Supplementary Information

Sequence-specific Post-Synthetic Oligonucleotide Labeling for Single-Molecule Fluorescence Applications

David Egloff[1], Igor A. Oleinich[1], Meng Zhao, Sebastian L. B. König, Roland K. O. Sigel, Eva Freisinger[*]

Department of Chemistry, University of Zurich, Winterthurerstrasse 190

CH-8057 Zurich, Switzerland

[1]*DE and IO contributed equally*

Table of contents:

I. List of abbreviations

A	adenine	NTP	nucleotide triphosphate
AcOH	acetic acid	PAGE	polyacrylamide gel electrophoresis
APS	ammonium persulfate	PDS	2,2'-dipyridyl disulfide
CTAB	cetyltrimethylammonium bromide	PNK	polynucleotide kinase
DCC	*N*,*N*'-dicyclohexylcarbodiimide	PPh_3	triphenylphosphine
DMAP	4-dimethylaminopyridine	PSA	polymerase stop assay
DMF	dimethylformamide	Rf	retention factor
DMSO	dimethylsulfoxide	RNA	ribonucleic acid
DNA	deoxyribonucleic acid	RP-HPLC	reverse phase high performance liquid chromatography
ε	etheno	*rt*	retention time
εA	etheno adenine	Taq	*Thermophilus aquaticus*
EtOAc	ethyl acetate	TBE	TRIS/borate/EDTA buffer
EtOH	ethanol	TBTA	tris(benzyltriazolylmethyl)amine
MALDI-TOF	matrix-assisted laser desorption/ionization - time of flight	TEAA	triethylammonium acetate
MeCN	acetonitrile	TEMED	tetramethylenediamine
MeOH	methanol	THF	tetrahydrofurane
NaOAc	sodium acetate	TLC	thin layer chromatography
NEt_3	triethylamine	TRIS	tris(hydroxymethyl)aminomethane
NHS	*N*-hydroxysuccinimide	UV/Vis	ultra violet/visible light
NMR	nuclear magnetic resonance		

II. General methods

Unless stated otherwise, starting materials were obtained in the highest commercial grades and used without further purification. Solvents for organic synthesis were distilled prior to use and, if necessary, dried under standard conditions. HPLC solvents were obtained in HPLC grade from Roth AG (Arlesheim, Switzerland) and Sigma-Aldrich (Buchs, Switzerland), deuterated solvents were purchased from Armar AG (Döttingen, Switzerland). Thin layer chromatography (TLC) was performed with Merck TLC Silicagel 60 F254 plates and visualized with UV light or potassium permanganate staining solution. Flash chromatography was carried out using high-purity grade silica gel (pore size 60 Å, 220-440 mesh particle size, 35-75 μm particle size, Sigma-Aldrich). H_2O used for all manipulations with oligonucleotides was purified with a TKA GenPure dispenser and water as well as buffer solutions for enzymatic reactions were sterilized either by autoclaving or sterile filtration. Unmodified and 3'-phosphorylated oligonucleotides were obtained from Microsynth AG (Balgach, Switzerland). Enzymes for radioactive labeling and polymerase stop assays were purchased from Promega AG (Dübendorf, Switzerland) (Calf Intestinal Alkaline Phosphatase, GoTaq DNA polymerase with 5X Colorless GoTaq Reaction Buffer) and New England Biolabs (Ipswich, MA) (T4 Polynucleotide Kinase). [γ-^{32}P]-ATP for radioactive labeling of primer **P1** was obtained from Perkin Elmer AG (Oftringen, Switzerland) at a concentration of 150 mCi/mL. AccuGel 29:1 (40 %) and 10 X TBE (0.89 M Tris borate pH = 8.3, 20 mM Na_2EDTA) for polyacrylamide gel electrophoresis (PAGE) was purchased from National Diagnostics (Chemie Brunschwig AG, Basel, Switzerland). TEMED and APS were obtained from Sigma-Aldrich and Roth AG, respectively.

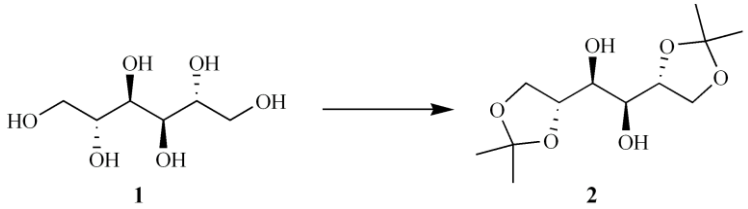
^{1}H and ^{13}C NMR spectra were recorded on a Bruker ARX-300 or AV-400 spectrometer. Chemical shifts in ^{1}H and ^{13}C NMR spectra, respectively, were reported in parts per million (ppm) on the δ scale from an internal standard of residual deuterated solvent. Data for ^{1}H NMR were reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad), coupling constant in Hertz (Hz) and integration. ESI mass spectra were recorded on a Bruker Esquire 6000 spectrometer. MALDI-TOF analyses were performed on an Autoflex time-of-flight mass spectrometer with a 337 nm nitrogen laser. All MALDI-MS spectra were recorded in negative ion mode using a 3-HPA matrix and calculated masses correspond to the molecular weight deducting a proton ($[MW-H]^-$). Reactions with oligonucleotides were performed using an Eppendorf Thermomixer Comfort and oligonucleotide concentrations were determined using a Thermo Scientific Nanodrop 2000 Spectrophotometer. Fluorescence measurements were carried out on a Varian Cary Eclipse Fluorescence Spectrophotometer using the software Cary Eclipse Scan Application in combination with a

thermostat Varian Cary Water Peltier System PCV 150 to maintain the temperature at 25°C. For HPLC a Dionex UltiMate 3000 HPLC system with a Waters X-Bridge C18 column (3.5 µm, 4.6 x 150 mm) and the software Chromeleon 7.1 were used. Visualization and quantification of radioactive gels were conducted on a Typhoon Scanner FLA 9500 with ImageQuant TL 1D (GE Healthcare AG, Glattbrugg, Switzerland).

III. Synthesis of the reactive strand RS1

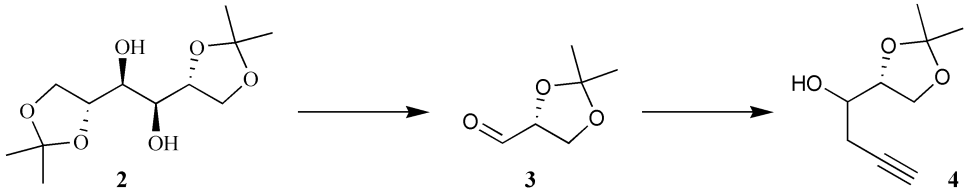
a. Organic synthesis of the reactive group

1,2-Bis-(2,2-dimethyl-[1,3]dioxolan-4-yl)-ethane-1,2-diol (2)[1]



$SnCl_2$ (10 mg, 0.05 mmol) was added to a solution of D-mannitol (5.0 g, 27 mmol) and 2,2-dimethoxypropane (8 mL, 65 mmol) in 12 mL freshly distilled dimethoxyethane under nitrogen atmosphere. The mixture was heated under reflux for 75 min to obtain a clear solution before cooling down below the reflux temperature and adding pyridine (10 µL). As soon as the solution had reached room temperature, the solvent was evaporated and 50 mL CH_2Cl_2 were added. The mixture was heated to 40 °C for 10 min, filtered hot and the filtrate was concentrated *in vacuo*. The white residue was recrystallized from di-n-butylether to obtain white crystals of **2** (2.6 g, 10 mmol, 36 % yield). TLC (CH_2Cl_2:MeOH, 95:5 v/v): Rf = 0.29; ^{1}H NMR (300 MHz, $CDCl_3$): δ 4.23-4.10 (m, 4H), 4.00-3.95 (dd, *J* = 8.4, 5.6 Hz, 2H), 3.75 (d, *J* = 6.3 Hz, 2H), 2.56 (s, *br*, 2H), 1.42 (s, 6H), 1.36 (s, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 109.5, 76.5, 71.4, 66.9, 26.9, 25.3; ESI-MS (*m/z*): $[M+Na]^+$ calcd. for $C_{12}H_{22}O_6$, 285.1; found, 285.0.

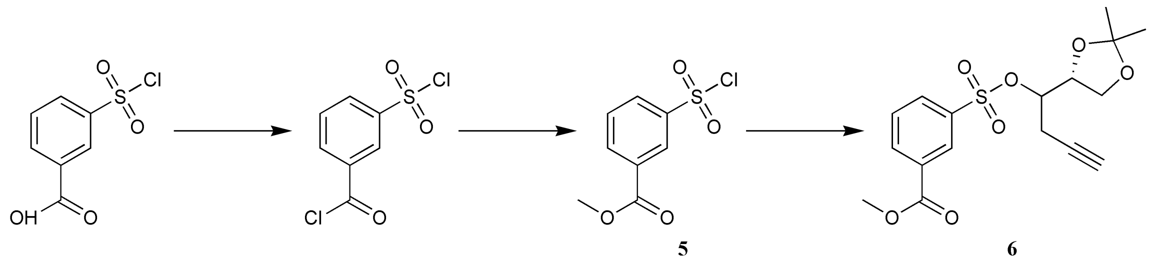
(2R)-1,2-O-Isopropylidenehex-5-yne-1,2,3-triol (4)[2]



A saturated $NaHCO_3$ solution in water (371 µL) was added to a solution of **2** (1.0 g, 3.9 mmol) in 10 mL CH_2Cl_2 and cooled in an ice bath. To this suspension $NaIO_4$ (1.2 g, 5.5 mmol) was added portion wise over a period of 25 min while maintaining the temperature below 35 °C. The resulting mixture was stirred for another 30 min in an ice bath and afterwards at room temperature. After 3 h TLC indicated complete consumption of **2** and $MgSO_4$ (356 mg, 3.0 mmol) was added and stirring continued for 30 min. The suspension was filtered, the filter cake suspended in 10 mL CH_2Cl_2 and filtered again. Finally, the combined filtrates were concentrated under reduced pressure (25 °C, 133 mbar) to yield crude glyceraldehyde **3**.

Propargyl bromide (1.72 mL, 80 % w/v in toluene) and Zn powder (1.04 g) were added to a solution of **3** in 12 mL THF and cooled to 0 °C. 3 mL sat. aq. NH_4Cl was added portion wise over a period of 30 min and stirring was continued at 0 °C for 1 h. Subsequently, the reaction mixture was allowed to warm to room temperature overnight. The reaction was stopped after 16 h by filtration. The filtrate was treated with 2 mL 5 % v/v HCl and the filter cake was washed with $CHCl_3$ (5 mL) and filtered. The two filtrates were combined and extracted with $CHCl_3$ (50 mL). The organic layer was washed with sat. $NaHCO_3$ (30 mL) and brine (30 mL), dried with $MgSO_4$ and concentrated *in vacuo*. Purification by flash chromatography (hexane:EtOAc, 4:1) affords a colourless oil of a diastereomeric mixture of **4** (890 mg, 5.2 mmol, 67 % yield rel. to **2**). TLC (hexane:EtOAc, 2:1 v/v): Rf = 0.47; ^{1}H NMR (400 MHz, $CDCl_3$, two isomers: erythro (E, major) and threo (T, minor)): δ 4.11-4.06 (m, 4H, E/T), 4.00-3.95 (m, 2H, E/T), 2.56-2.42 (m, 4H, E/T), 2.07 (t, *J* = 2.7 Hz, 1H, E), 2.05 (t, *J* = 2.7 Hz, 1H, T), 2.01 (s, *br,* 2H, E/T), 1.44 (s, 3H, T), 1.41 (s, 3H, E), 1.38 (s, 3H, T), 1.36 (s, 3H, E); ^{13}C NMR (100 MHz, $CDCl_3$, two isomers: erythro (E, major) and threo (T, minor)): δ 109.7 (E), 109.6 (T), 80.1 (E), 79.9 (T), 77.4 (E), 77.4 (T), 71,.4 (T), 70.9 (E), 70.4 (E), 70.3 (T), 66.2 (E), 66.0 (T), 26.8 (T), 26.7 (E), 25.3 (T), 25.3 (E), 24.0 (E), 23.7 (T); ESI-MS (*m/z*): $[M+Na]^+$ calcd. for $C_9H_{14}O_3$, 193.1; found, 193.0.

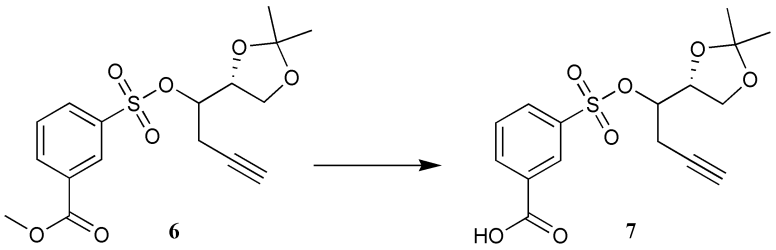
3-[1-(2,2-Dimethyl-[1,3]dioxolan-4-yl)-but-3-ynyloxysulfonyl]-benzoic acid methyl ester (6)



$SOCl_2$ (15 mL) was combined with 3-chlorosulfonylbenzoic acid (1.6 g, 7.3 mmol) and DMF (10 drops) under nitrogen atmosphere and heated under reflux for 4 h. As soon as the solution was cooled down to room temperature, it was concentrated *in vacuo*, twice co-evaporated with toluene to remove leftover $SOCl_2$ and vigorously dried under high vacuum to yield the crude acyl chloride. Then it was combined with 15 mL MeOH and stirred at room temperature. After 15 min, MeOH was removed under reduced pressure and the oily leftover was twice co-evaporated with CH_2Cl_2 to yield a pale yellow oil of crude **5**. **5** was dissolved in 25 mL anhydrous CH_2Cl_2, combined with **4** (623 mg, 3.7 mmol) and DMAP (50 mg, 0.4 mmol) and cooled to 0 °C under nitrogen atmosphere. After the addition of NEt_3 (2.5 mL), the solution was allowed to warm to room temperature overnight. After 16 h, when TLC indicated complete consumption of **4**, 10 mL CH_2Cl_2 were added and the solution was washed with H_2O (2 x 60

mL) and brine (30 mL). The organic layer was dried (Na_2SO_4), concentrated *in vacuo* and purified by flash chromatography (hexane:EtOAc, 6:1) to afford **6** (833 mg, 2.26 mmol, 62 % yield rel. to **4**). TLC (hexane:EtOAc, 2:1 v/v): Rf = 0.66; ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.60 (t, *J* = 1.6 Hz, 1H), 8.33 - 8.30 (dt, *J* = 7.9, 1.4 Hz, 1H), 8.14 - 8.11 (ddd, *J* = 7.9, 1.9, 1.2 Hz, 1H), 7.65 (t, *J* = 7.9 Hz, 1H), 4.68 - 4.64 (dt, *J* = 6.6, 4.9 Hz, 1H), 4.33 - 4.29 (td, *J* = 6.4, 5.4 Hz, 1H), 4.07 - 4.03 (dd, *J* = 8.9, 6.4 Hz, 1H), 3.96 (s, 3H), 3.90 - 3.86 (dd, *J* = 8.9, 5.4 Hz, 1H), 2.70 - 2.58 (m, 2H), 1.91 (t, *J* = 2.7 Hz, 1H), 1.32 (s, 3H) 1.30 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 165.2, 137.7, 134.9, 132.0, 131.6, 129.6, 129.2, 110.2, 79.5, 77.7, 74.9, 71.7, 65.9, 52.8, 26.6, 25.1, 21.8; ESI-MS (*m/z*): $[M+Na]^+$ calcd. for $C_{17}H_{20}O_7S$, 391.1; found, 391.0.

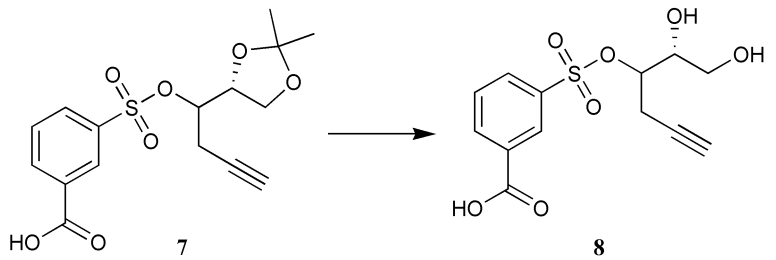
3-[1-(2,2-Dimethyl-[1,3]dioxolan-4-yl)-but-3-ynyloxysulfonyl]-benzoic acid (7)



6 **7**

3 mL 0.8 M $LiOH_*H_2O$ was added to a solution of **6** (532 mg, 1.4 mmol) in 3 mL THF and stirred at room temperature. After 80 min, TLC indicated complete consumption of **6** and the reaction was stopped by adding 5 % v/v HCl until a pH of 3-4 was reached. The solvent was removed by evaporation and the leftover residue was dissolved in 25 mL 1 M HCl. This aqueous solution was extracted with CH_2Cl_2 (3 x 15 mL), dried (Na_2SO_4) and concentrated *in vacuo* to obtain a viscous, colorless oil of **7** (439 mg, 1.2 mmol, 86 % yield). TLC (hexane:EtOAc:AcOH, 50:50:1 v/v): Rf = 0.56; ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.69 (t, *J* = 1.7 Hz, 1H), 8.40 - 8.38 (dt, *J* = 7.9, 1.4 Hz, 1H), 8.21 - 8.18 (m, 1H), 7.71 (t, *J* = 7.9 Hz, 1H), 4.73 - 4.69 (m, 1H), 4.36 - 4.31 (m, 1H), 4.10 - 4.06 (dd, *J* = 8.9, 6.4 Hz, 1H), 3.93 - 3.90 (dd, *J* = 8.9, 5.3 Hz, 1H), 2.68 - 2.65 (m, 2H), 1.93 (t, *J* = 2.7 Hz, 1H), 1.34 (s, 3H), 1.32 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.7, 138.0, 135.4, 132.9, 130.7, 129.9, 129.8, 110.3, 79.6, 77.7, 75.0, 71.8, 65.9, 26.6, 25.1, 21.8; ESI-MS (*m/z*): $[MW-H]^-$ calcd. for $C_{16}H_{18}O_7S$, 353.1; found, 353.4.

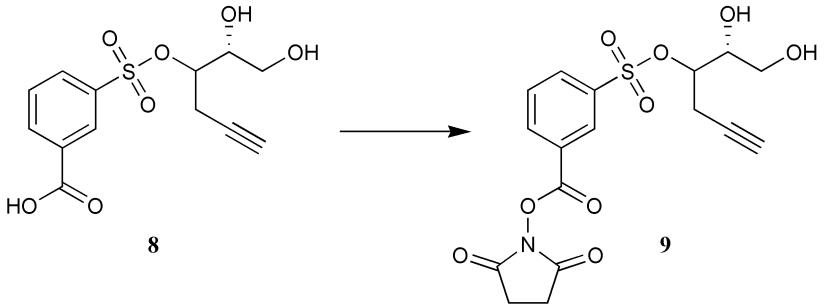
3-[1-(1,2-Dihydroxy-ethyl)-but-3-ynyloxysulfonyl]-benzoic acid (8)



7 **8**

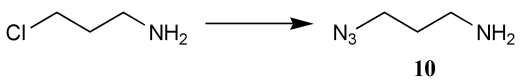
2 mL 3 M HCl were added to a solution of **7** (429 mg, 1.2 mmol) in 6 mL THF and stirred at room temperature for 25 h. THF was removed under reduced pressure (40 °C, 100 mbar), the leftover aqueous solution was combined with 40 mL 1 M HCl and extracted with EtOAc (2 x 40 mL). The combined organic layers were dried (Na_2SO_4) and concentrated *in vacuo* to yield **8** (326 mg, 1.0 mmol, 86 % yield). TLC (hexane:EtOAc:AcOH, 50:50:1 v/v): Rf = 0.13; ^{1}H NMR (400 MHz, MeOD-d4): δ 8.53 (t, *J* = 1.7 Hz, 1H), 8.33 - 8.31 (dt, *J* = 7.9, 1.3 Hz, 1H), 8.17 - 8.15 (ddd, *J* = 7.9, 1.9, 1.2 Hz, 1H), 7.73 (t, *J* = 7.9 Hz, 1H), 5.11 (s, *br*, 3H), 4.72 - 4.69 (m, 1H), 3.95 - 3.91 (m, 1H), 3.61 - 3.57 (dd, *J* = 11.5, 4.7 Hz, 1H), 3.48 - 3.44 (dd, *J* = 11.5, 6.2 Hz, 1H), 2.70 - 2.56 (m, 2H), 2.12 (t, *J* = 2.7 Hz, 1H); ^{13}C NMR (100 MHz, MeOD-d4): δ 167.6, 138.6, 135.8, 133.2, 133.0, 130.8, 130.0, 82.2, 79.4, 72.6, 72.1, 63.2, 20.9; ESI-MS (*m/z*): [MW-H]$^-$ calcd. for $C_{13}H_{14}O_7S$, 313.0; found, 312.9.

3-[1-(1,2-Dihydroxy-ethyl)-but-3-ynyloxysulfonyl]-benzoic acid 2,5-dioxo-pyrrolidin-1-yl ester (9)



DCC (2.7 mg, 13 µmol) in 20 µL DMF was added to a solution of NHS (1.5 mg, 13 µmol) and **8** (3.14 mg, 10 µmol) in 80 µL DMF in a 1.5 mL Eppendorf tube. After incubation in a thermomixer (25°C, 500 rpm) for 15 h the suspension was centrifuged ($16.1 \cdot 10^3$ rfc, 4 °C, 12 min) and the supernatant containing crude **9** was collected. The crude product solution (approx. 100 mM) was used for the synthesis of reactive strand **RS1** without further purification and stored at -20 °C until use.

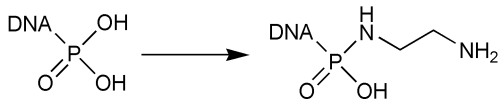
Synthesis of 3-azido-1-propylamine (10)[3]



Sodium azide (1.5 g, 23.0 mmol) was added to a solution of 3-chloropropyl-1-amine hydrochloride (1.03 g, 7.6 mmol) in water (15 mL) and the reaction mixture was heated to 80 °C for 36 h. The solution was basified with solid potassium hydroxide to pH = 10 and extracted with diethyl ether (3 × 25 mL), the organic layer was washed with water (10 mL) and brine (10 mL) and dried (Na_2SO_4). After concentration a volatile colourless oil was formed (**10**, 0.71 g, 95 %

yield). ^{1}H NMR (300 MHz, $CDCl_3$, δ): 3.34 (t, *J* = 6.8 Hz, 2H), 2.81 ppm (t, *J* = 6.8 Hz, 2H), 1.71 (m, *J* = 6.8 Hz, 2H), 1.46 ppm (t, *J* = 2.7 Hz, 2H).

b. Synthesis of the reactive strand RS1

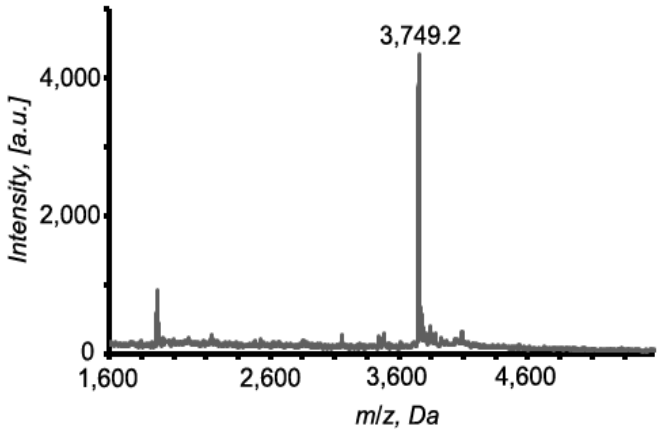


***CTAB precipitation of oligonucleotide*[4]**

A 3'-phosphorylated oligonucleotide (5'-GCC TGT TGT CGC-(PHOS)-3') in ddH_2O (25 µL, 4 mM) was treated with aq. cetyltrimethylammonium bromide (CTAB, 6 µL, 8 % w/v) and incubated for 30 min at room temperature to form a turbid solution. The solution was centrifuged ($16.1 \cdot 10^3$ rfc, 4 °C, 30 min), the supernatant was removed, and the precipitate was dried under vacuum.

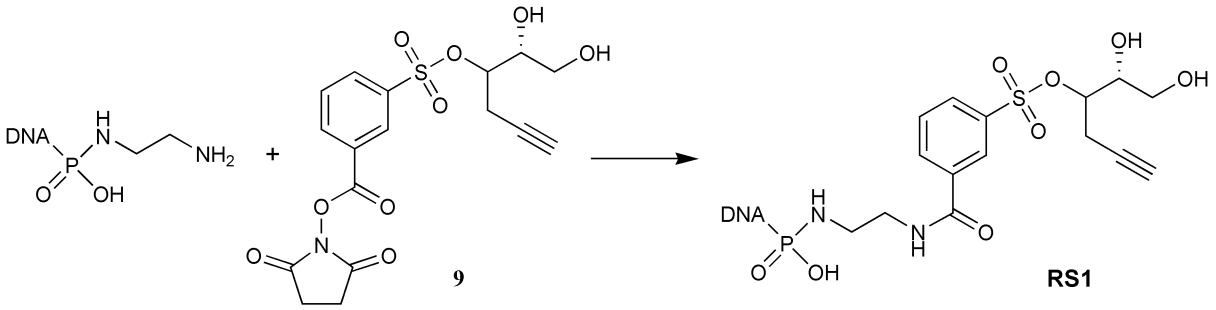
***Synthesis of 3'-amino modified oligonucleotide*[4]**

The CTAB precipitated oligonucleotide was dissolved in DMSO (50 µL, 2 mM) and combined with 2,2'-dipyridyl disulfide (PDS) in DMSO (50 µL, 1.6 M), 4-dimethylaminopyridine (DMAP) in DMSO (50 µL, 1.2 M) and triphenylphosphine (PPh_3) in DMF (50 µL, 1.2 M) and incubated in a thermomixer (25 °C, 500 rpm) for 25 min to form a yellow solution. 8 µL (approx. 1000 eqs.) ethylene diamine was added and incubation was continued for 30 min. The resulting amino-modified oligonucleotide was obtained quantitatively after precipitation with 2 % (w/v) $LiClO_4$ in acetone and washing with acetone. MALDI-TOF (*m/z*): $[MW-H]^-$ calcd. 3749.1 Da, found 3749.2 Da; RP-HPLC conditions were as follows: flow rate: 0.5 mL/min; solvent A: 0.1 M triethylammonium acetate (TEAA) (pH = 7.0), solvent B: MeCN; gradient: 0 - 35 min 7 - 13 % B, 35 - 38 min 13 % B, 38 - 48 min 13 - 100 % B; λ = 260 nm: *rt* = 14.9 min.

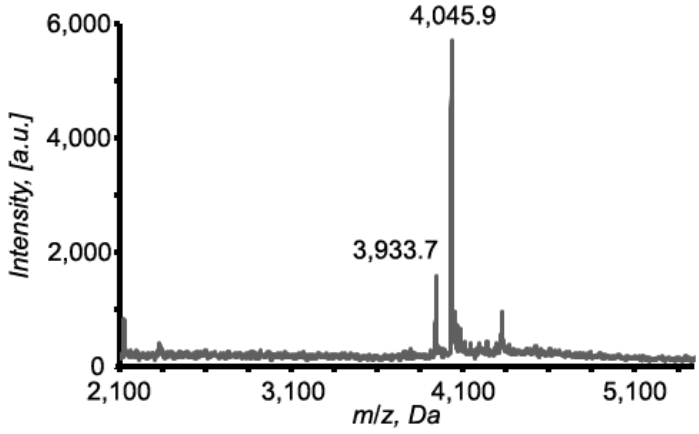


Supplementary Figure S1. MALDI-TOF spectrum (*m/z*) of amino-modified oligonucleotide, the precursor of **RS1**, $[MW-H]^-$ calcd. 3,749.1 Da, found 3,749.2 Da.

Synthetic procedure for the coupling of reactive group 9 to amino-modified oligonucleotide

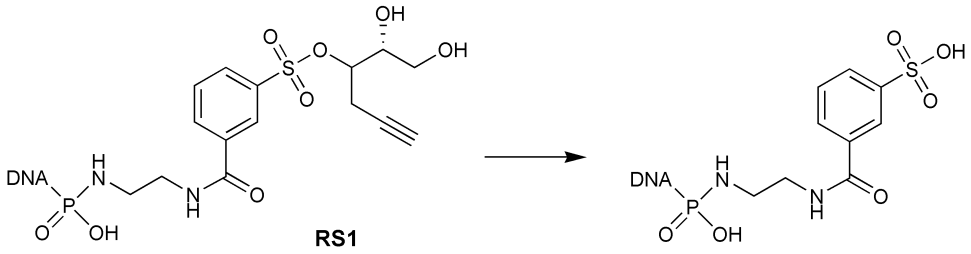


The crude NHS-ester **9** in DMF (2 µL, approx. 100 mM) was added to the amino-modified oligonucleotide (10 nmol) dissolved in 10 µL 100 mM potassium phosphate buffer (pH = 7.5). Incubation in a thermomixer (25 °C, 500 rpm) for 1 h and precipitation with NaCl and EtOH yielded reactive strand **RS1** in approx. 90-95 % purity. Degradation at the tosyl group takes place relatively fast at room temperature. MALDI-TOF (*m/z*): $[MW-H]^-$ calcd. 4045.7, found 4045.9; RP- HPLC conditions were as follows: flow rate: 0.5 mL/min; solvent A: 0.1 M TEAA (pH = 7.0), solvent B: MeCN; gradient: 0 - 35 min 7 - 13 % B, 35 - 38 min 13 % B, 38 - 48 min 13 - 100 % B; λ = 260 nm.: *rt* = 37.5 min.

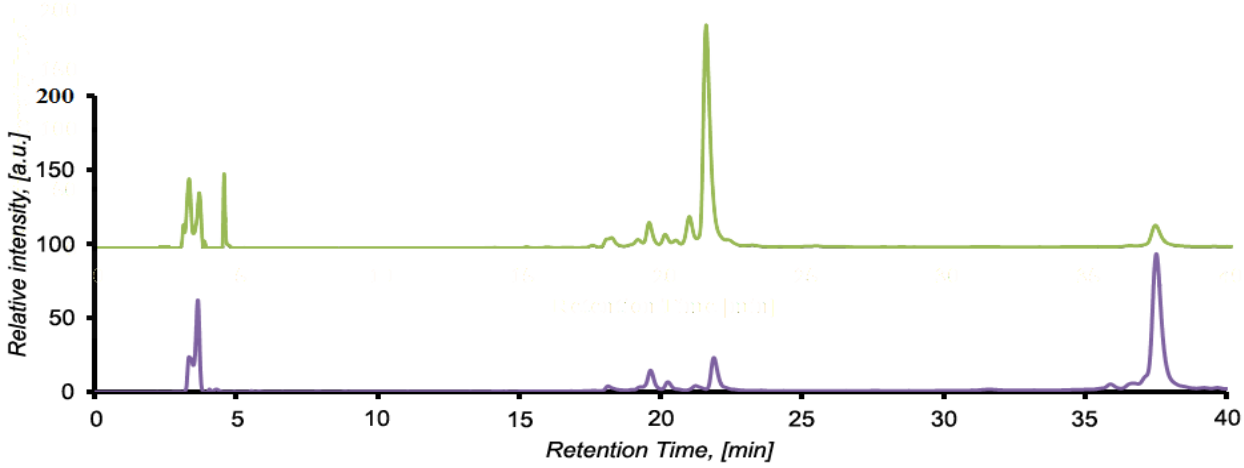


Supplementary Figure S2. MALDI-TOF spectrum (*m/z*) of **RS1**, $[MW-H]^-$ calcd. 4045.7, found 4045.9 Da. The signal at 3933.7 Da corresponds to the degraded strand **RS1**, $[MW-H]^-$ calcd. 3932.7 Da (see also section c below).

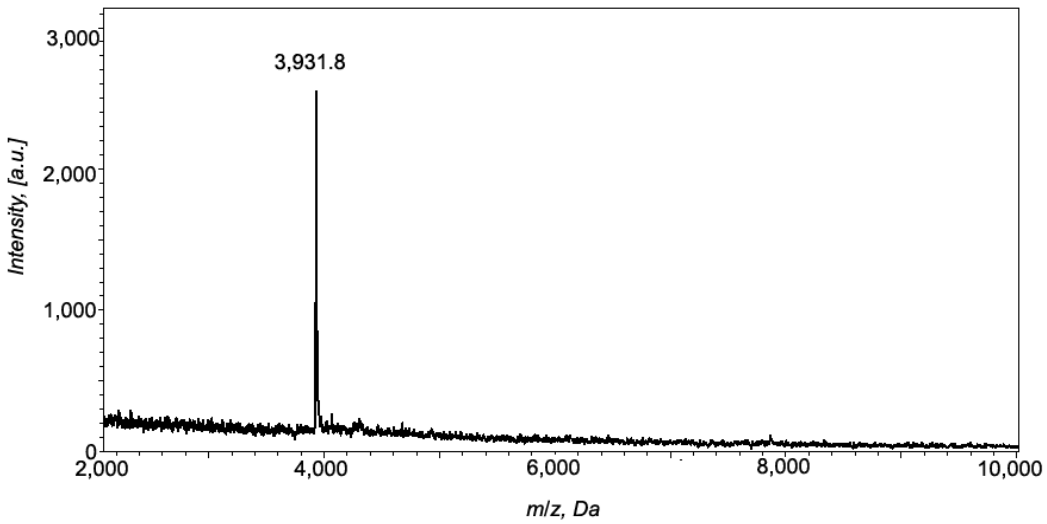
c. Degradation of reactive strand RS1 at higher temperatures



Using standard annealing protocols (2 min at 90 °C and slowly cooling down to room temperature in 45 min) leads to a drastically decreased yield of propargyl-ε-adenine formation. The reason for this lies in the susceptibility of the reactive strand to degradation at higher temperatures. Analysis by RP-HPLC before and after standard annealing conditions shows a significant decrease of the signal for reactive strand **RS1** and an increase of the signal for degraded **RS1**. The ratio between the two strands decreases from approximately 85:15 to 15:85. The degraded strand **RS1** was analyzed by MALDI-TOF analysis revealing the loss of the glycerol moiety leaving only the sulfonic acid moiety attached to the oligonucleotide. HPLC conditions were as follows: flow rate: 0.5 mL/min; solvent A: 0.1 M TEAA (pH = 7.0), solvent B: MeCN; gradient: 0 - 35 min 7-13 % B, 35 - 38 min 13 % B, 38 - 45 min 13-100 % B; λ = 260 nm.



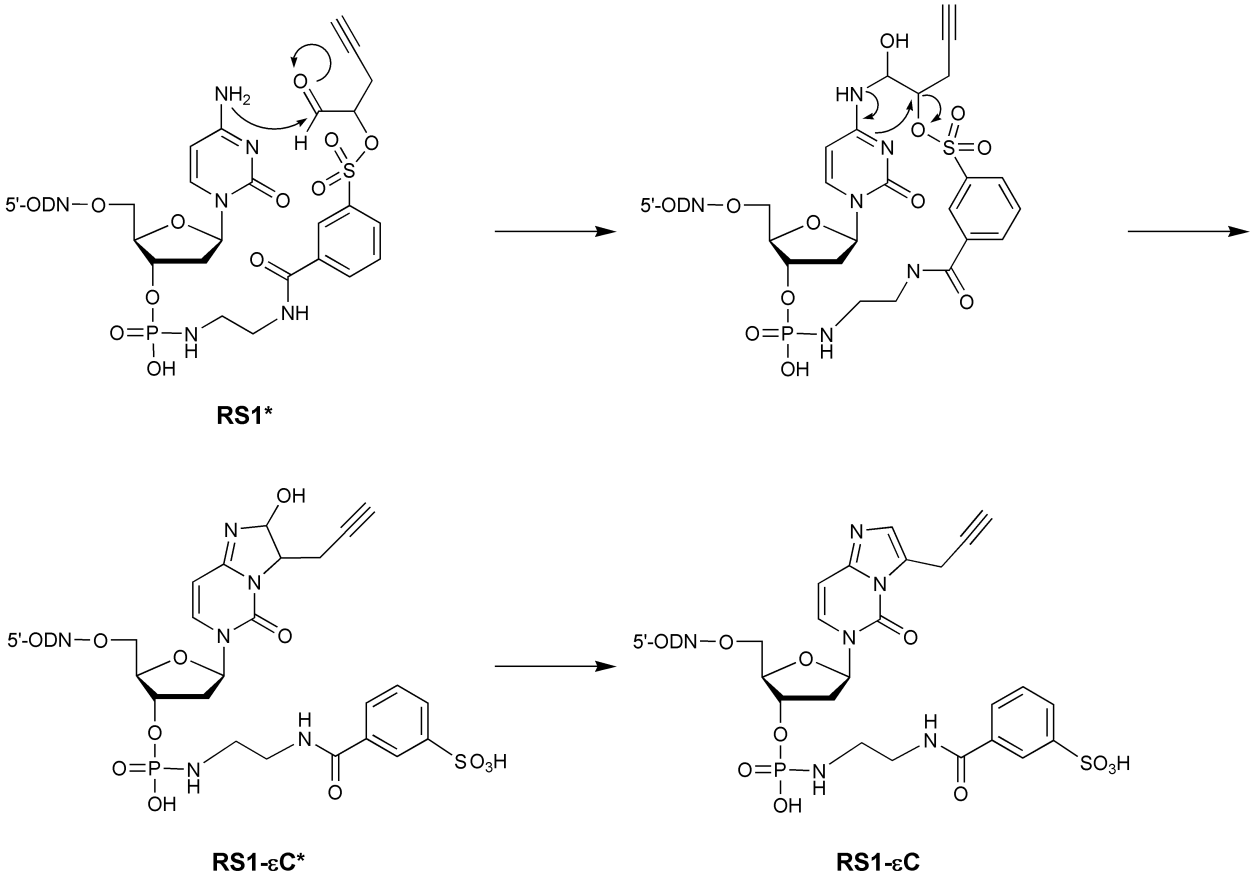
Supplementary Figure S3. HPLC chromatogram (λ = 260 nm) of reactive strand **RS1** before (purple, bottom) and after annealing conditions (2 min at 90°C, green, top). The peak with *rt* = 21.9 min corresponds to degraded **RS1** and the peak with *rt* = 37.5 min to intact **RS1**. The area under the peaks was used to measure the ratio of intact to degraded **RS1**, which decreased from approximately 85:15 to 15:85.



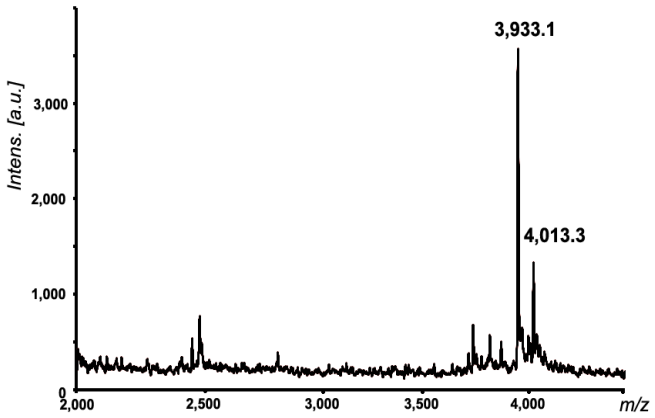
Supplementary Figure S4. MALDI-TOF spectrum (*m/z*) of peak at 21.9 min corresponds to the degraded strand **RS1**, $[MW-H]^-$ calcd. 3932.7, found 3931.8 Da.

d. Self-reaction of the reactive strand RS1

RS1 was incubated under standard conditions for etheno-formation (see III. ε-Base forming reaction). Oxidation with $NaIO_4$ yields an aldehyde moiety, the activated form of the reactive group, denoted strand **RS1***. If there is no target strand present, the aldehyde is able to attack the cytosine residues within the oligonucleotide it is attached to. The suggested mechanism of etheno-base formation[5] involves then the generation of a carbinolamine, followed by the cyclization to the hydroxyethano intermediate **RS1-εC***. Elimination of H_2O finally yields the inactivated reactive strand bearing an 12-propargyl-εC moiety (**RS1-εC**). The hydroxyethano intermediate is most likely responsible for the peak 4013.3 *m/z* detected in MALDI-MS (Supplementary Figures S6), the 12-propargyl-εC structure, however, was not detected, possibly because no heating step was performed.



Supplementary Figure S5. Possible intermediates of the self-reaction of reactive strand **RS1**.



Supplementary Figure S6. MALDI-TOF spectrum (*m/z*) of reaction mixture after incubation of reactive sequence **RS1**. Degraded **RS1**, $[MW\text{-}H]^-$ calcd. 3932.7, found 3933.1, Self-modified reactive strand **RS1-εC***, $[MW\text{-}H]^-$ calcd. 4013.7, found 4013.3.

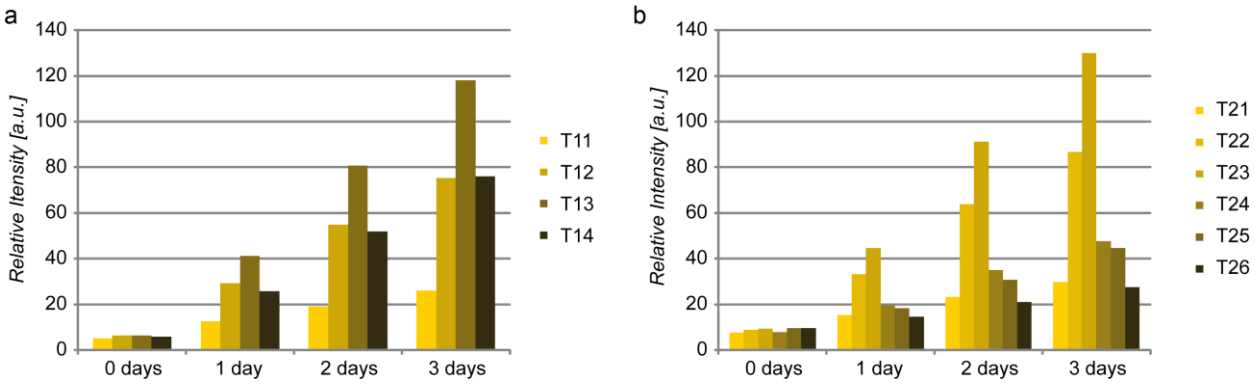
IV. ε-Base forming reaction

a. General ε-formation procedure with reactive and target strands

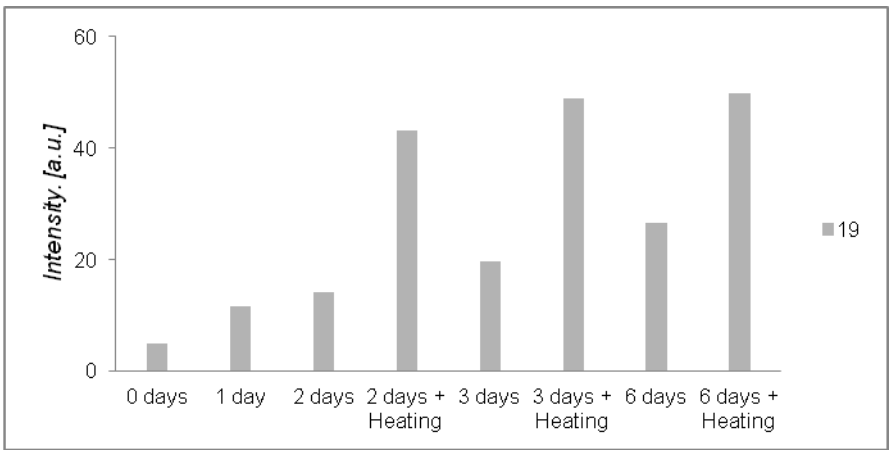
To generate the εA formation, reactive strand **RS1** (0.6 nmol, 5 µL) and the respective target strand (0.3 nmol, 5 µL) were mixed with 60 µL NaCl (1 M) and 50 µL NaOAc (20 mM, pH = 5.5) and incubated for 15 min. After the addition of 30 µL 10 mM $NaIO_4$ in NaOAc (20 mM, pH = 5.5) the mixture was incubated in a thermomixer (25 °C, 500 rpm) for 90 min. $NaIO_4$ was quenched with 30 µL 15 mM ethylene glycol in NaOAc (20 mM, pH = 5.5) and incubation was continued under the same conditions. After 3 days the reaction mixture was heated to 60 °C for 2 hours to accelerate dehydration of the hydroxyethanoA intermediate. Reaction mixtures were used without dilution or purification for fluorescence measurements and RP-HPLC analyses. Samples were purified with NAP-5 columns or with ZipTip C18 for PSA and MALDI-TOF MS measurements.

b. Fluorescence spectroscopy

Fluorescence emission profiles of ε-formation with short (**T11-T14**) and long target strands (**T21-T26**) were recorded from 200 to 800 nm upon excitation at 275 nm, the emission at 410 nm for samples taken at different time points is depicted in Fig. S7. Fluorescence quenching by stacking interactions of neighbouring nucleobases has to be considered for samples with low emission.



Supplementary Figure S7. Fluorescence emission at 410 nm upon excitation at 275 nm of the reactions of reactive strand **RS1** with target strands (**T11-T14**) (A) and (**T21-T26**) (B) after 0-3 days and without heating step at the end of the reaction.



Supplementary Figure S8. Fluorescence emission at 410 nm upon excitation at 275 nm of the reactions of reactive strand **RS1** with target strand **T11** monitored for 0-6 days with and without heating. Using a heating step at the end of the reaction promotes the conversion of hydroxyethano intermediates still present to etheno products (see Fig. 3). While the etheno products are fluorescent, the intermediates are not.

c. Denaturing polyacrylamide gel electrophoresis (PAGE)

Acrylamide solution for 12 % denaturing gels was prepared as follows: AccuGel (150 mL), urea (210 g) and TBE x 10 (50 mL) were combined while stirring and heating at 50-60 °C for 1-2 h to obtain a clear solution. ddH_2O was added to a final volume of 500 mL and stirring was continued for 10 min. The clear solution was set aside for approx. 1 h to cool down, filtered through a stericup and stored in a sterile bottle at 4 °C and protected from light. Polymerization of the gel solution was induced by addition of TEMED and APS: For 50 mL gel solution 20 µL TEMED and 200 µL APS were used. Gels were stored at 4 °C overnight prior to use.

Radioactive Labeling of primer (35)

A solution of primer **P1** (5 µM), [γ-^{32}P]-ATP (1:6 dilution of standard solution) and T4 PNK (10 U) in PNK buffer (1:6 dilution) in a total volume of 6 µL was incubated at 37 °C for 60 min without agitation. The reaction mixture was combined with 12 µL stop solution (82 % formamide, 0.16 % xylene cyanol and 0.16 % bromophenol blue) and purified with 12 % denaturing PAGE. The radioactively labeled primer **P1** was recovered by cutting out the corresponding band on the gel, extraction by "crush and soak" (10 mM MOPS pH = 6, 1 mM EDTA, 250 mM NaCl), and precipitation with EtOH. After drying, radiolabeled **P1** was dissolved in 70 µL ddH_2O and used for PSA.

d. Polymerase stop assay (PSA)

Reaction mixtures of reactive and target strands were heated to 60 °C for 2 h and purified with ZipTip C18, dried and re-dissolved in ddH_2O to get a concentration of 40 nM relative to the target strand. 5 µL of this solution was combined with ddH_2O (7.5 µL), GoTaq buffer (4 µL) and ^{32}P-labeled primer **P1** (1 µL). Annealing of **P1** and target strand was induced by incubation at 90 °C for 2 min and slowly cooling down. After addition of dNTPs (0.5 µL, 10 mM) and Taq DNA polymerase (2 µL, 0.2 U), the mixtures were incubated at 55 °C for 30 min. The reactions were quenched with 15 µL stop solution (82 % formamide, 0.16 % xylene cyanol and 0.16 % bromophenol blue) and incubated at 90 °C for 5 min prior to analysis by 12 % denaturing PAGE. Visualization and quantification was conducted on a Typhoon Scanner FLA 9500 with ImageQuant TL (1D gel analysis, background subtraction Rolling Ball: 500).

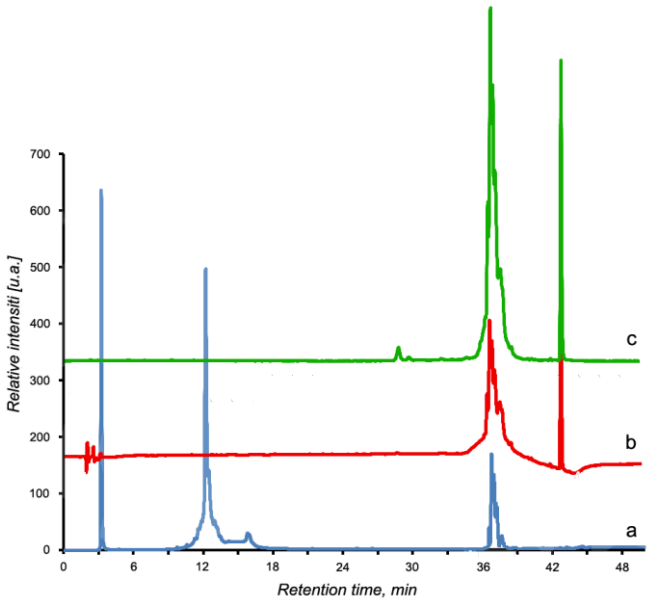
Supplementary Table S1. Yields with standard deviations of etheno-base formation in reactions of **RS1** with target strands (**T21**-**T26**) determined by PSA with radioactively labeled primer **P1**. (*n*+*y*) = position of target base.

Lane	Target strand	*n*+*y*	PSA Yield
2	**T21**	-	1 %
3	**T21**	1	66 ± 3 %
4	**T22**	2	67 ± 5 %
5	**T23**	3	56 ± 3 %
6	**T24**	4	41 ± 1 %
7	**T25**	6	11 ± 1 %
8	**T26**	9	10 ± 1 %

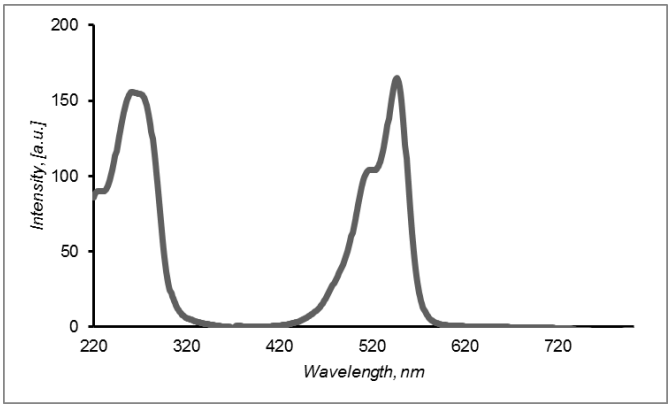
V. General procedure for the Cu(I)-catalysed Huisgen 1,3-dipolar cycloaddition.

For the Cu(I)-catalysed Huisgen 1,3-dipolar cycloaddition, a solution of the respective 12-propargyl-εA containing target strand (4 nmol, 150 µL) in 1 M TEAA buffer at pH = 7.2 was mixed with 50 µL of DMSO, 8 µL of 10 mM ascorbic acid and either 3-azido-1-propanamine or the azide-Cy3 derivative (80 nmol in 20 µL of DMSO). The solution was briefly vortexed and saturated with Argon for 30 s before 4 µL of a 10 mM solution containing $CuSO_4$ and TBTA in a 1:1 ratio in H_2O:DMSO (45:55) was added. The solution was saturated with Argon for 1 min and the vial closed. After incubation for 16 hours at 20 °C and precipitation with 700 µL of EtOH (=

3 volumes) in the presence of 10 µL of 3 M NaCl, the resulting product was dissolved in deionised water and purified by RP-HPLC: Waters X-Bridge C18 column (3.5 µm, 4.6 x 150 mm). Solvent A: TEAA 100 mM with 1% CH_3CN, pH 7.0. Solvent B: acetonitrile. Flow rate: 0.5 mL/min. Solvent gradient: 0-10 min, 5-10% solvent B; 10-30 min, 10-20% solvent B; 30-40 min, 20-80% solvent B; 40-50 min, 80% solvent B. Subsequently, traces of non-conjugated Cy3 dye were removed with a NAP-5 column (GE Healthcare).



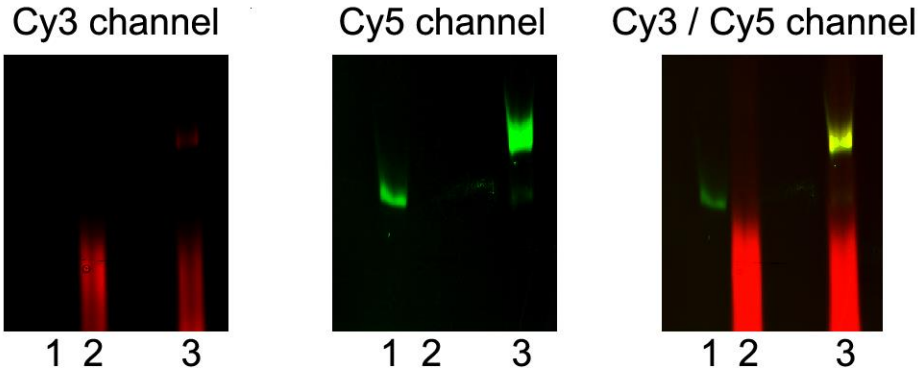
Supplementary Figure S9. RP-HPLC chromatograms of Huisgen 1,3-dipolar cycloaddition of azido-Cy3 fluorophore to 12-propargyl-εA modified oligonucleotide **T31-εA**. The reaction was analyzed at λ = 260 nm (**a**), λ = 550 nm (**b**) and with fluorescence detector $\lambda_{ex.}$ = 550 nm, $\lambda_{em.}$ = 570 nm (**c**). The observed peaks belong to the buffer components (*rt* = 3 min), reacted (or degraded) reactive strand **RS1**, non-modified **T31**, as well as **T31-εA** (*rt* = 12.2 min), **T31-εA-Cy3** as well as probably **T31-hydroxoethanoA-Cy3** (*rt* = 36.5-37.9 min), and free azide-Cy3 fluorophore (*rt* of 42.8 min). Additional broadening of the peaks at 12.2 and 36.5-37.9 min is attributed to impurities of the oligos (*i.e.*, (*n*-1), (*n*-2), etc., as they were purchased in desalted and not HPLC purified quality).



Supplementary Figure S10. UV-Vis spectrum of **T31-εA-Cy3** recorded with the diode array detector of the HPLC system. The intensity maxima at 260 nm and 550 nm are typical for DNA and Cy3, respectively.

a. Native polyacrylamide gel electrophoresis (PAGE)

Acrylamide solution for 12 % native gels was prepared as follows: AccuGel (30 mL), ddH_2O (40 mL) and TBE X 10 (20 mL) were combined while stirring and heating at 50-60 °C for 1-2 h to obtain a clear solution. ddH_2O was added to a final volume of 100 mL and stirring was continued for 10 min. The clear solution was set aside for approx. 1 h to cool down, filtered through a stericup and stored in a sterile bottle at 4 °C and protected from light. Polymerization of the gel solution was induced by addition of TEMED and APS: For 50 mL gel solution 20 µL TEMED and 200 µL APS were used. Gels were stored at 4 °C overnight prior to use. Running buffer was prepared as follows: To a solution of 20.6 g Tris-HCl were added 6.6 g HEPES and 5g of KCl. The pH was adjusted to 7.5 with acetic acid and ddH_2O was added to a final volume of 100 mL.



Supplementary Figure S11. Native PAGE to evaluate annealing of post-synthetically labeled strand **T31-εA-Cy3** with the complementary Cy5-labeled strand **P31**. Lane 1 contains just **P31**, lane 2 just **T31-εA-Cy3**, and lane 3 **P31** and **T31-εA-Cy3** in a 20:1 ratio. The position of fluorophore carrying strands was either detected in the Cy5-chanel (green, $\lambda_{ex.}$ = 650 nm, $\lambda_{em.}$ =

670 nm) or the Cy3 channel (red, $\lambda_{ex.}$ = 550 nm, $\lambda_{em.}$ = 570 nm). The overlay of both emission channels (right picture) reveals co-localization of the Cy3 and the Cy5 label in one band (yellow) that is upward shifted relative to the individual single strands demonstrating duplex formation between the two DNA labeled strands. In the left picture, the band of **T31-εA-Cy3** (lane 2) is masked by a smear due to unbound and difficult to remove free Cy3 dye.

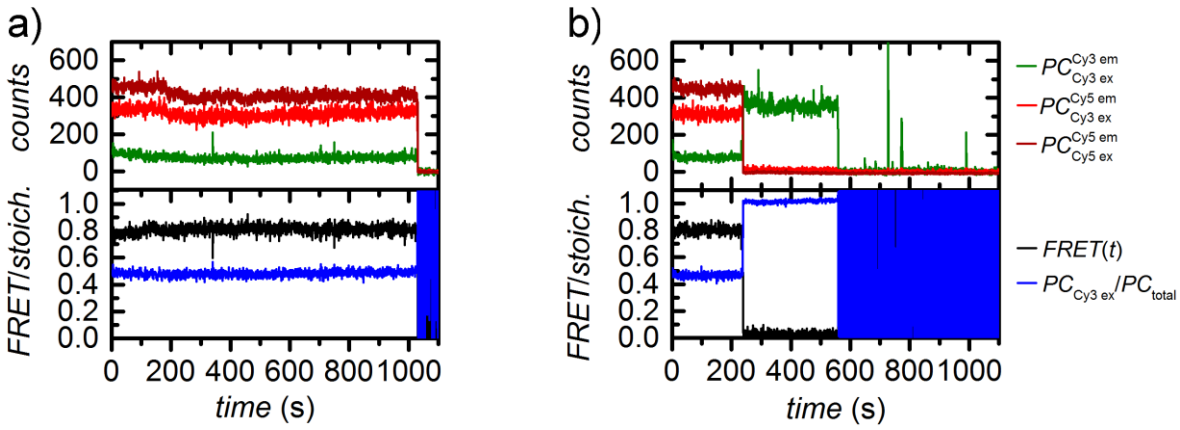
VI. smFRET measurements and data processing

For the smFRET imaging, sample chambers were built from quartz slides (Finkenbeiner, Waltham, MA) and passivated with biotinylated BSA. Strand **P31** was tethered to the surface via a biotin-streptavidin linkage, followed by annealing of strand **T31-εA-Cy3** at RT for 5 min. Excess dye was washed out and smFRET clips were recorded in 50 mM MOPS, 100 mM KNO_3, 1% w/v D-glucose, 1x enzymatic oxygen scavenging solution (165 U/mL glucose oxidase, 2170 U/mL catalase), and 1 mM Trolox, pH 6.90. Cy3 and Cy5 emissions were followed for 18 min at a time resolution of 300 ms using a home-built total internal reflection fluorescence microscope with a CCD camera and two lasers at alternating laser excitation at 532 and 640 nm as described.[6-8] Clips were corrected for background noise and cross-talk, followed by spot detection and co-localization of corresponding coordinates. Approximately 40 % of all spots detected in the Cy3 channel could be mapped onto a corresponding spot in the Cy5 channel, a finding that coincides with the fact that up to 55 % of all Cy5 molecules are known to be pre-bleached prior to the measurement.[9] Single fluorophore emission time traces were created by integration over 3x3 pixels around the central coordinates, yielding three emission time-dependent PCs: Cy3 emission upon Cy3 excitation, Cy5 emission upon Cy5 excitation and Cy5 emission upon Cy3 excitation ($\mathrm{PC}(t)^{Cy3\ em}_{Cy3\ ex}$, $\mathrm{PC}(t)^{Cy5\ em}_{Cy5\ ex}$, $\mathrm{PC}(t)^{Cy5\ em}_{Cy3\ ex}$). Single-fluorophore emission trajectories were used to calculate apparent FRET efficiencies and fluorophore stoichiometries over time using the following two equations:[6]

$$FRET(t) = \frac{\mathrm{PC}(t)^{Cy5\ em}_{Cy3\ ex}}{\mathrm{PC}(t)^{Cy3\ em}_{Cy3\ ex} + \mathrm{PC}(t)^{Cy5\ em}_{Cy3\ ex}} \quad (1)$$

$$stoichiometry(t) = \frac{\mathrm{PC}(t)^{Cy5\ em}_{Cy3\ ex} + \mathrm{PC}(t)^{Cy5\ em}_{Cy5\ ex}}{\mathrm{PC}(t)^{Cy3\ em}_{Cy3\ ex} + \mathrm{PC}(t)^{Cy5\ em}_{Cy3\ ex} + \mathrm{PC}(t)^{Cy5\ em}_{Cy5\ ex}} \quad (2)$$

Data analysis was performed using the MATLAB-based software package MASH.[10]

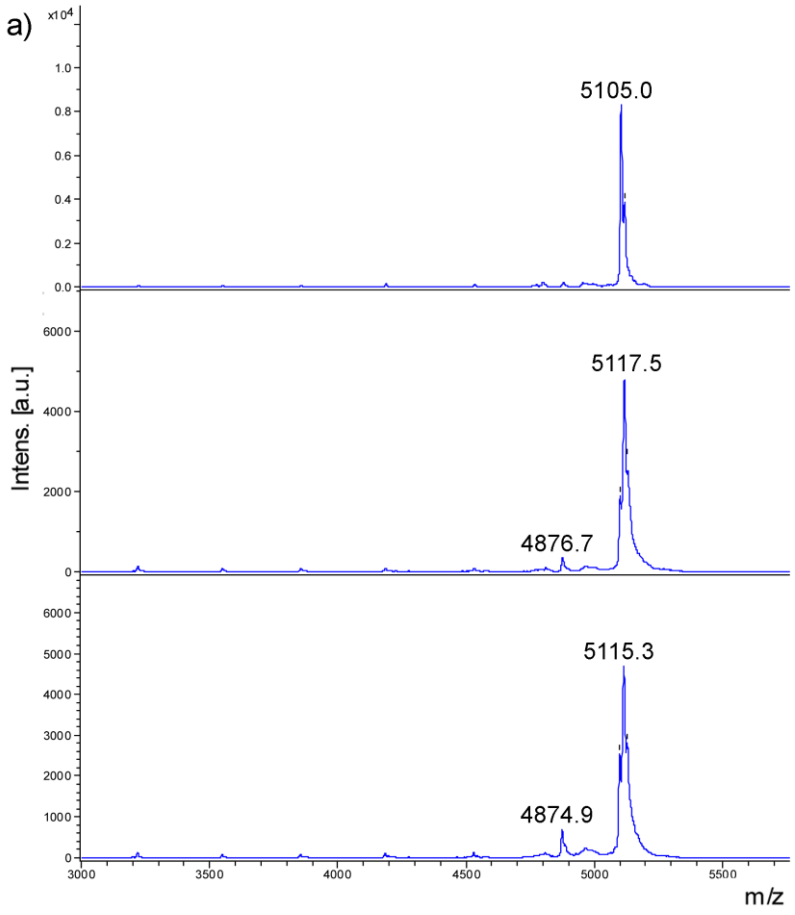


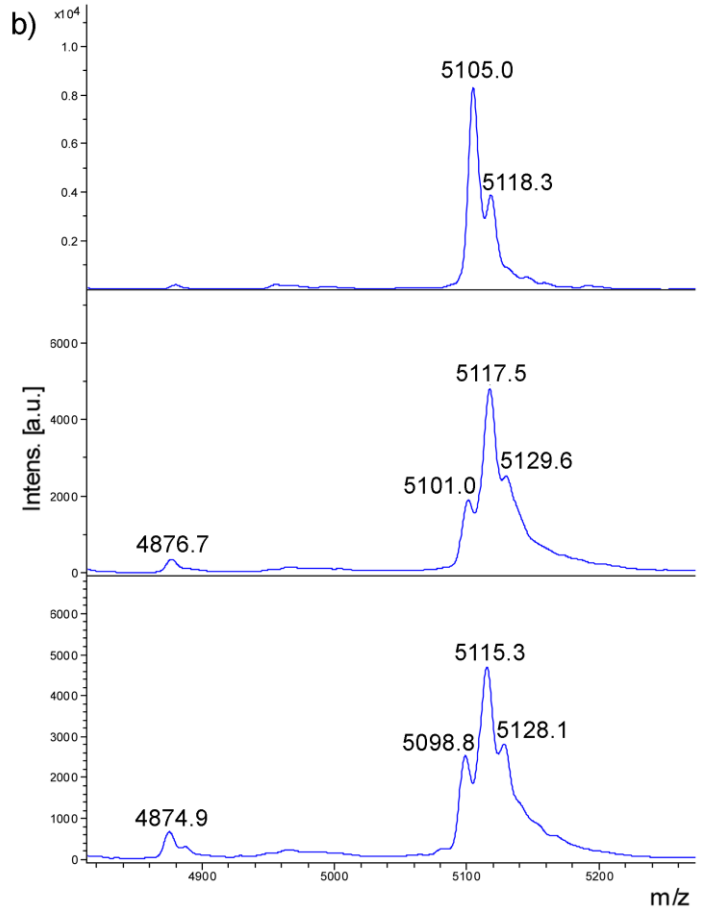
Supplementary Figure S12. Single-molecule imaging of surface-immobilized, post-synthetically labeled **P31/T31-εA-Cy3** duplex. Imaging conditions: 50 mM MOPS, 100 mM KNO_3, 1 % w/v D-glucose, 1x enzymatic oxygen scavenging solution, 1 mM 6-hydroxy-2,5,7,8-tetramethylchroman-2-cyclooctatetraene (Trolox). (a,b) Representative fluorophore emission, apparent FRET, and fluorophore stoichiometry time traces. Single-step photobleaching of Cy5 (b) results in the disappearance of FRET and a change in the Cy3:Cy5 stoichiometry from 1:1 to 1:0. Single-step photobleaching of Cy3 (a,b) is unambiguously characterized by high noise levels in FRET and stoichiometry trajectories.

VII. Modification of RNA target strand.

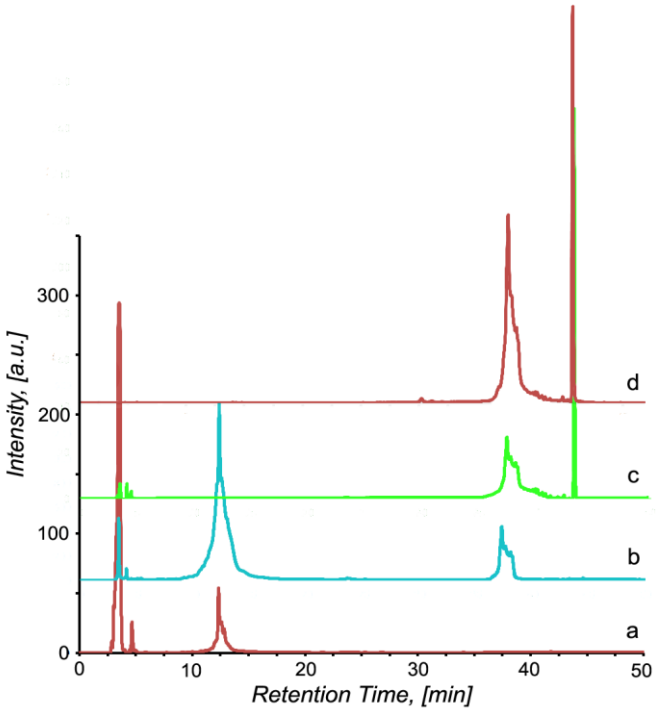
a. Stability of RNA target strand under ε-formation conditions.

The RNA target strand **RNA11** (20 µL, 200 µM) was combined with 60 µL NaCl (1 M) and 50 µL NaOAc (20 mM, pH = 5.5) and incubated for 15 min. After the addition of 30 µL 10 mM $NaIO_4$ in NaOAc (20 mM, pH = 5.5) the mixture was incubated in a thermomixer (25 °C, 500 rpm) for 90 min. $NaIO_4$ was quenched with 30 µL 15 mM ethylene glycol in NaOAc (20 mM, pH = 5.5) and incubation was continued under the same conditions. The reaction was stopped after 3 days. Samples were taken at different time points, desalted with ZipTip C18, and analyzed by MALDI-TOF (Figure S14).


a)
x10^4
1.0
0.8
0.6
0.4
0.2
0.0
5105.0
6000
4000
2000
0
5117.5
4876.7
6000
5000
4000
3000
2000
1000
0
5115.3
4874.9
Intens. [a.u.]
3000
3500
4000
4500
5000
5500
m/z


b)
x10^4
1.0
0.8
0.6
0.4
0.2
5105.0
5118.3
6000
4000
2000
0
5117.5
5129.6
5101.0
4876.7
6000
5000
4000
3000
2000
1000
0
5115.3
5128.1
5098.8
4874.9
Intens. [a.u.]
4900
5000
5100
5200
m/z

Supplementary Figure S13. MALDI-TOF spectra (*m/z*) collected in the negative ion mode of **RNA11** as obtained from the commercial source (Microsynth, HPLC purified; top spectra in a) and b)), after incubation with $NaIO_4$ for 90 min (spectra in the middle), and after incubation in the quenched mixture for 3 days as described above in the text (bottom spectra). a) depicts the entire mass range and b) shows a zoom of the main signals. Top spectra reveal the presence of unmodified **RNA11** (calcd. $[M\text{-}H]^-$ 5103.16 Da) and its NH_4^+ adduct (calcd. $[M+NH_4^+\text{-}H]^-$ 5120.19 Da). Oxidation produces a dialdehyde at the 3'-ribose (calcd. 5101.14 Da, 5118.17 Da (NH_4^+ adduct), 5123.11 Da (Na^+ adduct)) as well as a minor signal for **RNA11** without the 3-terminal cytosine (calcd. 4876.7 Da).



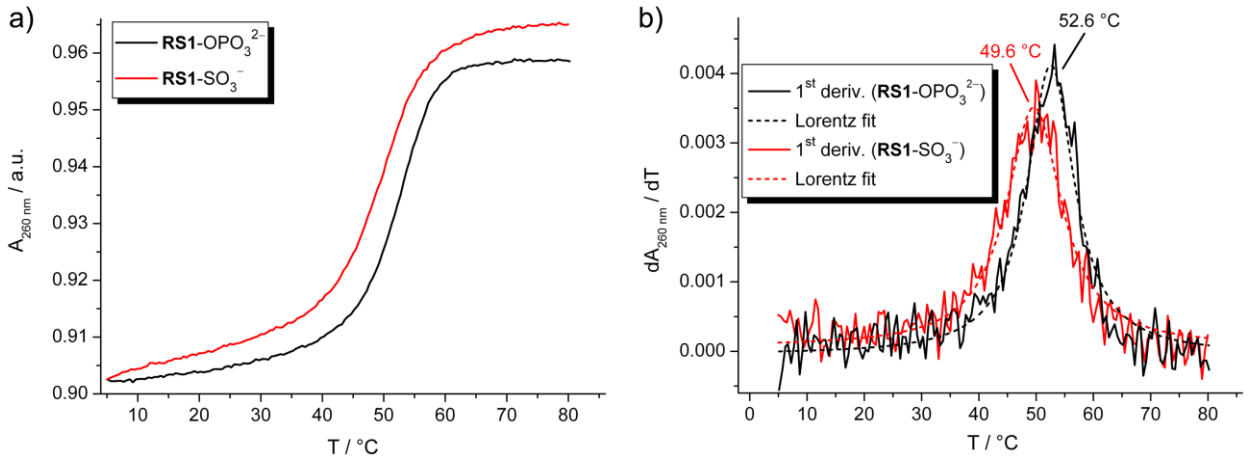
Supplementary Figure S14. RP-HPLC chromatograms of RNA strand modification. (**a**): **RNA11** was incubated at etheno-forming conditions in the absence of reactive strand **RS1** (monitoring wavelength $\lambda = 260$ nm). **B-D:** The Huisgen 1,3-dipolar cycloaddition reaction between the azido-Cy3 fluorophore and the 12-propargyl-εA modified RNA primer **RNA11-εA** was monitored at $\lambda = 260$ nm (**b**), $\lambda = 550$ nm (**c**), and $\lambda_{ex.} = 550$ nm, $\lambda_{em} = 570$ nm (**d**). The observed peaks belong to the buffer components (*rt* = 3 min), reacted (or degraded) reactive strand **RS1**, non-modified **RNA11**, as well as **RNA11-εA** (*rt* = 12.9 min), **RNA11-εA-Cy3** as well as probably **RNA11-hydroxoethanoA-Cy3** (*rt* = 37-39 min), and free azide-Cy3 fluorophore (*rt* of 42.8 min). Additional broadening of the peaks at 12.9 and 37-39 min is attributed to impurities of the oligonucleotides (*i.e.*, (*n*-1), (*n*-2), etc., as they were purchased in the desalted form and not HPLC purified).

VIII. Determination of melting temperatures with UV spectroscopy.

Melting curves of duplexes were recorded at λ = 260 nm with a heating rate of 1 °C/min using a quartz cuvette with 1 cm path length.

a. T_m of reactive sequence precursor RS1-PO_3^{2-} and degraded reactive sequence RS1-SO_3^-, respectively, with T12

The final concentrations of **T12** and **RS1-PO_3^{2-}** or **RS1-SO_3^-** were 5 and 1.6 µM, respectively, in 50 mM NaCl.

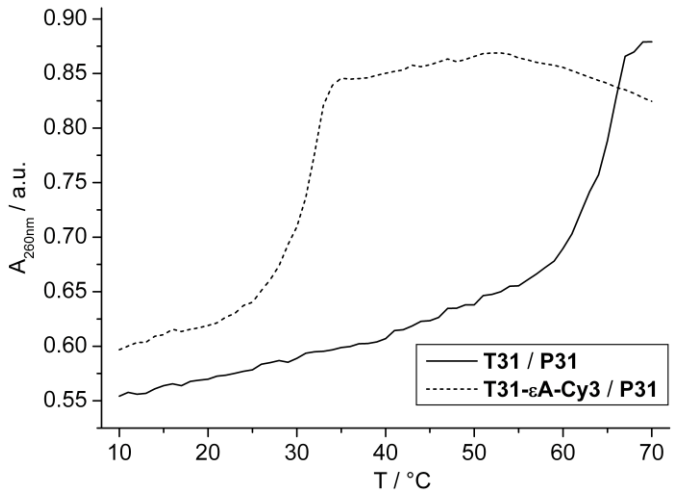


Supplementary Figure S15. Melting profiles of **T12/RS1-PO_3^{2-}** (black) and **T12/RS1-SO_3^-** (red). a) The absorption at 260 nm is plotted against the respective temperature. For better comparison, the starting point at 5 °C was arbitrarily set to 0.9025 for both curves. b) Plot showing the first derivatives (solid lines) of the melting curves depicted in a) as well as the Lorentzian fitting of the two curves (dashed lines) to determine the T_m values of 52.6 °C for **T12/RS1-PO_3^{2-}** and of 49.6 °C for **T12/RS1-SO_3^-**.

b. T_m of T31/P31 and T31-εA-Cy3/P31 used for smFRET studies

For all experiments cyanine dye-labeled DNA strands were used that were dissolved in degassed water and stored at -20 °C.

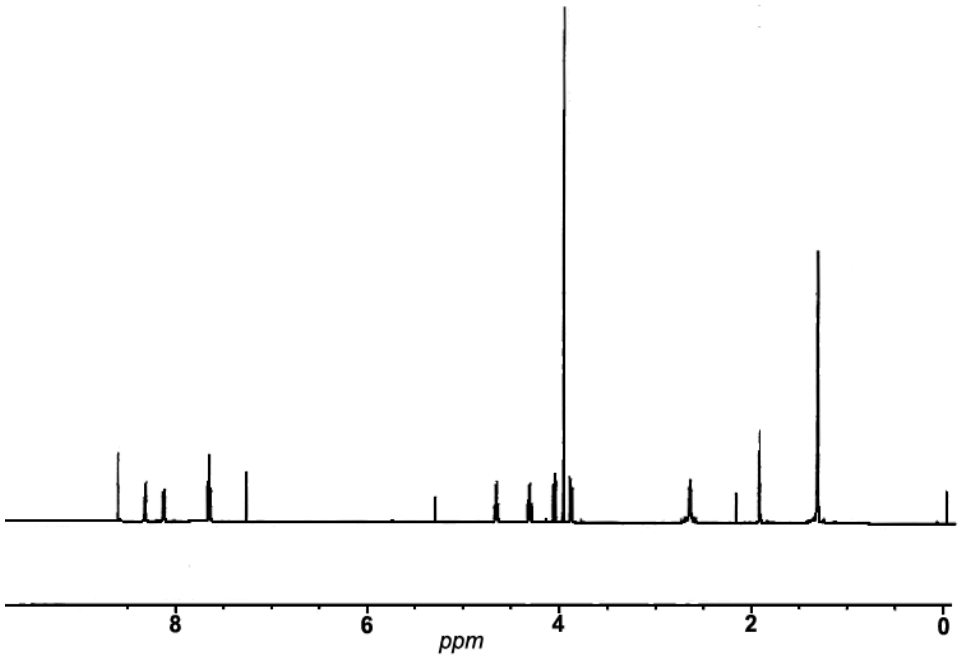
The final concentrations of **P31** and **T31** or **T31-εA-Cy3** were 2 and 1 µM, respectively, in 100 mM NaCl and 10 mM sodium acetate buffer at pH 7.



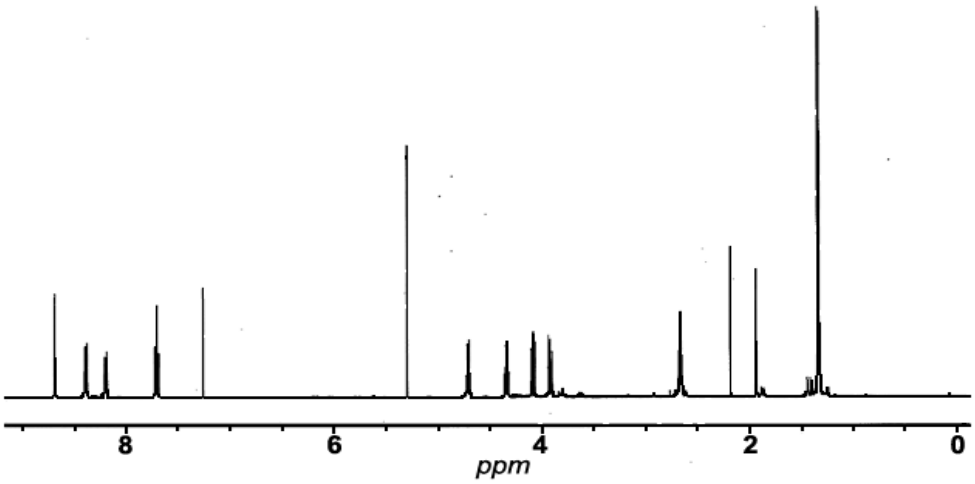
Supplementary Figure S16. Melting profiles of **T31**/**P31** and **T31-εA-Cy3**/**P31**. The absorption at 260 nm is plotted against the respective temperature resulting in a melting temperature T_m of 64 °C for the **T31**/**P31** duplex (solid line) and of 32 °C for the **T31-εA-Cy3**/**P31** pair (dashed line).

IX. Spectra

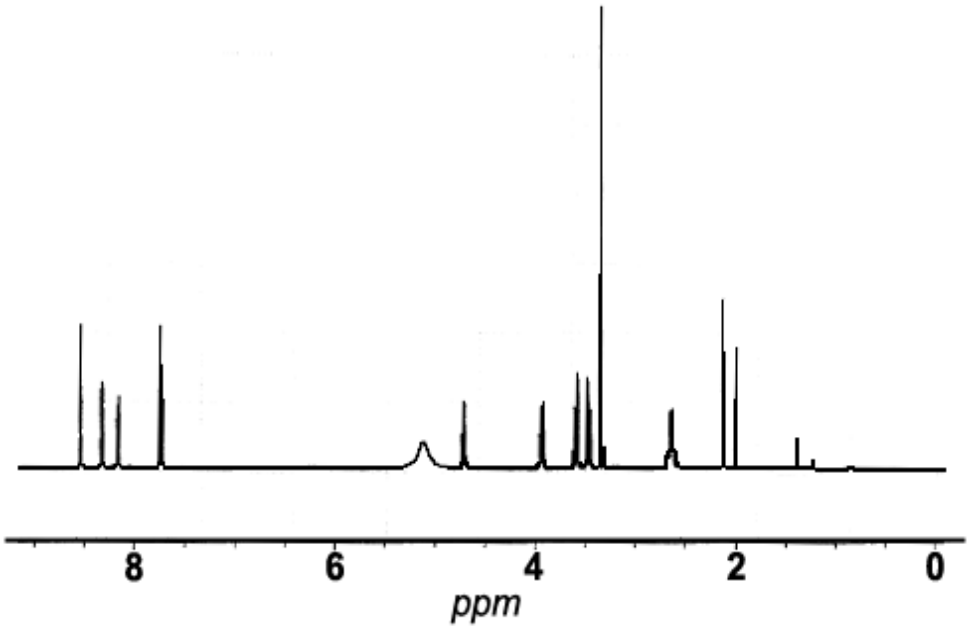
a. ^{1}H NMR spectra.



Supplementary Figure S17. ^{1}H NMR spectrum of **6** in $CDCl_3$.

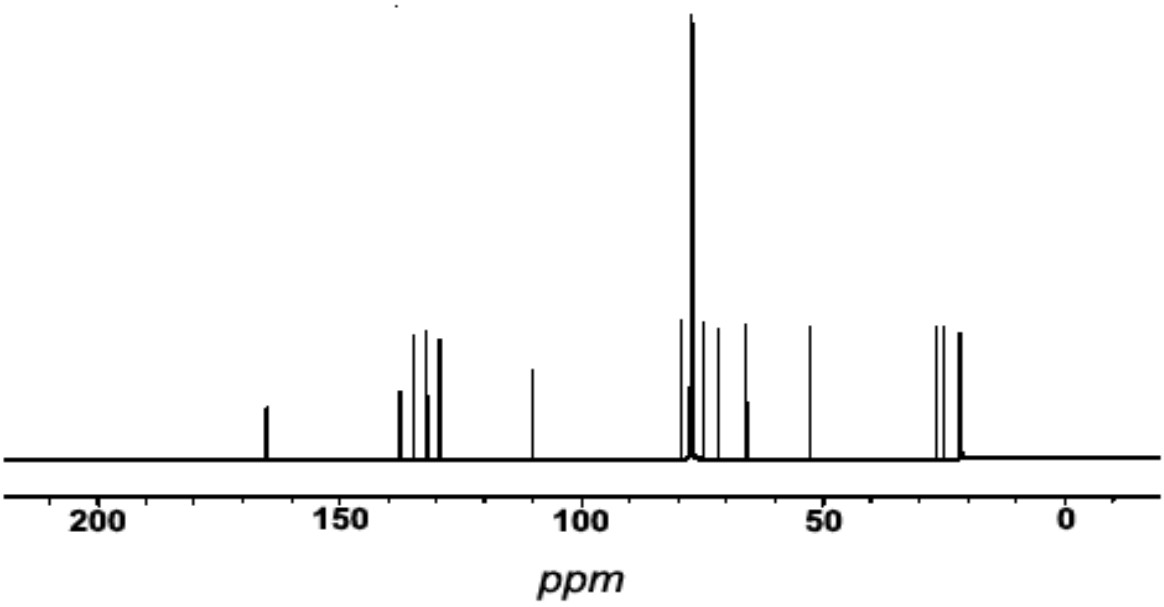


Supplementary Figure S18. ^{1}H NMR spectrum of **7** in $CDCl_3$.

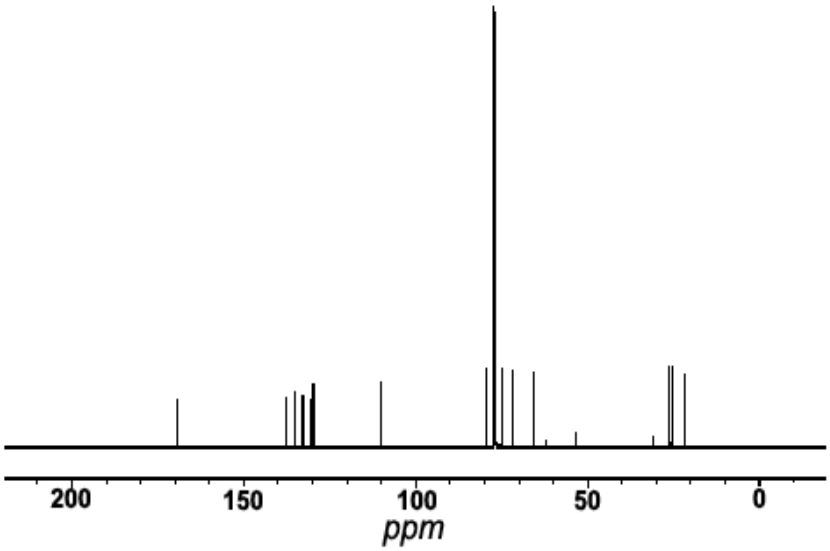


Supplementary Figure S19. ^{1}H NMR spectrum of **8** in MeOD-d4.

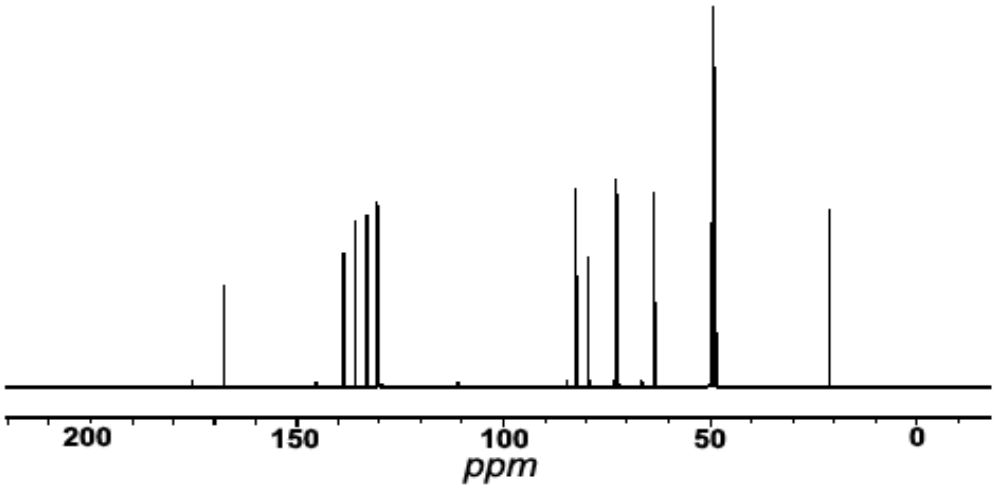
b. ^{13}C NMR spectra.



Supplementary Figure S20. ^{13}C NMR spectrum of **6** in $CDCl_3$.

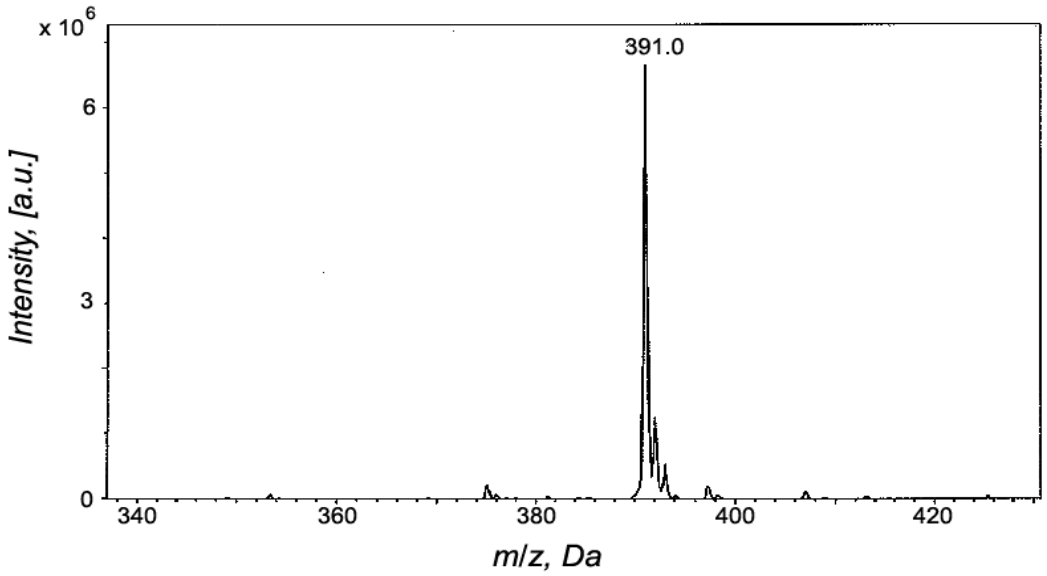


Supplementary Figure S21. ^{13}C NMR spectrum of **7** in $CDCl_3$.

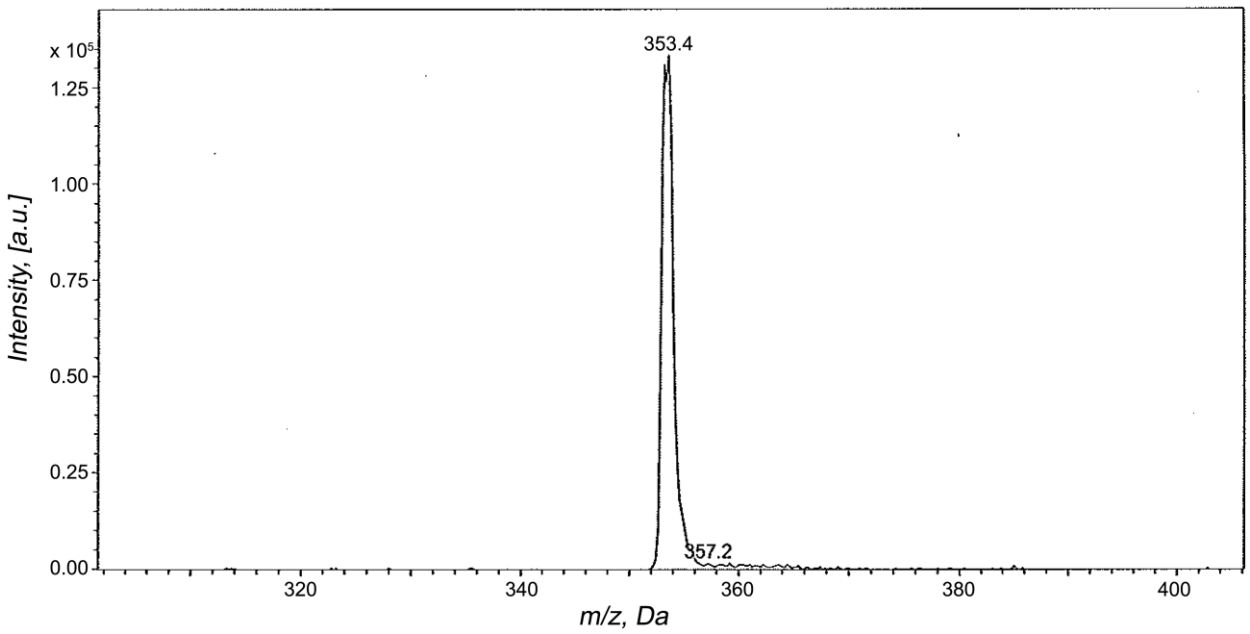


Supplementary Figure S22. ^{13}C NMR spectrum of **8** MeOD-d4.

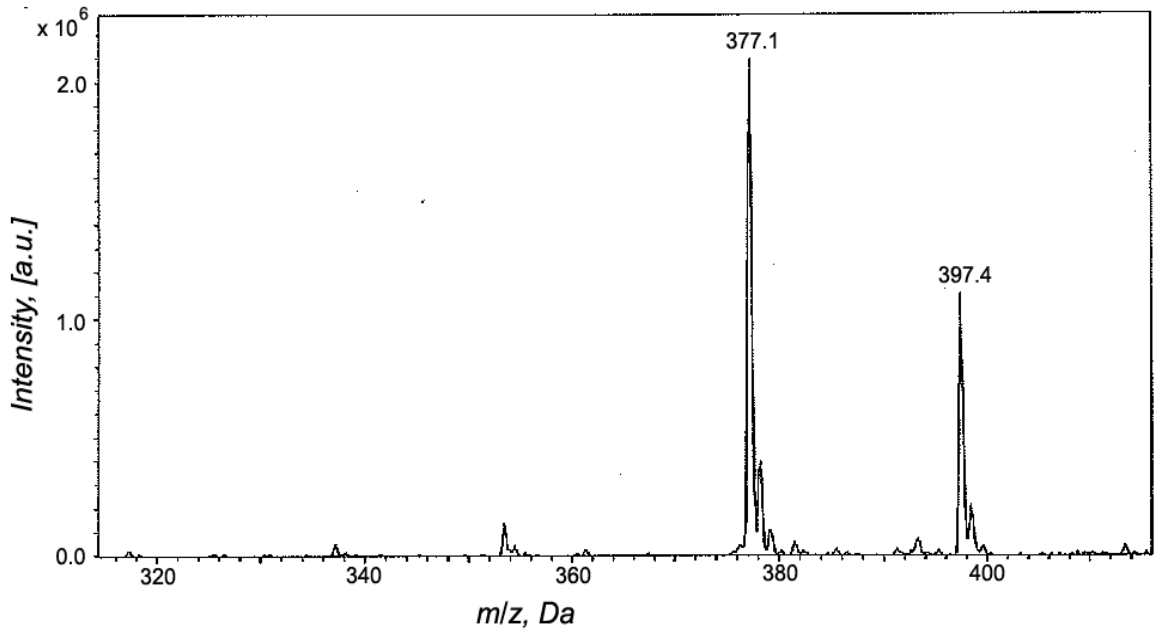
c. ESI MS spectra.



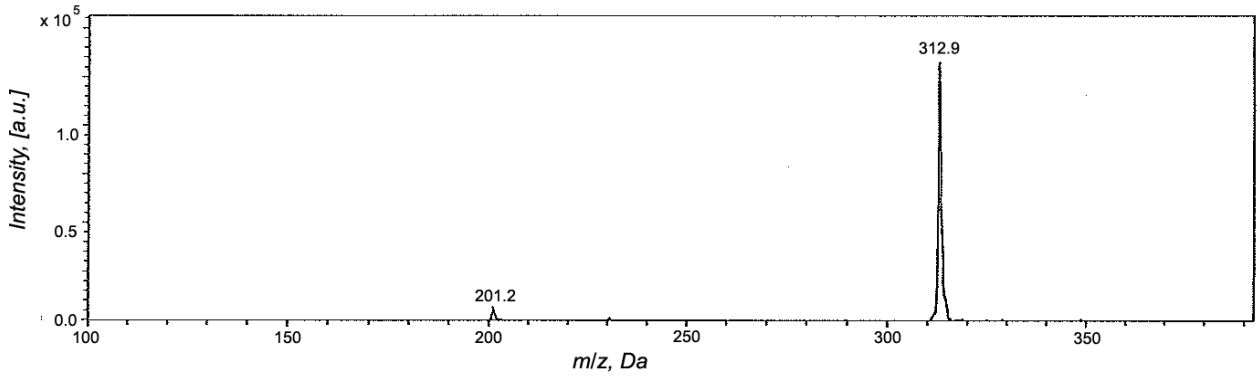
Supplementary Figure S23. ESI (+) MS spectrum of **6** $[M-H+Na]^+$ calcd. 391.1 Da, found 391.0 Da.



Supplementary Figure S24. ESI (-) MS spectrum of 7 $[MW-H]^-$ calcd. 353.3, found 353.4.



Supplementary Figure S25. ESI (+)MS spectrum of **7** $[M-H+Na]^+$ calcd. 377.2, found 377.1.



Supplementary Figure S26. ESI (-)MS spectrum of **8** $[MW-H]^-$ calcd. 313.1, found 312.9.

X. References

(1) Schmid, C. R., and Bryant, J. D. (1995) D-(R)-glyceraldehyde acetonide. *Org. Synth. 72*, 6-13.

(2) Krishna, P. R., Lavanya, B., and Sharma, G. V. M. (2003) Stereoselective synthesis of C-phenyl D- and L-glycero heptopyranosides. *Tetrahedron-Asymmetr. 14*, 419-427.

(3) Hannant, J., Hedley, J. H., Pate, J., Walli, A., Farha Al-Said, S. A., Galindo, M. A., Connolly, B. A., Horrocks, B. R., Houlton, A., and Pike, A. R. (2010) Modification of DNA-templated conductive polymer nanowires via click chemistry. *Chemical Communications 46*, 5870-5872.

(4) Grimm, G. N., Boutorine, A. S., and Helene, C. (2000) Rapid routes of synthesis of oligonucleotide conjugates from non-protected oligonucleotides and ligands possessing different nucleophilic or electrophilic functional groups. *Nucleos. Nucleot. Nucl. 19*, 1943-1965.

(5) Kusmierek, J. T., and Singer, B. (1982) Chloroacetaldehyde-treated ribo- and deoxyribopolynucleotides. 1. Reaction products. *Biochemistry 21*, 5717-5722.

(6) Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., and Weiss, S. (2004) Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. USA 101*, 8936-8941.

(7) Selvin, P. R., and Ha, T. *Single-molecule techniques - a laboratory manual*; 1st ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, USA, 2007.

(8) Zhao, R., and Rueda, D. (2009) RNA folding dynamics by single-molecule fluorescence resonance energy transfer. *Methods 49*, 112-117.

(9) Ha, T., and Tinnefeld, P. (2012) Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annu. Rev. Phys. Chem. 63*, 595-617.

(10) Kowerko, D., Hadzic, M., Börner, R., König, S. L. B., Heidernätsch, M., and Sigel, R. K. O. (2014) Multifunctional analysis software for heterogeneous single molecule FRET data. *submitted*.