.78 (m, 1H, H2″ of A1), 2.76 (m, 1H, H2″ of A2), 1.5−1.6 (d, 3H, P-CH_3).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 9.2−9.3 (s, 2H, N-H of A), 8.2−8.8 (s, 4H, H2 and H8 of A), 7.6−6.8 (m, 31H, Ar-H), 6.60 (t, 1H, $J = 6.8$ Hz, H1′ of A1), 6.50 (t, 1H, $J = 7.1$ Hz, H1′ of A2), 5.68 (m, 1H, H3′ of A2), 5.34 (m, 1H, H3′ of A1), 4.42 (m, 1H, H4′ of A2), 4.35 (m, 1H, H4′ of A1), 4.11 (d, 1H, 3′OH of A), 3.71 (s, 6H, O-CH_3), 3.45 (m, 2H, H5′ of A2), 3.35 (m, 2H, H5′ of A1), 3.08 (m, 1H, H2′ of A2), 3.06 (m, 1H, H2′ of A1), 2.78 (m, 1H, H2″ of A1), 2.76 (m, 1H, H2″ of A2), 1.5−1.6 (d, 3H, P-CH_3).

*N⁶-Benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyadenoylyl-(3′→5′)-3′-O-(2-cyanoethyl N,N-diisopropylamidophosphoramidite)-2′-deoxyadenosine 3′-Methylphosphonate (**8**).* Conversion of the nucleoside dimer of the methylphosphonate d[Ap(CH_3)A] to the corresponding phosphoramidite was accomplished by a procedure analagous to that described for **5**. From 110 mg (0.10 mmol) of **7**, 65 mg (0.05 mmol) of **8** was obtained (yield of 52% for the R_p diastereomer). A 45% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol/triethylamine, 94/5/1) = 0.35, and R_f of S_p isomer (chloroform/methanol/triethylamine, 94/5/1) = 0.35.

^{31}P NMR of R_p isomer ($CDCl_3$): δ 31.78, 31.84 (methylphosphonate), 149.04, 149.06 (phosphoramidite),

^{31}P NMR of S_p isomer ($CDCl_3$): δ 32.11, 32.27, (methylphosphonate), 149.25, 149.29 (phosphoramidite).

*N⁶-Benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyadenoylyl-(3′→5′)-N⁴-isobutyryl-3′-O′-tert-butyldiphenylsilyl-2′-deoxycytidine 3′-Methylphosphonate (**9**).* The procedure for dimer coupling to prepare the protected derivative of d[Ap(CH_3)C] was adapted from that described by Wozniak et al. (1994). To 1.1 g (1.26 mmol) of N^6-benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-adenosine hexafluoromethylphosphonate (**1**) (Cormier & Pannunzio, 1991) was added 0.45 g (0.84 mmol) of N^4-isobutyryl-3′-*O*-(*tert*-butyldiphenylsilyl)-2′-deoxycytidine corresponding to **2a** followed by 5 equiv (0.168 g) of LiCl (4.20 mmol) in 10 mL of anhydrous pyridine. After 5 min of stirring, 5 equiv (0.63 mL) of diazobiscycloundecene (4.20 mmol) was added and the reaction stirred overnight. TLC analysis (chloroform/methanol, 95/5) indicated that two new products were present. The reaction mixture was evaporated to dryness without any workup, dissolved in a minimum amount of dichloromethane, and chromatographed on a column of silica gel (2.5 × 15 cm) using a gradient of methanol (0 to 4%). After collection of the appropriate fractions and evaporation to dryness, 370 mg of the R_p diastereomer (35%) and 259 mg of the S_p diastereomer (25%) was obtained. An additional 70 mg of a mixture of both R_p

and S_p isomers (7%) was also isolated. A combined yield of 67% was obtained.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.20, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.18.

MS FAB calculated for $C_{68}H_{74}N_8O_{12}SiP$ (M + H^+): 1253.4933. Found: 1253.4940.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 9.1 (s, 1H, N-H of A), 8.6 (s, 1H, H2 of A), 8.1 (s, 1H, H8 of A), 7.97 (d, 2H, H5 and H6 of C), 7.6−6.8 (m, 23H, Ar-H), 6.39 (t, 1H, J = 6.3 Hz, H_1' of C), 6.31 (t, 1H, J = 6.3 Hz, H1′ of A), 5.18 (m, 1H, H3′ of C), 4.32 (m, 1H, H3′ of A), 4.23 (m, 1H, H4′ of C), 3.98 (m, 1H, H4′ of A), 3.8 (s, 6H, O-CH_3), 3.59 (m, 2H, H5′ of A), 3.38 (m, 2H, H5′ of C), 2.97 (m, 1H, H2′ of C), 2.73 (m, 1H, H2′ of A), 2.59 (m, 1H, C-H of isobutyryl group), 2.50 (m, 1H, H2″ of C), 1.87 (m, 1H, H2″ of A), 1.3−1.4 (d, 3H, P-CH_3), 1.2 (d, 6H, isobutyryl group), 1.05 (s, 9H, *tert*-butyl group).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 9.1 (s, 1H, N-H of A), 8.7 (s, 1H, H2 of A), 8.2 (s, 1H, H8 of A), 8.0 (d, 2H, H5 and H6 of C), 7.6−6.8 (m, 23H, Ar-H), 6.55 (t, 1H, J = 7.5 Hz, H_1' of C), 6.40 (t, 1H, J = 7.5 Hz, H1′ of A), 5.40 (m, 1H, H3′ of C), 4.45 (m, 1H, H3′ of A), 4.25 (m, 1H, H4′ of C), 3.98 (m, 1H, H4′ of A), 3.8 (s, 6H, O-CH_3), 3.50 (m, 2H, H5′ of A), 3.40 (m, 2H, H5′ of C), 2.90 (m, 1H, H2′ of C), 2.75 (m, 1H, H2′ of A), 2.59 (m, 1H, C-H of isobutyryl group), 2.55 (m, 1H, H2″ of C), 2.15 (m, 1H, H2″ of A), 1.5−1.6 (d, 3H, P-CH_3), 1.2 (d, 6H, isobutyryl group), 1.05 (s, 9H, *tert*-butyl group).

*N^6-Benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyadenoylyl-(3′→5′)-N^4-isobutyryl-2′-deoxycytidine 3′-Methylphosphonate (**10**).* To 300 mg of either the R_p or the S_p diastereomer of the d[Ap(CH_3)C] dimer **9** (0.24 mmol) was added 2 equiv of TBAF (0.48 mL) in freshly distilled THF (5 mL) and the reaction stirred for 1 h. The solution was poured into a solution of 5% sodium bicarbonate and extracted with chloroform. After drying over sodium sulfate and concentrating the solution, the mixture was chromatographed on a column of silica gel (1 × 10 cm) using a gradient of methanol (0 to 8%). After collection of the appropriate fractions and evaporation to dryness, 210 mg (0.21 mmol), 86% of the R_p (71% of the S_p) diastereomer, of **10** was obtained.

R_f of R_p isomer (chloroform/methanol, 90/10%) = 0.52, and R_f of S_p isomer (chloroform/methanol, 90/10%) = 0.50.

MS FAB calculated for $C_{52}H_{56}N_8O_{12}P$ (M + H^+): 1015.3755. Found: 1015.3765.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 9.1 (s, 1H, N-H of A), 8.6 (s, 1H, H2 of A), 8.1 (s, 1H, H8 of A), 7.97 (d, 2H, H5 and H6 of C), 7.6−6.8 (m, 23H, Ar-H), 6.39 (t, 1H, J = 6.3 Hz, H1′ of C), 6.31 (t, 1H, J = 6.3 Hz, H1′ of A), 5.4 (d, 1H, 3′-OH of C), 5.18 (m, 1H, H3′ of C), 4.32 (m, 1H, H3′ of A), 4.23 (m, 1H, H4′ of C), 3.98 (m, H4′ of A), 3.8 (s, 6H O-CH_3), 3.59 (m, 2H, H5′ of A), 3.38 (m, 2H, H5′ of C), 2.97 (m, 1H, H2′ of C), 2.73 (m, 1H, H2′ of A), 2.59 (m, 1H, C-H of isobutyryl group), 2.50 (m, 1H, H2″ of C), 1.87 (m, 1H, H2″ of A), 1.3−1.4 (d, 3H, P-CH_3), 1.2 (d, 6H, isobutyryl group).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 9.1 (s, 1H, N-H of A), 8.7 (s, 1H, H2 of A), 8.2 (s, 1H, H8 of A), 8.0 (d, 2H, H5 and H6 of C), 7.6−6.8 (m, 23H, Ar-H), 6.55 (t, 1H, J = 7.5 Hz, H1′ of C), 6.40 (t, 1H, J = 7.5 Hz, H1′ of A), 5.4 (d, 1H, 3′-OH of C), 5.20 (m, 1H, H3′ of C), 4.45 (m, 1H, H3′ of A), 4.25 (m, 1H, H4′ of C), 3.98 (m, 1H, H4′ of A), 3.8 (s, 6H, O-CH_3), 3.50 (m, 2H, H5′ of A), 3.40 (m, 2H, H5′ of C), 2.90 (m, 1H, H2′ of C), 2.75 (m, 1H, H2′ of A), 2.59 (m, 1H, C-H of isobutyryl group), 2.55 (m, 1H, H2″ of C), 2.15 (m, 1H, H2″ of A), 1.5−1.6 (d, 3H, P-CH_3), 1.2 (d, 6H, isobutyryl group).

*N^6-Benzoyl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyadenoylyl-(3′→5′)-N^4-isobutyryl-3′-O-(2-cyanoethyl N,N-diisopropylamidophosphoramidite)-2′-deoxycytidine 3′-Methylphosphonate (**11**).* Conversion of the nucleoside dimer of the methylphosphonate d[Ap(CH_3)C] to the corresponding phosphoramidite was accomplished by a procedure analagous to that described for **5**. From 170 mg (0.17 mmol) of **10**, 150 mg (0.12 mmol) of **11** was obtained (yield of 73% for the R_p diastereomer). A 63% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol/triethylamine, 94/5/1) = 0.35, and R_f of S_p isomer (chloroform/methanol/triethylamine, 94/5/1) = 0.35.

^{31}P NMR R_p isomer ($CDCl_3$): δ 32.18, 32.23 (methylphosphonate), 149.21, 149.68 (phosphoramidite),

^{31}P NMR S_p isomer ($CDCl_3$): δ 32.38, 32.50, (methylphosphonate), 149.34, 149.70 (phosphoramidite).

N^2-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyguanidyl-(3′→5′)-(3′-O′-tbutyldiphenylsilyl-thymidine)-3′-methylphosphonate (12). The corresponding *tert*-butyldimethylsilyl-protected dimer of the methylphosphonate of d[Gp(CH_3)T] was prepared by a procedure analagous to that described for **9**. From 610 mg (1.27 mmol) of **2b**, 390 mg (0.33 mmol) of **12** was obtained (yield of 26% for the R_p diastereomer). A 29% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.20, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.17.

MS FAB calculated for $C_{62}H_{71}N_7O_{13}SiP$ (M + H^+): 1180.4617. Found: 1180.4623.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.6 (d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.18 (t, 1H, J = 6.7 Hz, H1′ of G), 6.02 (t, 1H, J = 6.9 Hz, H1′ of T), 5.25 (m, 1H, H3′ of G), 4.42 (m, 1H, H3′ of T), 4.13 (m, 1H, H4′ of G), 4.15 (m, H4′ of T), 3.8 (s, 6H O-CH_3), 3.35 (m, 2H, H5′ of T), 3.24 (m, 2H, H5′ of G), 2.66 (m, 1H, H2′ of G), 2.60 (m, 1H, C-H of isobutyryl group), 2.47 (m, 1H, H2′ of G), 2.40 (m, 1H, H2′ of T), 2.30 (m, 1H, H2′ of T), 1.90 (s, 3H, CH_3 of T), 1.4 (d, 3H, P-CH_3), 1.18 (d, 6H, isobutyryl group), 1.07 (s, 9H, *tert*-butyl group).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.6 (d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.18 (t, 1H, J = 6.86 Hz, H1′ of T), 6.02 (t, 1H, J = 6.34 Hz, H1′ ofG), 5.34 (m, 1H, H3′ of G), 4.36 (m, 1H, H3′ of G), 4.24 (m, 1H, H4′ of G), 3.94 (m, H4′ of T), 3.8 (s, 6H, O-CH_3), 3.59 (m, 2H, H5′ of T), 3.27 (m, 2H, H5′ of G), 2.66 (m, 1H, H2′ of G), 2.60 (m, 1H, C-H of isobutyryl group), 2.40 (m, 1H, H2′ of G), 2.37 (m, 1H, H2′ of T), 2.0 (s, 3H, CH_3 of T), 1.9 (m, 1H, H2′ of T), 1.4 (d, 3H, P-CH3), 1.18 (d, 6H, isobutyryl group), 1.07 (s, 9H, *tert*-butyl group).

*N^2-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyguanidyl-(3′→5′)-thymidine 3′-Methylphosphonate (**13**).* Deprotection of the *tert*-butyldimethylsilyl-protected dimer of the methylphosphonate of d[Gp(CH_3)T] was accomplished by a procedure analagous to that described for **10**. From 320 mg (0.27 mmol) of **12**, 190 mg (0.20 mmol) of **13** was obtained (yield of 75% for the R_p diastereomer). A 75% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 90/10%) = 0.50, R_f of S_p isomer (chloroform/methanol, 90/10%) = 0.50.

MS FAB calculated for $C_{46}H_{53}N_7O_{13}P$ (M + H^+): 942.3439. Found: 942.3441.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.6 (d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.18 (t, 1H, J = 6.7 Hz, H1′ of G), 6.02 (t, 1H, J = 6.9 Hz, H1′ of T), 5.4 (d, 1H, 3′-OH of T), 5.1 (m, 1H, H3′ of G), 4.42 (m, 1H, H3′ of T), 4.13 (m, 1H, H4′ of G), 4.15 (m, H4′ of T), 3.8 (s, 6H O-CH_3), 3.35 (m, 2H, H5′ of T), 3.24 (m, 2H, H5′ of G), 2.66 (m, 1H, H2′ of G), 2.60 (m, 1H, C-H of isobutyryl group), 2.47 (m, 1H, H2′ of G), 2.40 (m, 1H, H2′ of T), 2.30 (m, 1H, H2′ of T), 1.90 (s, 3H, CH_3 of T), 1.4 (d, 3H, P-CH3), 1.18 (d, 6H, isobutyryl group).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.6 (d H, N-H), 7.6−6.8 (m, 23H, arom), 6.18 (t, 1H, J = 6.86 Hz, H1′ of T), 6.02 (t, 1H, J = 6.34 Hz, H1′ of G), 5.4 (d, 1H, 3′-OH of T), 5.25 (m, 1H, H3′ of G), 4.36 (m, 1H, H3′ of G), 4.24 (m, 1H, H4′ of G), 3.94 (m, H4′ of T), 3.8 (s, 6H O-CH_3), 3.59 (m, 2H, H5′ of T), 3.27 (m, 2H, H5′ of G), 2.66 (m, 1H, H2′ of G), 2.60 (m, 1H, C-H of isobutyryl group), 2.40 (m, 1H, H2′ of G), 2.37 (m, 1H, H2′ of T), 2.0 (s, 3H, CH_3 of T), 1.9 (m, 1H, H2′ of T), 1.4 (d, 3H, P-CH3), 1.18 (d, 6H, isobutyryl group).

*N^2-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxyguanidyl-(3′→5′)-3′-O-(2-cyanoethyl N,N-diisopropylamidophosphoramidite)-thymidine 3′-Methylphosphonate (**14**).* Conversion of the nucleoside dimer of the methylphosphonate d[Gp(CH_3)T] to the corresponding phosphoramidite was accomplished by a procedure analagous to that described for **5**. From 140 mg (0.15 mmol) of **13**, 100 mg (0.088 mmol) of **14** was obtained (yield of 59% for the R_p diastereomer). A 62% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.33, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.33.

^{31}P NMR of R_p isomer ($CDCl_3$): δ 31.97, 32.08 (methylphosphonate), 148.96, 149.08 (phosphoramidite).

^{31}P NMR of S_p isomer ($CDCl_3$): δ 32.16, 32.32 (methylphosphonate), 148.76, 149.34 (phosphoramidite).

*N^4-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxycytidylyl-(3′→5′)-3′-O′-tert-butyldiphenylsilylthymidine 3′-Methylphosphonate (**15**).* The corresponding *tert*-butyldimethylsilyl-protected dimer of the methylphosphonate of d[Cp(CH_3)T] was prepared by a procedure analagous to that described for **9**. From 340 mg (0.71 mmol) of **2a**, 200 mg (0.17 mmol) of **15** was obtained (yield of 25% for the R_p diastereomer). A 24% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.20, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.17.

MS FAB calculated for $C_{61}H_{71}N_5O_{13}SiP$ ($M + H^+$): 1140.4555. Found: 1140.4534.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.1 (d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.3 (t, 1H, J = 7.2 Hz, H1′ of T), 6.2 (t, 1H, J = 7.2 Hz, H1′ of C), 5.05 (m, 1H, H3′ of C), 4.3 (m, 1H, H3′ of T), 4.15 (m, 1H, H4′ of C), 4.10 (m, H4′ of T), 3.8 (s, 6H O-CH_3), 3.55 (m, 2H, H5′ of T), 3.35 (m, 2H, H5′ of C), 2.7 (m, 1H, H2′ of C), 2.50 (m, 1H, C-H of isobutyryl group), 2.40 (m, 1H, H2′ of T), 2.30 (m, 1H, H2′ of C), 1.90 (s, 3H, CH_3 of T), 1.90 (m, 1H, H2′ of T), 1.4 (d, 3H, P-CH_3), 1.10 (d, 6H, isobutyryl group), 1.05 (s, 9H, *tert*-butyl group)

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.1 (d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.4 (t, 1H, J = 7.5 Hz, H1′ of T), 6.2 (t, 1H, J = 6.6 Hz, H1′ of C), 5.1 (m, 1H, H3′ of C), 4.25 (m, 1H, H3′ of T), 4.15 (m, 1H, H4′ of C), 4.0 (m, H4′ of T), 3.77 (s, 6H O-CH_3), 3.65 (m, 2H, H5′ of T), 3.35 (m, 2H, H5′ of C), 2.8 (m, 1H, H2′ of C), 2.6 (m, 1H, H2′ of T), 2.50 (m, 1H, C-H of isobutyryl group), 2.30 (m, 1H, H2′ of C), 1.90 (s, 3H, CH_3 of T), 1.8 (m, 1H, H2′ of T), 1.4 (d, 3H, P-CH_3), 1.20 (d, 6H, isobutyryl group), 1.05 (s, 9H, *tert*-butyl group)

*N^4-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxycytidylyl-(3′→5′)-thymidine 3′-Methylphosphonate (**16**).* Deprotection of the *tert*-butyldimethylsilyl-protected dimer of the methylphosphonate of d[Cp(CH_3)T] was accomplished by a procedure analagous to that described for **10**. From 190 mg (0.17 mmol) of **15**, 115 mg (0.13 mmol) of **16** was obtained (yield of 75% for the R_p diastereomer). A 74% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 90/10%) = 0.50, and R_f of S_p isomer (chloroform/methano,l 90/10%) = 0.50.

MS FAB calculated for $C_{45}H_{53}N_5O_{13}P$ ($M + H^+$): 902.3378 Found: 902.3411.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.1(d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.3 (t, 1H, J = 7.2 Hz, H1′ of T), 6.2(t, 1H, J = 7.2 Hz, H1′ of C), 5.3 (d, 1H, 3′-OH of T), 4.9 (m, 1H, H3′ of C), 4.3 (m, 1H, H3′ of T), 4.15 (m, 1H, H4′ of C), 4.10 (m, H4′ of T), 3.8 (s, 6H O-CH_3), 3.55 (m, 2H, H5′ of T), 3.35 (m, 2H, H5′ of C), 2.7 (m, 1H, H2′ of C), 2.50 (m, 1H, C-H of isobutyryl group), 2.40 (m, 1H, H2′ of T), 2.30 (m, 1H, H2′ of C), 1.90 (s, 3H, CH_3 of T), 1.90 (m, 1H, H2′ of T), 1.4 (d, 3H, P-CH_3), 1.10 (d, 6H, isobutyryl group).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.1(d, H, N-H), 7.6−6.8 (m, 23H, arom), 6.4 (t, 1H, J = 7.5 Hz, H1′ of T), 6.2 (t, 1H, J = 6.6 Hz, H1′ of C), 5.3 (d, 1H, 3′-OH of T), 4.9 (m, 1H, H3′ of C), 4.25 (m, 1H, H3′ of T), 4.15 (m, 1H, H4′ of C), 4.0 (m, H4′ of T), 3.77 (s, 6H O-CH_3), 3.65 (m, 2H, H5′ of T), 3.35 (m, 2H, H5′ of C), 2.8 (m, 1H, H2′ of C), 2.6 (m, 1H, H2′ of T), 2.50 (m, 1H, C-H of isobutyryl group), 2.30 (m, 1H, H2′ of C), 1.90 (s, 3H, CH_3 of T), 1.8 (m, 1H, H2′ of T), 1.4 (d, 3H, P-CH_3), 1.20 (d, 6H, isobutyryl group).

*N^4-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxycytidylyl-(3′→5′)-3′-O-(2-cyanoethyl N,N-diisopropylamidophosphoramidite)thymidine 3′-Methylphosphonate (**17**).* Conversion of the nucleoside dimer of the methylphosphonate d[Cp(CH_3)T]-to the corresponding phosphoramidite was accomplished by a procedure analagous to that described for **5**. From 150 mg (0.17 mmol) of **16**, 120 mg (0.11 mmol) of **17** was obtained (yield of 65% for the R_p diastereomer). A 64% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.33, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.33.

^{31}P NMR of R_p isomer ($CDCl_3$): δ 31.72, 31.88 (methylphosphonate), 149.18, 149.29 (phosphoramidite).

^{31}P NMR of S_p isomer ($CDCl_3$): δ 32.01, 32.15 (methylphosphonate), 149.10, 149.26 (phosphoramidite).

*N^4-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxycytidylyl-(3′→5′)-N^2-isobutyryl-3′-O-tert-butyldiphenylsilyl-2′-deoxyguanosine 3′-Methylphosphonate (**18**).* The corresponding *tert*-butyldimethylsilyl-protected dimer of the methylphosphonate of d[Cp(CH_3)G] was prepared by a procedure analagous to that described for **9** with a yield of 35% for the R_p diastereomer and 22% for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.20, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.17.

MS FAB calculated for $C_{65}H_{75}N_8O_{13}PSiK$ ($M + K^+$): 1273.4598. Found: 1273.4623.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.0 (d H, H of C), 7.6−6.8 (m, 23H, arom), 6.25 (t, 1H, J = 6.0 Hz, H1′ of G), 6.15 (t, 1H, J = 6.0 Hz, H1′ of C), 5.00 (m, 1H, H3′ of

C), 4.39 (m, 1H, H3′ of G), 4.19 (m, 1H, H4′ of C), 4.00 (m, H4′ of G), 3.8 (s, 6H O-CH_3), 3.59 (m, 2H, H5′ of G), 3.30 (m, 2H, H5′ of C), 2.80 (m, 1H, H2′ of G), 2.75 (m, 1H, H2′ of C), 2.65 (m, 1H, C-H of isobutyryl group), 2.25 (m, 1H, H2′ of G), 2.2 (m, 1H, H2′ of C), 1.2−1.4 (d, 14H, isobutyryl group), 1.1 (s, 9H, *tert*-butyl group), 0.95 (d, 3H, P-CH_3).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.0 (d, H, H of C), 7.6−6.8 (m, 23H, arom), 6.25 (t, 1H, J = 5.6 Hz, H1′ of G), 6.20 (t, 1H, J = 5.3 Hz, H1′ of C), 5.11 (m, 1H, H3′ of C), 4.30 (m, 1H, H4′ of C), 4.16 (m, 1H, H3′ of G), 3.8 (s, 6H O-CH_3), 3.66 (m, 1H, H4′ of G), 3.17 (m, 2H, H5′ of G), 3.12 (m, 2H, H5′ of C), 2.88 (m, 1H, H2′ of C), 2.67 (m, 1H, H2′ of G), 2.65 (m, 1H, C-H of isobutyryl group), 2.60 (m, 1H, H2′ of G), 2.23 (m, 1H, H2′ of C), 1.2−1.4 (d, 12H, isobutyryl group), 1.1 (s, 9H, t-butyl group), 0.95 (d, 3H, P-CH_3).

*N^4-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxycytidylyl-(3′→5′)-N^2-isobutyryl-2′-deoxyguanosine 3′-Methylphosphonate (**19**).* Deprotection of the *tert*-butyldimethylsilyl-protected dimer of the methylphosphonate of d[Cp(CH_3)G] was accomplished by a procedure analagous to that described for **10**. From 300 mg (0.24 mmol) of **18**, 110 mg (0.11 mmol) of **19** was obtained (yield of 47% for the R_p diastereomer). A 49% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 90/10%) = 0.45, and R_f of S_p isomer (chloroform/methanol, 90/10%) = 0.45.

MS FAB calculated for $C_{49}H_{57}N_8O_{13}PK$ ($M + K^+$): 1035.3420. Found: 1035.3424.

^{1}H NMR of R_p isomer ($CDCl_3$): δ 8.0 (d, H, H of C), 7.6−6.8 (m, 13H, arom), 6.25 (t, 1H, J = 6.0 Hz, H1′ of G), 6.15 (t, 1H, J = 6.0 Hz, H1′ of C), 5.52 (d, 1H, OH3′ of G) 5.00 (m, 1H, H3′ of C), 4.39 (m, 1H, H3′ of G), 4.19 (m, 1H, H4′ of C), 4.00 (m, H4′ of G), 3.8 (s, 6H O-CH_3), 3.59 (m, 2H, H5′ of G), 3.30 (m, 2H, H5′ of C), 2.80 (m, 1H, H2′ of G), 2.75 (m, 1H, H2′ of C), 2.65 (m, 1H, C-H of isobutyryl group), 2.25 (m, 1H, H2′ of G), 2.2 (m, 1H, H2′ of C), 1.2−1.4 (d, 12H, isobutyryl group), 0.95 (d, 3H, P-CH_3).

^{1}H NMR of S_p isomer ($CDCl_3$): δ 8.0 (d, H, H of C), 7.6−6.8 (m, 23H, arom), 6.25 (t, 1H, J = 5.6 Hz, H1′ of G), 6.20 (t, 1H, J = 5.3 Hz, H1′ of C), 5.55 (d, 1H, OH3′ of G), 5.11 (m, 1H, H3′ of C), 4.30 (m, 1H, H4′ of C), 4.16 (m, 1H, H3′ of G), 3.8 (s, 6H O-CH_3), 3.66 (m, 1H, H4′ of G), 3.17 (m, 2H, H5′ of G), 3.12 (m, 2H, H5′ of C), 2.88 (m, 1H, H2′ of C), 2.67 (m, 1H, H2′ of G), 2.65 (m, 1H, C-H of isobutyryl group), 2.60 (m, 1H, H2′ of G), 2.23 (m, 1H, H2′ of C), 1.2−1.4 (d, 12H, isobutyryl group), 0.95 (d, 3H, P-CH_3).

*N^4-Isobutyryl-5′-O-(4,4′-dimethoxytrityl)-2′-deoxycytidylyl-(3′→5′)-N^2-isobutyryl-3′-O-(2-cyanoethyl N,N-diisopropylamidophosphoramidite)-2′-deoxyguanosine 3′-Methylphosphonate (**20**).* Conversion of the nucleoside dimer of the methylphosphonate d[Cp(CH_3)G] to the corresponding phosphoramidite was accomplished by a procedure analagous to that described for **5**. From 100 mg (0.10 mmol) of **19**, 70 mg (0.058 mmol) of **20** was obtained (yield of 60% for the R_p diastereomer). A 50% yield was obtained for the S_p diastereomer.

R_f of R_p isomer (chloroform/methanol, 95/5%) = 0.35, and R_f of S_p isomer (chloroform/methanol, 95/5%) = 0.35.

^{31}P NMR of R_p isomer ($CDCl_3$): δ 31.36, 31.48 (methylphosphonate), 148.89, 148.95 (phosphoramidite).

^{31}P NMR of S_p isomer ($CDCl_3$): δ 31.88, 32.00 (methylphosphonate), 149.33, 149.38 (phosphoramidite).

DNA Synthesis

The 20-mers containing a single diastereomeric methylphosphonate linkage were prepared by solid phase DNA synthesis under standard conditions (Matteucci & Caruthers, 1981). The dimer phosphoramidite building blocks could be incorporated into DNA strands with coupling efficiencies that were comparable to those of common nucleoside phosphoramidites. Deprotection of the methylphosphonate-containing oligonucleotides was performed as described by Hogrefe et al. (1993).

The fully deprotected sequences were isolated by (20%) polyacrylamide gel electrophoresis in the presence of 7 M urea. The correct band was extracted from the gel and isolated using the "crush and soak" technique with a 0.3 M NaOAc buffer. The eluant was decanted and the gel pieces washed with double-distilled water. The combined eluant was concentrated and desalted using a column of Sephadex G-10.

Nucleoside Analyses

Nucleoside composition was determined after P1 nuclease/snake venom phosphodiesterase/bacterial alkaline phosphatase hydrolysis: a 20 μL reaction mixture containing 0.5 A_{260} unit of oligomer in 25 mM sodium acetate, pH 5.3, was incubated for 1 h at 37 °C with nuclease P1. The reaction mixture was then rebuffered to pH 8 with a solution of 100 mM Tris·HCl, pH 8.0; 10 mM $MgCl_2$, 3 units of snake venom phosphodiesterase, and 2 units of alkaline phosphatase were added and the mixture incubated an additional 2 h at 37 °C. An aliquot containing approximately 0.2 A_{260} unit of material was analyzed by HPLC using a 4.6 × 250 mm column of ODS-Hypersil in 20 mM potassium phosphate, pH 5.5, and a gradient of 0 to 100% methanol (60 min) followed by isocratic elution with solvent B. The following retention times were observed (260 nm): dC, 6.5; dG, 8.4; dT, 9.2; dA, 10.4; R_p d[Ap(CH_3)A], 14.4; S_p d[Ap(CH_3)A], 14.9; R_p d[Ap(CH_3)C], 15.8; S_p d[Ap(CH_3)C], 16.6; R_p d[Cp(CH_3)T], 13.4; S_p d[Cp(CH_3)T], 14.9; R_p d[Gp(CH_3)T], 14.2; S_p d[Gp(CH_3)T], 15.0; R_p d[Tp(CH_3)A], 14.6; S_p d[Tp(CH_3)A], 14.8; R_p d[Cp(CH_3)G], 13.2; S_p d[Cp(CH_3)G], 13.6 min.

T_m Values

T_m values were obtained in 10 mM Tris·HCl (pH 7.4) and 250 mM sodium chloride at duplex concentrations of approximately 1 μM. Absorbance and temperature values were measured with an AVIV 14DS UV−visible spectrophotometer equipped with digital temperature control. The temperature of the cell compartment was increased in 1 °C steps (from 0 to 95 °C), and when thermal equilibrium was reached, temperature and absorbance data were collected. T_m values were determined both from first-order derivatives and by graphical analysis of the absorbance vs temperature plots.

Alkaline Phosphatase Protection Assay

Radioisotopic Labeling. Each 20-mer was 5′-end labeled with [γ-^{32}P]-ATP as follows: A 200 μL reaction mixture contained 0.5 A_{260} unit of 20-mer (∼0.1 mM), 40 mM Tris·-

HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 0.2 mM Na_2-EDTA, 0.005−0.025 mM ATP, ∼50 μCi [γ-^{32}P]ATP, 5 μg/mL BSA, and 10 units of T_4 polynucleotide kinase was incubated overnight at 37 °C. The product was isolated by adsorption on a C_{18} Sep-Pak cartridge. The cartridge was washed with 4% aqueous methanol to remove the excess ATP and then with 20% aqueous methanol to elute the product. The product was lyophylized to dryness and dissolved in 10 mM Tris·HCl, pH 7.4, and 250 mM sodium chloride at a duplex concentration of approximately 0.5 mM. This solution was heated to 80 °C for 5 min, slowly cooled to 5 °C, and finally diluted with buffer to a final duplex concentration of 0.5−1.0 μM.

Alkaline Phosphatase Binding Analysis. The assay as described by Marmorstein et al. (1991) was employed. Some of the details of the procedures are noted below. Both the *trp* repressor and the calf intestine alkaline phosphatase were dialyzed against 10 mM Tris·HCl, pH 7.4, and 250 mM sodium chloride. Approximately 1 mg of repressor in 30 μL of storage buffer (0.5 mM sodium phosphate and 250 mM sodium chloride, pH 7.0) was diluted to a final volume of 500 μl and dialyzed at 4 °C. Three buffer changes were completed at 4 h intervals, and after 16 h, the dialysis was stopped, the protein was transferred to a suitable container, and the final concentration was determined spectrophotometrically at 280 nm ($\epsilon = 1.48 \times 10^4$ M^{-1} cm^{-1}). One hundred units of alkaline phosphatase (1 unit/μL) was also dialyzed under the same conditions.

Alkaline Phosphatase Reaction in the Absence of Repressor. A 200 μL reaction solution at 22 °C containing 20 nM of duplex operator in 10 mM Tris·HCl, pH 7.4, and 250 mM sodium chloride and containing 10 μg/mL BSA was initiated by the addition of 2.5 or 5.0 units of alkaline phosphatase. Aliquots of 20−30 μL were removed at 15 s intervals (usually for a 2 min total period), and the reaction was quenched by addition of the aliquot into 25 μL of 0.15 M potassium phosphate, pH 7.0, containing 100 A_{260} units/mL of bulk tRNA. The nucleic acids were precipitated by the addition of 3 mL of 5% aqueous trichloroacetic acid (TCA) followed by incubation for 5 min at ambient temperature. The TCA solution was filtered through a nitrocellulose membrane, followed by filtration of an additional 3 mL "wash" solution of TCA. The membrane was dried and the radioactivity determined using scintillation fluid.

Alkaline Phosphatase Reaction in the Presence of Repressor. The alkaline phosphatase reaction in the presence of the *trp* repressor was performed by essentially the same procedure with the following exceptions. The 200 μL reaction solution contained 0.1, 1.0, or 10 μM repressor (dimer) and 0.5 mM L-tryptophan corepressor, and this mixture was incubated 30 min at 22 °C prior to the addition of the phosphatase. Control reactions were also performed in absence of the corepressor. The time intervals for removing aliquots of solution for precipitation and analysis varied depending upon the rate of reaction.

Data Analysis. The amount of radiolabeled duplex present at various time periods, determined from the precipitated radioactivity (as a ratio of the starting labeled duplex) was plotted logarithmically as a function of reaction time as described previously. In some reactions with the repressor, a biphasic reaction rate appeared to be present: an initial rapid dephosphorylation occurred, similar to the rate of the alkaline phosphatase reaction in the absence of the repressor, followed by a second slower rate. The initial rate accounted for approximately 10−15% of the labeled oligodeoxynucleotide and likely reflected the presence of nonduplex structures (e.g., hairpins). The presence of such structures is not generally considered in assays using self-complementary oligodeoxynucleotides, but are likely to occur, particularly when the concentration of the oligodeoxynucleotide is relatively low. In cases where a biphasic reaction rate appeared to be present, the second portion of the rate curve was used to extrapolate the concentration of "active" duplex 20-mer present when the reaction was initiated. Rate constants were obtained from the plots of the alkaline phosphatase reaction with (k_i') and without (k_i) repressor (R), and an apparent dissociation constant (K_D) was obtained from the relationship $K_D = [R]/(k_i/k_i' - 1)$, as has been described previously (Marmorstein et al., 1991). The K_D values reported are the result of a minimum of four independent assays and the standard deviations reported reflect deviations in these independent assays.

RESULTS

Each internucleotide phosphodiester linkage in a DNA sequence contains two prochiral unesterified phosphate oxygens (Figure 1a). One approach to probe interactions to a specific phosphate oxygen at such a prochiral site would be the stereospecific deletion of one of the oxygens atoms (and its replacement with hydrogen) by the introduction of a chiral H-phosphonate linkage (Figure 1b). Selected H-phosphonate derivatives have been used in one study (Jeltsch et al., 1993), but they are generally not stable to the conditions of DNA synthesis and deprotection (Zon & Stec, 1991). An alternative linkage that could alter the nature of the interactions to the phosphodiester linkage would be the chiral phosphorothioate (Figure 1c), and such chiral derivatives have been used in recent studies (Koziolkiewicz & Stec, 1992; Lesser et al., 1992; Kurpiewski et al., 1996; Thorogood et al., 1996). However, much of the charge associated with a phosphorothioate diester appears to be localized on the sulfur atom, and this functional group can in principle still participate in hydrogen bonding or ionic interactions. We wanted to employ an analogue that could have the effect of eliminating a specific interaction to one of the nonbonded oxygens of the phosphate diester. For this study, we chose a third possibility, that of the methylphosphonate derivatives in either the R_p or S_p configuration. For a given interaction involving an internucleotide phosphodiester, one methylphosphonate diastereomer offers a neutral oxygen as a potential hydrogen-bonding site, while the second diastereomer provides only the nonbonding methyl group. Two recent studies have employed sequences containing methylphosphonate linkages to probe for essential contacts in either a DNA−protein (Botfield & Weiss, 1994) or an RNA−protein (Pritchard et al., 1994) complex and an earlier study with the *lac* repressor involved a set of chiral derivatives (Noble et al., 1984).

Synthesis and Characterization of the Analogue Operator Sequences. A single racemic methylphosphonate linkage can be introduced into a DNA sequence (Botfield & Weiss, 1994) using a solid-phase-based protocol (Miller et al., 1991) and the methylphosphonamidite nucleoside building block (Dorman et al., 1984; Jager & Engels, 1984; Agrawal & Goodchild, 1987). However, the resulting diastereomeric mixture can only be resolved in the case of relatively short sequences. Recent reports suggest improvements in the use

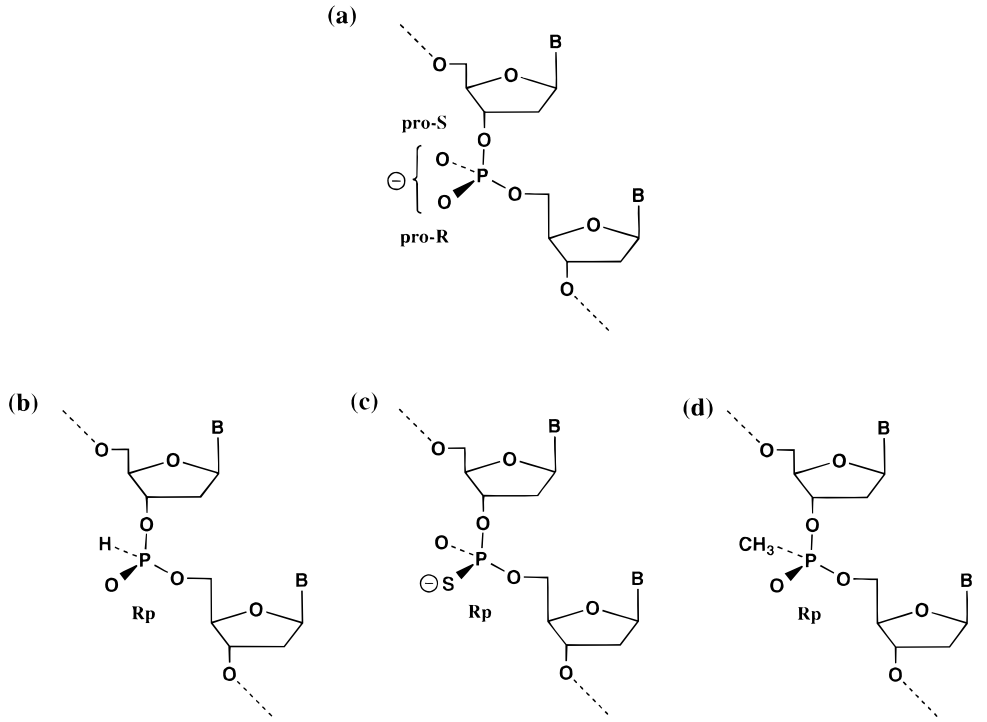
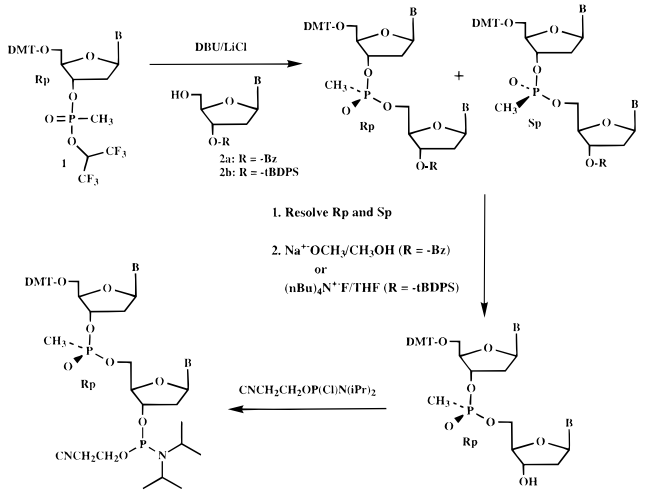


FIGURE 1: (a) A prochiral phosphodiester illustrated the absolute positions of the *pro-R* and *pro-S* nonbridging oxygens. Three phosphodiester analogues are illustrated: (b) an H-phosphonate, (c) a phosphorothioate, and (d) a methylphosphonate.

of chiral reagents as a route to the introduction of stereospecific methylphosphonate linkages by solid-phase-based protocols (LeBec & Wickstrom, 1994, 1996; Rosmanitz et al., 1994), but such procedures are not yet routine. The synthesis of chiral dinucleoside methylphosphonates has been described in a number of studies in which that racemic linkage is prepared and the two diastereomers resolved (Loschner & Engels, 1989; Vyazovkina et al., 1993; Rosmanitz et al., 1994). It should be possible using such an approach to prepare dinucleoside building blocks as stereospecific methylphosphonates and incorporate these chiral dimers into DNA sequences at selected sites. A similar approach has been used for the preparation of sequences containing chiral phosphorothioate linkages (Connolly et al., 1984; Stec et al., 1984; Fidanza et al., 1992) and has also been described for a single example of a sequence containing a methylphosphonate linkage (Schweitzer et al., 1995).

Each diastereomeric mixture of methylphosphonate dimers was prepared from the known hexafluoroisopropyl ester derivatives (Cormier & Pannunzio, 1991) by reaction with the 3′-protected nucleoside in the presence of DBU/LiCl (Scheme 1). Although this reaction proceeds with a high degree of stereospecificity, we always observed the presence of small quantities of the second diastereomer in reactions in which a stereochemically pure hexafluoroisopropyl ester was employed. Because these reactions were not absolutely stereospecific in our hands, it was more effective to use the racemic hexafluoroisopropyl esters in the coupling reactions and resolve the individual diastereomers by chromatography after dimer formation. Stereochemical assignments of the isolated dimers were based upon TLC and ^{31}P NMR characteristics as has been described elsewhere (Stec et al., 1985; Seela & Kretschmer, 1991). Two methylphosphonate dimers have been crystallized to date (Chacko et al., 1983; Han et al., 1990), 2-D NMR ROESY spectra have also been used (Löschner & Engels, 1990) to make assignments for the absolute stereochemistry, and both been used to correlate

Scheme 1



HPLC retention times for fully deprotected dimers. In all cases, the R_p diastereomer is the "faster" eluting species from the reversed phase HPLC column. Although there has been some confusion regarding this property for the d[ApCH$_3$A] dimer (Kan et al., 1980), more recent work (Löschner & Engels, 1990) has corrected the original stereochemical assignments for this dimer. After the coupling procedure, the faster moving diastereomer was isolated and correlated with the faster moving species by HPLC and was thereby assigned the R_p configuration. This analysis is consistent with that described previously (Seela & Kretschmer, 1991). The ^{31}P resonances for these faster moving species were all observed to be upfield from the corresponding "slower" migrating species, also consistent with stereochemical assignments as the R_p diastereomers (Lesnikowski et al., 1988; Loschner & Engels, 1989; Seela & Kretschmer, 1991). In the coupling reactions, the benzoyl group was used initially as the 3′-protecting group, but during the subsequent deprotection reaction, cleavage of the benzoyl ester resulted in the loss of the base-protecting group in dimers containing

Table 1. Binding of the *trp* Repressor to Analogue Operator Sequences Containing Diastereomeric Methylphosphonate Substitutions

-10 -5 -1 +1 +5 +10
CpGpTpApCpTpApGpTpTpApApCpTpApGpTpApCpG
GpCpApTpGpApTpCpApApTpTpGpApTpCpApTpGpC
+10 +5 +1 -1 -5 -10

entry	analogue[a]	flanking positions[b]	T_m (°C)	apparent K_D (1.0 μM *trp*)	apparent $\Delta\Delta G_{22}$ (kcal/mol)[c]	predicted interactions[d]
1	native sequence		65.0	0.024	−	
2	Ap(CH_3)A(R_p)	+1/+2	67.5	0.60 ± 0.12	0.96	*pro-S*−Thr_{44}
3	Ap(CH_3)A(S_p)	+1/+2	65.9	0.075 ± 0.02	0.34	*pro-R*−Arg_{84}
4	Ap(CH_3)C(R_p)	+2/+3	66.9	0.30 ± 0.04	0.75	*pro-S*−trp (H_2O)
5	Ap(CH_3)C(S_p)	+2/+3	64.9	0.69 ± 0.02	1.00	*pro-R*−Arg_{54} (H_2O)
6	Cp(CH_3)T(R_p)	+3/+4	67.5	0.032 ± 0.003	−[e]	
7	Cp(CH_3)T(S_p)	+3/+4	60.5	0.54 ± 0.08	0.93	*pro-R*−Gly_{78}/Thr_{81}
8	Gp(CH_3)T(R_p)	−3/−2	66.2	0.026 ± 0.008	−	
9	Gp(CH_3)T(S_p)	−3/−2	63.6	0.023 ± 0.004	−	
10	Tp(CH_3)A(R_p)	−8/−7	63.5	0.12 ± 0.02	0.47	
11	Tp(CH_3)A(S_p)	−8/−7	63.0	0.26 ± 0.05	0.71	*pro-R*−Asn_{87}
12	Gp(CH_3)T(R_p)	−9/−8	65.9	1.30 ± 0.40	1.19	*pro-S*−Lys_{90}
13	Gp(CH_3)T(S_p)	−9/−8	63.8	1.35 ± 0.40	1.20	*pro-R*−Gln_{68}/Ser_{86}
14	Cp(CH_3)G(R_p)	−10/−9	62.0	0.11 ± 0.04	0.45	
15	Cp(CH_3)G(S_p)	−10/−9	64.1	1.22 ± 0.01	1.17	*pro-R*−Gln_{68}

[a] Sequence indicates the methylphosphonates dimer incorporated into the 20-mer operator sequence. The sites replaced with methylphosphonates are noted with asterisks in the sequence of the heading. [b] Numbers indicate the nucleoside positions flanking the methylphosphonate linkage. [c] ΔΔG values are calculated from the apparent native and analogue K_D values and reported per operator half-site. [d] Predicted interactions based upon the crystal structure analysis (Otwinowski et al., 1988). Note: the R_p methylphosphonate places the methyl group of the linkage into the space normally occupied by the *pro-S* oxygen of the phosphodiester linkage. [e] The − notations indicate that the $\Delta\Delta G$ value per half-site was less than 0.1 kcal/mol.

2′-deoxycytidine. Therefore, after preparation of the diastereomers of d[Ap(CH_3)A] and d[Tp(CH_3)A] using 3′-*O*-benzoyl protection, the remaining four pairs of diastereomers were prepared using 3′-*O*-*tert*-butyldiphenylsilyl protection. After removal of the 3′-protecting group, each stereochemically pure dimer was converted to the corresponding phosphoramidite as illustrated in Scheme 1 for a methylphosphonate of the R_p configuration.

Oligonucleotides of 20 residues were prepared using standard phosphoramidite protocols. At the desired site, the methylphosphonate dimer of the appropriate sequence and stereochemistry could be incorporated into the DNA sequence as a phosphoramidite building block with no detectable change in coupling yield. After assembly, each 20-mer was deprotected in the described fashion (Hogrefe et al., 1993), purified by PAGE and stored at −20 °C. Confirmation of the stereochemical integrity of each methylphosphonate linkage was obtained by HPLC analysis of an enzyme digest of each analogue 20-mer sequence. The methylphosphonate linkage is refractory to cleavage by nucleases, so after digestion of each methylphosphonate-containing sequence, the methylphosphonate nucleoside dimer could be identified (Figure 2). As illustrated in Figure 2 for two diastereomeric methylphosphonate sequences, the two diastereomeric methylphosphonate dimers are eluted from the column after the four common 2′-deoxynucleosides with varying retention times. In each analysis, we could confirm that only one of the two possible diastereomeric linkages was present based upon the elution of a single methylphosphonate diastereomer.

Initial characterization of the 20-mer operator sequences was by a simple measurement of T_m values. These values were obtained in the buffer/salt conditions used for analysis of repressor binding. For virtually all of the sequences in this study, the introduction of two symmetrical methylphosphonate linkages (one in each half-site) had a measurable effect on the T_m for the 20-mer sequence. In all but one case (see entry 14, Table 1), the R_p diastereomers exhibited T_m values that were higher than those observed for the S_p

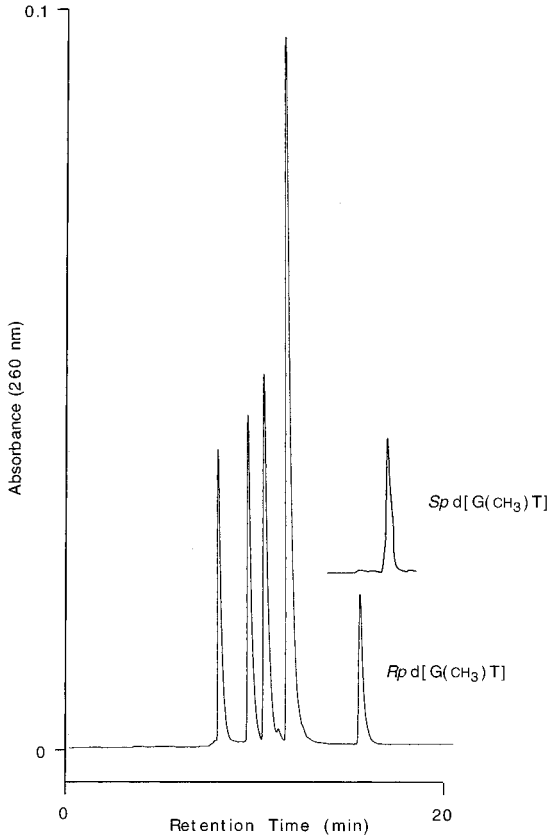


FIGURE 2: HPLC analysis of the digest (nuclease P1 followed by snake venom phosphodiesterase/alkaline phosphatase) of two diastereomeric operator sequences (entries 8 and 9, Table 1) containing either an R_p or an S_p methylphosphonate linkage. Peak identification from left to right is dC, dG, T, dA, and R_p d[G(CH_3)T]. Partial chromatogram illustrates the elution of the S_p d[G(CH_3)T] derivative from the diastereomeric methylphosphonate sequence.

diastereomers (Table 1). Additionally, the effect on T_m appeared more significant for linkages present nearer the center of the operator. For example, the R_p diastereomers of entries 2, 4, 6, 8, and 12 (Table 1) all exhibited higher T_m values than that obtained for the native sequence, while those of entries 10 and 14 had lower T_ms than did the native sequence. But whether this effect is positional in nature or simply related to the flanking nucleosides is beyond the scope of this study. The S_p diastereomers generally exhibited T_m values that were lower than that of the native, but two examples, entries 3 and 5 (Table 1), were equivalent or slightly higher than the value obtained for the native sequence.

Protein Binding Studies. We have used an enzyme protection assay described in other studies (Marmorstein et al., 1991; Mazzarelli et al., 1992) to measure relative binding constants for the native and analogue operator sequences. The assay relies upon binding by the *trp* repressor to the 20-mer operator sequence to protect a terminal phosphate from cleavage by alkaline phosphatase. By comparing the rate of phosphatase activity in the absence and in the presence of the repressor, an apparent dissociation constant can be obtained based upon substrate depletion kinetics. The results from this assay are quite sensitive to the conditions of the assay, particularly the concentration of the repressor. At 0.1 μM repressor, an apparent K_D of 3 nM is obtained, a value quite similar to that obtained by gel retardation assays using longer sequences (Carey, 1988). However, at this concentration of repressor, poorly bound operator sequences are not observed to provide any protection in the assay and relative binding affinities cannot be obtained. At higher concentrations of repressor (1.0 μM Trp), the apparent K_D for the native sequence artificially rises to 24 nM, but relative differences in binding effects for all sequences could be obtained (Table 1). In general, this assay is quite effective for measuring relative binding affinity for poorly bound complexes but is less effective in discriminating differences for tightly bound complexes. We chose to use this assay since the interactions to the negatively charged phosphodiesters were expected to provide significant contributions to complex stability such that loss of single contact could result in a very poorly bound complex. We also tried to employ gel retardation analyses (Carey, 1988), but this assay did not function well with the relatively short fragments used in this study. The alternative, a filter binding assay, was complicated by the observation that the radiolabeled DNA of the repressor−operator complex was not effectively retained on the filter.

In many cases, the presence of the diastereomeric methylphosphonate linkages resulted in variable binding by the *trp* repressor. For the d[Ap(CH_3)A] dimer, the S_p diastereomer exhibited a K_D that was near that of the native sequence, while the R_p diastereomer resulted in a K_D ca. 25-fold greater than that measured for the native operator (entries 2 and 3, Table 1). By comparison, the R_p diastereomers of the d[Cp(CH_3)T], d[Tp(CH_3)A], and d[Cp(CH_3)G] analogue sequences (entries 6, 10, and 14) resulted in dissociation constants more or less similar to that of the native operator ($\leq$5-fold), while the K_D values measured for the S_p diastereomers (entries 7, 11, and 15) were 22-, 12-, and 50-fold greater, respectively. In other cases, the two diastereomeric sequences were more similar in effects. Both diastereomers of d[Gp(CH_3)T] at sequence position −2/−3 (entries 8 and 9) exhibited K_D values that were indistinguishable from that of the native operator sequence, while the same linkage at sequence position −8/−9 (entries 12 and 13) resulted in the two poorest binding analogue operators of this study, with K_D values ca. 54-fold larger than that observed for the native sequence. Both diastereomers of the d[Ap(CH_3)C] (entries 4 and 5) analogues exhibited dissociation constants that were significantly greater than that of the native sequence, varying from 12- to 28-fold larger, respectively.

A comparison of the results for the native and the various analogue operator sequences permits an estimation of incremental binding energy ($\Delta\Delta G$) for the functional group eliminated by the substitution of the methylphosphonate analogue. For a number of analogue sequences (entries 2, 5, 7, 12, 13, and 15), this value approaches or exceeds 1 kcal/mol per half-site. Three analogue sequences (entries 6, 8, and 9) have values that are essentially identical with the native sequence, and $\Delta\Delta G$ values of $<$0.1 kcal/mol per half-site are obtained. The remaining five operator sequences (entries 3, 4, 10, 11, and 14) result in $\Delta\Delta G$ values that vary from 0.34 to 0.75 kcal/mol per half-site.

DISCUSSION

Nucleoside dimers containing a methylphosphonate derivative of known chirality can be converted to a phosphoramidite building block and incorporated into DNA sequences without loss of stereochemical integrity. The observation that only a single diastereomeric dimer is present in each case after digestion of the analogue operator sequences (Figure 2) confirms that the linkage does not undergo racemization during assembly, deprotection, or isolation procedures. The relative retention times obtained for these dimers after digestion of the 20-mer sequences are consistent with other work in which only the simple dimers have been prepared (Stec et al., 1985). This approach provides a general route for the introduction of chiral methylphosphonates at predetermined sites in DNA sequences. While unlikely to be efficient for sequences composed entirely of methylphosphonate linkages, it is effective for the production of sequences that are incrementally altered at relatively few sites.

Substitution of a methylphosphonate for a phosphodiester both introduces an additional chiral site into the sequence and eliminates one negative charge. The interaction between an amino acid side chain and a given prochiral nonbonded oxygen can, in principle, be affected by one or both changes in the nature of the linkage. Substitution of one of the methylphosphonate diastereomers for the prochiral phosphodiester linkage replaces one of the partially charged nonbonded oxygens with a methyl group, and this substitution should completely eliminate any hydrogen-bonding interactions at this prochiral site. By comparison, the complementary diastereomer replaces the partially charged oxygen of the phosphodiester with the neutral oxygen of the methylphosphonate linkage. In this case, the nature of the hydrogen-bonding interaction has been modified somewhat in that neither partner is now charged, and this distinction could contribute to lower overall binding energetics. On the other hand, the overall energetics of hydrogen bonding in water arise largely from entropy effects, both the solvated protein and the solvated DNA release hydrogen-bonded water into bulk solution to drive complex formation (Fersht et al., 1985; Fersht, 1987). This latter effect could be expected to be similar for the phosphodiester linkage and the appropriate

diastereomer of the methylphosphonate linkage that orients the oxygen toward the hydrogen bond donor. If a hydrogen-bonding interaction to a neutral methylphosphonate oxygen has similar overall energetics to that of a partially charged prochiral phosphate oxygen, then little difference in binding affinity should be observed when the correct methylphosphonate diastereomer (offering the P═O hydrogen-bonding site) is present at a position where only one of the prochiral oxygens is in contact with the protein. By comparison, incorporation of the isomeric methylphosphonate should completely disrupt such an interaction, and while the overall differences in energetics may vary significantly, the limiting value for the loss of a critical interactions can be assigned as 0.5 kcal/mol, the minimum value observed in other studies for the loss of a hydrogen bond between neutral partners (Fersht et al., 1985; Fersht, 1987b). Charge effects can increase this value substantially, and it has been previously estimated that the contribution to overall binding energetics by a single phosphate contact is as much as 1.3 kcal/mol (Lohman et al., 1980; Jen-Jacobson et al., 1986). These limiting values simply set a criterion for assessment of the variations in binding affinity by the operator analogues.

A comparison of the solution binding data obtained in this study with the predicted interactions based upon single crystal X-ray diffraction analysis helps to relate the two different approaches to analyze specific intermolecular contacts present at the interface of the protein−nucleic acid complex. In this analysis it must be noted that due to the vagaries in the Cahn−Ingold−Prelog system for the assignment of stereocenters, the R_p methylphosphonate replaces the *pro-S* oxygen of the phosphodiester with a methyl group and retains an oxygen in the *pro-R* position (see Figure 1). Similarly, the S_p methylphosphonate replaces the *pro-R* phosphate oxygen with a methyl group. Although both reported crystal structures (Otwinowski et al., 1988; Lawson & Carey, 1993) suggest that the protein makes very similar contacts to the operator sequence in both the dimeric and tetrameric forms, the latter structure employs a sequence containing a central dyad about which the two dimers are bound. This format for binding is not possible with the operator employed in the current study (containing two half-sites), so the following discussion will primarily relate the observations of the current work to the originally published dimer structure (Otwinowski et al., 1988).

The two analogue operator sequences that exhibited the poorest binding of any of the sequences examined in this study are those that involve the phosphodiester between G_{-9} and T_{-8} (entries 12 and 13, Table 1). Regardless of which methylphosphonate isomer is introduced at this site, the dissociation constant increases some 50-fold with a change in incremental binding energy of 1.2 kcal/mol in each half-site (2.4 kcal/mol for the complex). These values are consistent with the estimates of binding energetics for individual phosphate interactions (Lohman et al., 1980; Jen-Jacobson et al., 1986). These observations suggest that in the presence of either methylphosphonate diastereomer there is a disruption in critical interactions to this linkage. One explanation for this result is that the loss of the charged partner from the critical interaction, regardless of the diastereomeric character, significantly impacts binding affinity. But as can be observed in Table 1, and as will be noted below, there are a number of cases where differential binding effects are present. With one methylphosphonate isomer nearly native-like, binding occurred while with the other a significant loss of affinity was observed. The alternate and simpler interpretation for the results with this analogue is that both of the prochiral oxygens are involved in complex stabilizing contacts with the repressor. In fact, the crystal structure analysis (Otwinowski et al., 1988) suggests that both of the nonbonded oxygens of this phosphodiester are involved in direct contacts with the repressor. The *pro-R* oxygen interacts with both the amide nitrogen of Gln_{68} as well as with the hydroxyl of Ser_{86}, while the *pro-S* oxygen appears to contact the terminal amino group of Lys_{90}.

We cannot exclude from this first example of analogue-binding data that part of the loss of binding affinity arises from the introduction of the neutral linkage into the operator sequence with resultant detrimental nonspecific electrostatic effects. To examine this possibility, we introduced a methylphosphonate linkage, using the same dimer building block, d[Gp(CH_3)T], between the residues G_{-3} and T_{-2}. This site is near the center of the operator sequence, and both crystal structure analyses (Otwinowski et al., 1988; Lawson & Carey, 1993) have indicated the absence of any significant interactions to this residue in either half-site. The solution-binding study (entries 8 and 9, Table 1) indicated that both analogue operators were bound by the repressor with an affinity that was indistinguishable from the native sequence. These observations suggest that the reduction in binding affinity observed when the R_p or the S_p analogue linkage was placed between G_{-9} and T_{-8} more likely reflects a disruption of contacts to either of the prochiral oxygens rather than electrostatic effects resulting from the loss of charge, and this observation is consistent with the predictions of the crystal structures.

At two other sites, the substitution of either methylphosphonate diastereomer affected binding affinity in a similar manner. When placed between A_{+2} and C_{+3}, the measured K_D values increased by 12- and 20-fold for the R_p and S_p methylphosphonate analogues, respectively (entries 4 and 5, Table 1). These observations again suggest that both prochiral oxygens may make critical interactions with the repressor, although the $\Delta\Delta G$ values obtained are somewhat below the estimated values (0.75 and 1.00 kcal/mol). The crystal structure analysis (Otwinowski et al., 1988) implicates two water molecules that bridge each of the prochiral phosphate oxygens: one to the hydroxyl of Thr_{53} and the second to the hydroxyl of Thr_{81}. Additionally, two direct contacts are present involving Arg_{54} (to the *pro-R* oxygen) and the ring nitrogen of the tryptophan cofactor (interacting with the *pro-S* oxygen). The significant changes in binding affinity for either methylphosphonate diastereomer are consistent with the loss of critical contacts to one or the other of the prochiral phosphate oxygens with either isomeric methylphosphonate substitution and is consistent with the crystal structure analyses.

The remaining analogue operators exhibit diastereomeric variability in their interactions with the repressor. For example, substitution of the linkage between C_{+3} and T_{+4} with the diastereomeric methylphosphonates results in two quite different effects (entries 6 and 7, Table 1). The R_p diastereomer has binding properties that do not differ from those of the native sequence. By comparison, the S_p diastereomer results in over a 20-fold increase in the measured K_D value with a $\Delta\Delta G$ value near 1 kcal/mol. These results, in contrast to those noted in the above examples, suggest that the *pro-R* phosphate oxygen makes

a critical contact to the repressor protein, while the *pro-S* oxygen is not involved in any significant complex-stabilizing interactions (note: the R_p methylphosphonate maintains the oxygen of the P═O hydrogen bond acceptor in the *pro-R* position, see Figure 1). In the crystal structure analysis (Otwinowski et al., 1988), both the peptide bond nitrogen of Gly_{78} and the hydroxyl of Thr_{81} are both interactions with the *pro-R* oxygen of the phosphate residue between C_{+3} and T_{+4}. The *pro-S* oxygen does not appear to be involved in any contact with the repressor.

A second site with variable binding effects dependent upon the diastereomeric character of methylphosphonate is the linkage between C_{-10} and G_{-9}. As with the previous example, varying binding affinities are observed with the isomeric analogues (entries 14 and 15, Table 1). The R_p diastereomer exhibits a slight decrease in affinity (K_D increases ∼4-fold) while the S_p diastereomer exhibits a K_D value ca. 50-fold greater than that of the native sequence (0.45 and 1.17 kcal/mol per interaction, respectively). The crystal structure analysis (Otwinowski et al., 1988) implicates only a single interaction at this site, the amide nitrogen of Gln_{68} contacts the *pro-R* phosphate oxygen.

A third example of diastereomeric methylphosphonate linkages resulting in variable binding affinities occurs with the linkage between T_{-8} and A_{-7}. The presence of the R_p diastereomer results in a K_D that is increased 5-fold, while the S_p diastereomer has a slightly larger effect with a 10-fold increase in K_D. While there are qualitative differences in the results for the two diastereomeric methylphosphonate analogues, the quantitative $\Delta\Delta G$ values fall below the expected value in both cases. In the crystal structure (Otwinowski et al., 1988), contact to this phosphodiester is implicated by only a single interaction with the protein. The amide nitrogen of Asn_{87} appears to contact the *pro-R* phosphate oxygen of the operator sequence, while no interactions appear to involve the *pro-S* oxygen. It is noteworthy that introduction of the methyl group at the *pro-S* site (entry 10) at this linkage, as well at the corresponding substitution between C_{-10} and G_{-9} (entry 14), both result in moderate changes in binding affinity and nearly a $^1/_2$ kcal/mol incremental change in free energy. These analogues suggest that secondary effects induced by the analogue in either the structure of the DNA, or at the interface between the DNA and the protein, may negatively impact binding in a moderate way, although the details of such an effect are at present difficult to describe.

In the three examples noted here, variable binding by the isomeric methylphosphonate linkages permits, in each case, the identification of one of the prochiral phosphate oxygens as a critical contact for solution binding. These identified phosphate oxygens are in agreement with those implicated by the crystal structure analyses. The final analogue operator, that containing the diastereomeric methylphosphonate linkage between A_{+1} and A_{+2}, also exhibits differential binding effects (entries 2 and 3, Table 1). The R_p methylphosphonate results in a 25-fold decrease in binding affinity, while the S_p diastereomer exhibits nearly native-like binding (within 3-fold of the unmodified operator). These results suggest, in contrast with the previous three examples, the presence of a critical contact to the *pro-S* phosphate oxygen and the absence of such interactions with the *pro-R* oxygen. Two amino acids from the repressor have been implicated (Otwinowski et al., 1988) in interactions at this site: one involves the hydroxyl of Thr_{44} and the *pro-S* oxygen while a second to the *pro-R* oxygen has been suggested to involve the terminal guanidinium nitrogen of Arg_{84}. The contacts appear to be made to both of the prochiral oxygens of this linkage, yet differential binding affinities are observed in solution. One explanation for this slight discrepancy in suggested critical contacts is that the hydrogen-bonding interaction implicated between the *pro-R* phosphate oxygen and the guanidinium nitrogen of Arg_{84} is in excess of 3 Å in length. This length may reduce the contribution of this interaction significantly, such that the solution binding assay does not distinguish a contribution from Arg_{84}. Alternatively, with the interaction disrupted by the presence of a methyl group, the arginine at this site may be able to take part in an alternative binding interaction, which would modulate the apparent loss in binding energy, as observed.

CONCLUSIONS

DNA sequences containing diastereomeric methylphosphonate linkages can be prepared from the stereochemically pure dimer building block, and these analogue sequences can be used effectively to probe for interactions to either of the prochiral phosphate oxygens. These analogue linkages have been employed to probe critical interactions in the *trp* repressor−operator complex. In a number of instances, the diastereomeric analogue sequences exhibit variable binding affinities that can be used to identify one of the prochiral phosphate oxygens as a critical site for complex-stabilizing interactions, while the isomeric nonbonded oxygen remains unimportant for high-affinity binding. In all but one example, these identified sites for interactions can be correlated with contacts implicated by the crystal structure analysis of the *trp* repressor−operator complex.

REFERENCES

Agrawal, S., & Goodchild, J. (1987) *Tetrahedron Lett. 28*, 3539−3542.

Aiken, C. R., & Gumport, R. I. (1991) *Methods Enzymol. 208*, 433−457.

Bass, S., Sugiono, P., Arvidson, D. N., Gunsalvus, R. P., & Youderian, P. (1987) *Genes Dev. 1*, 565−572.

Bass, S., Sorrells, V., & Youderian, P. (1988) *Science 242*, 240−245.

Botfield, M. C., & Weiss, M. A. (1994) *Biochemistry 33*, 2349−2355.

Carey, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 975−979.

Carey, J. (1989) *J. Biol. Chem. 264*, 1941−1945.

Carey, J., Lewis, D. E. A., Lavoie, T. A., & Yang, J. (1991) *J. Biol. Chem. 266*, 24509−24513.

Chacko, K. K., Lindner, K., Saenger, W., & Miller, P. S. (1983) *Nucleic Acids Res. 11*, 2801−2814.

Connolly, B. A., Potter, B. V. L., Eckstein, F., Pingoud, A., & Grotjahn, L. (1984) *Biochemistry 23*, 3443−3453.

Cormier, J. F., & Pannunzio, T. (1991) *Tetrahedron Lett. 32*, 7161−7164.

Dorman, M. A., Noble, S. A., McBride, L. J., & Caruthers, M. H. (1984) *Tetrahedron 49*, 95−105.

Fersht, A. (1987) *Trends Biochem. Sci. 12*, 301−304.

Fersht, A., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature 314*, 235−238.

Fidanza, J. A., Ozaki, H., & McLaughlin, L. W. (1992) *J. Am. Chem. Soc. 114*, 5509−5517.

Han, F., Watt, W., Duchamp, D. J., Callahan, L., Kezdy, F. J., & Agarwal, K. (1990) *Nucleic Acids Res. 18*, 2759−2767.

Haran, T. E., Joachimiak, A., & Sigler, P. B. (1992) *EMBO J. 11*, 3021−3030.

Hogrefe, R. I., Vaghefi, M. M., Reynolds, M. A., Young, K. M., & Arnold, L. J., Jr. (1993) *Nucleic Acids Res. 21*, 2031−2038.

Iwai, S., Maeda, M., Shimada, Y., Hori, N., Murata, T., Morioka, H., & Ohtsuka, E. (1994) *Biochemistry 33*, 5581−5588.
Jager, A., & Engels, J. (1984) *Tetrahedron Lett. 25*, 1427−1440.
Jeltsch, A., Alves, J., Wolfes, H., Maas, G., & Pingoud, A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8499−8503.
Jen-Jacobson, L., Lesser, D., & Kurpiewski, M. (1986) *Cell 45*, 619−629.
Joachimiak, A., Haran, T. E., & Sigler, P. B. (1994) *EMBO J. 13*, 367−372.
Kan, L. S., Cheng, D. M., Miller, P. S., Yano, J., & Ts'o, P. O. P. (1980) *Biochemistry 19*, 2122−2132.
Kelley, R. L., & Yanofsky, C. (1982) *Proc. Natl. Acad. Sci. U.S.A. 82*, 483−487.
Koziolkiewicz, M., & Stec, W. J. (1992) *Biochemistry 31*, 9460−9466.
Kurpiewski, M. R., Koziolkiewicz, M., Wilk, A., Stec, W. J., & Jen-Jacobson, L. (1996) *Biochemistry 35*, 8846−8854.
Lawson, C. L., & Carey, J. (1993) *Nature 366*, 178−182.
LeBec, C., & Wickstrom, E. (1994) *Tetrahedron Lett. 35*, 9525−9528.
LeBec, C., & Wickstrom, E. (1996) *J. Org. Chem. 61*, 510−513.
Lesnikowski, Z. J., Jaworska, M., & Stec, W. J. (1988) *Nucleic Acids Res. 16*, 11675−11689.
Lesser, D. R., Grajkowski, A., Kurpiewski, M. R., Koziolkiewicz, M., & Jen-Jacobson, L. (1992) *J. Biol. Chem. 267*, 24810−24818.
Lohman, T. M., de Haseth, P. L., & Record, M. T., Jr. (1980) *Biochemistry 19*, 3522−3530.
Loschner, T., & Engels, J. (1989) *Tetrahedron Lett. 30*, 5587−5590.
Löschner, T., & Engels, W. (1990) *Nucleic Acids Res. 18*, 5083−5088.
Marmorstein, R. Q., Sprinzl, M., & Sigler, P. B. (1991) *Biochemistry 30*, 1141−1148.
Matteucci, M. D., & Caruthers, M. H. (1981) *J. Am. Chem. Soc. 103*, 3185−3191.
Mazzarelli, J. M., Rajur, S. B., Iadarola, P. I., & McLaughlin, L. W. (1992) *Biochemistry 31*, 5925−5936.
Miller, P. S., Cushman, C. D., & Levis, J. T. (1991) in *Oligonucleotides and Analogues: A Practical Approach* (Eckstein, F., Ed.) pp 137−154, IRL Press, Oxford.
Noble, A., Fisher, E. F., & Caruthers, M. H. (1984) *Nucleic Acids Res. 12*, 3387−3404.
Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321−329.
Pritchard, C. E., Grasby, J. A., Hamy, F., Zacharek, A. M., Singh, M., Karn, J., & Gait, M. (1994) *Nucleic Acids Res. 22*, 2592−2600.
Rosmanitz, P., Eisenhardt, S., Bats, J. W., & Engels, J. W. (1994) *Tetrahedron 50*, 5719−5734.
Schweitzer, M., Samstag, W., & Engels, J. W. (1995) *Nucleosides Nucleotides 14*, 817−820.
Seela, F., & Kretschmer, U. (1991) *J. Org. Chem. 56*, 3861−3869.
Smith, S., Rajur, S. B., & McLaughlin, L. W. (1994) *Nat. Struct. Biol. 1*, 18−23.
Stec, W. J., Zon, G., Egan, W., & Stec, B. (1984) *J. Am. Chem. Soc. 106*, 6077−6079.
Stec, W. J., Zon, G., Egan, W., Byrd, R. A., Phillips, L. R., & Gallo, K. A. (1985) *J. Org. Chem. 50*, 3908−3913.
Thorogood, H., Grasby, J. A., & Connolly, B. A. (1996) *J. Biol. Chem. 271*, 8855−8862.
Vyazovkina, E. V., Rife, J. P., Lebedev, A. V., & Wickstrom, E. (1993) *Nucleic Acids Res. 21*, 5957−5963.
Waters, T. R., & Connolly, B. A. (1994) *Biochemistry 33*, 1812−1819.
Wozniak, L. A., Pyzowski, J., Wieczorek, M., & Stec, W. J. (1994) *J. Org. Chem. 59*, 5843−5846.
Zhang, R., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., & Sigler, P. B. (1987) *Nature 327*, 591−597.
Zon, G., & Stec, W. J. (1991) in *Oligonucleotides and Analogues: A Practical Approach* (Eckstein, F., Ed.) pp 87−108, Oxford University Press, New York.

BI9700781