Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 107–119

BIOORGANIC & MEDICINAL CHEMISTRY

Synthesis and Characterization of a Tetranucleotide Analogue Containing Alternating Phosphonate-Amide Backbone Linkages

Peilin Yu,[a,†] Wei Wang,[a,†,‡] Hua Zhang,[a]
Xueyong Yang,[a] T. Chyau Liang[b] and Xiaolian Gao[a,*]

[a]*Department of Chemistry, University of Houston, Houston, TX 77204-5641, USA*
[b]*Department of Biochemistry and Molecular Biology, University of Texas, Medical School, Houston, TX 77225, USA*

Received 2 June 2000; accepted 11 August 2000

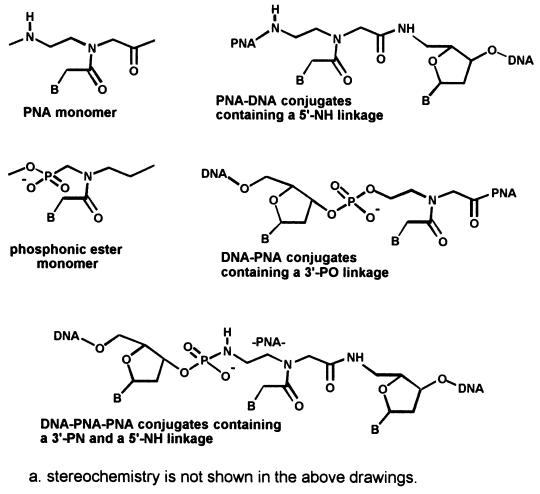
Abstract—Described herein is the synthesis and characterization of a tetranucleotide, 5′-dC-$_{phosphonate}$-T-$_{amide}$-T-$_{phosphonate}$-dC (**III**), in which the C-T and T-C steps contain a phosphonate backbone bond and T-T is a peptide nucleic acid dimer unit (neutral backbone). The 5′- and 3′-OH groups of the tetramer can be further derivatized and, thus, the compound is a potential building block for longer oligonucleotides which will contain alternating backbone modifications at designated positions. The synthesis involved first the preparation of two hybrid peptide-deoxyribose dinucleotides, CT-CO (**I**) and N-CT (**II**) (C and T are nucleobases; CO and N are carboxylic and amino terminal, respectively); each is linked through a phosphonate linkage. A condensation reaction between the two dimers, followed by deprotection, resulted in the formation of a peptide linkage to give the desired tetramer **III**. The reaction conditions used are mild to afford products in moderate to excellent yields. The DNA-PNA-DNA tetramer, d(CTTC), is a substrate for T4 kinase but fails to give a ligation product, even though NMR shows weak interactions between the tetramer **III** with its complementary sequence, d(GAAG).

Introduction

The general utility of oligonucleotides, which include ribo- and deoxyribo-oligonucleotides and their analogues (antisense oligonucleotides, AONs), as gene-specific ligands has been demonstrated in a wide range of applications. The binding of these molecules to selected sites in mRNA for inhibition of the initiation events of gene expression, both in vitro and in vivo, has in particular attracted great attention.[1–5] However, for in vivo antisense applications, a major hurdle yet to overcome is to improve the extra- and intra-cellular stability of AONs without sacrificing their cell permeability, binding affinity and specificity. Since the degradation of these compounds is often due to cleavage of the phosphodiester backbone bonds by nucleases, considerable effort has been invested to synthesize and test backbone modified oligonucleotides to achieve desirable chemical and biological properties.[6–11] Among the library of backbone modifications, phosphorothioate (PS) linkers[12] have been most promising in balanced chemical, biochemical and biomedical properties. However, the general applications of the PS oligonucleotides are hampered by the presence of stereo-isomers, the moderate binding affinity and the interactions of the PS sulfur with other biomolecules, such as proteins. Families of new analogous oligonucleotides are quickly expanding. Among these, several neutral and/or achiral backbone linkers, such as *N*-methylenemethyl imino (3′-$CH_2N(CH_3)O$-, MMI),[13] amide (3′-$CH_2HNC(O)$- or 3′-$CH_2C(O)NH$-),[14] formacetal (3′-OCH_2O-, FMA),[15] and 3′-thioformacetal (3′-SCH_2O-, TFMA)[16] moieties, have been prepared and shown to exhibit improved nuclease resistance and binding affinities comparable to natural oligonucleotides. Furthermore, fundamental progress has been made in the design of entirely new backbone and sugar moieties. The success in this area may be highlighted by the nucleotide analogues containing nucleobases attached to an *N*-(2-aminoethyl)glycine backbone via an acetoamide linkage (peptide linked nucleic acids or PNA) (Scheme 1(A)).[17] Reported in 1991, PNA has demonstrated great promise in meeting the criteria as stable, effective, and specific reagents, which recognize nucleic acids.[8] The PNA oligomers hybridize with single stranded or duplex DNA and RNA to form Watson–Crick

*Corresponding author. Tel.: +1-713-743-2805; fax: +1-713-743-2709; e-mail: gao@kitten.chemn.uh.edu
†These authors have equal contributions to this work.
‡Present address: Department of Neuro-Oncology, M. D. Anderson Cancer Center, University of Texas, Houston, TX 77030, USA.

0968-0896/01/$ - see front matter
PII: S0968-0896(00)00230-3



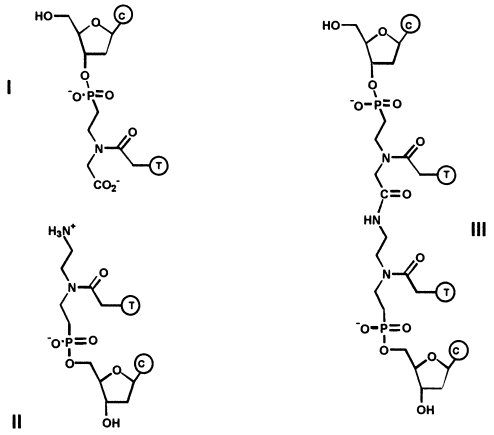
a. stereochemistry is not shown in the above drawings.

Scheme 1(A). PNA and its analogues reported in the literature.[a]

base paired hybrids. These PNA·DNA/RNA multiplexes are by far the most stable nucleic acids complexes containing synthetic nucleotide analogues. There is great expectation on PNA to become a new generation of antisense gene regulation therapeutics. To achieve this goal, some properties of PNA should be improved. PNA molecules without a charged cap/tail tend to aggregate in aqueous solutions, rendering difficulties in sample handling and quantitative binding studies. Additionally, PNA•RNA hybrids are no longer substrates for RNase H. Since the induced RNA degradation by RNase H is an important pathway of the antisense actions, the inhibition of the enzymatic activities by PNA binding has caused some concerns. To improve the solubility of the PNA compounds in aqueous solutions and to comply with the needs of biological applications, the preparations of a modified PNA and several conjugated PNA-DNA chimeras have been reported.[11,18–24] Most relevant to our work, in the phosphonic ester nucleic acids (Scheme 1(A)),[22] a charged phosphonic group partially substitutes the PNA neutral backbone. A T_9 sequence hybridizes with DNA A_9 to give a $T_m \sim 21\,^{\circ}C$, which is 1° lower than that of the natural DNA T_9 and A_9 duplex. The PNA-DNA chimeras are linked through $(PNA)_n$-C(O)-5′-NH-$(DNA)_n$, $(DNA)_n$-3′-PO-$(PNA)_n$ or $(DNA)_n$-3′-PNH-$(PNA)_n$ linkages (n is the monomer repeat number).[18–21] These PNA-DNA chimeras have shown several promising properties. These molecules form considerably more stable duplexes with RNA and DNA complementary strands; they exhibit improved cell uptake[24] and are more resistant to 3-*exo*-nucleases.[18] More importantly, PNA-DNA·RNA hybrids allow RNase H cleavage of the RNA residues that bind to the DNA segment.[21,24] Further studies are underway to characterize the physicochemical and biological properties of this new family of oligonucleotide analogues.[19,20]

Our research in antisense oligonucleotides has been focused on structural characterization of backbone modifications, which include neutral and achiral phosphodiester surrogates, such as FMA,[15] 3′-TFMA,[16] and MMI[13] backbone linkages. Using NMR spectroscopy and organic synthesis in combination, we investigate the relationship between chemical structures and properties of modified oligonucleotides and the effect of these properties on the recognition of DNA and RNA.[25] A series of AON oligonucleotides containing single or alternating backbone linkage modifications have been studied in detail. These comparisons of the related AON·DNA and AON·RNA duplexes have permitted a detailed view of the structures at the global and local levels. The analyses of the structures and the thermodynamic stabilities indicate that the continuity or the overall structural homogeneity of the modified oligonucleotides plays an important role in achieving their high affinity binding to RNA sequences. These results complement the current view that desirable antisense properties may be better achieved in molecular systems of chimeric charge/neutral sequences. Based on this assessment and in pursuit of new antisense oligonucleotides, we rationalized that a sequence comprised of PNA dimers separated by phosphodiester linkages (two or more units) may possess improved solubility, while they can maintain the binding affinity to natural nucleic acids and the susceptibility to RNase H. These chimeric molecules will be used to examine how the PNA dimer units are accommodated in various AONs and to derive information that is valuable for achieving optimal design of new backbone modified oligonucleotides.

In this paper, we report the synthesis and characterization of two phosphonate dimers (**I** and **II**) and the 5′-dC-$_{phosphonate}$-T-$_{amide}$-T-$_{phosphonate}$-dC tetramer (**III**), containing alternating phosphonate and peptide linkages (Scheme 1(B)). The possibility that the DNA-PNA-DNA tetramer **III** would serve as a substrate for various nucleic acid enzymes was investigated. Conceptually, the layout of the alternating backbone linkers in **III** here considerably differs from the reported DNA-PNA chimeras or the pseudo-PNA, or phosphonic ester backbones (Scheme 1). The modifications introduced in this work are located in alternating positions as dimer units. Our design maintains the 5′-phosphate group and employs a 3′-phosphonate linkage at the junction of the



a. C and T are nucleobases, CO and N denote the carboxylic and the amino terminus, respectively.

Scheme 1(B). B. Chemical structure of the CT-CO dimer, **I**, the N-TC dimer, **II**, and the CTTC tetramer **III**.[a]

PNA dimer and DNA monomers. It is, therefore, particularly interesting to gain knowledge on the basic physiochemical behaviors of the molecule.

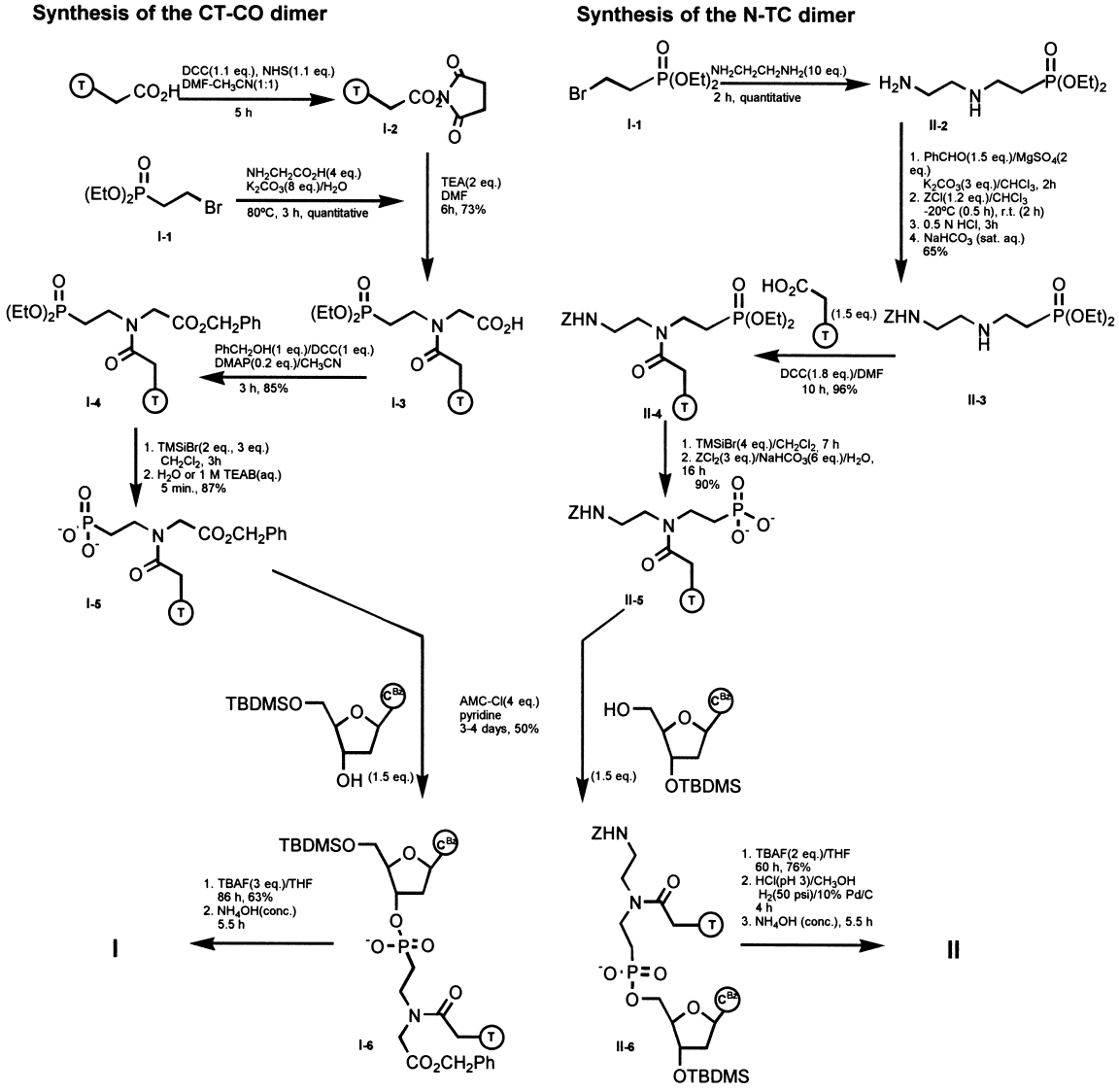
Results and Discussion

Synthesis

The synthetic procedures of the two dimers, **I** and **II**, were designed to resemble each other to simplify the overall synthesis [the intermediate products in the reactions described below are named according to whether the compound leads to the formation of **I** (**I-1** to **I-7**), **II** (**II-1** to **II-7**) or **III** (**III-1**)]. The initial work resulted in a useful method for the protection of a primary amino group in the presence of a secondary amino group via the formation of a semiaminal species.[26] The final synthesis of the tetramer **III** was achieved by joining the two hybrid dinucleotides, **I** and **II**, through an amide bond (Scheme 1(B)). Most reaction steps described herein have been carried out on multiple milligram scales in moderate to excellent yields. One- and two-dimensional (1-D and 2-D) NMR spectra were used to provide specific assignments of the ^{1}H, ^{13}C and ^{31}P resonances of the major intermediate compounds. The hybrid dimers (**I** and **II**) and the tetramer **III** were further characterized by electrospray mass analysis.

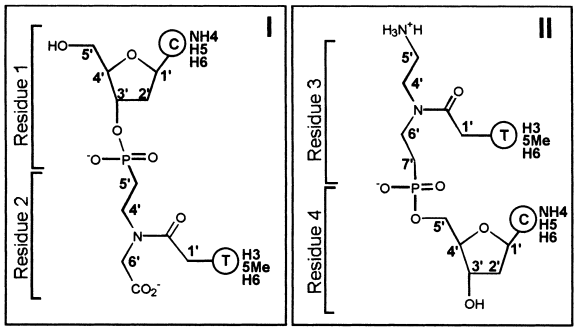
Synthesis of the DNA-PNA dimers

The synthesis of the CT-CO dimer, **I**, was achieved by preparing, first, the peptidyl phosphonate monomer, **I-4**, follwed by the fully protected dimer, **I-6** (Scheme 2). Thus, nucleophilic Br-substitution of diethyl-2-bromoethylphosphonate **I-1** with glycine was carried out under typical basic aqueous conditions to give a phosphonate with an extended alkyl chain. An activated thymin-1-yl-acetyl



Scheme 2. Synthetic procedures for the CT-CO and the N-TC dimers. AMC-CL = adamantanecarbonyl chloride; Bz = benzoyl; C = cytidine; DCC = dicyclohexylcarbodiimide; DMAP = 4-dimethylaminopyridine; DMF = dimethylformamide; Et = ethyl; NHS = *N*-hydroxysuccinimide; Ph = phenyl; T = thymine; TBAF = tetrabutylammonium flouride; TBDMS = *tert*-butydimethylsilyl; TEA = triethyl amine; TEAB = triethylammonium bicarbonate; THF = tetrahydrofuran; TMSiBr = trimethylsilyl bromide; Z = benzyloxycarbonyl.

ester **I-2** was prepared by reaction of thymin-1-yl-acetic acid with *N*-hydroxysuccinimide (NHS) in the presence of dicyclohexylcarbodiimide (DCC).[27] The condensation reaction between the chain length extended **I-1** and the activated ester **I-2** then proceeded in the presence of base TEA to give a peptidyl phosphonate T monomer **I-3,** with a carboxyl terminal group, in 73% yield. The carboxylic group of **I-3** was protected as a benzyl ester to yield the fully protected peptidyl phosphonate monomer **I.** The ethyl groups of the phosphonate were hydrolyzed by treatment with bromotrimethylsilane (TMSiBr)[28] followed by an aqueous solution to give, after reverse-phase HPLC purification, the phosphonic acid monomer **I-5** in 87% yield. To minimize the hydrolysis of the benzyl group it is important to finish the reaction in ~3 h and to slowly add a total of 5 equiv of TMSiBr. The coupling reaction involving **I-5** and 4-*N*-benzoyl-5′-*O*-(*tert*-butyl-dimethylsilyl)-dC[29] containing a free 3′-OH proceeded slowly in the presence of 1-adamantanecarbonyl chloride (AMC-Cl) to yield the protected CT-CO dimer, **I.** Our investigation indicated that, under similar conditions, other condensation reagents, such as pivaloyl chloride or triisopropylbenzenesulfonyl chloride, failed to produce **I.** Additionally, the reactions for synthesis of **I-6** involved formation of phosphonate-peptide linkages and alkylation of a phosphonic acid. Similar types of reactions were reported by Coward and co-workers for preparation of phosphonate glutamyl-γ-glutamate derivatives, in which a simpler alkylation route was to use Mitsunobu conditions.[30] These conditions were also employed by others for preparation of glycosyl phosphonate derivatives in moderate yields.[31] However, the Mitsunobu reaction in our hands did not generate the desired product (i.e. from **I-5** to **I-6**). The protected CT-CO dimer, **I-6**, was characterized by 1-D and 2-D ^{1}H NMR as described in the Experimental (the atom numbering system used in NMR spectral analysis is given in Scheme 3). Desilylation and ammonolysis of **I-6** afforded the CT-CO dimer, **I** (Scheme 2). **I** was thoroughly characterized using various 1-D and 2-D ^{1}H, ^{13}C and ^{31}P NMR experiments and mass spectrometry. The 1-D ^{1}H spectrum of **I** is displayed in Figure 1(A) and the ^{1}H, ^{13}C, and ^{31}P chemical shift assignments for **I** are summarized in Table 1. These results confirmed the formation of **I-6** and the presence of *cis*- and *trans*-amide bonds in this molecule.



a. shown for dimers **I** and **II**. The same numbering system is also applied to the tetramer **III**.

Scheme 3. ^{1}H resonance numbering system.

The strategy for synthesizing the N-TC dimer (**II**) was similar to what was described for the preparation of the CT-CO dimer **I** (Scheme 2). Thus, reacting the phosphonate **I-1** with ethylenediamine furnished a (2-*N*-(2-aminoethyl)aminoethyl)phosphonate, **II-2**, in quantitative yield. The formation of disubstituted ethylenediamine was negligible. The resultant primary amino group was selectively protected in the presence of a secondary amino group according to what was described previously[26] to give **II.** The condensation reaction was performed under standard conditions using DCC in the

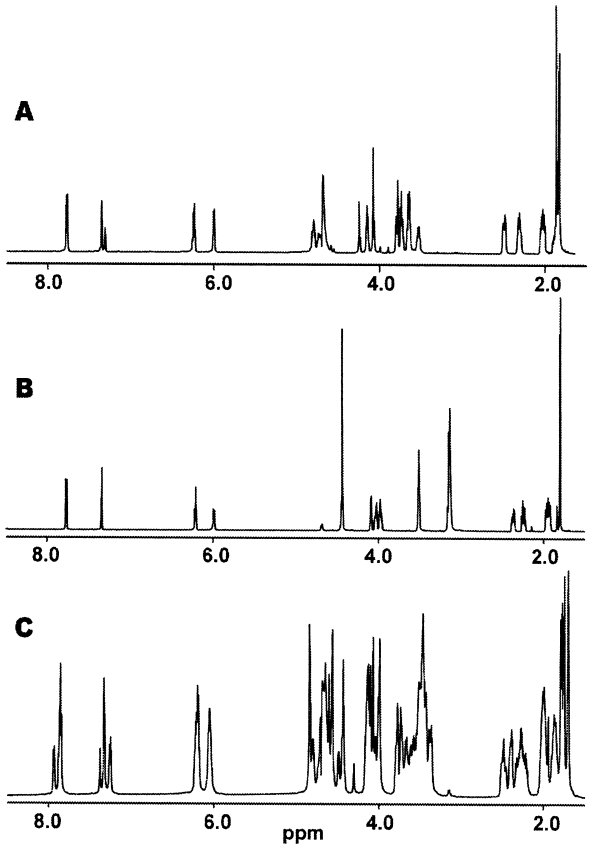


Figure 1. 1-D ^{1}H spectrum of (**A**) the CT-CO dimer **I**, (**B**) the N-TC dimer **II**, and (**C**) the CTTC tetramer **III**. All spectra were recorded in D_2O at 25 °C.

Table 1. ^{1}H, ^{13}C, and ^{31}P chemical shifts of the CT-CO dimer (**I**)[a]

Residue 1	^{1}H	^{13}C	Residue 2	^{1}H	^{13}C
1′	6.25	88.5	1′	4.82/4.57	51.3
2′,2″	2.31, 2.50	40.9	4′	3.65/3.51	46.4
3′	4.80	75.8	5′	2.03/1.88	28.4
4′	4.16	88.3	6′	4.25/4.08	52.1/51.3
5′,5″	3.74	63.3	5Me	1.87/1.84	13.6
5	6.00	98.7	6	7.36/7.31	145.7
6	7.78	144.1			
^{31}P	20.1/19.2				

[a]Resonance assignments were made from spectra recorded in D_2O at 25 °C. The atom numbering system is given in Scheme 3. ^{1}H assignments were made from TOCSY, DQF-COSY and ^{1}H–^{31}P COSY spectra, ^{13}C assignments were made from ^{1}H–^{13}C HMQC and HMBC spectra, ^{31}P chemical shifts were from 1-D ^{31}P and 2-D ^{1}H–^{31}P correlation spectra. Resonance assignments for *cis*- and *trans*-isomeric forms were separated by slashes.

presence of thymin-1-yl-acetic acid and **II-3** to yield a peptidyl phosphonate T monomer containing a terminal amino group, **II.** The phosphonate ester was hydrolyzed by treatment with TMSi-Br as described before (Scheme 2, formation of **I-5**). In this process, the amino Z-protecting group was also partially cleaved. However, the Z-group can be easily re-introduced in situ to give **II-5** in overall 90% yield. The formation of the protected N-TC dimer, **II-6**, was carried out under the conditions similar to those used for the CT-CO dimer, except that a 4-*N*-benzoyl-3′-*O*-(*tert*-butyl-dimethylsilyl)-dC was used to form a 5′-*O*-phosphonate linkage with the peptide T residue **II.** The protected dimer, **II-6**, was examined by 1-D and 2-D ^{1}H NMR to confirm the correct chemical structure. A small portion of **II-6** was completely deprotected to yield the N-TC dimer, **II** (Schemes 1 and 2). **II** was subject to the complete spin connectivity analysis based on the ^{1}H, ^{13}C and ^{31}P 1-D and 2-D correlation NMR spectra and was characterized by mass spectrometry. A 1-D ^{1}H spectrum of **II** is displayed in Figure 1(B) and the corresponding chemical shift assignments are provided in Table 2.

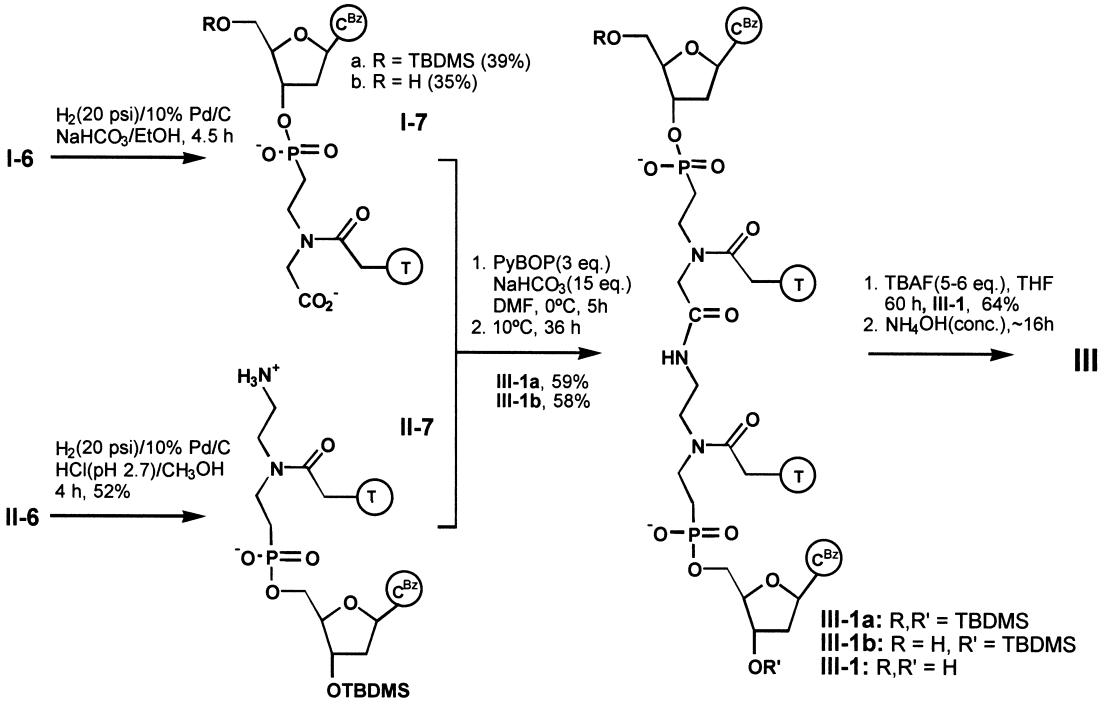
Table 2. ^{1}H, ^{13}C and ^{31}P chemical shifts of the N-TC dimer (**II**)[a]

Residue 3	^{1}H	^{13}C	Residue 4	^{1}H	^{13}C
1′	4.44	52.3	1′	6.20	88.7
4′	3.51	37.3	2′, 2″	2.24, 2.36	40.8
5′	3.15	44.9	3′	4.44	72.7
6′	3.14	48.6	4′	4.09	87.7
7′	1.95	24.4	5′,5″	3.97, 4.02	65.5
5Me	1.80	12.1	5	5.98	98.8
6	7.35	146.7	6	7.77	145.1
^{31}P	17.7				

[a]Experimental conditions and assignment definitions are given in Table 1. The assignments for 5′ and 6′ resonances in residue 3 are arbitrary.

Synthesis of the DNA-PNA-DNA tetramer III

In the synthesis of the tetramer **III**, the protecting groups on the C-terminal of the CT-CO dimer, **I-6**, or on the terminal NH_2 group of the N-TC dimer, **II-6**, were first removed (Scheme 4). The hydrogenolysis reaction conditions for removing the benzyl group from the CT-CO dimer, **I-6**, were optimized. Those that gave the highest yield were in a slightly basic alcohol solution using 10% Pd/C as catalyst and low pressure H_2 (Scheme 4). The presence of $NaHCO_3$ was found necessary for stabilizing the 4-*N*-benzoyl moiety of dC. The reaction was accompanied by partial desilylation of the 5′-*O*-TBDMS group to give **I-7a** as well as the 5′-*O*-desilylated **I-7b**, both of which contain a free terminal carboxyl group. The two products were separated using reverse-phase HPLC in total of 74% yield. Contrary to the basic hydrogenolysis of **I-6**, hydrogenolysis removal of the amino-Z-protecting group from the N-TC dimer, **II-6**, required acidic conditions. Under these conditions the protonated amino group presents no retardation effect on the reaction. Thus, in the presence of a methanol solution acidified to pH 2.7, the amino-deprotected N-TC dimer, **II-7**, was produced in reasonable yield. When the reaction solvent was a mixture of acetic acid and ethyl acetate (4:1), the product **II-7** was not observed, possibly due to the reduction of the C base. In contrast to our results, it was noted that hydrogenolysis did not cleave the Z-protecting group in the PNA compounds,[32] while acid cleavage using HBr/AcOH worked effectively. The condensation reaction between **I-7a** or **I-7b** with **II-7** was initiated at 0 °C in the presence of $NaHCO_3$ and a coupling reagent, PyBOP



Scheme 4. Synthetic procedures for the DNA-PNA-DNA tetramer **III**.

(benzotriazol-1-yl-oxy-trix-pyrrolidino-phosphonium-hexaflurophosphate) (Scheme 4). Under these conditions, cleaner products were obtained compared to those using diisopropyl ethylamine as base. The reaction was then allowed to continue at 10 °C, resulting in the formation of an amide linkage between the two dimers to give the phosphonate-amide tetramer with or without a 5′-*O*-TBDMS in 58–59% yields (Scheme 4). Neither the phosphonate group in **I-7** and **II-7** nor the 5′-OH in **I-7b** needed to be protected during the condensation reaction. The formation of a phosphoramidate bond was not observed as judged from the absence of the ^{1}H(NH)–^{31}P correlation in the ^{1}H–^{31}P 2-D COSY spectrum. Finally, deprotection of the 5′- and the 3′-*O*-TBDMS groups (64% yield) using TBAF and of the 4-*N*-benzoyl group using ammonolysis afforded 5′-dC-$_{\text{phosphonate}}$-T-$_{\text{amide}}$-T-$_{\text{phosphonate}}$-dC tetramer (CTTC), **III**. The HPLC purified product was analyzed by mass and a set of 1-D and 2-D multinuclear correlation experiments.

Characterization

NMR characterization of III

The CTTC tetramer **III**, which contains a neutral amide and charged phosphonate linkages, is readily soluble in aqueous solutions (most concentrated sample used was ~10 mM), permitting the acquisition of spectra under close to physiological conditions (Fig. 1(C)). The chemical shift assignments of the ^{1}H, ^{13}C, and ^{31}P resonances of **III** are listed in Table 3. The sugar protons of residues 1 and 4 were assigned by their *J*-coupling connectivities using DQF-COSY and TOCSY spectra. The base protons of C and T displayed COSY cross peaks along the aromatic proton frequencies (7–8 ppm, H6 of the C and T residues) due to coupling between C H6 and H5 at ~6 ppm and between T H6 and methyl protons at ~1.8 ppm. The methylene groups linked to ^{31}P in residues 2 and 3 were identified by their *J*-correlations in the ^{1}H–^{31}P COSY spectra, while the methylene groups adjacent to the carbonyl groups were assigned based on their two bond ^{1}H–^{13}C (C is carbonyl carbon) heteronuclear coupling connectivities. The heteronuclear correlation (^{1}H–^{13}C HMQC and HMBC and ^{1}H–^{31}P COSY) analyses also provided the assignments for the ^{13}C and ^{31}P resonances (Table 3).

Table 3. ^{1}H, ^{13}C, and ^{31}P chemical shifts of the CTTC tetramer (**III**)[a]

Residue 1	^{1}H	^{13}C	Residue 2	^{1}H	^{13}C
1′	6.25	88.6	1′	4.86	51.6
2′,2′	2.32/2.28, 2.48	41.2	4′	3.69/3.64	46.5
3′	4.82/4.74	75.6	5′	2.03/2.01	28.3
4′	4.16/4.14	88.2	6′	4.12/4.01	65.7
5′,5′	3.80/3.75	63.4	5Me	1.81/1.72	13.5
5	6.02	98.1	6	7.40/7.28	145.7
6	7.89/7.81	144.7			
^{31}P	20.9/20.2				

Residue 3	^{1}H	^{13}C	Residue 4	^{1}H	^{13}C
1′	4.61/4.58	51.4	1′	6.21	88.6
4′	3.44/3.37	39.1	2′,2′	2.27/2.22, 2.39	41.4
5′	3.60/3.47	39.1	3′	4.50/4.44	72.9
6′	3.55/3.51	47.9	4′	4.15/4.14	88.2
7′	2.01/1.89	27.9	5′, 5′	4.06,4.02	51.7
5Me	1.81/1.72	13.5	5	6.02	98.1
6	7.35/7.26	145.7	6	7.89/7.81	144.7
^{31}P	19.7/19.4/19.4				

[a]Experimental conditions and assignment definitions are given in Table 1. Some of the resonances displayed more than two isomeric forms and these are not all listed in the Table. For residues 2 and 3: H6 found at 7.4, 7.35, 7.34, 7.28, 7.27, and 7.26 ppm and 5Me was found at 1.81, 1.72, 1.77, 1.78, 1.81, and 1.71 ppm. The H5 and H6 assignments of residues 1 and 4 and 5Me and H6 of residues 2 and 3 are not residue-specific.

The presence of the correct backbone amide linkage between the two dimer units has been confirmed by a NOESY experiment undertaken at 5 °C in 90% H_2O–10% D_2O (Fig. 2). This spectrum demonstrated the interproton contacts linking the amide proton resonances at 8.1.4 ppm (F2 dimension, Figure 2) with T2 H6′ and T3 H4′ at 3.3.1 ppm (F1 dimension). These NOEs are unique for the given chemical structure of **III**, providing unambiguous identification of the indicated backbone linkage (Schemes 1 and 3). The existence of the *cis*- and *trans*-isomers due to the three amide bonds in **III** is evident as seen in the multiplicity of several resonances (Table 3 and Figure 2(C)). *cis*- And *trans*-isomers were also reported for the free PNA sequences.[33,34] When bound to a DNA or an RNA complementary strand, the *trans*-conformations around the backbone amide bonds become more stable. For **III** to

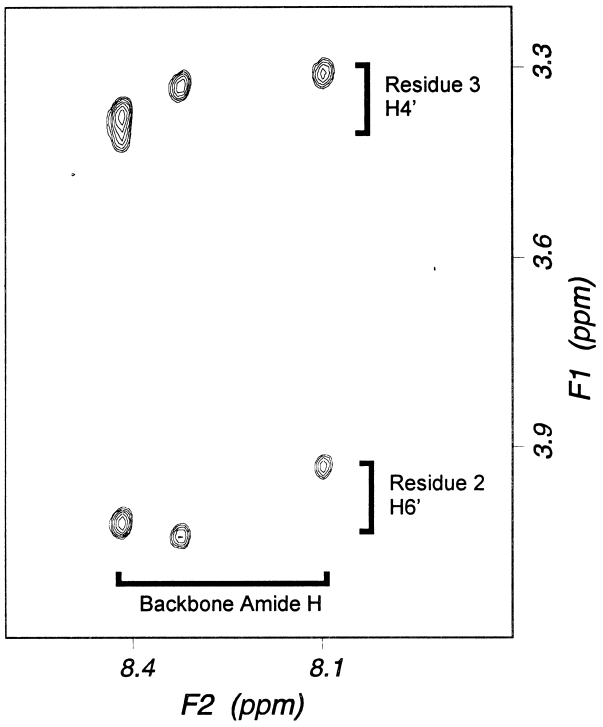


Figure 2. Expanded NOESY (150 ms mixing time) spectrum of the tetramer **III** recorded in 90% H_2O–10% D_2O at 278 K. The cross peaks in this spectral region are due to the contacts between the backbone amide proton (F2, 8.1.4 ppm) and adjacent methylene protons (F1, 3.3.1 ppm).

become a usable building block for preparation of longer oligonucleotide sequences, preliminary studies were undertaken to protect the phosphonate oxygen groups in this molecule or its precursors **I** and **II** using 2-cyanoethanol or methanol. The phosphonate groups were able to slowly react with these alcohol reagents to give phosphonate esters (Yu and Gao, unpublished results). The reaction conditions, however, remain to be optimized.

The hybridization properties of the phosphonate-amide CTTC tetramer **III** with the complementary d(GAAG) sequence were examined in a 90% H_2O–10% D_2O phosphate buffer solution for observation of all proton resonances. 1-D spectra for each of the strands and for a mixture of the two were recorded at 278 and 267 K (Fig. 3), because at these temperatures the interstrand interactions are more stabilized. As a comparison, the spectrum of the mixture of d(CTTC) and d(GAAG) was also recorded (our unpublished data). Inspection of Figure 3 indicates a number of prominent changes upon mixing **III** and d(GAAG) in the base proton resonance region. The resonances at 8.02 ppm of **III** (Fig. 3(A)) were assigned to C NH_2 protons and the spectral patterns in this region reflect the presence of *cis*- and *trans*-isomers due to the backbone amide and the base to backbone amide bonds. This spectral region was greatly simplified along with the disappearance of the broad resonance at 7.3 ppm (Fig. 3(A)) for the sample containing **III** and d(GAAG) as shown in Figure 3(C). The resonance at 6.4 ppm in the free d(GAAG) (Fig. 3(B)), which was assigned to an A H2 base proton, was shifted downfield by 0.1 ppm in the titrated sample (Fig. 3(C)). Comparing Figure 3A and B with 3C, changes in chemical shifts of other base proton resonances at the 7.05 ppm spectral region were also evident. Although no stable Watson–Crick base paired imino proton resonances were observed, the apparent spectral perturbations upon mixing of the two strands are encouraging. These changes involve exchangeable as well non-exchangeable base protons, providing clear indications that **III** interacts, albeit weakly, with d(GAAG). As a comparison, a natural DNA CTTC strand titrated with the complementary d(GAAG) strand under similar conditions used for **III** and d(GAAG) did not show resonances, such as the base paired T NH resonances at 13–14 ppm, which would indicate the formation of a duplex. This may be because these tetramer sequences are simply too short to form a stable duplex structure.

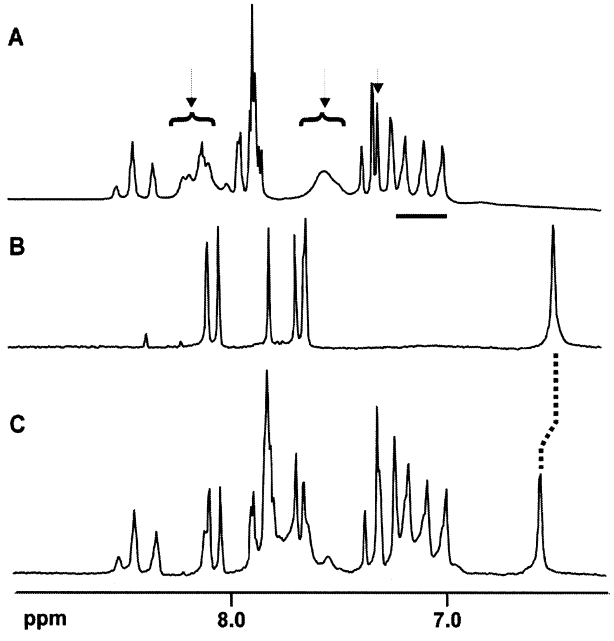


Figure 3. 1-D ^1H spectra of tetramers recorded in 90% H_2O–10% D_2O (50 mM NaCl, 5 mM phosphate, pH 6.3) at 278 K. The spectral region (6.3–9 ppm) contains nucleobase aromatic and amino proton resonances and amide proton resonances. (A) The CTTC tetramer **III**. The resonances which changed positions or disappeared upon mixing **III** with the complementary d(GAAG) strand are indicated by parentheses and arrows. The signals underlined are due to ammonium ion associated with the phosphate group of the tetramer. (B) The free d(GAAG). (C) The sample resulting from the titration of the CTTC **III** and d(GAAG) in a 1:1 ratio. The chemical shift comparison of the 6.5 ppm peak in this and in B is indicated by a dashed line.

Enzymatic treatment of III—in vitro experiments

Tetramer **III**, terminated on either side with natural DNA residues, maintains functional groups that may serve as anchor point to chain extension using enzymes, such as ligase. The resultant sequence would incorporate PNA dimer units at designated position depending on the template sequence used. In these experiments, **III** and natural d(CTTC) were first treated with T4 polynucleotide kinase to give 5′-phosphorylation.[35] The reactions were monitored using HPLC on a C_{18} reverse-phase column (data not shown). The non-phosphorylated sequence required a longer retention time compared to the more polar, phosphorylated counterpart. In comparison to natural tetramer, complete phosphorylation of **III** was observed at a prolonged reaction time (17 h versus 5 h for d(CTTC)) and with an additional amount of T4 kinase. The phosphorylated tetramers were purified using preparative HPLC and treated with T4 ligase at 15 °C (a temperature chosen to compromise low temperature required for hybridization of short oligonucleotides and yet to maintain sufficient enzymatic activity) using a template strand (bold lettering indicates site for CTTC binding), 5′-d(CGCAAGAA**GAAG**AG CAAACGC), and two shorter strands, which are complementary to the template strand or flank the d(pCT TC) sequence. The sequences for the short strands are: 5′-d(pTTCTTGCG) for the 5′-segment of the template strand and 5′-d(GCGTTTGCT) for the 3′-segment of the template strand. Nondenaturing polyacrylamide gel electrophoresis was used to monitor the formation of ligation products using EtBr staining (Fig. 4) or ^{32}P labeling (data not shown). After 24 h enzyme treatment, natural tetramer was shown to be covalently linked with the flanking sequences, and the 21-mer duplex formed (Figure 4, lane 4). This is consistent with the formation of a ligation product in the presence of the template strand. However, for **III**, under the same reaction conditions even after 144 h, the formation of a larger molecular weight compound could not be detected (Figure 4, lanes 2 and 6). The failure of **III** to produce the ligation product might be due to either low binding affinity of **III** to the template strand and/or low efficiency of T4

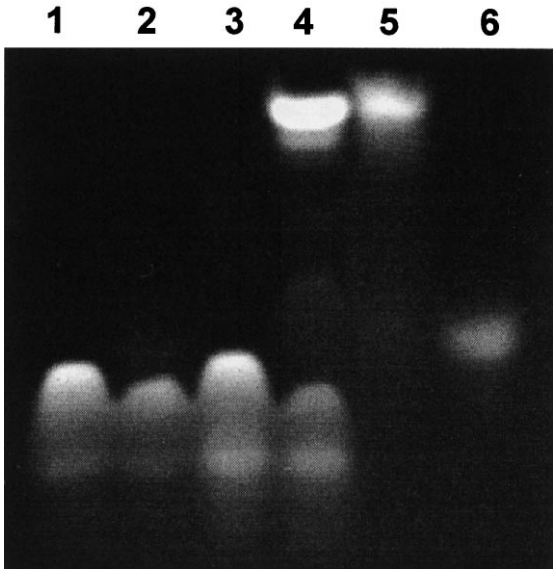


Figure 4. Denaturing PAGE of ligation experiments. Lane 1: DNA-PNA-DNA tetramer **III** and ligation reaction mixture except for T4 ligase; lane 2: DNA-PNA-DNA tetramer **III** ligation reaction; lane 3: d(CTTC) and ligation reaction mixture except for T4 ligase; line 4: d(CTTC) ligation reaction; lane 5: control duplex (synthesized complementary oligonucleotide 21-mer sequences containing CTTC); lane 6: control oligonucleotide 21-mer strand containing CTTC. Lanes 1 and 2 display ligation results of **III**. Lanes 3 and 4 are comparison experiments using natural DNA tetramer d(CTTC). Lane 5 provides a reference to the duplex formation due to successful ligation. Lane 6 provides a reference to single stranded ligation product. If the ligated sequence containing **III** forms, but the corresponding duplex is unstable, the ligated product would be detectable by referencing to the profile shown in lane 6.

ligase on modified sequences. However, these experiments show the DNA-PNA-DNA dC-TT-dC chimera can be phosphorylated using T4 polynucleotide kinase. We speculate that a longer chimera sequence might be able to form a more stable hybrid duplex, and thus could be a suitable substrate for DNA ligase enzymes.

Conclusion

In summary, reported herein is a convergent synthesis of a new type of tetranucleotide, **III**, containing alternating achiral, phosphonate (negatively charged) and *N*-ethyleneglycosyl amide (neutral) backbone linkages in dimer units. **III** was assembled by the condensation of the two dinucleotides, **I** and **II** (Fig. 1), each of which contains a peptide nucleic acid residue linked with a natural nucleotide. The reaction conditions developed are mild and thus, in principle, they should be suitable for the synthesis of dimer analogues containing purine bases. The terminal residues of **III** are natural deoxyribonucleotides, which provide 5′- and 3′-OH groups for chain length extension using standard phosphoramidite chemistry.[36] The tetramer **III**, which contains mixed backbone chemical structures, is readily soluble in aqueous buffer solutions containing 30–100 mM NaCl and is stable under these conditions for an extended period of time. The titration experiments suggest that **III** interacts with the complementary strand. The tetramer **III** is also a substrate for T4 polynucleotide kinase. These properties are promising for constructing longer AON sequences containing alternating phosphate-phosphonate-amide backbone structures using the tetramer **III** or, more likely, its analogues as building blocks. One example of such analogues would contain a PNA dimer, which has an inverted N-CO linkage compared to **III**, in between DNA residues. This would shift the alignment of an analogue sequence with the counter strand DNA to achieve better match of nucleobases of the two strands.

Materials and Methods

Materials

The starting material, diethyl-2-bromoethylphosphonate **I-1**, was purchased from ACROS. CH_3CN, pyridine, and CH_2Cl_2 were dried by distillation from CaH_2; THF was dried over CaH_2; TEA was distilled from KOH; and DMF was freshly distilled under reduced pressure. Chromatographic purifications were carried out using the following conditions: silica gel (Silica gel 60, 230–400 mesh, EM Separations Technology) using organic solvents (see experimental description) as eluant; low pressure reverse-phase C_{18} column (Lichroprep RP-18, 40–63 μm, EM Separations Technology) using H_2O (A) and CH_3OH (B) gradients, both solvents contain 0.1% TFA; and HPLC reverse-phase columns (Waters, C_8 or C_{18}, μBondapak or Nova-Pak) using solvent A: 0.1 M NH_4OAc (pH 6.5) and B: CH_3CN, unless otherwise indicated. Desalting or size exclusion columns used were packed with Bio-gel PG-60 (Bio-Rad). Cation exchange (Na^+) was carried out using Dowex-50 X2 resin (H^+-form, Aldrich). Double filtered H_2O was used as eluant for desalt and cation exchange chromatography. T4 polynucleotide kinase and T4 ligase were purchased from Gibco BRL.

Experimental

NMR spectra were recorded on QE 300 MHz and AMX-II 600 MHz spectrometers (University of Texas Medical School at Houston and University of Houston). NMR resonance assignments follow the numbering systems shown in Scheme 3 and residue numbers 1–4 are based on the tetramer **III**. The reference for ^1H resonances is external tetramethylsilane (0.0 ppm) or HOD in an aqueous solution (4.70 ppm), for ^{13}C is $CHCl_3$ (77.0 ppm) or DMSO-d_6 (39.5 ppm), and for ^{31}P is external trimethylphosphate in a 0.1 M NaCl aq solution (−4.0 ppm at 25 °C).

Benzyl *N*-(2-(diethylphosphonyl)ethyl)-*N*-(thymin-1-yl-acetyl)glycinate I-4. To a suspension of glycine (4.50 g, 60 mmol) and K_2CO_3 (16.56 g, 120 mmol) in H_2O (100 mL) at 80 °C was added diethyl-2-bromoethylphosphonate **I-1** (2.73 mL, 15 mmol). After stirring for 3 h at 80 °C, the reaction mixture was cooled to room temperature, acidified with concd HCl to pH 3.5, and concentrated

to ca. 20 mL. The precipitate was removed and the filtrate was purified on an RP-18 column using solvent A with a gradient of 0–20% B. *N*-(2-Diethylphosphonylethyl)glycine was obtained in quantitative yield. ^{1}H NMR (D_2O, ppm). Residue 2 [ethyl group: 1.30 (t, 6H), 4.07–4.22 (m, 4H); H4′: 3.30–3.42 (m, 2H); H5′: 2.30–2.45 (m, 2H); H6′: 3.78 (s, 1H)].

To thymin-1-yl-acetic acid (2.76 g, 15 mmol) in DMF (20 mL) and CH_3CN (20 mL) were added *N*-hydroxysuccinimide (1.90 g, 16.5 mmol) and DCC (3.40 g, 16.5 mmol). The activated ester, **I-2**, was formed after stirring for 5 h as indicated by TLC (EtOAc:CH_3OH:AcOH, 79:20:1) and directly used for next step reaction. To this reaction mixture were added *N*-(2-diethylphosphonylethyl)glycine (2.39 g, 10 mmol) and TEA (3 mL, 20 mmol) in DMF (7 mL). The reaction continued at room temperature for additional 6 h. The solvents were then removed and H_2O (15 mL) was added to the residue. The solution was filtered and the filtrate was purified on an RP-18 column using solvent A with a gradient of B (10–50%). The fractions containing *N*-(2-(diethylphosphonyl)ethyl)-*N*-(thymin-1-ylacetyl)glycine **I-3** were combined and dried (yield 73%). ^{1}H NMR ($CDCl_3$, ppm). Residue 2 [ethyl group: 1.12–1.28 (m, 6H), 4.02–4.10 (m, 4H); H1′ and H6′: 4.10, 4.20, 4.30, 4.41 (s, 4H); H4′: 3.78–3.92 (m, 2H); H5′: 2.04–2.22 (m, 2H); H6: 7.12,7.18 (s, 1H); 5Me: 1.90 (s, 3H)].

To **I-3** (4.05 g, 10 mmol) and benzyl alcohol (1.05 mL, 10 mmol) in CH_3CN (30 mL) were added DCC (2.06 g, 10 mmol) and catalytic amount of DMAP (0.24 g, 2 mmol). The reaction was completed in 3 h at room temperature The solution was filtered and the filtrate was concentrated. The residue was dissolved in $CHCl_3$ (50 mL), washed with 0.5 N HCl, 5% $NaHCO_3$ and brine, dried over $MgSO_4$ and then evaporated in vacuo (yield 85%). ^{1}H NMR (DMSO-d_6, ppm). Residue 2 [ethyl group: 1.20 (m, 6H), 3.98 (m, 4H); H1 and H6′: 4.16, 4.48, 4.70 (s, d, s, 4H); H4′: 3.43, 3.59 (m 2H); H5′: 1.98, 2.24 (m, 2H); H6 and benzyl aromatic H: 7.24–7.40 (m, 6H); H3: 11.29; 5Me: 1.74 (s, 3H); benzyl methylene H: 5.12, 5.20 (s, 2H)]; ^{13}C NMR (DMSO-d_6, ppm). 169.4, 168.9, 167.4, 167.2, 164.4, 151.0, 142.0, 135.7, 135.5, 128.4, 128.2, 128.1, 128.0, 127.9, 108.2, 108.1, 66.6, 66.0, 61.3, 61.2, 61.1, 48.6, 47.9, 47.5, 47.48, 42.0, 16.2, 11.9; ^{31}P NMR (DMSO-d_6, ppm). 28.50, 28.57. Anal. calcd for $C_{22}H_{30}N_3O_8P$: C, 53.33; H. 6.10; N, 8.48. Found: C, 53.62; H, 6.04; N, 8.65.

Benzyl *N*-(2-((4-*N*-benzoyl-5′-*O*-(*tert*-butyl-dimethylsilyl)-2′-deoxycytidin-3′-*O*-yl)phosphonyl) ethyl)-*N*-(thymin-1-yl-acetyl)glycinate I-6. To **I-4** (dried by co-evaporation with CH_3OH–toluene) (0.4 g, 0.8 mmol) dissolved in dry CH_2Cl_2 (3 mL) was added under N_2 bromotrimethylsilane (total 0.53 mL, 4 mmol) in two portions (3 and 2 equiv each time) at 1.5 h intervals. The solvent was evaporated, and to this mixture 1 M TEAB was added. The solution was then washed with ethyl acetate. The aqueous layer was collected, concentrated and purified on an RP-18 column eluted using solvent A with a gradient of 20–40% B. Benzyl *N*-(2-phosphonylethyl)-*N*-(thymin-1-yl-acetyl)glycinate **I-5** was obtained in 87% yield. ^{1}H NMR ($CDCl_3$, ppm). Residue 2 [H1′ and H6′: 4.10, 4.30, 4.40, 4.65, (s, 4H); H4′: 3.44.70 (m, 2H); H5′: 1.82–1.18 (m, 2H); H6: 7.00, 7.10 (s, 1H); benzyl aromatic H: 7.20–7.30 (m, 5H): 5Me: 1.85 (s, 3H); benzyl methylene H: 5.10, 5.20 (s, 2H)].

I-5 (0.74 g, 1.69 mmol) was dissolved in 1 M aq TEAB (38 mL) and dried under vacuum overnight. The salt formed and 4-*N*-benzoyl-5-*O*-TBDMS-dC (1.13 g, 2.53 mmol) were co-evaporated with dry pyridine twice and dissolved in dry pyridine (15 mL), to which 1-adamantanecarbonyl chloride (AMC-Cl, 1.99 g, 6.74 mmol) was added at room temperature under N_2. The reaction mixture was stirred for 93 h under N_2 and quenched with 1 M TEAB (21 mL). The mixture was extracted with CH_2Cl_2, which was washed with 0.5 M TEAB and brine and dried over Na_2SO_4. After evaporating CH_2Cl_2 in vacuo, the residue was purified on silica gel eluted with 7–15% CH_3OH in CH_2Cl_2 containing 1% TEA to give **I-6** (yield 50%). ^{1}H NMR (DMSO-d_6, DQF-COSY, ppm). Residue 1 [TBDMS H: 0.06 (s, 6H), 0.84 (s, 9H); benzoyl amide H: 11.31 (s, 1H); H6: 8.27 (d, 1H); H5 and benzoyl aromatic H: 7.23–7.98 (m, 6H); H1′: 6.12 (t, 1H); H2′ and H2″: 2.06, 2.49 (m, 2H); H3′: 4.64 (m, 1H); H4′: 4.15 (m, 1H); H5′ and H5″: 3.77–3.84 (m, 2H)]. Residue 2 [ethyl H: 1.14 (t, 9H), 2.98 (m, 6H); benzyl methylene H: 5.11, 5.19 (s, 2H); H1′ and H6′: 4.15, 4.40, 4.47, 4.76 (s, 4H); H4′: 3.32, 3.45, 3.53 (m, 2H); H5′: 1.57, 1.70 (m, 2H); H6 and benzyl aromatic H: 7.23–7.98 (m, 6H); 5Me: 1.72 (s, 3H); H3: 11.31 (s, 1H)].

CT-CO dimer I

Desilylation: A solution of **I-6** (180 mg, 0.186 mmol) in dry THF (2 mL) was stirred with 1 M TBAF (0.56 mL, 0.56 mmol, 3 equiv) at room temperature for 86 h under N_2. Upon completion the solvent was removed from the reaction mixture under reduced pressure. The residue was purified using silica gel chromatography (15–30% CH_3OH in CH_2Cl_2 containing 1% TEA) to give partially protected **I** in 63% yield. ^{1}H NMR (DMSO-d_6, DQF-COSY, ppm). Residue 1 [ethyl: 1.14 (t, 9H), 2.98 (q, 6H); H1′: 6.10 (t, 1H); H2′ and H2″: 2.16, 2.43 (m, 2H); H3′: 4.71 (m, 1H); H4′: 3.97 (m, 1H); H5′ and H5″: 3.63 (m, 2H); 5′OH; 5.66 (brs, 1H); H6: 8.33 (d, 1H); H5 and benzoyl aromatic H: 7.33–7.98 (m, 6H); benzoyl amide H: 11.30 (s, 1H)]. Residue 2 [H4′: 3.52, 3.40 (m, 2H); H5′: 1.57, 1.76 (m, 2H); H1′ and H6′: 4.15, 4.44, 4.76 (s, d, s, 4H); benzyl methylene H: 5.10, 5.18 (s, 2H); H6 and benzyl aromatic H: 7.23–7.98 (m, 6H); 5Me: 1.72 (s, 3H); H3: 11.30 (s, 1H)]. ^{31}P NMR (DMSO-d_6, ppm) 17.80, 18.27.

Debenzoylation: The above product (10 mg, 0.012 mmol) was treated with concd NH_4OH (3 mL) at room temperature for 5.5 h and then evaporated to dryness. The residue was purified by preparative reverse-phase HPLC using 0.1 M TEAA (pH 6.5) with a gradient of 2–5% CH_3CN. The purified fractions were combined and desalted on a size exclusion column to afford 107 OD (260 nm, OD = optical density per mL) of **I**. NMR: **I** was completely characterized by various 1-D and 2-D ^{1}H, ^{13}C, and ^{31}P NMR experiments (DQF-COSY, TOCSY,

^1H–^{13}C HMQC, HMBC, ^1H–^{31}P COSY) and these results are summarized in Table 1. ESI-MS calcd for $C_{20}H_{27}N_6O_{11}P$ 558.5, found 558.7.

Diethyl-(2-*N*-(2-*N*-benzylocarbonylaminoethyl)-2-*N*-(thymin-1-yl-acetyl)aminoethyl)phosphonate II-4. To ethylenediamine (6.7 g, 100 mmol) stirred at rt was added dropwise diethyl-2-bromoethylphosphonate **I-1** (1.82 mL, 10 mmol). The reaction was completed in 2 h, to which was added ethyl acetate (100 mL). The solution was decanted and the solvent was evaporated at 60 °C in vacuo. The residue was re-suspended in ethyl acetate (60 mL). After standing at −4 °C for overnight, the solution was decanted and concentrated at 60 °C to give diethyl-(2 - *N* - (2 - aminoethyl)aminoethyl)phosphonate **II-2** in quantitative yield. ^1H NMR ($CDCl_3$, ppm). Residue 3 [ethyl H: 1.30 (t, 6H), 4.05–4.15 (m, 4H); H4′, H5′ and H6′: 2.60.95 (m, 6H); H7′: 1.90.05 (m, 2H)].

The synthesis of diethyl-(2-*N*-(2-*N*-benzylocarbonylaminoethyl)aminoethyl)phosphonate **II-3** followed procedures similar to what were discussed previously.[26] To a suspension of **II-2** (0.22 g, 1 mmol), $MgSO_4$ (0.24 g, 2 mmol) and K_2CO_3 (0.41 g, 3 mmol) in $CHCl_3$ (1 mL) was added benzaldehyde (0.15 mL, 1.5 mmol). After stirring at rt for 2 h, the mixture was cooled to −20 °C, and benzyl chloroformate (0.17 mL, 1.2 mmol) was added dropwise to the mixture. The reaction was stirred for 0.5 h at −20 °C and for an additional 2 h at rt. The mixture was filtered and the solvent was evaporated. HCl (6 mL, 0.5 N) was added to the residue and the resultant solution was allowed to stand at rt for 3 h. The reaction mixture was washed with diethyl ether and dissolved in cold, saturated $NaHCO_3$. The solution was extracted with ethyl acetate and the organic layer was washed again with H_2O and brine and concentrated. The residue was purified on silica gel eluted with 5–10% CH_3OH in $CHCl_3$ to give **II-3** in 65% yield. ^1H NMR (DMSO-d_6, ppm). Residue 3 [ethyl group: 1.20 (t, 6H), 3.95 (m, 4H); H4′, H5′ and H6′: 2.54, 2.67, 3.03 (m, 6H); H7′: 1.85 (m 2H); benzyl methylene H: 4.99 (s, 2H); benzyl aromatic H: 7.33 (s, 5H); benzyloxy amide H: 7.17 (m, 1H)].

To **II-3** (0.79 g, 2.2 mmol) dissolved in DMF (8 mL) were added thymin-1-yl-acetic acid[27] (0.61 g, 3.3 mmol) and DCC (0.81 g, 3.9 mmol). The mixture was stirred at rt overnight and monitored by TLC (5% MeOH in CH_2Cl_2). The precipitate generated during the reaction was filtered and the filtrate was evaporated in vacuo. The residue was dissolved in CH_2Cl_2, washed with H_2O and brine, and dried over $MgSO_4$. The final product **II-4** was purified using silica gel chromatography eluted with 4% CH_3OH in CH_2Cl_2 (yield 96%). ^1H NMR (DMSO-d_6, ppm). Residue 3 [ethyl group: 1.22 (m, 6H), 3.97 (m, 4H); H1: 4.54 (d, 2H); H4, H5 and H6: 3.10.59 (m, 6H); H7: 2.00, 2.23 (m 2H); benzyl methylene H: 5.01 (d, 2H); H6 and benzyl aromatic H: 7.24–7.33 (m, 6H); H3: 11.27 (brs, 1H); benzyloxy amide H: 7.46 (m,1H); 5Me: 1.72 (d, 3H)]. ^{13}C NMR (DMSO-d_6, ppm). 166.7, 166.6, 164.4, 156.4, 156.2, 151.0, 142.2, 137.0, 128.3, 127.8, 127.7, 108.0, 65.5, 65.4, 61.3, 61.2, 47.8, 47.7, 46.0, 45.4, 41.5, 40.7, 38.8, 38.0, 16.2, 11.9. ^{31}P NMR (DMSO-d_6, ppm). 28.46, 28.72. Anal. calcd for $C_{23}H_{33}N_4O_8P$: C, 52.67; H, 6.34; N, 10.68. Found: C, 52.35; H, 6.33; N, 10.81.

(4-*N*-Benzoyl-3′-*O*-*tert*-butyl-dimethylsilyl)-2′-deoxycytidin-5′-*O*-yl-(2-*N*-(2-*N*-benzylocarbonylaminoethyl)-2-*N*-(thymin-1-yl-acetyl)aminoethyl)phosphonic acid II-6. To **II-4** (2.2 g, 4.2 mmol) in dry CH_2Cl_2 (30 mL) stirred under N_2 was added TMSiBr (2.2 mL, 16.5 mmol) in three portions with 2 h intervals. After stirring for 7 h under N_2 the reaction mixture was evaporated. One M TEAB (30 mL) was added and the mixture was stirred for an additional 5 min. The solvent was evaporated to dryness and H_2O (30 mL) was added. To this solution were added benzyl chloroformate (1.8 mL, 12.5 mmol) and $NaHCO_3$ (2.1 g, 25 mmol). The acylation reaction was kept stirring overnight. The solution was then acidified to pH 2.5 with concd HCl, and washed with ethyl acetate. The aqueous mixture was concentrated and purified on an RP-18 column using 40% B in A. 2-*N*-(2-*N*-Benzylocarbonylaminoethyl)-2-*N*-(thymin-1-yl-acetyl)aminoethylphosphonic acid **II-5** was obtained in 90% yield. ^1H NMR (DMSO-d_6, ppm). Residue 3 [H1′: 4.60, 4.65 (s, 2H); H4′, H5′ and H6′: 3.20.60 (m, 6H); H7′ and 5Me: 1.80.10 (m, 5H); benzyl methylene H: 5.05, 5.10 (s, 2H); H6 and benzoxyl amide and aromatic H: 7.20–7.60 (m, 7H); H3: 11.35 (s, 1H)].

II-5 (0.71 g, 1.52 mmol) was dissolved in 1 M aq TEAB (38 mL) and dried in vacuo overnight. The salt formed and 4-*N*-benzoyl-3′-*O*-TBDMS-dC[29] (1.02 g, 2.3 mmol) were co-evaporated with dry pyridine and dissolved in dry pyridine (12 mL). To the mixture was added AMC-Cl (1.21 g, 6.1 mmol) at rt under N_2. The reaction was stirred for 72 h under N_2 and quenched with 1 M TEAB (19 mL). The mixture was extracted with CH_2Cl_2, washed with 0.5 M TEAB and brine, and dried over Na_2SO_4. The solvent was evaporated and the residue was purified on silica gel eluted with 7–15% CH_3OH in CH_2Cl_2 containing 1% TEA to give **II-6** in 50% yield. ^1H NMR (DQF-COSY, DMSO-d_6, ppm). Residue 3 [H1′: 4.51, 4.65 (s, 2H); H4′ and H5′: 3.08, 3.18 (m, 4H); H6′: 3.29, 3.34 (m, 2H); H7′: 1.65, 1.75 (m, 2H); benzyloxy amide H: 7.43, 7.52 (m, 1H); benzyl methylene H: 4.98, 5.00 (s, 2H); H6 and benzyl aromatic H: 7.31.97 (m, 6H); 5Me: 1.66, 1.69 (s, 3H); H3: 11.24 (s, 2H)]. Residue 4 [TBDMS H: 0.06 (s, 6H), 0.85 (s, 9H); H6: 8.50, 8.54 (d, 1H); H5: 7.32 (d, 1H); H1′: 6.16 (m, 1H); H2′, H2″: 2.16, 2.27 (m, 2H); H3′: 4.44, 4.49 (m, 1H); H4′: 3.97 (m, 1H); H5′, H5″: 3.87, 3.92 (m, 2H); benzoyl amide H: 11.24 (s, 1H); benzoyl aromatic H: 7.31–7.97 (m, 5H)].

The N-TC dimer II

Desilylation: **II-6** (111 mg, 0.111 mmol) in dry THF (1 mL) was treated with 1 M TBAF (0.33 mL, 0.33 mmol, 3 equiv) at rt for 18 h. Purification procedures used were similar to what were described for **I-6** to **I** conversion to give desilylated **II-6** (63 mg, 64%). ^1H NMR (DMSO-d_6, ppm). Residue 3 [H6, benzyloxy amide and benzyl aromatic H: 7.2.9 (m, 7H); H3: 11.23 (brs, 1H); H1′: 4.52, 4.63 (s, 2H); H7′ and 5Me: 1.67 (m, 5H); benzyl methylene H: 4.98 (s, 2H)]. Residue 4 [ethyl H: 1.15 (t,

9H), 3.01 (q, 6H); H6: 8.47 (d, 1H); H5 and benzoyl H: 7.2.9 (m, 6H); H1′: 6.15 (t, 1H); H2′ and H2″: 2.06, 2.29 (m 2H); H3′ 4.26 (m, 1H); H4′, H5′ and H5″: 3.82–3.96 (m, 3H); 3′OH: 5.58 (brs, 1H); benzoyl amide H: 11.23 (brs., 1H)]; H4′, H5′ and H6′ of Residue 3 were obscured by residual H_2O in DMSO-d_6.

Hydrogenolysis: To **II-6** (15 mg, 0.017 mmol) in CH_3OH (3 mL), adjusted to pH 3.2 with HCl (10%), was added Pd/C (10%, 8 mg). The mixture was shaken on a hydrogenation apparatus under H_2 (50 psi) for 4 h. Pd/C was then removed by filtration and washed with CH_3OH. The combined filtrate was evaporated and the residue was purified by reverse-phase HPLC using 0.1 M TEAA buffer solution (pH 6.5) with a gradient of solvent B (2–20% in 25 min). The product fractions were freeze-dried and applied to a size exclusion column to give the T(4-*N*-benzoyl)C dimer (106 $UV_{260\ nm}$ absorption unit). This compound was treated with concd NH_4OH (1.5 mL) at rt for 12 h to give **II** after purification on a reverse-phase HPLC column using buffer A with a gradient of solvent B from 2–5% in 40 min and 5–15% in 20 min in 50% yield. NMR: **II** was characterized by various 1-D and 2-D ^1H, ^{13}C and ^{31}P NMR experiments (DQF-COSY, TOCSY, ^1H–^{13}C HMQC and HMBC, ^1H–^{31}P COSY). These results are summarized in Table 2. ESI-MS cacld for $C_{20}H_{30}N_7O_9P$ 543.2, found 543.7.

Synthesis of tetramer III

To **I-6** (0.24 g, 0.25 mmol) in EtOH (24 mL) were added $NaHCO_3$ (84 mg, 1 mmol) and Pd/C (120 mg, 10%). The mixture was shaken on a hydrogenation apparatus under H_2 (20 psi) at rt for 4.5 h. The catalyst powder was then removed by filtration and washed with ethanol, and the filtrate was evaporated. The residue was purified by prep-C_{18} reverse-phase HPLC chromatography using buffer solution A with a gradient of solvent B (5–15% B in 15 min, 15–55% B in 80 min). The pure products were obtained in 5-*O*-silylated (**I-7a**, 39%) and desilylated (**I-7b,** 35%) forms. ^1H NMR (DMSO-d_6, DQF-COSY, ppm). **I-7a** Residue 1 [TBDMS H: 0.08 (m, 6H), 0.84 (d, 9H); H6: 8.27 (d, 1H); H5: 7.33.27 (m, 1H); benzoyl aromatic H: 7.51.00 (m, 5H); H1: 6.13 (t, 1H); H2′ and H2″: 2.11, 2.54 (m 2H); H3′: 4.16, 4.68 (m, 1H); H4′: 4.16 (m, 1H); H5′ and H5″: 3.79, 3.88 (m, 2H)]. Residue 2 [H1′ and H6″: 3.84, 4.42, 4.92 (s, 4H); H4′: 3.55 (m, 2H); H5′: 1.67 (m, 2H); H6: 7.33.27 (m, 1H), 5Me: 1.73 (s, 3H)]. ^1H NMR (DMSO-d_6, DQF-COSY, ppm). **I-7b**. Residue 1 [H6: 8.33 (d, 1H); H5: 7.32 (m, 1H); benzoyl aromatic H: 7.50–7.99 (m, 5H); H1′: 6.11 (t, 1H); H2′ and H2″: 2.18, 2.46 (m, 2H); H3′: 4.65, 4.70 (m, 1H); H4′: 3.97, 4.01(m, 1H); H5′ and H5″: 3.63 (m, 2H)]. Residue 2 [H1′ and H6′: 3.86, 3.91, 4.43, 4.87 (s, 4H); H4′: 3.49, 3.56 (m, 2H); H5′: 1.66. 1.63 (m, 2H); H6: 7.32 (m, 1H); 5Me: 1.73 (s, 3H)].

To **II-6** (230 mg, 0.23 mmol) dissolved in CH_3OH (23 mL), adjusted to pH 2.7 with HCl (0.5 N), was added Pd/C (10%, 115 mg) and the mixture was shaken under H_2 (20 psi) at rt for 4 h. The reaction mixture was filtered and the catalyst Pd/C was washed with CH_3OH. The combined filtrate was evaporated in vacuo. The crude product was purified by preparative reverse-phase HPLC using buffer A with a gradient of B (5–20% B in 15 min, 20–50% in 75 min) to give **II-7** in 52% yield. ^1H NMR (DMSO-d_6, DQF-COSY, ppm). Residue 3 [H1′: 5.02 (s, 2H); H4′ and H5″: 3.47, 3.05 (m, 4H); H6′: 3.57 (m, 2H); H7′: 1.76 (m, 2H); H6: 7.33 (m, 1H); 5Me: 1.68 (s, 3H)]. Residue 4 [H6: 8.47 (d, 1H); H5: 7.34 (m, 1H); benzoyl aromatic H: 7.49–7.97 (m, 5H); H1: 6.16 (t, 1H); H2 and H2″: 2.14, 2.26 (m, 2H); H3: 4.46 (m, 1H); H4′: 3.99 (m, 1H); H5′ and H5″: 3.89 (m, 2H); TBDMS H: 0.08 (s, 6H), 0.86 (s, 9H)].

III-1. To **II-7** (30 mg, 0.04 mmol) dissolved in DMF (1.1 mL) were added **I-7a** (R = TBDMS, 42 mg, 0.06 mmol), PyBOP (58 mg, 0.11 mmol) and $NaHCO_3$ (46 mg, 0.55 mmol). The mixture was stirred at 0 °C for 5 h, then at 10 °C for 36 h. The reaction solution was evaporated to dryness in vacuo. The residue was dissolved in CH_3OH and the precipitate was filtered. The filtrate was evaporated to dryness and the residue was purified by reverse-phase HPLC using buffer A and a gradient of B (5–15% in 10 min, 15–50% in 70 min, and 50% in 10 min). The pure fractions of **III-1** were combined and freeze dried (yield 59%).

III-1b. To **II-7** (43 mg, 0.05 mmol) in DMF (1.7 mL) were added **I-7b** (R = H, 61 mg, 0.09 mmol), PyBOP (88 mg, 0.17 mmol), and $NaHCO_3$ (71 mg, 0.84 mmol). The reaction was carried out in a manner similar to what was described for preparation of **III-1a**. **III-1** (45 mg) was obtained in a 58% yield after purification by reverse-phase HPLC using buffer A and a gradient of B (5–15% in 10 min and 15–50% in 70 min) and freeze-drying.

Desilylation of **III-1a** and **III-1a**: **III-1a** (32 mg, 0.02 mmol) in THF (0.8 mL) was stirred with TBAF (1.0 M in THF, 0.12 mL) at rt for 60 h. The solvent was evaporated and the residue was purified repetitively by reverse-phase HPLC using buffer A with a gradient of B from 5 to 35% in 70 min to give a pure, desilylated tetramer, **III-1**, in 63% yield. Using reaction conditions similar to what was described for **III-1a**, **III-1a** (20 mg, 0.015 mmol) produced the same desilylated tetramer, **III-1**, which was purified by reverse-phase HPLC using solvent A with a gradient of B from 5 to 40% in 70 min, in 65% yield. ^1H NMR (DMSO-d_6, DQF-COSY, ppm). Residues 1 and 4 [H6: 8.40 (m, 2H); H5: 7.31 (m, 2H); H1′: 6.14 (m, 2H); C1 H2′ and H2″: 2.17, 2.49 (m, 2H); C4 H2′ and H2″: 2.09, 2.28 (m, 2H); C1 H3′: 4.64, 4.78 (m, 1H); C4 H3′: 4.31 (m, 1H); C1 H4′: 4.05 (m, 1H); C4 H4′: 3.97 (m, 1H); C1 H5′ and H5″: 3.64 (m, 2H), C4 H5′ and H5″: 3.82, 3.91 (m, 2H)]. Residues 2 and 3 [H6: 7.38, 7.41 (s, 2H); 5Me: 1.59, 1.65, 1.69, 1.74 (s, 6H); H1′ and T2 H6′: 4.42, 4.57, 4.81, 4.92 (s, m, 6H); T2 H4′ and T3 H6′: 3.38, 3.42, 3.55 (m, 4H); T3 H4′: 3.22 (m, 2H); T2 H5′ and T3 H7′: 1.60, 1.72, 1.80 (m, 4H); T3 H5′: 3.15, 3.29 (m, 2H)]. Exchangeable H [T2 and T3 H3 and benzoyl amide H: 11.2 (brd, 4H); amide H: 8.80, 9.50 (brd, 1H); 3′OH and 5′OH: 5.95 (brd, 2H)]. Protecting group resonances [benzoyl aromatic H: 7.98, 7.59, 7.47 (m, 10H)]. ^{31}P NMR (DMSO-d_6, ppm). 16.68, 16.90, 17.47, 17.73, 18.30.

III-1 (~5 mg) obtained from the above reactions was dissolved in an aqueous NH_4OH solution (28%, 2 mL) and kept in a sealed vial at rt overnight. NH_4OH was removed under reduced pressure and then in vacuo and the residue was applied on to a preparative reverse-phase HPLC column using buffer A with a gradient of B from 2 to 20% in 55 min to afford desired **III** (3.5 mg). NMR: **III** is fully characterized by ^1H, ^{13}C and ^{31}P 1-D and 2-D correlation NMR spectra, such as DQF-COSY, TOCSY, ^1H–^{13}C HMQC, HMBC, ^1H–^{31}P COSY, etc., and these results are reported in Table 3. ESI-MS calcd for $C_{40}H_{55}N_{13}O_{19}P_2$ 1083.3, found 1083.3.

Enzymatic Reactions of III

Phosphorylation of III

In the presence of T4 polynucleotide kinase (10 u), **III** (0.2 μmol) or CTTC (0.2 μmol, used in a reference reaction) in a buffer solution (0.1 mL, 70 mM Tris–HCl, 10 mM $MgCl_2$, 100 mM KCl, 1 mM 2-mercaptoethanol), pH 7.6, was incubated with ATP (3 mM) at 37 °C. The reaction was monitored by reverse-phase HPLC using 50 mM triethylammonium acetate and CH_3CN as eluant. CTTC phosphorylation was complete within 4 h, while 55% **III** was phosphorylated after 7 h of incubation. Additional enzyme (5 u) was added at this point and HPLC profile indicated 99% phosphorylation after total 17 h reaction.

Ligation reaction: In the presence of T4 DNA ligase (5 u), the flanking strands [d(GCGTTTGCT) and d(pTT CTTGCG)], and the template strand [d(CGCAAGAA-**GAAG**AGCAAACGC)], p**III** or pCTTC (50 μM) in a buffer solution (0.01 mL, 50 mM Tris–HCl, 10 mM $MgCl^2$, 1 mM ATP, 1 mM DTT, 5% (w/v) polyethylene glycol-8000), pH 7.6, was incubated at 15 °C. The reaction was monitored using UV melting measurement (0–80 °C, Varian CAREY 3E) and denaturing polyacrylamide gel electrophoresis (20% acrylamide, Bis: acrylamide 1:29, 100 mM Tris–HCl, pH 7.4, 4 °C).

Acknowledgements

The authors thank Drs. Dawei Li for helping 2-D NMR data collection and Dr. Jefferey S. Rice for proofreading of the manuscript. The authors are also grateful to Dr. Rich H. Griffey (Isis Pharmaceutical, Inc.) for mass analysis. The 600 MHz NMR spectrometer at the University of Houston was funded by the W. M. Keck Foundation. This research is supported by Grants from NIH and the Welch Foundation (E-1270).

References

1. Crooke, S. T. *Annu. Rev. Pharmacol. Toxicol.* **1992**, *32*, 329.
2. Agrawal, S.; Iyer, R. P. *Curr. Opin. Biotechnol.* **1995**, *6*, 12.
3. Milligan, J. F.; Matteucci, M. D.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 1923.
4. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 280.
5. Wagner, R. W.; Matteucci, M. D.; Grant, D.; Huang, T.; Froehler, B. C. *Nature Biotech.* **1996**, *14*, 840.
6. (a) Uhlmann, E.; Peyman, A. *Chemical Reviews* **1990**, *90*, 544. (b) Uhlmann, E.; Peyman, A.; Ryte, A.; Schmidt, A.; Buddecke, E. *Methods Enzymol.* **2000**, *313*, 268.
7. Sanghvi, Y. S. In *Comprehensive Natural Products Chemistry*; Barton, D. H. R.; Nakanishi, K., Eds.; DNA and aspects of molecular biology. Ed. Kool, E. T.; Pergamon: New York, 1998; Vol. 7.
8. Nielsen, P. E.; Haaima, G. *Chem. Soc. Rev.* **1997**, 73.
9. Matteucci, M. *Ciba Found. Symp.* **1997**, *209*, 5.
10. Sczakiel, G. *Front. Biosci.* **2000**, *5*, D194.
11. Barawkar, D. A.; Thomas, Y. K.; Bruice, T. W.; Bruice, T. C. *J. Am. Chem. Soc.* **2000**, published on web.
12. Neilson, J.; Birll, W. K.-D.; Caruthers, M. H. *Tetrahedron Lett.* **1988**, *29*, 2911.
13. (a) Vasseur, J.-J.; Debart, F., Sanghvi, Y. S.; Cook, P. D. *J. Am. Chem. Soc.* **1992**, *114*, 4006. (b) Morvan, F.; Sanghvi, Y. S.; Perbost, M.; Vasseur, J.-J.; Bellon, L. *J. Am. Chem. Soc.* **1996**, *118*, 255.
14. De Mesmaeker, A.; Waldner, A.; Lebreton, J.; Hoffman, P.; Fritsch, V.; Wolf, R. M.; Freier, S. M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 226.
15. Matteucci, M. *Tetrahedron Lett.* **1990**, *31*, 2385.
16. Jones, R. J.; Lin, K.-Y.; Milligan, J. F.; Wadwani, S.; Matteucci, M. D. *J. Org. Chem.* **1993**, *58*, 2983.
17. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Science **1991**, *254*, 1497. (b) Ericksson, M.; Nielsen, P. E. *Quart. Rev. Biophys.* **1996**, *29*, 369. (c) Nielsen, P. E. Applications of peptide nucleic acids. *Curr. Opin. Biotech.* **1999**, *10*, 71.
18. Bergmann, F.; Bannwarth, W.; Tam, S. *Tetrahedron Lett.* **1995**, *36*, 6823.
19. Stetsenko, D. A.; Lubyako, E. N.; Potapov, V. K.; Azhikina, T. L.; Sverdlov, E. D. *Tetrahedron Lett.* **1996**, *37*, 3571.
20. van der Laan, A. C.; Meeuwenoord, N. J.; Kuyl-Yeheskiely, E.; Oosting, R. S.; Brands, R.; van Boom, J. H. *Recl. Trav. Chim. Pays-Bas* **1995**, *114*, 295.
21. Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Ryte, A. *Angew. Chem. Int. Engl.* **1996**, *35*, 2632.
22. Peyman, A.; Uhlmann, E.; Wagner, K.; Augustin, A.; Breipohl, G.; Will, D. W.; Shafer, A.; Wallmeier, H. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2636.
23. Ferrer, E.; Shevchenki, A.; Eritja, R. *Bioorg. Med. Chem.* **2000**, *8*, 291.
24. Uhlmann, E. *Biol. Chem.* **2000**, *379*, 1045.
25. (a) Gao, X.; Brown, F. K.; Jeffs, P.; Bischofberger, N.; Lin, K.-Y.; Pipe, A. J.; Noble, S. A. *Biochemistry* **1992**, *31*, 6228. (b) Gao, X.; Jeffs, P. W. *J. Biomol. NMR* **1994**, *4*, 17. (c) Gao, X.; Jeffs, P. W. *J. Biomol. NMR* **1994**, *4*, 367. (d) Veal, J. M.; Gao, X; Brown, F. K. *J. Am. Chem. Soc.* **1993**, *115*, 7139. (e) Rice, J. S.; Gao, X. *Biochemistry* **1997**, *36*, 399. (f) Cross, C.; Rice, J. S.; Gao, X. *Biochemistry* **1997**, *36*, 4096. (g) Gao, X.; Rice, J. S.; Cross, C. W.; Yang, X. *Nucleosides Nucleotides* **1997**, *16*, 1599. (h) Yang, X.; Sanghvi, Y. S.; Gao, X. *J. Biomol. NMR* **1997**, *10*, 383. (i) Yang, X.; Hang, X.; Cross, C.; Bare, S.; Sanghvi, Y.; Gao, X. *Biochemistry* **1999**, *38*, 12586.
26. Wang, W.; Liang, T. C.; Zheng, X.; Gao, X. *Tetrahedron Lett.* **1995**, *36*, 1181.
27. Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett.* **1994**, *35*, 5173.
28. McKenna, C. E.; Higa, M. T.; Cheung, N. H.; McKenna, M. *Tetrahedron Lett.* **1977**, *18*, 155.
29. (a) Ti, G. S.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1982**, *104*, 1316. (b) McBride, L. J.; Caruthers, M. H. *Tetrahedron Lett.* **1983**, *24*, 245.
30. Malchowski, W. P.; Coward, J. K. *J. Org. Chem.* **1994**, *59*, 7625.

31. Imamura, M.; Hashimoto, H. *Tetrahedron Lett*. **1996**, *37*, 1451.
32. Farese, A.; Patino, N.; Condom, R.; Dalleu, S.; Guedj, R. *Tetrahedron Lett*. **1996**, *37*, 1413.
33. Leijon, M.; Gralund, A.; Nielen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry* **1994**, *33*, 9820.
34. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777.
35. Kozlov, I. A.; Nielsen, P. E.; Orgel, L. E. *Bioconjug. Chem*. **1998**, *9*, 415.
36. McBride, L. J.; Caruthers, M. H. *Tetrahedron Lett*. **1983**, *24*, 245.