Synthesis of 3′,5′-Dipeptidyl Oligonucleotides

Ina Schwope, Colleen F. Bleczinski, and Clemens Richert*

Department of Chemistry, Tufts University, Medford, Massachusetts 02155

Received January 7, 1999

Peptide−DNA hybrids are richly functionalized analogues of biomolecules that may have applications as hybridization probes and antisense agents with tunable binding and targeting properties. So far, synthetic efforts have mainly focused on hybrids bearing a single peptide chain, either at the 5′- or the 3′-terminus. Such singly modified analogues are vulnerable to nuclease attack at the unmodified terminus. Here we report a convenient and high-yielding solid-phase synthesis of 3′- and 5′-modified analogues of DNA with aminoacyl and peptidyl appendages at both termini. Using MALDI-TOF mass spectra of crude products as the criterion, serine, glycolic acid, hydroxylauric acid, and dimethyl hydroxypropionic acid were tested as 3′-linker residues. The latter, together with a direct amide link at the 5′-terminus, gave the highest yields of hybrids. The optimized procedure assembles hybrids on a controlled pore glass support bearing three consecutive ω-hydroxy lauric acid linkers. This support greatly reduces side reactions observed with conventional supports, probably due to its ability to increase steric accessibility during coupling ("swelling") and its rapid hydrolysis during deprotection with ammonium hydroxide. Dihybrids with aromatic, basic, and acidic amino acid residues were prepared, including H-Phe-Gly-TGCGCA-DP-Phe-OH, where DP denotes the dimethyl hydroxypropionic acid linker, whose structure was confirmed via mass spectrometry and one- and two-dimensional NMR. Further, a mixed coupling with seven Fmoc-protected amino acids was shown to produce a combinatorial library of dipeptidyl DNA hybrids.

Introduction

Modified oligonucleotides have applications as hybridization probes, bioorganic model compounds, and potential drugs for antisense and antigene applications.[1] They are also analogues of complex biopolymers that pose synthetic challenges. This is particularly true for peptide−DNA and peptide−RNA hybrids, where the peptidyl portion introduces additional reactive groups to an already richly functionalized core molecule. A number of methods for the assembly of oligonucleotides with a peptide appendage to either the 5′- or the 3′-terminus have been published.[2] Besides numerous methods that use bifunctional linkers, there are techniques that lead to directly linked peptide nucleic acid hybrids, with either an amide link to amino-terminal DNA[3−6] or a phosphodiester linkage to a hydroxyl group of an amino acid side chain.[7−9]

For biomedical applications, a modification of both the 3′- and the 5′-terminus of oligonucleotides is desirable, as both termini are vulnerable to nuclease attack. If the modification is a peptide appendage, this requires the synthesis of 3′,5′-dipeptidyl oligonucleotides. While, to the best of our knowledge, no general synthetic methodology for the preparation of dipeptidyl DNA hybrids has been reported, other constructs, such as PNA−DNA−PNA chimera[10] and fluorochrome-bearing "molecular beacons"[11] are known, as well as 3′,5′-bis-conjugates, where the 3′-appendage is the tripeptide GGH.[12] As part of our study on the tuning of the affinity of oligonucleotides for

* To whom correspondence should be addressed. Telephone: (617) 627-3475. Fax: (617) 627-3443. E-mail: crichert@tufts.edu.

(1) Selected reviews: (a) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544−584. (b) Milligan, J. F.; Matteucci, M. D.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 1923−1937. (c) Stein, C. A.; Cheng, Y.-C. *Science* **1993**, *261*, 1004−1012. (d) Sanghvi, Y. S., Cook, P. D., Eds.; *Carbohydrate Modifications in Antisense Research*; ACS Symposium Series; American Chemical Society: Washington, DC, 1994; pp 1−22. (e) Hunziker, J.; Leumann, C. In *Modern Synthetic Methods*; Ernst, B., Leumann, C., Eds.; Verlag Helvetica Chimica Acta: Basel, 1995; pp 333−417. (f) DeMesmaeker, A.; Haener, R.; Martin, P.; Moser, H. E. *Acc. Chem. Res.* **1995**, *28*, 366−374. (g) Agrawal, S.; Iyer, R. P. *Pharmacol. Ther.* **1997**, *76*, 151−160. (h) Curcio, L. D.; Bouffard, D. Y.; Scanlon, K. J. *Pharm. Therapeutics* **1997**, *74*, 317−332. (i) Crooke, S. T. *Biotechnol. Genet. Eng. Rev.* **1998**, *15*, 121−157. (j) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem.* **1998**, *110*, 2954−2983; *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2797−2823.

(2) Selected references: (a) Haralambidis, J.; Duncan, L.; Tregear, G. W. *Tetrahedron Lett.* **1987**, *28*, 5199−5202. (b) Haralambidis, J.; Duncan, L.; Angus, K.; Tregear, G. W. *Nucleic Acids Res.* **1990**, *18*, 493−499. (c) Juby, C. D.; Richardson, C. D.; Brousseau, R. *Tetrahedron Lett.* **1991**, *32*, 879−882. (d) Eritja, R.; Pons, A.; Escarceller, M.; Giralt, E.; Albericio, F. *Tetrahedron* **1991**, *47*, 4113−4120. (e) Arar, K.; Monsigny, M.; Mayer, R. *Tetrahedron Lett.* **1993**, *34*, 8087−8090. (f) Bongartz, J. P.; Aubertin, A. M.; Milhaud, P. G.; Lebleu, B. *Nucleic Acids Res.* **1994**, *22*, 4681−4688. (g) Truffert, J.-C.; Lorthioir, O.; Asseline, U.; Thuong, N. T.; Brack, A. *Tetrahedron Lett.* **1994**, *35*, 2353−2356. (h) Zhu, T.; Wei, Z.; Tung, C.-H.; Dickerhof, W. A.; Breslauer, K.; Georgopoulos, D. E.; Leibowitz, M. J.; Stein, S. *Antisense Res. Dev.* **1993**, *3*, 265−275. (i) Wei, Z.; Tung, C.-H.; Zhu, T.; Dickerhof, W. A.; Breslauer, K. J.; Georgopoulos, D. E.; Leibowitz, M. J.; Stein, S. *Nucleic Acids Res.* **1996**, *24*, 655−661. (j) Truffert, J. C.; Asseline, U.; Thuong, N. T.; Brack, A. *Protein Pept. Lett.* **1995**, *2*, 419−424. (k) Corey, D. *J. Am. Chem. Soc.* **1995**, *117*, 9373−9374. (l) DelaTorre, B. G.; Avino, A.; Tarrason, G.; Piulats, J.; Albericio, F.; Eritja, R. *Tetrahedron Lett.* **1994**, *35*, 2733−2736. (m) Peyrottes, S.; Mestre, B.; Burlina, F.; Gait, M. J. *Tetrahedron* **1998**, *54*, 12513−12522. (n) Tung, C. H.; Rudolph, M. J.; Stein, S. *Bioconjugate Chem.* **1991**, *2*, 464−465. (o) Robles, J.; Pedroso, E.; Grandas, A. *Tetrahedron Lett.* **1991**, *32*, 4389−4392. (p) Dreef-Tromp, C. M.; van Dam, E. M. A.; van ven Elst, H.; van der Marel, G. A.; Van Boom, J. H. *Nucleic Acids Res.* **1990**, *18*, 6491−6495. (q) Kuyl-Yheskiely, E.; Dreef-Tromp, C. M.; Geluk, A.; van der Marel, G. A.; Van Boom, J. H. *Nucleic Acids Res.* **1989**, *17*, 2897−2905. (r) Kuyl-Yeheskiely, E.; Tromp, C. M.; Lefeber, A. W. M.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron* **1988**, *44*, 6515−6523. (s) Beltran, M.; Pedroso, E.; Grandas, A. *Tetrahedron Lett.* **1998**, *39*, 4115−4118.

(3) Bergmann, F.; Bannwarth, W. *Tetrahedron Lett.* **1995**, *36*, 1839−1842.

(4) Guibourdenche, C.; Seebach, D.; Natt, F. *Helv. Chim. Acta* **1997**, *80*, 1−13.

(5) Bruick, R. K.; Dawson, P. E.; Kent, S. B. H.; Usman, N.; Joyce, G. F. *Chem. Biol.* **1996**, *3*, 49−56.

(6) Tetzlaff, C. N.; Schwope, I.; Bleczinski, C. F.; Steinberg, J. A.; Richert, C. *Tetrahedron Lett.* **1998**, *39*, 4215−4218.

10.1021/jo990036d CCC: $18.00
Published on Web 05/29/1999

target strands,[13] we became interested in expanding our methodology for directly linked peptide−DNA hybrids from 5′-peptidyl oligonucleotides[6] to 3′,5′-dipeptidyl DNA hybrids. For these, a linker-free or minimal linker approach was sought, as this would ensure that little entropic penalty would have to be paid for engaging a linker in complex formation. This is expected to lead to "triple hybrids" that are able to bind RNA and DNA with their nucleic acid portions and both of their peptide appendages.

Here we report how 3′,5′-dipeptidyl DNA hybrids can be synthesized from commercially available Fmoc−amino acid building blocks, phosphoramidites, and a dimethyl hydroxypropionic acid moiety mimicking the linker portion of coenzyme A, a naturally occurring nucleic acid hybrid. Besides this biomimetic linker, our optimization study led to triester-derivatized controlled pore glass as the solid support giving the highest yield of dipeptidyl hybrids. With this support, our entirely solid-phase-based technique was shown to produce hybrids with aliphatic, aromatic, basic, and acidic amino acid residues, both on the level of single compounds and as a small chemical library.

Results

The synthetic work reported here focused on the construction of the 3′-terminal portion of 3′,5′-dipeptidyl oligonucleotides and relied on previously reported methodology for appending the 5′-peptidyl moiety.[6] To allow for release from the support during the one-step basic deprotection established for solid-phase DNA synthesis, an ester linkage between the solid support and the carboxy terminus of the peptide was chosen. The hydroxyl group of a hexaethyleneglycol-derivatized surface[14,15] of controlled pore glass (cpg), available from earlier work in these laboratories,[16] provided the hydroxyl component for the ester linkage. Recent reports on 3′-phosphodiester-linked peptide−oligonucleotide hybrids[17] suggested that the β-elimination expected[7a] to cleave the phosphodiester linkages between the 3′-terminus of oligonucleotides and the side chain hydroxyls of serine and threonine are manageable. We chose Fmoc-Cys(S-*t*-Bu)-OH as a terminal residue, since its side chain protecting group was believed to be resistant to ammonium hydroxide treatment and removable under neutral aqueous conditions compatible with DNA.[18] Coupling of this and the following residue was induced with a peptide coupling mixture of 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexaflourophosphate (HBTU), hydroxybenzotriazole monohydrate (HOBT), and Hünig's base (DIEA) in DMF. A capping step involving acetic anhydride and DMAP was performed after the esterification reaction to suppress DNA synthesis on any unreacted hydroxyl groups. Piperidine deprotection followed by coupling of Fmoc-Ser-OH provided solid support **1** (Scheme 1) with the primary β-hydroxyl group of serine as the starting point for DNA chain assembly.

Two oligonucleotides were synthesized on solid support **1**, the all-thymidine tetranucleotide **2** and the self-complementary, mixed sequence hexamer TGCGCA (**3**), where T denotes a 5′-amino-5′-deoxythymidine residue. The success of these and the following syntheses was judged by analyzing crude products (obtained after ammonium hydroxide treatment) by MALDI-TOF mass spectrometry. MALDI-TOF MS has previously been demonstrated to be well suited for the analysis of peptide−DNA conjugates[19,20] and gives more detailed structural information than amino acid analysis[21] or electrophoretic mobility. We employ MALDI conditions that have been shown to allow quantitative detection of compounds whose desorption/ionization efficiency is known.[22,16] Under the experimental and peak picking conditions used, the absolute mass accuracy in these spectra is ca. ±0.2% and the relative accuracy of assigned peak maxima at 2000 Da is ca. ±1−2 Da. Spectra of crude **2** showed two major side products, one whose molecular weight was 105 Da lower than that of the desired product at *m*/*z* 1510 and one whose molecular weight was 277 or 278 Da lower (Figure 1a). Both were tentatively identified as resulting from β-eliminations. The former as loss of *t*BuSSH from the side chain of the protected cysteine, followed by conjugate addition of OH^- to the resulting enone, and the latter as the elimination of the entire peptidyl substituent from the phosphodiester-linked hybrid via β-elimination at the serine side chain. Judged by MALDI spectra, the yield of the desired full-length hybrid was even lower for the mixed sequence hexamer **3** ($m/z = 2146$) than for **2** (Figure 1b). In addition to the chains with truncated 3′-appendages, a ladder of peaks corresponding to shortened oligonucleotide chains was found. These were most probably the result of cleavage reactions along the backbone, as the trityl signature of the DNA synthesis indicated high coupling yields.

These results prompted us to revise the synthetic scheme. To prevent β-eliminations, we decided to switch to a non-amino acid linker residue that does not bear a β-hydrogen and to employ glycine as the terminal residue. Hakala and collaborators have reported that glycolic

(7) (a) Kuyl-Yeheskiely, E.; van der Klein, P. A. M.; Visser, G. M.; van der Marel, G. A.; van Boom, J. H. *Recl. Trav. Chim. Pays-Bas* **1986**, *105*, 69−70. (b) Kuyl-Yeheskiely, E.; Tromp, C. M.; Lefeber, A. W. M.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron* **1988**, *44*, 6515−6523. (c) Dreef-Tromp, C. M.; van der Maarel, J. C. M.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1992**, 4015−4029.

(8) Ueno, Y.; Saito, R.; Hata, T. *Nucleic Acids Res.* **1993**, *21*, 4451−4457.

(9) (a) Robles, J.; Pedroso, E.; Grandas, A. *J. Org. Chem.* **1994**, *59*, 2482−2486. (b) Robles, J.; Pedroso, E.; Grandas, A. *Nucleic Acids Res.* **1995**, *23*, 4151−4161.

(10) van der Laan, A. C.; Brill, R.; Kuimelis, R. G.; Kuyl-Yeheskiely, E.; van Boom, J. H.; Andrus, A.; Vinayak, R. *Tetrahedron Lett.* **1997**, *38*, 2249−2252.

(11) Tyagi, S.; Kramer, F. R. *Nature Biotechnol.* **1996**, *14*, 303−308.

(12) Kahl, J. D.; McMinn, D. L.; Greenberg, M. M. *J. Org. Chem.* **1998**, *63*, 4870−4871.

(13) Sarracino, D. A.; Steinberg, J. A.; Vergo, M. T.; Woodworth, G. F.; Tetzlaff, C. N.; Richert, C. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2511−2516.

(14) Southern, E. M.; Case-Green, S. C.; Elder, J. K.; Johnson, M.; Mir, K. U.; Wang, L.; Williams, J. C. *Nucleic Acids Res.* **1994**, *22*, 1368−1373.

(15) Maskos, U.; Southern, E. M. *Nucleic Acids Res.* **1992**, *20*, 1679−1684.

(16) Bleczinski, C. F.; Richert, C. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1737−1743.

(17) Robles, J.; Maseda, M.; Beltran, M.; Concernau, M.; Pedroso, E.; Grandas, A. *Bioconjugate Chem.* **1997**, *8*, 785−788.

(18) Soukchareun, S.; Haralambidis, J.; Tregear, G. *Bioconjugate Chem.* **1998**, *9*, 466−475.

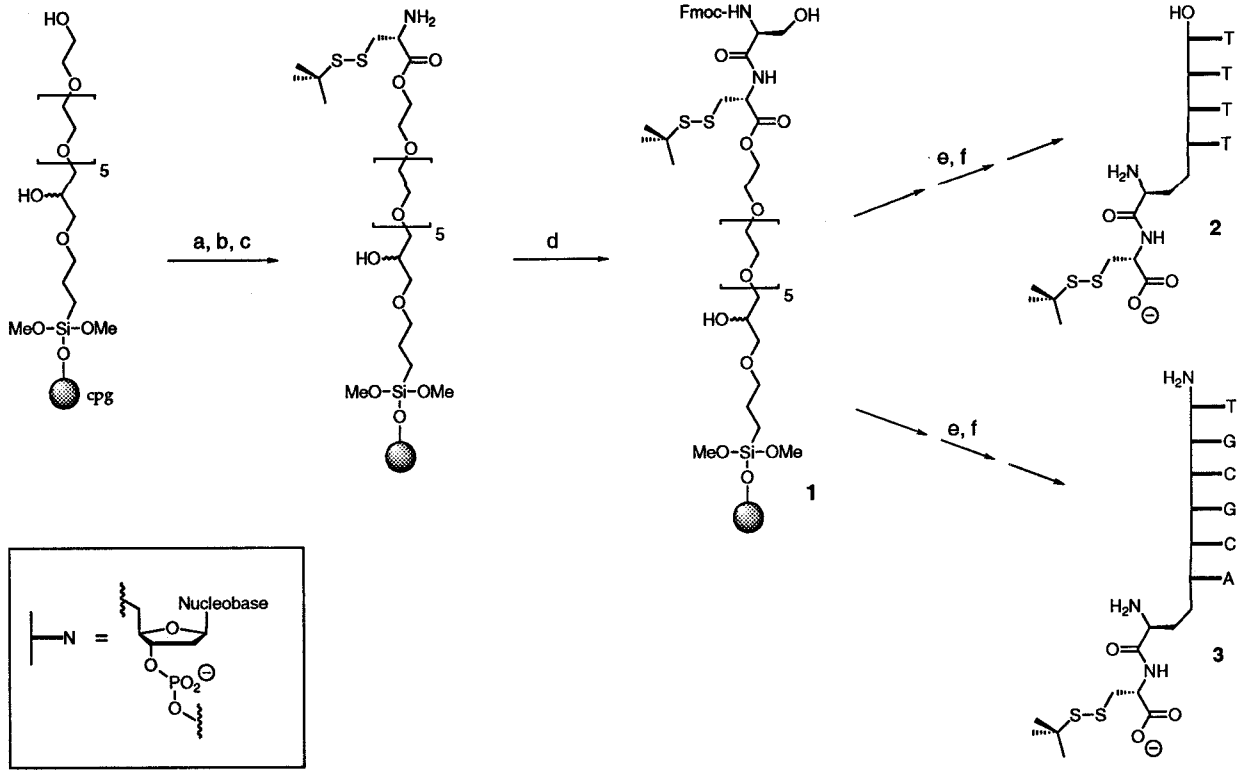
(19) Jensen, O. N.; Kulkarni, S.; Aldrich, J. V.; Barofsky, D. F. *Nucleic Acids Res.* **1996**, *24*, 3866−3872.

(20) Berlin, K.; Jain, R. K.; Tetzlaff, C.; Steinbeck, C.; Richert, C. *Chem. Biol.* **1997**, *4*, 63−77.

(21) Zhu, T. M.; Peng, Y.; Lackland, H.; Stein, S. *Anal. Biochem.* **1993**, *214*, 585−587.

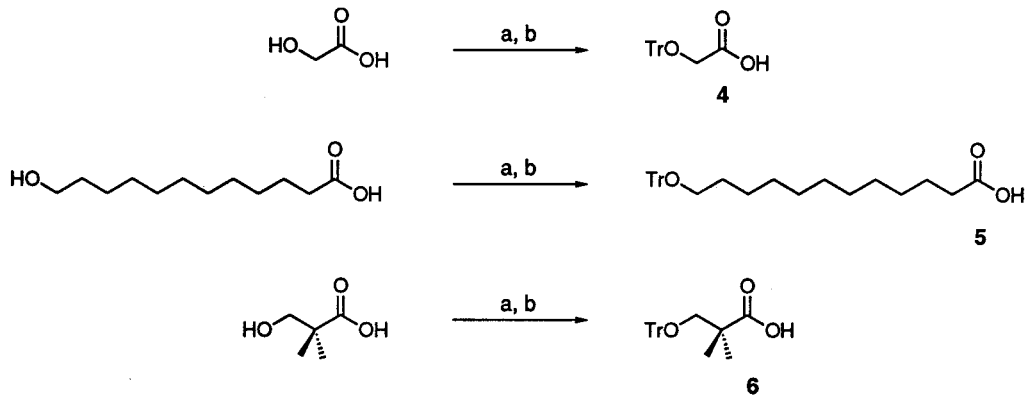
(22) Sarracino, D.; Richert, C. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2543−2548.

Scheme 1[a]



[a] Key: (a) Fmoc-Cys(*t*-BuS)-OH, HBTU, HOBT, DIEA, DMF; (b) Ac_2O, DMAP, lutidine, 1-methylimidazole, THF; (c) piperidine, DMF; (d) Fmoc-Ser-OH, HBTU, HOBT, DIEA, DMF; (e) DNA assembly via phosphoramidite protocol; (f) NH_4OH; cpg is controlled pore glass.

Scheme 2[a]



[a] Key: (a) Tr−Cl, pyridine; (b) aqueous AcOH.

acid can be used as a base-resistant linking moiety for solid-phase bound oligonucleotides.[23] This linker has the advantage of being short so that the peptidyl chain thus linked may engage in complex formation with target strands without major entropic cost. To prepare the glycolic acid for coupling reactions, we tritylated its hydroxyl group, leading to **4** as the building block for chain assembly (Scheme 2). To make the solid-phase syntheses more efficient, the diol solid support used in the initial protocol was replaced with commercially available aminopropyl-derivatized controlled pore glass (Scheme 3). 12-Hydroxy lauric acid was tritylated to give **5** and coupled to the derivatized glass. Detritylation with TFA/CH_2Cl_2 provided the hydroxyl group needed for temporary immobilization of the carboxy terminus of peptide−DNA hybrids.

Esterification of the hydroxy support with Fmoc−glycine, followed by deprotection with piperidine produced **7** (Scheme 3), whose amino group was coupled to glycolic acid building block **4**. Detritylation furnished **8**, which was subjected to automated DNA synthesis via the phosphoramidite protocol. Mixed hexamer **9** was obtained in higher yield than **3**, with the −105 Da peaks ascribed to the elimination/addition absent, but truncated oligomers were again abundant (see Supporting Information). Further, a MALDI signal at 58 Da lower than that of **9** was strong, indicating that chains lacking either the glycine or the glycolic acid residue were produced. This could have been due to a cleavage of the peptidic bond

(23) Hakala, H.; Heinonen, P.; Iitiae, A.; Loennberg, H. *Bioconjugate Chem.* **1997**, *8*, 378−384.

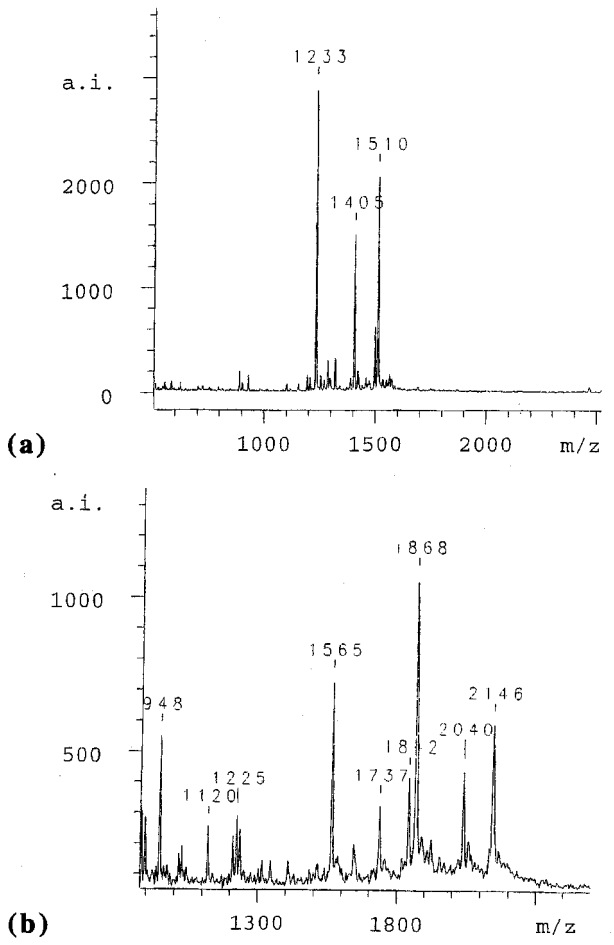


Figure 1. MALDI-TOF mass spectra of crude products (a) **2** and (b) **3**. Calculated molecular weights of products are given in the Experimental Section.

between the glycine and glycolic acid residue. Alternatively, DNA synthesis on surface positions without the glycolic acid or without the glycine residue could have occurred. To elucidate which of these two hypotheses was incorrect, we synthesized a sequence with a linker moiety whose molecular weight was sufficiently different from that of glycine. It was convenient to use tritylated hydroxylauric acid derivative **5**, which was already available from derivatization of the aminopropyl cpg. The 3′-modified oligonucleotide **11** resulting from chain assembly on the support derivatized with **5** was formed in much higher yield than **9** and **3** (Figure 2a). Not only was the amount of oligomers of 57/58 Da lower molecular weight considerably reduced, but the products of truncation reactions were also almost absent. This indicated that the long alkyl chain of the hydroxylauric acid linker suppressed fragmentation reactions, either because it positioned the terminal carboxylate such that intramolecular attack and/or catalysis could no longer occur through entropically favored transition states or because it provided a sufficiently long extension of the immobilization chain to make intermolecular attack between individual oligonucleotide chains less likely. A combination of changes in the synthetic protocol aimed at reducing the chain degradation reactions without resorting to the long alkyl chain of lauric acid proved successful.

First, a linker moiety that would reduce intramolecular attack over the level observed for the glycolic acid residue

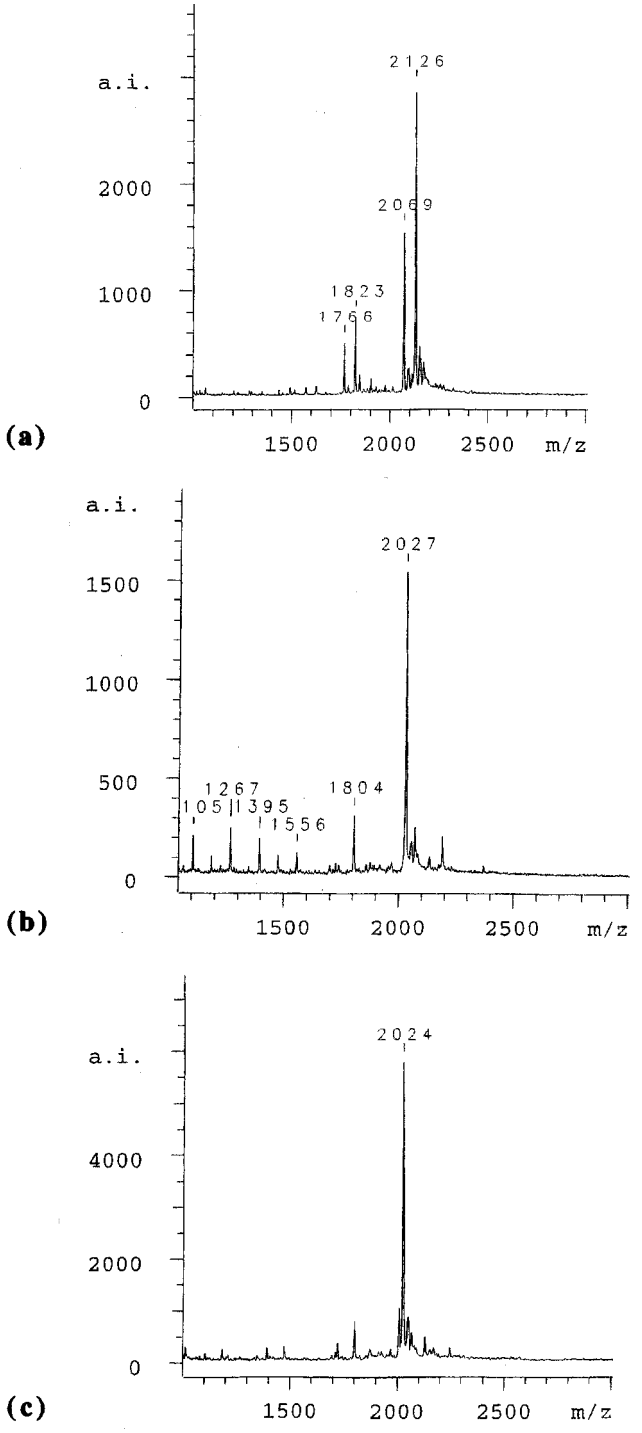
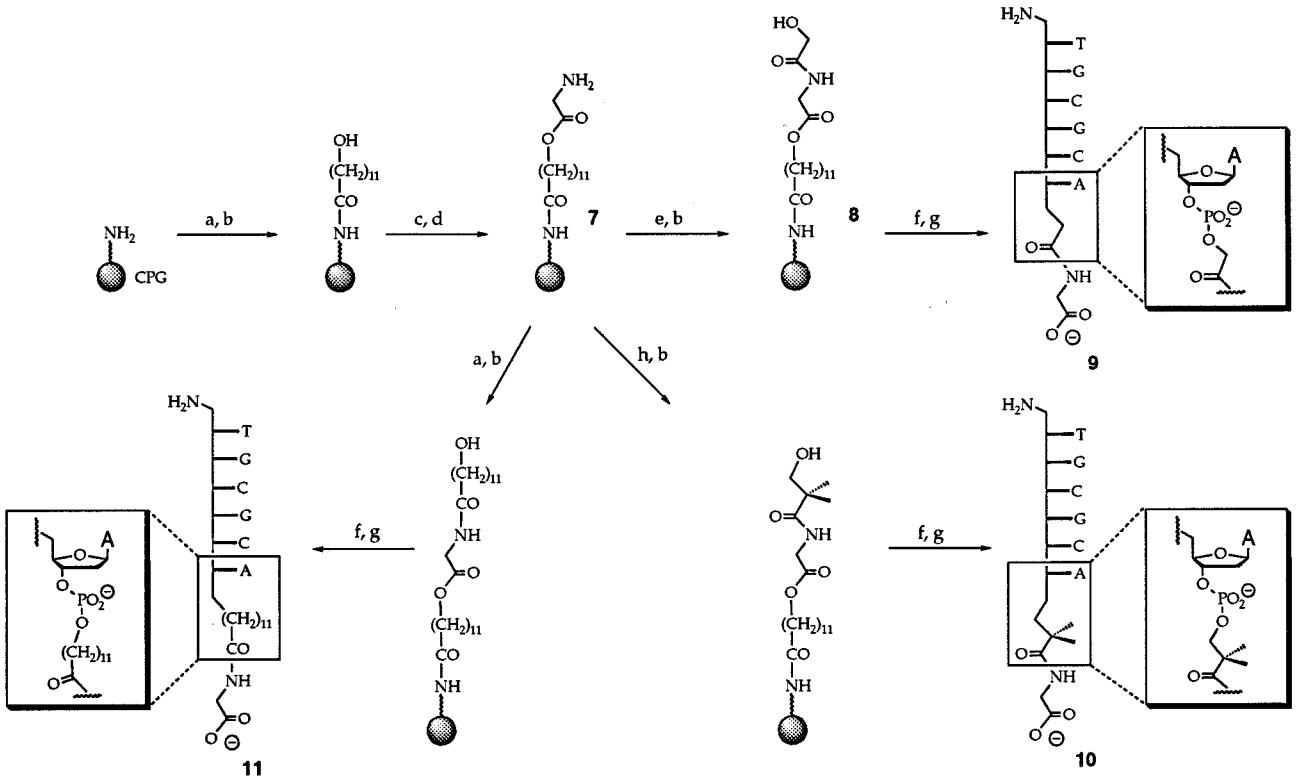


Figure 2. MALDI-TOF mass spectra of crude products (a) **11**, (b) **10** as obtained from a synthesis on monoester solid support **7**, and (c) **10** as obtained by synthesis on triester support **13** via intermediates **18** and **25**.

had to be identified. A survey of phosphate-linked hybrid structures found in nature led to coenzyme A, whose nucleic acid portion is linked to the functional β-mercaptoethylamine moiety via a pantothenic acid residue. The portion of this residue that separates the β-phosphate from the first nucleophilic group (the hydroxyl group of the pantothenic acid) is formally a *tert*-butyl group ($-O-CH_2C(CH_3)_2-$). Mimicking this group such that the carbonyl group of the peptide link in the appendage takes the position of pantothenic acid's hydroxyl group led to 3-hydroxy-2,2-dimethylpropionic acid as a short, "biomi-

Scheme 3[a]



[a] Key: (a) **5**, HBTU, HOBT, DIEA, DMF; (b) TFA, CH_2Cl_2; (c) Fmoc-Gly-OH; HOBT, HBTU, DIEA, DMF; (d) piperidine, DMF; (e) **4**, HBTU, HOBT, DIEA, DMF; (f) DNA assembly via the phosphoramidite protocol; (g) NH_4OH; (h) **6**, HBTU, HOBT, DIEA, DMF. See Scheme 1 for definition of DNA backbone shorthand.

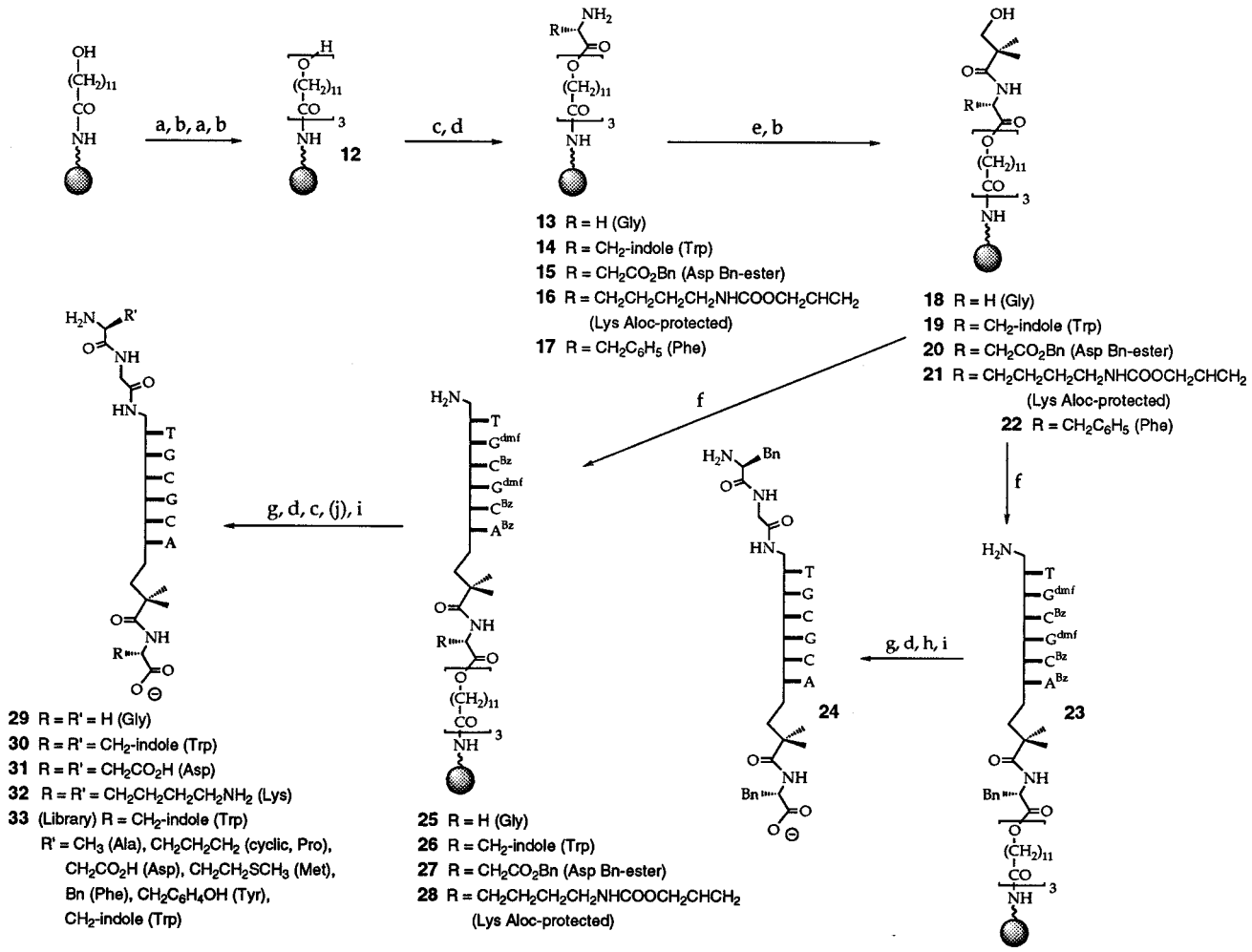
metic" linker between the DNA and the aminoacyl portion of our hybrids. Oligonucleotide synthesis on a solid support bearing this moiety (introduced as trityl derivative **6**) gave a high yield of **10** (Scheme 3), as judged by the MALDI spectrum of the crude product (Figure 2b).

To reduce degradation reactions originating in interchain attack and to ensure high coupling yields for peptide assembly at the 5′-terminus, a solid support bearing three, rather than one, hydroxylauric acid spacers was synthesized (**12**, Scheme 4). This was expected to have several effects. First, the longer spacer puts the oligonucleotide chains further away from the curved surface of the cpg beads, decreasing the packing density and thus the likelihood of interchain reactions on the support. Second, the three ester linkages result in a more rapid release of oligonucleotides during the deprotection, making degradation based on nucleophilic attack between liberated nucleophilic groups and phosphotriester moieties on other partially deprotected chains less prominent. Third, the long spacer should be well solvated in organic solvents, such as the DMF used for peptide couplings, thus increasing steric accessibility in a fashion much like the swelling of typical peptide resins does. This was expected to lead to high yields in the coupling reactions. Since the activated esters formed with the HBTU/HOBT/DIEA activation mixture are only moderately reactive toward hydroxyl groups, the three consecutive esterification reactions required for the preparation of **12** also decreased the loading of the cpg, again increasing steric accessibility and thus favoring high yields. The confirmation that at least one of these assumptions was correct came from the multistep solid-phase synthesis of **10** on solid support **13** via intermediates **18** and **25** (Scheme 4). When analyzed by MALDI-TOF mass spectrometry, the crude product of the ammonium hydroxide deprotection of **25** showed the full-length oligomer as the species dominating the spectrum (Figure 2c).

To validate the methodology, a series of dipeptidyl oligonucleotides of the general sequence H-Aa-Gly-DNA-DP-Aa-OH, where Aa stands for an amino acid residue, were synthesized via intermediates **14**−**17**, **19**−**22**, and **23** or **26**−**28** (Scheme 4). This included hybrids with aromatic residues (Trp, **30**), acidic residues (Asp, **31**), basic residues (Lys, **32**), and glycine residues (**29**). For **32**, the allyloxycarbonyl-protected building block Fmoc-Lys(Aloc)-OH was employed, necessitating a two-step deprotection, first with $Pd(PPh_3)_4$ and then with ammonium hydroxide. While **29**, **30,** and **31** were obtained in excellent purity (Figure 3a−c), the MALDI spectrum of crude **32** (Figure 3d) showed the presence of several low-level side products. When **28** was deprotected with intact allyloxycarbonyl groups, most of these side products were absent (see Supporting Information), demonstrating that they were not due to poor yields during assembly of the hybrid but due to side reactions during the deprotection sequence. Therefore, removal of the lysine side chain protection after the ammonium hydroxide treatment may be advantageous.[24] Hybrids **29**−**32** were purified by RP-HPLC, followed by confirmation of product identity by MALDI-TOF MS.

(24) (a) Lemaire-Audoire, S.; Savignac, M.; Blart, E.; Pourcelot, G.; Genet, J. P.; Bernard, J. M. *Tetrahedron Lett.* **1994**, *35*, 8783−8786. (b) Lemaire-Audoire, S.; Savignac, M.; Blart, E.; Bernard, J. M.; Genet, J. P. *Tetrahedron Lett.* **1997**, *38*, 2955−2958.

Scheme 4[a]



[a] Key: (a) **5**, HBTU, HOBT, DIEA, DMF; (b) TFA, CH_2Cl_2; (c) Fmoc-Aa-OH; HOBT, HBTU, DIEA, DMF; (d) piperidine, DMF; (e) **6**, HBTU, HOBT, DIEA, DMF; (f) DNA assembly via the phosphoramidite protocol; (g) Fmoc-Gly-OH, HBTU, HOBT, DIEA, DMF; (h) Fmoc-Phe-OH, HBTU, HOBT, DIEA, DMF; (i) NH_4OH; (j) for **28**, $Pd(PPh_3)_4$, $[Et_2NH_2]^+HCO_3^-$, CH_2Cl_2. Aa, amino acid residue; Aloc, allyloxycarbonyl. See Scheme 1 for definition of DNA backbone shorthand.

In one case (Aa = Phe, **24**), a dipeptidyl−DNA hybrid of the general sequence H-Aa-Gly-DNA-DP-Aa-OH was prepared on a scale sufficient to analyze the product by NMR. The one-dimensional proton NMR spectrum and the MALDI-TOF mass spectrum of purified **24** are shown in Figure 4. Expansions of relevant regions of two-dimensional NMR spectra are shown in Figure 5. Assignment of the nucleic acid portion of **24** started from the NOESY cross-peak between H6 and the methyl group of the only thymidine residue in the sequence (T3, Scheme 5). This was unambiguously discernible in a spectrum acquired in D_2O at 300 ms mixing time. The cross-peak between H6 and H1′ of the same residue provided an entry into the sequential assignment of the H8/H6 and H1′ resonances of all nucleotides. The H1′ resonances could then be connected to H2′/2″, H3′, and H4′-protons via cross-peaks in COSY and TOCSY spectra. For T3, C5, and A8, at least one H5′ resonance could be assigned based on TOCSY cross-peaks, together with an isochronous resonance for H5′/5″ of G4 and/or G5. While the phenylalanine residues clearly manifested themselves in the AA′BB′C systems of the aromatic rings (Figure 4b), linking these to their nonaromatic resonances was nontrivial. The expected cross-peaks between the Hα and Hβ protons were detected in COSY and TOCSY spectra of D_2O solutions, but only the TOCSY cross-peaks originating from the NHα of Phe10, detected in a spectrum acquired in H_2O/D_2O, clarified the assignment. Since the amino-terminal Phe1 does not give rise to such a cross-peak, the tentative assignment of its Hα, which was based on the chemical shift information, could be unambiguously confirmed. Another NH assignment based on the H_2O/D_2O spectra was that at the 5′-position of T3. Its cross-peaks to the already known resonances of H5′/5″ of the same residue were decisive. The linking residue DP9 contains two isolated spin systems, of which the methyl groups were identified based on chemical shift and integration, leaving the glycine residue at position 2 as the last moiety whose presence had to be confirmed. The only remaining unassigned NOESY cross-peak in the NHα/Hα region led to signals for Hα/Hα′ of this residue (Gly2), completing the NMR characterization of **24**.

Since Robles and collaborators have reported that 3′-modified oligonucleotides with phosphodiester linkages to the nonnucleic acid terminus can be susceptible to exonuclease attack,[17] we synthesized non-self-complementary hybrid **35** via **34** (Scheme 6). This compound was mixed with an equimolar amount of the unmodified DNA octamer[6] 5′-TGGTTGAC-3′ and then exposed to snake venom phosphodiesterase (EC 3.1.4.1). The disap-

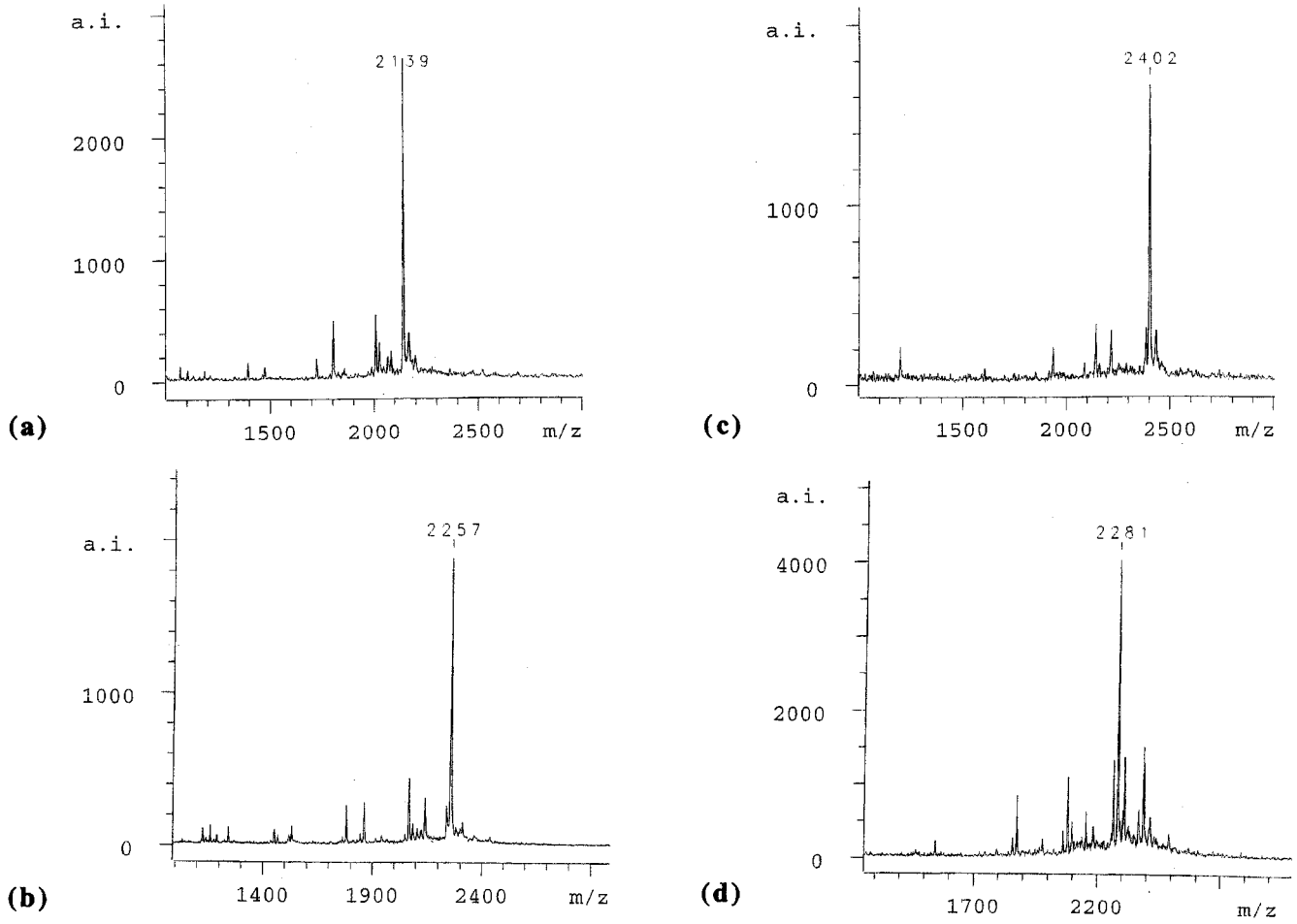


Figure 3. MALDI-TOF mass spectra of crude products obtained with optimized procedures: (a) **29**; (b) **31**; (c) **30**; (d) **32**.

pearance of the full-length oligomers was monitored by MALDI-TOF mass spectrometry. While the unmodified octamer was readily degraded by the nuclease, more than 75% of hybrid **35** was intact at the end of the experiment (Figure 6).

Finally, a combinatorial library of dipeptidyl−DNA hybrids was prepared. To this end, intermediate **26** with a 3′- and carboxy-terminal tryptophan residue was 5′-extended with a glycine residue, followed by piperidine deprotection and coupling with a mixture of the Fmoc−amino acid building blocks of alanine, proline, aspartic acid (side chain benzyl protected), methionine, phenylalanine, tyrosine, and tryptophan (Scheme 4). Except for Fmoc-Pro-OH, for which twice the amount was used, all amino acids were employed at near equimolar amounts. The MALDI-TOF mass spectrum of the crude library (**33**, Figure 7) showed all components of the expected library with greater peak intensity than the uncoupled oligomer with only the single glycine residue at the 5′-terminus (at *m*/*z* 2212), demonstrating a high yielding coupling reaction. Further, no major side products were detected in the mass range of the full-length oligomers, making the crude library suitable for mass spectrometrically monitored selection experiments (SMOSE).[20]

Discussion

Some of the difficulties encountered when optimizing the immobilization method and the peptidyl−DNA linkage at the 3′-terminus were unexpected. The installation of a base-labile ester linkage to the solid support and the assembly of an oligonucleotide on a hydroxyl group displayed by the peptide portion are seemingly straightforward synthetic goals. In fact, a so-called "3′-amino-modifier" solid support with an Fmoc-protected amino group, a DMT-protected hydroxyl group, and an ester link to a succinyl anchor is commercially available,[25] albeit as a racemic mixture. We were able to synthesize 3′-amino-modified oligonucleotides on this support in good yield, but no peptidyl hybrids were obtained when we attempted peptide synthesis followed by DNA assembly and deprotection with NH_4OH (results not shown), probably because of a lack of reactivity of the amino groups on the tightly packed controlled pore glass surface. This highlights the need for a compromise between loading and reactivity, but also explains why, despite the availability of the "3′-amino-modifier", the development of solid supports for the synthesis of 3′-hybrids has become an active field of research.[26−28] Compared to other methods, our methodology has the advantage of including 5′-modifications, being versatile and entirely solid-phase based, employing commercially available standard DNA

(25) 3′-Amino-modifier C3 CPG, Glen Research, Sterling VA,; see also: Zendegui, J. G.; Vasquez, K. M.; Tinsley, J. H.; Kessler, D. J.; Hogan, M. E. *Nucleic Acids Res.* **1992**, *20*, 307−314.

(26) (a) McMinn, D. L.; Hirsch, R.; Greenberg, M. M. *Tetrahedron Lett.* **1998**, *39*, 4155−4158. (b) Matray, T. J.; Yoo, D. J.; McMinn, D. L.; Greenberg, M. M. *Bioconjugate Chem.* **1997**, *8*, 99−102.

(27) Shchepinov, M. S.; Stetsenko, D. A. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1181−1184.

(28) McMinn, D. L.; Greenberg, M. M. *J. Am. Chem. Soc.* **1998**, *120*, 3289−3294.

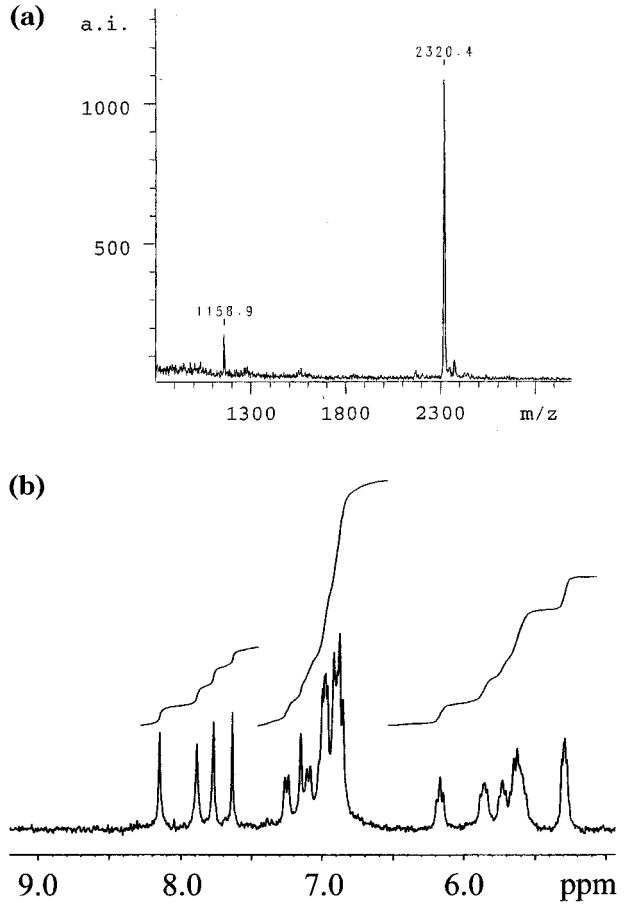


Figure 4. Selected spectra of purified **24**: (a) MALDI-TOF spectrum showing both $[M - H]^-$ and $[M - 2H]^{2-}$ ions, (b) nucleobase and H1′ region of the ^{1}H NMR spectrum in D_2O at 300 MHz and 22 °C.

and amino acid building blocks, and producing the hybrids in conventional deprotection reactions.

The dependence of hybrid yield on the nature of the linking residue at the 3′-terminus is also noteworthy. Our results with serine confirm that the propensity for β-elimination is too high to make this amino acid residue suitable for deprotection with ammonium hydroxide, the base for which protecting groups on the commercially available nucleic acid building blocks have been optimized. It is unclear to us how acetylation of the amino terminus may suppress the elimination,[17] and it is reassuring that homoserine, which does not position the phosphate leaving group next to the acidic α-hydrogen, is suitable for hybrid synthesis.[29] According to the MALDI spectra of crude **9**−**11**, the glycolic acid linker, but not the hydroxylauric acid or the dimethylhydroxypropionic acid (DP) residue, causes fragmentation of the DNA backbone during deprotection. This assumption is based on consistently high trityl release during DNA assembly and the presence of terminated chains in the crude product of **9**. Intramolecular side reactions in oligonucleotide conjugates are not unknown.[30,31] In our case, the attack probably occurs on phosphotriesters, as

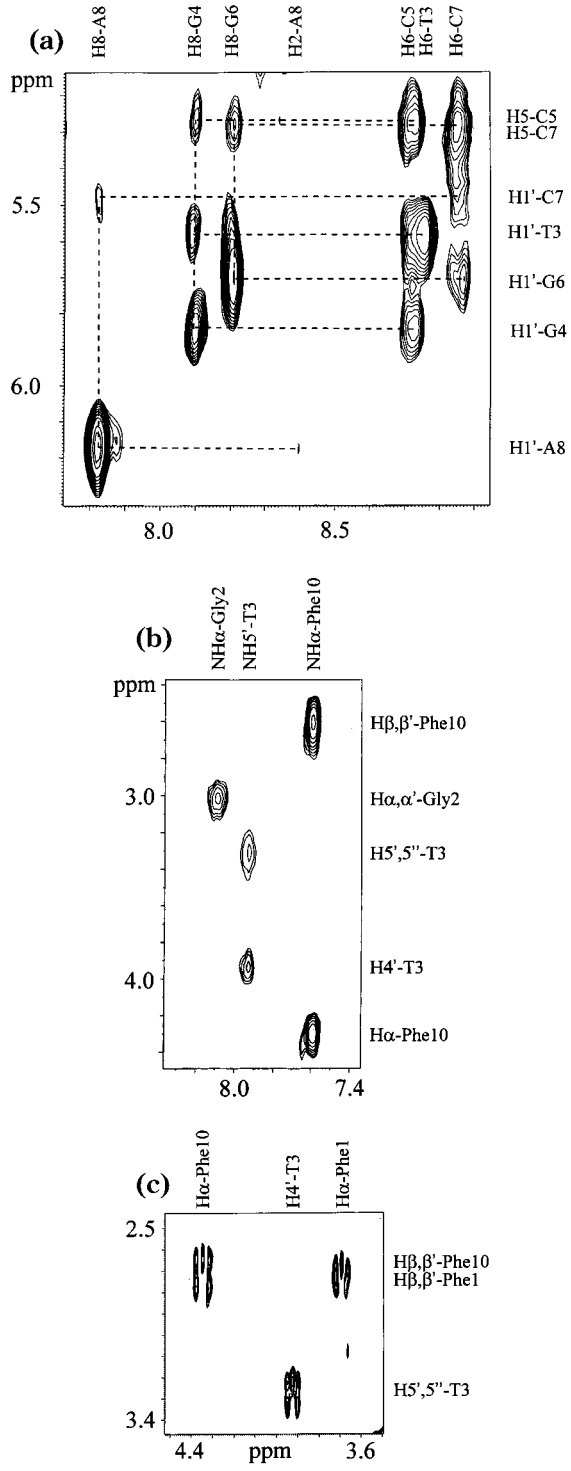
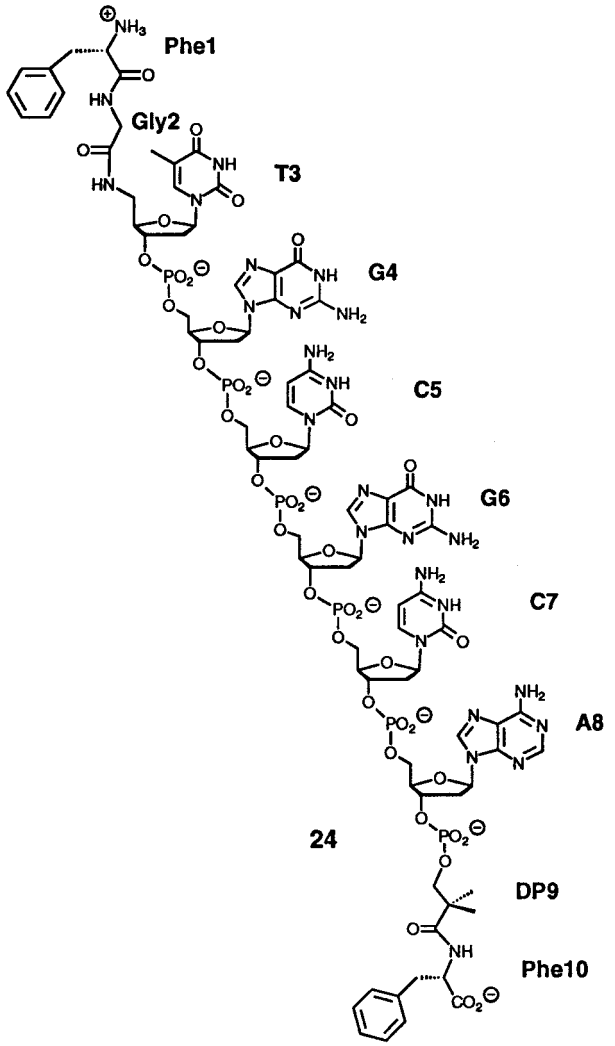


Figure 5. Selected regions of 2D NMR spectra of H-Phe-Gly-TGCGCA-DP-Phe−OH (**24**): (a) cross-peaks between H6/H8 and H1′ and H5 resonances of the nucleotides in a NOESY spectrum acquired in H_2O/D_2O (9:1) at 7 °C, 100 mM NaCl, and 300 ms mixing time; connectivities are indicated by broken lines, lines for intraresidue connectivities in pyrimidines are left out for clarity; (b) cross-peaks between backbone amide protons and their respective spin systems in the TOCSY spectrum at 7 °C in H_2O/D_2O (9:1), 100 mM NaCl, t_{mix} = 60 ms; (c) TOCSY cross-peaks between protons of the aliphatic spin systems of Phe1, T3, and Phe10 as measured at 22 °C in D_2O.

prolonged treatment of the fully deprotected hybrids with ammonium hydroxide does not induce cleavage. The long alkyl chain of the lauric acid linker, which is typical for

(29) Beltran, M.; Maseda, M.; Perez, Y.; Robles, J.; Pedroso, E.; Grandas, A. *Nucleosides Nucleotides* **1997**, *16*, 1487−1488.
(30) Acedo, M.; Tarrason, G.; Piulats, J.; Mann, M.; Wilm, M.; Eritja, R. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1577−1580.
(31) Juodka, B.; Sasnauskiene, S.; Shabarova, Z. *J. Carbohydr. Nucleosides Nucleotides* **1981**, *8*, 519−535.

Scheme 5. Structural Formula and Numbering of Residues for Compound 24

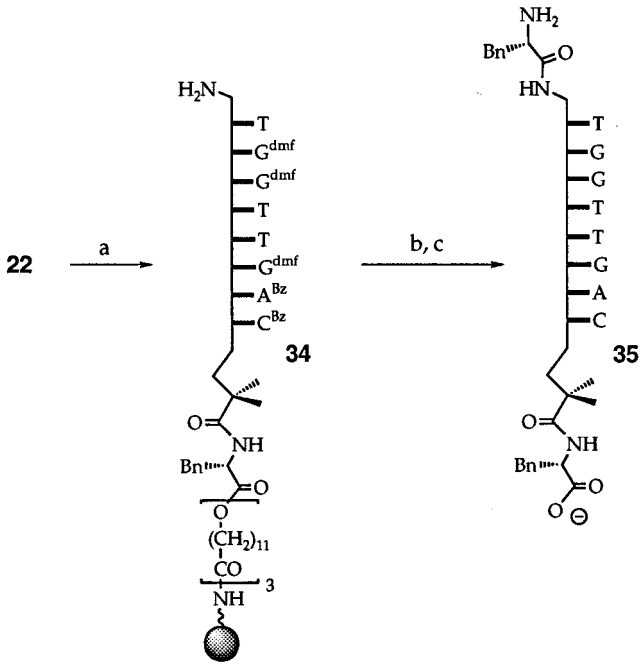


linkers used in bioconjugation techniques,[32] apparently disfavors such attacks. The dimethylhydroxypropionic acid (DP) residue accomplishes this with a much shorter contour length, shorter, in fact, than that of homoserine and with less flexible a hydrocarbon chain. The residual low-level fragmentation reactions may be abolished by removing the cyanoethyl groups from the phosphotriesters with anhydrous base prior to deprotecting the oligomers. Gait and collaborators have found that this increases the yield for 5′-peptidyl−DNA hybrids.[2m] In doing so, one runs the risk of premature strand release due to intramolecular cleavage reactions at the linker[33] or one has to resort to differently designed solid supports.[34]

Our study also demonstrates the usefulness of MALDI-TOF mass spectrometry as a routine tool for optimizing syntheses of modified oligonucleotides. Judging product purity by measuring the peak intensities of the full-length product and fragmentation side products under conditions known to allow quantitative analysis[16,20,22] yields a lower limit of product purity, as shorter oligonucleotides are known to produce stronger MALDI signals than longer chains of the same nucleobase composition.[35] The NMR analysis of purified **24**, together with the HPLC analysis of selected other crude hybrids, complemented the mass spectrometric technique and allowed a more detailed structure proof. At the same time, it provided chemical shift information for the phenylalanine residues that suggests little or no stacking of the phenyl rings on either nucleobases or each other. Preliminary results from experiments with 3′,5′-hybrids with longer and more lipophilic peptidyl appendages indicate that this is not always the case.[36]

The synthesis of 3′,5′-dipeptidyl−DNA hybrid **24** also provided anecdotal evidence for the increasingly protein-

Scheme 6[a]



[a] Key: (a) DNA assembly via the phosphoramidite protocol; (b) Fmoc-Phe-OH, HBTU, HOBT, DIEA, DMF; (c) NH_4OH.

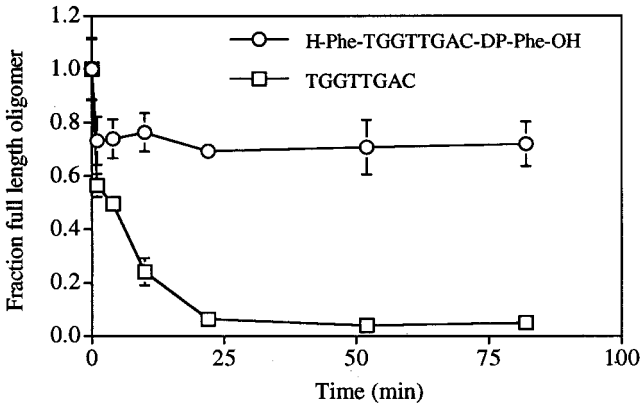


Figure 6. Nuclease stability of 3′,5′-dipeptidyl-hybrid. Kinetics of nuclease degradation of **35** and its unmodified DNA counterpart 5′-TGGTTGAC-3′ by snake venom phosphodiesterase (EC 3.1.4.1), as measured by MALDI-TOF mass spectrometry.

(32) Hermanson, G. T. *Bioconjugation Techniques*; Academic Press: San Diego, 1996.

(33) Palom, Y.; Grandas, A.; Pedroso, E. *Nucleosides Nucleotides* **1998**, *17*, 1177−1182.

(34) (a) Brown, T.; Pritchard, C. E.; Turner, G.; Salisbury, S. A. *J. Chem. Soc., Chem. Commun.* **1989**, 891−893. (b) Stengele, K.-P.; Pfleiderer, W. *Tetrahedron Lett.* **1990**, *31*, 2549−2552.

(35) Schneider, K.; Chait, B. T. *Org. Mass Spectrom.* **1993**, *28*, 1353−1361.

(36) Schwope, I.; Richert, C. Unpublished results.

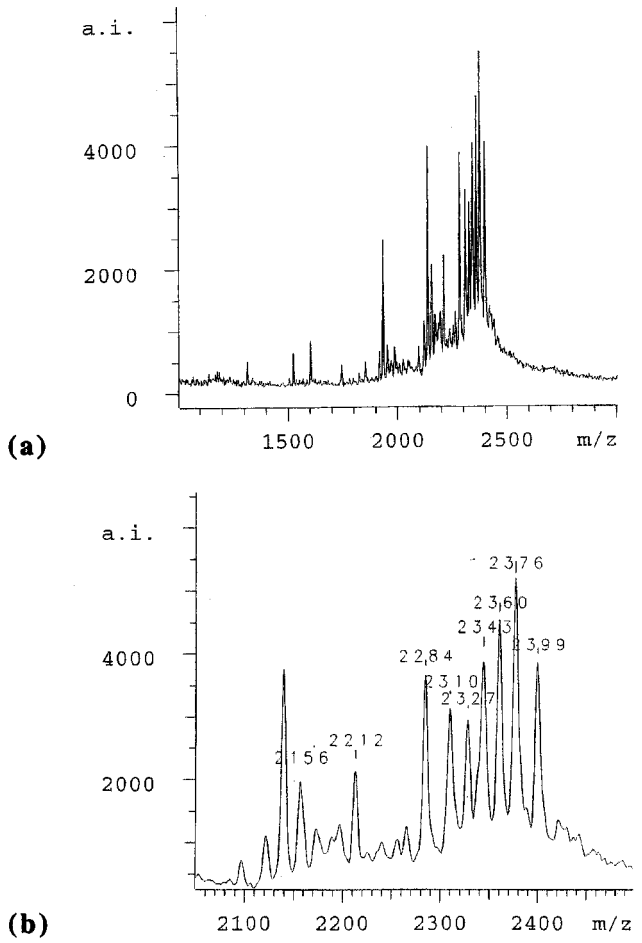


Figure 7. MALDI-TOF mass spectrum of crude library **33**: (a) overview; (b) expansion of the mass range of the pseudomolecular ions of the library components. Peaks at lower molecular weight originate from uncoupled sequences without 5′-substituent (*m*/*z* 2212) and with only the glycine residue (*m*/*z* 2156).

like properties of such biomolecular hybrids. Adsorption to glass and plastic surfaces was noted when quantified material was handled and then requantified. Further, highly concentrated solutions of lipophilically modified **24** and **30**, but not the more polar **31** and **32**, gave a small second product-containing peak in HPLC traces. These were tentatively identified as duplexes or aggregates, as reinjection of the material collected from these peaks again produced two peaks, one at the expected retention time for the hybrid and one at the retention time of the duplexes/aggregates. This indicates that the terminal appendages significantly change the molecular recognition properties of the oligonucleotides. For RP_{18} HPLC, this change produces a technological annoyance, but in terms of molecular targets, it may be beneficial and may be used to tune target affinity. Since selection studies are just beginning, one can only speculate on applications for 3′,5′-hybrids, but antisense and hybridization probes with fine-tuned target affinities and nuclease stabilities, improved molecular beacons, and bioorganic model compounds do come to mind. Peptide sequences to be appended to the 3′-terminus could include viral fusion peptides,[37] nuclear target signals,[38] or photoactivable residues,[39] whereas 5′-residues may be chosen to increase the affinity for RNA.[13]

Conclusion

In summary, we have developed a synthesis of 3′,5′-dipeptidyl oligonucleotides with a short linking residue using building blocks, reagents, and instrumentation readily available to nucleic acid chemists. This synthesis may be fully automated using solid-phase synthesizers. The results from our optimization study indicate that a nonoptimized, short linker can position reactive groups on the 3′-appendage such that they attack the DNA chains. The resulting fragmentation reactions can be suppressed with a linker mimicking a portion of coenzyme A and a triester anchor to the support. The methodology is applicable to a broad set of Fmoc−amino acid building blocks and nucleoside phosphoramidites, allowing the synthesis of combinatorial libraries. Since peptidyl−DNA hybrids access both nucleic acid and peptide structure space, many new functional molecules may emerge from this class of biomolecular analogues.[40]

Experimental Section

General. The following anhydrous solvents were purchased and used without further purification: acetonitrile and pyridine from Fisher, dichloromethane from VWR (EM brand), and DMF from Fluka. Reagents and building blocks were also used as provided by the supplier; 12-hydroxylauric acid and 2,2-dimethyl-3-hydroxypropionic acid were from Fluka, glycolic acid, triethylamine, and diisopropylethylamine were from Aldrich, and trityl chloride was from Sigma. Phosphoramidites for dA^{Bz}, dC^{Bz}, and T were obtained from Perseptive Biosystems; dG^{dmf} was from ABI/Perkin-Elmer. The protected 5′-amino-5′-deoxythymidine phosphoramidite was prepared as described previously.[6] Amino acid building blocks Fmoc-Gly-OH, Fmoc-Trp-OH, Fmoc-Cys(S-tBu)-OH, Fmoc-Asp(OBzl)-OH, Fmoc-Phe-OH Fmoc-Pro-OH, Fmoc-Met-OH, Fmoc-Tyr-OH, and Fmoc-Ala-OH were purchased from Advanced Chem Tech, as well as HOBT and HBTU. Fmoc-Lys(Alloc)-OH was from Perseptive Biosystems, and Fmoc-Lys(Fmoc)-OH was from Novabiochem. Underivatized controlled pore glass was native-500-cpg from Prime Synthesis, mean pore size 559 Å, surface area 63.6 m^2/g, particle size 80−200 mesh. Aminopropyl-derivatized controlled pore glass was AMP-cpg from CPG Inc., mean pore diameter 1049 Å, surface area 22.7 m^2/g, particle size 120−200 mesh, loading 130 μmol/g. Analytical thin-layer chromatography, column chromatography, and high-resolution mass spectrometry were performed as previously described.[39] Trityl signals were measured in solutions containing 0.1% *p*-toluenesulfonic acid on a Perkin-Elmer Lambda 10 spectrophotometer. Oligonucleotides were purified by HPLC on analytical C18 or C4 reversed-phase columns (Alltech or Vydac) using a gradient of CH_3CN in 0.1 M triethylammonium acetate and detection at 260 nm. NMR spectra were acquired on 300 MHz spectrometers. For 2D spectra, 256 increments of 2000 data points were acquired in phase-sensitive mode using presaturation during the recycle delay as water suppression. Data were zero-filled to 1k in f1, followed by apodization and analysis in XWINNMR. MALDI-TOF spectra were acquired on a Bruker BIFLEX spectrometer in negative, linear mode at 20 kV with a delayed extraction voltage of 17.5 kV and 180 ns delays for laser shots at 60−70 J/shot. A mixture of 2,2,6-trihydroxy acetophenone (0.3 M in ethanol), diammonium citrate (0.1 M in water), and CH_3CN (5:2:3) was

(37) Soukchareun, S.; Tregear, G. W.; Haralambidis, J. *Bioconjugate Chem.* **1995**, *6*, 43−53.
(38) Reed, M. W.; Fraga, D.; Schwartz, D. E.; Scholler, J.; Hinrichsen, R. D. *Bioconjugate Chem.* **1995**, *6*, 101−108.
(39) Berlin, K.; Jain, R. K.; Simon, M. D.; Richert, C. *J. Org. Chem* **1998**, *63*, 1527−1535.
(40) Service, R. F. *Science* **1998**, *282*, 1020−1021.

used as matrix. Calculated masses are average masses, and detected *m*/*z* values are maxima of unresolved isotope envelopes. The names of hybrids are written from amino to carboxy terminus, using peptide short hand, with the DNA sequence from 5′- to 3′-terminus inserted as one letter code and GA, LA, DP, and Aa denoting glycolic acid, hydroxylauric acid, dimethyl hydroxypropionic acid, and unspecified amino acid residues, respectively.

2-*O*-Tritylglycolic Acid (4). Glycolic acid (1.00 g, 13.10 mmol) and trityl chloride (7.33 g, 26.20 mmol, 2 equiv) were dried under high vacuum for 1 h and then dissolved in pyridine (30 mL), followed by addition of DMAP (4 mg, 0.03 mmol, 2.5 mol %). The reaction was allowed to proceed for 48 h at 22 °C. The solvent was evaporated, and the residue was redissolved in CH_2Cl_2 (20 mL). The solution was washed with 5% aqueous AcOH (2 × 120 mL) and with brine (4 × 120 mL) and dried over Na_2SO_4. After evaporation of the solvent in vacuo, the residue was purified by column chromatography (silica, CH_2Cl_2, MeOH 5−20%) to give tritylglycolic acid **4**, 2.50 g (60% yield) as a colorless solid. ^{1}H NMR ($CDCl_3$): δ 3.94 (s, 2H), 7.30−7.43 (m, 9H), 7.43−7.50 (m, 6H). ^{13}C NMR ($CDCl_3$): δ 62.5, 88.3, 127.9, 128.6, 129.0, 143.3, 175.5. MS (EI+) *m*/*z* 317 (M^+, 35%). HRMS (FAB, NOBA/LiOAc) for $C_{21}H_{18}O_3Li$ [M + Li]$^+$: calcd 325.1411, found 325.1415.

12-Trityloxylauric Acid (5). 12-Hydroxylauric acid (1.0 g, 4.6 mmol) and trityl chloride (2.58 g, 9.2 mmol) were dissolved in dry dichloromethane (75 mL). Triethylamine (6.4 mL, 46 mmol) was added, and the reaction was stirred under argon at room temperature for 36 h. Methanol (50 mL) was added, followed by stirring for an additional 4 h. The solution was treated with 5% acetic acid, then washed with saturated NaCl to a pH of 7 and dried over Na_2SO_4. Following evaporation under reduced pressure, the yellow residue was subjected to two consecutive flash chromatography columns using silica with petroleum ether/ethyl acetate (4:1). The pure fractions were combined, yielding 1.22 g (2.7 mmol, 58%) of the target compound as a slightly yellow oil. ^{1}H NMR ($CDCl_3$): δ 1.20−1.45 (m, 14H), 1.59−1.73 (m, 4H), 2.38 (t, J = 7.5 Hz, 2H), 3.07 (t, J = 6.6 Hz, 2H), 7.22−7.38 (m, 9H), 7.45−7.53 (m, 6H), 10.3 (br s, 1H). ^{13}C NMR ($CDCl_3$): δ 25.1, 26.7, 29.5, 29.7, 29.87, 29.96, 29.99, 30.00, 30.5, 34.5, 64.1, 86.7, 127.2, 128.1, 129.2, 145.0, 180.7. FABMS *m*/*z* 481 (M + Na^+, 5%), 243 (Tr^+, 100%). HRMS (FAB, NOBA/LiOAc) for $C_{31}H_{38}O_3Li$ [M + Li]$^+$: calcd 465.2971, found 465.2981.

2,2-Dimethyl-3-trityloxypropionic Acid (6). 2,2-Dimethyl-3-tritylhydroxypropionic acid (5.0 g, 42.3 mmol) and trityl chloride (23.6 g, 84.6 mmol, 2 equiv) were dried under high vacuum for 1 h. The reagents were dissolved in pyridine (80 mL), and the mixture was allowed to react for 48 h at room temperature with stirring. The solvent was evaporated, and the residue was taken up in CH_2Cl_2 (80 mL). The solution was washed with water (80 mL), 5% aqueous AcOH (2 × 80 mL), and brine (4 × 50 mL) and dried over Na_2SO_4. Evaporation produced a residue that was purified by two consecutive column chromatographies (silica, CH_2Cl_2, MeOH 0−20%) to afford the title compound as a white solid, 7.69 g (21.4 mmol, 51%). ^{1}H NMR ($CDCl_3$): δ 1.20 (s, 6H), 3.14 (s, 2H), 7.18−7.30 (m, 9H), 7.38−7.47 (m, 6H). ^{13}C NMR ($CDCl_3$): δ 23.0, 44.0, 70.0, 86.8, 127.4, 128.2, 129.2, 144.3, 183.2; MS (EI+) *m*/*z* 359 (M^+, 20%), 282 (M^+ − Phe, 53%), 243 (Tr + H^+, 100%). HRMS (FAB, NOBA/LiOAc) for $C_{24}H_{24}O_3Li$ [M + Li]$^+$: calcd 367.1879, found 367.1885.

Solid-Phase Synthesis, General Procedures. Coupling To Produce Amide or Ester Linkages (Step A). The following procedure for the coupling of 12-trityloxylauric acid (**5**) to aminopropyl cpg is representative. Aminopropyl cpg (35 mg, 4.5 μmol) was placed in a polypropylene reaction column for ABI DNA synthesizers (100 mg capacity). A mixture of 12-trityloxylauric acid **5** (45.9 mg, 100 μmol), HOBT (13.5 mg, 100 μmol), and HBTU (34.1 mg, 90 μmol) was dried at 0.1 Torr and dissolved in DMF (600 μL). This was treated with diisopropylethylamine (40 μL, 30.2 mg, 230 μmol), leading to a slight darkening of the solution. The coupling solution was immediately injected into the cpg-charged column with the aid of two syringes, such that all air was displaced from the reservoir. The coupling was allowed to proceed for 1 h at room temperature with mixing through syringe-induced flow every 5 min. The solution was pushed out of the column and the solid support was carefully rinsed with CH_3CN (10 mL), followed by drying at 0.1 Torr.

Removal of Trityl Protecting Groups (Step B). The following protocol for the detritylation of aminopropyl cpg derivatized with **5** (prepared according to the protocol in the preceding paragraph) is representative. The trityloxylauric acid-bearing glass support in the reaction column was treated with a mixture of CH_2Cl_2/TFA (1:1, 1 mL) for 2 min at room temperature, leading to the development of a yellow color. The reaction solution was pushed out of the column, and the solid support was washed with CH_3CN (10 mL), followed by drying at 0.1 Torr.

Removal of Fmoc Protecting Groups (Step C). This procedure for the Fmoc-deprotection of solid support **12** derivatized with Fmoc-Gly-OH is representative. The derivatized solid support (40 mg, ca. 0.7 μmol Fmoc groups) in a reaction column was treated with a mixture of piperidine/DMF (1:4, 1 mL) for 30 min at room temperature. The deprotection mixture was removed from the column, and the solid support was rinsed with CH_3CN (10 mL), followed by drying at 0.1 Torr.

Capping (Step D). The following representative protocol was employed for the capping of unreacted amino and hydroxyl groups on solid support **13** derivatized with tritylated linker **6**. The reaction column bearing the cpg with the terminal linker moiety was attached to an ABI model 381 DNA synthesizer, and the capping program for DNA synthesis cycles (as recommended by the manufacturer, system software version 1.5) was run four consecutive times. This program uses a mixture of acetic anhydride and 2,6-lutidine in THF ("cap A") and 1-methylimidazole in THF ("cap B") and includes washing steps with CH_3CN. After the reaction cycles, residual solvent was removed and the support was dried by exposing the column to reduced pressure (0.1 Torr).

Synthesis of Oligonucleotides (Step E). The following protocol was used for the preparation of glycine-containing immobilized hybrid **25** on solid support **18** and is representative for other syntheses. The derivatized controlled pore glass **18** (35 mg, 0.65 μmol terminal hydroxyl groups) in a reaction column was attached to an ABI DNA synthesizer, model 381. The synthesis cycles for a "trityl-off" oligonucleotide recommended by the manufacturer (system software 1.5) for assembly of 5′-TGCGCA-3′ on a 1.0 μmol scale was employed without modifications, using MMT-protected 5′-amino-5′-deoxythymidine phosphoramidite[6] for the thymidine channel. After DNA synthesis, the glass support was dried under reduced pressure.

Deprotection and Release from Support (Step F). Given below is a representative protocol applied to the deprotection of **29** and used analogously for other hybrids, unless otherwise stated. The hybrid-bearing controlled pore glass (35 mg, ca. 0.5 μmol hybrid) in a polypropylene cup was treated with ammonium hydroxide (30% NH_3, 1.5 mL) for 16 h. The supernatant was aspired, the residue was washed with water (0.5 mL), and excess NH_3 was removed from the combined solutions with an air stream. The resulting solution was lyophilized to dryness, taken up in water, and filtered (0.2 μm) for HPLC purification and/or MALDI analysis.

Solid Support 1. Native controlled pore glass (3 mg) was modified using the procedure of Southern and collaborators[15] and derivatized with a sequence of steps A (using Fmoc-Cys-(S-*t*Bu)-OH), D, C, and A (using Fmoc-Ser-OH).

H-TTTT-Ser-Cys(S-*t*Bu)-OH (2) and H-TGCGCA-Ser-Cys(S-*t*Bu)-OH (3). Solid support **1**, prepared as described in the preceding paragraph, was subjected to DNA synthesis (step E) using the 0.2 μmol scale protocol and deprotected with NH_4OH at room temperature (step F) or at 50 °C for 1.5 h or with NH_4OH/dioxane (1:1) for 16 h at 22 °C. (**2**) MALDI-TOF MS for $C_{50}H_{71}N_{10}O_{32}P_4S_2$ [M − H]$^-$: calcd 1512.2, found 1510. (**3**) MALDI-TOF MS for $C_{68}H_{93}N_{26}O_{39}P_6S_2$ [M − H]$^-$: calcd 2148.6, found 2146.

Solid Support 7. Aminopropyl controlled pore glass (15 mg, 1.95 μmol amino groups) was reacted with steps A (using **5**) and B, followed by coupling with Fmoc-Gly-OH (step A) and Fmoc removal (step C).

Solid Support 8. A sample of support **7** (5 mg, 0.6 μmol amino groups) was reacted with **4** using step A and deprotected with step B.

H-TGCGCA-GA-Gly-OH (9). Derivatized controlled pore glass **8**, prepared as described in the preceding paragraph, was subjected to DNA synthesis following protocols for steps E and F. MALDI-TOF MS for $C_{62}H_{80}N_{25}O_{39}P_6$ [M − H]$^-$: calcd 1985.3, found 1985.

H-TGCGCA-LA-Gly-OH (11). A sample of support **7** (5 mg, 0.6 μmol amino groups) was reacted with **5** using step A, followed by steps D and B. DNA synthesis and deprotection was performed with the protocol for steps E (manufacturer's protocol for 0.2 μmol scale) and F. MALDI-TOF MS for $C_{72}H_{100}N_{25}O_{39}P_6$ [M − H]$^-$: calcd 2125.6, found 2126.

H-TGCGCA-DP-Gly-OH (10). Compound **10** was prepared analogously to the synthesis of **11**, above, except that **6** was used in the first coupling reaction. Alternatively, **10** was prepared through deprotection of solid support **25**, using step F. MALDI-TOF MS for $C_{65}H_{86}N_{25}O_{39}P_6$ [M − H]$^-$: calcd 2027.4, found 2027.

Solid Support 12. Controlled pore glass (5 mg, 0.65 μmol amino groups) was derivatized to triester support **12** by performing three consecutive rounds of step A (with **5**) and step B, with step D prior to the final detritylation. This decreased the loading to 0.2 μmol (31%), as measured by trityl release.

Solid Supports 13−17, 18−22, 23, and 25−28. A sample of support **12** (0.5−0.7 μmol of terminal hydroxyl groups), prepared as described in the preceding paragraph, was twice subjected to step A with the respective Fmoc−amino acid building block, followed by step C and steps A (twice, with **6**) and B. DNA synthesis (step E) was performed with the DNA synthesis program for 1.0 μmol syntheses. Solid support **23** was prepared on a 4 μmol scale to provide sufficient material for NMR spectroscopy. Release of the dimethoxytrityl cation during DNA syntheses indicated between 81% and 90% overall yield (97−98.3% per step) for the chain assembly.

H-Aa-Gly-TGCGCA-DP-Gly-OH (24, 29−33). The amino-terminal glass supports, prepared as described in the preceding protocol, were subjected to step A with Fmoc-Gly-OH (repeated once), Fmoc removal with step C, and two consecutive couplings A with the appropriate Fmoc−amino acid building block. For **24**, this was performed at a 4-fold scale. For library **33**, a mixture of Fmoc-Ala-OH (3.5 mg, 11 μmol), Fmoc-Pro-OH (7.1 mg, 21 μmol), Fmoc-Asp(Bn)-OH (5.1 mg, 11 μmol), Fmoc-Met-OH (4.8 mg, 13 μmol), Fmoc-Phe-OH (3.5 mg, 9 μmol), Fmoc-Tyr-OH (3.9 mg, 10 μmol), and Fmoc-Trp-OH (4.5 mg, 10 μmol) was used with the activation agents as given in the general protocol for step A. The last coupling step was followed by deprotection according to step F. For **32**, allyloxycarbonyl groups were removed by exposure of **28** to PPh_3 (1.5 mg, 23 μmol), $Pd[PPh_3]_4$ (7.5 mg, 20.5 μmol), and diethylammonium hydrogen carbonate (7.5 mg, 56 μmol) in CH_2Cl_2 (0.7 mL) for 2.5 h at room temperature under light protection, followed by washing with CH_3CN (10 mL). Yields of hybrids were determined from the integration of the product peaks in the HPLC trace of the crude. For lipophilic sequences, no full recovery from RP_{18} columns was achieved, leading to lowered yields.

H-Gly-Gly-TGCGCA-DP-Gly-OH (29). Yield 84%. HPLC: CH_3CN gradient 0% for 5 min to 25% in 45 min, R_t = 39.2 min. MALDI-TOF MS for $C_{69}H_{92}N_{27}O_{41}P_6$ [M − H]$^-$: calcd 2142.5, found 2139.

H-Trp-Gly-TGCGCA-DP-Trp-OH (30). Yield 69%. HPLC: CH_3CN gradient 0% for 5 min to 25% in 45 min, R_t = 47.1/49.8 min. MALDI-TOF MS for $C_{87}H_{106}N_{29}O_{41}P_6$ [M − H]$^-$: calcd 2399.8, found 2402.

H-Asp-Gly-TGCGCA-DP-Asp-OH (31). Yield 93%. HPLC: CH_3CN gradient 0% for 5 min to 22% in 45 min, R_t = 36.2 min. MALDI-TOF MS for $C_{73}H_{96}N_{27}O_{45}P_6$ [M − H]$^-$: calcd 2257.6, found 2257.

H-Lys-Gly-TGCGCA-DP-Lys-OH (32). Yield 38%. HPLC: CH_3CN gradient 0% for 5 min to 25% in 50 min, R_t = 42.2 min. MALDI-TOF MS for $C_{77}H_{110}N_{29}O_{41}P_6$ [M − H]$^-$: calcd 2283.7, found 2281.

Library H-Aa-Gly-TGCGCA-DP-Trp-OH (33). Aa = Ala: MALDI-TOF MS for $C_{79}H_{101}N_{28}O_{41}P_6$ [M − H]$^-$ calcd 2284.7, found 2284. Aa = Pro: MALDI-TOF MS for $C_{81}H_{104}N_{28}O_{41}P_6$ [M − H]$^-$ calcd 2311.7, found 2310. Aa = Asp: MALDI-TOF MS for $C_{80}H_{101}N_{28}O_{43}P_6$ [M − H]$^-$ calcd 2328.7, found 2327. Aa = Met: MALDI-TOF MS for $C_{81}H_{105}N_{28}O_{41}P_6S$ [M − H]$^-$ calcd 2344.8, found 2343. Aa = Phe: MALDI-TOF MS for $C_{85}H_{105}N_{28}O_{41}P_6$ [M − H]$^-$ calcd 2360.8, found 2360. Aa = Tyr: MALDI-TOF MS for $C_{85}H_{105}N_{28}O_{42}P_6$ [M − H]$^-$ calcd 2376.8, found 2376. Aa = Trp (**30**), found 2399.

H-Phe-Gly-TGCGCA-DP-Phe-OH (24). Yield 41%. HPLC: CH_3CN gradient 0% for 5 min to 25% in 45 min, R_t = 47.0/49.7 min. MALDI-TOF MS for $C_{83}H_{104}N_{27}O_{41}P_6$ [M − H]$^-$: calcd 2321.7, found 2320. ^{1}H NMR (300 MHz, D_2O, 22 °C) δ 0.93 (s, 6H, CH_3−DP9), 1.59 (s, 3H, CH_3−T3), 1.61 (m, 1H, H2′−C7), [1.7−4.9 ppm range, selected resonances by residue, 2.70/2.76 (Hβ/Hβ′−Phe1), 3.70 (Hα−Phe1), 3.02 (Hα/Hα′−Gly2), 1.92/2.24 (H2′/H2″−T3), 3.22/3.33 (H5′/H5″−T3), 3.93 (H4′−T3), 4.49 (H3′−T3), 2.62/2.66 (H2′/H2″−G4), 3.93 (H5′/H5″−G4/G6), 4.26 (H4′−G4/G6), 4.85 (H3′−G4), 1.91/2.29 (H2′/H2″−C5), 3.907 (H5′/H5″−C5), 4.07 (H4′−C5), 4.74 (H3′−C5), 2.50/2.53 (H2′/H2″−G6), 4.84 (H3′−G6), 2.05 (H2″−C7), 3.90 (H4′−C7), 4.64 (H3′−C7), 2.66/2.48 (H2′/H2″−A8), 3.96 (H5′/H5″−A8), 4.21 (H4′−A8), 4.83 (H3′−A8), 2.66/2.78 (Hβ/Hβ′−Phe10), 4.34 (Hα−Phe10)], 5.28 (d, J = 7 Hz, 1H, H5−C5), 5.30 (d, J = 6.8 Hz, 1H, H5−C7), 5.59 (m, 1H, H1′−C7), 5.62 (m, 1H, H1′−C5), 5.65 (m, 1H, H1′−T3), 5.73 (t, J = 6.7 Hz, 1H, H1′−G6), 5.85 (dd, J = 5.9 Hz, 9.2 Hz, 1H, H1′−G4), 6.17 (t, J = 6.7 Hz, 1H, H1′−A8), 6.82−7.06 (m, 10H, aromatic H−Phe1/Phe10), 7.09 (d, J = 7.4 Hz, 1H, H6−C7), 7.15 (s, 1H, H6−T3), 7.26 (d, J = 7.4 Hz, 1H, H6−C5), 7.63 (s, 1H, H2−A8), 7.78 (s, 1H, H8−G6), 7.89 (s, 1H, H8−G4), 8.15 (s, 1H, H8−A8); exchangeable protons observed in H_2O/D_2O (88:12) at 22 °C δ 6.40 (br s, 1H, H42−C5), 6.54 (br s, 1H, H42−C7), 7.45 (d, J = 7.5 Hz, 1H, NHα−Phe10), 7.76 (NH5′−T3), 8.30 (br s, 1H, H41−C7), 8.34 (br s, 1H, H41−C5), 12.92 (br s, 1H, NH1−G4); exchangeable proton observed in H_2O/D_2O (88:12) and 100 mM NaCl at 7 °C δ 8.09 (NHα−Gly2), 12.84 (br s, 1H, NH1−G6).

H-Phe-TGGTTGAC-DP-Phe-OH (35). Starting from support **22** (0.1 μmol free hydroxyl groups), **34** was prepared with the DNA synthesis program for 0.2 μmol scale analogously to the preparation of **25**−**28**. A single coupling step (A) with Fmoc-Phe-OH was followed by treating 20% of the protected hybrid with NH_4OH according to protocol **F**. MALDI-TOF MS for $C_{102}H_{127}N_{30}O_{54}P_8$ [M − H]$^-$: calcd 2902.1, found 2906.

Nuclease Stability Assay. The exonuclease stability assay was performed similarly to those described previously.[22,20] Specifically, a solution of **35** (300 pmol) in water (1.3 μL) was mixed with a solution of the unmodified DNA control octamer 5′-TGGTTGAC-3′ (300 pmol) in water (1.7 μL), followed by addition of additional water (7 μL) and 0.5 M NH_4OAc buffer pH 6.0 (1.8 μL). The sample was annealed at 55 °C, and the nuclease reaction was performed at 37 °C with phosphodiesterase I, type VII (EC 3.1.4.1, 0.5 units). The unmodified DNA decamer 5′-CGCATTAGCA-3′ was used as internal standard. Data acquisition[16] and analysis[22,20] were performed as previously described.

Acknowledgment. The authors are grateful to David Sarracino for help with the acquisition of MALDI spectra, Charles Tetzlaff for computer support and help with DNA syntheses, and Dr. Dwayne Daniels and Robert Altman for helpful discussions. I.S. was supported by a postdoctoral fellowship from the Deutsche

Forschungsgemeinschaft, FRG, under the Post-Doc-Stipendien program. This work was supported by the NIH, National Institute of General Medical Science (Grant GM54783 to C.R.). The NMR facility at the Chemistry Department of Tufts University is supported by NSF Grant CHE-9723772 (to C.R. and M. d'Alarcao).

Supporting Information Available: Carbon and proton NMR spectra of **4**, **5**, and **6**, MALDI-TOF mass spectra of **9** and the deprotection product of **28**, and an HPLC trace of crude **31**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO990036D