Oligonucleotides with Cationic Backbone and Their Hybridization with DNA: Interplay of Base Pairing and Electrostatic Attraction

Boris Schmidtgal, Arne Kuepper, Melissa Meng, Tom N. Grossmann, and Christian Ducho


DOI: 10.1002/chem.201704338
Abstract: Non-natural oligonucleotides represent important (bio)chemical tools and potential therapeutic agents. Backbone modifications altering hybridization properties and bio-stability can provide useful analogues. Here, we employ an artificial nucleosyl amino acid (NAA) motif for the synthesis of oligonucleotides containing a backbone decorated with primary amines. An oligo-T sequence of this cationic DNA analogue shows significantly increased affinity for complementary DNA. Notably, hybridization with DNA is still governed by Watson–Crick base pairing. However, single base pair mismatches are tolerated and some degree of sequence-independent interactions between the cationic NAA backbone and fully mismatched DNA are observed. These findings demonstrate that a high density of positive charges directly connected to the oligonucleotide backbone can affect Watson–Crick base pairing. This provides a paradigm for the design of therapeutic oligonucleotides with altered backbone charge patterns.

Introduction

Oligonucleotides have unique binding properties, rendering them essential molecules in living organisms and providing a platform for the modulation of biological functions through antigene, antisense, or RNA interference approaches. In natural nucleic acids (DNA and RNA), nucleotides are linked by phosphate diester moieties, thus leading to a polyanionic backbone at physiological pH. The polyanionic character contributes to the low cellular uptake of such oligonucleotides, thereby compromising in vivo applications. In addition, DNA and RNA possess low stability toward nucleases, which are ubiquitous in biological systems. These limitations have led to the development of artificial, biostable nucleic acids (e.g., phosphorothioates and “locked” nucleic acids (LNA)). In many cases, these analogues exhibit altered binding properties, providing insights into the structural contributions to duplex formation. For instance, the charge pattern of the oligonucleotide backbone can influence both hybridization and pharmacokinetic properties. Therefore, artificial electroneutral internucleotide linkages, for example, amide,[6] sulfone,[5] or triazole[6] motifs, have been developed. Triazole linkers have even been shown to be biocompatible, that is, triazole-modified DNA can be recognized by polymerases in cells and can therefore be employed to construct genes.[6b–e] Peptide nucleic acid (PNA) is an example of an electroeutral nucleic acid mimic with an amide-based backbone exhibiting only limited resemblance to DNA and RNA.[7] However, fully electroeutral nucleic acid analogues often suffer from low water solubility and a tendency to aggregate in aqueous solution. These limitations and the quest for oligonucleotides with fundamentally different properties have led to the development of positively charged nucleic acids. In most cases, positive charges were introduced through a modification of the 2’-hydroxy groups (in RNA) or nucleobases leaving the phosphate diester backbone unchanged. These strategies furnished zwitterionic structures,[8] but resulted in densely charged oligonucleotides.

An alternative approach involves the replacement of phosphate diester units by non-natural positively charged linkers, thus providing an oligomer with a retained overall number of charges, but reversed polarity. This may be advantageous for biomedical applications, in particular with respect to cellular uptake, as indicated by the favorable properties of cationic cell-penetrating peptides (CPPs).[9] Only a few positively charged internucleotide linkages have been reported so far: 1) Bruce’s guanidine[10] and S-methylthiourea[13] linkages; 2) Letsinger’s phosphoramidate linkages, in which amines were connected to the backbone through alkyl linkers,[13] and 3) our recently reported NAA-modified oligonucleotides.[14]

Bruce’s rather rigid guanidine linkage and Letsinger’s flexible aminooalkyl moieties provide cationic oligonucleotides with high affinity for DNA, but conformational properties that deviate significantly from native nucleic acids. This raises the question of how moderately flexible internucleoside linkages would impact the hybridization properties of cationic oligonucleotide analogues. In principle, the NAA modification (with its rigid amide bond and the adjacent 5’-C-6’-C single bond, Figure 1) could be used to assemble corresponding oligonucleotides. However, previously reported “dimeric” phosphoramidite building blocks only allow incorporation of the NAA modification adjacent to phosphate diester units, thus resulting in partially zwitterionic DNA analogues.[14]

The favorable properties of these partially zwitterionic NAA-modified DNAs and the general interest in fully cationic oligonucleotides have inspired us to design oligomers of type 1,
which are completely assembled from NAA internucleoside linkages (Figure 1). The availability of such cationic oligonucleotide analogues provides the basis for answering some of the questions highlighted above. In particular, it would allow an assessment of the influence of the fully cationic backbone on duplex stability and sequence specificity of hybridization with native DNA.

Results and Discussion

For the synthesis of nucleoside-derived δ-peptide-like oligomers 1, we envisioned connecting monomeric units of type 2 through amide formation in analogy to solid-phase peptide synthesis (SPPS, Figure 1). Building block 2 should be obtained either from the 3'-azido-substituted nucleoside-5'-aldehyde 3, or via its protected 3'-amino analogue 4. We decided to focus on cationic oligomers with thymine as nucleobase, aiming to prepare both the all-(S)- and the all-(R)-configured oligomers (with respect to the stereochemical configuration at the 6'-position). For subsequent hybridization experiments and biophysical characterizations, we designed 14-mer oligomers 1a and 1b (Figure 1).

The synthesis of building blocks (S)-2 and (R)-2 started from 3'-N-benzyloxymethyl-(BOM)-protected 3'-azido-3'-deoxy-thymidine 5 and its 3'-N-Cbz-protected 3'-amino congener 6, respectively, to compare the routes via azide 3 and protected amine 4 (Scheme 1; see Supporting Information for synthesis of 5 and 6). Aldehydes 3 and 4 were obtained by IBX oxidation of 5 and 6 in quantitative yields. On the basis of our previously reported syntheses of nucleosyl amino acids,[15] we applied a sequence of Wittig–Horner olefination and asymmetric hydrogenation to introduce the amino acid motif. Wittig–Horner transformations of aldehydes 3 and 4 with phosphonate 7 (see Supporting Information) furnished didehydro amino acids 8 and 9 in yields of 73% and 68%, respectively, with high stereoselectivities toward the desired Z-isomers (93:7 and 91:9, respectively). However, the concomitantly formed E-isomers could not be fully removed, and thus, asymmetric hydrogenations were performed with the Z/E-mixtures. Hydrogenation of 8 and 9 in the presence of chiral Rh catalysts (S,S)- and (R,R)-Me-DuPHOS-Rh[16] afforded the nucleosyl amino acid products 10 and 11 in yields of 54–92%, with the major isomer (6'S or 6'R) depending on the employed catalyst ([S,S]-catalyst for (6'S), [R,R]-catalyst for (6'R); for stereochimical assignments see Experimental Section). In contrast to our previous syntheses of nucleosyl amino acids,[15] the hydrogenation products were not obtained in diastereomerically pure form, but with diastereomeric ratios ranging from 85:15 to 95:5. HPLC purification of (S)-11 and (R)-11 gave the pure 6'-epimers (d.r. > 99:1), to be followed by the efficient concomitant hydrogenolytic removal of Cbz, Bn, and BOM groups. As this hydrogenation step proceeded less satisfactorily for the 3'-azido congeners, the synthesis of the target structures through the 3'-N-Cbz-amino route (using aldehyde 4) was superior overall. Subsequent 3'-N-Fmoc protection furnished diastereomerically pure building blocks (S)-2 and (R)-2 in yields of 58% and 48%, respectively, over the last two steps (Scheme 1). Fmoc-based SPPS employing either (S)-4 or (R)-4, followed by final acidic cleavage and

![Figure 1. NAA-derived fully cationic oligonucleotides 1a and 1b including their retrosynthesis. SPPS: solid-phase peptide synthesis; BOM = benzyloxy-methyl.](image1)

![Scheme 1. Synthesis of building blocks (S)-2 and (R)-2 for the preparation of cationic target oligomers 1a and 1b.](image2)
deprotection (reactions not shown), gave the two 14-mer dia-
t stereomers 1a and 1b, respectively.

For investigation of the hybridization properties with DNA, fully cationic 6′-all-(S)-14-mer 1a was incubated with comple-
mentary 14-mer DNA (A14). The influence of ion strength on
duplex formation was investigated by varying the concentra-
tion of NaCl (50–125 mM, all in phosphate buffer, pH 7.4). Under all conditions, duplex formation was observed, as indi-
cated by hyperchromicity upon heating, leading to the typical
sigmoidal melting curves (Fig-
ure S2, Supporting Information) allowing the determination of melting temperatures (Tm) (Table 1, entries 1 to 4, and Fig-
ure S2). As expected,[13] we observed a decrease in Tm with increasing
NaCl concentration, which can be ascribed to salt-
mediated shielding of the backbone in 1a (positive charges)
and in DNA (negative charges). We then decided to use
100 mM NaCl, as this represents a commonly applied concen-
tration. As a reference, the native DNA–DNA duplex (Tm
DNA) was used. At 100 mM NaCl, the duplexes of both cationic ana-
logues (1a and 1b) with fully complementary DNA (A14) were
more stable than the corresponding DNA–DNA duplex, as indi-
cated by the differences in melting temperatures (ΔTm) (9°C
and 17°C, respectively; Table 1, entry 3; melting curves shown
in Figure 2, solid lines). The cationic 14-mers 1a and 1b con-
tained 13 non-native internucleoside linkages (“modifications”, 
mod.), so these results were equivalent to ΔTm/mod. values of
+0.7°C (1a) and +1.3°C (1b), respectively.

Subsequently, we studied the base-specificity of duplex for-
mation of 1a and 1b with partially mismatched DNA strands.
Upon introduction of a single base mismatch (C, G, or T in-
stead of A in the middle of the DNA sequence), sigmoidal UV
melting curves with 1a and 1b were observed, indicating
duplex formation (Figures S3 and S4, Supporting Information).
Remarkably, both cationic oligomers 1a and 1b showed nearly
retained duplex stability upon incorporation of this single mis-
match (Table 1, entries 5–7 vs. entry 3). In contrast, the corre-
sponding DNA–DNA duplexes encountered the expected
destabilization of approximately 13°C. This led to ΔTm values
(difference from the corresponding single mismatched native
DNA duplex) of up to around 30°C. In the case of T-C and T-G
mismatches, isomer 1a furnished slightly more stable duplexes
with the single-mismatched DNA than with the fully comple-
mentary A14 strand. For the same base mismatches, the 1b–
DNA duplexes were slightly destabilized relative to the fully
complementary analogues (Table 1, entries 5 and 6 vs. entry 3).
Overall, the (6′R)-configured NAA linkage (1b) furnished more
stable duplexes than the (6′S)-configured congener in the case
of the fully matched sequence and of the T-T mismatch
(Table 1, entries 3 and 7, 1b vs. 1a). Remarkably, 1b shows
some preference for fully complementary DNA, resulting in
Tm values that are 0.3–4.1°C lower with the single mismatched
counterstrands.

These results indicate that oligonucleotide analogues 1a
and 1b were relatively insensitive to single base mismatches in
the DNA counterstrand. Hence, we aimed to probe whether
Watson–Crick base pairing contributes to duplex formation, or
if hybridization merely results from electrostatic attraction of
the two backbones (oligocation 1ab with oligoanionic DNA).
Therefore, melting curves were recorded for equimolar mix-
tures of thymidine-derived oligomers 1a or 1b with a fully
mismatched 14-mer DNA (G6TTG6, Figure 2, Table 1, entry 8).
The resultant curves indicate no specific melting process, that
is, no defined transition between an aggregated and a non-ag-
ggregated state was observed (Figure 2B, C, dashed lines). How-
ever, 1a in particular and also 1b to some extent showed con-
siderable hyperchromicity with G6TTG6 upon heating (up to

<p>| Table 1. Tm values (°C ± SD) of cationic oligonucleotide analogues 1a and 1b (as well as the Tm DNA reference) with native DNA strands. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>DNA</th>
<th>NaCl [mM]</th>
<th>Tm ref. 1a (6′S)</th>
<th>Tm ref. 1b (6′R)</th>
<th>ΔTm 1a (6′S)</th>
<th>Tm ref. 1b (6′R)</th>
<th>ΔTm 1b (6′R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A14</td>
<td>50</td>
<td>53.6 ± 0.9</td>
<td>53.6 ± 0.9</td>
<td>-</td>
<td>n.d.</td>
<td>-</td>
</tr>
<tr>
<td>A14</td>
<td>75</td>
<td>51.1 ± 0.6</td>
<td>51.1 ± 0.6</td>
<td>-</td>
<td>n.d.</td>
<td>-</td>
</tr>
<tr>
<td>A14</td>
<td>100</td>
<td>45.1 ± 0.9</td>
<td>45.1 ± 0.9</td>
<td>+9</td>
<td>53.8 ± 0.4</td>
<td>+17</td>
</tr>
<tr>
<td>A14</td>
<td>125</td>
<td>43.1 ± 0.4</td>
<td>43.1 ± 0.4</td>
<td>-</td>
<td>n.d.</td>
<td>-</td>
</tr>
<tr>
<td>A14′CHA6</td>
<td>100</td>
<td>49.8 ± 0.3</td>
<td>49.8 ± 0.3</td>
<td>-26</td>
<td>49.7 ± 0.5</td>
<td>+26</td>
</tr>
<tr>
<td>A14′CHA6</td>
<td>100</td>
<td>50.0 ± 0.3</td>
<td>50.0 ± 0.3</td>
<td>-26</td>
<td>50.3 ± 1.4</td>
<td>+27</td>
</tr>
<tr>
<td>A14′TA6</td>
<td>100</td>
<td>45.9 ± 0.0</td>
<td>45.9 ± 0.0</td>
<td>-23</td>
<td>53.5 ± 0.3</td>
<td>+30</td>
</tr>
<tr>
<td>A14′TA6</td>
<td>100</td>
<td>45.9 ± 0.0</td>
<td>45.9 ± 0.0</td>
<td>-23</td>
<td>53.5 ± 0.3</td>
<td>+30</td>
</tr>
<tr>
<td>A14′GA6</td>
<td>100</td>
<td>45.9 ± 0.0</td>
<td>45.9 ± 0.0</td>
<td>-23</td>
<td>53.5 ± 0.3</td>
<td>+30</td>
</tr>
<tr>
<td>A14′GA6</td>
<td>100</td>
<td>45.9 ± 0.0</td>
<td>45.9 ± 0.0</td>
<td>-23</td>
<td>53.5 ± 0.3</td>
<td>+30</td>
</tr>
</tbody>
</table>

[a] In aqueous 10 mM NaH2PO4 (pH 7.4) and NaCl. [b] Base mismatches underlined and in bold. [c] Tm(A1-1a-DNA)-Tm(A1-1b-DNA); Tm(A1-DNA)-Tm(A14-DNA). [d] Tm(1b-DNA)-
Tm(1b-DNA). [e] n.d. = not determined. [f] No sigmoidal melting curve was observed.

Figure 2. Melting curves (average of triplicates) for A) native DNA oligonucleotide Tm, B) cationic oligonucleotide analogue 1a (6′S), and C) cationic oligonu-
cleotide analogue 1b (6′R) with native complementary DNA (A14, solid lines), as well as with native fully mismatched DNA (G6TTG6, dashed lines).
with fully mismatched DNA (6-mer duplex, Figure 2A, solid line). In contrast, the hyperchromicity of the mixture 1a + G6TTG6 starts to increase at around 35 °C, but this is less pronounced for 1b + G6TTG6 setting in at approximately 70 °C. A possible explanation for this behavior is a charge-mediated unspecific formation of aggregates at lower temperatures for both 1a and 1b with fully mismatched DNA. Elevated temperatures can then be expected to induce the disassembly of these structures. An additional indication of the presence of charge-mediated unspecific interactions are the less defined transitions in the melting curves of 1a and 1b with the fully matched A6, DNA counterstrand (Figure 2B, C, solid lines). These duplexes appear to undergo complex temperature-induced transitions, with several changes in UV absorbance in addition to the main transition (that is, melting of the duplex). In comparison, the native DNA–DNA (T6-A6) reference duplex (Figure 2A, solid line) shows a sharp transition.

To obtain insights into the structural properties of cationic oligomers 1a and 1b and the resultant duplexes, we performed circular dichroism (CD) spectroscopy. Initially, the single-stranded oligonucleotides were investigated (Figure 3A), including the corresponding native DNA (T6, black) and complementary native DNA (A6, grey). A comparison of their CD spectra reveals differences between all oligomeric thymidine analogues (T6, 1a and 1b). Notably, the spectra of native T6 (black) and cationic 1a (blue) are more similar in comparison to 1b (orange), which differs particularly at lower wavelengths. This indicates substantial structural differences between both cationic oligomers in their single-stranded form (note that 1a and 1b are diastereomers).

Subsequently, duplexes of both cationic oligomers (1a and 1b) and of native DNA (T6) with complementary DNA (A6) were investigated (Figure 3B). The CD spectrum of the 1a-A6 duplex (blue) shows resemblance to the CD signals of the native T6-A6 DNA–DNA duplex[17] (black, Figure 3B). On the other hand, differences are observed for the 1b-A6 duplex (orange, Figure 3B): although the pattern of signals for λ < 260 nm is rather similar to that of the native duplex, it significantly differs for λ > 260 nm. Instead of a single maximum at approximately 275 nm, 1b-A6 exhibits two maxima at around 260 nm and 300 nm, respectively. Overall, the similarities in the CD spectra suggest that binding of both isomers 1a and 1b to complementary DNA probably furnished duplexes with mainly DNA-like helical topologies. However, 1b (which, remarkably, formed the most stable duplex with A6) displayed the most pronounced deviations both in its single-stranded form and in its complex with complementary native DNA A6.

We then studied the CD spectra of equimolar mixtures of oligomers 1a or 1b with fully mismatched 14-mer DNA (G6TTG6, Figure 4). For the (6'S)-configured oligomer 1a, the CD spectrum of the mixture with mismatched DNA (Figure 4A, solid line) does not show any signal pattern indicative of a helical duplex structure (as compared with that of the corresponding matched duplex, Figure 3B). To assess the possibility of nonspecific interactions between both single strands in this mixture, we determined the CD spectrum of G6TTG6 alone (Figure S5, Supporting Information) and added it to the spectrum of single-stranded 1a. This combined spectrum (Figure 4A, dashed line) differs significantly from the experimentally obtained spectrum of the mixture. This suggests Watson–Crick-independent nonspecific interactions between 1a and fully mismatched G6TTG6 at ambient temperature, and is in line with the aforementioned melting behavior (Figure 2). Similar results were obtained for the mixture of (6'R)-configured oligomer 1b with fully mismatched DNA (Figure 4B).

Conclusion

We report the synthesis of a novel amide-linked cationic oligonucleotide analogue based on NAA internucleoside linkages. Two cationic oligomers (1a and 1b, 6'-epimers) were synthe-
sized and investigated by UV melting studies and CD spectroscopy. Both oligomers form very stable, presumably helical, duplexes with native complementary DNA strands. The high duplex stability (compared with the native DNA–DNA duplex) resembles that of some previously reported cationic oligomers, in particular Bruce’s guanidine-linked oligonucleotides.[11] Notably, hybridization of 1a and 1b with DNA was insensitive to single base mismatches, thus indicating robustness of hybridization toward limited local perturbations within the duplex. However, duplex formation was not detected in the case of a fully mismatched DNA counterstrand, demonstrating Watson–Crick base pairing to be a requirement for hybridization and the occurrence of a defined topology. The most likely explanation for our observations is that electrostatic attraction can compensate for single base mismatches, but is not sufficient to foster the formation of a defined structure in the absence of Watson–Crick base pairing. In the latter case, sequence-unspecific charge-mediated aggregation phenomena occur. This behavior is in sharp contrast to the pronounced mismatch sensitivity of partially zwitterionic oligonucleotides containing up to four NAA linkages.[14a]

Interestingly, both isomers differ moderately in their DNA-binding affinity as well as selectivity, and on the basis of their CD spectra, also in their structural topology. Apparently, the (6'-R)-configuration in 1b is slightly more beneficial for hybridization. Notably, an analogous behavior had also been observed for previously reported partially zwitterionic NAA-modified oligonucleotides.[14a] Both isomers 1a and 1b appear to adopt different structures in their single-stranded form. Furthermore, the CD spectrum of duplex 1b-1a reveals a maximum at approximately 300 nm, which is unusual for fully helical oligonucleotide duplexes.

Overall, our findings will contribute to the future design of oligonucleotides for potential biomedical applications. The favorable properties of cationic cell-penetrating peptides (CPPs)[30] indicate that the introduction of positive charges into the oligonucleotide backbone might be beneficial for their therapeutic or diagnostic use, in particular with respect to cellular uptake. However, as demonstrated in this work, fully cationic (in contrast to partially zwitterionic) oligonucleotides can suffer from impaired base pairing fidelity and unspecific aggregation in the absence of Watson–Crick base pairing. In our future work, we will therefore study how the ratio of negatively and positively charged linkages impacts base-pairing fidelity. In addition, more detailed structural studies will be performed, with the long-term goal of obtaining a thorough understanding of the interplay of conformation, base pairing, and electrostatic attraction in duplexes of (partially) cationic and anionic oligonucleotide strands.

Experimental Section

General methods

The syntheses of starting materials 5 and 6 and of phosphonate 7 are described in the Supporting Information. All other chemicals were purchased from standard suppliers. Reactions involving oxygen- and/or moisture-sensitive reagents were performed under an atmosphere of argon using anhydrous solvents. Anhydrous solvents were obtained in the following manner: THF was dried over sodium/benzophenone and distilled, CHCl3 was dried over CaH2 and distilled, MeOH was dried over activated molecular sieves (3 Å) and degassed, MeCN was dried over P2O5 and distilled, pyridine was dried over CaH2 and distilled, toluene was dried over sodium/benzophenone and distilled. The thus-obtained solvents were stored over molecular sieves (4 Å; in case of MeOH and MeCN, 3 Å). All other solvents were of technical quality and distilled prior to use, and deionized water was used throughout. Column chromatography was performed on silica gel 60 (0.040–0.063 mm, 230–400 mesh ASTM, VWR) under flash conditions unless otherwise indicated. TLC was performed on aluminum plates precoated with silica gel 60 F254 (VWR). Visualization of the spots was achieved using UV light (254 nm) and/or staining under heating (H2SO4 staining solution: 4 g vanillin, 25 mL conc. H2SO4, 80 mL AcOH, and 680 mL MeOH; KMnO4 staining solution: 1 g KMnO4, 6 g K2CO3, and 1.5 mL NaOH (1.25 M) solution, all dissolved in 100 mL H2O; ninhydrin staining solution: 0.3 g ninhydrin, 3 mL AcOH, and 100 mL 1-butanol). Analytical chiral HPLC was performed on a Jasco system equipped with a pu 2080 Plus pump, an AS 2055 Plus autosampler, an MD 2010 Plus multivavelength detector, and an IB ChiralpakTM column (0.8×27.5 cm) purchased from Diacel. Method: isocratic eluent: 2-propanol–MeOH–EtOAc; flow 0.8 mL min−1; injection volume 10 μL (c=4 mg·mL−1) in EtOAc. Preparative chiral HPLC was performed on a Jasco system equipped with a pu 2080 Plus pump, an MD 2010 Plus multivavelength detector, and an IB ChiralpakTM column (1.5×28 cm) purchased from Diacel. Method: isocratic eluent: 73.27 n-hexane–EtOAc; flow 5 mL·min−1; injection volume 100 μL (c≈100 mg·mL−1) in EtOAc. 300 MHz- and 500 MHz-1H, 75 MHz- and 126 MHz-13C, and 121 MHz-31P NMR spectra were recorded on Varian MERCURY 300, UNITY 300, INOVA 500, and INOVA 600 spectrometers. All 13C NMR spectra were H-decoupled. All spectra were recorded at room temperature unless indicated otherwise, and were referenced internally to solvent reference frequencies. For calibration of 31P NMR signals, 85% phosphoric acid was used as an external standard. Chemical shifts (δ) are quoted in ppm and coupling constants (J) are reported in Hz. Signals were assigned by using H,H-COSY, H,SQC, and HMBC spectra obtained on the spectrometers detailed above. Mass spectra of small molecules was performed on a PerkinElmer Lambda 2 spectrometer (in the range 190–500 nm. Wavelengths of maximum absorption (λmax) are reported in nm with the corresponding logarithmic molar extinction coefficient (log ε) given in parentheses (ε in dm3·mol−1·cm−1).
6'-N-Boc-3'-N-Fmoc-amino-3'-deoxy-(S)-thymidinyl amino acid (R)-2: The synthesis of (R)-configured NAA (R)-2 was performed according to the procedure for the synthesis of (S)-configured NAA (S)-2 with diastereomerically pure NAA (R)-11 (vide infra, 370 mg, 0.49 mmol), Pd-black (528 mg, 4.98 mmol), n-butylamine (731 mg, 1.00 mL, 10.0 mmol), MeOH (12 mL), NaOAc (153 mg, 0.50 mmol), Fmoc-OSu (169 mg, 0.50 mmol), and acetone/water (3:2) at room temperature. The resultant solution was stirred for 5 min under reduced pressure. The resultant crude product was purified by column chromatography (9:1 CHCl3:MeOH, 0.5% AcOH). The obtained product was coevaporated with toluene (3 × 5 mL) to give (S)-2 as a fine white powder (393 mg, 58% over two steps from (S)-11). M.p. decomposition >110°C; TLC: Rf = 0.26 (9:1 CHCl3:MeOH); δ13Cppm = +23.5 (c 1.1, CHCl3); 1H NMR (500 MHz, CD3OD): δ = 1.44 (s, 9H, C(CH3)3), 1.93 (s, 3H, 6-H), 2.06–2.16 (m, 1H, 5'-H), 2.21–2.39 (m, 3H, 2'-H, 4'-H, 5'-H), 3.89–3.97 (m, 1H, 4'-H), 4.05–4.14 (m, 1H, 3'-H), 4.18–4.27 (m, 2H, Fmoc-CH2), 4.32–4.48 (m, 3H, 9-H, 6'-H, 6'-NH), 6.08–6.15 (m, 1H, 1'-H), 7.31 (dd, J = 7.3, 7.3 Hz, 2'-H, 7'-H), 7.34 (dd, J = 7.3, 7.3 Hz, 2'-H, 3'-H, 4'-H), 7.53 (brs, 1H, 6-H), 7.65 (dd, J = 7.3 Hz, 2'-H), 8.42 (m, 5'-H, 7'-H), 7.79 ppm (d, J = 7.3 Hz, 2'-H, 1'-H, 8'-H); 13C NMR (126 MHz, CD3OD): δ = 11.1 (C-7), 27.3 (C(CH3)3), 35.3 (C-5'), 36.8 (C-2'), 46.8 (Fmoc-CH2), 51.3 (C-6), 54.2 (C-3), 64.6 (C-9'), 79.2 (C(CH3)3), 81.0 (C-8'), 84.4 (C-1), 110.5 (C-5), 119.5 (C-1'), 124.7, 124.8 (C-4', C-5'), 126.8, 127.4 (C-2', C-7'), 127.8, 128.5 (C-3', C-6'), 134.6 (C-6), 141.3 (C-8'a-C-9'a), 143.9 (C-4'a-C-4'b), 150.8 (C-12), 153.6 (Bo-C-O), 156.3 (Fmoc-C-O), 160.5 (C-4'), 174.4 ppm (COOH); IR (ATR): ν = 1685, 1519, 1445, 1255, 1161, 1070, 1052, 722, 759, 736 cm−1; UV (MeCN): λmax (log ε) = 206 (4.58), 264 nm (4.25); HRMS (ESI) calcd for C14H13N5O2Na; found: 263.2374; [M+Na]+.

2'-NAMB-cbz-amino-3'-deoxy-thymidinyl-5'-aldehyde 8: A solution of phosphonate 7 (1.67 g, 4.49 mmol) in THF (34 mL) was added to a cooled (–78°C) solution of K2Bu (504 mg, 4.49 mmol) in THF (43 mL) at –78°C. The solution was stirred for 5 min at –78°C. Subsequently, a solution of aldehyde 3 (1.45 g, 2.95 mmol) in THF (18 mL) was added. The reaction mixture was stirred for 16 h and was allowed to warm slowly to RT during this time period. The resultant solution was cooled to 0°C and MeOH (5 mL) was added, after which the solution was diluted with EtOAc (200 mL). It was then washed with water (1 × 100 mL) and brine (1 × 100 mL), dried over Na2SO4, filtered, and evaporated under reduced pressure. The resultant crude product was purified by column chromatography (3:2 iso-hexanes-EtOAc) to give 8 as a colorless foam (2.06 g, 73%, diastereomeric mixture Z/E 93:7). As described before,14 the stereochemical assignment (Z/E) was based on the coordinates at C-1 and C-2.
empirical rules for NMR data established by Mazurkiewicz et al.\(^{[14,15,16]}\)

\(Z\)-6’-N-Boc-3-N-BOM-3’-N-Cbz-amino-5,6-didehydro-3-deoxy-thymidinyl amino acid benzyl ester 9: A solution of phosphate 7 (1.21 g, 3.25 mmol) in THF (25 mL) was added to a precooled (–78 °C) solution of KOBu (330 mg, 2.95 mmol) in THF (28 mL) at –78 °C. The resultant solution was stirred for 5 min at –78 °C. Subsequently, a solution of aldehyde 4 (1.45 g, 2.95 mmol) in THF (12 mL) was added. The reaction mixture was stirred for 16 h and was allowed to warm slowly to RT during this time period. The resultant suspension was cooled to 0 °C and MeOH (3 mL) was added, after which the solution was diluted with EtOAc (150 mL). It was then washed with water (1:80 mL) and brine (1:80 mL), dried over Na$_2$SO$_4$, filtered, and evaporated under reduced pressure. The resultant crude product was purified by column chromatography (3:2 iso-hexanes–EtOAc) to give 9 as a colorless foam (1.48 g, 68%, diastereomeric mixture Z/E 91:9). As described before,\(^{[14,15,16]}\) the stereochemical assignment (Z/E) was based on the empirical rules for NMR data established by Mazurkiewicz et al.\(^{[19]}\)

\(Z\)-9: TLC := tR$_1$ := 0.31 (1:1 iso-hexanes–EtOAc); \(\delta \) NMR (300 MHz, CDCl$_3$): \(\delta = 1.46 \) (s, 9H, C(CH$_3$)$_3$), 1.88 (d, J = 1.1, 3H, 7'-H), 2.40 (d, J = 13.9, 7.0, 7.0 Hz, 1H, 2’-H), 2.44 (d, J = 13.9, 8.0 Hz, 1H, 2’-H), 6.22 Hz, 1H, 2’-H), 4.06–4.13 (m, 1H, 3’-H), 4.69 (s, 2H, 1’-H), 4.79 (dd, J = 8.4, 6.4 Hz, 1H, 4’-H), 5.11 (d, J = 13.9, 8.0 Hz, 1H, 4’-H), 5.13 (d, J = 12.7 Hz, 1H, 1’-H), 5.23 (d, J = 12.6 Hz, 1H, 1’-H), 5.26 (d, J = 12.6 Hz, 1H, 1’-H), 5.47 (d, J = 9.7 Hz, 1H, 2’-H), 5.49 (d, J = 9.7 Hz, 1H, 2’-H), 6.36 Hz, 1H, 1’-H), 6.31 (d, J = 8.4 Hz, 1H, 5’-H), 6.85 (s, 1H, 6’-H), 7.09 (s, 1H, 6’-H), 7.21–7.40 ppm (m, 15H, aryl-H); \(\nu\) max, 17.43 (C-2’-O), 150.3 (Boc-C), 150.0 (Boc-C), 149.2 (C-2’), 161.5 (C-2’), 154.3 (C-1’), 138.7 (C-2’’), 150.4 (Boc-C), 124.9 (C-1’’), 120.0, 120.2, 128.3 (aryl-C), 135.4 (C-3’), 138.7 (C-2’’), 150.4 (Boc-C), 124.5 (C-1), 171.1 ppm (ester-C=O); IR (KBr): \(\nu\) max = 2101, 1704, 1652, 1420, 1250, 1156, 774, 736, 698 cm$^{-1}$; UV (MeCN): \(\lambda_{\text{max}}\) (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

- \(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].

\(\lambda\)max (log $\varepsilon$) = 206 (4.65), 266 nm (4.19); HRMS (ESI) calcd for C$_{13}$H$_{13}$NO$_2$; 633.2678; found: 633.2675 [M–H].
HPLC to give the pure (6'S)-diastereomer (S)-11 (900 mg from 1.0 g of diastereomeric mixture, obtained from several reactions). As described before,[13,14,16] the stereochromatic assignment (6'S/6'R) was based on the catalyst-controlled nature of the reaction and on an X-ray crystal structure of a nucleosyl amino acid derivative.[17b] (S)-11: M.p. 69 °C; TLC: Rf = 0.31 (3:2 isoo-hexanes-ETOAc); HPLC (analytical): tR = 30.5 min; HPLC (preparative): tR = 32.0 min; [α]280 [η]20 × c = +34.4 (c 1.1, CHCl3); 1H NMR (500 MHz, CD3OD, 50°C): δ = 1.39 (s, 9H, C(CH3)3), 1.68 (d, J = 13.9, 6.9, 6.9 Hz, 1H, 2′-H), 1.74 (d, J = 13.9, 8.4, 5.9 Hz, 1H, 2′-H), 1.88 (s, 3H, 7-H), 1.95 (dd, J = 14.4, 7.3, 7.2 Hz, 1H, 5′-H), 2.17–2.26 (m, 1H, 5′-H), 3.69 (dd, δ = 6-7.3, 4.2, 4.2 Hz, 1H, 4′-H), 3.83–3.94 (m, 1H, 3′-H), 4.35–4.75 (m, 1H, 6′-H), 4.71 (s, 2H, 1′-H), 4.91 (d, J = 12.5 Hz, 1′-H, 1′-H), 4.95 (d, J = 12.5 Hz, 1′-H, 1′-H), 5.01 (d, J = 12.3 Hz, 1′-H, 1′-H), 5.06 (d, J = 12.3 Hz, 1′-H, 1′-H), 5.11–5.29 (m, 1H, 3′-H), 5.51 (s, 2H, 2′-H), 5.91 (s, 2H, 1′-H), 5.97 (dd, δ = 5.9, 5.9 Hz, 1H, 1′-H), 6.93 (brs, 1H, 6-H), 6.99–7.19 (m, 11H, aryl-H), 7.23–7.27 (m, 2H, aryl-H), 7.33–7.37 ppm (m, 2H, aryl-H); 13C NMR (126 MHz, CD3OD, 50°C): δ = 12.8 (C-7), 28.0 (C(CH3)3), 35.6 (C-5), 37.3 (C-2), 51.2 (C-6), 54.1 (C-3), 66.8 (C-1′), 70.6 (C-′1), 72.1 (C-′2), 79.5 (C(CH3)3), 80.8 (C-4′), 84.7 (C-1′), 110.3 (C-3′), 122.7, 127.5, 127.7, 127.9, 128.0, 128.1, 128.1, 128.3, 128.4 (aryl-C), 133.4 (C-6), 135.6, 136.7, 138.7 (C-3′, C-′2, C-′3′), 150.7 (C-2), 155.2 (Boc=C–O), 155.7 (Boc=C–O), 162.7 (C-4′), 171.5 ppm (ester=O); IR (ATR): ν = 1709, 1647, 1528, 1270, 1237, 1121, 1012, 723, 693 cm−1; UV (MeCN): λmax (log ε) = 206 (4.66), 261 nm (4.27); HRMS (ESI) calcd for C32H30N4O7S: 741.3141; found: 741.3144 [M+H].

6′-N-Boc-3′-N-BOM-3′-N-Cbz-amino-3′-deoxy-(R)-thymidylidino acid amino benzyl ester (R)-11: The synthesis of (R)-11 was performed according to the procedure for the synthesis of (6′R)-diastereomer (R)-11 (370 mg from 430 mg of diastereomeric mixture, obtained from several reactions). As described before,[13,14,16] the stereochromatic assignment (6′S/6′R) was based on the catalyst-controlled nature of the reaction and on an X-ray crystal structure of a nucleosyl amino acid derivative.[17b] (R)-11: M.p. 64°C; TLC: Rf = 0.31 (3:2 isoo-hexanes-ETOAc); HPLC (analytical): tR = 37.5 min; HPLC (preparative): tR = 39.0 min; [α]280 [η]20 × c = +39.1 (c 0.87, CHCl3); 1H NMR (500 MHz, CD3OD, 50°C): δ = 1.38 (s, 9H, C(CH3)3), 1.60–1.69 (m, 1H, 2′-H), 1.70–1.80 (m, 1H, 5′-H), 1.76 (s, 3H, 7-H), 1.88–2.00 (m, 1H, 5′-H), 2.00–2.09 (m, 1H, 5′-H), 3.52–3.59 (m, 1H, 3′-H), 3.77–3.87 (m, 1H, 3′-H), 4.60–4.70 (m, 1H, 6′-H), 4.71 (s, 2H, 1′-H), 4.89 (d, J = 12.2 Hz, 1H, 1′-H), 4.97–5.04 (m, 1H, 1′-H), 5.01 (d, J = 12.5 Hz, 1′-H, 1′-H), 5.06 (d, J = 12.5 Hz, 1′-H, 1′-H), 5.19–5.29 (m, 1H, 6′-H), 5.48 (d, J = 9.5 Hz, 1H, 2′-H), 5.51 (d, J = 9.5 Hz, 1H, 2′-H), 5.75–5.85 (m, 1H, 1′-H), 6.93 (brs, 1H, 6-H), 7.00–7.19 (m, 11H, aryl-H), 7.23–7.27 (m, 2H, aryl-H), 7.33–7.37 ppm (m, 2H, aryl-H); 13C NMR (126 MHz, CD3OD, 50°C): δ = 12.9 (C-7), 28.0 (C(CH3)3), 35.7 (C-5), 37.3 (C-2), 52.0 (C-′3), 54.5 (C-′3), 70.6 (C-1′), 72.1 (C-′2), 79.5 (C(CH3)3), 80.9 (C-′4), 84.5 (C-1′), 110.3 (C-3′), 122.7, 127.5, 127.7, 127.9, 128.0, 128.1, 128.1, 128.3, 128.4 (aryl-C), 133.4 (C-6), 135.6, 136.7, 138.7 (C-3, C-′2, C-′3), 150.7 (C-2), 155.2 (Boc=C–O), 155.7 (Boc=C–O), 162.7 (C-4′), 171.5 ppm (ester=O); IR (ATR): ν = 1709, 1647, 1528, 1270, 1237, 1121, 1012, 723, 693 cm−1; UV (MeCN): λmax (log ε) = 206 (4.66), 261 nm (4.27); HRMS (ESI) calcd for C32H30N4O7S: 741.3141; found: 741.3139 (M+H).
The authors declare no conflict of interest.

Keywords: backbone modifications • DNA • oligonucleotides • peptides • stereoselective synthesis


Manuscript received: September 14, 2017
Accepted manuscript online: October 19, 2017
Version of record online: December 27, 2017