Supporting Information

Optimized reaction conditions for amide bond formation in DNA-encoded combinatorial libraries

Yizhou Li[1], Elena Gabriele[1], Florent Samain[2], Nicholas Favalli[1], Filippo Sladojevich[2], Jörg Scheuermann[1], and Dario Neri[1]*

[1] Department of Chemistry and Applied Biosciences, Swiss Federal Institute of Technology (ETH Zürich), Vladimir-Prelog-Weg 4, CH-8093 Zürich (Switzerland)

[2] Philochem AG, Libernstrasse 3, CH-8112 Otelfingen (Switzerland)

Table of Contents

1. General Procedures.

Materials and Instruments. Unless otherwise noted, all reagents and solvents were purchased from commercial sources and used as received. Water was purified with a Thermo Scientific Barnstead Nanopure system. Oligonucleotides were purchased from DNA Technology (Denmark). Carboxylic acid building blocks were purchased from several commercial suppliers including ABCR (Karlsruhe, Germany), ChemBridge (SanDiego, CA), Sigma-Aldrich (St. Louis, MO), TCI Europe (Zwijndrecht, Belgium), Alfa Aesar (Ward Hill, MA), Matrix Scientific (Columbia, SC), and Acros Organics (Geel, Belgium).
Universal indicator paper (pH 1-11) was purchased from Macherey-Nagel (Germany).
For UPLC-MS analysis, an X-Bridge Oligonucleotide BEH C18 2.1 x 50 mm column, using an A= TEA (10 mM), HFIP (5 mM) in water to B= methanol, gradient from 10% B to 40% B in 6.5 minutes. Xevo G2-XS Q-TOF with electrospray ionization source was used for detection. MaxEnt 1 software was used to deconvolute the multiple charge states.
HPLC purifications were performed on a CT18-XTerra 10 x 150 mm column, using an A= TEAA (0.1 M in water) to B= CH_3CN (80% in water) gradient from 10% B to 40% B in 15 minutes.

Ethanol Precipitation. To aqueous DNA solutions, 10 % by volume of 5 M NaCl was added, followed by 2.5-3 volumes of cold ethanol. The colloidal solution was allowed to sit for 16-18 h at −20 °C, prior to a centrifugation step at 17,000 g for 15 minutes at 4 °C. The resulting supernatant was discarded and the pellet was rinsed once with cold 70% ethanol. After centrifugation at 17, 000 g for another 5 minutes at 4 °C, the supernatant was discarded and the pellet was dried on a speedvac. The recovered samples were dissolved in appropriate buffer for subsequent analysis or experiments. This procedure was generally performed after each chemical reaction.

General Protocol for Functionalization of Amino-modified DNAs with Carboxylic Acids by the EDC/HOAt/DIPEA method. To a solution of DNA (500 pmol) in MOPS buffer (50 mM, pH 8.0, 0.5 M NaCl, 72 μL) was added a mixture of carboxylic acid (60 mM, 45 μL), EDC (300 mM, 4 μL), HOAt (60 mM, 4 μL) and DIPEA (300 mM, 4 μL) in DMSO, previously activated for 15 minutes at room temperature. The reaction was agitated at room temperature for 16 h. The reaction solution was then treated with a second addition of freshly activated carboxylic acid in DMSO (same activation mixture as above) and it was agitated for further 6 h at room temperature. The reaction was quenched by addition of NH_4OAc (500 mM, 25 μL) at room temperature for 30 minutes and the DNA-conjugates were isolated by ethanol precipitation. The pellet was re-dissolved in water (200 μL) and analyzed by UPLC-MS.

- **pH values at different stages of the EDC/HOAt/DIPEA method:**

	MOPS Buffer (50 mM, pH 8.0, 0.5 M NaCl)			MOPS Buffer (500 mM, pH 8.0, 0.5 M NaCl)			No Buffer (pH 7.0, 0.5 M NaCl)		
Carboxylic Acid	A	B	C	A	B	C	A	B	C
CA-1	8	8	8	8	8	8	7	7	7
CA-2	8	8	8	8	8	8	7	7	7
CA-3	8	8	8	8	8	8	7	7	7
CA-4	8	8	8	8	8	8	7	7	7
CA-5	8	8	8	8	8	8	7	7	7
CA-6	8	8	8	8	8	8	7	7	7
CA-7	8	8	8	8	8	8	8	8	8
CA-8	8	8	8	8	8	8	7	7	7

A: pH value at 5 minutes after the first time addition of activation mixture. **B**: pH value at 5 minutes after the second time addition of activation mixture. **C**: pH value before the quenching of reaction mixture. All pH values were measured by universal indicator paper (pH 1-11).

- **Reaction conversions with different buffer systems:**

Carboxylic Acid	MOPS Buffer (50 mM, pH 8.0, 0.5 M NaCl)	MOPS Buffer (500 mM, pH 8.0, 0.5 M NaCl)	No Buffer (pH 7.0, 0.5 M NaCl)
CA-1	90%	90%	90%
CA-2	90%	90%	90%
CA-3	90%	90%	90%
CA-4	90%	90%	0%
CA-5	90%	90%	60%
CA-6	90%	90%	90%
CA-7	90%	90%	90%
CA-8	90%	90%	90%

2. Oligonucleotide Sequences.

- In Figure 2, **ODN-1**: 5'-(NH_2-C_{12})-GGA GCT TGT GAA TTC TGG ATC TTA GGA CGT GTG TGA ATT GTC;
- In Figure 5, **ODN-2**: 5'-(NH_2-C_6)-GGA GCT TGT GAA TTC TGG ATC TTA GGA CGT GTG TGA ATT GTC;

  ODN-3: 5'-(NH_2-C_3)-GGA GCT TGT GAA TTC TGG ATC TTA GGA CGT GTG TGA ATT GTC;

  ODN-4: 5'-(NH_2-PEG)-GGA GCT TGT GAA TTC TGG ATC TTA GGA CGT GTG TGA ATT GTC;

  ODN-5: GGA GCT TGT GAA TTC TGG ATC TTA GGA CGT GTG TGA ATT GTC-(C_6-NH_2)-3';

  ODN-6: 5'- CGT CGA TCC GGC GCC AT*G GGA CTC G; T* =Amino-C6-dT

- **ESI-MS characterization of DNAs used in this section:**

Oligonucleotide	Expected mass (Da)	Observed mass (Da)
ODN-1	13324.9	13325.4
ODN-2	13240.8	13240.9
ODN-3	13198.6	13199.0
ODN-4	13229.3	13228.8
ODN-5	13240.8	13241.0
ODN-6	7814.2	7814.5

3. Representative Mass Spectrum of DNA-Conjugates Characterization.

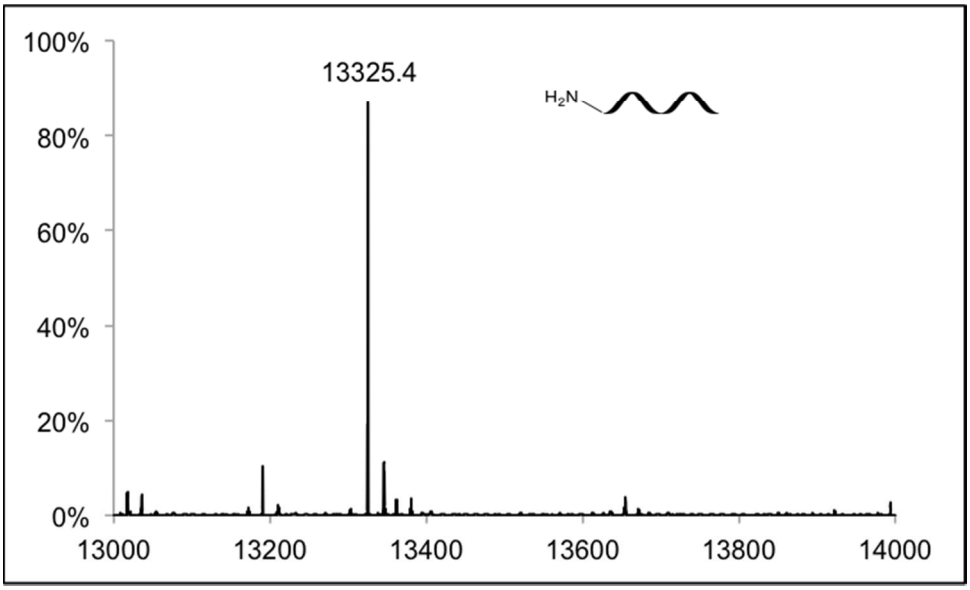

Figure S1. Deconvoluted mass spectrum of DNA-conjugate at 4.36-minute in figure 2a. Observed: 13325.4; expected: 13324.9.

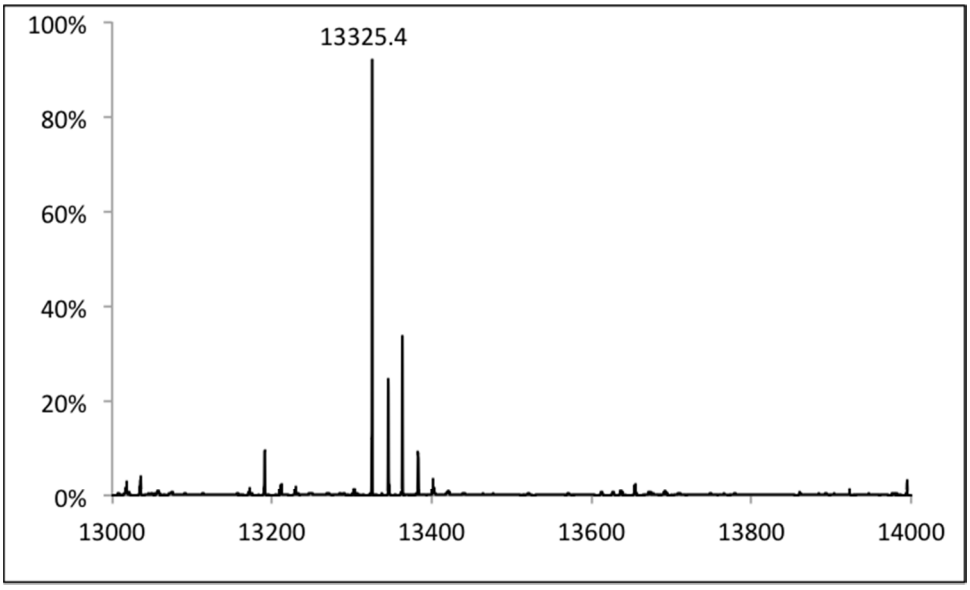

Figure S2. Deconvoluted mass spectrum of DNA-conjugate at 4.35-minute peak in figure 2b. Observed: 13325.4; expected: 13324.9.

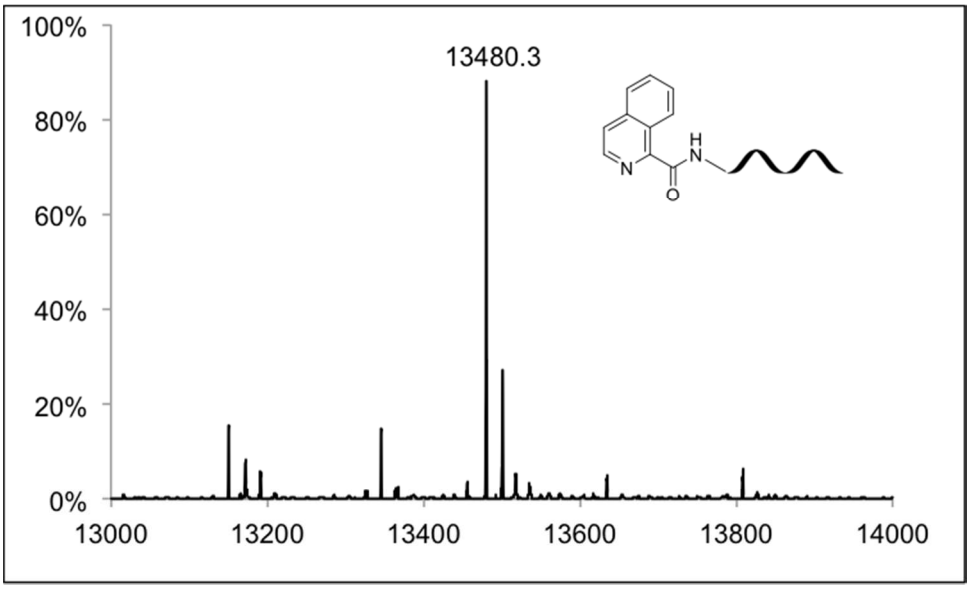

Figure S3. Deconvoluted mass spectrum of DNA-conjugate at 5.03 minute peak in figure 2c. Observed: 13480.3; expected: 13480.5.

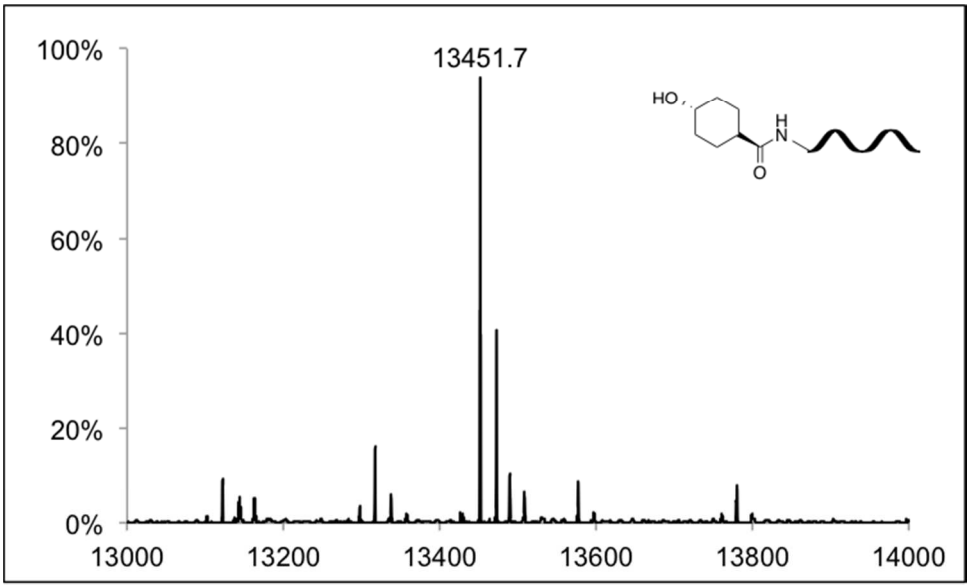

Figure S4. Deconvoluted mass spectrum of DNA-conjugate yielded from CA-9 (trans-4-hydroxycyclohexane-1-carboxylic acid) and ODN-1 by the DEC/HOAt/DIPEA method. Observed: 13451.7; expected: 13451.2. There is no detectable EDC adduct (expected: 13605.0) was observed for CA-9, which gives significant EDC adduct alongside the expected DNA-conjugate by the pseudo-solid phase method. [1]

4. DNA-Conjugates Characterization in Figure 1.

Carboxylic Acid	Expected mass (Da)	Observed mass (Da)
CA-1	13448.8	13449.0
CA-2	13521.3	13521.0
CA-3	13534.9	13535.0
CA-4	13488.8	13489.0
CA-5	13480.3	13480.5
CA-6	13511.8	13512.0
CA-7	13515.9	13516.0
CA-8	13511.9	13512.8
CA-9	13451.2	13451.7

5. DNA-Scaffold Conjugates' Structures, Synthesis, Purification, and Characterization.

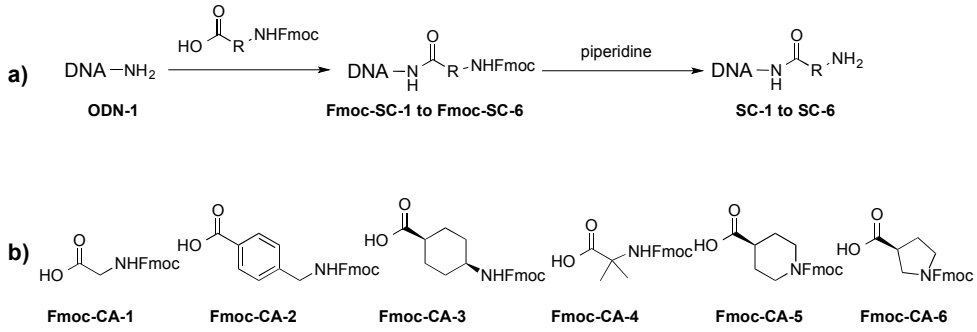

Figure S5. a) Synthesis scheme of DNA-scaffold conjugates **SC-1** to **SC-6**. b) Structures of the Fmoc-protected amino acids used for conjugation.

Fmoc-protected amino acid (100 mM in DMSO, 50 µL), *sulfo*-NHS (100 mM stock solution in 2:1 DMSO/H_2O, 20 µL) and EDC*HCl (100 mM in DMSO, 50 µL) were added to DMSO (100 µL) and allowed to stand at room temperature for 15 minutes. Subsequently, a mixture of the amino-modified oligonucleotide (**ODN-1**, 50 nmol) dissolved in TEA*HCl buffer (250mM, pH 10.0, 100 µL) was added and the reaction kept at room temperature for additional 8 h. The reaction was precipitated with ethanol. To the resulting pellet, piperidine was added (10% v/v aq stock solution, 100 µL) and the de-protection reaction was allowed to stand at room temperature for 30 minutes. The DNA-scaffold conjugate was precipitated with ethanol and purified by HPLC. The separated and collected conjugate was vacuum-dried overnight, re-dissolved in H_2O and characterized by ESI-MS.

- **ESI-MS characterization of DNA-conjugates used in this section:**

SC	Yield (%)	Expected mass (Da)	Observed mass (Da)
SC-1	53	13381.6	13382.4
SC-2	18	13457.0	13457.9
SC-3	34	13449.8	13450.6
SC-4	38	13410.4	13411.7
SC-5	57	13436.1	13436.9
SC-6	80	13421.9	13422.5

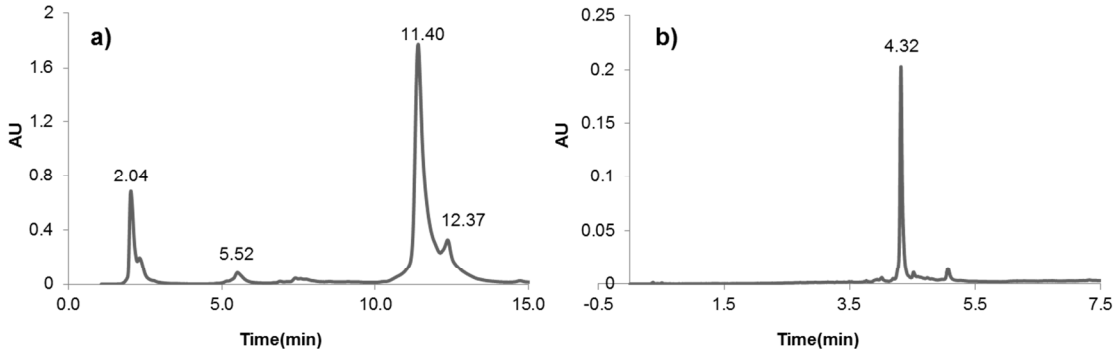

Figure S6. a) HPLC trace of the purification of DNA-scaffold conjugate **SC-3**.The fraction at 11.40 min was collected. b) UPLC profile of the DNA-scaffold conjugate **SC-3**. Detection at 260 nm.

6. DNA-Conjugates Characterization in Figure 4 and Figure 5.

- **ESI-MS characterization of DNA-conjugates used in Figure 4:**

SC	CA	Expected mass (Da)	Observed mass (Da)
SC-1	**CA-1**	13505.8	13506.8
SC-1	**CA-2**	13578.2	13579.4
SC-1	**CA-3**	13591.8	13593.1
SC-1	**CA-4**	13545.7	13546.7
SC-1	**CA-5**	13536.7	13537.8
SC-1	**CA-6**	13568.0	13569.7
SC-1	**CA-7**	13572.0	13573.9
SC-1	**CA-8**	13568.9	13570.0
SC-2	**CA-1**	13582.0	13583.0
SC-2	**CA-2**	13654.4	13655.3
SC-2	**CA-3**	13668.0	13669.1
SC-2	**CA-4**	13622.0	13622.9
SC-2	**CA-5**	13612.9	13613.9
SC-2	**CA-6**	13644.8	13645.7
SC-2	**CA-7**	13648.8	13650.1
SC-2	**CA-8**	13645.0	13646.2
SC-3	**CA-1**	13574.0	13575.2
SC-3	**CA-2**	13646.4	13647.7
SC-3	**CA-3**	13660.0	13661.2
SC-3	**CA-4**	13614.0	13615.0
SC-3	**CA-5**	13605.0	13606.2
SC-3	**CA-6**	13636.8	13637.9
SC-3	**CA-7**	13640.8	13642.4
SC-3	**CA-8**	13637.1	13638.3
SC-4	**CA-1**	13533.7	13534.9
SC-4	**CA-2**	13606.1	13607.3

SC-4	**CA-3**	13620.2	nd
SC-4	**CA-4**	13573.2	nd
SC-4	**CA-5**	13565.1	13566.2
SC-4	**CA-6**	13596.3	nd
SC-4	**CA-7**	13600.8	nd
SC-4	**CA-8**	13596.7	13598.1
SC-5	**CA-1**	13559.8	13560.5
SC-5	**CA-2**	13631.3	13632.7
SC-5	**CA-3**	13645.1	13646.3
SC-5	**CA-4**	13488.8	13600.3
SC-5	**CA-5**	13590.1	13591.5
SC-5	**CA-6**	13622.7	13623.7
SC-5	**CA-7**	13627.7	13628.1
SC-5	**CA-8**	13622.1	13623.5
SC-6	**CA-1**	13546.1	13547.1
SC-6	**CA-2**	13618.6	13618.0
SC-6	**CA-3**	13632.2	13633.1
SC-6	**CA-4**	13586.1	13587.1
SC-6	**CA-5**	13577.1	13578.3
SC-6	**CA-6**	13608.9	13609.7
SC-6	**CA-7**	13612.9	13614.1
SC-6	**CA-8**	13609.2	13610.1

nd: Masses were determined only for conversion >10%, because of limited sensitivity of the instrument.

- **ESI-MS characterization of DNA-conjugates used in Figure 5:**

ODN	CA	Expected mass (Da)	Observed mass (Da)
ODN-2	**CA-1**	13365.0	13365.0
ODN-2	**CA-2**	13437.5	13437.0
ODN-2	**CA-3**	13451.1	13451.0
ODN-2	**CA-4**	13405.0	13405.0
ODN-2	**CA-5**	13396.0	13396.0
ODN-2	**CA-6**	13427.8	13428.4
ODN-2	**CA-7**	13431.8	13432.7
ODN-2	**CA-8**	13428.1	13428.0
ODN-3	**CA-1**	13320.8	13322.9
ODN-3	**CA-2**	13435.6	13437.0
ODN-3	**CA-3**	13408.1	13409.1
ODN-3	**CA-4**	13362.3	13363.0
ODN-3	**CA-5**	13394.7	13396.0
ODN-3	**CA-6**	13385.1	13386.6
ODN-3	**CA-7**	13389.2	13390.7
ODN-3	**CA-8**	13384.8	13386.1
ODN-4	**CA-1**	13353.5	13353.2
ODN-4	**CA-2**	13426.0	13425.7
ODN-4	**CA-3**	13439.6	13439.2
ODN-4	**CA-4**	13393.5	13393.6
ODN-4	**CA-5**	13384.5	13384.4
ODN-4	**CA-6**	13416.3	13416.5
ODN-4	**CA-7**	13420.3	13420.8
ODN-4	**CA-8**	13416.6	13416.4
ODN-5	**CA-1**	13365.0	13365.2
ODN-5	**CA-2**	13437.5	13437.6

ODN-5	**CA-3**	13451.1	13451.2
ODN-5	**CA-4**	13405.0	134505.2
ODN-5	**CA-5**	13396.0	13396.3
ODN-5	**CA-6**	13427.8	13428.6
ODN-5	**CA-7**	13431.8	13432.8
ODN-5	**CA-8**	13428.1	13428.2
ODN-6	**CA-1**	7938.4	7938.6
ODN-6	**CA-2**	8010.9	8010.5
ODN-6	**CA-3**	8024.5	8024.6
ODN-6	**CA-4**	7978.4	7978.6
ODN-6	**CA-5**	7969.4	7969.7
ODN-6	**CA-6**	8001.2	8002.1
ODN-6	**CA-7**	8005.2	8006.2
ODN-6	**CA-8**	8001.5	8001.7

7. Reference

(1) Franzini, R. M., Samain, F., Abd Elrahman, M., Mikutis, G., Nauer, A., Zimmermann, M., Scheuermann, J., Hall, J., and Neri, D. (2014) Systematic Evaluation and Optimization of Modification Reactions of Oligonucleotides with Amines and Carboxylic Acids for the Synthesis of DNA-Encoded Chemical Libraries. *Bioconjug. Chem. 25*, 1453-1461.