136. Triple-Helix Formation by Pyrimidine Oligonucleotides Containing Nonnatural Nucleosides with Extended Aromatic Nucleobases: Intercalation from the Major Groove as a Method for Recognizing C·G and T·A Base Pairs

by **Thomas E. Lehmann, William A. Greenberg, David A. Liberles, Carol K. Wada,** and **Peter B. Dervan***

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

(4.VII.97)

The sequence-specific recognition of double-helical DNA by oligonucleotide-directed triple helix formation is limited primarily to purine tracts. To identify potential lead compounds which are able to extend the sequence repertoire of triple helical complexes, we designed two carbocyclic nucleosides with nucleobases attached *via* amide bonds. N^5-[(1*R*,2*S*,3*R*,4*R*)-3-hydroxy-4-(hydroxymethyl)-2-methoxycyclopentyl]-2-{[1*H*-pyrrol-2-yl)carbonyl]-amino}thiazole-5-carboxamide (**L1**) and 2-benzamido-N^5-[(1*R*,2*S*,3*R*,4*R*)-3-hydroxy-4-(hydroxymethyl)-2-methoxycyclopentyl]thiazole-5-carboxamide (**L2**) were synthesized and incorporated into pyrimidine oligonucleotides. The 2-(trimethylsilyl)ethoxymethyl (SEM) protecting group for the 1*H*-pyrrole NH was found to be compatible with DNA solid-phase synthesis of pyrimidine oligonucleotides. By quantitative DNase I footprinting analysis, both nonnatural nucleosides **L1** and **L2** showed preferential binding of pyrimidine over purine bases: **L1/2** · (C·G) ≈ **L1/2** · (T·A) > **L1/2** · (G·C) ≈ **L1/2** · (A·T). Comparison with the previously reported nonnatural nucleosides with extended aromatic nucleobases 1-(2-deoxy-β-D-ribofuranosyl)-4-(3-benzamidophenyl)-imidazole (**D3**) and N^4-[6-(benzamido)pyridin-2-yl]-2′-deoxycytidine (^{bz}**M**) suggests that the observed binding selectivity C·G ≈ T·A > G·C ≈ A·T for the nucleoside analogs **L1**, **L2**, **D3**, and ^{bz}**M** is derived from sequence-specific intercalation with preferential stacking of their nucleobases over pyrimidine · purine *Watson-Crick* base pairs.

1. Introduction. – Oligodeoxyribonucleotide-directed triple-helix formation is a useful method for the sequence-specific recognition of predetermined sequences of double-helical DNA [1] [2][1]). Triple-helical complexes can be classified into pyrimidine and purine structural motifs, which both bind into the major groove of double-helical DNA. In the pyrimidine motif, pyrimidine-rich oligonucleotides bind in an orientation that is parallel to the purine *Watson-Crick* strand through formation of specific *Hoogsteen* H-bonds to the purine *Watson-Crick* base. Specificity is derived from thymine (T) recognition of adenine thymine (A·T) base pairs (T · (A·T) triplex) and protonated cytosine (C^+) recognition of guanine · cytosine (G·C) base pairs (C^+ · (G·C) triplex) (*Fig. 1*). In the purine motif, purine-rich oligonucleotides bind in an antiparallel orientation with respect to the purine *Watson-Crick* strand through the formation of specific reverse *Hoogsteen* H-bonds to the purine *Watson-Crick* base. Specificity is derived from G recognition of G·C base pairs (G · (G·C) triplex) and A or T recognition of A·T base pairs (A · (A·T) or T · (A·T) triplexes).

[1]) For recent reviews on triple helix formation, see [3–5].

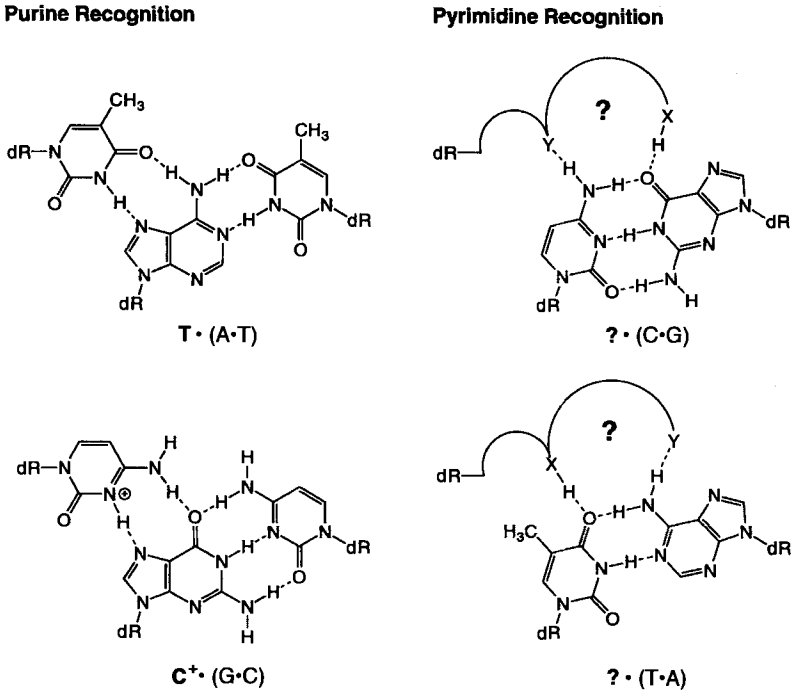


Fig. 1. *Triple-helix-mediated recognition of double-helical DNA in the pyrimidine · purine · pyrimidine motif.* Selective recognition of purine · pyrimidine base pairs is derived from T · (A · T) and C^+ · (G · C) triplex formation. No natural bases code for the recognition of pyrimidine · purine base pairs. dR = 1-(2′-deoxy-β-D-ribofuranosyl).

Due to the selectivity and affinity of triple-helix-directed recognition of DNA, this methodology has found applications in the regulation of gene expression [6–9], in the site-specific cleavage of megabase DNA [10–14], as well as in the sequence-specific bending of DNA [15–17]. Triple-helical complexes are sensitive to single mismatches, which typically result in a lowering of binding affinity of $\sim 10^2$–10^3-fold for a representative 15 base pair binding site [18][19].

Although oligonucleotide-directed triplex formation offers a powerful chemical approach for the sequence-specific recognition of double-helical DNA, target sequences for both the pyrimidine and purine triple-helical motifs are primarily restricted to purine tracts. Efforts directed to the design of nonnatural nucleosides to extend triple-helix methodology towards the recognition of mixed sequences containing all four base pairs have met with limited success [20].

We describe continuing efforts to design heterocycles that recognize C · G or T · A base pairs selectively within the pyrimidine triple helix motif. We sought to engineer a nucleobase that *i*) sterically matches the edges of pyrimidine · purine *Watson-Crick* base pairs in the major groove, *ii*) positions H-bond donors and acceptors to form H-bonds in the major groove with C · G or T · A base pairs, *iii*) maintains a sugar-phosphate backbone geometry compatible with the pyrimidine triple-helix motif, and *iv*) allows energetically favorable stacking of the bases in the third strand.

Molecular-modeling analysis[2]) suggests that the nonnatural nucleosides **L1** and **L2** might each form a selective H-bond interaction with C·G base pairs (*Fig. 2*). For **L1**, the formation of two H-bonds between the carbonyl group of the thiazole-pyrrole-connecting amide bond to the amino group of cytosine and from the pyrrole NH to the carbonyl group of guanine are indicated. In nucleoside **L2** the pyrrole is replaced by a benzene ring; therefore, the formation of one H-bond is anticipated. The attachment of the nucleobase *via* an amide bond to the carbohydrate moiety and the linkage of nucleobase building blocks *via* amide bonds appears desirable for several reasons: *i*) spanning a longer distance from the sugar C(1′) towards the H-bond acceptor and donors on the pyrimidine purine *Watson-Crick* base pair, *ii*) incorporating a planar linker with potentially good stacking as well as H_2O-solvation properties, *iii*) offering a semiflexible element that can adopt a favorable geometry for the interaction with the pyrimidine · purine *Watson-Crick* base pair, and *iv*) allowing a modular, flexible attachment of heterocyclic nucleobase fragments with versatile synthetic accessibility. A carbocyclic ribofuranose analog was chosen in order to make the amino group at C(1′) of the ribose configurationally stable (*Fig. 2*). The replacement of the $O^{4'}$-atom by a CH_2 group in the ribofuranose ring in thymidine (T_d) and 5-methyl-2′-deoxycytidine (Me^5C_d), was shown by *Froehler* and *Ricca* to be compatible with triple-helix formation in the pyrimidine motif [22]. A 2′-MeO group in ribofuranoses in the third strand is beneficial for triple-helix stability in ribofuranoses and does not impose unfavorable steric interactions [23] [24].

2. Synthesis of Carbocyclic Nucleosides. – Enantiomerically pure building block **1** was obtained by diastereoselective OsO_4-catalyzed dihydroxylation of enantiomerically pure (−)-(1*R*,4*S*)-2-azabicyclo[2.2.1]hept-5-en-3-one (*Scheme 1*) [25]. The dihydroxy-lactam **1** was hydrolyzed by refluxing in 3N HCl. The resulting acid was converted to the corresponding methyl ester by refluxing in anhydrous MeOH in the presence of catalytic

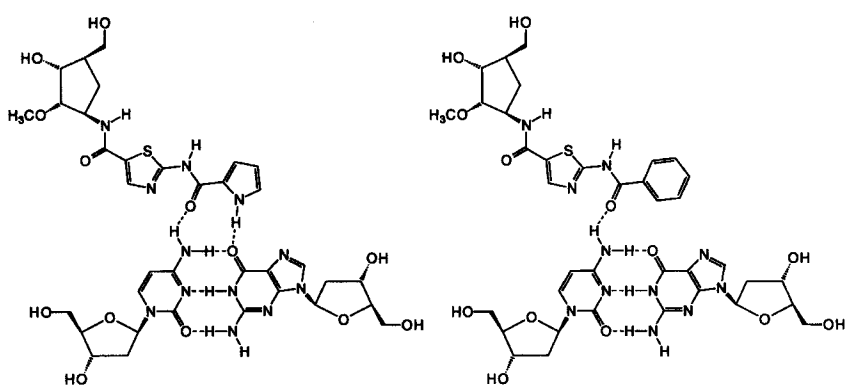
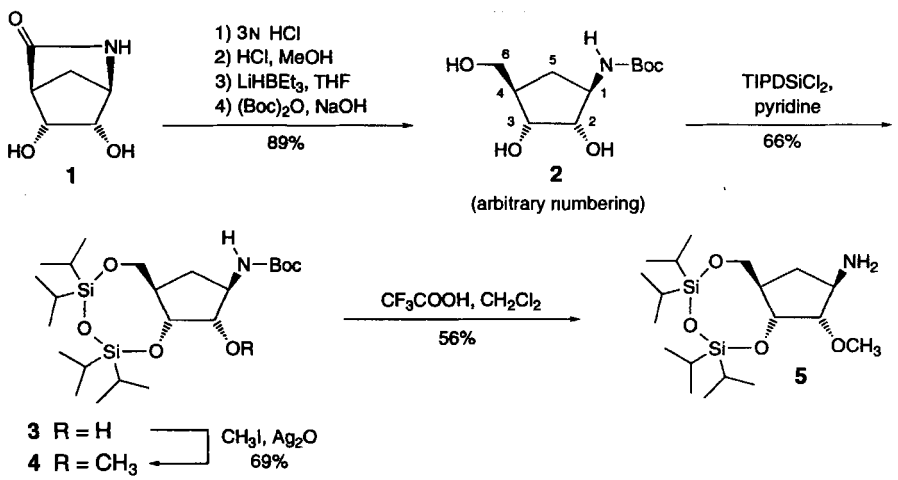


Fig. 2. *Design rationale for recognition of* Watson-Crick *C·G base pairs by heterocyclic bases* **L1** *and* **L2** *within a pyrimidine · purine · pyrimidine triple-helical motif*

[2]) Starting with structural coordinates from a pyrimidine · purine · pyrimidine triple-helical complex determined by NMR [21].

amounts of HCl. Reduction with $LiHBEt_3$ afforded a highly hydrophilic carbocyclic aminotriol[3]). Selective protection of the amino group with di(*tert*-butyl) dicarbonate ($(Boc)_2O$) yielded the triol **2** (89% from **1**). Protection of the 3-OH and the 4-CH_2OH group with 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane ($TIPDSiCl_2$) in pyridine and dimethylformamide (DMF) at room temperature afforded the hydroxycyclopentane derivative **3** in 66% yield. Methylation of the 2-OH group in refluxing MeI in the presence of Ag_2O afforded the fully proteced 2-methoxycyclopentane derivative **4** in 69% yield. Selective cleavage of the Boc protecting group by treatment with CF_3COOH and phenol in CH_2Cl_2 produced the carbocyclic ribofuranosylamine **5** in 56% yield. Cleavage of the disiloxane protecting group was not observed under these conditions.

Scheme 1. *Synthesis of Carbocyclic Building Block* **5**

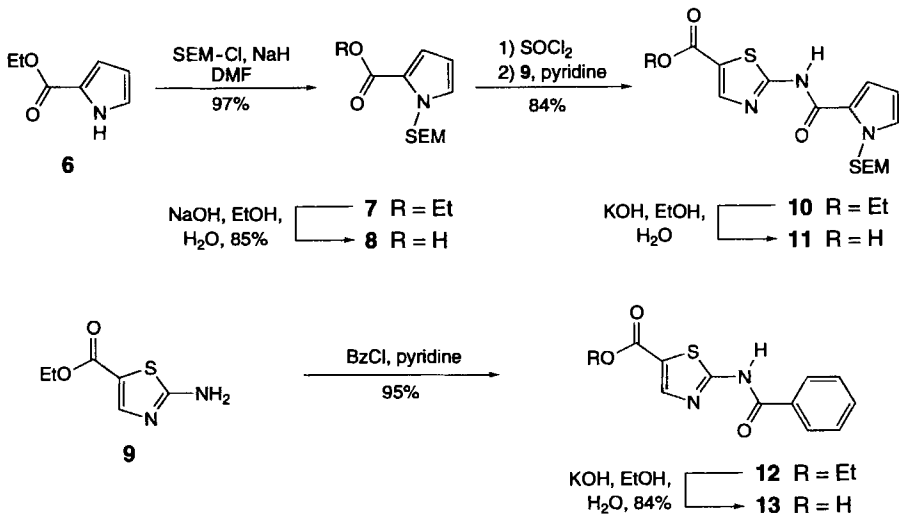


Protection of the 1*H*-pyrrole NH with a 2-(trimethylsilyl)ethoxymethyl (SEM) protecting group [27][28] was chosen because of its very good base stability and good stability towards nucleophiles and acids (*Schemes 2* and *3*). Although there is no precedent for its use in oligonucleotide synthesis, we have found the SEM group to be very suitable for the solid-phase synthesis of pyrimidine oligonucleotides (*Sect. 4*). Starting from ethyl 1*H*-pyrrole-2-carboxylate **6** [29], deprotonation with NaH and alkylation with SEM-Cl afforded the SEM-protected ester **7** in 97% yield (*Scheme 2*). Hydrolysis of **7** by treatment with NaOH in $EtOH/H_2O$ provided acid **8** in 85% yield. The amide **10** was formed in 84% yield by converting the acid **8** to the acyl chloride by refluxing in $SOCl_2$ followed by coupling with the free amino group of ethyl 2-aminothiazole-5-carboxylate **9** [30]. The carboxylic acid **11** of the nucleobase **L1** was prepared by saponification of **10** with KOH in $EtOH/H_2O$.

Ethyl 2-aminothiazole-5-carboxylate **9** was benzoylated with benzoyl chloride in pyridine affording the amide **12** in 95% yield. Saponification of **12** with KOH in $EtOH/H_2O$ provided the carboxylic acid **13** of the nucloebase **L2** in 84% yield (*Scheme 2*).

[3]) This reaction sequence was adopted from a synthesis of the racemic amino-triol building block [26].

Scheme 2. *Synthesis of Nucleobases* **11** *and* **13**

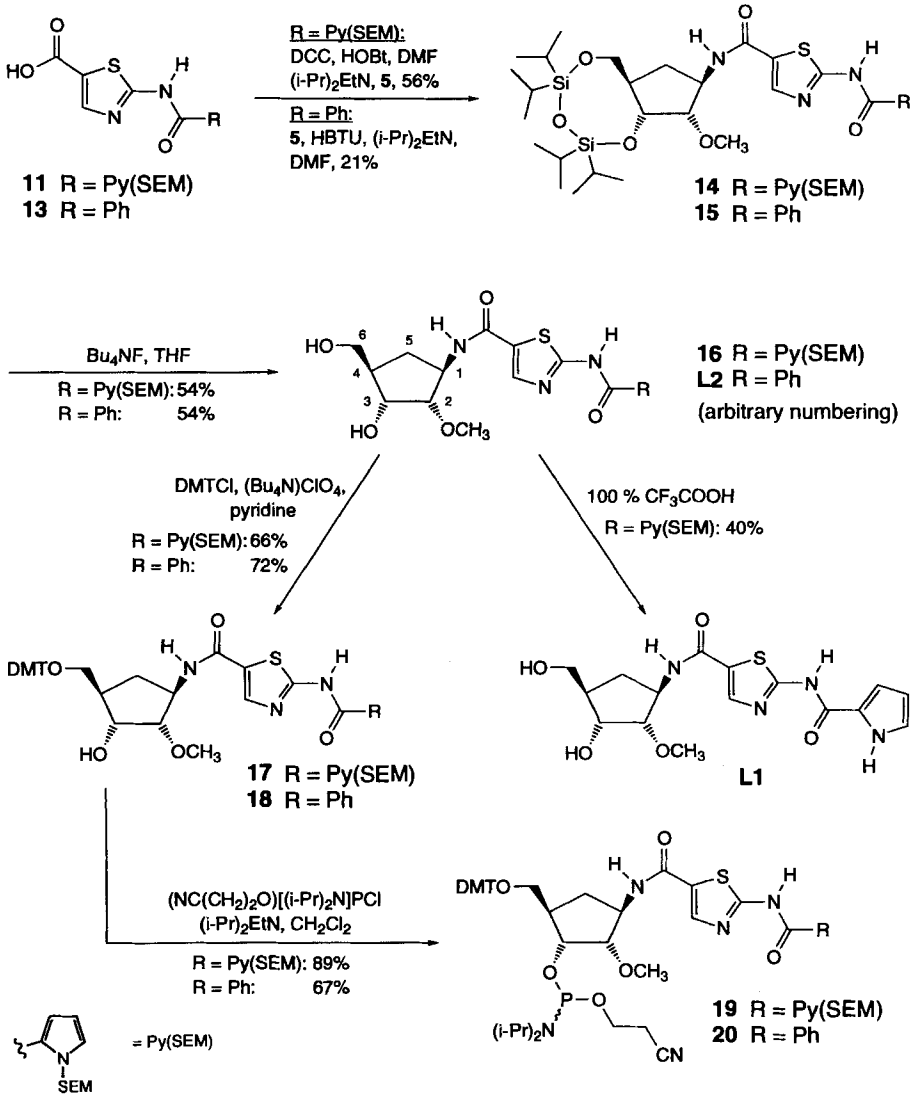


The carboxylic acids **11** and **13** of the nucleobases were activated with *N*,*N*′-dicyclohexylcarbodiimide (DCC)/1-hydroxybenzotriazole (HOBt) and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), respectively, and coupled with the secondary amine **5** in ethyldiisopropylamine ($EtN(i\text{-}Pr)_2$) and DME in yields of 56 and 21%, respectively (*Scheme 3*). The tetraisopropyldisiloxane protecting group in the resulting diamides **14** and **15** was cleaved with Bu_4NF in THF at room temperature with yields of 54%. No cleavage of the SEM protecting group was observed under these reaction conditions. The 4-CH_2OH group of **16** and **L2** was selectively protected with dimethoxytrityl chloride (DMT-Cl) and Bu_4NClO_4 [31] in anhydrous pyridine at room temperature in yields of 67 and 72%, respectively. The 3-OH group of the DMT-protected carbocyclic nucleosides **17** and **18** was reacted with chloro(2-cyanoethoxy)(diisopropylamino)phosphine in $EtN(i\text{-}Pr)_2$ in anhydrous CH_2Cl_2 to give the phosphoramidites **19** and **20** in 72 and 67% yields, respectively. An authentic sample of the carbocyclic nucleoside analog **L1** with free pyrrole NH was prepared in 40% yield by deprotection of **16** with neat CF_3COOH at room temperature[4]).

3. Synthesis of Oligonucleotides Containing the Nucleosides L1 and L2. – The syntheses of the oligonucleotides **OL-1** and **OL-2** were carried out on a 1.0-μmol scale using 2-cyanoethyl-phosphoramidite chemistry (*Figs. 3* and *4*) [32–34]. Coupling reactions of the phosphoramidites of T_d and $^{5Me}C_d$ were carried out by automated DNA synthesis. The phosphoramidites **19** and **20** of **L1** and **L2**, respectively, were coupled manually for 1 h to reduce the quantities of phosphoramidites needed.

[4]) The low yield is most likely due to the aqueous workup, which resulted in some loss of the hydrophilic product **L1**. TLC of the crude reaction mixture showed a complete conversion to product.

Scheme 3. *Synthesis of Carbocyclic Phosphoramidites* **19** *and* **20**



Following the solid-phase synthesis of oligonucleotide **OL-1**, the 5′-DMT group was removed by standard automated deprotection. The SEM protecting group was cleaved by treating the controlled pore glass (CPG) bound oligonucleotide with neat CF_3COOH. Monitoring of the reaction by HPLC and matrix-assisted laser-desorption-ionization time-of-flight mass spectrometry (MALDI-TOF-MS) showed a clean conversion to the SEM-deprotected oligomer[5]). Deprotection of the amino group of $^{5Me}C_d$, and of the

[5]) Test treatment of the CPG-attached oligonucleotide 5′-d($TT^{5Me}CT^{5Me}CTTT^{5Me}CT^{5Me}C^{5Me}GTTT$)-3′ with anhydrous CF_3COOH at room temperature for 2, 5, and 15 min, deprotection with conc. aqueous NH_3, solution and analysis by HPLC and MALDI-TOF-MS showed no detectable decomposition.

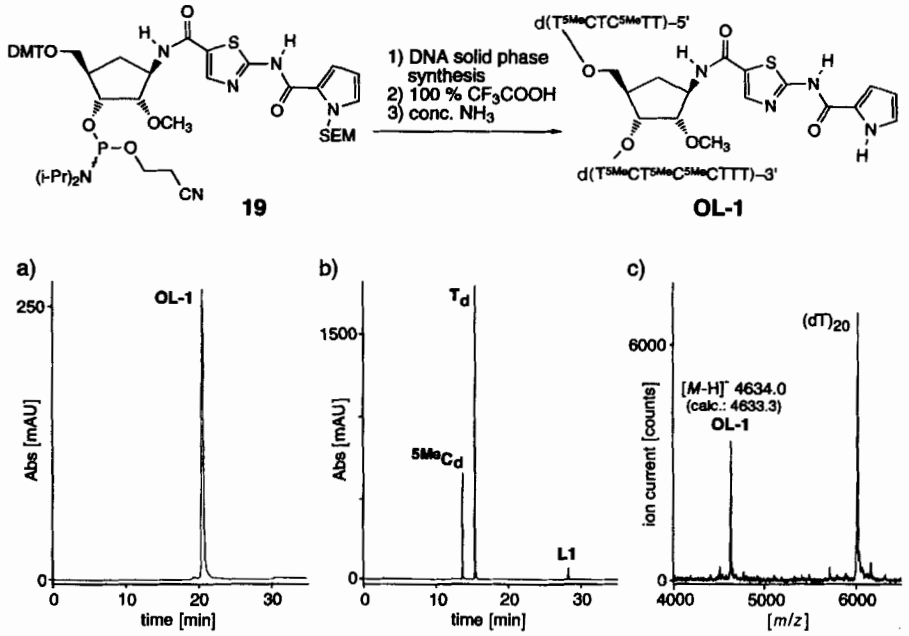


Fig. 3. *Synthesis and characterization of oligonucleotide* **OL-1**: a) *HPLC analysis of purified* **OL-1** (260 nm); b) *HPLC analysis of digested* **OL-1** (snake venom phosphodiesterase and alkaline phosphatase, 260 nm); c) *MALDI-TOF-MS analysis of purified* **OL-1**

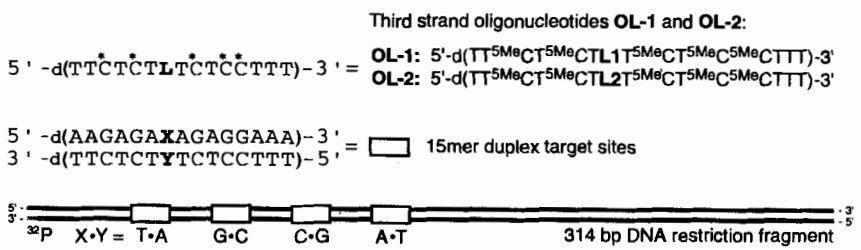


Fig. 4. *Sequences of third strand oligonucleotides* **OL-1** *and* **OL-2** *and duplex target sites within a 3'-end-*^{32}*P-labeled 314 base pair (bp)* AfI II/Fsp I *restriction fragment*

phosphate groups, and cleavage of the oligonucleotide from the solid support was carried out by treatment of the CPG-bound oligomer with concentrated ammonia at room temperature for 24 h. The oligonucleotide **OL-1** was purified by 20% polyacrylamide-gel electrophoresis (*Fig. 3*).

The 5'-DMT protected oligonucleotide **OL-2** was cleaved from the solid support and partially deprotected by treatment with concentrated ammonia at room temperature for 24 h. The still DMT-protected oligonucleotide **OL-2** was purified by reversed-phase FPLC, the 5'-DMT protecting group was cleaved by acidolysis, and the 5'-deprotected oligomer **OL-2** was purified a second time by reversed-phase FPLC.

The oligonucleotides **OL-1** and **OL-2** were characterized by HPLC and MALDI-TOF-MS (*Fig. 3*). Digestion of the purified oligonucleotides **OL-1** and **OL-2** with snake-

venom phosphodiesterase and calf-intestine alkaline phosphatase afforded the corresponding nucleoside monomers. HPLC Analysis of the digestion mixtures confirmed the identity of each nucleoside by correlation with the retention times and the UV spectra of authentic samples $^{5Me}C_d$,T_d, **L1**, and **L2** (*Fig. 3*). Integration of the peak areas confirmed the base composition of the oligomers. The amide bonds of **L1** and **L2** were found to be stable to the deprotection conditions. The SEM protecting group was shown to be compatible with the solid-phase synthesis of pyrimidine oligonucleotides.

4. Analysis of Binding Selectivity and Affinity by Quantitative DNase I Footprinting. – The binding affinities and selectivities of the two oligonucleotides **OL-1** and **OL-2** containing the nonnatural nucleosides **L1** and **L2** for a 314 base pair (bp) restriction fragment containing all four possible 15 bp double-helical binding sites (5′-d(AAGAGAXAGAGGAAA)-3′; **X** = T, G, C, A) [19] for **OL-1** and **OL-2** were determined by quantitative DNase I footprinting (*Fig. 4*) [35]. A detailed description of the employed quantitative footprinting titration experiments has been described previously [19]. The 3′-end-^{32}P-labeled 314 base pair *Afl II*/*Fsp I* restriction fragment was allowed to equilibrate with oligonucleotides **OL-1** or **OL-2** over the concentration range of 1 nM to 80 μM. Standardized experimental condition (100 mM NaCl, 10 mM 'bis-tris' (2,2-bis(hydroxymethyl)-2,2′,2″-nitrilotriethanol), 250 μM spermine, pH 7.0, room temperature) were used to allow the comparison of the binding data between different nucleobases within a given sequence context [19][36][37]. The conditions were chosen such that the association constants for the triple-helical complexes span the range which can be measured by DNase I footprinting [19]. Following 72 h of equilibration, DNase I was added and digestion was allowed to proceed for 5 min. After quenching, the reaction mixtures were separated by denaturing polyacrylamide-gel electrophoresis and the resulting gels were analyzed by storage phosphor imaging (*Fig. 5*). Integration of site and reference blocks allowed the determination of the apparent fractional occupancy (θ_{app}) of the site at each oligonucleotide concentration ($[O]_{tot}$). A binding isotherm was fitted to the resulting pairs of θ_{app}, $[O]_{tot}$ values (see *Exper. Part*) and the equilibrium association constant (K_T) was calculated (*Fig. 6*). K_T Values from three independent determinations were averaged to obtain each of the association constants reported in *Fig. 6*.

The highest affinities and selectivities were observed for oligonucleotide **OL-2** (*Fig. 6*). Triple-helical binding sites with the pyrimidine · purine base pairs C·G ($K_T = 2.4 \cdot 10^6$ M^{-1}) and T·A ($K_T = 2.6 \cdot 10^6$ M^{-1}) opposite to **L2** were bound 13- to 26-fold more tightly than binding sites with the purine · pyrimidine base pairs G·C ($K_T = 1.0 \cdot 10^5$ M^{-1}) and A·T ($K_T = 1.9 \cdot 10^5$ M^{-1}) opposite to **L2**. Slightly decreased affinities and selectivities were observed for oligonucleotide **OL-1** (*Fig. 6*). Triple-helical binding sites with the pyrimidine · purine base pairs C·G ($K_T = 9.2 \cdot 10^5$ M^{-1}) and T·A ($K_T = 1.6 \cdot 10^6$ M^{-1}) opposite to **L1** were bound 6- to 11-fold more tightly than binding sites with the purine · pyrimidine base pairs GC ($K_T = 1.4 \cdot 10^5$ M^{-1}) and AT ($K_T = 1.5 \cdot 10^5$ M^{-1}) opposite to **L1**. Both oligonucleotides showed no significant discrimination between the two binding sites which incorporated pyrimidine · purine (**L** · (C·G) compared to **L** · (T·A) triplexes) and purine · pyrimidine (**L** · (G·C) compared to **L** · (A·T) triplexes) base pairs, respectively (**L** = **L1** or **L2**).

Affinities for oligonucleotides 5′-d(TT^{5Me}CT^{5Me}CT**Z**T^{5Me}CT^{5Me}C^{5Me}CTTT)-3′ containing **Z** = ^{5Me}C, T, G, or A instead of **L1** or **L2** range from $K_T \geq 1.4–2.1 \cdot 10^8$ M^{-1} for

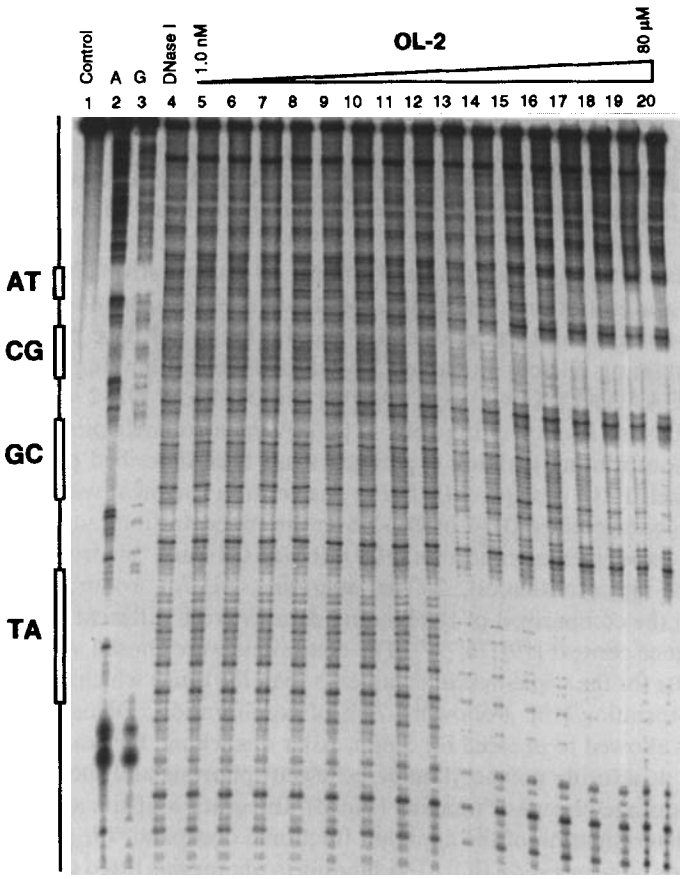


Fig. 5. *Storage phosphor autoradiogram of an 8% denaturing polyacrylamide gel used to separate the fragments generated by DNase I digestion in a quantitative footprinting titration experiment:* Lane 1: *intact 3′-labeled DNA after incubation in the absence of the third strand oligonucleotide;* Lanes 2–3: *products of adenine/guanine specific sequencing reactions;* Lanes 4–20: *DNase I digestion products obtained in the presence of varying concentrations of oligonucleotide* **OL-2** (no oligonucleotide (*Lane 4*), 1 nM (*Lane 5*), 2 nM (*Lane 6*), 4 nM (*Lane 7*), 10 nM (*Lane 8*), 20 nM (*Lane 9*), 40 nM (*Lane 10*), 100 nM (*Lane 11*), 200 nM (*Lane 12*), 400 nM (*Lane 13*), 1 μM (*Lane 14*) 2 μM (*Lane 15*), 4 μM (*Lane 16*), 10 μM (*Lane 17*), 20 μM (*Lane 18*), 40 μM (*Lane 19*), 80 μM (*Lane 20*)). The bar drawn to the left of the audioradiogram indicates the positions of the 15 bp binding sites within the 314 bp restriction fragment.

the two natural match triplexes $^{5Me}C \cdot (G \cdot C)$ and $T \cdot (A \cdot T)$ to $K_T < 1 \cdot 10^4\ M^{-1}$ for most mismatch triplexes [19][6]). High-to-intermediate affinities were previously determined with the same test system for the specific interactions of *i*) N^7-glycosylated guanine[7]),

[6]) Association constants greater than $1 \cdot 10^8\ M^{-1}$ are reported as lower limits due to the slow kinetics of triple helix formation under these conditions. The true association constants are probably 2- to 4-fold greater [19].

[7]) Abbreviations (also for corresponding nucleobases): 7-(2-deoxy-β-D-ribofuranosyl)guanine (H^7G), 5-amino-1-(2′-deoxy-β-D-ribofuranosyl)-3-methyl-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one (P), 1-(2′-deoxy-β-D-ribofuranosyl)-4-(benzamidophenyl)-1*H*-imidazole (D3), N^4-(6-aminopyridin-2-yl)-2′-deoxycytidine (M), [3-(2′-deoxy-β-D-ribofuranosyl)-2-methyl-3*H*-naphtho[1,2-*d*]imidazol-8-yl]urea (Z), and 3-(2′-deoxy-β-D-ribofuranosyl)-3*H*-imidazo[4,5-*h*]quinoline (Q).

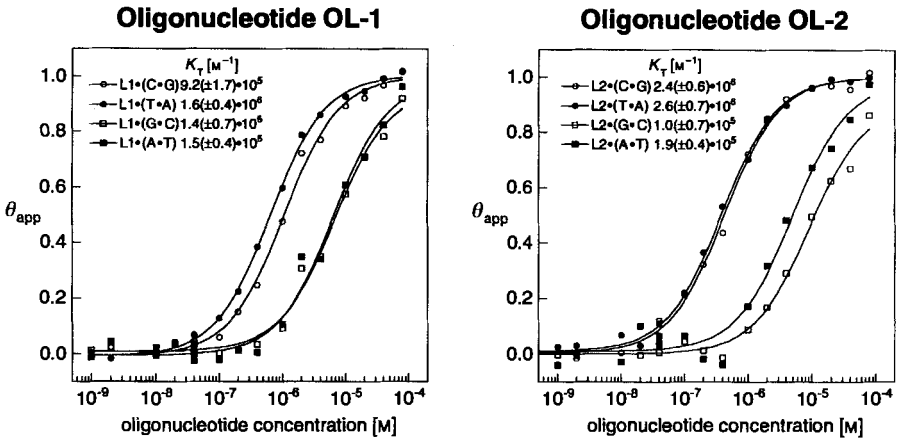
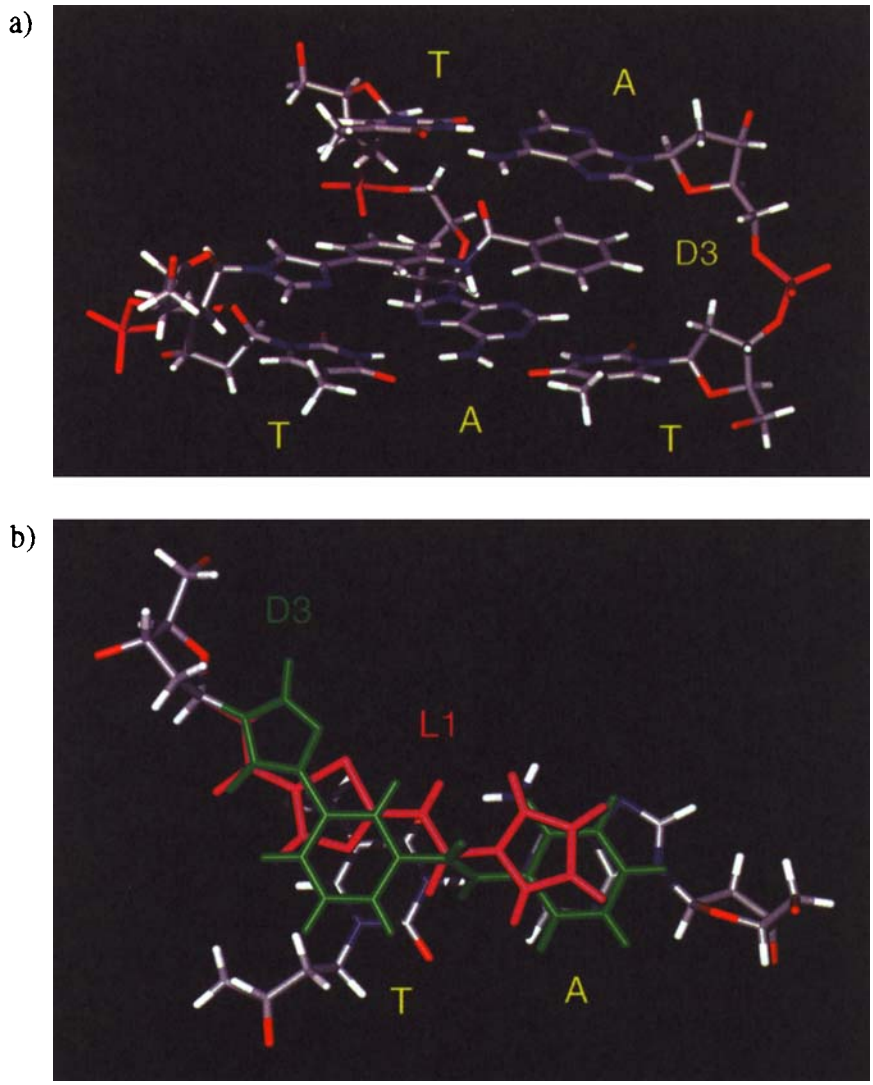


Fig. 6. *Binding isotherms derived from the DNase I footprinting gels of the oligonucleotides* **OL-1** *and* **OL-2** *to the four different 15 bp target sites* (*Fig. 5*). Each isotherm represents the average of three experiments ± s.e.m. performed in 100 mM NaCl, 10 mM 'bis-tris', 250 μM spermine, and 5 μM bp calf thymus DNA at pH 7.0 and 23°.

$N^7G \cdot (G \cdot C\ K_T \geq 1.6 \cdot 10^8\ M^{-1})$ [36–38][6]), *ii*) 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one, $P \cdot G \cdot C$ ($K_T = 3.1 \cdot 10^7\ M^{-1}$) [19][39][40], and *iii*) $G \cdot (T \cdot A)$ ($K_T = 1.1 \cdot 10^6\ M^{-1}$) [19][41][42].

The oligonucleotides **OL-1** and **OL-2** exhibit slightly higher affinities to binding sites with pyrimidine · purine base pairs (C · G and T · A) than an oligonucleotide with a G opposite to a T · A *Watson-Crick* base pair. The observation that **L1** and **L2** both bind with similar affinities to either C · G and T · A base pairs suggests that most of the binding affinity is not derived from the formation of selective H-bonds as anticipated in the initial design (*Fig. 2*). Selectivities similar to **L1** and **L2** had been observed previously for the nonnatural nucleoside 1-(2′-deoxy-β-D-ribofuranosyl)-4-(3-benzamidophenyl)-1*H*-imidazole (D3)[7]) [43] by quantitative DNase I footprinting: $D3 \cdot (T \cdot A)$ ($K_T \geq 4.8 \cdot 10^8\ M^{-1}$) ≈ $D3 \cdot (C \cdot G)$ ($K_T \geq 2.1 \cdot 10^8\ M^{-1}$) >> $D3 \cdot (A \cdot T)$ ($K_T = 8.8 \cdot 10^6\ M^{-1}$) ≈ $D3 \cdot (G \cdot C)$ ($K_T = 4.6 \cdot 10^6\ M^{-1}$)[8]). A solution structure determined by NMR of a pyrimidine · purine · pyrimidine triplex containing D3 in the third strand showed that the observed selectivity for pyrimidine · purine *Watson-Crick* base pairs is due to selective intercalation of the nucleobase of D3 into the *Watson-Crick* base pairs of the DNA duplex (*Fig. 7,a*) [45]. Selectivity is most likely derived from preferential stacking of D3 over the pyrimidine · purine base pairs T · A and C · G compared to the purine · pyrimidine base pairs A · T and G · C [45]. The nucleobases of **L1** and **L2** possess a system of aromatic rings fused by amide bonds comparable to the nucleobase of D3 and exhibit the same selectivity pattern for pyrimidine · purine over purine · pyrimidine *Watson-Crick* base pairs. This suggests that **L1**, **L2**, and D3 share sequence-selective intercalation as a common binding mechanism. This hypothesis seems plausible when overlaying the nucleobases of D3 and **L1** within the three dimensional model derived from the NMR

[8]) Measurements were performed in 100 mM NaCl, 10 mM 'bis-tris', 250 μM spermine, 5 μM bp calf thymus DNA at pH 7.0 and 22°. The sequence 5′-d(TTTTT^{5Me}CT^{5Me}CTD3T^{5Me}CT^{5Me}CT)-3′ within a 242 bp restriction fragment was used [44].



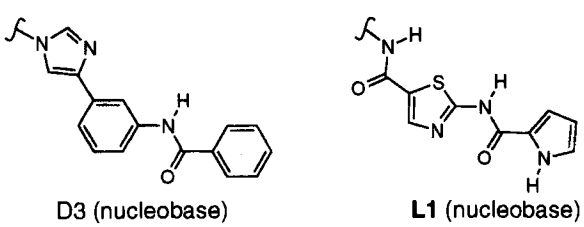


Fig. 7. *The selective recognition of T · A and C · G* (not shown) *base pairs by the nonnatural nucleoside D3 within the pyrimidine · purine · pyrimine triple-helical binding motif mediated by intercalation of the nucleobase of D3 into double-helical DNA:* a) *intercalation of the nucleobase of D3 into double-helical DNA within a pyrimidine · purine · pyrimidine triple helix determined by NMR* [45]; b) *superposition of the nucleobase of* **L1** *with the structure shown in* a (view from the top). The upper T · A *Watson-Crick* base pair has been omitted for clarity. D3 (nucleobase) is shown in green, **L1** (nucleobase) is shown in magenta.

structural determination of the triplex containing D3 (*Fig.7,b*) [45]. Herein, the nucleobase of **L1** was docked at the C(1′) position of the ribofuranose of D3 in a similar conformation to that of D3[9]). The terminal pyrrole ring of **L1** extends over the purine of the pyrimidine · purine *Watson-Crick* base pair, suggesting that selectivity is derived from preferred stacking of the pyrrole (or benzene in the case of **L2** and D3) moiety on a purine rather than a pyrimidine base.

5. Comparison of Different Nonnatural Nucleosides that Target Pyrimidine · Purine Base Pairs. – We were interested in comparing the above results within the same test system (*Fig. 4*) with other nonnatural nucleosides which possess nucleobases with extended aromatic structures. *Miller* and coworkers reported that N^4-(6-aminopyridin-2-yl)-2′-deoxycytidine (**M**)[7]) shows preferential binding of C·G > A·T > T·A ≈ G·C *Watson-Crick* base pairs within a pyrimidine · purine · pyrimidine triple-helical complex as determined by melting-temperature profiles and CD spectroscopy [46][47]. In model experiments *Zimmerman* and *Schmitt* showed by NMR titration experiments that the isolated heterocycle *N*-butyl-*N*′-(3-butyl-2-methyl-3*H*-naphtho[1,2-*d*]imidazol-8-yl)urea (^{Bu}Z)[7]) binds to a C·G *Watson-Crick* base pair [48].

As part of our efforts to extend the recognition code for triple-helix formation, we synthesized and tested the nucleoside analogs 3-(2-deoxy-β-D-ribofuranosyl)-3*H*-imidazol[4,5-*h*]quinolin (Q)[7]) and 3-(2-deoxy-β-D-ribofuranosyl)-3*H*-imidazol[4,5-*h*]quinolin-8-amine (^{NH_2}Q) [49] along with the nucleoside **M** described by *Miller* and coworkers [46][47], the *N*-benzoylated derivative of M, ^{Bz}M [50], and the nucleoside analog [3-(2-deoxy-β-D-ribofuranosyl)-2-methyl-3*H*-naphtho[1,2-*d*]imidazol-8-yl]urea Z [50] of the isolated butylated nucleobase analog ^{Bu_2}Z described by *Zimmerman* and *Schmitt* (*Table*) [48].

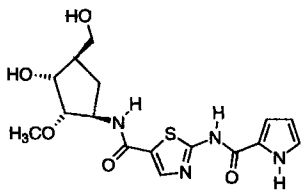
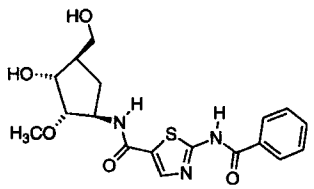
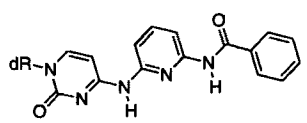
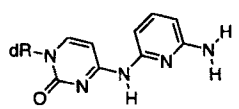
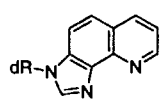
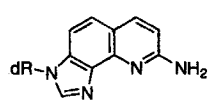
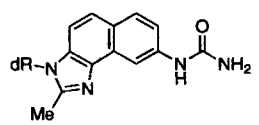
Binding affinities and specificities of oligonucleotides containing the nucleoside analogs ^{Bz}M, M, Q, ^{NH_2}Q, and ^{NH_2}Z were obtained as described in *Fig. 4* and the *Exper. Part*. The benzoylated nucleoside analog ^{Bz}M showed a similar selectivity pattern compared to **L1**, **L2**, and D3 towards pyrimidine · purine base pairs (*Table*). These results, analog with the similar extended aromatic structures of the four nucleoside analogs, suggest an intercalative binding mode for ^{Bz}M. Further evidence for the intercalation of ^{Bz}M is provided by the binding data for M, in which the benzoyl group of ^{Bz}M has been removed. As a result of this modification, specific binding of C·G and T·A is lost. Only a low-affinity binding for G·C is observed. This result suggests that the shortened nucleobase of M cannot extend over the purine base in pyrimidine · purine *Watson-Crick* base pairs to form a stabilizing interaction.

A comparison of the binding affinities and selectivities of **L1** and **L2** with those of ^{Bz}M and M suggests that attaching a nucleobase to a 3-hydroxy-4-(hydroxymethyl)-2-methoxycyclopent-1-yl moiety *via* an amide bond does not impose a serious energetic penalty on the formation of a triple-helical complex. The versatile synthetic accessibility and structural control in the design of these carbocycles suggests that this approach may be useful in the future development of modified nucleosides.

---

[9]) Molecular modeling and molecular-dynamics simulations suggest that this is also the favored conformation of **L1** and **L2**.

Table. *Equilibrium Association Constants (K_T) Determined by Quantitative DNase I Footprinting for 28 Triple-Helical Complexes of 7 Nonnatural Nucleosides*[a])

Compound		Association constant K_T [M^{-1}]				Ref.
		CG	TA	GC	AT	
L1		$9.2 \cdot 10^5$	$1.6 \cdot 10^6$	$1.4 \cdot 10^5$	$1.5 \cdot 10^5$	[51]
L2		$2.4 \cdot 10^6$	$2.6 \cdot 10^6$	$1.0 \cdot 10^5$	$1.9 \cdot 10^5$	[51]
^{Bz}M		$8.5 \cdot 10^6$	$5.8 \cdot 10^6$	$5.7 \cdot 10^5$	$8.8 \cdot 10^5$	[50]
M		$\sim 1 \cdot 10^5$	$< 1 \cdot 10^5$	*$7.0 \cdot 10^5$*	$< 1 \cdot 10^5$	[46][47][50]
Q		$\leq 1 \cdot 10^5$	$\leq 1 \cdot 10^5$	$\leq 1 \cdot 10^5$	$\leq 1 \cdot 10^5$	[49]
^{NH_2}Q		$\leq 1 \cdot 10^5$	$\leq 1 \cdot 10^5$	*$2.6 \cdot 10^5$*	$\leq 1 \cdot 10^5$	[49]
Z		$< 1 \cdot 10^5$	$< 1 \cdot 10^5$	*$1 \cdot 10^5$*	$< 1 \cdot 10^5$	[48][50]

[a]) Each reported K_T value was determined by quantitative DNase I footprinting using the test system described in *Fig. 4* and in the *Exper. Part.* Measurements were performed in 100 mM NaCl, 10 mM 'bis-tris', 250 μM spermine, and 5 μM bp calf thymus DNA at pH 7.0 and 22 ± 1°; dR = 1-(2′-deoxy-β-D-ribofuranosyl).

The nucleobases Q, ^{NH_2}Q, and Z with fused heterocyclic ring systems show low binding affinities, with a slight preference for binding to G·C base pairs. It is likely that no specific H-bonds are formed between Q, ^{NH_2}Q, and Z, and a *Watson-Crick* base pair in the context of a triple-helical complex in H_2O, in contrast to the results observed in $CHCl_3$ for ^{Bu_2}Z. This suggests that model systems in organic solvents do not necessarily reflect the energetics of triple-helix formation in an aqueous environment.

6. Conclusions. – Two designed carbocyclic nucleoside analogs **L1** and **L2**, together with the cytosine derivative ^{Bz}M, show selective binding of pyrimidine · purine base pairs (C · G and T · A) over purine · pyrimidine base pairs (G · C and A · T) within a triple-helical complex. A comparison of the binding affinities and the structure of the nucleobases of **L1**, **L2**, ^{Bz}M, and the previously described D3 suggest that all four nucleoside analogs likely share a common binding mode of sequence-specific intercalation. Although all four nucleoside analogs **L1**, **L2**, ^{Bz}M, and D3, do not distinguish between C · G and T · A, it may be possible to optimize selectivity by modifying the structure of the intercalating moiety.

It is remarkable that the design of novel bases for specific recognition of C · G or T · A base pairs by triple helix formation has continued to resist any definitive solution. It is diagnostic that the field of molecular recognition particularly as it pertains to the aqueous environment and nucleic acids is still at an early stage of development.

We are grateful for financial support from the *Office of Naval Research*, the *National Institutes of Health*, the *Beckman Institute*, and the *Studienstiftung des deutschen Volkes* for an *BASF* postdoctoral fellowship to *T. E. L.*

Experimental Part

1. General. – Reagents were purchased from *Aldrich* or *Fluka* of the highest quality available and used without further purification, unless otherwise noted. (−)-(1*R*,4*S*)-2-azabicyclo[2.2.1]hept-5-en-3-one (ee > 99%) was obtained from *Chiroscience*, Cambridge, UK. Ethyl 2-aminothiazole-5-carboxylate (**9**) was prepared according to [30]. 5-Methyl-2′-deoxycytidine (N-Bz) phosphoramidite was purchased from *BioGenex*. All other DNA-synthesis reagents were provided by *Glen Research*. Enzymes were purchased from *Boehringer Mannheim* or *New England Biolabs*. Solvents were obtained from *E. Merck* (GR grade). Anh. solvents were purchased from *Fluka* over molecular sieves. TLC: on *E. Merck* silica-gel plates (*60* F_{254}), and were visualized by UV at 254 nm or by staining with vanillin (0.4 g) in conc. H_2SO_4 (32 ml) and EtOH (436 ml) followed by heating. Flash chromatography (FC). *E. Merck* silica gel *60*, 40–63 μm. Solvents were removed under reduced pressure at 40° using a *Büchi Rotavapor*. M.p.: *Thomas Hoover* capillary melting point apparatus; open capillaries, not corrected. NMR: *General Electric QE* 300-NMR spectrometer, δ in ppm relative to residual undeuterated solvent $CDCl_3$ (δ(H) 7.26, δ(C) 77.0), (D_6)DMSO (δ(H) 2.50, δ(C) 39.7), CD_3OD (δ(H) 3.30, δ(C) 49.0). IR: *Perkin-Elmer FTIR* spectrometer series *1600* or *Spectrum 1000*. UV: *Hewlett-Packard 8452A* diode-array spectrophotometer; λ_{max} in nm, ε in $M^{-1}cm^{-1}$. $[\alpha]_D$: *Jasco DIP-181* digital polarimeter, d = 5 cm, c in g/100 ml. HR-MS: California Institute of Technology Mass Spectrometry Facility (*Vacuum Generators ProSpecE*) or University of California, Riverside Mass Spectrometry Facility (*Vacuum Generators ZAB2SE* or *7070*); EI: ionization energy 70 eV; FAB: matrix 3-nitrobenzyl alcohol (*m*NBA). MALDI-TOF-MS: California Institute of Technology Protein/Peptide Micro Analytical Laboratory, *Voyager-Elite* research-grade instrumentation, matrix 6-aza-2-thiothymine in 50% MeCN with 20 mM diammonium citrate. FPLC: *Pharmacia*, 2 × *P-500* pumps, *LCC-500* controller, *FRAC 100* fraction collector, *UV-2* dual path monitor; *ProRPC* 15 μm HR 10/10 column; buffer *A*: 0.1M (Et_3NH)OAc, pH 7.0; buffer *B*: 40% MeCN in 0.1M (Et_3NH)OAc, pH 7.0; flow rate 2.0 ml/min; detection at 254 nm. HPLC: *Hewlett-Packard 1090* liquid chromatograph; *Rainin Microsorb-MV C18*, 5 μm, 100 Å, 4.6 × 250 mm column; linear gradient: 0 min 0% *B*, 5 min 0% *B*, 25 min 40%, 40 min 100% *B* (buffer *A*: 20 mM NH_4OAc, pH 5.0, buffer *B*: 50% MeCN in 20 mM NH_4OAc, pH 5.0); flow rate 1.0 ml/min; oven temp. 40°; detection at 260 nm. Storage phosphor imaging was carried out on a *Molecular Dynamics PhosphorImager 425E* using ImageQuant 3.3 Software. Molecular modeling and visualization was carried out on a *Silicon Graphics Indigo*^{2}) *IMPACT R10000* workstation using MacroModel 5.5 (Amber and MM3 force fields) [52] and *BIOSYM/Molecular Simulations* Insight II 95.0.

2. Synthesis of Nucleosides. – tert-*Butyl* N-*[(1*R,*2*S,*3*R,*4*R*)*-2,3-Dihydroxy-4-(hydroxymethyl)cyclopentyl]carbamate* (**2**). A soln. of lactam **1** (4.07 g, 28.4 mmol) in 3N HCl (20 ml) was stirred at reflux for 3 h [25][26]. Evaporation of the solvent and lyophilization overnight afforded a pale yellow residue (5.83 g). The residue was dissolved in anh. MeOH (30 ml). After addition of HCl-sat. anh. MeOH (350 μl), the clear, yellow soln. was stirred at reflux for 2.5 h. Evaporation and drying in vacuum overnight afforded a yellow, crystalline solid (6.34 g). Under Ar, a portion of the crude methyl ester (2.61 g) was suspended in anh. THF (50 ml). $Li(HBEt_3)$ (95 ml of a 1M

soln. in THF, 95.0 mmol) was added dropwise at 0° with stirring over 50 min. After stirring at r.t. for 25 min, the solvent was evaporated affording a clear, pale yellow oil. The oil was dissolved in MeOH (150 ml) and was acidified at 0° with 1M HCl to pH 1. Evaporation and lyophilization afforded of a pale red residue (9.73 g). A portion of the crude triol (9.35 g) was dissolved in H_2O (30 ml). The pH of the clear, pale red soln. was adjusted to pH 13 by the addition of NaOH. At r.t. di(*tert*-butyl) dicarbonate (6.11 g, 28.0 mmol) was added portionwise over a period of 2 h and the mixture was stirred for 3 d. The pH was readjusted to 13 by addition of NaOH pellets, and another portion of di(*tert*-butyl) dicarbonate (3.00 g, 13.7 mmol) was added over a period of 6.5 h. After stirring for 24 h, sat. NH_4Cl (50 ml) was added. Evaporation and purification by FC ($MeOH/CH_2Cl_2$ 1:9) afforded **2** (2.63 g, 89% overall yield). White, crystalline solid. TLC ($MeOH/CH_2Cl_2$ 1:9): R_f 0.18. M.p. 105.0–106.5°; $[\alpha]_D^{25} = -6.3$ ($c = 1.17$, MeOH). IR (KBr): 3365, 2964, 2924, 2873, 1684, 1528, 1292, 1179, 1070, 1051. ^{1}H-NMR (300 MHz, (D_6)DMSO): 0.91–1.01 (*m*, 1 H); 1.37 (*s*, 9 H); 1.78–1.85 (*m*, 1 H); 1.93–2.00 (*m*, 1 H); 3.25–3.40 (*m*, 2 H); 3.48–3.63 (*m*, 3 H); 4.31 (*d*, $J = 4.7$, 1 H); 4.43 (*d*, $J = 3.3$, 1 H); 4.52 (Ψ_t, $J = 5.0$, 1 H); 6.71 (*d*, $J = 7.8$, 1 H). ^{13}C-NMR (75 MHz, (D_6)DMSO): 28.3; 30.6; 44.9; 55.0; 63.0; 72.1; 75.9; 77.4; 155.2. HR-MS (FAB, *m*NBA): 248.1500 ($C_{11}H_{22}NO_5^+$, MH^+; calc. 248.1498).

tert-*Butyl* N-*[(6a*R,*8*R,*9S,9aR)-Hexahydro-9-hydroxy-2,2,4,4-tetraisopropylcyclopenta[f][1,3,5,2,4]trioxadisilocin-8-yl]carbamate* (**3**). To a soln. of **2** (1.20 g, 4.83 mmol) in anh. DMF (5 ml) and anh. pyridine (3.9 ml), was added 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (1.55 ml, 4.83 mmol) dropwise under Ar at r.t. After stirring for 5 h, MeOH (5 ml) was added and stirring was continued for 20 min. The solvent was evaporated and residual pyridine was removed by azeotropic distillation with toluene (2 × 25 ml). Purification by FC AcOEt/hexane 2:8) afforded **3** (1.57 g, 66%). White, crystalline solid. TLC (AcOEt/hexane 2:8): R_f 0.23. M.p. 58.0–62.0°. $[\alpha]_D^{25} = -11.5$ ($c = 1.22$, MeOH). IR (NaCl, film): 3540, 3500, 3317, 2943, 2866, 1668, 1538, 1463, 1248, 1161, 1117, 1034, 884, 676. ^{1}H-NMR (300 MHz, (D_6)DMSO): 0.95–1.07 (*m*, 29 H); 1.37 (*s*, 9 H); 1.83–1.98 (*m*, 2 H); 3.45–3.53 (*m*, 1 H); 3.62–3.69 (*m*, 2 H); 3.77–3.82 (*m*, 1 H); 3.88–3.93 (*m*, 1 H); 4.22 (br. *s*, 1 H); 6.97 (*d*, $J = 7.2$, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): 12.6; 12.9; 13.3; 13.4; 17.0; 17.1; 17.2; 17.3; 17.4; 17.5; 28.3; 30.0; 45.7; 55.8; 63.1; 72.5; 76.8; 79.5; 155.5. HR-MS (FAB, *m*NBA): 490.3032 ($C_{23}H_{48}NO_6Si_2^+$, MH^+; calc. 490.3020).

tert-*Butyl* N-*[(6a*R,*8*R,*9S,9aR)-Hexahydro-2,2,4,4-tetraisopropyl-9-methoxycyclopenta[f][1,3,5,2,4]trioxadisilocin-8-yl]carbamate* (**4**). A suspension of **3** (245 mg, 500 µmol) and Ag_2O (579 mg, 2.5 mmol) in MeI (2.5 ml) was stirred at reflux for 2 d. The solvent was removed by distillation, the residue was dissolved in CH_2Cl_2 and the resulting soln. was filtered through *Celite*. The *Celite* was washed with AcOEt and the combined filtrate was evaporated. Purification by FC (AcOEt/hexane 1:9) afforded **4** (173 mg, 69%). Clear, colorless oil. TLC (AcOEt/hexane 2:8): R_f 0.32. $[\alpha]_D^{25} = +3.7$ ($c = 1.07$, MeOH). IR (NaCl, film): 3454, 3337, 2933, 2868, 1706, 1499, 1466, 1365, 1249, 1173, 1127, 1041, 886, 695. ^{1}H-NMR (300 MHz, $CDCl_3$): 0.93–1.14 (*m*, 29 H); 1.42 (*s*, 9 H); 2.06–2.19 (*m*, 2 H); 3.37–3.40 (*m*, 1 H); 3.50 (*s*, 3 H); 3.67 (*dd*, $J = 11.7$, 1.9, 1 H); 3.74–3.79 (*m*, 1 H); 3.90 (*dd*, $J = 11.7$, 2.7, 1 H); 3.91–3.97 (*m*, 1 H); 4.35 (*s*, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): 12.5; 12.7; 13.3; 13.5; 17.0; 17.1; 17.2; 17.3; 17.4; 28.3; 29.6; 43.1; 52.7; 58.2; 59.7; 72.0; 79.3; 86.4; 154.8. HR-MS (FAB, *m*NBA): 504.3179 ($C_{24}H_{50}NO_6Si_2^+$, MH^+; calc. 504.3177).

*(6a*R,*8*R,*9S,9a*R)-*Hexahydro-2,2,4,4-tetraisopropyl-9-methoxycyclopenta[f][1,3,5,2,4]trioxadisilocin-8-amine* (**5**). To a soln. of **4** (620 mg, 1.23 mmol) and phenol (1.70 g, 18.1 mmol) in anh. CH_2Cl_2 (10 ml), was added CF_3COOH (566 µl, 7.4 mmol) at r.t. After stirring for 27 h at r.t., the pH was adjusted to 10 by the addition of sat. Na_2CO_3 soln. and the solvent was evaporated. Purification by FC ($AcOEt/Et_3N$/hexane 1:1:9) afforded **5** (277 mg, 56%). Clear, pale-yellow oil. TLC ($MeOH/CH_2Cl_2$ 5:95): R_f 0.07. IR (NaCl, film): 3374, 2945, 2867, 1465, 1134, 1082, 1040, 886, 693. ^{1}H-NMR (300 MHz, $CDCl_3$): 0.79–1.09 (*m*, 29 H); 1.78 (br. *s*, 2 H); 1.98–2.11 (*m*, 1 H); 2.12–2.21 (*m*, 1 H); 3.19–3.30 (*m*, 2 H); 3.51 (*s*, 3 H); 3.64 (*dd*, $J = 11.6$, 5.0, 1 H); 3.92 (*dd*, $J = 11.6$, 3.7, 1 H); 4.17 (*dd*, $J = 7.9$, 5.0, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): 12.5; 12.8; 13.4; 13.5; 16.9; 17.0; 17.1; 17.2; 17.3; 17.4; 31.8; 45.3; 54.3; 58.3; 61.8; 72.3; 89.0. HR-MS (FAB, mNBA): 404.2649 ($C_{19}H_{42}NO_4Si_2^+$, MH^+; calc. 404.2652).

Ethyl 1-{[2-(Trimethylsilyl)ethoxy]methyl}-1H-pyrrole-2-carboxylate (**7**). Under Ar, NaH (660 mg of 60% NaH in mineral oil, 16.5 mmol) was washed with anh. hexane (2 × 4 ml) and suspended in anh. DMF (20 ml). At r.t., **6** (2.09 g, 15.0 mmol) was added portionwise within 5 min under Ar. After stirring for 25 min, the soln. was cooled to 0° and SEM-Cl (2.65 ml, 15.0 mmol) was added dropwise over 5 min. After stirring for 25 min at 0° and 1 h at r.t., the mixture was poured into cold sat. $NaHCO_3$ soln. (100 ml). After addition of H_2O (100 ml), the aq. phase was extracted with Et_2O (4 × 100 ml). The combined org. layers were washed with H_2O (50 ml) and brine (50 ml), dried ($MgSO_4$), and evaporated. Purification by FC (AcOEt/hexane 5:95) afforded **7** (3.90 g, 94%). Clear, colorless oil. TLC (AcOEt/hexane): R_f 0.33. IR (NaCl, film): 3111, 2954, 2896, 1706, 1420, 1316, 1233, 1113, 1082, 860, 837, 743. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.42 (*s*, 9 H); 0.89 (*t*, $J = 8.2$, 2 H); 1.34 (*t*, $J = 7.1$, 3 H); 3.51 (*t*, $J = 8.2$, 2 H); 4.28 (*q*, $J = 7.1$, 2 H); 5.69 (*s*, 2 H); 6.17 (*dd*, $J = 3.6$, 2.8, 1 H); 6.98–7.01 (*m*, 2 H). ^{13}C-NMR

(75 MHz, $CDCl_3$): −1.5; 14.4; 17.7; 59.9; 65.8; 76.7; 108.6; 118.8; 122.4; 128.3; 161.0. HR-MS (CI, NH_3): 270.1525 ($C_{13}H_{24}NO_3Si^+$, MH^+; calc. 270.1525).

1-{[2-(Trimethylsilyl)ethoxy]methyl}-1H-pyrrole-2-carboxylic Acid (**8**). A soln. of **7** (2.67 g, 9.90 mmol) and NaOH (1.98 g, 49.5 mmol) in EtOH (9 ml) and H_2O (9 ml) was stirred at reflux for 5 h. After the addition of H_2O (40 ml), the aq. phase was washed with Et_2O (2 × 20 ml). The aq. layer was acidified by the addition of 3N HCl (30 ml) and extracted with AcOEt (3 × 70 ml). The combined org. layers were washed with brine (30 ml) and dried ($MgSO_4$). Evaporation and purification of the crude product by FC (AcOEt/hexane 3:7) afforded **8** (2.03 g, 85%). White solid. TLC (AcOEt/hexane 3:7): R_f 0.14. M.p. 86.0–87.0°. IR (NaCl, film): 3111, 2953, 2627, 1676, 1437, 1252, 1065, 836, 751. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.02 (*s*, 9 H); 0.92 (*t*, *J* = 8.2, 2 H); 3.56 (*t*, *J* ≈ 8.2, 2 H); 5.70 (*s*, 2 H); 6.24 (*dd*, *J* = 3.8, 2.7, 1 H); 7.09 (*dd*, *J* = 2.7, 1.8, 1 H); 7.16 (*dd*, *J* = 3.8, 1.8, 1 H); 12.16 (br. *s*, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): −1.5; 17.8; 66.1; 76.8; 109.2; 121.2; 121.5; 129.8; 166.1. HR-MS (EI): 241.1133 ($C_{11}H_{19}NO_3Si^+$, M^+; calc. 241.1134).

Ethyl 2-{[(1-{[2-(Trimethylsilyl)ethoxy]methyl}-1H-pyrrol-2-yl)carbonyl]amino}thiazole-5-carboxylate (**10**). A soln. of **8** (1.66 g, 6.86 mmol) in $SOCl_2$ (28 ml) was stirred at reflux for 6 min under Ar. The solvent was evaporated and residual $SOCl_2$ was co-evaporated with anh. benzene (10 ml) under Ar. The resulting clear, pale-brown oil was dissolved in anh. benzene (6.9 ml), and 4.0 ml of the soln. was added dropwise to a soln. of **9** (689 mg, 4.00 mmol) in anh. pyridine (6 ml) at r.t. under Ar. After stirring for 21 h at r.t., the mixture was poured into 3N HCl (30 ml) and AcOEt (30 ml). The layers were separated, and the aq. phase was extracted with AcOEt (3 × 30 ml). The combined org. phases were washed with sat. Na_2CO_3 soln. (30 ml) and brine (30 ml), dried ($MgSO_4$), and evaporated. Purification by FC afforded **10** (1.33 g, 84%). Pale-yellow, crystalline solid. TLC (AcOEt/hexane 2:8): R_f 0.14. M.p. 117.0–118.0°. IR (NaCl, film): 3128, 2953, 1707, 1635, 1508, 1400, 1314, 1248, 1082, 861, 749. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.06 (*s*, 9 H); 0.93 (*t*, *J* = 8.4, 2 H); 1.34 (*t*, *J* = 7.1, 3 H); 3.60 (*t*, *J* = 8.4, 2 H); 4.32 (*q*, *J* = 7.1, 2 H); 5.76 (*s*, 2 H); 6.28 (*dd*, *J* = 3.8, 2.8, 1 H); 7.04 (*dd*, *J* = 3.8, 1.4, 1 H); 7.12 (*dd*, *J* = 2.8, 1.4, 1 H); 7.94 (*s*, 1 H); 11.91 (br. *s*, 0.5 H); 12.13 (br. *s*, 0.5 H). ^{13}C-NMR (75 MHz, $CDCl_3$): −1.5; 14.2; 17.7; 61.1; 66.3; 77.1; 109.2; 117.4; 122.3; 123.5; 129.5; 144.4; 158.7; 161.9; 163.6. HR-MS (CI, NH_3): 396.1416 ($C_{17}H_{26}N_3O_4SSi^+$, MH^+; calc. 396.1413).

2-{[(1-{[2-(Trimethylsilyl)ethoxy]methyl}-1H-pyrrol-2-yl)carbonyl]amino}thiazole-5-carboxylic Acid (**11**). A soln. of **10** (1.08 g, 2.72 mmol) and KOH (1.53 g, 27.2 mmol) in EtOH (13 ml) and H_2O (10 ml) was stirred at reflux for 3 h. The mixture was poured into H_2O (20 ml) and Et_2O (20 ml). The layers were separated, and the aq. phase was acidified to pH 1 with 3N HCl. The white precipitate was collected by filtration and dried under vacuum, affording 2.06 g of crude **11** which was used without further purification. M.p. 265.0–267.0°. IR (KBr): 3422, 3147, 2953, 1670, 1514, 1315, 1101, 1080, 862, 744. ^{1}H-NMR (300 MHz, (D_6)DMSO): −0.10 (*s*, 9 H); 0.79 (*t*, *J* = 8.2, 2 H); 3.45 (*t*, *J* = 8.2, 2 H); 5.71 (*s*, 2 H); 6.22 (*dd*, *J* = 3.9, 2.7, 1 H); 7.36 (*dd*, *J* = 2.7, 1.4, 1 H); 7.45 (*dd*, *J* = 3.9, 1.4, 1 H); 8.09 (*s*, 1 H); 12.59 (br. *s*, 1 H); 13.05 (br. *s*, 1 H). ^{13}C-NMR (75 MHz, (D_6)DMSO): −1.4; 17.1; 64.9; 76.4; 108.4; 118.0; 122.4; 122.6; 131.0; 144.5; 158.8; 162.5; 162.9. HR-MS (CI, NH_3): 368.1088 ($C_{15}H_{22}N_3O_4SSi^+$, M^+; calc. 368.1100).

Ethyl 2-(Benzamido)thiazole-5-carboxylate (**12**). To a soln. of **9** (861 mg, 5.0 mmol) in anh. pyridine (5 ml), was added PhCOCl (697 µl, 6.0 mmol) dropwise at r.t. After stirring for 2 d at r.t., the mixture was poured into 3N HCl (40 ml) and AcOEt (40 ml). The layers were separated, and the aq. layer was extracted with AcOEt (3 × 40 ml). The combined org. layers were washed with sat. Na_2CO_3 soln. (40 ml) and brine (40 ml), dried ($MgSO_4$), and evaporated. Purification by FC (AcOEt/hexane 2:8, 3:7) afforded **12** (1.32 g, 95%). White solid. TLC (AcOEt/hexane 3:7): R_f 0.15. M.p. 189.5–190.0°. IR (NaCl): 3287, 3062, 2953, 1706, 1542, 1490, 1235, 1172, 1090, 752. ^{1}H-NMR (300 MHz, (D_6)DMSO): 1.31 (*t*, *J* = 7.1, 3 H); 4.30 (*q*, *J* = 7.1, 2 H); 7.56 (*dd*, *J* = 7.5, 7.3, 2 H); 7.65 (*d*, *J* = 7.3, 1 H); 8.10 (*d*, *J* = 7.5, 2 H); 8.23 (*s*, 1 H); 13.10 (*s*, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): 14.2; 61.3; 122.7; 128.1; 129.0; 132.1; 133.3; 144.0; 161.7; 164.1; 166.2. HR-MS (CI, NH_3): 277.0658 ($C_{13}H_{13}N_2O_3S^+$, MH^+; calc. 277.0647).

2-(Benzamido)thiazole-5-carboxylic Acid (**13**). A soln. of **12** (1.16 g, 4.21 mmol) and KOH (2.36 g, 42.1 mmol) in H_2O (30 ml) and EtOH (55 ml) was stirred at reflux for 5 h. The mixture was poured into E_2O (70 ml) and H_2O (50 ml). The layers were separated and 3N HCl (70 ml) was added to the aq. phase. The white precipitate was collected by filtration and washed with cold H_2O. Drying under vacuum afforded **13** (878 mg, 84%). White solid. M.p. 285.0–286.0°. IR (KBr): 3179, 3067, 2842, 1677, 1495, 1301, 1237, 1165, 893, 706. ^{1}H-NMR (300 MHz, (D_6)DMSO): 7.55 (*dd*, *J* = 7.8, 7.3, 2 H); 7.65 (*d*, *J* = 7.8, 1 H); 8.11 (*d*, *J* = 7.3, 2 H); 8.15 (*s*, 1 H); 13.03 (br. *s*, 2 H). ^{13}C-NMR (75 MHz, (D_6)DMSO): 122.9; 128.3; 131.5; 132.9; 144.5; 162.7; 162.9; 165.7. HR-MS (CI, NH_3): 249.0330 ($C_{11}H_9N_2O_3S^+$, MH^+; calc. 249.0334).

N^5-[(6aR,8R,9S,9aR)-Hexahydro-2,2,4,4-tetraisopropyl-9-methoxycyclopenta[f][1,3,5,2,4]trioxadisilocin-8-yl]-2-{[(1-{[2-(trimethylsilyl)ethoxy]methyl}-1H-pyrrol-2-yl)carbonyl]amino}thiazole-5-carboxamide (**14**). A

soln. of **11** (120 mg, 327 μmol), *N,N'*-dicyclohexylcarbodiimide (67 mg, 325 μmol) and 1-hydroxybenzotriazole · H_2O (50 mg, 326 μmol) in DMF (2 ml) was stirred for 9 h at r.t. under Ar. The precipitate was removed by filtration and $EtN(i\text{-}Pr)_2$ (200 μl, 1.22 mmol), followed by **5** (44 mg, 109 μmol) in anh. CH_2Cl_2 (200 μl) were added to the filtrate dropwise at r.t. After stirring at r.t. for 20 h, the mixture was poured into sat. NH_4Cl soln. (25 ml) and extracted with AcOEt (4 × 15 ml). The combined org. layers were washed with sat. Na_2CO_3 soln. (10 ml) and brine (10 ml), dried ($MgSO_4$), and evaporated. Purification by FC afforded 45 mg (56%) of **14** (45 mg, 56%). White solid. TLC (AcOEt/hexane 1:1): R_f 0.29. M.p. 200.0–206.0°. IR (NaCl, film): 3262, 3097, 2945, 2867, 1667, 1621, 1538, 1312, 1104, 1082, 1041, 861, 743. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.06 (*s*, 9 H); 0.93 (*t*, *J* = 8.3, 2 H); 0.99–1.07 (*m*, 28 H); 1.24–1.34 (*m*, 1 H); 2.16–2.33 (*m*, 2 H); 3.46–3.62 (*m*, 6 H); 3.72 (*d*, *J* = 11.7, 1 H); 3.95 (*d*, *J* = 11.7, 1 H); 4.01 (*dd*, *J* = 9.5, 4.7, 1 H); 4.19–4.25 (*m*, 1 H); 5.62 (*d*, *J* = 5.4, 1 H); 5.73 (*s*, 2 H); 6.24–6.26 (*m*, 1 H); 7.06–7.10 (*m*, 2 H); 7.75 (*s*, 1 H); 11.70 (*s*, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): −1.5; 12.5; 12.7; 13.3; 13.5; 17.0; 17.1; 17.2; 17.3; 17.4; 17.5; 17.6; 17.7; 29.4; 43.2; 52.3; 58.3; 59.6; 66.3; 72.0; 77.0; 86.2; 109.2; 117.4; 123.5; 126.0; 129.5; 140.1; 158.4; 160.5; 161.3.

*2-Benzamido-*N^5*-[(6a*R*,8*R*,9*S*,9a*R*)-Hexahydro-2,2,4,4-tetraisopropyl-9-methoxycyclopenta[f][1,3,5,2,4]-trioxadisilocin-8-yl]thiazole-5-carboxamide* (**15**). To a soln. of **13** (340 mg, 1.37 mmol) in anh. DMF (10 ml) were added **5** (277 mg, 685 μmol) in anh. CH_2Cl_2 (2 ml), $EtN(i\text{-}Pr)_2$ (349 ml, 2.06 mmol), and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (520 mg, 1.37 mmol) at r.t. After stirring for 1 d at r.t., the mixture was poured into sat. NH_4Cl soln. (40 ml) and AcOEt (50 ml). The layers were separated, and the aq. layer was extracted with AcOEt (3 × 50 ml). The combined org. layers were washed with H_2O (30 ml) and brine (30 ml), dried ($MgSO_4$), and evaporated. Purification by FC (AcOEt/hexane 4:6, 1:1, 8:2) afforded **15** (93 mg, 21%). White solid. TLC ($MeOH/CH_2Cl_2$ 5:95): R_f 0.40. M.p. 179.0–180.5°. IR (NaCl): 3362, 3148, 3064, 2944, 2867, 1673, 1634, 1530, 1297, 1039, 887, 702. ^{1}H-NMR (300 MHz, $CDCl_3$): 0.95–1.06 (*m*, 28 H); 1.23–1.35 (*m*, 1 H); 2.19–2.32 (*m*, 2 H); 3.52–3.55 (*m*, 1 H); 3.55 (*s*, 3 H); 3.72 (*dd*, *J* = 11.8, 1.5, 1 H); 3.94 (*dd*, *J* = 11.8, 2.6, 1 H); 4.02 (*dd*, *J* = 9.7, 4.8, 1 H); 4.19–4.25 (*m*, 1 H); 5.74 (*d*, *J* = 6.4, 1 H); 7.47–7.64 (*m*, 4 H); 7.98 (*d*, *J* = 7.2, 2 H); 12.60 (br. *s*, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): 12.5; 12.7; 13.3; 13.5; 17.0; 17.1; 17.2; 17.3; 17.4; 17.6; 29.4, 43.3; 52.4; 58.4; 59.6; 73.0; 86.1; 126.4; 128.0; 128.9; 131.9; 133.2; 139.9; 160.4; 161.8; 165.8. HR-MS (FAB, *m*NBA): 634.2808 ($C_{30}H_{48}N_3SSi_2^+$, MH^+; calc. 634.2802).

N^5*-[(1*R*,2*S*,3*R*,4*R*)-3-Hydroxy-4-(hydroxymethyl)-2-methoxycyclopentyl]-2-{[(1-{[2-(trimethylsilyl)ethoxy]methyl}-1*H*-pyrrol-2-yl)carbonyl]amino}thiazole-5-carboxamide* (**16**). To a soln. of **14** (345 mg, 458 μmol) in anh. THF (2.5 ml) was added a soln. of $Bu_4NF \cdot 3\,H_2O$ (318 mg, 1.01 mmol) in anh. THF (1.5 ml) at 0°. After stirring for 1 h at 0°, the mixture was poured into sat. $NaHCO_3$ soln. (30 ml) and extracted with AcOEt (4 × 20 ml). The combined org. layers were washed with brine (20 ml), dried ($MgSO_4$), and evaporated. Purification of the crude product by cation-exchange chromatography (sulfopropyl *Sephadex C25*, $MeOH/H_2O$ 1:1) and FC ($MeOH/CH_2Cl_2$ 1:40, 1:20, 1:9) afforded **16** (126 mg, 54%). White, solid foam. TLC ($MeOH/CH_2Cl_2$ 1:9): R_f 0.32. M.p. 82.0–84.0°. IR (NaCl, film): 3303, 2950, 1666, 1633, 1537, 1313, 1101, 1081, 860, 742. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.07 (*s*, 9 H); 0.92 (*t*, *J* = 8.3, 2 H); 1.41–1.48 (*m*, 1 H); 2.11–2.16 (*m*, 1 H); 2.39–2.50 (*m*, 1 H); 2.99 (br. *s*, 1 H); 3.52 (*s*, 3 H); 3.60 (*t*, *J* = 8.3, 2 H); 3.59–3.65 (*m*, 1 H); 3.70 (*dd*, *J* = 10.0, 3.5, 1 H); 3.90 (*d*, *J* = 9.5, 1 H); 4.18–4.22 (*m*, 1 H); 4.40–4.45 (*m*, 1 H); 5.68 (*d*, *J* = 10.2, 1 H); 5.75 (*d*, *J* = 10.2, 1 H); 6.26 (*dd*, *J* = 3.6, 2.9, 1 H); 7.05–7.08 (*m*, 2 H); 7.32 (*d*, *J* = 7.5, 1 H); 7.65 (*s*, 1 H); 11.58 (br. *s*, 1 H). ^{13}C-NMR (75 MHz, $CDCl_3$): −1.5; 17.7; 30.1; 45.1; 51.6; 57.7; 62.7; 66.3; 73.1; 77.0; 86.2; 109.1; 117.3; 123.4; 127.2; 129.4; 139.2; 158.4; 160.7; 161.5. HR-MS (CI, NH_3): 511.2041 ($C_{22}H_{35}N_4O_6SSi^+$, MH^+; calc. 511.2047).

*2-Benzamido-*N^5*-[(1*R*,2*S*,3*R*,4*R*)-3-Hydroxy-4-(hydroxymethyl)-2-methoxycyclopentyl]thiazole-5-carboxamide* (**L2**). To a soln. of **15** (86.6 mg, 137 μmol) in anh. THF (0.75 ml) was added $Bu_4NF \cdot 3H_2O$ (172 mg, 545 μmol) in anh. THF (0.5 ml) dropwise at 0°. After stirring for 70 min at r.t., the mixture was poured into sat. $NaHCO_3$ soln. (5 ml) and extracted with AcOEt (4 × 15 ml). The combined org. layers were washed with brine (5 ml), dried ($MgSO_4$), and evaporated. Purification of the crude product by FC ($MeOH/CH_2Cl_2$ 1:9) yielded **L2** (47.3 mg) which contained residual quaternary ammonium salt. Further purification by cation-exchange chromatography (sulfopropyl *Sephadex C25*, $MeOH/H_2O$ 1:1) afforded pure **L2** (29.2 mg, 54%). White solid. TLC ($MeOH/CH_2Cl_2$): R_f 0.25. M.p. 234.0–235.0°. UV ($EtOH/H_2O$ 1:29): 224 (34200), 260 (8155), 302 (32600). IR (NaCl, film): 3318, 3062, 2906, 1682, 1614, 1532, 1307, 1184, 1116, 1029, 889, 723. ^{1}H-NMR (300 MHz, CD_3OD): 1.28–1.38 (*m*, 1 H); 2.07–2.14 (*m*, 1 H); 2.07–2.14 (*m*, H); 2.27–2.37 (*m*, 1H); 3.47 (*s*, 3 H); 3.54–3.63 (*m*, 3 H); 3.99–4.02 (*m*, 1 H); 4.35–4.42 (*m*, 1 H); 7.51–7.65 (*m*, 3 H); 8.01–8.04 (*m*, 3 H). ^{13}C-NMR (75 MHz, CD_3OD): 31.0; 46.6; 53.8; 58.2; 64.2; 72.8; 87.2; 128.7; 129.1; 129.8; 133.8; 133.9; 141.0; 163.4; 163.9; 168.0. HR-MS (CI, NH_3): 392.1292 ($C_{18}H_{22}N_3O_5S^+$, MH^+; calc. 392.1280).

N^5*-((1*R*,2*S*,3*R*,4*R*)-4-{[Bis(4-methoxyphenyl)(phenyl)methoxy]methyl}-3-hydroxy-2-methoxycyclopentyl)-2-{[(1-{[2-(trimethylsilyl)ethoxy]methyl}-1*H*-pyrrol-2-yl)carbonyl]amino}thiazole-5-carboxamide* (**17**). A

soln. of **16** (84 mg, 164 μmol) in anh. pyridine (1.0 ml) was added to dried (2 h oil-pump vacuum) DMT-Cl (72 mg, 213 μmol) and $(Bu_4N)ClO_4$ (73 mg, 213 μmol) under Ar at r.t. After stirring for 3 h at r.t., the mixture was poured into sat. $NaHCO_3$ soln. (5 ml) and extracted with CH_2Cl_2 (3 × 10 ml). The combined org. layers were washed with brine (5 ml), dried ($MgSO_4$), and evaporated. Purification of the crude product by FC ($MeOH/Et_3N/CH_2Cl_2$ 1:2:97) and drying for 2 d under vacuum afforded **17** (120 mg) as a white foam which contained residual amounts of Et_3N and Bu_4N^+ salt as impurities (by NMR: 88 mg (66%) pure **17**)[10]. TLC ($MeOH/CH_2Cl_2$ 1:19): R_f 0.29. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.05 (*s*, 9 H); 0.93 (*t*, *J* = 8.3, 2 H); 1.21–1.31 (*m*, 1 H); 2.21–2.27 (*m*, 1 H); 2.41–2.51 (*m*, 1 H); 3.37–3.51 (*m*, 1 H); 3.51 (*s*, 3 H); 3.56–3.62 (*m*, 2 H); 3.59 (*t*, *J* = 8.3, 2 H); 3.37–3.76 (*m*, 1 H); 3.76 (*s*, 6 H); 4.00–4.04 (*m*, 1 H); 4.38–4.43 (*m*, 1 H); 5.70 (*s*, 2 H); 6.02 (*d*, *J* = 7.4, 1 H); 6.24 (*dd*, *J* = 3.9, 2.7, 1 H); 6.80 (*d*, *J* = 8.3, 4 H); 7.02 (*dd*, *J* = 3.9, 1.6, 1 H); 7.06 (*dd*, *J* = 2.7, 1.6, 1 H); 7.16–7.31 (*m*, 7 H); 7.40 (*d*, *J* = 8.5, 2 H); 7.66 (*s*, 1 H). HR-MS (FAB, *m*NBA, NaCl): 835.3213 ($C_{43}H_{52}N_4NaO_8SSi^+$, *M*$Na^+$; calc. 835.3173).

*2-Benzamido-*N^5*-((1*R*,2*S*,3*R*,4*R*)-4-{[bis(4-methoxyphenyl)(phenyl)methoxy]methyl}-3-hydroxy-2-methoxycyclopentyl)thiazole-5-carboxamide* (**18**). A soln. of **L2** (43 mg, 110 μmol) in anh. pyridine (1.0 ml) was added to dried (2 h oil-pump vacuum) DMT-Cl (48 mg, 142 μmol) and $(Bu_4N)ClO_4$ (49 mg, 142 μmol) under Ar at r.t. After stirring for 3 h at r.t., the mixture was poured into sat. $NaHCO_3$ soln. (5 ml) and extracted with CH_2Cl_2 (3 × 10 ml). The combined org. layers were washed with brine (5 ml), dried ($MgSO_4$), and evaporated. Purification by FC ($MeOH/Et_3N/CH_2Cl_2$ 1:4:40) and drying for 2 d under vacuum afforded **18** (77 mg) as a clear, pale-yellow oil which contained residual Et_3N as an impurity (by NMR: 55 mg (72%) of pure **18**)[10]. TLC ($MeOH/Et_3N/CH_2Cl_2$ 1:4:40): R_f 0.09. ^{1}H-NMR (300 MHz, $CDCl_3$): 1.51–1.61 (*m*, 1 H); 2.18–2.26 (*m*, 1 H); 2.35–2.44 (*m*, 1 H); 3.11–3.19 (*m*, 3 H); 3.44 (*s*, 3 H); 3.60–3.64 (*m*, 1 H); 3.71 (*s*, 6 H); 3.96–4.00 (*m*, 1 H); 4.34–4.42 (*m*, 1 H); 5.25 (*d*, *J* = 8.7, 1 H); 6.75 (*d*, *J* = 8.7, 4 H); 7.11–7.27 (*m*, 7 H); 7.36–7.51 (*m*, 5 H); 7.71 (*s*, 1 H); 7.96 (*d*, *J* = 7.2, 2 H).

N^5*-((1*R*,2*S*,3*R*,4*R*)-4-{[Bis(4-methoxyphenyl)(phenyl)methoxy]methyl}-3-{[(2-cyanoethoxy)(diisopropylamino)phosphino]oxy}-2-methoxycyclopentyl)-2-{[(1-{[2-(trimethylsilyl)ethoxy]methyl}-1*H*-pyrrol-2-yl)carbonyl]amino}thiazole-5-carboxamide* (**19**). At 0° under Ar, chloro(2-cyanoethoxy)(diisopropylamino)phosphine (37 μl, 166 μmol) was added dropwise to a soln. of **17** (73 mg, 89 μmol) and $EtN(i\text{-}Pr)_2$ (75 μl, 442 μmol) in anh. CH_2Cl_2 (1 ml). After stirring for 20 min at 0° and 6.5 h at r.t., the mixture was directly subjected to FC (10% Et_3N in AcOEt/hexane 1:9, 4:6, 1:1, 8:2) affording **19** (80 mg, 89%) as a white foam (mixture of the two diastereoisomers). TLC (10% Et_3N in AcOEt/hexane 2:8): R_f 0.40. IR (NaCl, film): 3279, 3059, 2965, 2876, 2252, 1657, 1627, 1538, 1509, 1465, 1312, 1250, 1083, 1034, 837, 733. ^{1}H-NMR (300 MHz, $CDCl_3$): −0.04 (*s*, 9 H); 0.95 (*t*, *J* = 8.2, 2 H); 1.04–1.36 (*m*, 13 H); 2.38–2.44 (*m*, 1 H); 2.62 (*t*, *J* = 6.3, 2 H); 2.70–2.77 (*m*, 1 H); 3.61 (*d*, *J* = 5.1, 1 H); 3.47 (*s*, 3 H); 3.51–3.62 (*m*, 4 H); 3.61 (*t*, *J* = 8.2, 2 H); 3.77 (*s*, 6 H); 3.78–3.91 (*m*, 2 H); 4.17–4.23 (*m*, 1 H); 4.36–4.41 (*m*, 1 H); 5.71 (*s*, 2 H); 6.02 (*d*, *J* = 6.7, 1 H); 6.26 (*dd*, *J* = 3.8, 2.6, 1 H); 6.78–6.82 (*m*, 4 H); 7.02 (*dd*, *J* = 3.8, 1.6, 1 H); 7.07 (*dd*, *J* = 2.6, 1.6, 1 H); 7.17–7.34 (*m*, 7 H); 7.39–7.46 (*m*, 2 H); 7.70 (*s*, 1 H). HR-MS (FAB, *m*NBA, NaCl): 1035.4213 ($C_{52}H_{69}N_6NaO_9PSSi^+$, *M*$Na^+$; calc. 1035.4251).

*2-Benzamido-*N^5*-((1*R*,2*S*,3*R*,4*R*)-4-{[Bis(4-methoxyphenyl)(phenyl)methoxy]methyl}-3-{[2-cyanoethoxy)(diisopropylamino)phosphino]oxy}-2-methoxycyclopentyl)thiazole-5-carboxamide* (**20**). At 0° under Ar, chloro(2-cyanoethoxy)(diisopropylamino)phosphine (27 μl, 119 μmol) was added dropwise to a soln. of **18** (55 mg, 80 μmol) and $EtN(i\text{-}Pr)_2$ (54 μl, 317 μmol) in anh. CH_2Cl_2 (1 ml). After stirring for 20 min at 0° and 3 h at r.t., the mixture was directly subjected to FC (10% Et_3N in AcOEt/hexane 1:9, 3:7, 1:1, 8:2) affording **20** (48 mg, 67%) as a white solid (mixture of the two diastereoisomers). TLC (10% Et_3N in AcOEt/hexane 8:2): R_f 0.13. M.p. 113.0–117.0°. IR (NaCl, film): 3264, 3060, 2966, 2934, 2253, 1646, 1622, 1538, 1506, 1464, 1297, 1253, 1180, 1074, 1032, 829. ^{1}H-NMR (300 MHz, $CDCl_3$): 0.97–1.63 (*m*, 13 H); 2.40–2.84 (*m*, 4 H); 3.15–3.90 (*m*, 8 H); 3.47 (*s*, 3 H); 3.67 (*s*, 6 H); 4.18–4.24 (*m*, 1 H); 4.35–4.40 (*m*, 1 H); 6.11 (*d*, *J* = 6.7, 1 H); 6.80 (*d*, *J* = 8.8, 1 H); 7.16–7.62 (*m*, 13 H); 7.98 (*d*, *J* = 7.4, 2 H). HR-MS (FAB, *m*NBA, NaCl): 916.3498 ($C_{48}H_{56}N_5NaO_8PS^+$, *M*Na^+; calc. 916.3485).

N^5*-[(1*R*,2*S*,3*R*,4*R*)-3-Hydroxy-4-(hydroxymethyl)-2-methoxycyclopentyl]-2-[(1*H*-pyrrol-2-ylcarbonyl)amino]thiazole-5-carboxamide* (**L1**). A soln. of **16** (20 mg, 39 mmol) in anh. CF_3COOH (1 ml) was stirred for 3 min at r.t. The CF_3COOH was evaporated at 40° under reduced pressure. After repeating this procedure once, the residue was dissolved in sat. $NaHCO_3$ soln. (2 ml) and washed with CH_2Cl_2 (3 × 3 ml). The aq. phase was acidified to pH 6.5 by the addition of sat. NH_4Cl soln. (30 ml) and conc. HCl soln. (2 drops), and extracted with AcOEt (3 × 50 ml). The combined org. phases were washed with brine (10 ml), dried ($MgSO_4$), and evaporated. Purifica-

[10]) Although Et_3N was added, some cleavage of the DMT group was observed during FC. Due to the lability of the tritylated product, it was used without further purification for the subsequent reaction.

tion by FC afforded **L1** (6 mg, 40%). White solid. TLC ($MeOH/CH_2Cl_2$ 1:9): R_f 0.19. M.p. > 280°. IR (NaCl, film): 3418, 3226, 2923, 1660, 1615, 1538, 1393, 1325, 1131, 1042, 861, 738. ^{1}H-NMR (300 MHz, (D_6)DMSO): 1.12–1.23 (*m*, 1 H); 1.87–1.96 (*m*, 1 H); 2.00–2.10 (*m*, 1 H); 3.32 (*s*, 3 H); 3.32–3.45 (*m*, 3 H); 3.83–3.87 (*m*, 1 H); 4.23–4.28 (*m*, 1 H); 4.42 (br. *s*, 1 H); 4.63 (br. *s*, 1 H); 6.19–6.21 (*m*, 1 H); 7.08–7.10 (*m*, 1 H); 7.32–7.34 (*m*, 1 H); 8.11 (*s*, 1 H); 8.34 (*d*, J = 8.1, 1 H); 11.98 (br. *s*, 1 H); 11.41 (br. *s*, 1 H). HR-MS (CI, NH_3): 381.1221 ($C_{16}H_{21}N_4O_5S^+$, MH^+; calc. 381.1223).

3. Synthesis of Oligonucleotides. – Oligonucleotides were synthesized on a 1-μmol scale using an *Applied Biosystems 394* DNA/RNA synthesizer employing 2-cyanoethyl phosphoramidite chemistry. All couplings of T_d and $^{5Me}C_d$ phosphoramidites were performed according to the manufacturer's 1-μmol cyanoethyl phosphoramidite chemistry protocol [34]. Phosphoramidites **19** and **20** were coupled manually by passing 150 μl of a soln. of the phosphoramidite in anh. MeCn (**19**: 0.13M, **20**: 0.18M) and of 0.45M 1*H*-tetrazole in anh. MeCN (150 μl) *via* syringes through the columns containing the controlled pore glass (CPG) attached 5′-deprotected oligonucleotide. After 1 h coupling at r.t., the columns were washed with anh. MeCN (2 × 2.5 ml) and dried with Ar. The subsequent capping and oxidation steps were carried out by standard automated protocols.

Following the solid-phase synthesis of oligonucleotide **OL-1**, the 5′-DMT group was removed by standard automated deprotection[11]). The columns were washed with anh. MeCN (2 × 2.5 ml) and dried with Ar[12]). Cleavage of the SEM protecting group was carried out by treatment of the CPG-bound oligonucleotide with anh. CF_3COOH (0.5 ml) for 7 min at r.t., washing with anh. MeCN (3 × 2.5 ml) and drying with Ar. This procedure was repeated once[5]). The oligonucleotide was deprotected and cleaved from the solid support by treatment with conc. aq. NH_3 for 24 h at r.t. Purification by 20% polyacrylamide-gel electrophoresis and desalting with a *Waters Sep-Pak® Plus C18* cartridge afforded **OL-1** in 17% yield (168 nmol). HPLC: t_R 21.4. MALDI-TOF-MS: 4634.1 ($C_{156}H_{206}N_{37}O_{98}P_{14}S^-$, $[M - H]^-$; calc. 4633.3).

The 5′-DMT protected oligonucleotide **OL-2** was cleaved from the solid support and deprotected for 24 h at r.t. with conc. aq. NH_3[13]). The strill DMT-protected oligonucleotide **OL-2** was purified by reversed-phase FPLC (0 min 0% *B*, 10 min 0% *B*, 15 min 35% *B*, 40 min 75% *B*, 45 min 100% *B*, 50 min 100% *B*) and the DMT protecting group was cleaved by treatment with 80% aq. AcOH for 15 min at r.t. Purification by reversed-phase FPLC (0 min 0% *B*, 10 min 0% *B*, 40 min 50% *B*, 45 min 100% *B*, 58 min 100% *B*) and desalting on a *Pharmacia NAP-5* column afforded **OL-2** in 5% yield (45 nmol). HPLC: t_R 21.7. MALDI-TOF-MS: 4643.6 ($C_{158}H_{207}N_{36}O_{98}P_{14}S^-$, $[M - H]^-$; calc. 4644.3).

Oligonucleotides **OL-1** and **OL-2** were quantitated by UV absorbance at 260 nm, using the extinction coefficients of 8800 $M^{-1}cm^-$ for thymidine, 5700 $M^{-1}cm^-$ for 5-methyl-2′-deoxycytidine, 3500 $M^{-1}cm^-$ for **L1** and 8200 $M^{-1}cm^{-1}$ for **L2**.

4. Enzymatic Oligonucleotide Digest and HPLC Analysis. – The purified oligonucleotides (5 nmol) were digested with 3 units of calf-intestine alkaline phosphatase and 0.009 units of snake-venom phosphodiesterase from *Crotalus durissus* (both from *Boehringer Mannheim*) in 50 μl of 50 mM *Tris* · HCl, 10 mM $MgCl_2$ at pH 8.0 and 37° for 12 h and analyzed by HPLC. Comparison with $^{5Me}C_d$ (t_R 13.6), T_d (t_R 15.3), **L1** (t_R 27.9), and **L2** (t_R 32.7) as authentic standards confirmed the identity of the nucleosides, and integration of the peak areas confirmed the nucleoside composition of the oligonucleotides **OL-1** and **OL-2**.

5. Preparation and Labeling of DNA Restriction Fragment. – A plasmid containing a 314 bp *Afl II/Fsp I* restriction fragment with four 15 base pair target sites for triple-helix formation (see *Fig. 4*) was digested, 3′-labeled with [α-^{32}P]deoxyadenosine 5′-triphosphate and [α-^{32}P]thymidine 5′-triphosphate and purified as previously described in [19].

6. DNase I Footprinting. – DNase I footprinting titrations, quantitation of the denaturing polyacrylamide gels using storage phosphor imaging and data analysis were carried out as described in [19]. Briefly, the oligonucleotides **OL-1** and **OL-2** were diluted to a total of 20 concentrations spanning a range from 0.1 nM to 80 μM (see *Figs. 5*

---

[11]) Coupling yields by DMT cation assay: 55% overall; 65% for the coupling of **L1**.

[12]) A sample of the SEM-protected oligonucleotide was taken out of the column, deprotected with conc. aq. NH_3 for 24 h at r.t. and analyzed by HPLC (t_R 24.3 min) and MALDI-TOF-MS: 4763.9 ($C_{162}H_{221}N_{37}O_{99}P_{14}SSi^-$, $[M - H]^-$; calc. 4763.5).

[13]) Coupling yields by DMT cation assay: 23% overall; 30% for the coupling of **L2**.

and 6). A buffered stock soln. containing the radiolabeled 314 bp restriction fragment was added to give the final soln. conditions: 100 mM NaCl, 10 mM 'bis-tris', 250 μM spermine, 5 μM (in base pairs) calf-thymus DNA, 20000 cpm labeled restriction fragment, pH 7.0. After equilibration for 72 h at 23°, 5 μl of a soln. containing 1.25 milli units DNase I in 50 mM $CaCl_2$, 50 mM $MgCl_2$, 10 mM 'bis-tris', pH 7.0, and 5% glycerol was added, and digestion was allowed to proceed for 5 min at r.t. The DNase I reactions were quenched by the addition of 8.3 μl of a soln. containing 1.4M NaCl, 0.14M EDTA, pH 8.0, and 0.35 mg/ml of glycogen. The DNA reaction products were precipitated with EtOH and separated on an 8% denaturing polyacrylamide sequencing gel run in 1 × TBE (100 mM *Tris*, 90 mM boric acid, 1.0 mM EDTA, pH 8.3). The gel was dried for 1 h at 80° on a slab drier and exposed to a storage phosphor screen (*Molecular Dynamics*) overnight at r.t.

Data from the footprint titrations were obtained by quantitation of the sequencing gels using storage phosphor imaging. Volume integrations of rectangles encompassing the footprint site(s) and a reference site at which DNase reactivity was invariant across the titration generated values for the site intensity (I_{site}) and reference intensity (I_{ref}). These I_{site} and I_{ref} values were corrected by subtracting out background intensity as described in the detailed protocol. The apparent fractional occupancy (θ_{app}) at the site at each oligonucleotide concentration ($[O]_{tot}$) was calculated from the corrected I_{site} and I_{ref} values using the following equation:

$$\theta_{app} = 1 - \frac{I_{site}/I_{ref}}{I^0_{site}/I^0_{ref}}$$

from the where I^0_{site} and I^0_{ref} are the site and the reference intensities, respectively, from the control lane to which no third strand oligonucleotide has been added. The resulting pairs of (θ_{app}, $[O]_{tot}$) values were plotted on a semilogarithmic scale (see *Fig. 6*). The following binding isotherm was fit to the experimental data using a nonlinear least squares algorithm in the program KaleidaGraph 3.0.1 (*Abelbeck Software*):

$$\theta_{app} = \theta_{min} + (\theta_{max} - \theta_{min}) \frac{K_T[O]_{tot}}{1 + K_T[O]_{tot}}$$

where θ_{min} is the apparent fractional occupancy at the lowest oligonucleotide concentration, θ_{max} is the apparent fractional occupancy at saturation, and K_T is the equilibrium association constant. The report data are the average of three determinations plus or minus the standard error of the mean (± s.e.m.).

REFERENCES

[1] H. E. Moser, P. B. Dervan, *Science* **1987**, *238*, 645.
[2] T. LeDoan, L. Perrouault, D. Praseuth, N. Habhoub, J.-L. Decout, N. T. Thuong, J. Lhomme, C. Helene, *Nucleic Acids Res.* **1987**, *15*, 7749.
[3] N. T. Thuong, C. Helene, *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666.
[4] V. N. Soyfer, V. N. Potaman, 'Triple-Helical Nucleic Acids', Springer-Verlag, New York, 1996.
[5] J. S. Sun, T. Garestier, C. Helene, *Curr. Opin. Struct. Biol.* **1996**, *6*, 327.
[6] L. J. Maher, B. Wold, P. B. Dervan, *Science* **1989**, *245*, 725.
[7] G. Degols, J. P. Clarenc, B. Lebleu, J. P. Leonetti, *J. Biol. Chem.* **1994**, *269*, 16933.
[8] C. Giovannangeli, L. Perrouault, C. Escude, S. Gryaznov, C. Helene, *J. Mol. Biol.* **1996**, *261*, 386.
[9] M. Kochetkova, M. F. Shannon, *J. Biol. Chem.* **1996**, *271*, 14438.
[10] T. J. Povsic, P. B. Dervan, *J. Am. Chem. Soc.* **1990**, *112*, 9428.
[11] J.-P. Shaw, J. F. Milligan, S. H. Krawczyk, M. Matteucci, *J. Am. Chem. Soc.* **1991**, *113*, 7765.
[12] S. A. Strobel, P. B. Dervan, *Nature (London)* **1991**, *350*, 172.
[13] S. A. Strobel, L. A. Doucettestamm, L. Riba, D. E. Housman, P. B. Dervan, *Science* **1991**, *254*, 1639.
[14] K. B. Grant, P. B. Dervan, *Biochemistry* **1996**, *35*, 12313.
[15] D. A. Liberles, P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9510.
[16] T. Akiyama, M. E. Hogan, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12122.
[17] T. Akiyama, M. E. Hogan, *Biochemistry* **1997**, *36*, 2307.
[18] W. A. Greenberg, P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 5016.
[19] E. S. Priestley, P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 4761.
[20] S. O. Doronina, J.-P. Behr, *Chem. Soc. Rev.* **1997**, 63.
[21] I. Radhakrishnan, D. J. Patel, *Structure* **1994**, *2*, 17.
[22] B. C. Froehler, D. J. Ricca, *J. Am. Chem. Soc.* **1992**, *114*, 8320.
[23] C. Escude, J.-S. Sun, M. Rougee, T. Garestier, C. Helene, *C. R. Acad. Sci., Ser. III* **1992**, *315*, 521.

[24] M. Shimizu, A. Konishi, Y. Shimada, H. Inoue, E. Ohtsuka, *FEBS Lett.* **1992**, *302*, 155.
[25] R. C. Cermak, R. Vince, *Tetrahedron Lett.* **1981**, *22*, 2331.
[26] B. L. Kam, N. J. Oppenheimer, *J. Org. Chem.* **1981**, *46*, 3268.
[27] B. H. Lipshutz, J. J. Pegram, *Tetrahedron Lett.* **1980**, *21*, 3343.
[28] J. M. Muchowski, D. R. Solas, *J. Org. Chem.* **1984**, *49*, 203.
[29] D. E. Bailey, R. E. Johnson, N. F. Albertson, *Org. Synth.* **1971**, *51*, 100.
[30] O. Dann, *Chem. Ber.* **1943**, *76*, 419.
[31] M. P. Reddy, J. B. Rampal, S. L. Beaucage, *Tetrahedron Lett.* **1987**, *28*, 23.
[32] M. J. Gait, 'Oligonucleotide Synthesis', IRC Press, Oxford, U.K., 1984.
[33] N. D. Sinha, J. Biernat, J. McManus, H. Koster, *Nucleic Acids Res.* **1984**, *12*, 4539.
[34] 'Models 392 and 394 DNA/RNA Synthesizers User Manual, System Software Version 1.01', Applied Biosystems, Foster City, California, USA, 1991.
[35] M. Brenowitz, D. F. Senear, M. A. Shea, G. K. Ackers, *Methods Enzymol.* **1986**, *130*, 132.
[36] J. Hunziker, E. S. Priestley, H. Brunar, P. B. Dervan, *J. Am. Chem. Soc.* **1995**, *117*, 2661.
[37] H. Brunar, P. B. Dervan, *Nucleic Acids Res.* **1996**, *24*, 1987.
[38] K. M. Koshlap, P. Schultze, H. Brunar, P. B. Dervan, J. Feigon, *Biochemistry* **1997**, *36*, 2659.
[39] J. S. Koh, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 1470.
[40] I. Radhakrishnan, D. J. Patel, E. S. Priestley, H. M. Nash, P. B. Dervan, *Biochemistry* **1993**, *32*, 11228.
[41] L. C. Griffin, P. B. Dervan, *Science* **1989**, *245*, 967.
[42] I. Radhakrishnan, D. J. Patel, J. M. Veal, X. L. Gao, *J. Am. Chem. Soc.* **1992**, *114*, 6913.
[43] L. C. Griffin, L. L. Kiessling, P. A. Beal, P. Gillespie, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 7976.
[44] D. A. Liberles, Ph. D. Thesis, California Institute of Technology, 1997.
[45] E. Wang, K. M. Koshlap, P. Gillespie, P. B. Dervan, J. Feigon, *J. Mol. Biol.* **1996**, *257*, 1052.
[46] C. Y. Huang, P. S. Miller, *J. Am. Chem. Soc.* **1993**, *115*, 10456.
[47] C. Y. Huang, G. X. Bi, P. S. Miller, *Nucleic Acids Res.* **1996**, *24*, 2606.
[48] S. C. Zimmermann, P. Schmitt, *J. Am. Chem. Soc.* **1995**, *117*, 10769.
[49] C. K. Wada, P. B. Dervan, unpublished results.
[50] W. A. Greenberg, Ph. D. Thesis, California Institute of Technology, 1997.
[51] T. E. Lehmann, P. B. Dervan, unpublished results.
[52] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, C. G. T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440.